 Nucleic Acids Research, 1999, Vol. 27, No. 19 **3771–3778**

The *Drosophila* poly(A)-binding protein II is ubiquitous throughout *Drosophila* development and has the same function in mRNA polyadenylation as its bovine homolog *in vitro*

Béatrice Benoit, Annette Nemeth[1], Nathalie Aulner[2], Uwe Kühn[1], Martine Simonelig*, Elmar Wahle[1] and Henri-Marc Bourbon[2]

Dynamique du Génome et Evolution, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris, France, [1]Institut für Biochemie, Justus Liebig Universität Giessen, Heinrich Buff Ring 58, 35392 Giessen, Germany and [2]Centre de Biologie du Développement, UMR 5544 du CNRS, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse, France

Received August 6, 1999; Revised and Accepted August 13, 1999 DDBJ/EMBL/GenBank accession no. AF116341

ABSTRACT

The poly(A)-binding protein II (PABP2) is one of the polyadenylation factors required for proper 3'-end formation of mammalian mRNAs. We have cloned *Pabp2*, the gene encoding the *Drosophila* homolog of mammalian PABP2, by using a molecular screen to identify new *Drosophila* proteins with RNP-type RNA-binding domains. Sequence comparison of PABP2 from *Drosophila* and mammals indicates that the most conserved domains are the RNA-binding domain and a coiled-coil like domain which could be involved in protein–protein interactions. *Pabp2* produces four mRNAs which result from utilization of alternative poly(A) sites and encode the same protein. Using an antibody raised against *Drosophila* PABP2, we show that the protein accumulates in nuclei of all transcriptionally active cells throughout *Drosophila* development. This is consistent with a general role of PABP2 in mRNA polyadenylation. Analysis of *Drosophila* PABP2 function in a reconstituted mammalian polyadenylation system shows that the protein has the same functions as its bovine homolog *in vitro*: it stimulates poly(A) polymerase and is able to control poly(A) tail length.

INTRODUCTION

The 3'-end formation of virtually all eukaryotic mRNA precursors (pre-mRNAs) occurs in a two-step reaction, which includes endonucleolytic cleavage of pre-mRNA, followed by the addition of a poly(A) tail of 200–250 nt to the upstream cleavage product (reviewed in 1–4). In mammals, both poly(A) site selection and polyadenylation require a multicomponent complex whose assembly is dependent on the highly conserved AAUAAA polyadenylation signal found 10–35 nt upstream of the cleavage site. Most of the 3'-end processing factors have been purified from mammalian cells and cDNAs coding for them have been cloned (reviewed in 2). Accurate cleavage of mammalian pre-mRNA substrates *in vitro* requires five factors: cleavage and polyadenylation specificity factor (CPSF), which consists of four subunits and binds the AAUAAA element; cleavage stimulation factor (CstF), which consists of three subunits; cleavage factors I and II (CFI and CFII); poly(A) polymerase (PAP). Poly(A) addition to a 'pre-cleaved' RNA can be reconstituted from three purified factors: PAP, CPSF and poly(A)-binding protein II (PABP2). PAP by itself has a low affinity for RNA and is unable to recognize a pre-mRNA specifically. Specificity requires the AAUAAA element and CPSF (5,6), but even in the presence of CPSF, the activity of PAP remains weak. PAP activity is again stimulated by binding of PABP2 to the poly(A) tail when the tail has reached a length of 10 adenylate residues (7,8). Together, CPSF and PABP2 stimulate PAP activity by holding PAP on the RNA such that a full-length poly(A) tail is synthesized in a single processive event (9). PABP2 also appears to be involved in control of poly(A) tail length. When the poly(A) tail has reached a length of ~250 residues, its elongation is no longer processive and becomes slow and distributive. PABP2 binds the poly(A) tail stoichiometrically; therefore, it is thought to be involved in a counting mechanism that determines the number

*To whom correspondence should addressed at present address. Tel: +33 4 99 61 99 59; Fax: +33 4 99 61 99 57; Email: Martine.Simonelig@igh.cnrs.fr
Present addresses:
Béatrice Benoit and Martine Simonelig, Génétique du Développement de la Drosophile, Institut de Génétique Humaine, 141 rue de la Cardonille, 34396 Montpellier Cedex 5, France
Annette Nemeth, Uwe Kühn and Elmar Wahle, Institut für Biochemie, Universität Halle, 06099 Halle, Germany

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

of residues in the poly(A) tail and allows processive elongation to terminate when the length of the tail is ~250 nt (10).

cDNA and genomic clones encoding bovine PABP2 were previously isolated (11). The sequence of cDNAs predicts a 306 amino acid protein with a calculated molecular weight of 33 kDa. This protein contains a single RNP-type consensus RNA-binding domain (12,13) between an acidic N-terminal and a basic C-terminal domain. Recently, cDNAs and genomic clones encoding PABP2 have also been isolated from mouse (14) and human (15). The protein is highly conserved between the three species. In contrast, the RNA-binding domain of PABP2 is clearly different from any of the four RNA-binding domains found in the cytoplasmic poly(A)-binding protein (PABP) (16,17).

To confirm the role of PABP2 as a key polyadenylation factor *in vivo*, we have started a functional analysis of the *Drosophila melanogaster* PABP2 protein. The *Drosophila* gene encoding PABP2 (*Pabp2*) produces several mRNAs which all encode the same protein. This protein is mostly nuclear and present ubiquitously throughout *Drosophila* development. In addition, we show that *Drosophila* PABP2 has the same function as its bovine homolog in reconstituted mammalian polyadenylation assays. This study is a first step toward a genetic analysis which would allow determination of the biological function of *Drosophila* PABP2.

MATERIALS AND METHODS

Cloning of *Pabp2* genomic DNA and cDNAs

A *D.melanogaster* genomic library in λEMBL3 (18) was screened with a probe corresponding to the 2C3 cDNA insert (19). One phage, λ2G2, was recovered and subcloned into pBluescript II (Stratagene). A 3.7 kb *Nsi*I–*Pst*I subclone was sequenced on both strands by the chain termination method using Sequenase 2.0 (US Biochemical) and either T3, T7 or specific primers. The *Pabp2* genomic sequence was deposited in GenBank (accession no. AF116341). Imaginal disc cDNAs were isolated by screening an oligo(dT)-primed imaginal disc cDNA library in λZAPII, provided by Konrad Basler (University of Zürich), with a probe corresponding to 144 bp of *Pabp2* cDNA spanning the RNA-binding domain. Sequence data were analyzed with the FASTA, BLAST, BESTFIT and PRETTY programs of the GCG package.

RNA blots and RNase protection assays

RNA blots were performed as reported previously (20) using the wild-type stock Canton S. The RNase protection assays were performed as reported by Audibert and Simonelig (21), with total RNA from the Canton S stock. To produce the antisense probe a *Eag*I–*Bbs*I genomic fragment of *Pabp2* (nt 1452–3068) was cloned into pBluescript II; the resulting clone was digested with *Bfa*I (position 2040) and transcribed with T3 RNA polymerase. The probe was 1095 nt long and after hybridization at 42°C and RNase digestion, protected fragments with sizes of 597 nt for 2C3 and DD68 cDNAs, 424 nt for 2C7 cDNA, 393 nt for 2C1 cDNA and 273 nt for 2C2 and DD16 cDNAs were expected.

Expression and purification of *Drosophila* and bovine PABP2-His$_6$ proteins

The complete ORF of *Drosophila Pabp2* was PCR amplified with an upstream primer introducing an *Nde*I site as part of the ATG initiation codon and a downstream primer introducing a *Bam*HI site after the stop codon. The PCR product was digested with *Nde*I and *Bam*HI and cloned into the T7 expression vector pGM10 (22). The ORF of bovine *PABP2* was cut out as a *Nde*I–*Bam*HI fragment from the expression construct described earlier (11) and cloned into pGM10. Both fusion proteins carry the sequence Met-Ala-His$_6$ at their N-termini. Transformed cells were grown as described (11), except ampicillin was substituted by 100 μg/ml carbenicillin and the culture medium was changed every 2 h to guarantee stringent selection conditions. The cells were harvested and resuspended in 50 mM Tris–HCl, pH 8.0, 300 mM KCl and stored at –70°C. While the cells were allowed to thaw, 10% (v/v) glycerol, 1 mM phenylmethylsulfonyl fluoride, 2 μg/ml leupeptin and 2 μg/ml pepstatin were added. The suspension was sonicated, centrifuged, and the supernatants were incubated overnight with 500 μl/l culture of a 50% slurry of Ni^{2+}–NTA agarose (Qiagen) on a rotating platform in a cold room. The mix was packed into a column and washed with 10 column vol 50 mM Tris–HCl pH 8.0, 50 mM KCl, 10% (v/v) glycerol, 10 mM imidazole-HCl. The protein was eluted in 3 ml of the same buffer containing 250 mM imidazole-HCl. *Drosophila* PABP2-His$_6$ was further purified on a 1 ml MonoQ FPLC column and bovine PABP2-His$_6$ on a 1 ml MonoS FPLC column (Pharmacia).

Poly(A) binding and polyadenylation assays

Poly(A) binding activity was determined by a nitrocellulose filter-binding assay. A reaction mix with 5'-end-labeled RNA in filter-binding buffer (50 mM Tris–HCl pH 8.0, 100 mM KCl, 1 mM EDTA, 0.5 mM DTT) was assembled, divided into aliquots of 40 μl each, and the binding reaction was started by the addition of protein in a volume of 10 μl. For the K_d determination 0.7 nM A_{14} was used and for the competition experiments 150 nM A_{70} (concentrations refer to AMP). The incubation time was 30 min at room temperature. Immediately before use, each nitrocellulose filter (0.2 μm pore size; Schleicher & Schuell NC 20) was washed with 1 ml wash buffer (50 mM Tris–HCl pH 8.0, 100 mM KCl, 1 mM EDTA) supplemented with 4 μg/ml *Escherichia coli* rRNA, then 80% of the binding reaction was applied to the filter and the filter was washed with 5 ml ice-cold wash buffer. Bound radioactivity was counted in a liquid scintillation counter. Polyadenylation assays were carried out as described (10), except the reaction volume was reduced to 15 μl. PABP2 and CPSF from calf thymus and recombinant bovine PAP were purified as described (10). RNA substrates were prepared as described (8,10).

Antibody preparation

For the generation of rabbit polyclonal antibodies, the recombinant *Drosophila* PABP2 protein was purified under denaturing conditions. Cells were resuspended in 8 M urea, 100 mM NaH_2PO_4, 10 mM Tris–HCl, pH 8.0, stirred for 2 h at room temperature, sonicated and centrifuged. An aliquot of 2 ml of a 50% slurry of Ni^{2+}–NTA agarose was added to the supernatant

for every liter of culture, and the mixture was incubated for 2 h on a rotating platform at room temperature and then packed into a column. The column was washed with 20 vol of the resuspension buffer and 60 vol of the same buffer adjusted to pH 6.3. The protein was eluted with this buffer containing 500 mM imidazole-HCl and dialyzed extensively against a series of buffers with decreasing concentrations of urea. The protein was concentrated in a Centricon 10 cartridge (Amicon), washed twice with 1 ml 50 mM NH_4HCO_3 and lyophilized. Antibodies were prepared by Eurogentec (Belgium). Antibody affinity purification was performed as follows. Approximately 500 µg purified, recombinant *Drosophila* PABP2 was dissolved in 1 ml of coupling buffer (0.2 M $NaHCO_3$, 0.5 M NaCl pH 8.3) and covalently linked to a 1 ml NHS-activated HiTrap column (Pharmacia). Activation, protein coupling and inactivation of remaining reactive groups were carried out according to the manufacturer's protocols. After coupling, the column was equilibrated with PBS. Antibody-containing rabbit serum was centrifuged for 2 min at 14 000 *g*. An aliquot of 1.1 ml of the cleared serum was diluted with 4 ml PBS and passed over the column five times. After washing with 15 ml PBS, bound antibodies were eluted with 0.2 M glycine-HCl, pH 2.5. Fractions (1 ml) were collected into tubes containing 350 µl of ice-cold 1 M Tris–HCl and immediately mixed. After pooling of antibody-containing fractions, they were concentrated and the buffer was changed to PBS in a Centriplus 10 vial. Antibodies were stabilized by the addition of acetylated BSA to a final concentration of 0.1 mg/ml.

Western blots and antibody staining

The wild-type stock Canton S was used. For western blots, tissues were homogenized in 3× sample buffer (2% SDS, 0.125 M Tris–HCl pH 6.9, 5% β-mercaptoethanol, 10% glycerol, bromophenol blue) and proteins were separated by SDS–PAGE. After electrotransfer to nitrocellulose, the blot was blocked in PBS, 0.5% Tween-20, 5% milk. The *Drosophila* PABP2 protein was detected using the affinity-purified antibody at a dilution of 1:1000 in PBS, 0.5% Tween-20, 5% milk and revealed using anti-rabbit Ig horseradish peroxidase (1:1500) and an ECL kit (Amersham). For immunostaining, ovaries were fixed for 20 min in 4% paraformaldehyde, 0.5% NP-40 (200 µl) and heptane (800 µl), then washed for 3 min with methanol and three times with PBS. Tissues were blocked for 2 h in PBT (PBS, 1% BSA, 0.1% Triton X-100) and incubated overnight at 4°C with the affinity-purified antibody 1:250 in PBT. After four washes over 2 h in PBS, 0.1% BSA, 0.1% Triton X-100, 2% horse serum, tissues were incubated with preabsorbed FITC-conjugated anti-rabbit IgG (Institut Pasteur Production) 1:50 in PBS, 0.1% BSA, 0.1% Triton X-100 for 2 h at room temperature. Following four washes in PBS, 0.1% BSA, 0.1% Triton X-100 over 2 h and one wash in PBS, tissues were mounted in Citifluor. Embryos were dechorionated for 5 min in bleach and fixed for 30 min in 0.1 mM PIPES, 2 mM $MgSO_4$, 1 mM EGTA (2.1 ml), 37% formaldehyde (0.4 ml) and heptane (2.5 ml). Following devitellinization in heptane (3 ml) and methanol–EGTA (9:1) (3 ml), embryos were rehydrated in PBS, blocked for 2 h in PBT with 0.5 mg /ml RNase A at room temperature and treated like ovaries except that the affinity-purified antibody was at a dilution of 1:500. Embryos were then incubated for 3 min in PBS, 0.1% BSA, 0.1% Triton X-100 with 1 µg/ml propidium iodide, before being washed in PBS and mounted in Citifluor. Confocal microscopy was carried out using a MCR 600 Bio-Rad confocal microscope.

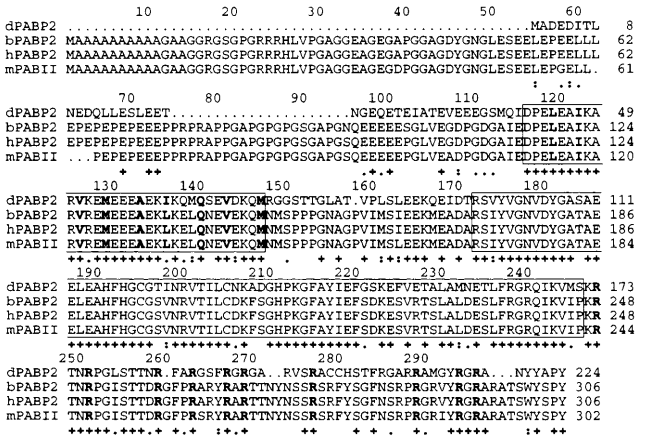

Figure 1. Comparison of *Drosophila* PABP2 to mammalian homologs. Alignment of the amino acid sequence of *Drosophila* PABP2 (dPABP2) with sequences of bovine (bPABP2), human (hPABP2) and mouse (mPABII) PABP2 (accession nos: *Drosophila*, L34934; bovine, X89969; human, AF026029; mouse, U93050). Plus signs indicate the positions of conserved residues shared by the four proteins. Colons represent similarities and single dots represent less similar amino acids. The domain similar to α-helical coiled-coil domains and the RNP-type RNA-binding domain are boxed. Within the coiled-coil domain, hydrophobic residues present at conserved positions in the four 7-residue repeats are in bold. In the C-terminal domain, arginine residues which are conserved between mammals and *Drosophila* are in bold.

RESULTS

The *Drosophila* PABP2 protein

By a PCR-based approach, we previously cloned cDNAs encoding three novel putative RNA-binding proteins of the RNP-type superfamily from *D.melanogaster* (19). Database searches revealed that one of these proteins, referred to as Rox2, shares extensive similarity with the bovine polyadenylation factor PABP2 (11). The two proteins show 58% identity and 74% similarity. On the basis of this homology, the *rox2* gene was renamed *Pabp2*. An alignment of the PABP2 protein from *Drosophila*, cow, human and mouse is shown in Figure 1. The shorter size of the *Drosophila* protein (224 amino acids) compared to the mammalian homologs (306 amino acids in cow) results from the lack in *Drosophila* PABP2 of an N-terminal region rich in alanine, glycine and proline residues. This region contains, in particular, an alanine stretch (amino acids 2–11) which is missing in *Drosophila* PABP2. In addition to the RNP-type RNA-binding domain, which is well conserved (77% identity and 87% similarity between the bovine and *Drosophila* proteins), another portion of the protein (amino acids 41–72 in *Drosophila*) in the N-terminal region is strikingly conserved (81% identity and 97% similarity between the bovine and *Drosophila* proteins). This conserved region shows a regular seven-residue pattern in which hydrophobic residues are found at alternate intervals of three and four residues. This pattern is characteristic of α-helical coiled-coil domains known to be involved in homotypic and/or

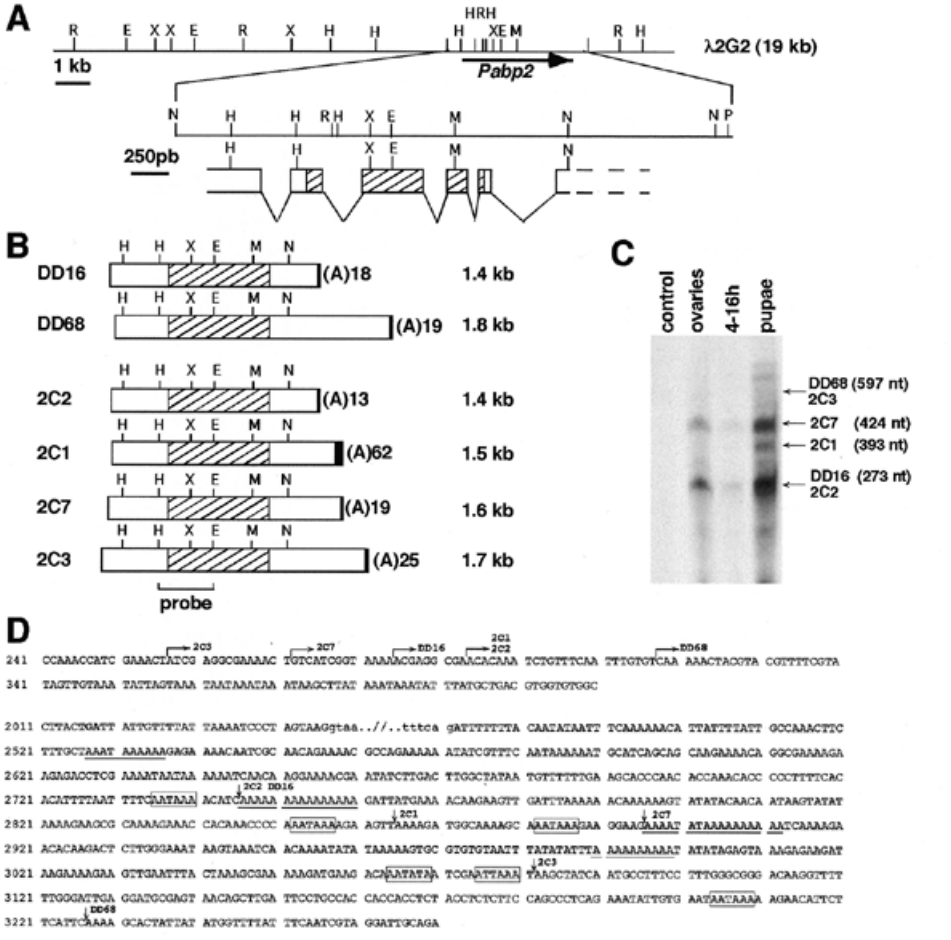

Figure 2. Structural organization of *Pabp2* genomic and cDNA clones. (**A**) Restriction map of the λ2G2 clone and structure of the *Pabp2* gene (E, *Eag*I; H, *Hin*dIII; M, *Sma*I; N, *Nsi*I; P, *Pst*I; R, *Eco*RI; X, *Xho*I). Hatched boxes represent coding sequences and open boxes are 5'- and 3'-UTRs. (**B**) Schematic representation of *Pabp2* cDNAs from imaginal discs (DD16 and DD68) and from 4–8 h embryos (2C2, 2C1, 2C7 and 2C3). The sizes of the clones and of their poly(A) tracts are indicated. (**C**) RNase protection assays with total RNA from ovaries, 4–16 h embryos and pupae. Positions of the protected fragments are indicated. Control was with 10 µg of tRNA. (**D**) Sequences of *Pabp2* 5'- and 3'-UTRs. Intron 5 is in lower case letters and the stop codon (TGA) is in bold. Arrows indicate the start and the end of each cDNA clone. Poly(A) tracts in the 3'-UTR are underlined and consensus poly(A) signals are boxed.

heterotypic protein–protein interactions (23,24). Bovine PABP2 can form oligomers *in vitro* (11); therefore, the conservation of this N-terminal domain and its similarity to coiled-coil domains indicate that it could be involved in oligomerization of PABP2. Alternatively, it might mediate an interaction with a different protein, e.g. PAP. Finally, 10 out of 15 arginine residues present in the basic C-terminal domain of the bovine protein are conserved in its *Drosophila* homolog (Fig. 1). Most of these arginines are methylated in bovine PABP2 (25).

Organization and expression of the *Drosophila Pabp2* gene

We used the largest cDNA isolated previously (2C3) (19) as a probe to screen a *Drosophila* genomic library in λEMBL3. One phage, λ2G2, which covers the whole *Pabp2* transcribed sequence, was recovered (structure shown in Fig. 2A). A 3.7 kb *Nsi*I–*Pst*I fragment from λ2G2, which spans the largest *Pabp2* cDNA, was subcloned and sequenced entirely. Comparison of this sequence to that of *Pabp2* cDNAs allowed determination of the structure of the gene (Fig. 2A). *Drosophila* genomic Southern blots showed that *Pabp2* is in single copy in the *Drosophila* genome (data not shown). The four cDNAs isolated previously from an embryonic library (19) and two cDNAs newly isolated from an imaginal disc library are shown in Figure 2B. Three cDNAs (2C2, 2C7 and 2C3) contain an extra G residue at their 5'-end which is not present in the corresponding genomic sequence. This may reflect a cap structure (26), suggesting that these cDNAs could be complete at their 5'-end and that alternative transcription start sites, clustered within 38 bp, could be used (Fig. 2D).

Figure 2A shows that the 5'- and 3'-untranslated regions (UTRs) are each interrupted by a single intron and that three introns occur in the coding sequence. Strikingly, the third intron of *Pabp2* and the fifth intron of its bovine counterpart disrupt the corresponding protein sequences at the same

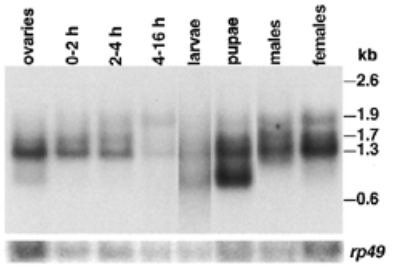

Figure 3. mRNA profile of *Pabp2* during *Drosophila* development. Northern blot of poly(A)$^+$ RNA from ovaries, 0–2, 2–4 and 4–16 h embryos, third instar larvae, pupae, adult males and adult females. Hybridization was with an RNA probe spanning the 5'-half of the *Pabp2* coding sequence (Fig. 2B). The blot was reprobed with the *rp49* clone as a loading control.

position, namely after an AAG codon specifying Lys167 in *Drosophila* and Lys243 in cow (Fig. 1).

The six cDNAs that we have isolated differ in the length of their 3'-UTRs. All six have a poly(A) stretch at their 3'-end and a consensus polyadenylation signal upstream of this oligo(A) (Fig. 2D). This suggests that all cDNAs could have a mature 3'-end. However, like its bovine homolog, *Pabp2* contains oligo(A) stretches in its 3'-UTR. cDNAs 2C2, DD16 and 2C7 stop upstream of one of these oligo(A) stretches (Fig. 2D) and could result from internal priming within these stretches. We performed RNase protection assays to determine whether 2C2/DD16 and 2C7 have a mature 3'-end (Fig. 2C). A probe complementary to a region of the 3'-UTR (position 2040–3068) was hybridized to total RNA from ovaries, embryos and pupae. This probe spans three poly(A) sites mapped from the sequence of 2C2/DD16, 2C1 and 2C7 cDNAs. The size of the probe-protected fragments at different developmental stages confirms the utilization of these three poly(A) sites (Fig. 2C). Therefore, analysis of cDNAs and RNase protection assays indicate utilization of at least five poly(A) sites in *Pabp2*. mRNAs produced by utilization of these alternative poly(A) sites have different 3'-UTRs but encode the same protein.

Pabp2 mRNAs were examined at various developmental stages by northern blot using a fragment of cDNA 2C3 as a probe (see Fig. 2B). Consistent with the structure of *Pabp2* cDNAs, several transcripts were detected (Fig. 3). Their size is in rough agreement with the size of the cDNAs. Two mRNA species (1.3 and 1.7 kb) are present at all stages. Two other mRNAs are also detected. The 1 kb species is abundant in ovaries, larvae and pupae and the 1.9 kb species is detected at all stages except in ovaries and early embryos. Production of these mRNAs could result from utilization of the alternative poly(A) sites and/or from retention of introns. In cow and human, the first and sixth introns of *PABP2* appear to be retained in some cDNA clones (11,15). We tested, by RT–PCR using RNA from males, females or ovaries as template, whether introns could be retained in *Pabp2* RNAs and we found that they are not (data not shown). This suggests that the different transcripts result from utilization of the alternative poly(A) sites and possibly from utilization of different transcription start sites.

Immunodetection of the PABP2 protein during *Drosophila* development

An antibody against the *Drosophila* PABP2 protein was raised in rabbit and affinity purified. The specificity of this antibody was tested by western blot using protein extracts from various developmental stages. Figure 4A shows that a single 33 kDa protein is revealed by the antibody at each stage. The cDNA sequence predicts a 25 kDa protein. A difference between the predicted and the measured molecular mass of bovine PABP2 was also found and is thought to result from the unusual amino acid composition of the protein (11).

Using this specific antibody, we examined the localization of the PABP2 protein during *Drosophila* development. Figure 4B shows the distribution of PABP2 during oogenesis. Adult ovaries contain both somatic and germline cells. In the germarium, each germline stem cell produces a progeny of 15 nurse cells and one oocyte. This cluster is surrounded by somatic follicle cells to form the egg chamber. Egg chambers at successive developmental stages are arrayed within an ovariole, later stages of development being located more posteriorly (27). The PABP2 protein is detected throughout oogenesis in germline and follicle cells, and it mostly accumulates in nuclei, including oocyte nuclei. PABP2 also appears to be mostly nuclear in the germarium, but starting with egg chamber formation, nuclear accumulation in nurse cells and oocyte is accompanied by a widespread distribution of the protein in the cytoplasm. At stage 10, the protein is abundant in nuclei of nurse cells, oocyte and follicle cells and in nurse cell cytoplasm, but its amount is lower in oocyte cytoplasm. This could indicate that protein that is transferred to the oocyte from the nurse cells is translocated to the nucleus.

The distribution of the PABP2 protein during embryogenesis (28) was examined by double staining with the anti-PABP2 antibody and propidium iodide to reveal DNA and is shown in Figure 4C–N. In freshly laid embryos, PABP2 is uniformly distributed in the cytoplasm (Fig. 4C). It then progressively concentrates in nuclei during the preblastoderm and syncytial blastoderm stages (Fig. 4D, E and G–I), before the onset of zygotic transcription. After cellularization, the protein is mostly nuclear (Fig. 4F and J–L). It remains mostly nuclear and appears to be present in all cells at later stages of embryogenesis and in larvae (data not shown). In Figure 4J–N the first mitotic domain after cellularization of the embryo is also visible (29). Figure 4M and N shows that, during mitosis, the amount of PABP2 appears to be reduced in prophase nuclei; during metaphase, PABP2 is found in the cytoplasm and a large amount of the protein remains close to the metaphase plate.

Immunodetection of PABP2 in HeLa cells indicates that the protein is present in the nucleus both in a widespread pattern and in more intensely stained speckles (30) which correspond to sites in the nucleus where splicing factors accumulate. Figure 4O shows that nuclear staining with anti-PABP2 in cellularized embryos is non-homogeneous and also appears as a widespread staining with more intense regions.

Biochemical characterization of *Drosophila* PABP2

Drosophila PABP2 was expressed in *E.coli* with an N-terminal His_6-tag. The protein was purified on a Ni^{2+}–NTA column, followed by chromatography on a MonoQ column. The

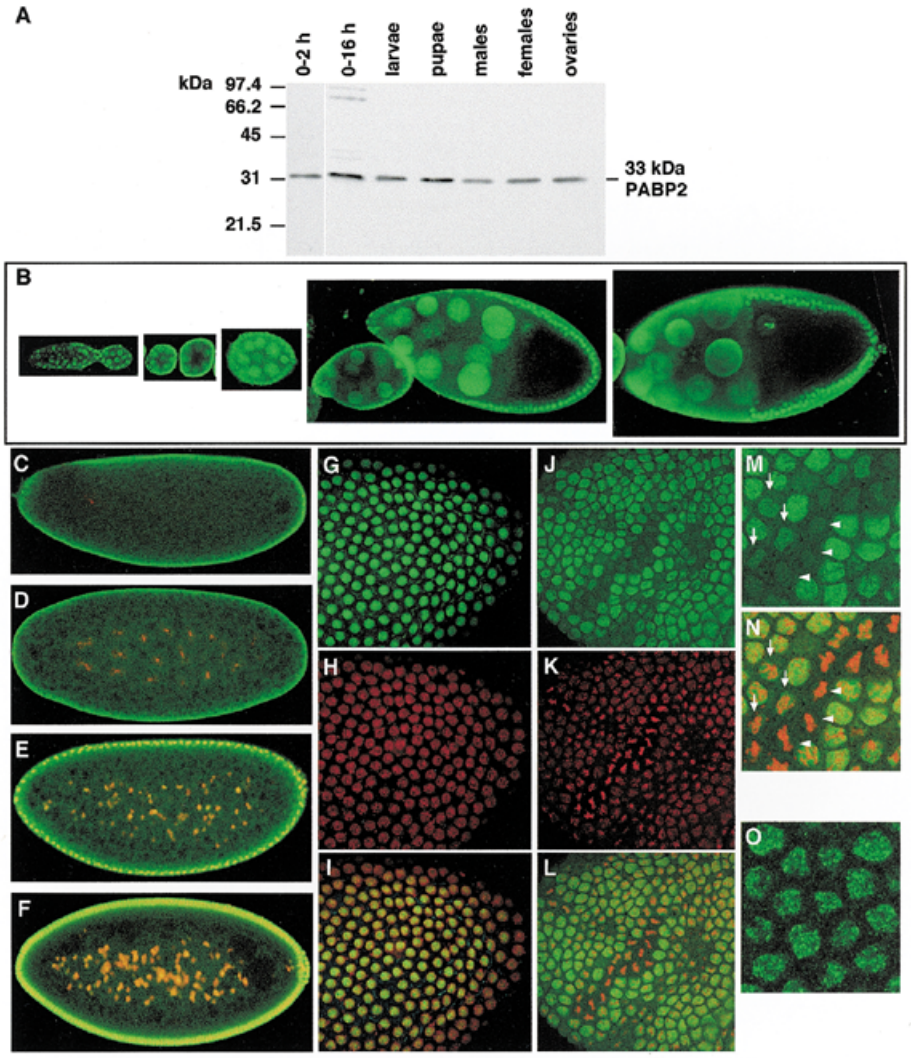

Figure 4. Detection of PABP2 protein during *Drosophila* development. (**A**) Western blot with protein extracts from 0–2 and 0–16 h embryos, third instar larvae, pupae, adult males, adult females and ovaries. (**B**) Immunodetection of the PABP2 protein during oogenesis; germarium and stages 1 and 2, stages 4 and 5, stage 7, stages 8 and 9 and stage 10B egg chambers. (C–N) Double staining of embryos with anti-PABP2 antibody (green) and propidium iodide (red). Posterior of the embryos is to the right. (**C**) Freshly laid embryo; (**D**) preblastoderm; (**E**) syncytial blastoderm; (**F**) cellularized blastoderm. Zygotic transcription starts during the syncytial blastoderm stage. (G–I) Posterior pole of a syncytial blastoderm stage embryo which shows accumulation of PABP2 in nuclei. (**G**) Anti-PABP2 staining; (**H**) propidium iodide staining; (**I**) superimposition of both patterns. (J–L) Anterior pole of an embryo after cellularization which shows accumulation of PABP2 in nuclei and mitotic domain 1. (**J**) Anti-PABP2 staining; (**K**) propidium iodide staining; (**L**) superimposition of both patterns. (M–N) Higher magnification of mitotic domain 1 shown in (J–L). (**M**) Anti-PABP2 staining; (**N**) superimposition of anti-PABP2 and propidium iodide staining patterns. Condensation of chromosomes in prophase (arrows) and metaphase plates (arrowheads) are visible. (**O**) High magnification of cells in a cellularized embryo which shows the non-homogeneous staining of anti-PABP2 in nuclei.

recombinant protein was active in RNA binding. In nitrocellulose filter binding experiments, oligo(A) with a chain length of 14 nt was bound with an apparent K_d of 10 nM (data not shown). This is very similar to the affinity of His-tagged recombinant bovine PABP2. In experiments designed to examine the RNA binding specificity of *Drosophila* PABP2, poly(A) with a chain length of ~70 nt was used as a radiolabeled ligand and binding was competed with unlabeled polynucleotides (Fig. 5). Unlabeled poly(A) competed as expected, binding being reduced to 50% by the addition of an equimolar amount of competitor. Poly(G) competed as efficiently as poly(A). *Escherichia coli* rRNA was a weaker competitor by a factor of about 10, whereas poly(C) had no significant effect even at a 100-fold excess. Thus, *Drosophila* PABP2 binds specifically to purine polyribonucleotides, as was previously reported for mammalian PABP2 (7,8).

Drosophila PABP2 was also assayed for its ability to replace mammalian PABP2 in the *in vitro* polyadenylation system. *Drosophila* PABP2 stimulated recombinant bovine PAP when assayed in the absence of CPSF with a simple oligo(A) primer (data not shown). The *Drosophila* protein was also able to replace bovine PABP2 in a complete reconstituted system including both mammalian CPSF and recombinant mammalian PAP with a primer RNA containing an AAUAAA

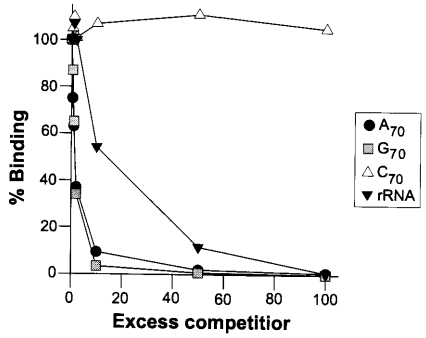

Figure 5. Specificity of poly(A) binding by *Drosophila* PABP2-His_6. For the poly(A) binding experiment radioactively labeled A_{70} and unlabeled competitor RNA in the indicated amounts were mixed before the protein was added. Binding of *Drosophila* PABP2-His_6 to A_{70} without competitor was set to 100%.

sequence as well as an oligo(A) tail. Under these conditions, authentic bovine PABP2, recombinant His-tagged bovine PABP2 and recombinant His-tagged *Drosophila* PABP2 all led to rapid and processive poly(A) tail extension, whereas the reaction lacking a PABP2 showed a slow elongation (Fig. 6, compare the 20 s time points). The *Drosophila* PABP2-containing reaction also exerted a length control similar to the mammalian protein: poly(A) tail extension was rapid up to a length of ~250 nt and then proceeded more slowly, as described in detail before (10) (Fig. 6).

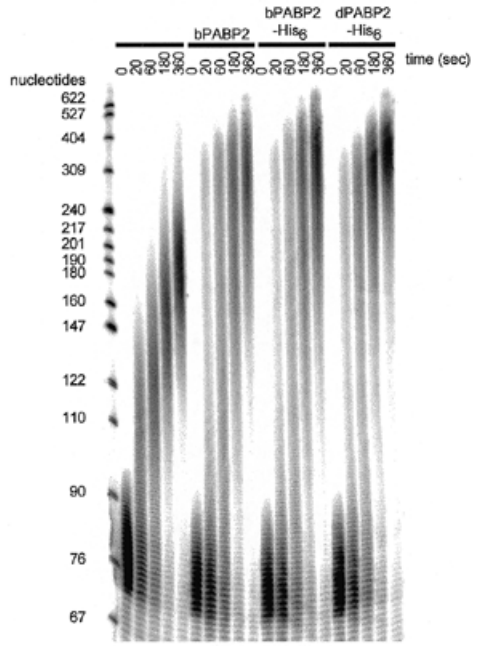

Figure 6. *Drosophila* PABP2 stimulates bovine PAP. The assay was carried out with 4 fmol PAP, 60 fmol CPSF, 50 fmol L3pre$(A)_{15}$ substrate and with or without 100 fmol PABP2 purified from calf thymus (bPABP2), 200 fmol bovine PABP2-His_6 (bPABP2-His_6) or 500 fmol *Drosophila* PABP2-His_6 (dPABP2-His_6), both expressed in *E.coli*. Reactions, assembled on ice in the absence of ATP, were prewarmed, started by the addition of ATP and stopped after the indicated times by the addition of SDS-containing buffer. After purification the reaction products were analyzed on a denaturing 10% polyacrylamide gel.

DISCUSSION

In this paper, we show that the *rox2* cDNA, which we cloned previously in a molecular screen to isolate new RNP-type RNA-binding proteins, encodes the *Drosophila* homolog of mammalian PABP2. Biochemical characterization of the *Drosophila* protein shows that it behaves as mammalian PABP2 in *in vitro* assays: it binds specifically to purine polyribonucleotides and is able to replace bovine PABP2 in a reconstituted mammalian polyadenylation system. In the presence of bovine CPSF, recombinant bovine PAP and a primer RNA containing an AAUAAA signal and an oligo(A) tail, *Drosophila* PABP2 leads to rapid and processive extension of the poly(A) tail up to 250 nt and then the reaction becomes slow. This indicates that *Drosophila* PABP2 is able to stimulate PAP and to control the poly(A) tail length *in vitro*, as does bovine PABP2. These results show that *Drosophila* PABP2 has the same functions as bovine PABP2 *in vitro*. Functionality of the heterologous combination of mammalian and *Drosophila* proteins is also of interest with respect to the identification of interaction surfaces. Sequence comparison of PABP2 from *Drosophila* and mammals indicates that the protein is highly conserved throughout its entire length apart from the N-terminus. Three domains can be identified in the protein. The first domain is a 32-residue sequence which shows similarity to coiled-coil domains and is the most conserved region in the protein (81% identity between cow and *Drosophila*). This region could be involved in oligomerization of PABP2 (11), as similar coiled-coil domains are known to be involved in homotypic or heterotypic protein–protein interactions (23). Alternatively, this domain might mediate an interaction with PAP. The RNP-type RNA-binding domain shows 77% identity between the bovine and *Drosophila* proteins. The third domain, at the C-terminus of the protein, is rich in arginine residues, most of which are conserved between *Drosophila* and mammals. This region contributes to RNA binding (A.Nemeth, U.Kühn and E.Wahle, unpublished data). The lack of conservation of the N-terminal part of the protein in *Drosophila*, as well as the shorter size of this region (40 residues versus 115 residues in cow) indicates that this part of the protein is not essential for function of PABP2 *in vitro*. In mammals, this region contains an alanine stretch the expansion of which is responsible for oculopharyngeal muscular dystrophy in humans (15). This part of the protein, which is lacking in *Drosophila*, could have a role in PABP2 function *in vivo* in mammals.

Cloning of the *Pabp2* gene and analysis of its RNA pattern indicate that it produces at least four RNA species, the relative amounts of which vary at different developmental stages. Analysis of cDNAs reveals that these different RNA species do not result from alternative splicing but from utilization of alternative poly(A) sites. cDNAs we have isolated stop at five different positions within 480 nt in the 3'-UTR of *Pabp2*. These five poly(A) sites are preceded by poly(A) signals and their utilization has been confirmed by RNase protection assays (Fig. 2C and data not shown). Use of at least two alternative poly(A) sites has also been described for the mouse *PABII* gene (14). In mouse as in *Drosophila*, mRNAs produced by utilization of the different poly(A) sites encode the

same protein, but vary in their 3'-UTRs; this could influence their stability or translatability. Four stretches of oligo(A) are present in the 3'-UTR of *Pabp2*; this is reminiscent of the bovine *PABP2* 3'-UTR, which contains six stretches of oligo(A). These oligo(A) stretches are potential binding sites for PABP2. Their conservation in *Drosophila* reinforces the hypothesis that an autoregulatory loop could regulate PABP2 expression (11). In contrast, retention of the first and last introns of the *PABP2* message, which occurs in human (15) and in cow (11) and could also regulate the level of PABP2 protein, does not appear to be conserved in *Drosophila*.

Western blots and immunodetection of *Drosophila* PABP2 indicate that the protein is ubiquitously expressed in *Drosophila* and accumulates mostly in nuclei. Nuclear localization is in agreement with a role of PABP2 in polyadenylation and the presence of the protein in all tissues suggests that it is involved in polyadenylation of most mRNAs. Immunodetection experiments did not allow us to exclude the possibility that a low amount of the protein is present in the cytoplasm of somatic cells. A low level of PABP2 was also detected in the cytoplasm of HeLa cells by electron microscopy (30). The presence of PABP2 in the cytoplasm would corroborate recent findings suggesting that PABP2 shuttles between nucleus and cytoplasm (31; A.Calado and M.Carmo-Fonseca, personal communication).

The distribution of PABP2 during embryogenesis suggests that the protein is provided maternally to the embryo since the protein is present before zygotic transcription. This is not unexpected as the protein should be present at the onset of zygotic transcription for 3'-end processing of zygotic mRNAs. These experiments also reveal that PABP2 is quite abundant in the cytoplasm of early embryos where cytoplasmic polyadenylation takes place. Cytoplasmic elongation of maternal mRNA poly(A) tails allows the translational activation of these maternal mRNAs, which are stored as dormant untranslated mRNAs during oogenesis (32). In *Drosophila*, at least *bicoid* and *Toll* mRNAs have been shown to undergo this regulation which activates the production of Bicoid and Toll proteins during early embryogenesis, where they play a key role in the anterior and dorsal–ventral patterning systems, respectively (33,34). The presence of PABP2 in the cytoplasm of early embryos is consistent with a possible role of PABP2 in this regulatory process. To test this hypothesis and to analyze the role of PABP2 in poly(A) tail formation and in poly(A) tail length control *in vivo*, a genetic analysis of *Drosophila Pabp2* is in progress.

ACKNOWLEDGEMENTS

The work of A.N. and E.W. was initiated while they were at the Biozentrum, University of Basel. We are grateful to Diana Blank for expert technical assistance and Walter Keller for support. This work was supported by an ATIPE grant (no. 3) from the Centre National de la Recherche Scientifique and a grant from the Association pour la Recherche sur le Cancer (no. 1184 to H.M.B.), by the Deutsche Forschungsgemeinschaft (to E.W. and U.K.) and the Fonds der Chemischen Industrie (to E.W.) and by the Centre National de la Recherche Scientifique (UMR 9922), the Universities P. et M. Curie et D. Diderot and the Association pour la Recherche sur le Cancer (no. 9885 to M.S.). B.B. held an award from the Ministère de l'Enseignement Supérieur et de la Recherche.

REFERENCES

1. Colgan,D.F. and Manley,J.L. (1997) *Genes Dev.*, **11**, 2755–2766.
2. Keller,W. and Minvielle-Sebastia,L. (1997) *Curr. Opin. Cell Biol.*, **9**, 329–336.
3. Wahle,E. and Kühn,U. (1997) *Prog. Nucleic Acids Res. Mol. Biol.*, **57**, 41–71.
4. Wahle,E. and Ruegsegger,U. (1999) *FEMS Microbiol. Rev.*, **23**, 277–295.
5. Bienroth,S., Wahle,E., Suter-Crazzolara,C. and Keller,W. (1991) *J. Biol. Chem.*, **266**, 19768–19776.
6. Murthy,K.G.K. and Manley,J.L. (1992) *J. Biol. Chem.*, **267**, 14804–14811.
7. Wahle,E. (1991) *Cell*, **66**, 759–768.
8. Wahle,E., Lustig,A., Jenö,P. and Maurer,P. (1993) *J. Biol. Chem.*, **268**, 2937–2945.
9. Bienroth,S., Keller,W. and Wahle,E. (1993) *EMBO J.*, **12**, 585–594.
10. Wahle,E. (1995) *J. Biol. Chem.*, **270**, 2800–2808.
11. Nemeth,A., Krause,S., Blank,D., Jenny,A., Jenö,P., Lustig,A. and Wahle,E. (1995) *Nucleic Acids Res.*, **23**, 4034–4041.
12. Kenan,D.J., Query,C.C. and Keene,J.D. (1991) *Trends Biochem. Sci.*, **16**, 214–220.
13. Burd,C.G. and Dreyfuss,G. (1994) *Science*, **265**, 615–621.
14. Lee,Y.J., Lee,J., Yang,I.C., Hahn,Y., Lee,Y. and Chung,J.H. (1998) *Biochim. Biophys. Acta*, **1395**, 40–46.
15. Brais,B., Bouchard,J.-P., Xie,Y.-G., Rochefort,D.L., Chrétien,N., Tomé,F.M.S., Lafrenière,R.G., Rommens,J.M., Uyama,E., Nohira,O., Blumen,S., Korcyn,A.D., Heutink,P., Mathieu,J., Duranceau,A., Codère,F., Fardeau,M. and Rouleau,G.A. (1998) *Nature Genet.*, **18**, 164–167.
16. Lefrère,V., Vincent,A. and Amalric,F. (1990) *Gene*, **96**, 219–225.
17. Sachs,A.B. (1990) *Curr. Opin. Cell Biol.*, **2**, 1092–1098.
18. Dura,J.M., Randsholt,N.B., Deatrick,J., Erk,I., Santamaria,P., Freeman,J.D., Freeman,S.J., Weddell,D. and Brock,H.W. (1987) *Cell*, **51**, 829–839.
19. Brand,S.F., Pichoff,S., Noselli,S. and Bourbon,H.-M. (1995) *Gene*, **154**, 187–192.
20. Simonelig,M., Elliott,K., Mitchelson,A. and O'Hare,K. (1996) *Genetics*, **142**, 1225–1235.
21. Audibert,A. and Simonelig,M. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 14302–14307.
22. Martin,G. and Keller,W. (1996) *EMBO J.*, **15**, 2593–2603.
23. McLachlan,A.D. and Karn,J. (1982) *Nature*, **299**, 226–231.
24. Lupas,A., Dyke,M.V. and Stock,J. (1991) *Science*, **252**, 1162–1164.
25. Smith,J.J., Rucknagel,K.P., Schierhorn,A., Tang,J., Nemeth,A., Linder,M., Herschman,H.R. and Wahle,E. (1999) *J. Biol. Chem.*, **274**, 13229–13234.
26. Hirzmann,J., Luo,D., Hahnen,J. and Hobom,G. (1993) *Nucleic Acids Res.*, **21**, 3597–3598.
27. King,R.C. (1970) *Ovarian Development in Drosophila melanogaster*. Academic Press, New York, NY.
28. Foe,V.E., Odell,G.M. and Edgar,B.A. (1993) *The Development of Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 149–300.
29. Foe,V.E. (1989) *Development*, **107**, 1–22.
30. Krause,S., Fakan,S., Weis,K. and Wahle,E. (1994) *Exp. Cell Res.*, **214**, 75–82.
31. Chen,Z., Li,Y. and Krug,R.M. (1999) *EMBO J.*, **18**, 2273–2283.
32. Richter,J.D. (1996) In Hershey,J.W.B., Mathews,M. and Sonenberg,N. (eds), *Translational Control*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 481–503.
33. Salles,F.J., Lieberfarb,M.E., Wreden,C., Gergen,J.P. and Strickland,S. (1994) *Science*, **266**, 1996–1999.
34. Schisa,J.A. and Strickland,S. (1998) *Development*, **125**, 2995–3003.