4108–4115 *Nucleic Acids Research, 1998, Vol. 26, No. 18*

Specific binding of *Drosophila* nuclear protein PEP (protein on ecdysone puffs) to hsp70 DNA and RNA

Stefan Hamann and Wolf H. Strätling*

Institut für Physiologische Chemie, Universitäts-Krankenhaus Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

Received July 1, 1998; Revised and Accepted July 29, 1998

ABSTRACT

The *Drosophila* protein PEP (protein on ecdysone puffs), a component hnRNP complexes, was previously immunocytologically localized on *Drosophila* giant chromosomes to puffs induced by ecdysone and to some heat shock-induced puffs (e.g. at the hsp70 locus at 87A7). Here, PEP was purified to homogeneity and characterized in its DNA and RNA binding features with specific reference to the hsp70 locus. In south-western blotting assays, PEP was found to bind with high affinity to the hsp70 coding region, but not to a flanking region nor to the boundary elements scs and scs′, and non-specifically to the intergenic hsp70 SAR. In UV cross-linking assays, PEP binds with even higher affinity to hsp70 transcripts, but not to transcripts of a flanking region or of a nearby gene, *aurora*. Finally, competition experiments indicate that PEP recognizes specific sequences within hsp70 mRNA; in these sequences two distinct motifs were found to be enriched. In summary, our results suggest the recognition of specific transcripts as a molecular basis for the association of the protein with specific hnRNP complexes.

INTRODUCTION

RNA binding proteins are at least as diverse as their DNA binding counterparts with respect to biological functions, binding specificities and binding motifs. Some RNA binding proteins fulfil defined roles in developmental processes through binding to specific RNA elements. For instance, the *Drosophila* morphogenic proteins bicoid and pumilio act as translational repressors through binding to discrete target sequences in caudal and hunchback mRNA, respectively (1,2). The *Drosophila* female-specific RNA-binding protein sex-lethal (SXL) associates with the 5′ and 3′ untranslated regions (UTRs) of male-specific lethal-2 (msl2) mRNA and represses its translation in females, a process important for dosage compensation (3,4). In contrast, the role of most of the proteins, that associate with pre-mRNA [heterogeneous nuclear (hn)RNA] to form heterogeneous nuclear ribonucleoprotein (hnRNP) particles, is not as clear. They bind to RNA concomitantly during transcription (5) and are mostly restricted to the cell nucleus (6), though some have been shown to shuttle between the nucleus and the cytoplasm (7).

It has been proposed that hnRNPs package hnRNAs in a transcript-specific manner so that the processing events and transport processes treat each transcript as an individual entity (8). Theoretically this is achievable through the sequence-specific manner, in which hnRNPs bind RNA, and secondly, through a combinatorial assembly of the ~20 major and the many more minor hnRNPs. Earlier studies indicated that various hnRNPs have different relative affinities for ribohomopolymers *in vitro*. For instance, the hnRNP Ms bind avidly to poly(G) and poly(C) polymers (8), and hnRNPs F and H bind only to poly(G) (9). The hnRNP C prefers polypyrimidine stretches that are found at the 3′ end of most introns (10). Recently the sequence specificities of some hnRNPs have been determined in more detail. By selection/amplification from pools of random sequence RNA it was shown that hnRNP A1 binds most avidly to the sequence UAGGGA/U, which resembles the consensus sequences of 5′ and 3′ splice sites (11). The hnRNPs K and E1 bind to a control element in the 3′ UTR of 15-lipoxygenase mRNA, a process that silences this mRNA early during erythroid cell differentiation (12).

An increasing number of RNA binding proteins is reported to have high affinity to DNA as well. For instance, the hnRNP U binds to matrix/scaffold attachment regions (MAR/SARs) (13,14), that are thought to fasten chromosomal loops to the nuclear matrix (15) and were experimentally found to insulate transgene expression from position effects of the chromatin structure at the site of integration (16–19). The zinc finger protein MOK2, which is mainly associated with hnRNP complexes, recognizes a specific DNA sequence, suggesting a role in transcription (20). Members of the hnRNP D and E groups have been proposed to bind to chromosomal telomers (21). The hnRNP K, in addition to its role in hnRNP complexes, acts as a transcription factor; it binds to a *cis*-element (CT element) in specific promoters and interacts with the TATA box-binding protein (22). The *Drosophila* homeodomain protein bicoid transcriptionally activates target genes at different threshold concentrations (23). Furthermore, it binds caudal mRNA and acts as a translational repressor (1).

The *Drosophila* protein PEP (protein on ecdysone puffs) was originally identified through an immunological approach. It is associated preferentially with active ecdysone-inducible puffs on *Drosophila* polytene chromosomes and is furthermore found on some but not all heat shock-induced puffs, e.g. at locus 87A (24).

*To whom correspondence should be addressed. Tel: +49 40 4717 2392; Fax: +49 40 4717 4862; Email: straetli@plexus.uke.uni-hamburg.de

Chromosomal immunostaining of PEP *in situ* is RNase-sensitive (25). The 110 kDa protein co-purifies with hnRNP complexes immunoprecipitated with an antibody directed against a lower molecular weight hnRNP protein. Through a sequential immunostaining procedure, it was furthermore shown that PEP associates with a specific subset of active chromosomal sites. While these results and the presence of four zinc finger motifs in PEP suggest that an important signal for hnRNP deposition resides within RNA, the molecular basis of the deposition on specific transcripts is unknown. Here, we identified PEP through its ability to bind DNA. It preferentially recognizes the coding region of the *hsp70* genes at locus 87A7, that was previously decorated in giant chromosomes by anti-PEP antibodies after heat shock (24). Furthermore, it preferentially recognizes hsp70 transcripts and specific sequences within these. These results show that PEP is a sequence-specific DNA and RNA binding protein.

MATERIALS AND METHODS

Materials

Drosophila Kc cells (26) were grown in 75 cm^2 tissue culture flasks or in 1–3 l spinner culture flasks in D-22 insect medium (Sigma) supplemented with 5% fetal calf serum (Boehringer Mannheim), 100 U/ml penicillin and 100 µg/ml streptomycin at room temperature.

Plasmid constructs and probes

The 446 bp chicken lysozyme 5′ MAR fragment H1–*Hae*II and the 657 bp *Drosophila* histone SAR fragment *Hin*fI–*Eco*RI have been described (27,28). The 992 bp scs fragment *Pvu*II–*Pvu*II and the 500 bp scs′ fragment *Eco*RI–*Hin*cII derived from plasmid ELBA 6 (a gift of A. F. Stewart) and were subcloned into pBSIISK+ (29,30). Plasmids pKSaur1 and pGEM-Sgs4, containing cDNAs for *aurora* and Sgs-4, respectively, have been described (31,32). The hsp70 SAR fragment *Xba*I–*Bam*HI (1023 bp) and the hsp70 fragments A (*Xba*I–*Bgl*I, 632 bp), B (*Bgl*I–*Pst*I, 797 bp), C (*Pst*I–*Sal*I, 873 bp) and D (*Sal*I–*Bgl*I, ~650 bp) were obtained by appropriate digestion of plasmid 122X14 (33). Furthermore, the *Sal*I–*Sal*I fragment CO (2183 bp) containing nearly the complete *hsp70* gene, and fragments B, C, D′ (*Sal*I–*Xho*I, ~ 900 bp) and E (*Xho*I–*Sal*I, 486 bp), which derived from plasmid 122 (33), were subcloned into pBSIISK+. Figure 4 shows a map of the hsp70 locus 87A7 with the location of all DNA probes used, as well as the RNA probes synthesized *in vitro* after linearization of the relevant plasmids by appropriate restriction. Synthetic RNA oligomers were purchased from MWG-Biotech (Ebersberg).

Purification of PEP

Kc cell nuclei were prepared as described by Marzluff *et al*. (34). Nuclei were pre-extracted in buffer PE (20 mM Tris–HCl, pH 7.5, 100 mM NaCl, 0.5 mM $MgCl_2$, 0.5 mM phenylmethylsulfonyl fluoride) and extracted in buffer E (20 mM Tris–HCl, pH 7.5, 500 mM NaCl, 1 mM $MgCl_2$, 140 mM 2-mercaptoethanol). The extract was diluted to 250 mM NaCl with buffer D (20 mM Tris–HCl, pH 7.5, 1 mM $MgCl_2$, 140 mM 2-mercaptoethanol) and incubated with DEAE-cellulose in a batch procedure. The supernatant was applied to a P11 phosphocellulose column (10 ml), equilibrated in Puffer P [20 mM Tris–HCl, pH 7.5, 70 mM 2-mercaptoethanol, 0.05% (v/v) NP-40] containing 250 mM NaCl (P250). The column was washed with buffer P250 and eluted with a linear gradient from 250 to 900 mM NaCl in buffer P. Fractions containing PEP were pooled, diluted to 300 mM NaCl and loaded onto a heparin–Sepharose column, equilibrated in buffer P containing 300 mM NaCl (P300). The column was washed with buffer P300 and eluted in one step with 700 mM NaCl in buffer P. The volume of the eluate was reduced to 2.5 ml using a centriprep-10-concentrator (Amicon) and passed over a PD-10 column (Pharmacia). Protein was eluted with 3.5 ml of buffer M [20 mM MOPS, pH 7.5, 70 mM 2-mercaptoethanol, 0.05% (v/v) NP-40] containing 200 mM NaCl (M200). The eluate was loaded onto a Mono S HR5/5 column (Pharmacia), equilibrated in buffer M200. After washing the column with buffer M200, a linear gradient from 200 to 1000 mM NaCl in buffer M was applied. PEP eluted as a symmetrical peak at 450 mM NaCl. MAR binding activity was monitored throughout the purification by southwestern blotting assay using fragment H1–*Hae*II of the chicken lysozyme 5′ MAR as a probe (35).

Peptide sequencing

The Mono S fraction of PEP was blotted onto a nylon membrane and digested with trypsin as described previously (14). Resulting peptides were separated on a Vydac C4-RP microbore column by reverse-phase high performance liquid chromatography. Selected fractions were submitted to automated Edman degradation on an Applied Biosystems 473A protein sequencer.

UV cross-linking assay

Photoreactive ^{32}P-labeled RNA probes (specific radioactivity 300 000 c.p.m./ng) were transcribed *in vitro* in 20 µl reactions containing 150 ng of linearized plasmid, 2 µl NTP-mix (5 mM each of ATP, GTP and CTP, 100 µM UTP containing 10% 5-BrUTP), 2 µl 10× transcription buffer (0.4 M Tris–HCl, pH 8.0, 60 mM $MgCl_2$, 100 mM dithiothreitol, 20 mM spermidine), 1 µl RNasin (40 U/µl, Boehringer Mannheim), 50 µCi [α^{32}P]UTP (800 Ci/mmol, Hartmann Analytic), and 1 µl T7 RNA polymerase (10 U/µl, Boehringer Mannheim). Unlabeled competitor RNAs were synthesized in reactions of the same composition except containing 5 mM of each NTP, and purified by denaturing polyacrylamide gel electrophoresis. The samples were incubated at 37°C for 1 h. Template DNA was removed by digestion with 1 µl RNase-free DNase I (10 U/µl, Boehringer Mannheim) at 37°C for 15 min. After addition of H_2O to a final volume of 200 µl, samples were extracted once with phenol and once with phenol/chloroform. RNA was precipitated with 0.2 vol of 4 M LiCl and 2.5 vol of ethanol at –80°C, and washed RNA pellets were redissolved in 50 µl H_2O containing 1 µl RNasin. Binding reactions were set up in PCR-reaction tubes in the following order (total volume 50 µl): 10 µl 10× binding buffer [50 mM HEPES, pH 7.5, 15 mM $MgCl_2$, 0.2 M KCl, 25% (v/v) glycerol, 5 mM dithiothreitol], 3 µg protein (phosphocellulose fraction), 3 µg yeast tRNA as non-specific competitor, 0–20 µl non-labeled specific competitor RNA, and 0.5–1 ng radiolabeled RNA. Reactions were incubated at 22°C for 25 min. Samples were then irradiated in a UV Stratalinker using a total energy of 1.5 J/cm^2 at 254 nm. UV cross-linked samples were digested with 20 µg RNase A (Type XII-A, Sigma) at 37°C for 20 min. After addition of 12.5 µl sample buffer [10% (w/v) SDS, 321.5 mM Tris–HCl, pH 7.5, 50% (v/v) glycerol, 700 mM 2-mercaptoethanol, 0.12% (w/v) bromophenol blue] and incubation at 65°C for 5 min,

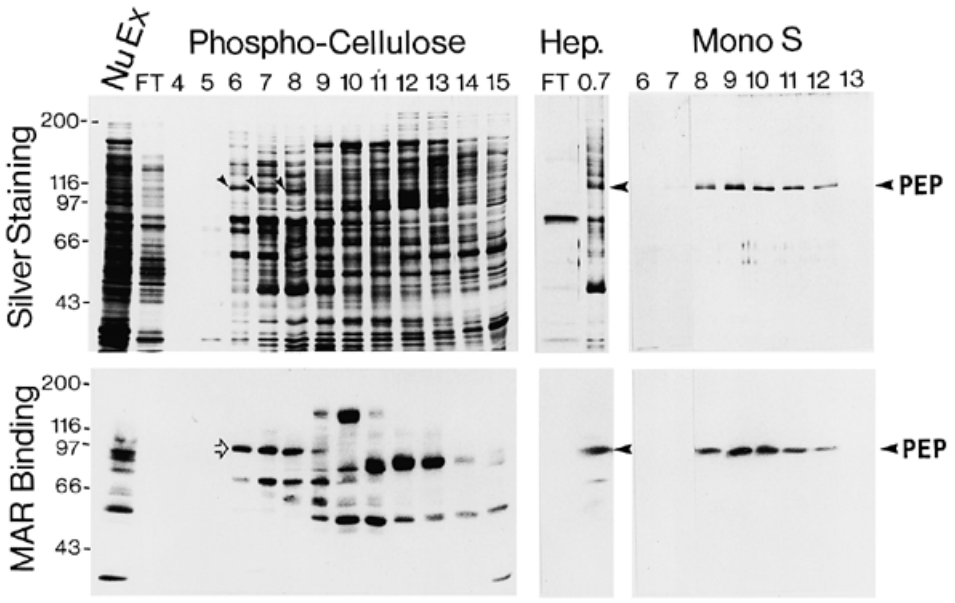

Figure 1. Purification of a 110 kDa MAR binding polypeptide (PEP). Protein patterns of fractions at different stages of purification are visualized by SDS–PAGE and silver staining (**top**). MAR binding activity was monitored using a southwestern blotting assay (**bottom**). A nuclear extract (NuEx) of *Drosophila* Kc cells was applied to P11 phosphocellulose and eluted with a gradient of 250–900 mM NaCl. Fractions 6–8, containing a 110 kDa MAR binding polypeptide (PEP), were loaded onto heparin–Sepharose (Hep.). After discarding the flow-through (FT), PEP was eluted with 700 mM NaCl (0.7). Final purification was achieved on Mono S (200–1000 mM NaCl gradient). PEP was detected in fractions 8–12. The large arrowheads and the open arrow indicate the position of PEP. The small arrowheads mark a stained band that probably represents PEP.

complexes were resolved on SDS–7% polyacrylamide gels, and dehydrated gels were exposed to X-ray film.

RESULTS

Purification and identification of PEP

A southwestern blotting assay was utilized to screen MAR/SAR binding proteins from *Drosophila* for a protein that might be related to previously characterized chicken MAR binding proteins (14,35). Following separation of a crude nuclear extract from Kc cells on an SDS–polyacrylamide gel and blotting onto a nitrocellulose filter, incubation with a labeled chicken lysozyme MAR probe indicated that multiple MAR binding proteins were present in the extract (Fig. 1, lane NuEx). Fractionation of the extract by chromatography on a phosphocellulose column resolved six polypeptides with prominent MAR binding activity (Fig. 1). A polypeptide with an apparent molecular mass of ~110 kDa, marked by an open arrow, was chosen for further study. Using fractions 6–8 of the phosphocellulose column, this polypeptide was purified to homogeneity by successive passages on heparin–Sepharose and Mono S columns. Silver stained gels and southwestern blots were performed following each step of purification (Fig. 1). We then separated the purified polypeptide on an SDS–polyacrylamide gel, excised the blotted 110 kDa band from the filter, digested it with trypsin, and resolved the resulting mixture of peptides by reverse-phase high performance liquid chromatography. Sequences of three internal peptides were determined. Comparison of these with the GenBank database showed them to be 100% identical to the deduced sequence of a previously cloned *Drosophila* protein, PEP (24,25). In detail, peptides T20 (PYASVPNDMFY), T14 (IDYDTHLLSAEHLK) and T8 (AAAPAAVASPAA) correspond to amino acids 206–216, 331–344 and 664–675, respectively, in PEP.

The recovery of PEP was estimated to be 2.7%, as shown in Table 1. From the yield and the number of cells extracted, we calculate an abundance of ~1.5 × 10^6 molecules of PEP in Kc cells. Thus PEP is a highly abundant nuclear protein in this cell type, as it is in *Drosophila* Schneider II cells, as previously reported (25). As early as after fractionation of the nuclear extract on phosphocellulose, PEP was identifiable as a distinct band in SDS–polyacrylamide gels (see the small arrowheads in Fig. 1); this furthermore testifies the great abundance of PEP.

Table 1. Purification of PEP

Fraction	Total protein (µg)	MAR binding activity[a] Total (fmol)	Specific (fmol/100 µg)	Yield (%)
Crude nuclear extract	15 120	255	1.7	(100)
Phosphocellulose	1 150	190	17.3	74
Heparin–Sepharose	294	102	37.6	40
Mono S	8	6.8	85	2.7

[a]Purification of PEP was quantitated through MAR binding activity in a southwestern blotting assay. Following incubation of blotted proteins with labeled MAR fragment H1–*Hae*II, the 110 kDa band was excised from the filter and the amount of bound radioactivity was measured by liquid scintillation counting. Binding activity is expressed in fmol of bound fragment H1–*Hae*II.

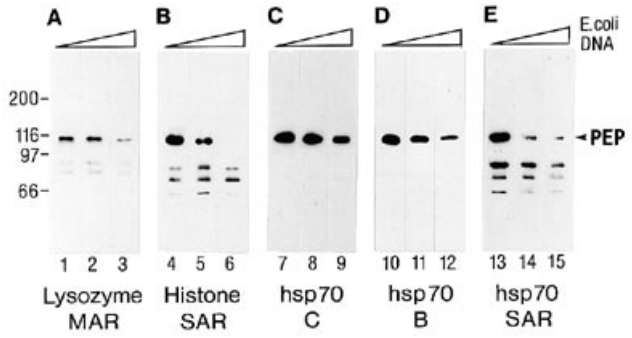

Figure 2. DNA binding features of PEP. A phosphocellulose fraction of PEP was separated on an SDS–polyacrylamide gel, blotted and incubated with the indicated labeled DNA probes in the presence of increasing concentrations (50, 100 and 200 µg/ml) of *E.coli* competitor DNA. The positions of PEP and marker proteins are shown.

DNA binding properties of PEP

Since we identified PEP through its ability to bind to a chicken MAR probe, we first analysed its DNA binding activity for various probes by southwestern assays utilizing a partially purified protein preparation (phosphocellulose fraction). In contrast to the use of a homogeneous protein preparation, this allows us to evaluate the specificity of the DNA binding activity of PEP in comparison to contaminating DNA binding proteins. Figure 2A shows that the chicken lysozyme MAR probe bound strongly to PEP and much weaker to contaminating polypeptides throughout the range of *Escherichia coli* competitor DNA concentration used (50–200 µg/ml). Two *Drosophila* SAR fragments, the histone SAR and the hsp70 SAR (28), also bound efficiently to PEP at 50 µg/ml of competitor DNA (Fig. 2B and E, lanes 4 and 13). However, binding to PEP drastically decreased at 100 and 200 µg/ml of competitor DNA to levels below those of binding to contaminating DNA binding polypeptides (lanes 5, 6, 14 and 15). These results were very reproducible, as shown e.g. by a southwestern assay using a different partially purified preparation of PEP. The hsp70 SAR probe bound to at least nine polypeptides in this preparation (Fig. 3, panel SAR). Thus we conclude that PEP does not specifically bind to the family of MAR/SAR elements.

In the southwestern assay shown in Figure 2 we also included two probes from the coding region of the hsp70 locus (see map in Fig. 4). To our surprise, these two probes, fragments B and C, strongly bound to solely PEP but not to any contaminating polypeptide (Fig. 2C and D, lanes 7 and 10). Furthermore, binding was stable at elevated competitor DNA concentrations (100 and 200 µg/ml; lanes 8, 9, 11 and 12). This prompted us to systematically monitor the affinity of PEP to various sequences of the hsp70 locus. Eight selected fragments of the locus were labeled to equimolar specificities (see map in Fig. 4) and subsequently incubated with blots of a partially purified phosphocellulose fraction. As mentioned above, the fragment containing the intergenic hsp70 SAR bound to a number of polypeptides in an apparently unspecific manner (Fig. 3). On the contrary, fragments A–D, containing coding hsp70 sequences solely bound to PEP, even at the lowest competitor DNA concentration (50 µg/ml). However, fragment E located 3′ of the *hsp70* gene had no affinity to PEP nor to any other polypeptide. Similarly, fragments containing the boundary elements scs and scs′ did not bind to PEP (29). Thus these southwestern assays reveal high affinity of PEP to preferentially the coding sequence of *hsp70*.

RNA binding activity of PEP

In giant chromosomes of heat-shocked larvae, PEP is associated with the cytological locus 87A containing the divergently transcribed *hsp70* genes at 87A7 (24). Furthermore, immunoprecipitation experiments had revealed that PEP is a component of a subset of hnRNP complexes (25). We therefore considered the possibility that PEP would specifically bind to hsp70 RNA. Radiolabeled transcripts CO were synthesized *in vitro* from the coding sequence of the *hsp70* gene (see map in Fig. 4), incubated with a partially purified preparation of PEP, cross-linked by UV-irradiation, and the complexes were resolved by SDS–PAGE. We detected a protein–RNA complex with an estimated molecular mass of ~110 kDa in the presence of excess unlabeled non-specific competitor RNA (Fig. 5, lane 2). Complex formation was completely abolished, when a 100-fold molar excess of unlabeled transcript CO was added, but not when the same molar excess of unlabeled transcripts from the 3′ located sequence E was added (lanes 3–6). Furthermore, an antisense CO transcript competed only weakly (data not shown). We conclude that a polypeptide of ~110 kDa, PEP, binds specifically to hsp70 transcripts *in vitro*.

To investigate whether PEP binds tighter to RNA or to DNA, we again used the cross-linking assay with labeled CO RNA as probe. While an excess of unlabeled CO RNA as low as 50-fold effectively competed the reaction (Fig. 6, lane 2), an excess of CO DNA as high as 200–400-fold was unable to compete (lane 7 and

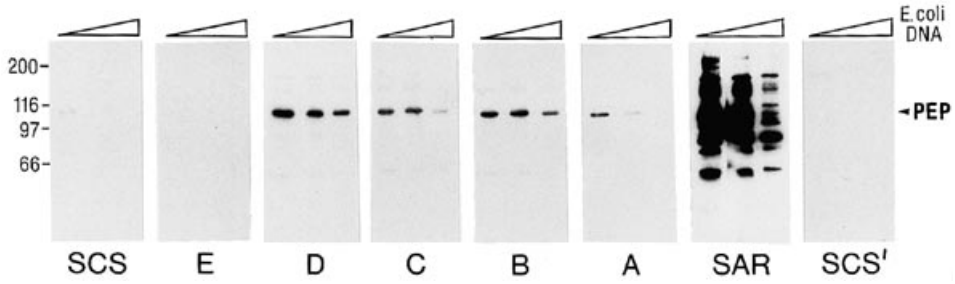

Figure 3. Specific binding of PEP to coding hsp70 fragments. A phosphocellulose fraction of PEP was electrophoretically separated, blotted and incubated with the indicated DNA probes as in Figure 2. The positions of PEP and marker proteins are shown.

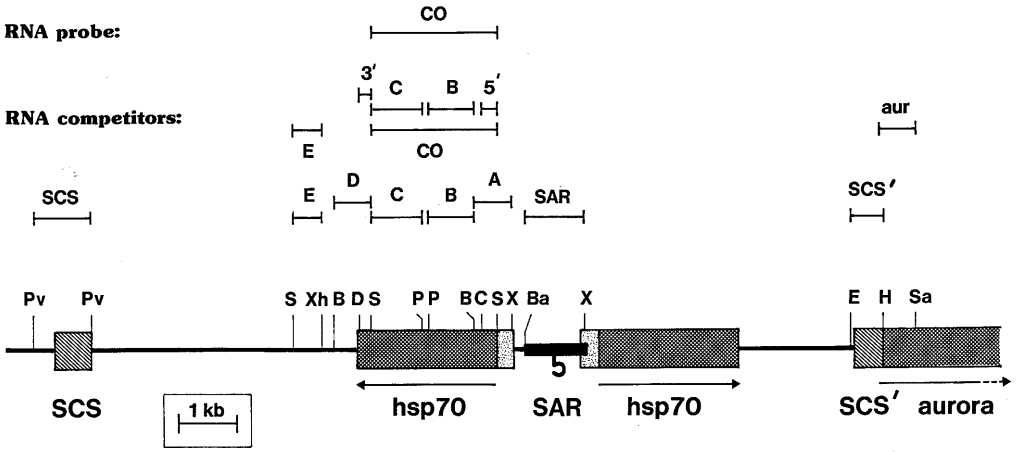

Figure 4. Schematic representation of the hsp70 locus 87A7 and the probes used. The locus contains two divergently transcribed *hsp70* genes. The coding sequences are densely stippled, while the promoter regions are lightly stippled. The spacer in between harbors a SAR element (bar). The locus is flanked by two sites of specialized chromatin structures, scs and scs′, (hatched bars) and furthermore contains the gene *aurora*. DNA probes used for southwestern blotting are depicted above the sequence, while the RNA probe and competitors used are shown in the uppermost part of the figure. Relevant restriction sites are: B, *Bgl*I; Ba, *Bam*HI; C, *Cla*I; D, *Dra*I; E, *Eco*RI; H, *Hin*cII; P, *Pst*I; Pv, *Pvu*II; S, *Sal*I; X, *Xba*I; Xh, *Hho*I.

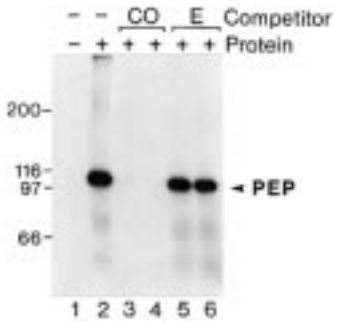

Figure 5. A UV cross-linking assay detects specific binding of PEP to hsp70 RNA. A phosphocellulose fraction of PEP was incubated with labeled transcript CO in the absence or presence of a 100-fold excess of the indicated competitor RNAs. The complexes were then UV cross-linked, digested with RNase, and separated on an SDS–polyacrylamide gel.

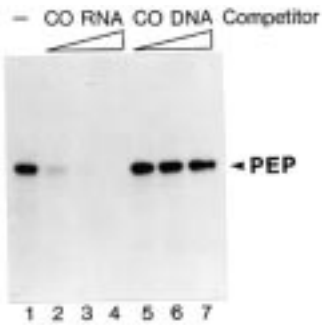

Figure 6. PEP binds tighter to hsp70 RNA than to DNA. PEP (phosphocellulose fraction) was incubated with labeled transcript CO in the absence or presence of increasing molar excesses (50-, 100- and 200-fold) of CO RNA or CO DNA, and analysed as in Figure 5.

data not shown). Densitometric scanning of the autoradiogram indicates that PEP has at least a 30-fold higher affinity to hsp70 RNA than to the coding DNA.

Since the cross-linking assays in Figures 5 and 6 utilized a partially purified preparation of PEP, we performed a reaction comparing this preparation (phosphocellulose fraction) with a highly purified preparation of PEP (Mono S fraction). Comparison of lanes 2 and 3 in Figure 7 shows that both preparations cross-linked CO RNA with equal efficiency. This verifies that PEP but not a contaminating polypeptide bound to hsp70 mRNA. We then investigated whether PEP binds to multiple sites within the hsp70 mRNA or to a single site, potentially present in the UTRs. Transcripts were prepared from the 5′ and 3′ UTRs and from the promoter-proximal and -distal halves (segments B and C, respectively) of the translated region, and used as competitors in a cross-linking assay with labeled transcript CO as probe (Fig. 7, lanes 4–11). PEP efficiently bound to the translated fragments B and C but weakly to the 5′ and 3′ UTRs, indicating that PEP does not belong to the group of RNA binding proteins that recognize specific sequences in UTRs (3,4,12,36).

Specificity of the RNA binding activity

Since it was previously concluded that PEP is a component of a subset of hnRNP complexes (25), it was important to study whether PEP would bind to any translated RNA sequence or display some binding specificity. In a cross-linking assay with labeled CO RNA as probe, transcripts B and C of the *hsp70* gene effectively competed the reaction (Fig. 8), confirming our previous results. In sharp contrast, a transcript from the gene *aurora*, which is located very close to the *hsp70* genes (see map in Fig. 4), was unable to compete. Furthermore, a transcript from the *Sgs-4* gene, encoding for one of the glue proteins that attach the pupa to the substrate during metamorphosis, competed only weakly (32,37). *Aurora* encodes a serine–threonine protein kinase necessary for centrosome separation and is transcribed

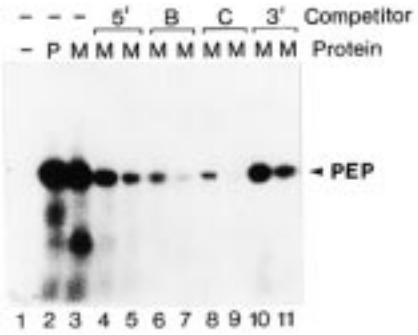

Figure 7. PEP preferentially binds to the coding region of hsp70 mRNA. A phosphocellulose (P) or Mono S (M) fraction of PEP was incubated with labeled transcript CO in the absence or presence of a 50- or 100-fold molar excess of the indicated competitor RNAs. Probes were analysed as in Figure 5.

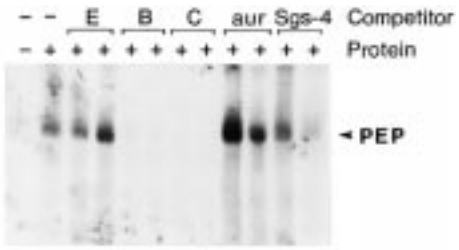

Figure 8. PEP preferentially binds hsp70 RNA. PEP (phosphocellulose fraction) was incubated with labeled transcript CO in the absence or presence of a 100- or 200-fold molar excess of the indicated competitor transcripts. Probes were then analysed as in Figure 5.

from a promoter, which colocalizes with scs′ (31). *Sgs-4* is encoded by the cytological locus 3C11, that was not detectable with anti-PEP antibodies by chromosomal immunostaining (25). Thus our results show that, in *in vitro* cross-linking assays, PEP binds with high affinity to hsp70 mRNA, but not to aurora mRNA, that is transcribed from the same locus. We conclude that PEP displays a surprisingly high degree of specificity in its RNA binding features. This is supported by the very weak binding of PEP to Sgs-4 mRNA. Figure 8 furthermore shows that transcripts from fragment E did not bind PEP (see also Fig. 4), again documenting that PEP does not bind any RNA.

Since the cross-linking assays showed that PEP binds to hsp70 mRNA with high specificity, we attempted to investigate whether it would also recognize much shorter fragments of hsp70 mRNA. We performed cross-linking reactions utilizing CO RNA as a probe and five selected RNA oligonucleotides as specific competitors. These oligonucleotides were chosen from hsp70 mRNA fragment C for their varying nucleotide composition (Table 2) and their varying potential to re-fold into secondary structures (not shown). The G-rich oligo 4, oligo 2, which is slightly enriched for A and C, and the A-rich oligo 3 effectively competed with the binding of PEP to CO RNA, while oligo 5, enriched for A, C and G, competed weakly (Table 2). On the other hand, oligo 1 did not compete at all. First of all, these results show that PEP can recognize specific sequences within hsp70 mRNA. Secondly, they give some insight into the determinants that govern the binding of PEP to hsp70 mRNA. The overall nucleotide composition, a major determinant for the RNA binding of many hnRNPs (8–10,38), does not seem to be a crucial factor for the recognition of RNA sequences by PEP. A high potential to form secondary structures might have some inhibitory influence, since the non-competing oligo 1 exhibited the highest capacity to build a loop–stem structure. Since we were thus unable to find a general feature that could satisfactorily explain the differential oligo binding ability of PEP, we looked for sequence motifs that might be contained in binding oligos 2–4 but not in weakly or non-binding oligos 1 and 5. We found two such motifs, GAU and GRRCG (R indicates a purine), that are emphasized in Table 2. Interestingly, these motifs are enriched in hsp70 mRNA fragments B and C, that efficiently bind PEP, but sparely present in the non-binding aurora mRNA and RNA fragment E (Table 3). They occur at an intermediate density in the weakly binding Sgs-4 mRNA. This seems to suggest that the frequency of these two motifs is a major determinant, among potentially others, in the binding of PEP to hsp70 mRNA.

DISCUSSION

Here we identified PEP in a screen for MAR/SAR binding proteins in *Drosophila* nuclear extracts by use of a southwestern blotting assay. However, several lines of evidence indicate that binding of PEP to MARs is of low affinity and of low specificity. First, binding of PEP to two *Drosophila* SARs, the histone SAR and the hsp70 SAR (28), drastically decreased, as the concentration of *E.coli* competitor DNA was raised. In contrast, binding of these SARs to contaminating polypeptides decreased much less. Thus at high concentrations of competitor DNA, the affinity of PEP to these SARs was lower than that of contaminating polypeptides. On the other hand, we found that PEP binds to the coding region of the *hsp70* gene with high preference and affinity. First, *E.coli* DNA competed binding to hsp70 DNA much less than binding to SARs. Second, solely PEP, but not any contaminating polypeptide in the phosphocellulose fraction, bound to hsp70 DNA. Third, three flanking fragments, scs, scs′ and fragment E, had no affinity at all to PEP (29,30).

In UV cross-linking assays, we furthermore show that PEP binds with high specificity to hsp70 transcripts. It does not bind a transcript of the *aurora* gene, that is located very close to the *hsp70* genes (31), nor a transcript of the flanking sequence E. Competition assays indicate that binding of PEP to hsp70 mRNA is at least 30-fold stronger than binding to hsp70 DNA. We thus conclude that PEP for its own can recognize hsp70 RNA. It was previously shown that PEP-containing hnRNP complexes assemble preferentially on transcripts of ecdysone-regulated genes and some but not all heat-shocked genes (24). Our results suggest that the molecular basis of this selectivity is the specific recognition of the RNA moieties. Yet it has to be considered that binding of PEP to specific sequences is probably modulated, when PEP is present in complexes with other hnRNPs. The *in vivo* organization of hnRNP particles and the proteins, which directly interact with PEP within the particles, possibly influence its RNA binding activity.

Most hnRNPs have been reported to recognize distinct features in their RNA targets. Human hnRNP Ms bind avidly to poly(G) and poly(U) homopolymers, while mammalian hnRNPs F and H bind only to poly(G) (8,9). Poly(C) homopolymer is tenaciously bound by mammalian hnRNP K/J (33). A defined high affinity

Table 2. Oligo competition assays

Oligo	Position	Sequence	Composition (%)				Inhibition of cross-linking at	
			U	A	C	G	200-fold excess	400-fold excess
1	1449-1481	CACUUUCAUUGGGAAUUGAGACCGCUGGAGGUG	27.2	21.2	18.1	33.3	0	6
2	1482-1511	UAAUGACCAAGCU**GAU**CGAGCGCAACUGCC	16.6	30.0	30.0	23.3	38	90
3	1749-1778	CCAAGAACAUCAC**GAU**CAAGAACGACAAGG	6.6	46.6	26.6	20.0	30	80
4	2140-2169	GGCGGGAG**GAU**UUGGAGGCUACUCUGGACG	20.0	16.6	16.6	46.6	43	100
5	1940-1669	GACGAGGCUGACAAGAACUCCGACUUGGAC	13.3	30	26.6	30.0	5	31

Cross-linking assays were performed with labeled CO RNA in the presence of 0, 200-, 300-, 400- and 500-fold weight excesses of oligos 1–5. Only those values obtained with 200- and 400-fold excesses of competitor are shown. They represent means of two experiments. Two sequence motifs, GAU (bold) and GRRCG (boxed), shared by competing oligos 2–4, are highlighted.

binding site has been reported for hnRNP A1 (11,21). Using five oligonucleotides as specific competitors in a cross-linking assay, we showed that PEP could recognize specific sequences within hsp70 mRNA. However, we were unable to identify any general sequence characteristic, such as nucleotide composition or the potential to fold into secondary structures, as the molecular basis for this recognition. Instead, we found that binding-competent oligos contain two sequence motifs, which are lacking in the non-binding or weakly binding oligos (Table 2). Furthermore, these motifs are enriched in hsp70 mRNA relative to three non-binding transcripts (Table 3). Although we have thus identified two motifs that might play a role in the recognition of hsp70 mRNA by PEP, we have to admit that an extension of the competition cross-linking assays by use of a greater number of oligos might refine and modify our present view of the sequence features recognized by PEP. Furthermore, it is likely that, within hnRNP particles, binding of PEP is influenced by the overall architecture of the complexes and by adjacent proteins. We also note that PEP most likely binds through its zinc fingers, and that zinc fingers in general are distinguished by a remarkably great propensity to recognize specific nucleic acid sequences (24,39).

Though binding of PEP to *hsp70* DNA proved to be at least 30-fold weaker than to hsp70 mRNA, the DNA binding ability might have a physiological role. Amero *et al.* (25) pointed out that PEP is present at some ecdysone-regulated sites prior to puff appearance at those sites. Thus it is possible that, before the onset of transcription, PEP is bound to the *hsp70* gene, and as soon as transcription starts, PEP jumps onto its mRNA target. This process would be analogous to a recently proposed reaction, that couples splicing and 3′-processing to transcription. The C-terminal domain of RNA polymerase II is thought to provide a platform for various pre-mRNA processing factors that associate with the nascent transcript as it emerges from the polymerase (40,41). Several nuclear proteins have been described, whose DNA as well as RNA binding ability is of high physiological significance. The most extensively studied example is the transcription factor IIIA (TFIIIA) from *Xenopus laevis*. TFIIIA binds to the internal control region of the 5S rRNA gene and forms a highly specific complex with the 5S rRNA itself (42,43). The zinc finger protein MOK2 is, like PEP, a component of hnRNP complexes and has the dual activity to bind distinct RNA polymers and to recognize a defined DNA sequence (20). The hnRNP U, another abundant component of hnRNP particles, has the ability to bind certain RNA homopolymers as well as MAR sequences (13,14).

Finally, we would like to raise the possibility that binding of PEP to specific transcripts influences gene expression. Many reports document that hnRNP proteins have diverse functions in the processing and stability of mRNAs. The hnRNPs C, F and M and the *Drosophila* hnRNPs hrp 45 and 48 have been involved in splicing (44–48). The group of hnRNP Ds (AUF1 protein family) recognizes A+U-rich elements in the 3′ UTR of mRNAs and promote their degration (36). Thus it is very likely that the association of PEP with specific hnRNP complexes affects processes like mRNA transport or stability in an as yet unknown fashion.

Table 3. Frequency of sequence motifs GAU and GRRCG in various transcripts

Transcript	Sequence motifs		Length (n)	Motifs/ 1000 n
	GAU	GRRCG		
hsp70 RNA fragment C	11	14	875	28.6
hsp70 RNA fragment B	10	6	798	20.1
Sgs-4 RNA	11	1	741	16.1
Transcript E	3	2	492	10.2
aurora RNA	4	1	514	9.7

ACKNOWLEDGEMENTS

We thank F. Buck for excellent peptide sequencing, M. Rehbein for help with cross-linking experiments, H. Saumweber for several supplies of Kc cells, and S. Beckendorf, W. J. Gehring, D. M. Glover and A. F. Stewart for generous gifts of plasmid DNA. This work was supported by grants to W.H.S. from the Deutsche Forschungsgemeinschaft. This article is based on the doctoral thesis of S. Hamann in the Faculty of Biology, University of Hamburg.

REFERENCES

1 Dubnau,J. and Struhl,G. (1996) *Nature*, **379**, 694–699.
2 Murata,Y. and Wharton,R.P. (1995) *Cell*, **80**, 747–756.
3 Kelley,R.L., Wang,J., Bell,L. and Kuroda,M.I. (1997) *Nature*, **387**, 195–199.
4 Bashaw,G.J. and Baker,B.S. (1997) *Cell*, **89**, 789–798.
5 Economidis,I.V. and Pederson,T. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 4296–4300.
6 Jones,K.E., Okamura,C.S. and Martin,T.E. (1980) *J. Cell. Biol*., **86**, 235–243.
7 Piñol-Roma,S. and Dreyfuss,G. (1992) *Nature*, **355**, 730–732.
8 Datar,K.V., Dreyfuss,G. and Swanson,M.S. (1993) *Nucleic Acids Res*., **21**, 439–446.
9 Matunis,M.J., Xing,J. and Dreyfuss,G. (1994) *Nucleic Acids Res*., **22**, 1059–1067.
10 Piñol-Roma,S. and Dreyfuss,G. (1993) *Mol. Cell. Biol*., **13**, 5762–5770.
11 Burd,C.G. and Dreyfuss,G. (1994) *EMBO J*., **13**, 1197–1204.
12 Ostareck,D.H., Ostareck-Lederer,A., Wilm,M., Thiele,B.J., Mann,M. and Hentze,M.W. (1997) Cell, **89**, 597–606.
13 Kiledjian,M. and Dreyfuss,G. (1992) *EMBO J*., **11**, 2655–2664.
14 von Kries,J.P., Buck,F. and Strätling,W.H. (1994) *Nucleic Acids Res*., **22**, 1215–1220.
15 Paulson,J.R. and Laemmli,U.K. (1977) *Cell*, **12**, 817–828.
16 Stief,A., Winter,D.M., Strätling,W.H. and Sippel,A.E. (1989) *Nature*, **341**, 343–345.
17 Phi-Van,L., von Kries,J.P., Ostertag,W. and Strätling,W.H. (1990) *Mol. Cell. Biol*., **10**, 2302–2307.
18 McKnight,R.A., Shamay,A., Sankaran,L., Wall,R.J. and Henninghausen,L. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 6943–6947.
19 Dietz,A., Kay,V., Schlake,T., Landsmann,J. and Bode,J. (1994) *Nucleic Acids Res*., **22**, 2744–2751.
20 Arranz,V., Harper,F., Florentin,Y., Puvion,E., Kress,M. and Ernoult-Lange,M. (1997) *Mol. Cell. Biol*., **17**, 2116–2126.
21 Ishikawa,F., Matunis,M., Dreyfuss,G. and Cech,T.R. (1993) *Mol. Cell. Biol*., **13**, 4301–4310.
22 Michelotti,E.F., Michelotti,G.A., Aronsohn,A.I. and Levens,D. (1996) *Mol. Cell. Biol*., **16**, 2350–2360.
23 Driever,W. and Nüsslein-Volhard,C. (1989) *Nature*, **337**, 138–143.
24 Amero,S.A., Elgin,S.C.R. and Beyer,A.L. (1991) *Genes Dev*., **5**, 188–200.
25 Amero,S.A., Matunis,M.J., Matunis,E.L., Hockensmith,J.W., Raychaudhuri,G. and Beyer,A.L. (1993) *Mol. Cell. Biol*., **13**, 5323–5330.
26 Echalier,G. and Ohanessian,A. (1970) *In Vitro*, **6**, 162–172.
27 von Kries,J.P., Phi-Van,L., Diekmann,S. and Strätling,W.H. (1990) *Nucleic Acids Res*., **18**, 3881–3885.
28 Mirkovitch,J., Mirault,M.-E. and Laemmli,U.K. (1984) *Cell*, **39**, 223–232.
29 Udvardy,A., Maine,E. and Schedl,P. (1985) *J. Mol. Biol*., **185**, 341–358.
30 Farkas,G. and Udvardy,A. (1992) *Nucleic Acids Res*., **20**, 2604.
31 Glover,D.M., Leibowitz,M.H., McLean,D.A. and Parry,H. (1995) *Cell*, **81**, 95–105.
32 Mougneau,E., von Seggern,D., Fowler,T., Rosenblatt,J., Jongens,T., Rogers,B., Gietzen,D. and Beckendorf,S.K. (1993) *Mol. Cell. Biol*., **13**, 184–195.
33 Goldschmidt-Clermont,M. (1980) *Nucleic Acids Res*., **8**, 235–252.
34 Marzluff,W.F., Murphy,E.C. and Huan,R.C.C. (1973) *Biochemistry*, **12**, 3440–3446.
35 von Kries,J.P., Buhrmester,H. and Strätling,W.H. (1991) *Cell*, **64**, 123–135.
36 Kiledjian,M., DeMaria,C.T., Brewer,G. and Novick,K. (1997) *Mol. Cell. Biol*., **17**, 4870–4876.
37 Beckendorf,S.K. and Kafatos,F.C. (1976) *Cell*, **9**, 365–373.
38 Matunis,M.J., Michael,W.M. and Dreyfuss,G. (1992) *Mol. Cell. Biol*., **12**, 164–171.
39 Greisman,H.A. and Pabo,C.O. (1997) *Science*, **275**, 657–661.
40 Cordon,J.L. and Patturajan,M. (1997) *Trends Biochem. Sci*., **22**, 413–416.
41 Steinmetz,E.J. (1997) *Cell*, **89**, 491–494.
42 Engelke,D.R., Ng,S.Y., Shastry,B.S. and Roeder,R.G. (1980) *Cell*, **19**, 717–728.
43 Pelham,H.R. and Brown,D.D. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 4170–4174.
44 Gamberi,C., Izaurralde,E., Beisel,C. and Mattaj,I.W. (1997) *Mol. Cell. Biol*., **17**, 2587–2597.
45 Choi,Y.D., Grabowski,P.J., Sharp,P.A. and Dreyfuss,G. (1986) *Science*, **231**, 1534–1539.
46 Gattoni,R., Mahé,D., Mahl,P., Fischer,N., Mattai,M.-G., Stevenin,J. and Fuchs,J.-P. (1996) *Nucleic Acids Res*., **24**, 2535–2542.
47 Alzhanova-Ericsson,A.T., Sun,X., Visa,N., Kiseleva,E., Wurtz,T. and Daneholt,B. (1996) *Genes Dev*., **10**, 2881–2893.
48 Hammond,L.E., Rudner,D.Z., Kanaar,R. and Rio,D.C. (1997) *Mol. Cell. Biol*., **17**, 7260–7267.