The EMBO Journal vol.10 no.2 pp.459-466, 1991

EAP, a highly conserved cellular protein associated with Epstein-Barr virus small RNAs (EBERs)

David P.W.Toczyski and Joan A.Steitz

Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, 333 Cedar Street, PO Box 3333, Yale University School of Medicine, New Haven, CT 06510, USA

Communicated by J.A.Steitz

Human B lymphocytes latently infected with Epstein-Barr virus (EBV) synthesize very large amounts (5×10^6/cell) of two small nuclear RNAs called EBERs (Epstein-Barr encoded RNAs). These RNAs are of unknown function and, like many RNA polymerase III (Pol III) transcripts, bind the La autoantigen. We have discovered that the EBERs also associate with a second highly abundant host-encoded protein designated EAP (EBER associated protein). Human EAP is a small (14 777 dalton, 128 amino acid) polypeptide that binds both EBER 1 and EBER 2. EAP is also found in association with one or both of two analogous virally-encoded RNAs found in baboon cells infected with herpesvirus papio (HVP). We have devised a purification procedure for EAP and have cloned its cDNA from a human placental cDNA library using amino acid sequence data and the polymerase chain reaction (PCR). The predicted amino acid sequence of EAP shows a strong resemblance (77% identity) to an endodermal, developmentally regulated sea urchin protein called 217 (Dolecki *et al.*, 1988). EAP contains a potential nuclear localization signal and a highly acidic carboxy terminus, but does not display marked similarity to any other RNA binding proteins.

Key words: EBERs/Epstein-Barr virus/La protein/RNA binding protein/small nuclear ribonucleoprotein

Introduction

Epstein-Barr virus (EBV) is a large (172 kb genome), transforming, lymphotrophic γ-herpesvirus associated with infectious mononucleosis, Burkitt's lymphoma and nasopharyngeal carcinoma. EBV readily enters a nonreplicative latent phase during which only very few viral genes are expressed (Klein, 1989), including both latent membrane proteins (LMPs) (Fennewald *et al.*, 1984; Laux *et al.*, 1988; Longnecker and Kieff, 1990) and a number of EBNAs (Epstein-Barr nuclear antigens) (Kieff and Liebowitz, 1990; Miller, 1990). In addition, two small RNAs, the EBERs (Epstein-Barr encoded RNAs), are transcribed in very large amounts (each at $\sim 5 \times 10^6$ per cell) (Lerner *et al.*, 1981a; van Santen *et al.*, 1981).

EBER 1 and 2 (166 and 172 nucleotides respectively) are transcribed by Pol III (RNA polymerase III) (Rosa *et al.*, 1981; Jat and Arrand, 1982) and have been shown to be nuclearly located by *in situ* hybridization (Howe and Steitz, 1986). The EBER genes are unique in that they employ both tRNA-like intragenic Pol III promoter elements (boxes A and B) and upstream promoter elements normally associated with genes transcribed by RNA polymerase II (Howe and Shu, 1989). Like most Pol III transcripts, EBERs are uncapped (Lerner *et al.*, 1981a) and end with a poly U tract (Rosa *et al.*, 1981) that is specifically recognized by the La autoantigen, a 47 kd phosphoprotein implicated in transcription termination (Mattioli and Reichlin, 1974; Lerner *et al.*, 1981b; Habets *et al.*, 1983; Pizer *et al.*, 1983; Chambers *et al.*, 1988; Gottlieb and Steitz, 1989a,b). Whereas most Pol III transcribed RNAs lose their association with La as a result of 3′ end trimming [e.g. tRNAs and 5S (Rinke and Steitz, 1982)], the EBERs remain stably and quantitatively bound (Howe and Shu, 1988).

A number of other viral small RNAs have also been shown to bind the La protein. Herpesvirus papio (HVP) encodes two small RNAs, HVP-1 and HVP-2, which are 83% and 65% similar to EBER 1 and EBER 2 respectively (Howe and Shu, 1988). The HVP RNAs are Pol III products which bind La and can be folded into stable secondary structures that closely resemble the empirically determined EBER structures (Glickman *et al.*, 1988). The VA (virus associated) RNAs of adenovirus are also transcribed by Pol III and associate with La (Lerner *et al.*, 1981b; van Eekelen *et al.*, 1982; Francoeur and Mathews, 1982). Finally, the leader RNAs of vesicular stomatitis virus (VSV) (which are transcribed by the VSV polymerase) have been shown to bind La (Kurilla and Keene, 1983).

Although it was previously believed that the EBER 1 and EBER 2 ribonucleoproteins (RNPs) contained only the La protein, we have found that a second protein is also present in these particles. This protein, called EAP (EBER associated protein), is a 14 777 dalton (128 amino acids) cellular protein that can bind to both EBER 1 and EBER 2. EAP exhibits an extremely high degree of similarity with a previously analyzed sea urchin protein called 217 (Fregien *et al.*, 1983; Dolecki *et al.*, 1988). Both its conservation and abundance suggest that EAP may play an important role in cellular metabolism.

Results

EAP is precipitable by anti-La antibodies only in the presence of EBERs

In uninfected B lymphocytes, a variety of cellular RNAs exist as La containing RNPs, including pre-tRNAs, pre-5S rRNA, and Ro RNAs, as well as other less abundant RNAs (Hendrick *et al.*, 1981; Rinke and Steitz, 1982). In EBV infected cells, however, the EBER RNPs represent the majority of the La containing RNPs (Howe and Shu, 1988). This permits the use of anti-La antibodies, produced by individuals with autoimmune disease (Provost, 1979), as a relatively specific probe for the EBER RNPs.

Seven cell lines were radiolabeled in culture with [^{35}S]amino acids (predominantly methionine) and the resulting cell sonicates were immunoprecipitated with either

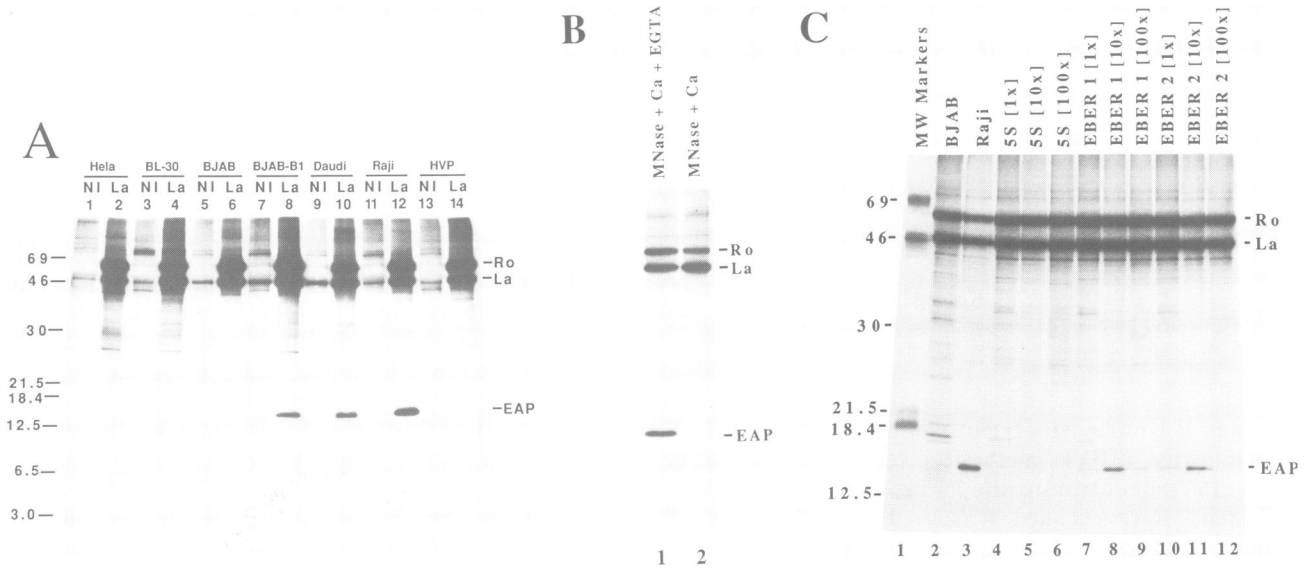

Fig. 1. EAP is precipitated by anti-La antibodies via its association with the EBERs or HVP RNAs. (**A**) Whole cell sonicates from ^{35}S-labeled cell lines were immunoprecipitated with either a non-immune human serum (NI) or patient anti-La serum (La). HeLa, BL-30 and BJAB (lanes 1–6) are EBV negative human lines; BJAB-B1, Daudi and Raji (lanes 7–12) are EBV positive human lines; HVP (lanes 13 and 14) are HVP positive baboon cells. (**B**) A ^{35}S-labeled Raji whole cell sonicate was incubated with micrococcal nuclease and 4 mM $CaCl_2$ in the presence (lane 1) or absence (lane 2) of 8 mM EGTA, followed by immunoprecipitation with anti-La antibodies. (**C**) A ^{35}S-labeled whole cell sonicate was made from either BJAB cells (lanes 2 and 4–12) or Raji cells (lane 3). Gel purified Raji cell 5S rRNA, EBER 1 or EBER 2 was incubated with BJAB sonicates (lanes 4–6, 7–9 and 10–12 respectively). The RNA was obtained from a number of Raji cells that was equal to ([1×]), 10 times greater than ([10×]) or 100 times greater than ([100×]) the number of BJAB cells used to make the sonicate. After 45 min at 25°C, the reactions were precipitated with anti-La antibody. No RNA was added to the sonicates shown in lanes 2 and 3. The mol. wts of the markers are given in kd. Reduced samples (containing β-mercaptoethanol) were run on 15% SDS–polyacrylamide gels, which were treated with Amplify (Amersham) and exposed to X-ray film.

anti-La (La) serum or human nonimmune (NI) serum (Figure 1A). The protein profiles of the immunoprecipitates from the three uninfected cell lines (HeLa, a human cervical carcinoma, and BL-30 and BJAB, two human EBV-negative B cell lymphomas) showed two predominant bands: the La protein (47 kd) and the Ro protein (60 kd) (lanes 2, 4 and 6). The latter polypeptide appears both because it coexists with the La protein in the Ro RNPs (Hendrick *et al.*, 1981) and because anti-La sera commonly contain anti-Ro antibodies as a dual specificity (Mattioli and Reichlin, 1974). Anti-La immunoprecipitates from three EBV-positive cell extracts, however, contain an additional protein with an apparent mol. wt of 17 kd (lanes 8, 10 and 12). These EBV-positive human B cell lines include two derived from individuals with Burkitt's lymphoma (Daudi and Raji) and one created by infecting the BJAB line with EBV (BJAB-B1). A band comigrating with the human protein is also present, albeit at lower levels, in anti-La immunoprecipitates from baboon B lymphocytes that have been infected with HVP and thereby contain HVP 1 and 2 (lane 14).

To determine whether the 17 kd band appears in the immunoprecipitates because of a direct association with La (in infected cells) or through its inclusion in a La containing RNP, we treated a Raji cell sonicate with micrococcal nuclease before exposure to anti-La antibodies. When this pretreatment was performed in the presence of free Ca^{2+}, which is required for micrococcal nuclease activity, the protein was no longer precipitated with anti-La antibodies (Figure 1B, compare lane 1 with lane 2). We conclude from this result that the 17 kd polypeptide is tethered to La only by virtue of each protein's independent association with a common RNA, most likely the EBERs. We have designated this protein EAP, for EBER associated protein.

The EBERs can be reconstituted with the La protein *in vitro* using gel-fractionated EBER 1 or EBER 2 and cell extract as a source of La (Rosa *et al.*, 1981). A protein which comigrates with EAP (Figure 1C, lane 3) becomes precipitable by anti-La antibodies when a ^{35}S-labeled cell sonicate from the EBV-negative cell line BJAB is incubated for 45 min at room temperature with gel-purified EBER 1 or EBER 2 (lanes 7–12), but not with 5S rRNA (lanes 4–6). Thus, EAP appears to bind both EBER 1 and EBER 2. Furthermore, since EAP is present in a cell line which does not contain the EBV genome, it must be a host-encoded protein which is expressed, at least to some extent, in uninfected cells. The variation in the amount of EAP precipitated from reconstitutions containing different amounts of EBERs (lanes 7–12) cannot be ascribed to the addition of inhibitors since La binding to gel-purified EBERs does not diminish at any RNA concentration used (data not shown). An alternative explanation for this phenomenon will be considered in the Discussion.

EAP copurifies with EBER RNPs

To obtain adequate EAP for amino acid sequence determination, we developed a purification protocol whereby the EBER RNPs are first isolated and then dissociated into their RNA and protein components. This strategy capitalizes on the distinct properties of EAP in two separate forms: as an RNP and as a free protein. Since no specific antibody or biological assay for EAP is yet available, it was not possible to assess yields or changes in specific activity during the purification procedure. Instead, peaks were pooled based upon the presence of the EBERs in all steps prior to the disruption of the EBER particles. Assuming that the ratio of EBER to EAP is one to one in an EBER RNP, we estimate

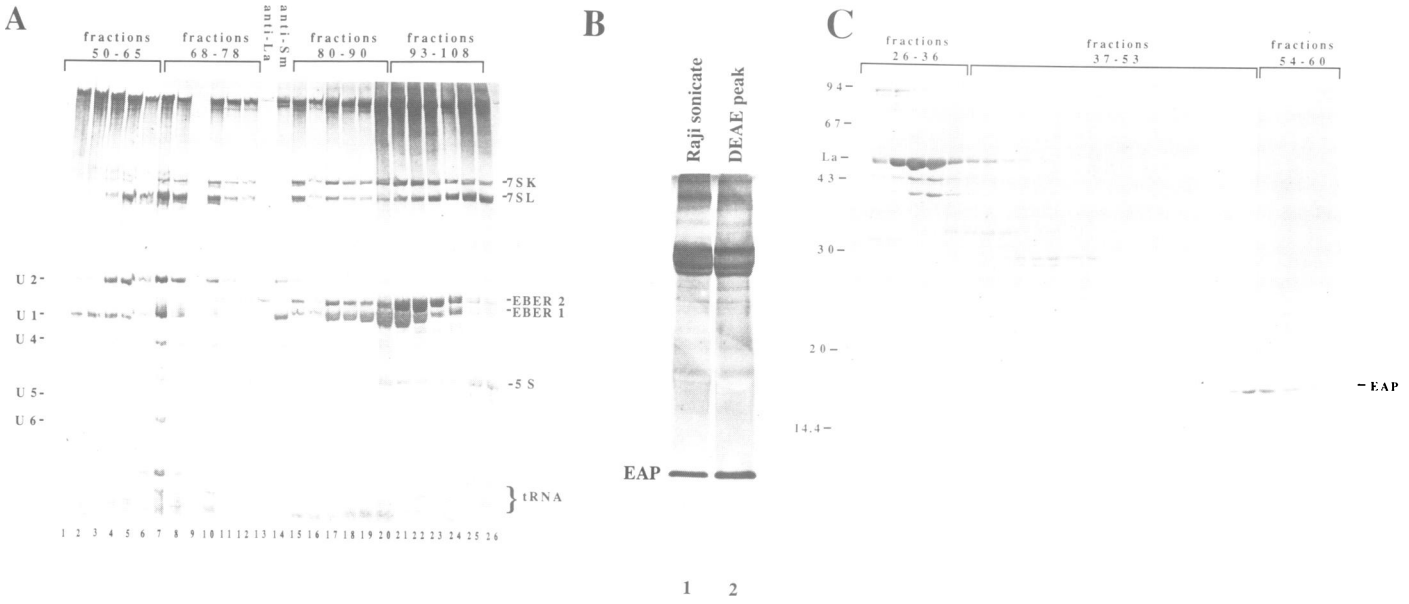

Fig. 2. Purification of EBER RNPs and EAP. (**A**) RNA from DEAE column fractions were run on a 5% polyacrylamide/8.3 M urea gel and silver stained (see Materials and methods). Spliceosomal snRNPs (containing U1, U2, U5 and U4/U6 RNAs) eluted predominantly in fractions 53–74 (lanes 2–10). The EBER RNPs peaked in fractions 84–102 (lanes 17–24) followed by the tRNAs in fractions 105–108 (lanes 25 and 26). Due to their negative staining in this particular experiment, the tRNAs (normally very dark) are not visible in the figure. Identification of the EBER and spliceosomal RNAs was based upon their comigration with RNAs immunoprecipitated with anti-La (for EBER 1 and EBER 2) or anti-Sm (for U1, U2, U4, U5 and U6) antibodies (lanes 13 and 14 respectively). 7SK, 7SL, 5S and the tRNAs were identified by mobility. Northern blotting was also performed in some experiments to distinguish EBER 1 from U1 RNA (data not shown). Lane 9 (fraction 72) was lost. (**B**) A limiting amount of anti-La antibody was used to immunoprecipitate a Raji whole cell sonicate (lane 1) and peak EBER containing DEAE fractions (e.g. fractions 84–102 in A) (lane 2) Samples were run on a 15% SDS–polyacrylamide gel and stained with silver (see Materials and methods). Silver stained bands representing La, Ro and the EBERs are obscured by the large amount of immunoglobulin heavy chain and cannot therefore be distinguished in the gel. (**C**) 5 μl from 0.8 ml fractions from a G-100 superfine sizing column were subjected to electrophoresis on a 13% SDS–polyacrylamide gel and stained with Coomassie blue. Even numbered fractions were analyzed for fractions 26–36 and 54–60; all fractions were analyzed for fractions 37–53. The bands seen in the peak La containing fractions (30–36) at 43 and 40 kd probably represent La breakdown products, which have been reported to occur (Stefano, 1984). Marker sizes are given in kd.

our final recovery of EBER-associated EAP to be 2–5%.

The starting material for the purification of EAP was a cytoplasmic extract from the Burkitt's lymphoma derived cell line Raji. Although the EBER RNPs are nuclear (Howe and Steitz, 1986), they leak readily into the cytoplasmic fraction upon aqueous fractionation (Hendrick *et al.*, 1981). Chromatography of the cytoplasmic extract on DEAE (Figure 2A) resulted in coelution of the two EBER particles at ~0.4 M NaCl, after the spliceosomal snRNPs (containing U1, U2, U4, U5 and U6 RNAs), but before the tRNAs. Because preliminary experiments (not shown) had revealed that the EBER RNPs are somewhat unstable at this salt concentration, two precautions were taken to avoid the disruption of the particles in this step: (i) the pH of the NaCl gradient was lowered to 6.7 so as to protonate the phosphates on the EBERs and thereby lower the salt concentration required to remove the RNPs from the column; and (ii) the fractions (2.5–3.0 ml) were collected into tubes which contained 1 ml of 25 mM Tris, pH 8.0, thereby immediately lowering the salt concentration. In some purifications, this DEAE fractionation was repeated (not shown) a second time using a narrower salt gradient (0.15–0.75 M NaCl; see Materials and methods).

To ensure that the EBER RNPs were indeed intact after the DEAE column, we immunoprecipitated peak EBER containing fractions with a limiting amount of anti-La antibody (Figure 2B, lane 2). Simultaneously, an equal amount of antibody was used to immunoprecipitate an excess of Raji cell sonicate (lane 1). These precipitates were analyzed in parallel on a polyacrylamide gel and stained with silver. The appearance of the EAP band in lane 2 shows that EAP remains associated with the fractionated EBER RNPs. The further observation that EAP is enriched in the immunoprecipitate from the DEAE-purified EBER RNP fractions (lane 2) relative to the cell sonicate (lane 1), which contains both free La and many different La RNPs, is consistent with our prior observations that EAP is not associated with free La protein or with non-EBER containing La RNPs. It should be noted that EAP stains with silver very darkly, which aids in its identification but also exaggerates its degree of purity.

The pooled EBER RNP fractions from the DEAE column (Figure 2A, fractions 81–104) were next fractionated on a 10–25% glycerol gradient (not shown). Based on the profiles of both the protein and RNA, the gradient fractions were pooled and concentrated. EBER RNPs were then disrupted in 2 M NaCl and fractionated by gel filtration on a superfine G-100 column. EAP eluted at ~15 kd as a relatively pure protein as judged by staining with Coomassie blue (Figure 2C, fractions 52–58).

Isolation and analysis of an EAP cDNA

Microsequencing of the N terminus (which was not blocked) and of cyanogen bromide–trypsin fragments (amino acids in bold; Figure 5) was carried out on purified EAP. Ten amino acids of N-terminal sequence had been previously obtained by immunoprecipitating a Raji cell sonicate with anti-La antibodies and transferring the immunoprecipitated

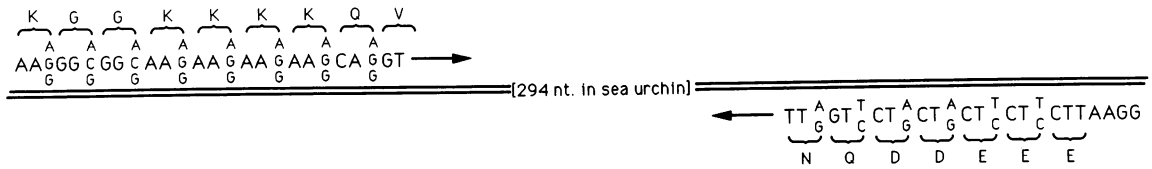

Fig. 3. Oligonucleotides used to PCR amplify EAP sequences from a human placental library. Two degenerate oligonucleotide probes were designed based upon peptide sequences from purified EAP and codon usage tables (Aota *et al.*, 1988). When two or three nucleotides are shown, they are equally represented in the mixture. Amino acids corresponding to each oligonucleotide are shown in the one letter code and represent amino acids 10–18 and 118–124. Knowledge of the sea urchin 217 protein sequence (Figure 5) allowed us to estimate the expected size of the PCR product (245 nucleotides). Original peptide sequencing predicted an aspartic acid at position 121 instead of a glutamic acid (as determined by the sequence of the cDNA clone; Figure 4).

protein from a polyacrylamide gel to a PVDF membrane, from which EAP was eluted and sequenced. This N-terminal sequence matched the sequence determined for the purified material, verifying that the isolated protein is the same as that in the anti-La immunoprecipitate.

When the partial amino acid sequences obtained from the purified protein were compared to the protein data base, we found that each of the fragments was either identical or very similar to amino acid sequences previously deduced for a 15 kd sea urchin protein designated 217 (Fregien *et al.*, 1983; Dolecki *et al.*, 1988). Based upon the expected locations of the sequenced EAP peptides extrapolated from the 217 cDNA sequence, we designed two degenerate oligonucleotides for use in polymerase chain reaction (PCR) amplification from a human placental cDNA library (Figure 3). A PCR product of the predicted size (345 bp) was the most prominent band present when high annealing temperatures were used (data not shown).

This PCR product was used to screen the same placental cDNA library for a full length cDNA clone. Figure 4 shows the EAP cDNA sequence determined from two overlapping clones. One of the clones isolated contained a single nucleotide deletion in a stretch of eight adenosines near the 5′ end (see legend to Figure 4). This mutation causes a frameshift that would lead to translation termination after only eight amino acids. Given that this mutation occurs in only one of the two clones (where all other overlapping sequences were identical) and that it resides within a stretch of eight adenosines, we suspect that it arose as a reverse transcription error which occurred during construction of the library.

The predicted translation product of the EAP cDNA contains nucleotide sequences corresponding to all of the sequenced peptides, the only discrepancy being that the glutamic acid at position 121 was originally determined to be an aspartic acid. The nucleotides in the vicinity of the initiation codon (GGTGCCATGG) (Figure 4) display a reasonably good match of the Kozak consensus sequence [GCC($^{A}/_{G}$)CCATGG] (Kozak, 1987). The initiating methionine appears to be post-translationally removed from the EAP polypeptide (see legend to Figure 5). One of the two cDNAs sequenced contained at its 3′ end a non-coding region extending >1 kb beyond the sequence shown in Figure 4 (data not shown). Since even this clone did not contain a polyadenylation signal (AAUAAA) followed by a poly(A) tail, the EAP message may possess an unusually long 3′ untranslated region.

Discussion

The fact that EBV employs only a small repertoire of gene products to transform B lymphocytes and maintain a latent infection suggests that the EBERs (by far the most abundant

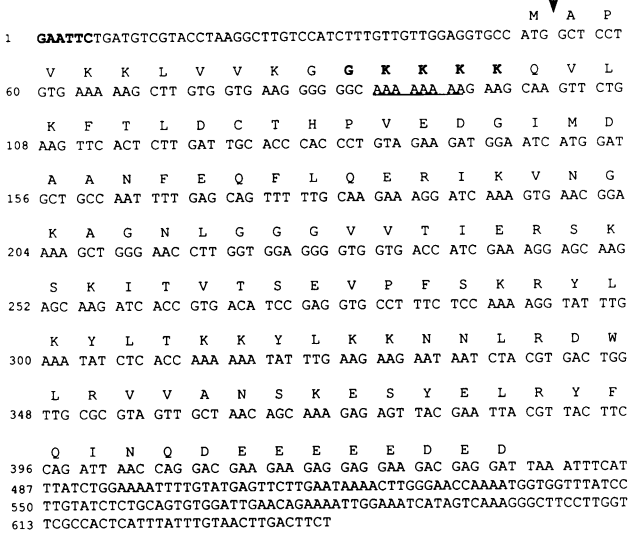

Fig. 4. The EAP cDNA sequence. Two distinct cDNA clones were obtained from a human placental library. The 5′ end sequence displayed begins with the *Eco*RI site of insertion into the λgt11 vector (nucleotides 1–6, in bold). The 3′ untranslated region shown is contained in both clones. A single nucleotide deletion in the eight adenosines (underlined nucleotides, 86–94) represents the only difference between the overlapping sections of the two cDNAs. The amino acids similar to the SV40 large T class of nuclear localization signals (Kalderon *et al.*, 1984; Moreland *et al.*, 1987) are shown in bold. The arrow drawn after the first methionine indicates a post-translational cleavage as determined by N-terminal protein sequencing, which showed alanine to be the first amino acid.

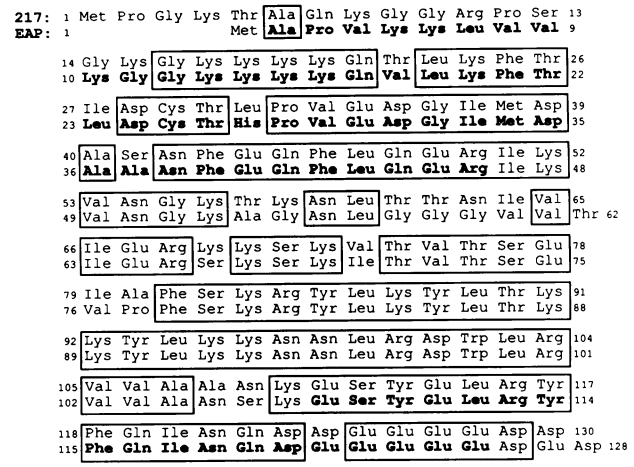

Fig. 5. Comparison of EAP and 217 protein sequences (Dolecki *et al.*, 1988). Boxed regions indicate sequence identity (77% overall). EAP amino acids shown in bold were determined by amino acid sequencing of either the entire protein (the N terminus was not blocked) (residues 2–39) or cyanogen bromide–trypsin fragments (amino acids 35–46, 108–113 and 114–125).

transcripts) play some vital role in one of these processes. To determine the function of these viral small RNAs, we set out to identify cellular macromolecules that interact with the EBER RNPs. In the course of this investigation, we found that the EBER RNPs contain not only the previously identified La protein, but also a 14 777 dalton protein, EAP. An EAP homolog in baboon cells likewise binds to one or both of the analogous HVP RNAs. Examination of the cloned cDNA encoding EAP showed it to be highly similar in amino acid sequence to 217, a protein deduced from a sea urchin cDNA isolated because of its developmental stage specific expression.

In vitro reconstitution experiments (Figure 1C) show that EAP binds specifically to both EBER 1 and EBER 2. Yet, the amount of EAP precipitated by anti-La antibodies in these reconstitutions (Figure 1C, lanes 7–9 and 10–12) was not proportional to the amount of gel-purified EBER added. Increasing the EBER concentration 10-fold resulted in the precipitation of ~10× as much EAP (compare lane 8 with 7 and lane 11 with 10), but a 100-fold increase in the amount of EBER precipitated approximately the same amount as did [1×] (compare lane 9 with 7 and lane 12 with 10). This is probably because when the concentration of EBER exceeds the concentration of available La or EAP, the two proteins reconstitute into separate particles containing only one protein each; since EAP is precipitable through its association with La in an EBER RNP, the amount of precipitable EAP decreases as excess RNA is added. These observations strongly support the hypothesis that the binding of EAP to EBERs is independent of La. The conclusion that EAP is itself an RNA binding protein is also consistent with the finding that EAP can be crosslinked to the EBERs with short-wave UV light (D.Toczyski, unpublished data).

Previous studies (Glickman *et al.*, 1988) identified sequences in the EBERs that are more protected from chemical and enzymatic modifying agents in the RNP than in the naked RNA. Except for regions adjacent to the 3′ poly U tract, which are presumed to be bound by La (Mathews and Francoeur, 1984; Stefano, 1984), no additional protection occurred in sequences conserved between the two EBERs and the two HVPs. Thus, identification of sequences or secondary structures in the EBERs that are recognized by EAP must await the availability of anti-EAP antibodies (now in preparation).

The ratio of EAP to La in the EBER and HVP RNPs is difficult to ascertain from the available data. The ^{35}S-labeling experiments (Figure 1) suggest that multiple EAP molecules may be present in each EBER RNP, since there are seven methionines in the La protein (Chambers *et al.*, 1988) and only one in EAP (Figure 5). However, the apparent ratio of the EAP and La bands varied from experiment to experiment and the turnover rates of the two proteins could be quite different. In contrast to Figure 1, the Coomassie stained gel of purified EBER RNPs (Figure 2C) suggests that there may be only one or two EAP molecules, assuming a single La protein, on each EBER. It also appears from the immunoprecipitation data that the HVP particles contain significantly less EAP than do the EBER particles (Figure 1A, compare lanes 8, 10 and 12 with lane 14). This is surprising given the sequence and potential secondary structure similarities between the EBERs and the HVP RNAs (Howe and Shu, 1988). The discrepancy could be caused by a difference in the methionine and cysteine content of the baboon and human EAP sequence. Alternatively, the binding of EAP to the HVPs may be weaker and therefore disrupted by the washing conditions used. It is also possible that EAP only associates with HVP 2, which is at a 10-fold lower concentration than HVP 1, EBER 1 or EBER 2. Further experiments will be required to distinguish between these possibilities and to determine the absolute EAP:RNA:La ratio for each viral RNP particle. In any case, it is apparent that EAP must be at least as abundant as the EBERs in infected cells.

The mechanism by which the EBERs achieve and maintain their nuclear localization is unknown. In all examples studied so far, protein components have been found to be required, although not necessarily sufficient, for the nuclear localization of RNPs (Mattaj, 1988). Although the La protein appears predominantly nuclear in indirect immunofluorescence studies (Deng *et al.*, 1981; Hendrick *et al.*, 1981; Smith *et al.*, 1985), it does exhibit significant cytoplasmic staining and is associated with several RNAs which are believed to reside primarily in the cytoplasm [the Ro's (Hendrick *et al.*, 1981) and the adenovirus VAI (Howe and Steitz, 1986)]. These observations suggest that association with the La protein does not, in and of itself, confer nuclear localization to the EBERs. EAP, on the other hand, contains a conserved cluster of four lysines near its N-terminal end (amino acids 13–16, Figure 5), which resembles the SV40 large T antigen class of nuclear localization signals in that a stretch of basic residues is preceded by a helix breaker (glycine or proline) (Kalderon *et al.*, 1984; Moreland *et al.*, 1987). The idea that EAP contains a nuclear localization signal is consistent with the observation that the VAs, which are believed to function in the cytoplasm (Kitajewski *et al.*, 1986), do not detectably bind EAP (D. Toczyski, unpublished data). Experiments are therefore under way to determine whether the EBERs are retained in (or are shuttled to) the nucleus through their association with EAP, much in the same way that the Sm proteins act, at least in part, to direct U RNAs to the nucleus (Mattaj and De Robertis, 1985).

The primary structure of EAP shows a number of highly charged regions which alternate in charge. Amino acids 5–20, which include the potential nuclear localization signal, comprise a basic domain containing eight lysines and no acidic residues (aspartic or glutamic acid) (Figure 5). Immediately following this is an acidic stretch (amino acids 24–45) containing six acidic, but no basic amino acids (lysines or arginines). A second basic region having 14 lysines and arginines and only two acidic residues extends from amino acid 46 to 97. Finally, the carboxy-terminal nine amino acids are entirely acidic. This stretch of acidic residues at the carboxy terminus is reminiscent of sequences seen in a number of nucleic acid binding proteins, such as the HMG proteins (Walker *et al.*, 1978; Watson and Dixon, 1981), and may constitute the EAP RNA binding domain. EAP does not, however, contain any other previously suggested nucleic acid binding motifs (Miller *et al.*, 1985; Adam *et al.*, 1986; Lazinski *et al.*, 1989; Milburn *et al.*, 1990).

The striking homology (77% identity) between EAP and the sea urchin 217 protein suggests that both are derived from a very old protein which may carry out some basic cellular function. The cDNA encoding the 217 protein was isolated from a pluteus cDNA library by selecting clones which were more abundant in the pluteus than in the blastula stage of development (Fregien *et al.*, 1983). Subsequent analyses showed the message encoding 217 to be enriched in the

endoderm and ~15× more abundant at the pluteus stage (the last stage examined) than in the blastula. As cautioned by the authors (Dolecki *et al.*, 1988), a number of genes isolated in this manner, such as the actin and tubulin genes, may not have a specific role in development. Preliminary analyses (D.Toczyski, unpublished data) comparing the abundance of the EAP message in a variety of human cell lines do not provide evidence for tissue specificity of the EAP message in humans.

Knowledge of the proteins of the EBER RNP particle will hopefully aid in deciphering the contribution of the EBERs to the EBV life cycle. Most virus encoded small RNAs can be assigned to pre-existing cellular RNP classes, which contain common protein constituents and carry out related functions (Birnstiel, 1988). For example, five recently discovered small RNAs (the HSURs) encoded by herpesvirus saimiri belong to the Sm class (Lee *et al.*, 1988; Wassarman *et al.*, 1989), which includes RNPs involved in mRNA biogenesis. Hence, the HSURs seem most likely to participate in some aspect of pre-mRNA modification. There are also viral small RNAs that form particles of the La class, which includes many host encoded RNAs that play roles in protein synthesis [e.g. tRNAs and 5S rRNA (Rinke and Steitz, 1982)]. Viral La RNAs include the EBERs of EBV, the HVP RNAs of HVP, the two VA RNAs of adenovirus, and potentially a small RNA encoded by the human cytomegalovirus [in that it is a Pol III transcript (Marschalek *et al.*, 1989)].

VAI RNA is the only viral small RNA to which a function has been ascribed. This RNA impedes an interferon-induced block to translation by inhibiting the dsRNA-dependent autophosphorylation of DAI, a kinase which phosphorylates and thereby inactivates the translation initiation factor eIF-2 (Thimmappaya *et al.*, 1982; Schneider *et al.*, 1984; Reichel *et al.*, 1985; Siekierka *et al.*, 1985; Kitajewski *et al.*, 1986; O'Malley *et al.*, 1986; Katze *et al.*, 1987). Similarities between the EBERs and the VA RNAs have led to the suggestion that they have similar functions (Lerner *et al.*, 1981a; Rosa *et al.*, 1981). Both the VAs and the EBERs are uncapped, abundant Pol III transcripts that associate permanently with the La protein (Reich *et al.*, 1966; Söderlund *et al.*, 1976; Lerner *et al.*, 1981a,b). Moreover, it has been observed that addition of the EBER genes can partially rescue the growth of adenovirus strains from which the VAI gene has been deleted (Bhat and Thimmappaya, 1985). *In vitro* results indicate that the EBERs, like VAI, can block the autophosphorylation of DAI, but higher EBER concentrations are required to achieve the same degree of inhibition (Garrett, 1990). This is difficult to reconcile with the fact that the abundance of the EBERs (~5 × 10^6/cell) in EBV-transformed cells is 10× lower than that of VAI (~10^8/cell) (Reich *et al.*, 1966; Söderlund *et al.*, 1976) in adenovirus-infected cells. Moreover, the EBERs are nuclear whereas VAI is predominantly cytoplasmic (Howe and Steitz, 1986). A function for the EBERs in blocking some aspect of interferon action would, none the less, be consistent with the recent finding that the EBER genes can be deleted from the EBV genome without disabling the virus's ability to transform B lymphocytes *in vitro* (where the host anti-viral response is attenuated) (Swaminathan *et al.*, 1991).

Just as VAI functions by binding DAI kinase (Siekierka *et al.*, 1985; Kitajewski *et al.*, 1986; O'Malley *et al.*, 1986; Katze *et al.*, 1987; Mellits *et al.*, 1990), it is conceivable that the EBERs act via their association with EAP. Given the high concentration of EBERs in EBV infected cells, it may be that EAP is not free to associate with its normal cellular ligand (probably an RNA). This could dramatically influence whatever cellular process this EAP–host RNA complex is involved in. Determining which nucleic acid associates with EAP in uninfected cells will be a key focus of further studies aimed at elucidating EBER function.

Materials and methods

Cells and general techniques

All cells were grown in RPMI 1640 medium (GIBCO) supplemented with 10% (unless noted otherwise) fetal bovine serum (GIBCO), 60 μg/ml penicillin, 100 μg/ml streptomycin and 300 μg/ml glutamine. SDS and 8.3 M urea–polyacrylamide gels, agarose gels, PCA (phenol/chloroform/isoamyl alcohol in a ratio of 25:25:1) extractions, ethanol precipitations, T7 RNA polymerase transcriptions, Coomassie staining and cloning techniques were performed as described in Ausubel *et al.* (1989). RNA gels were stained with silver using the procedure of Merril *et al.* (1981). Protein gels were stained with silver using a Rapid Ag-Stain kit purchased from ICN Radiochemicals.

In vivo labeling and immunoprecipitations

Approximately 4 × 10^6 cells were labeled with 100 μCi of ^{35}S (trans-label; ICN Radiochemicals) for ~18 h in MEM minimal medium (Sigma) supplemented as was RPMI 1640. Whole cell sonicates were prepared as previously described (Lerner *et al.*, 1981b). Immunoprecipitations with human anti-La sera (kindly provided by M.Mamula and J.Hardin) or normal human sera (cheerfully volunteered by D.Black and S.Siewert) were performed as described by Lee *et al.* (1988) except that pellets were washed 7× with NET-2 (150 mM NaCl, 50 mM Tris, pH 7.5, 0.05% NP-40).

EBER RNP reconstitution

Raji cells (10^8) were harvested by spinning for 5 min in a Beckman JS-4.2 rotor at 2000 r.p.m. (1100 *g*) at 4°C and resuspended in 5.0 ml of buffer containing 10 mM Tris, pH 7.5, 10 mM NaCl and 3 mM $MgCl_2$. After a 5 min incubation on ice, NP-40 was added to 0.5% and the cells were vortexed vigorously. NaCl (from a 4 M stock) was added to 150 mM and the cells were again incubated on ice for 5 min. The nuclei were removed by centrifugation for 15 min at 5000 r.p.m. (2580 *g*) in a Sorvall SA-600 rotor. The supernatant was then incubated with 1% SDS and 0.5 mg/ml proteinase K (Beckman) at 65°C for 15 min. After PCA extraction and ethanol precipitation, the RNA was run on a 5% denaturing polyacrylamide gel. The gel was stained with EtBr and the EBER 1, EBER 2 and 5S rRNA bands were excised and eluted overnight in pH 7.5 buffer containing 0.5 M NaAc, 10 mM EDTA, 10 mM Tris, 0.1% SDS. Each RNA was ethanol precipitated, resuspended in 200 μl of renaturation buffer (10 mM Tris, pH 7.5, and 5 mM $MgCl_2$) and heated to 65°C for 5 min followed by slow cooling. 150 μl, 15 μl or 1.5 μl of these RNAs (the equivalent of 7.5 × 10^7, 7.5 × 10^6 or 7.5 × 10^5 Raji cells respectively) were added in a volume of 150 μl of renaturation buffer to 250 μl of ^{35}S-labeled BJAB whole cell sonicate (representing 7.5 × 10^5 cells) and incubated at room temperature for 45 min with gentle agitation.

Micrococcal nuclease treatment

^{35}S-labeled whole cell sonicates were incubated with 4 mM $CaCl_2$ and 1 U/μl of micrococcal nuclease (Cooper Biomedical) for 15 min at 37°C. EGTA was added (to a final concentration of 8 mM) either before (Figure 1B, lane 1) or after (lane 2) this incubation.

Cloning and PCR

Oligonucleotides for sequencing and PCR were synthesized by J.Flory on an Applied Biosystems DNA Synthesizer. 1.5 μg of each of these oligonucleotides was used in a 100 μl PCR reaction containing 0.5 μl of 2% gelatin, 10 μl 2 mM dNTPs, 1.0 μl 1 M Tris, pH 8.4, 1.5 μl 100 mM $MgCl_2$, 2.5 μl 2 M KCl, 2.5 units of *Taq* DNA polymerase (Perkin Elmer Cetus) and 0.15 μg of λgt11 DNA. The λgt11 human placental cDNA library was kindly provided by S.Baserga. Samples were initially denatured by heating for 6 min at 94°C, subjected to 30 incubation cycles (94°C for 2 min; 55°C for 3 min; 72°C for 2 min) and finally incubated at 72°C for 6 min.

The PCR products were digested with *Eco*RI and, after gel fractionation, the 245 bp band was excised and cloned into pGem3Z between the *Eco*RI site and the *Pvu*II site. This construct allowed the production of a T7

transcribed RNA probe specific for EAP. 10^6 c.p.m. of this probe was used to screen the same λgt11 library (SSC protocol, unit 6.4; Ausubel *et al.*, 1987). Twelve plates containing 30 000 plaques/plate were screened in duplicate at 58°C overnight. Filters were washed (25°C, 6 × SSC, 1 h; 25°C, 0.5 × SSC, 1 h; 65°C, 0.5 × SSC, 1 h) and exposed to X-ray film for 2 days. DNA sequencing was performed using a Sequenase kit (US Biochemicals). All DNA sequences shown represent analysis of both strands.

Protein purification and amino acid sequencing

One liter of Raji cells was grown in T-flasks to a density of 5×10^5 cells/ml, transferred to 3 l spinner bottles, and split daily to 2×10^5 cells/ml with medium containing 8% fetal bovine serum. Cells were harvested (typically 25–30 l of cells at 7.5×10^5 cells/ml) by centrifugation for 10 min at 2000 r.p.m. (1100 *g*) in a Beckman JS-4.2 rotor. Cells were washed and resuspended (5 ml/l of cells) in pH 7.5 DE buffer (25 mM Tris, 3.0 mM $MgCl_2$, 0.1 mM EDTA, 1.0 mM DTT, 0.5% NP-40, 0.5 mM PMSF/0.5% isopropanol) with 50 mM NaCl. After homogenizing 10–15× with a Kontes Dounce homogenizer (pestle B), the nuclei were removed by centrifugation in a Sorvall SS-34 rotor for 15 min at 10 000 r.p.m. (10 300 *g*). If the supernatant was to be stored for later use it was brought to 20% glycerol and kept at −80°C. The frozen or fresh supernatant was then spun in a Beckman SW41 rotor at 4°C for 8–12 h at 40 000 r.p.m. (200 000 *g*). The supernatant from this spin was loaded onto an 80 ml column containing DE-52 resin (Whatman) previously equilibrated with pH 8.0 DE buffer. After loading, the column was washed with pH 8.0 DE buffer until the OD_{280} of the effluent was <0.1. A 400 ml linear gradient from 0 to 1.0 mM NaCl in pH 6.7 DE buffer containing 10 mM PIPES was then used to elute the bound RNPs. Fractions (2.5–3.0 ml) were collected into tubes which contained 1.0 ml of 25 mM Tris, pH 8.0. Selected fractions (25 μl) were treated with proteinase K (as described above), extracted twice with PCA, run on a 8.3% urea 5% polyacrylamide gel and stained with silver. Fractions were pooled based upon the position of the EBERs, which eluted at ~0.4 M salt as determined by conductivity (CDM 80; Radiometer America). An effort was made not to include early EBER fractions since they contained a much higher proportion of contaminating proteins than later EBER fractions. In some experiments, pooled fractions were diluted with an equal volume of water and reloaded onto a second DE-52 column (15–20 ml) and eluted as before but using a 0.15–0.75 M NaCl gradient. Pooled peak fractions were diluted 2-fold with water and concentrated to 2.0 ml on a Centriprep C-30 (Amicon). This material was loaded onto a 38 ml 10–25% glycerol gradient containing 25 mM Tris, pH 7.5, 3 mM MgCl and 1 mM DTT. After centrifugation in a SW28 rotor at 28 000 r.p.m. (103 000 *g*) for ~40 h, 1.2 ml fractions were collected. EBER RNP containing fractions were pooled as described in the text, diluted 2-fold with water, and concentrated on a C-10 Centricon apparatus (Amicon). On reaching 0.5 ml, one volume of 4 M NaCl, 10 mM EDTA, 50 mM Tris, pH 7.5, was added to the concentrated material and this mixture was gently agitated for 45 min at room temperature. The disrupted particles were then loaded onto a G-100 superfine (Pharmacia) sizing column (diameter 1 cm, length 1 m) previously equilibrated with 25 mM Tris, pH 7.5, 2.0 M NaCl, 1.0 mM EDTA. Fractions (0.8 ml) were collected at a rate of 5 fractions/h and analyzed on SDS–polyacrylamide gels. Purified protein was transferred to a PVDF membrane (Millipore) (Matsudaira, 1987). Subsequent cyanogen bromide/trypsin cleavage and amino acid sequencing were performed by the Yale University School of Medicine Protein and Nucleic Acid Chemistry Facility.

Acknowledgements

We thank members of the Steitz laboratory for protocols, materials and discussions, especially J.G.Howe and M.-D.Shu for cell lines and plasmids and S.Baserga for advice and reagents for PCR and DNA sequencing. E.Grogan (Yale University) kindly supplied the BL-30 cell line. We also thank E.Kieff, D.Wassarman, E.Sontheimer and V.Myer for critical reading of the manuscript and E.Garrett, T.Shenk, S.Swaminathan, B.Tomkinson and E.Kieff for sharing unpublished data. This work was supported by NIH grant CA 16038.

References

Adam,S.A., Nakagawa,T., Swanson,M.S., Woodruff,T.K. and Dreyfuss,G. (1986) *Mol. Cell. Biol.*, **6**, 2932–2943.

Aota,S., Gojobori,T., Ishibashi,F., Maruyama,T. and Ikemura,T. (1988) *Nucleic Acids Res.*, **16**, r315–r402.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1989) *Current Protocols in Molecular Biology*. Wiley, New York.

Bhat,R.A. and Thimmappaya,B. (1985) *J. Virol.*, **56**, 750–756.

Birnstiel,M.L. (ed.) (1988) *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer-Verlag, Berlin.

Chambers,J.C., Kenan,D., Martin,B.J. and Keene,J.D. (1988) *J. Biol. Chem.*, **263**, 18043–18051.

Deng,J.S., Takasaki,Y. and Tan,E.M. (1981) *J. Cell Biol.*, **91**, 654–660.

Dolecki,G.J., Lum,R. and Humphreys,T. (1988) *DNA*, **7**, 637–643.

Fennewald,S., van Santen,V. and Kieff,E. (1984) *J. Virol.*, **51**, 411–419.

Francoeur,A.M. and Mathews,M.B. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6772–6776.

Fregien,N., Dolecki,G.J., Mandel,M. and Humphreys,T. (1983) *Mol. Cell. Biol.*, **3**, 1021–1031.

Garrett,E.D. (1990) Ph.D. Thesis, Princeton University, Princeton, NJ.

Glickman,J., Howe,J.G. and Steitz,J.A. (1988) *J. Virol.*, **62**, 902–911.

Gottlieb,E. and Steitz,J.A. (1989a) *EMBO J.*, **8**, 841–850.

Gottlieb,E. and Steitz,J.A. (1989b) *EMBO J.*, **8**, 851–861.

Habets,W.J., den Brok,J.H., Boerbooms,A.M.Th., van der Putte,L.B.A. and van Venrooij,W.J. (1983) *EMBO J.*, **2**, 1625–1631.

Hendrick,J.P., Wolin,S.L., Rinke,J., Lerner,M.R. and Steitz,J.A. (1981) *Mol. Cell. Biol.*, **1**, 1138–1149.

Howe,J.G. and Shu,M.D. (1988) *J. Virol.*, **62**, 2790–2798.

Howe,J.G. and Shu,M.D. (1989) *Cell*, **57**, 825–834.

Howe,J.G. and Steitz,J.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9006–9010.

Jat,P. and Arrand,J.R. (1982) *Nucleic Acids Res.*, **10**, 3407–3425.

Kalderon,D., Roberts,B.L., Richardson,W.D. and Smith,A.E. (1984) *Cell*, **39**, 499–509.

Katze,M.G., DeCorato,D., Safer,B., Galabru,J. and Hovanessian,A.G. (1987) *EMBO J.*, **6**, 689–697.

Kieff,E. and Liebowitz,D. (1990) In Fields,B.N. and Knipe,D.M. (eds), *Virology*. Raven Press, New York, pp. 1889–1920.

Kitajewski,J., Schneider,R.J., Safer,B., Munemitsu,S.M., Samuel,C.E., Thimmappaya,B. and Shenk,T. (1986) *Cell*, **45**, 195–200.

Klein,G. (1989) *Cell*, **58**, 5–8.

Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8132.

Kurilla,M.G. and Keene,J.D. (1983) *Cell*, **34**, 837–845.

Laux,G., Perricaudet,M. and Farrell,P.J. (1988) *EMBO J.*, **7**, 769–774.

Lazinski,D., Grzadzielska,E. and Das,A. (1989) *Cell*, **59**, 207–218.

Lee,S.I., Murthy,S.C.S., Trimble,J.J., Desrosiers,R.C. and Steitz,J.A. (1988) *Cell*, **54**, 599–607.

Lerner,M.R., Andrews,N.C., Miller,G. and Steitz,J.A. (1981a) *Proc. Natl. Acad. Sci. USA*, **78**, 805–809.

Lerner,M.R., Boyle,J.A., Hardin,J.A. and Steitz,J.A. (1981b) *Science*, **211**, 400–402.

Longnecker,R. and Kieff,E. (1990) *J. Virol.*, **64**, 2319–2326.

Marschalek,R., Amon-Böhm,E., Stoerker,J., Klages,S., Fleckenstein,B. and Dingermann,T. (1989) *Nucleic Acids Res.*, **17**, 631–643.

Mathews,M.B. and Francoeur,A.M. (1984) *Mol. Cell. Biol.*, **4**, 1134–1140.

Matsudaira,P. (1987) *J. Biol. Chem.*, **262**, 10035–10038.

Mattaj,I.W. (1988) In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer-Verlag, Berlin, pp. 100–114.

Mattaj,I.W. and De Robertis,E.M. (1985) *Cell*, **40**, 111–118.

Mattioli,M. and Reichlin,M. (1974) *Arthritis and Rheumatism*, **17**, 421–429.

Mellits,K.H., Kostura,M. and Mathews,M.B. (1990) *Cell*, **61**, 843–852.

Merril,C.R., Goldman,D., Selman,S. and Ebert,M.H. (1981) *Science*, **211**, 1437–1438.

Milburn,S.C., Hershey,J.W.B., Davies,M.V., Kelleher,K. and Kaufman, R.J. (1990), EMBO J., **9**, 2783–2790.

Miller,G. (1990) In Fields,B.N. and Knipe,D.M. (eds), *Virology*. Raven Press, New York, pp. 1921–1958.

Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609–1614.

Moreland,R.B., Langevin,G.L., Singer,R.H., Garcea,R.L. and Hereford,L.M. (1987) *Mol. Cell. Biol.*, **7**, 4048–4057.

O'Malley,R.P., Mariano,T.M., Siekierka,J. and Mathews,M.B. (1986) *Cell*, **44**, 391–400.

Pizer,L.I., Deng,J.S., Stenberg,R.M. and Tan,E.M. (1983) *Mol. Cell. Biol.*, **3**, 1235–1245.

Provost,T.T. (1979) *J. Invest. Dermatol.*, **72**, 110–113.

Reich,P.R., Rose,J., Forget,B. and Weissman,S.M. (1966) *J. Mol. Biol.*, **17**, 428–439.

Reichel,P.A., Merrick,W.C., Siekierka,J. and Mathews,M.B. (1985) *Nature*, **313**, 196–200.

Rinke,J. and Steitz,J.A. (1982) *Cell*, **29**, 149–159.

Rosa,M.D., Gottlieb,E., Lerner,M.R. and Steitz,J.A. (1981) *Mol. Cell. Biol.*, **1**, 785–796.

Schneider,R.J., Weinberger,C. and Shenk,T. (1984) *Cell*, **37**, 291–298.
Siekierka,J., Mariano,T.M., Reichel,P.A. and Mathews,M.B. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1959–1963.
Smith,P.R., Williams,D.G., Venables,P.J.W. and Maini,R.N. (1985) *J. Immunol. Methods*, **77**, 63–76.
Söderlund,H., Pettersson,U., Vennström,B., Philipson,L. and Mathews,M.B. (1976) *Cell*, **7**, 585–593.
Stefano,J.E. (1984) *Cell*, **36**, 145–154.
Swaminathan,S., Tomkinson,B. and Kieff,E. (1991) *Proc. Natl. Acad. Sci. USA*, in press.
Thimmappaya,B., Weinberger,C., Schneider,R.J. and Shenk,T. (1982) *Cell*, **31**, 543–551.
van Eekelen,C., Buijtels,H., Linné,T., Ohlssons,R., Philipson,L. and van Venrooij,W. (1982) *Nucleic Acids Res.*, **10**, 3039–3052.
van Santen,V., Cheung,A. and Kieff,E. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 1930–1934.
Walker,J.M., Hastings,J.R.B. and Johns,E.W. (1978) *Nature*, **271**, 281–282.
Wassarman,D.A., Lee,S.I. and Steitz,J.A. (1989) *Nucleic Acids Res.*, **17**, 1258.
Watson,D.C. and Dixon,G.H. (1981) *Bioscience Reports*, **1**, 167–175.

Received on October 31, 1990