

The Multiple RNA-Binding Domains of the mRNA Poly(A)-Binding Protein Have Different RNA-Binding Activities

CHRISTOPHER G. BURD, ERIKA L. MATUNIS, AND GIDEON DREYFUSS*

Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6148

Received 28 November 1990/Accepted 5 April 1991

The poly(A)-binding protein (PABP) is the major mRNA-binding protein in eukaryotes, and it is essential for viability of the yeast *Saccharomyces cerevisiae*. The amino acid sequence of the protein indicates that it consists of four ribonucleoprotein consensus sequence-containing RNA-binding domains (RBDs I, II, III, and IV) and a proline-rich auxiliary domain at the carboxyl terminus. We produced different parts of the *S. cerevisiae* PABP and studied their binding to poly(A) and other ribohomopolymers in vitro. We found that none of the individual RBDs of the protein bind poly(A) specifically or efficiently. Contiguous two-domain combinations were required for efficient RNA binding, and each pairwise combination (I/II, II/III, and III/IV) had a distinct RNA-binding activity. Specific poly(A)-binding activity was found only in the two amino-terminal RBDs (I/II) which, interestingly, are dispensable for viability of yeast cells, whereas the activity that is sufficient to rescue lethality of a PABP-deleted strain is in the carboxyl-terminal RBDs (III/IV). We conclude that the PABP is a multifunctional RNA-binding protein that has at least two distinct and separable activities: RBDs I/II, which most likely function in binding the PABP to mRNA through the poly(A) tail, and RBDs III/IV, which may function through binding either to a different part of the same mRNA molecule or to other RNA(s).

The functional form of eukaryotic mRNA is a ribonucleoprotein (RNP) complex (7). The predominant protein of this complex is the poly(A)-binding protein (PABP), and it is associated preferentially with the 3' poly(A) tail of most mRNAs (4). Though the function of this protein has remained elusive, it is essential for viability of the yeast *Saccharomyces cerevisiae* and is thought to influence a wide variety of cellular processes involving mRNA, including ribosome biogenesis, translation, and mRNA stability (3, 12, 17-19).

The gene for the PABP has been cloned from several organisms, and sequence analysis revealed that it contains four consensus sequence-type RNA-binding domains (RBDs I, II, III, and IV) and a proline-rich carboxyl-terminal auxiliary domain (1, 6, 9, 15, 16, 21, 23, 24). A comparison of PABP sequences from divergent organisms indicated that the individual RBDs have been conserved individually; each is more similar to the corresponding domain in other species than to any of the other RBDs in the same protein, suggesting that the individual RBDs are not functionally equivalent (2, 8). To test this hypothesis and to characterize and localize the poly(A)-binding activity of this protein, we carried out in vitro RNA-binding experiments with the *S. cerevisiae* PABP. We assayed for the ability of the in vitro-translated protein, or parts of it, to bind to immobilized ribohomopolymers (22). In this assay, the *S. cerevisiae* PABP binds poly(A) specifically in 1 M NaCl (22). Though binding to poly(A) is the hallmark of the PABP, at salt concentrations which more closely approximate physiological conditions (0.1 M KCl) it also binds to poly(U) and poly(G) but not to poly(C) (22; Fig. 1 and 2). In fact, more PABP binds to poly(U) and to poly(G) than to poly(A) under these conditions (Fig. 1 and 2). In contrast, parallel binding experiments with heterogeneous nuclear RNP C1 and C2 proteins, avid poly(U)-binding proteins that contain a single

RBD, showed a much more restricted RNA-binding activity; only limited binding was observed to poly(G) (in 0.5 M NaCl), and no binding to poly(A) or to poly(C) was observed under any conditions (22).

MATERIALS AND METHODS

PCR. The template for the polymerase chain reactions (PCR) was a plasmid containing a *Kpn*I fragment that encodes the *S. cerevisiae* PABP gene (pYEA3) (1). Top-strand PCR primers encoded six amino acids amino terminal to each RBD and, for RBDs II, III, and IV, included an AUG initiator codon. Bottom-strand primers were complementary to the region encoding six amino acids carboxyl terminal to each RBD. (See the legend to Fig. 2 for the specific positions of the primers.) Oligonucleotides were synthesized by the Northwestern University Biotechnology Center (Northwestern University, Evanston, Ill.). PCR were done using a GeneAmp (Cetus/Perkin Elmer) kit according to the manufacturer's instructions and were done in a Perkin Elmer thermal cycler. PCR products were cloned into pGEM7Z, and their identities were confirmed by DNA sequencing.

RNA-binding assays. Each pairwise combination of RBDs was produced by transcription of the cloned PCR fragments with T7 RNA polymerase (Promega), followed by translation of the resultant RNA in rabbit reticulocyte lysate (Promega) in the presence of [³⁵S]methionine (Amersham). All constructs yielded polypeptides of the expected sizes. The relative efficiency of each translation was monitored by trichloroacetic acid precipitation and liquid scintillation counting of duplicate samples. The amount of labeled polypeptide used in each binding reaction was varied such that equal trichloroacetic acid-precipitable counts were used for each binding reaction. RNA-binding assays were performed in 0.1 M salt as previously described (22) except that KCl was substituted for NaCl. Briefly, in vitro-translated polypeptides were added to agarose-bound RNA homopolymers (Pharmacia) in a total volume of 300 μl of binding buffer (10

* Corresponding author.

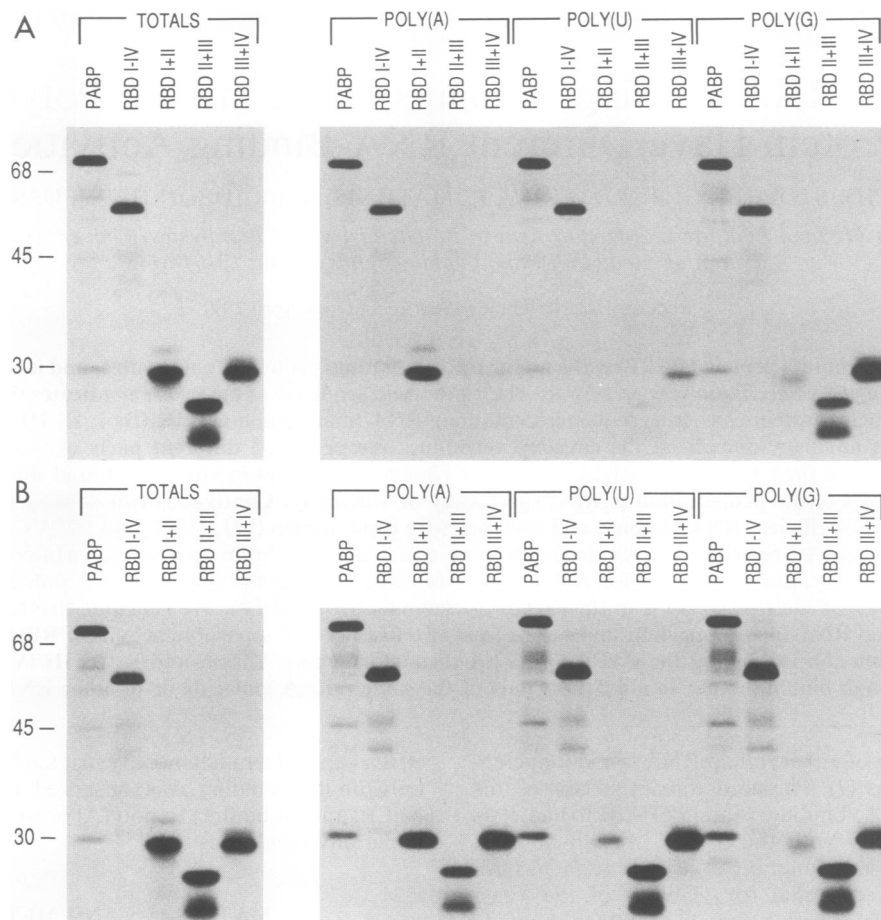


FIG. 1. Binding of PABP and polypeptide fragments of PABP to ribohomopolymers. The PABP and a fragment containing just the RBDs (I to IV) was produced by *in vitro* transcription-translation of pYEA3 truncated either outside the coding region (*Pst*I; lane PABP) or at an internal *Bst*NI site that truncates the cDNA at nucleotide 1388 (relative to the 5' *Kpn*I site) (lane RBD I-IV). Pairwise combinations of RBDs were produced by PCR of pYEA3 (see Fig. 2) and *in vitro* transcription-translation. An amount equivalent to 40% of the material used for each binding reaction is shown in the lanes marked TOTALS. RNA-binding assays were performed as described in Materials and Methods. In panel A, heparin (1 mg/ml) was added to the binding reaction; in panel B, yeast tRNA (0.5 mg/ml) was included in the binding reaction. The minor higher-molecular-weight polypeptides are artifacts of the *in vitro* transcription-translation. No binding to poly(C)-agarose was ever observed (data not shown).

mM Tris [pH 7.4], 2.5 mM MgCl₂, 100 mM KCl, 0.5% Triton X-100). Either heparin (1 mg/ml; Sigma) or yeast tRNA (0.5 mg/ml; Boehringer Mannheim) was included in the binding reaction. After a 10-min incubation, the beads were washed four times with 1 ml of binding buffer and the bound material was eluted by boiling in sodium dodecyl sulfate (SDS) sample buffer. The products were electrophoresed in 12.5% SDS-polyacrylamide gels and visualized by fluorography (22). To assess the relative binding activity of each polypeptide, aliquots of the material eluted from the RNA-agarose beads were quantitated by liquid scintillation counting.

Cloning of *Schizosaccharomyces pombe* PABP. A partial-length cDNA clone for the *S. pombe* PABP was isolated by immunological screening of a lambda gt11 cDNA library (a generous gift from JoAnne Wise, University of Illinois) with antibodies to human RNP proteins. A 1.8-kb cDNA was obtained and used to screen a different *S. pombe* lambda gt11 cDNA library (provided by Norbert Käufer, Drexel University) by hybridization. The authenticity of the clones was established by DNA sequence analysis.

DNA sequence analysis. Cloned PCR DNAs and the *S. pombe* PABP cDNA were sequenced by the dideoxy method

using T7 DNA polymerase (Pharmacia) after cloning into mp18/19 or pGEM7Z. *S. pombe* PABP-specific primers were also used in some cases.

Nucleotide sequence accession number. The DNA and amino acid sequences for *S. pombe* PABP have been submitted to the GenBank data base and have been assigned the accession number M64603.

RESULTS AND DISCUSSION

The cDNA sequence of the PABP revealed that the protein contains four RBDs (1). The RNA-binding activity of the PABP is due to the RBDs because a mutant that deletes the carboxyl-terminal auxiliary domain binds RNA with properties identical to those of the full-length protein (Fig. 1, lanes RBD I-IV; Fig. 2). To test the possibility that the binding by the protein to poly(A), poly(G), and poly(U) results from the binding of different RBDs (or combinations of RBDs) with different RNA-binding specificities, we produced each of the RBDs by PCR cloning and *in vitro* transcription-translation. Each construct contained several additional amino acids on both the amino- and carboxyl-

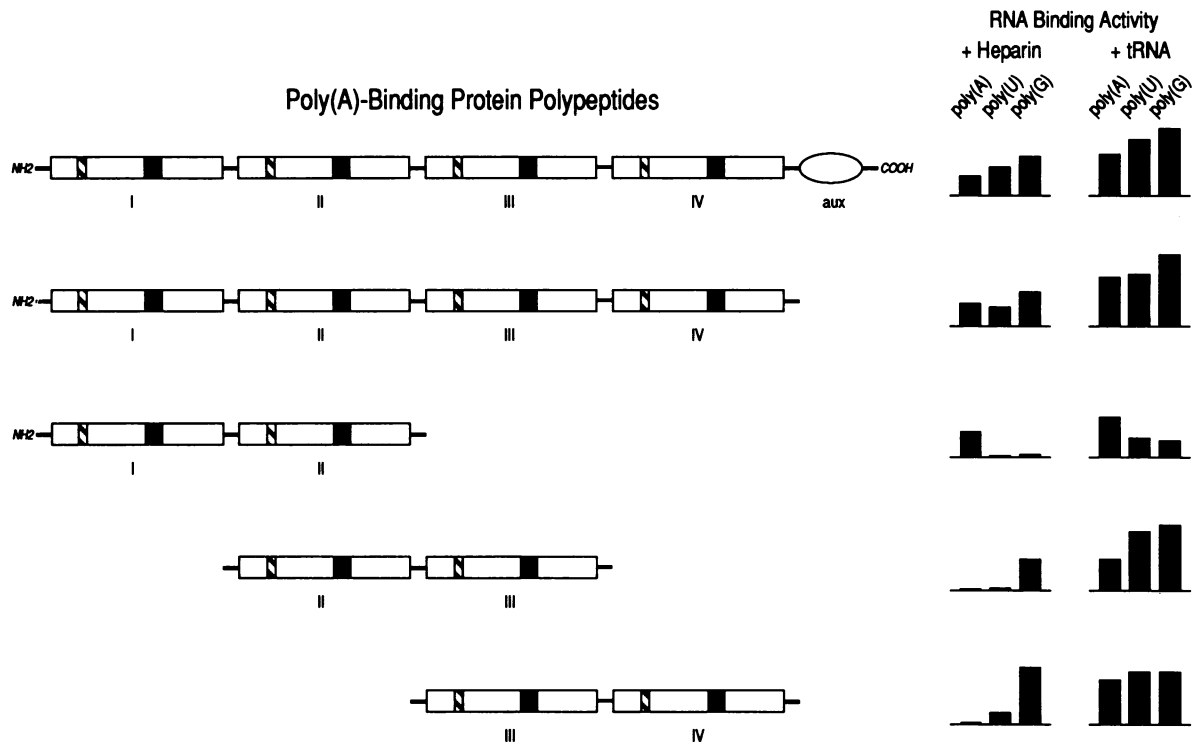


FIG. 2. PABP polypeptides and summary of RNA-binding activity. In the schematic representations of the various *S. cerevisiae* PABP fragments used in the RNA-binding assays, each box represents an RBD that is identified by a Roman numeral below it. The black region in the approximate center of each RBD represents the RNP consensus sequence, and the hatched box represents RNP2. RBDs I to IV encompass amino acids 1 to 433, RBDs I/II encompass amino acids 1 to 210, RBDs II/III span amino acids 116 to 312, and RBDs III/IV span amino acids 204 to 421. The relative RNA-binding activity of each peptide was determined by liquid scintillation counting of an aliquot of the material eluted from the RNA-agarose beads. The height of each bar indicates the percent of input counts per minute that bound to the RNA-agarose beads, and all bars are drawn to scale. The maximal value found was approximately 60% for the PABP to poly(G) with tRNA as the competitor.

terminal sides of the predicted RBD (based on sequence homology) (see legend to Fig. 2). These polypeptides were assayed for RNA-binding activity with either heparin or tRNA as the nonspecific competitor. With single-domain constructs, we did not observe any specific poly(A)-binding activity for any of the RBDs. Significant single-domain binding was detected only by RBD IV for poly(G) (data not shown). However, for contiguous pairwise combinations of RBDs (I/II, II/III, III/IV), we found that these polypeptides differ in their RNA-binding specificities (Fig. 1 and 2). The most striking difference in the RNA-binding activities is observed in the presence of the nonspecific polyanion competitor heparin (Fig. 1A). Under these conditions, the only poly(A)-binding activity resides in the amino-terminal RBDs I/II. When tRNA is used as a nonspecific RNA competitor (Fig. 1B), it is evident that the other combinations of RBDs (II/III and III/IV) have the capacity to bind poly(A), but this interaction is not selective for poly(A) and is not the strongest activity of these RBDs. Results nearly identical to those observed with tRNA as competitor were obtained in the absence of any competitor (data not shown). Most importantly, each pairwise combination, whether in the presence of tRNA competitor or heparin, had a distinct RNA-binding profile. These results support an earlier prediction (2) that proteins containing multiple RBDs have the capacity to interact with more than one RNA sequence [e.g., poly(A) and non-poly(A)] simultaneously.

The differences in the binding profiles observed with or

without heparin indicate that there are two distinct modes of RNA binding for the PABP. RBDs I/II bind poly(A) in a heparin-resistant manner but bind poly(G) in a heparin-sensitive manner. The opposite is true for RBDs III/IV: binding to poly(G) is heparin resistant but binding to poly(A) is heparin sensitive. No heparin-resistant binding to poly(U) was observed for any RBD combination. While we do not know the basis for the effect of heparin, it is consistent with competition for nonspecific electrostatic interactions with the highly charged phosphate backbone by virtue of its polyanionic character (10, 14). The PABP may, however, also be capable of making specific electrostatic contacts with the RNA because it binds poly(U) in a heparin-sensitive manner but not poly(C). By the criterion of heparin sensitivity, specific RNA binding by the PABP with poly(A) is made only by RBDs I/II. This is also apparent in the experiment using tRNA and is in agreement with previous studies that determined that the PABP contains only one high-affinity poly(A)-binding site and at least one low-affinity binding site (18). The binding to poly(U) and poly(G) is significant in that it reflects distinct types of RNA binding; it does not necessarily mean that these homopolymeric stretches are the specific and physiologically relevant target binding sites. In support of this view, several of the heterogeneous nuclear RNA-binding proteins do not bind single-stranded DNA but bind tenaciously to poly(G) (14, 22). With respect to poly(U) binding, however, we note that A+U-rich sequence elements located just upstream of the poly(A) tail

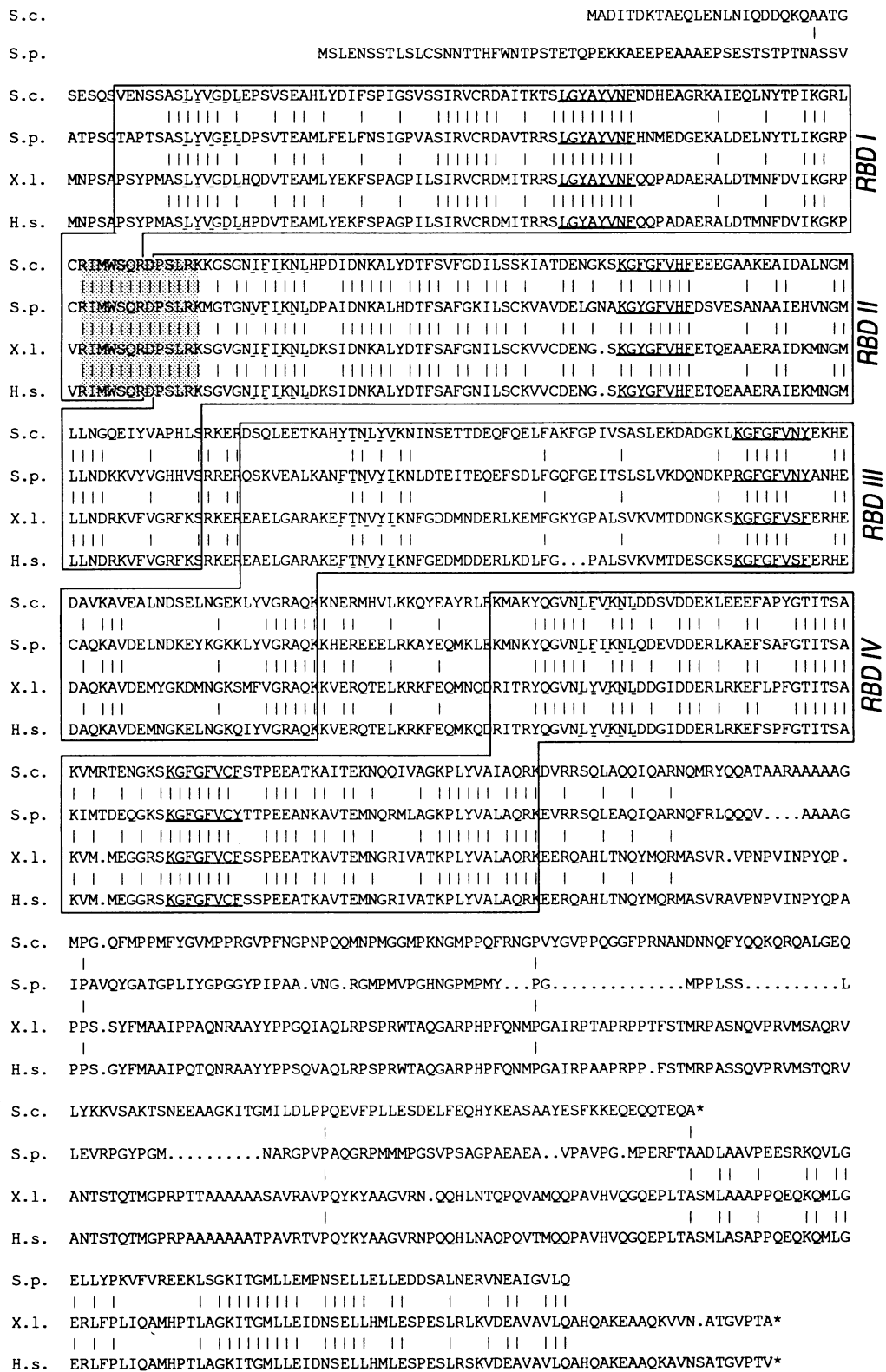


FIG. 3. Sequence comparison of *S. cerevisiae* (S.c.), *S. pombe* (S.p.), *X. laevis* (X.l.), and *Homo sapiens* (H.s.) PABPs. Amino acid sequences are shown in one-letter code. Identities between all four proteins are indicated by a vertical line. Within each RBD (boxed), the RNP consensus sequence is underlined with a solid line and RNP2 is underlined with a broken line. The shaded region between RBD I and RBD II is the most highly conserved element between all four proteins. The *S. cerevisiae* sequence is from reference 1, the *X. laevis* sequence is from reference 22, and the *H. sapiens* sequence is from reference 9.

in certain unstable mRNAs (e.g., *c-myc*) may destabilize the mRNA by decreasing the residency time of the PABP on the poly(A) tail, thus allowing nucleolytic attack of the RNA (3, 5).

The binding experiments described above indicated that RBDs I/II are required for specific poly(A) binding, since each of these two RBDs alone does not have such specificity. A polypeptide spanning amino acids 1 to 173 (encompassing RBD I and the amino-terminal part of RBD II ending immediately after the RNP consensus sequence), produced by transcription-translation of a *Bst*B1-truncated PABP cDNA, did not have specific poly(A)-binding activity (data not shown). Thus, RBD II amino acids 174 to 210 are essential for this specificity. However, RBD I, or at least part of it, is also essential for specific poly(A)-binding activity. Why are both amino-terminal RBDs (I/II) necessary for specific poly(A) binding? To address this question, we examined the previously determined amino acid sequences of PABP as well as that of the PABP of *S. pombe* presented here. The amino acid sequence of the *S. pombe* PABP aligned with all the other available PABP sequences is shown in Fig. 3. We note that the region of greatest identity includes the junction between RBDs I and II (shaded region of Fig. 3). There is a contiguous stretch of 13 identical amino acids between these two RBDs; in contrast, the junctions between the other RBDs have at most two contiguous identical amino acids. It appears that the binding activity of multiple RBD-containing RNP proteins is derived and specified not only by the particular determinants of the individual RBDs but also by their precise interactions and spatial positioning. Furthermore, the spatial positioning of RBDs may be influenced by the region that connects the RBDs and/or by specific RBD-RBD interactions. We cannot, however, rule out the possibility that this region itself also makes specific contacts with the RNA. As pointed out previously for the human and *S. cerevisiae* PABPs, the multiple RBDs appear to have been conserved individually, since each is more similar to the corresponding domain in each of the three other species than to any of the other domains, even within the same protein (8). This observation is strengthened when one compares the sequences of four divergent organisms (Fig. 3). The fourth domain is the most highly conserved (~60%), while RBDs I/II are also highly conserved (~55%). RBD III does not appear to be well conserved (~25%). In the less conserved auxiliary domain, the *S. pombe* sequence is more similar to the human and *Xenopus laevis* sequences than to the *S. cerevisiae* sequence.

After this report was submitted for publication, Nietfeld et al. (13) published experiments using the *X. laevis* PABP which were similar to those reported here. Our data agree with their conclusion that the minimal specific poly(A)-binding part of the PABP is RBDs I/II. There are, however, some differences between their findings and ours. By fusing individual RBDs to β -galactosidase and assaying for the ability of the fusion protein to bind soluble radiolabeled poly(A) or poly(U), Nietfeld et al. found that individual RBDs have the ability to bind RNA, with the strongest activity exhibited by RBD IV (13). This binding was neither efficient nor specific. In our assays, we detected a low level of RNA-binding activity by RBD IV but no RNA-binding activity for the other RBDs (data not shown). The different assays may have different shortcomings. For example, it is possible that different RBDs produced alone (i.e., not as a fusion protein) behave differently in solution, and this could affect their RNA-binding properties. On the other hand, in the experiments of Nietfeld et al. (13), the fusion of the

different RBDs to β -galactosidase could affect the RNA-binding properties of the RBDs. Compared with the assay system of Nietfeld et al. (13), our assay system apparently detects only relatively strong RNA-binding activity. Sachs et al. found that a peptide consisting of RBD IV and the auxiliary domain of the *S. cerevisiae* PABP was able to bind oligo(A) with a binding constant similar to that of the full-length PABP (18). These authors concluded that RBD IV and the auxiliary domain contain the (single) high-affinity binding site of the protein for poly(A). However, these measurements were made in the absence of competitor, and the binding constants of other parts of the PABP were not determined. We have found that in the absence of competitor, several different fragments of the PABP are able to bind to RNA with an efficiency that is very close to that of the full-length PABP (Fig. 1B).

The PABP is essential for viability of *S. cerevisiae* (18). It is quite surprising that the strongest poly(A)-binding activity is found in the two amino-terminal RBDs (I/II) which are dispensable for viability, whereas the activity that is sufficient to rescue lethality of a PABP-deleted strain is in the carboxyl-terminal RBDs (III/IV). Furthermore, RBD IV, the most highly conserved RBD of the protein, binds better to non-poly(A) ribohomopolymers than to poly(A) (Fig. 1 and unpublished data). Thus, it is possible that the poly(A)-binding activity of the PABP is not the essential function of this protein in *S. cerevisiae*. Poly(A) binding may simply ensure that the PABP is bound to every polyadenylated mRNA, allowing the essential function of the PABP, mediated by RBDs III/IV, to be efficiently executed. The poly(G)-binding activity of RBDs III/IV probably represents a binding specificity for an RNA sequence other than poly(A). We conclude that the PABP is a multifunctional RNA-binding protein that has at least two separable activities: RBDs I/II, which most likely function in binding the PABP to mRNA through the poly(A) tail, and RBDs III/IV, which may function through binding either to a different part of the same mRNA or to other RNAs. The presence of multiple non-equivalent RBDs in the PABP suggests that it may be able to bind to different segments of the RNA or to multiple RNA molecules (2). Nonequivalence of RBDs within the same protein has been recently shown for the small nuclear RNP proteins U1A and U2B' (11, 20, 21). It will be of great interest to find the specific target RNA(s) other than the poly(A) tail of the mRNA to which the PABP may bind, since these interactions are likely to be important for the essential function of this protein. Finally, our experiments demonstrate that the overall RNA-binding activity of proteins containing multiple RBDs may not simply be the additive composite of the individual RNA-binding activities; rather, it may be influenced by the interplay between the RBDs to produce an RNA-binding activity not present within the individual RBDs.

ACKNOWLEDGMENTS

We thank Matthias Görlach for helpful discussions and Maurice Swanson for advice with the RNA-binding assays. We also thank members of our laboratory for critical reading of the manuscript.

C.G.B. is a National Service Award trainee. This research was supported by grants from the National Institutes of Health and the Howard Hughes Medical Institute.

REFERENCES

1. Adam, S. A., T. Nakagawa, M. S. Swanson, T. K. Woodruff, and G. Dreyfuss. 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol. Cell. Biol.* 6:2932-2943.

2. Bandziulis, R. J., M. S. Swanson, and G. Dreyfuss. 1989. RNA-binding proteins as developmental regulators. *Genes Dev.* 3:431-437.
3. Bernstein, P., S. W. Peltz, and J. Ross. 1989. The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. *Mol. Cell. Biol.* 9:659-670.
4. Blobel, G. 1973. A protein of molecular weight 78,000 bound to the polyadenylate region of eukaryotic messenger RNAs. *Proc. Natl. Acad. Sci. USA* 70:924-928.
5. Brewer, G., and J. Ross. 1988. Poly(A) shortening and degradation of the 3' A+U-rich sequences of the human *c-myc* mRNA in a cell-free system. *Mol. Cell. Biol.* 8:1697-1708.
6. Burd, C. G., M. S. Swanson, M. Görlich, and G. Dreyfuss. 1989. Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. *Proc. Natl. Acad. Sci. USA* 86:9788-9792.
7. Dreyfuss, G. 1986. Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. *Annu. Rev. Cell Biol.* 2:459-498.
8. Dreyfuss, G., M. S. Swanson, and S. Piñol-Roma. 1988. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem. Sci.* 13:86-91.
9. Grange, T., C. Martins de Sa, J. Oddos, and R. Pictet. 1987. Human mRNA polyadenylate binding protein: evolutionary conservation of a nucleic acid binding protein. *Nucleic Acids Res.* 15:4771-4787.
10. Konarska, M. M., and P. A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* 46:845-855.
11. Lutz-Freyermuth, C., C. C. Query, and J. D. Keene. 1990. Quantitative determination that one of two potential RNA-binding domains of the A protein component of the U1 small nuclear ribonucleoprotein complex binds with high affinity to stem-loop II of U1 RNA. *Proc. Natl. Acad. Sci. USA* 87:6393-6397.
12. Munroe, D., and A. Jacobson. 1990. mRNA poly(A) tail, a 3' enhancer of translational initiation. *Mol. Cell. Biol.* 10:3441-3455.
13. Nietfeld, W., H. Mentzel, and T. Pieler. 1990. The *Xenopus laevis* poly(A) binding protein is composed of multiple functionally independent RNA binding domains. *EMBO J.* 9:3699-3705.
14. Piñol-Roma, S., Y. D. Choi, M. J. Matunis, and G. Dreyfuss. 1989. Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes Dev.* 2:215-227.
15. Query, C. C., R. C. Bently, and J. D. Keene. 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* 57:89-101.
16. Sachs, A. B., M. W. Bond, and R. D. Kornberg. 1986. A single gene from yeast for both nuclear and cytoplasmic polyadenylate-binding proteins: domain structure and expression. *Cell* 45:827-835.
17. Sachs, A. B., and R. W. Davis. 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell* 58:857-867.
18. Sachs, A. B., R. W. Davis, and R. D. Kornberg. 1987. A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. *Mol. Cell. Biol.* 7:3268-3276.
19. Sachs, A. B., and R. W. Davis. 1990. Translational initiation and ribosomal biogenesis: involvement of a putative rRNA helicase and RPL46. *Science* 247:1077-1079.
20. Scherly, D., W. Boelens, N. A. Dathan, W. J. van Venrooij, and I. W. Mattaj. 1990. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B' and their cognate RNAs. *Nature (London)* 345:502-506.
21. Scherly, D., W. Boelens, W. J. van Venrooij, N. A. Dathan, J. Hamm, and I. W. Mattaj. 1989. Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. *EMBO J.* 8:4163-4170.
22. Swanson, M. S., and G. Dreyfuss. 1988. Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. *Mol. Cell. Biol.* 8:2237-2241.
23. Swanson, M. S., T. Y. Nakagawa, K. LeVan, and G. Dreyfuss. 1987. Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA binding proteins. *Mol. Cell. Biol.* 7:1731-1739.
24. Zelus, B. D., D. H. Giebelhaus, D. W. Eib, K. Kenner, and R. T. Moon. 1989. Expression of the poly(A)-binding protein during development of *Xenopus laevis*. *Mol. Cell. Biol.* 9:2756-2760.