32–40 *Nucleic Acids Research, 1994, Vol. 22, No. 1* © *1994 Oxford University Press*

The yeast protein encoded by *PUB1* binds T-rich single stranded DNA

Moira Cockell, Severine Frutiger[1], Graham J.Hughes[1] and Susan M.Gasser*
Swiss Institute for Experimental Cancer Research (ISREC), Chemin des Boveresses 155, CH-1066 Epalinges/Lausanne and [1]Centre Médical Universitaire, Université de Genève, Département de Biochimie Médicale, 1 rue Servet, CH-1211 Geneva 4, Switzerland

Received October 18, 1993; Revised and Accepted December 6, 1993

ABSTRACT

We have characterized binding activities in yeast which recognise the T-rich strand of the yeast ARS consensus element and have purified two of these to homogeneity. One (ACBP-60) is detectable in both nuclear and whole cell extracts, while the other (ACBP-67) is apparent only after fractionation of extracts by heparin-sepharose chromatography. The major binding activity detected in nuclear extracts was purified on a sequence-specific DNA affinity column as a single polypeptide with apparent mobility of 60kDa (ACBP-60). This protein co-fractionates with nuclei, is present at several thousand copies per cell and has a K_d for the T-rich single strand of the ARS consensus between 10^{-9} and 10^{-10} M. Competition studies with simple nucleic acid polymers show that ACBP-60 has marginally higher affinity for poly dT_{30} than for a 30 nt oligomer containing the T-rich strand of ARS 307, and approximately 10 fold higher affinity for poly rU. Internal sequence information of purified p60 reveals identity with the open reading frames of genes *PUB1* and *RNP1* which encode polyuridylate binding protein(s). The second binding activity, ACBP-67, also binds specifically to the T-rich single strand of the ARS consensus, but with considerably lower affinity than ACBP-60. Peptide sequence reveals that the 67kDa protein is identical to the major polyA binding protein in yeast, *PAB1*.

INTRODUCTION

In the yeast *Saccharomyces cerevisiae*, replication origins on plasmids and chromosomes are composed of at least two functionally distinct domains (reviewed in 1). One of these, the 11 bp ARS core consensus, is an essential element, although it is not sufficient for the initiation of DNA synthesis on autonomously replicating plasmids (2–5). The importance of the conserved ARS consensus was confirmed by mutagenesis studies which show that single point mutations reduce the replicative is likely to be a binding site for a sequence specific DNA binding protein involved in the formation of replication initiation complexes.

There is now convincing evidence that the integrity of the cis-acting ARS consensus is important for the binding both *in vitro* and *in vivo* of a multiprotein complex called ORC (origin recognition complex), which recognizes double stranded ARS elements, contacting DNA at the conserved ARS consensus (A domain) and in the flanking B domain (8, 9). By analogy with phage, bacterial and viral replication systems that have been reconstituted *in vitro*, initiation of replication in eukaryotic genomes is likely to be a multistep process in which origin recognition, helix opening, assembly of the replication complex and initiation of elongation are sequential events involving different components. The identification of one protein–DNA interaction involving the ARS consensus does not preclude the existence of others. Moreover, *in vivo* footprinting experiments suggest that the ORC complex is bound to potential origins throughout the cell cycle and may be present on both active and inactive origins, implying that mechanisms other than ORC-origin binding may regulate the timing and choice of sequences destined to be used in a given cell cycle (9).

In the last few years, there have been a number of reports identifying single stranded binding activities in yeast, with apparent sequence specificity for the T-rich strand of the ARS core consensus (10–13). To date, the biological role of such protein(s) remains to be demonstrated, and it remains unclear whether the *in vitro* single-stranded binding activities cited in these independent reports are due to the same or to different proteins. In an attempt to resolve this issue, we have carried out an exhaustive search for single-stranded ARS binding activities. We detected several different activities in yeast whole cell and nuclear extracts that recognise the T-rich strand of the ARS consensus element, and we have purified to homogeneity the two that bind the single strand consensus with highest affinity. Sequence analysis identifies the two proteins as products of the yeast genes *PAB1* and *PUB1*.

* To whom correspondence should be addressed

MATERIALS AND METHODS

Preparation of extracts

Protein extracts were prepared from the protease deficient *Saccharomyces cerevisiae* strain 62-5c (14). For whole cell extracts, cells were harvested and washed 2× in cold distilled water. Cells were resuspended in breakage buffer (200 mM Tris−Cl pH 8, 5 mM EDTA, 1 mM DTT, 10% glycerol, 0.5 mM PMSF, 20 μg/ml pepstatin A and 20 μg/ml leupeptin) and broken with glass beads. The broken cells were filtered through nylon gause and the extract spun at 10,000×g for 45 min. Proteins were precipitated from the supernatant by addition of 0.35 g of ammonium sulphate per ml of extract and incubation for 1 hr on ice. The precipitate was recovered by centrifugation at 12,000×g for 1 hr and resuspended in A50 buffer (20 mM Tris−Cl pH 8, 1 mM EDTA, 0.5 mM DTT, 10% glycerol, 50 mM ammonium sulphate, 0.5 mM PMSF, 20 μg/ml pepstatin A, 20 μg/ml leupeptin) then dialysed overnight against the same buffer. All procedures were carried out at 4°C. Nuclear extracts and scaffold extracts were prepared as described by Hofmann *et al.* (15).

Purification of ACBP-60 and ACBP-67

Purified ACBP-60 and ACBP-67 for peptide sequencing were obtained from whole cell extracts by the following procedure; dialysed whole cell extract (prepared as described above) from 10 L of yeast grown to an OD_{600} of 1.5 was passed over a 40 ml heparin−sepharose column. The column was washed with 5 column volumes of A 50 buffer and then bound proteins were eluted with a linear gradient of ammonium sulphate (A50−A1000). Each of the different binding activities was monitored by UV crosslinking to single stranded end-labelled T-rich ARS box oligonucleotides, followed by analysis on SDS-PAGE.The flowthrough from the heparin−sepharose column (containing the bulk of the ACBP-60 binding activity) and elution fractions A300−A400 (containing the bulk of the ACBP-67 binding activity were then each subjected to DNA affinity chromatography. DNA affinity columns were prepared by coupling the biotinylated T-rich strand of the H4 ARS box oligonucleotide to a streptavin−agarose matrix. Fractions containing ACBP-60 or ACBP-67 binding activities were dialysed in K50 buffer (20 mM Tris−Cl pH 8, 1 mM EDTA, 0.5 mM DTT, 10% glycerol, 50 mM KCl, 0.5 mM PMSF, 20 μg/ml pepstatin A, 20 μg/ ml leupeptin) and then incubated for 2hr at 4°C with 1ml of streptavidin−agarose matrix containing approximately 50 μg of the H4 ARS box oligo. The slurry was collected in a Biorad econocolumn, washed extensively with K50 buffer and then eluted with a stepwise gradient of KCl (K100−K1M in 100 mM increments). Fractions containing ACBP-60 binding activity were incubated with the streptavidin−agarose−DNA in the presence of 5 mg of sonicated single stranded *E.coli* DNA and 5 mg poly A RNA. Purified ACBP-60 and ACBP-67 were reduced and carboxymethylated before CnBr cleavage or trypsin digestion. Resulting peptides were separated by two dimensional reverse phase HPLC as described by Hughes *et al.* (16). Sequence determination was carried out using Applied Biosystems models 473a and 477a, pulsed liquid phase microsequencers.

DNA probes, competitor deoyribonucleotides and ribonucleotides

Oligonucleotides used in this study were generated by chemical synthesis on an Applied Biosystems DNA synthesiser. For DNA affinity chromatography, the 31 nt oligonucleotide, GTGTTTATTTATTTTTATGTTTTGTATTTGG containing the T-rich strand of the H4 ARS box was biotinylated at the 3′ end using terminal transferase and biotin-16 dUTP. The T-rich H4 ARS oligonucleotide and the 22 nt, T-rich, ARS 307 oligonucleotide TTTATTTATGTTTTCTTCTGCA were 5′ end labelled with g^{32}P-ATP using polynucleotide kinase, for use as probes in gel retardation and UV crosslinking assays. The oligonucleotides containing 9/11 and 10/11 ARS consenses which were used as competitors in figure 6 had the sequence ACC-<u>TTAAATATTTA</u>CTCA and AACAT<u>CTTTATATTTA</u>TCCG, respectively. Matches to the ARS consensus are underlined. The 200 bp DNA fragment containing the H4 ARS box was generated from pAB9 (4) by digestion with HindIII followed by 5′ end-labelling and redigestion with RSAI. This fragment and a 230 bp 5′ end-labelled fragment from the pUC8 vector were gel purified. The double stranded DNA fragments were denatured by heating to 95°C followed by rapid cooling on ice immediately prior to incubation with protein extracts. Oligo dT_{14}, dA_{14} and dT_{30} were generated by chemical synthesis. Poly rA, poly rU, poly dC and poly d(IC) were purchased from Sigma Chemicals.

Gel retardation assays and UV crosslinking

Binding reactions were done in bandshift buffer (20 mM Tris−Cl pH 8, 50 mM KCl, 0.5 mM EDTA, 0.5 mM DTT, 10% glycerol). Routinely, 1ng of 5′ end-labelled DNA was incubated with protein extracts in a final volume of 20 μl. Binding reactions were carried out for 10 min at 4°C followed by incubation for 5 min at room temperature. Where indicated, proteins were crosslinked to DNA by UV irradiation (254 nm wavelength, at a distance of 5 cm for 5 min using a hand-held Camag UV lamp). Samples were electrophoresed at 200v for 1 hr at room temperature through 4% polyacrylamide gels in 0.5× TBE for assay of native protein−DNA complexes by EMSA (electrophoretic migration shift assay). For analysis of denatured complexes UV irradiated samples were dissolved in 2×FSB (125 mM Tris−Cl pH 6.8, 4% SDS, 10% glycerol, 0.02% bromophenol blue, 4% b mercaptoethanol), incubated at 96°C for 5 min and loaded on 10% SDS polyacrylamide gels (17).

Determination of K_d for ACBP-60

Gel retardation assays were performed with varying amounts of purified ACBP-60 and 1 ng of the T-rich strand of the H4 ARS consensus. Protein concentration was estimated from Coomassie staining of purified ACBP-60 and molecular weight marker proteins of known concentration. Bands corresponding to bound and free DNA were excised and the amount of radioactivity present in each was determined. The data were used to calculate the K_d of binding from the equation $1/r = K_d/(\text{free protein})$ where r = (protein−DNA complex)/ (protein) + (protein−DNA complex).

RESULTS

Budding yeast contains multiple binding activities with affinity for T-rich single stranded DNA

In initial attempts to identify ARS consensus binding activities, a single protein DNA complex was detected in yeast nuclear scaffold extracts using the labelled T-rich strand of the yeast H4 ARS consensus as probe (Figure 1A, lane 3). No complexes were formed when the complementary A-rich strand of the ARS consensus was used as a probe (Figure 1A, lanes 2 and 4), and

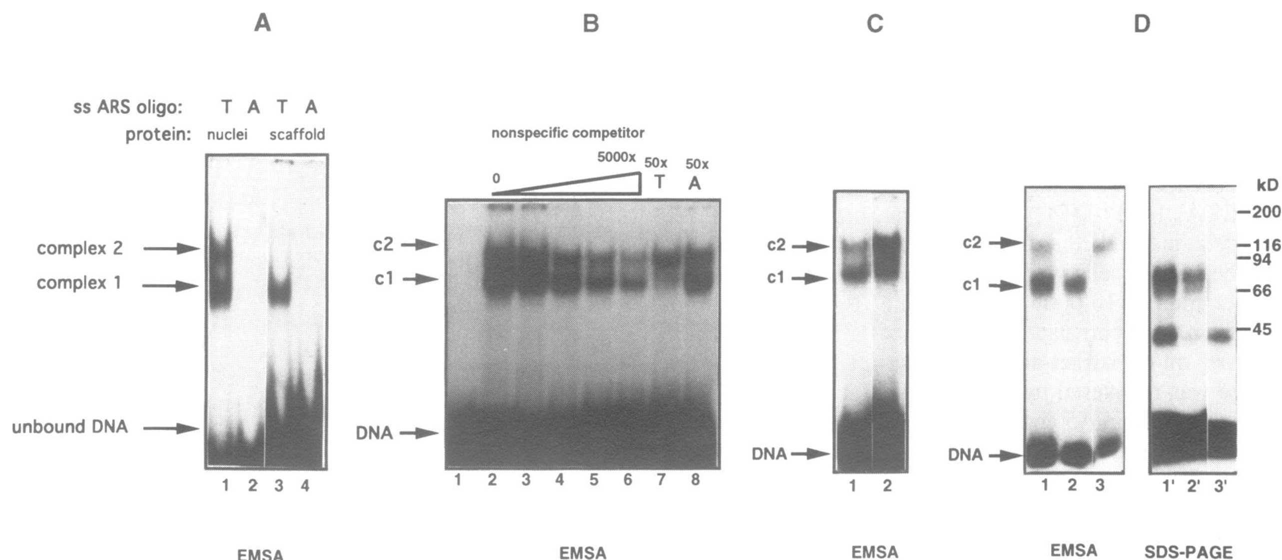


Figure 1. Yeast nuclear extracts give rise to at least two different protein–DNA complexes with single-stranded DNA containing the T-rich strand of the ARS consensus. **a)** Gel retardation assay performed with 1ng of 5′ end-labelled H4 ARS box oligo (T-rich strand, lanes 1 and 3, A-rich strand, lanes 2 and 4), incubated with 3 μg of nuclear (lanes 1 and 2) or scaffold (lanes 3 and 4) extract in the presence of 1μg of single-stranded *E.coli* competitor DNA. **b)** Gel retardation assay performed with 1ng of 5′ end-labelled H4 ARS T-rich oligo incubated with 3 μg of nuclear extract in the presence of increasing amounts of single-stranded *E.coli* competitor DNA (lane 1; no protein extract, lane 2; no competitor DNA , lane 3; 0.5 μg, lane 4; 1 μg , lane 5; 2.5 μg, lane 6; 5 μg) or a 50×molar excess of ARS oligo (lane 7;T-rich strand, lane 8; A-rich strand). **c)** Gel retardation assay performed with 5′ end-labelled H4 ARS T-rich oligo (lane 1; 1 ng, lane 2; 5 ng) incubated with 3 μg of nuclear extract. **d)** Gel retardation assay performed with 1ng of 5′ end-labelled H4 ARS T-rich oligo incubated with 3 μg of nuclear extract and single-stranded competitor DNA (lanes 1and 1′; 1μg *E.coli* DNA, lanes 2 and 2′; 5μg *E.coli* DNA, lanes 3 and 3′; 1μg *E.coli* DNA and 50ng of H4 ARS T-rich oligo. Binding assays were subjected to crosslinking by exposure to UV light. Samples were split in two and migration of protein–DNA complexes was assayed on non-denaturing (lanes 1–3) and denaturing (lanes 1′–3′) gels.

the complex with T-rich DNA resisted nonspecific competitor in 5000-fold weight excess (Figure 1B and ref. 10). In nuclear extracts a second, slower migrating complex also specific for the T-rich strand of the ARS consensus was detected (complex 2, Figure 1A, lane 1). This second complex is not detected in nuclear scaffold extracts, and it is less resistant to competition by nonspecific competitor DNA than complex 1 (Figure 1B). While binding with lower specificity, the protein responsible for the formation of complex 2 appears to be a protein of very high abundance, since complex 2 is not competed by a 50 fold excess of unlabelled T-rich probe, while complex 1 is (Figure 1B, lane 7). Consistent with this, a titration of ^{32}P-labelled probe into the binding assays with a constant amount of nuclear extract, demonstrates that complex 1 binding activity is saturated by probe DNA before complex 2 binding activity (Figure 1C, lane 2).

To identify the size of the proteins implicated in these complexes, we analysed on SDS gels the polypeptides crosslinked by UV to the labelled oligonucleotide, under conditions chosen to favor formation of either complex 1 (Figure 1D, lane 2), or complex 2 (Figure 1D, lane 3). When the UV crosslinked complex 1 was analysed by SDS-PAGE, we detected a denatured protein–DNA complex with an apparent molecular weight of 65–70kDa (Fig. 1D, lane 2'), while complex 2 correlated with detection of a crosslinked protein band migrating around 42–45kDa in the denaturing gel (Fig. 1D, lane 3'). Although complex 2 binding activity shows distinct preference for the T-rich over the A-rich strand of the ARS probe, we surmise from its relatively high abundance and low affinity that its *in vitro* DNA binding activity is unlikely to be of functional significance. We did not pursue the identity of the factor(s) forming complex 2 further (see Discussion).

During purification of the binding activity that produced complex 1 (see Figure 1A), whole cell extracts were fractionated by heparin-sepharose chromatography. Analysis of the column elution profile by detection of denatured, UV crosslinked protein DNA complexes on SDS gels revealed the presence of not one, but two differently migrating activities (Fig. 2A, labelled ACBP-60 and ACBP-67), which under certain conditions comigrate as complex 1 in bandshift assays. A major fraction of the faster-migrating ACBP-60 activity either did not bind heparin–sepharose or eluted from the column between 0.2M and 0.3M $(NH_4)_2SO_4$ (Fig 2A lanes 2–6). The complex formed in these fractions revealed a denatured product of an apparent molecular weight of 65–70kDa, after UV crosslinking of the labelled T-strand ARS sequence of 31 nt and analysis by SDS-PAGE. The slower migrating complex that eluted from heparin-sepharose between 0.26M and 0.4M $(NH_4)_2SO_4$ (Fig. 2A, lanes 5–11), revealed a UV crosslinked product that migrates with apparent molecular weight 75–80kDa. Subsequent purification by affinity chromatography shows that these are two distinct polypeptides with different binding properties. One polypeptide has an apparent molecular weight of 60kDa and is called ACBP-60; the other an apparent mobility of 67kDa, and was called ACBP-67, against which an antibody was raised (see Figures 2B and C).

ARS consensus binding protein ACBP-67 is polyadenylate binding protein 1

Fractions that contained the slower migrating binding activity and were free of contamination by other binding activities (Fig 2, lanes 7–11), were pooled and further purified over a DNA affinity column of the 31 nt T-rich single strand oligonucleotide

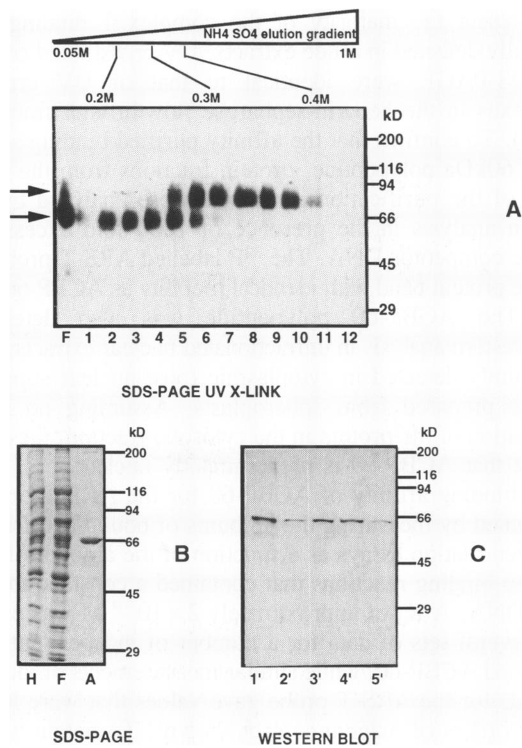


Figure 2. The 67 kDa ssARS binding activity revealed by heparin–sepharose chromatography of yeast whole cell extracts is polyadenylate binding protein-1. **a)** Yeast extracts were subjected to heparin-sepharose chromatography as described in Material and Methods. The column was eluted with a linear gradient of $(NH_4)_2SO_4$ (labelled NH4SO4) and fractions were tested for ARS binding activity by UV crosslinking to the H4 ARS T-rich oligonucleotide followed by SDS-PAGE. The majority of the binding activity detected in crude whole cell and nuclear extracts eluted in the flowthrough from the heparin-sepharose column (lane 1). A minor portion (less than 10 %) of binding activity that comigrated with the one in the flowthrough, eluted between 0.2 and 0.3M $(NH_4)_2SO_4$(lanes 2 to 7). In addition, a previously undetected, slower migrating activity eluted from the column between 0.3 and 0.4M $(NH_4)_2SO_4$ (lanes 6 to 12). The position of the two major crosslinked complexes are indicated by arrows. (Note that the apparent molecular weights of the denatured complexes are due to contributions from both protein and DNA components). **b)** The heparin-sepharose column fractions containing the higher molecular weight binding activity (0.3 to 0.4M $(NH_4)_2SO_4$) were pooled and further subjected to affinity chromatography on a streptavidin–agarose matrix linked to the T-rich strand of the H4 ARS oligonucleotide. The column was eluted with a stepwise gradient of KCl. The binding activity eluted at 0.3M KCl and coincided with the presence of a single polypeptide of 67 kDa. The figure shows a Coomassie stain of protein fractions from three stages of the purification scheme. H; heparin–sepharose 0.3 to 0.4 M pooled fractions, F; DNA affinity column flowthrough, A; affinity purified ACBP-67. **c)** Elution fractions containing the peak of ACBP–67 after DNA affinity chromatography were electrophoresed alongside bacterially expressed yeast PAB. The proteins were transfered to nitrocellulose and the filter was incubated with a polyclonal antibody raised against ACBP-67. Lane 1'; yPAB, lanes 2' to 5' DNA affinity elution fractions containing ACBP-67.

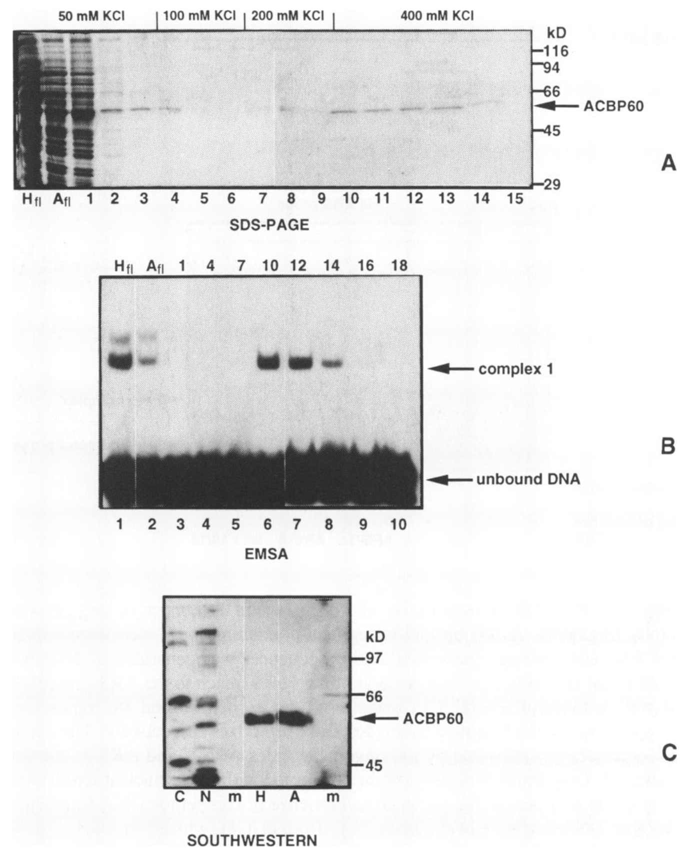


Figure 3. Single strand ARS binding activity coelutes with a 60 kDa polypeptide after sequence specific DNA affinity chromatography. **a)** Flowthrough from the heparin-sepharose column was passed over a matrix of streptavidin-agarose to which the T-rich strand of the H4 ARS was attached. Proteins bound to the column were eluted with a stepwise gradient of KCl as described in Material and Methods and assayed by electrophoresis on SDS-PAGE followed by Coomassie staining. Lane H_{fl}; heparin–sepharose flowthrough, lane A_{fl}; DNA affinity flowthrough, lanes 1 to 3; 50 mM KCl elution, lanes 4 to 6; 100 mM KCl elution, lanes 7 to 9; 200 mM KCl elution, lanes 10 to 15; 400 mM KCl elution. **b)** Gel retardation assays were performed with 1 ng of 5′ end-labelled ARS 307 T-rich oligo incubated with 1μl of each protein fraction. Lane 1; heparin–sepharose flowthrough, lane 2; DNA affinity flowthrough, lane 3; (fraction 1) 50 mM KCl elution, lane 4; (fraction 2) 100 mM KCl elution, lane 5; (fraction 7) 200 mM KCl elution, lanes 6 to 9; (fractions 10, 12, 14 and 16) 400 mM KCl elution, lane 10; (fraction 18) 1M KCl elution. **c)** Fractions from each stage of the purification procedure were electrophoresed on SDS-PAGE and then transferred to nitrocellulose filter. The filter-bound proteins were incubated with 5′ end-labelled H4 ARS T-rich strand in the presence of a large excess of single strand *E.coli* DNA. Filters were washed to remove unbound DNA and exposed to film. C; cytoplasmic extract, N; nuclear extract, H; heparin–sepharose flowthrough, A; affinity purified ACBP-60, M; C^{14} molecular weight markers.

containing the H4 ARS consensus (see Materials and Methods). The binding activity, as assayed by gel retardation of native complexes and by SDS-PAGE of UV crosslinked complexes, copurified with a single polypeptide of M_r = 67 as previously reported (10). Sequence information from three internal peptides (TAEQLENLXI; GFGFVXFX, and FGPIVXAX) generated from the purifed protein allowed a positive identification of the factor as the yeast polyadenylate binding protein (yPAB, 18). To confirm that yPAB was indeed responsible for the binding activity on ARS-T and not just a minor contaminant, we repeated the bandshift and crosslinking assays using the yeast *PAB1* gene expressed in bacteria (gift of A. Sachs) and the ARS-T probe. We found identical bandshifts and crosslinked forms as shown in Figure 2A. Moreover, a polyclonal antibody raised against affinity purified ACBP-67, recognised bacterial expressed yPAB with equal efficiency (Fig 2C). We conclude therefore that the 67kDa T-strand binding activity in yeast is the abundant cytoplasmic and nuclear protein encoded by *PAB1*.

The affinity of yPAB for DNA is significantly lower (100-fold, $K_d > 10^{-7}$ M) than its affinity for poly A RNA. Curiously, the protein will not bind polydA, dC, rU or dG polymers with appreciable affinity. The bandshift observed with purified protein is readily displaced by an excess of random yeast RNA or ssDNA, and it was necessary to perform the affinity chromato-

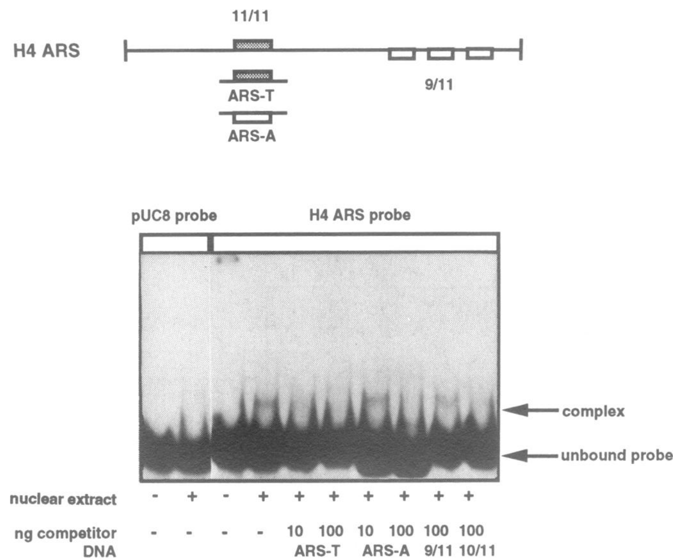


Figure 4. ACBP-60 recognises the ARS consensus in the context of long genomic DNA fragments. A 200 bp DNA fragment containing the T-rich strand of the H4 ARS and authentic genomic flanking sequences was generated as described in Material and Methods. The position of the 11/11 match to the ARS consensus, as well as those of several 9/11 matches on the opposite strand are represented on a diagram of the fragment by shaded and open boxes respectively. The extent of sequences encompassed by the oligonucleotides ARS-T and ARS-A are also indicated. Gel retardation assays were performed with crude nuclear extract and 1 ng of the fragment in the presence and absence of competitor DNA fragments. Another DNA fragment of similar length, (from pUC8), which contained no good match to the ARS consensus, gave no binding. Lane 1; pUC8 DNA probe alone. Lane 2; pUC8 probe incubated with nuclear extract. Lane 3; H4 ARS-T DNA probe alone. Lanes 4 to 10; ARS-T probe incubated with nuclear extract in the absence of competitor (lane 4), in the presence of 10 ng ARS-T oligo (lane 5), 100 ng ARS-T oligo (lane 6), 10 ng ARS-A oligo (lane 7), 100 ng ARS-A oligo (lane 8), 100 ng unrelated oligo containing a 9/11 ARS consensus match (lane 9), 100 ng unrelated oligo containing a 10/11 ARS consensus match (lane 10).

graphy in the absence of competitor DNA to quantitatively bind yPAB. Moreover, a very small proportion of the protein is found in the nuclear fraction (data not shown; 19). While the gene is essential, analysis of ts mutants and extragenic supressors do not suggest a role in DNA replication, but an essential role in the regulation of translation (20). Together these observations suggest that yPAB's interaction with ssDNA is probably not biologically significant.

ARS consensus binding protein ACBP-60 has a $K_d < 10^{-9}$ for T-rich DNA

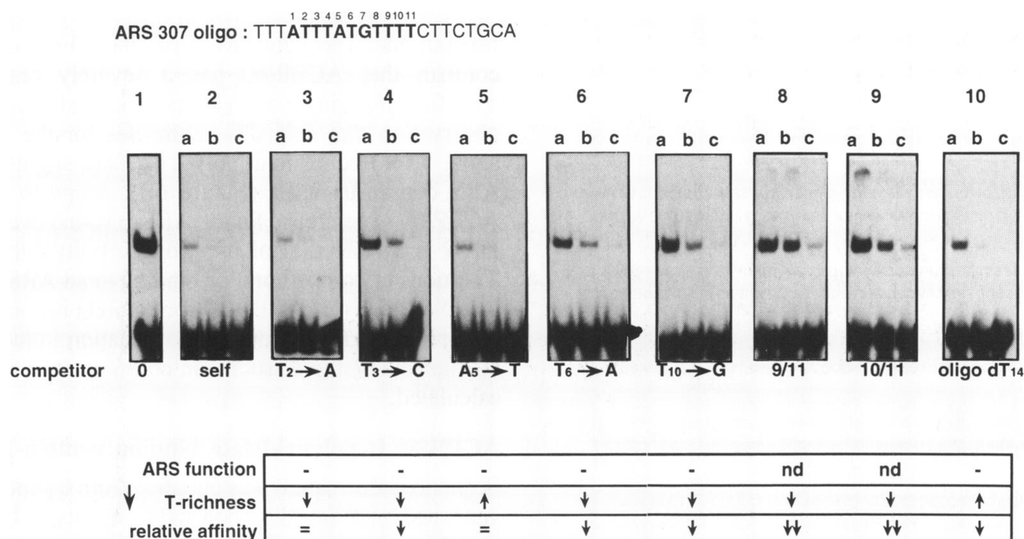
We next purified to homogeneity the second single strand DNA binding protein which upon crosslinking gave rise to the 65−70 kDa complex with the T-rich strand of the H4 ARS consensus (Figure 2A). The flowthrough fraction from the heparin−sepharose was used as a source from which to further purify this activity, since the flowthrough was depleted for yPAB and many other RNA and DNA binding proteins. Heparin-sepharose flowthrough fractions were passed over an ARS-T strand affinity column in the presence of a large excess of non specific single stranded DNA and poly A RNA. Upon sequential elution with increasing concentrations of KCl the binding activity was quantitatively recovered between 0.4 and 0.6M KCl with a single polypeptide of approximately 60kDa (Figures 3A and 3B). This fraction gave rise to a complex in the band shift assay identical to complex 1 shown in Figure 1 (Figure 3B), and appears to represent the majority of the complex 1 binding activity originally detected in crude extracts. UV crosslinked complexes on SDS-PAGE were identical to that of UV crosslinked complexes in the heparin-sepharose flowthrough fraction (not shown). To confirm that the affinity purified binding activity is due to 60kDa polypeptide, protein fractions from the different stages of the purification procedure were analysed by Southwestern analysis in the presence of 1000 fold excess of non specific competitor DNA. The ^{32}P labelled ARS-T probe bound a single protein band with identical mobility as ACBP-60 (Figure 3C). The ACBP-60 polypeptide was also detected by Southwestern analysis in unfractionated nuclear extracts, but was only faintly detected in cytoplasmic (post nuclear supernatant) extracts prepared from spheroplasts. Assuming no selective degradation of the protein in the cytosolic fraction, these results suggest that ACBP-60 is predominantly nuclear.

The binding affinity of ACBP-60 for the ARS-T probe was determined by measuring the amounts of bound and free probe in gel retardation assays as a function of the amount of protein added to binding reactions that contained a constant amount of probe DNA. A K_d of approximately 2×10^{-9} M was estimated from several sets of data for a number of independent batches of purified ACBP-60, while similar measurements of the affinity of yPAB for the ARS-T probe gave values that were lower by several orders of magnitude (not shown). To obtain a relative assessment of the affinity of ACBP-60 compared to that of a known single strand binding protein RF-A (Replication Factor-A, see 1, for review), the two were compared in the presence of a titration of non-specific competitor DNA on the same gel. The ACBP-60 complex migrates slightly faster than the RFA−DNA complex (data not shown), while its affinity for the T-rich strand of the ARS consensus is comparable to that previously determined for RF-A binding to T-rich ssDNA (21). Thus the dissociation constant of ACBP-60 is within the range expected for a single strand DNA−protein interaction of biological significance.

Affinity for a T-rich oligonucleotide does not necessarily mean that the protein can find its target in the midst of a longer stretch of DNA. For this reason we screened for an activity that recognises the ARS core consensus within the context of longer genomic fragments in nuclear extracts. The binding activity is sequence specific since it is detected in the presence of 5000×(w/w) of non-specific single stranded competitor DNA, but is competed by only 10 fold excess of the unlabelled ARS-T oligonucleotide (Fig 4). In addition, the observation that the protein is displaced by annealing of the complementary ARS-A oligonucleotide (Fig 4), suggests that it binds only the ss T-rich ARS consensus. Unrelated T-rich single strand oligonucleotides which also contain 9/11 and 10/11 matches to the ARS consensus compete for binding but with sigificantly reduced efficiency than bona fide ARS consensus (Fig. 4, last two lanes). One hundred-fold excess of oligo dT also competes, whereas oligo dA does not (data not shown). Purified ACBP-60 shows a similar activity, while purified ACBP-67 produces no complex with long ARS containing genomic fragments.

ACBP-60 binds poly-dT and poly-rU with high affinity

In previous publications on yeast ss ARS binding proteins, the sensitivity of binding to point mutation in the ARS consensus was invoked as evidence that the protein might be involved in ARS function (10−12). For this reason we have carefully compared the abilities of various oligonucleotides of different



competitor	T2→A	T3→C	A5→T	T6→A	T10→G	9/11	10/11	oligo dT14
ARS function	-	-	-	-	-	nd	nd	-
↓ T-richness	↓	↓	↑	↓	↓	↓	↓	↑
relative affinity	=	↓	=	↓	↓	↓↓	↓↓	↓

Figure 5. The effect of point mutations in the ARS 307 on ACBP-60 binding. Gel retardation assays were performed with 1 ng of 5′ end-labelled 22 nucleotide oligomer containing the T-rich strand of the ARS consensus of ARS 307 surrounded by its authentic flanking sequences and 0.5 ng of purified ACBP-60. The sequence of the oligonucleotide is shown above with the 11 nucleotide ARS consensus indicated in bold letters. Binding reactions were done in the presence of single strand *E.coli* DNA either without the inclusion of sequence specific competitor DNA (panel 1), in the presence of unlabelled oligonucleotide containing the wildtype ARS 307 sequence (panel 2), in the presence of unlabelled oligonucleotide containing single point mutants of the ARS 307 sequence (panels 3 to 7), in the presence of unrelated oligonucleotides containing a 9/11 or a 10/11 match (panels 8 and 9) to the ARS consensus, or in the presence of a 14 nucleotide oligomer of dT (panel 10). Lanes marked a, b and c represent a molar excess of sequence specific competitor of 10, 50 and 200 fold respectively. In other experiments where 44 nucleotide oligomers containing either wildtype or point mutant versions of the T-rich strand of the ARS consensus of ARS 307 were used, similar results were obtained (not shown). The effects of the individual mutations with regard to performance in functional assays (VanHouten and Newlon, 1990), T-richness and p60 binding affinity relative to the wildtype ARS 307 oligo are indicated in the table.

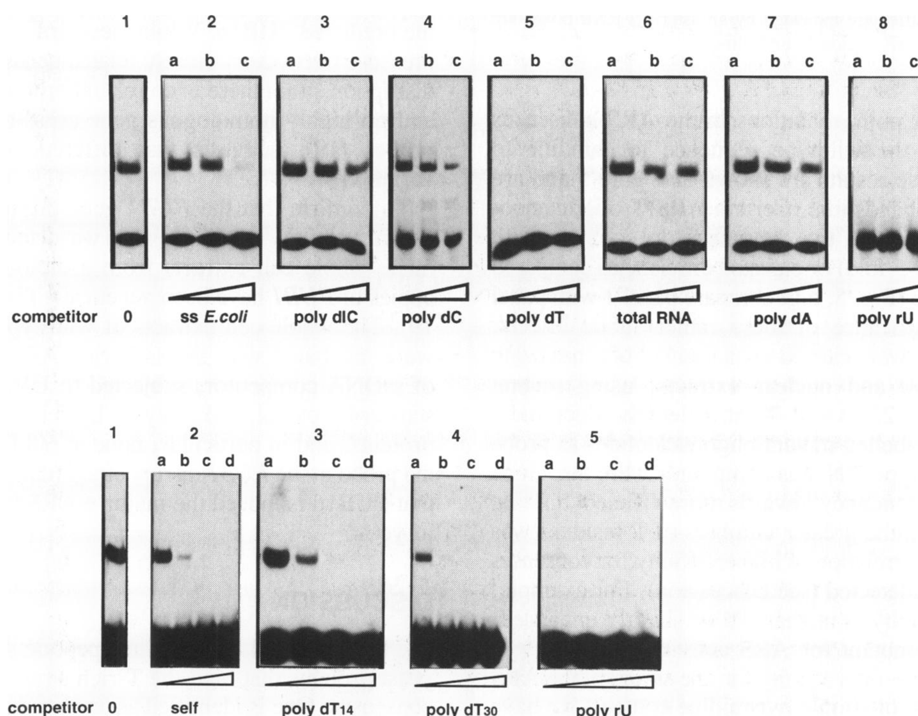


Figure 6. ACBP-60 binding to the ARS consensus element is efficiently competed by poly dT and poly rU. Gel retardation assays were performed with 1 ng of 5′ end-labelled 31 nucleotide oligomer containing the T-rich strand of the H4 ARS concsensus surrounded by its authentic flanking sequences and 0.5 ng of purified ACBP-60. **a)** Binding reactions were done either in the absence of competitor (panel 1), or in the presence of competitor DNA or RNA, (single strand *E.coli* DNA, panel 2), poly dIC (panel 3), poly dC (panel 4), poly dT (panel 5), total yeast RNA (panel 6), poly dA (panel 7), poly rU (panel 8). Lanes marked a, b and c represent 1, 5 and 10 μg of competitior respectively. **b)** Binding reactions were carried out either in the absence of competitor (panel 1) or in the presence of unlabelled H4 ARS oligo (panel 2), poly dT_{14} (panel 3), poly dT_{30} (panel 4), poly rU (panel 5). Lanes marked a, b, c and d represent 10, 50, 100 and 200 ng of competitor respectively.

```
PUB1     MSENNEEQHQQQQQQPVAVETPSAVEAPASADPSSEQSVAVEGNSEQAE
RNP1     MSENNEEQHQQQQQQPVAVETPSAVEAPASADPSSEQSVAVEGNSEQAE

                            *     *     RNP-2
PUB1     DNQGENDPSVVPANAITGGRETSDRVLYVGNLDKAITEDILKQYFQVGGP
RNP1     DNQGENDPSVVPANAITGGRETSDRVLYVGNLDKAITEDILKQYFQVGGP
ACBP60                  XIIGGREESDRVLYVGNXDK        KQYFQVGGP
                        RNP-1
PUB1     IANIKIMIDKNNKNVNYAFVEYHQSHDANIALQTLNGKQIENNIVKINWA
RNP1     IANIKIMIDKNNKNVNYAFVEYHQSHDANIALQTLNGKQIENNIVKINWA
ACBP60   IANIK
                       RNP-2
PUB1     FQSQQSSSDDTFNLFVGDLNVNVDDETLRNAFKDFPSYLSGHVMWDMQTG
RNP1     FQSQQSSSDDTFNLFVGDLNVNVDDETLRNAFKDFPSYLSGHVMWDMQTG

           RNP-1                                   *
PUB1     SSRGYGFVSFTSQDDAQNAMDSMQGQDLNGRPLRINWAAKRDNNNNNNYQ
RNP1     SSRGYGFVSFTSQDDAQNAMDSMQGQDLNGRPLRINWAAKRDNNNNNNYQ
ACBP60                                     INWADK
                  GAR   *
PUB1     QRRNYGNNNRGGFRQYNNNNNNNMNMGMNMNMNMNMNNSRGMPPSSMGMP
RNP1     QRRNYGNNNRGGFRQYNSNNNNNMNMGMNMNMNMNMNNSRGMPPSSMGMP
ACBP60                                                  GMP
                                                   RNP-2
PUB1     IGAMPLPSQGQPQQSQTIGLPPQVNPQAVDHIIRSAPPRVTTAYIGNIPH
RNP1     IGAMPLPSQGQPQQSQTIGLPPQVNPQAVDHIIRSAPPRVTTAYIGNIPH
ACBP60   IGAMPLPXQ
                                   RNP-1
PUB1     FATEADLIPLFQNFGFILDFKHYPEKGCCFIKYDTHEQAAVCIVALANFP
RNP1     FATEADLIPLFQNFGFILDFKHYPEKGCCFIKYDTHEQAAVCIVALANFP

                                      *
PUB1     FQGRNLRTGWGKERSNFMPQQQQQGGQPLIMNDQQQPVMSEQQQQQQQQ
RNP1     FQGRNLRTGWGKERSNFMPQQQQQGGQHL
ACBP60                   SNFMPQXXLXGGQPL

PUB1     QQQ
```

Figure 7. ACBP-60 is homologous to genes *PUB1* and *RNP1*. Six internal peptide sequences were obtained from tryptic digestion and CnBr cleavage of affinity purified ACBP-60. The peptide sequences have been aligned with those of the predicted poly U binding proteins encoded by *PUB1* and *RNP1*. Positions at which any of the three sequences differ are indicated by an asterisk. RNP boxes are indicated in italics and the GAR region is underlined.

sizes and with different point mutations in the ARS consensus core to compete with the wildtype sequence for binding to ACBP-60. Some of these results are shown in Figure 5 and are summarized below the bandshifts. Certain ARS-T oligonucleotides with point mutations in the consensus do compete with slightly reduced (at most 5-fold) or elevated efficiencies compared to the wildtype sequence (Fig. 5, e.g. compare panel 1 with panel 8). To assess the significance of these rather subtle effects, competition experiments were carried out a number of times using both purified ACBP-60 and nuclear extracts, using mutant oligonucleotides of both 22 nt and 44 nt in length, Reciprocal experiments using ^{32}P labelled mutant oligonucleotides as probe and the unlabelled wildtype DNA as competitor were also done (not shown). A general tendancy towards more efficient binding by oligonucleotides with the greater numbers of T residues was observed, but no strict correlation of higher affinity for consenses with ARS function was detected (see Discussion). For example, the A5 to T mutant affinity is unchanged or slightly enhanced, although this is a null mutant for ARS activity (6).

To test whether ACBP−60's affinity for one strand of the ARS consensus is a reflection of simple pyrimidine content, we have screened a series of simple nucleotide polymers as competitors for the ARS T-strand binding activity (Fig. 6A and 6B). Both poly dT and poly rU were extremely effective competitiors of ACBP-60 binding to ARS-T strands, competing efficiently at 10 to 50-fold (w/w) excess over the probe (Figure 6B). All other homopolymers of ribo- and deoxyribonucleic acids tested were very inefficient competitors with little or no competition at 1000 to 10000-fold weight excess of competitor, as observed for random mixed sequence RNA or DNA (Figure 6A). The results confirm that ACBP-60 has a severely restricted sequence specificity for single stranded nucleic acid (poly dT or rU). The observation that poly dT_{30} competes for the ARS T-sequence about 5-fold better than poly dT_{14}, and slightly better than the ARS consensus itself (Figure 6B), suggests that the basis of ACBP-60's preference for the ARS-T strand over mixed sequence DNA is a reflection of its strong affinity for oligo dT tracts. Titrations of competitors for binding to an ARS-T probe suggest that the affinity of ACBP-60 for poly rU is about 10 fold higher than for poly dT, but due to the variation in the size of the poly rU molecules, the dissociation constant could not be accurately calculated.

ACPB-60 is polyuridylate binding protein-1

Sequence information was obtained from 6 peptides of ACBP-60 generated by tryptic digestion and by cyanogen bromide cleavage. Comparison of these sequences with the theoretical translation of DNA sequences in the EMBL data base suggest ACBP-60 is highly homologous, if not identical, to two recently identified open reading frames (*RNP1* and *PUB1*, probably defining the same gene), encoding RNA binding protein(s) of unknown function (19, 22, 23). The alignment of these peptides with the predicted proteins based on sequencing the two independently cloned genes is shown in Figure 7. Asterisks indicated the variations in translation product between *RNP1*, *PUB1* and the peptide sequence results from ACBP-60. In three cases the ACBP-60 amino acid sequence diverges from the two published genes, and in two instances (the termination codon and the codon two amino acids upstream) RNP1p diverges from ACBP-60 and the predicted PUB1p. While these differences could indicate the presence of two very closely related genes the fact that in a *PUB1* disruption strain there is no related protein detected by antibody, and no highly homologous gene cross-hybridizing with *PUB1* probes (19), suggests that differences reflect strain-specifc variation.

To confirm that the *PUB1* gene product is the only ss ARS binding activity of 60kDa that we detect in yeast extracts, we made whole cell extracts from a diploid strain in which both copies of *PUB1* have been genetically disrupted (strain YJA504, 19). The whole cell extracts of wild type and disrupted strains were incubated with the ss T-rich ARS probe in the presence of ssDNA competitor, subjected to UV crosslinking using our standard protocol, and analysed by gel electrophoresis. No major proteins, and in particular, none at 60kDa, were crosslinked to the probe in the *PUB1* disrupt strain (data not shown), confirming that PUB1p is indeed the major ssARS binding protein detected in yeast.

DISCUSSION

We have purified and obtained peptide sequence data from two yeast proteins that bind the T-rich strand of the conserved ARS consensus. One is identified as an abundant cytosolic protein that binds poly A RNA with high affinity, and poly dT selectively, but with relatively low affinity. This protein is poly-adenylate binding protein 1, which associates *in vivo* with polysomes and polyadenylated messenger RNA. Its selective affinity for T-rich single stranded DNA (i.e. it will not bind poly dA, dG, nor dC) is unexpected and probably fortuitous. The second protein purified binds both selectively and with high affinity to the T-rich strand of the yeast ARS consensus, although it will also bind

poly dT and poly rU single stranded nucleic acids. Sequence analysis of this polypeptide shows that it is encoded by the recently cloned yeast gene *PUB1* (19, 23), or *RNP1* (22). The protein was identified in UV crosslinking studies as one of the major proteins associated with messenger RNA in yeast cells (19, 23). In contrast to yPAB, PUB1p is found in both the nuclear and cytosolic compartments, and does not copurify with polysomes, nor to a significant extent with monosomes (19). The majority of the cytosolic PUB1 protein is in a fraction <30S, which could represent either free protein or message-bound protein, but it is unlikely to be in a ribosomal complex. Bandshift studies on gels that determine complex size, suggest that the protein binds as a monomer (M.C., data not shown). Although it has recently been demonstrated that PUB1p binds tightly to polyU RNA (as opposed to rC, rG, or rA) its affinity for ssDNA was previously unrecorded.

PUB1 binds poly dT sequences with high affinity

We have demonstrated that PUB1p binds with high affinity (10^{-10} M > K_d < 10^{-9} M) to the T-rich strand of naturally occurring sequences containing the ARS consensus ($^{A}/_{T}$ TTTAT $^{G}/_{A}$ TTT $^{A}/_{T}$) of *S.cerevisiae*, while it will not bind with any detectable affinity to the A-rich complementary strand nor to homopolymers of dA, dC or dIC. We can detect the PUB1p activity in whole cell and in nuclear extracts of yeast, in the presence of 5000× excess of ss *E.coli* DNA and polyA RNA. PUB1p can bind the T-rich ARS consensus in a large stretch of single stranded genomic DNA (see Figure 4; and ref 10), and point mutations diminish the affinity with which PUB1p binds the ARS consensus, although in most cases these also reduce the T-content of the DNA probe. Consistently, we have observed that the purified *PUB1* protein binds a thymidine 30-mer with slightly higher affinity than the ARS consensus.

What is the cellular role of PUB1p?

It is not clear whether the DNA binding activity of PUB1p is of physiological significance. Three recent publications have reported the sequence and gene disruption of the *PUB1* gene (or *RNP1*) and none has found any phenotype resulting from the absence of the protein product (19, 22, 23) Either the protein plays no essential role in cell division, or other genes provide complementing activities. Although PUB1 protein can be crosslinked to the polyA messenger population, it does not bind the polyA tail itself, nor does not bind the same messenger RNA population as yPAB (19). Its affinity for poly rU as opposed to any other ribohomopolymer, and its low affinity for random RNA sequence, suggests that it may bind to specific sequence signals or secondary structure in the messenger RNA population, suggesting a role in transcriptional termination, splicing or mRNA transport. Although the protein binds to the T-rich strand of the ARS consensus, there is no strong evidence implicating *PUB1* in the mechanism of DNA replication, and any role it might have in S-phase, must be redundant. Nonetheless, the selectivity and affinity of PUB1p for ss T-rich DNA is sufficiently high that it is reasonable to propose a dual RNA/DNA binding activity *in vivo*. This has been proposed for a series of proteins such as the bacteriophage T4 gene 32 protein (reviewed in 24), nucleolin (e.g. 25), the yeast SSB1 protein (26), Lupus antigen (La; 27, 28), and the *Drosophila* DssRP (29).

Structural motifs

PUB1p has three extended RNA binding domains (RBD), which each contain two short motifs called RNP boxes. These domains are found in many RNA and ssDNA binding proteins (reviewed in 30). The RNP-1 and the RNP-2 motifs are indicated by italics in Figure 7. Those from a.a. 86 to 125 and from aa. 163 to 213 are the more conventional RNP boxes. In addition *PUB1* contains an extensive stretch of asparagine in the middle of the protein, and of glutamine at its C-terminus, which are not found in any of the reported RNP proteins analyzed to date (e.g 30, 31). Finally the GAR motif that is common to a number of RNA and ssDNA binding proteins is located between the second and third RBD's of PUB1p (aa. 260–265).

The RNP-1 motif appears to be involved directly in ss nucleic acid binding and, with its flanking loop regions, in the determination of specificity (32). It has been demonstrated, for instance, that one repeat of the RBD motif is sufficient in the yeast PAB1 protein to maintain polyA binding activity (33). In the U1 70kDa protein, as well, a single RBD with minimal flanking sequence binds specifically to U1RNA (34). When the RNP-1 box of *PUB1* is compared with other RNP-containing proteins, the most homologous to the first and second copies of the RNP-1 motif is either that of the *Drosophila* protein *ELAV*, or repeat three of mouse nucleolin (see summary, 31). Although *PUB1* is clearly not the yeast homologue of nucleolin, it is noteworthy that the region flanking RNP-2 also has high homology to that found in nucleolin, both contain a GAR motif (glycine and arginine rich stretch, see 22), and both bind DNA and RNA (35). In strong contrast to nucleolin, however, the C-terminus of PUB1p is proline-rich, and the N-terminus negatively charged.

Roles for ssDNA and RNA binding proteins

An ever-increasing number of presumably sequence-specific single strand DNA binding proteins have been reported in the last three years. Examples include MF3 and myoD (36), which are thought to function in transcriptional activation; RNP factors binding upstream of *c-myc* (37); two d(TC)$_n$ binding proteins (38, 39); PYBP, a pyrimidine-rich ss binding protein from human cells (40); the *Drosophila* DssRP binding activity which binds preferentially G and U residues (29), as well as a series of ss binding proteins thought to recognize sequences at replication origins. These include: ssARS binding factors from yeast (10–13); the human purine binding protein, *pur* (41); a single stranded DNA binding protein from *Crithidia* (42); and a factor that binds the inverted repeat domain of the SV40 origin in single stranded form (43). In none of the above cases has the presumed function of the protein, based on its sequence preference, been confirmed by *in vivo* experiments.

Better substantiated is the biological role for poly rU binding proteins, such as the lupus antigen La (27, 28), or factors that bind to premessenger RNA that determine splicing pattterns, such as the fly genes *sex-lethal*, or *transformer 2* (44, 45). Since PUB1p binds poly rU it is sensible to decipher its role *in vivo* by asking where one finds uridinyl tracts. Recently 3′ terminal uridylyltransferase activity has been demonstrated in the nucleus of eukaryotic cells (46), and natural substrates appear to be the U6 small nuclear RNA, other primary transcripts from pol III in *Xenopus* (47), transcripts of Alu families (48), 7sRNA (49), and certain viral transcripts (e.g. 50). In the case of PolIII transcripts, the poly rU tract appears to be a signal for transcriptional termination, involving the La protein (27, 28, 51). In the RNA editing reaction in *Tetrahymena* mitochondria, the role of terminal uridylyltransferase in extending the poly U tail of guide RNAs has been demonstrated *in vitro* (52). In view of its binding to both ss DNA and RNA, perhaps the most likely

role of PUB1p is in message termination, for which, however, there must exist redundant activities.

Are there other single-strand ARS binding proteins in yeast?

Reports of ssDNA binding proteins with affinity for sequences found near origins of replication have been published but in every case the identification of the protein is based exclusively on its affinity *in vitro* for the single stranded sequence, and no convincing evidence has yet been put forth showing that these proteins play a role in DNA replication *in vivo*. It was to this end that we purified and sequenced the major ssARS sequence binding activities, ACBP-60 and ACBP-67, both of which were present in the earlier preparations of ACBP published from this laboratory (10). Although ACBP-60 was thought to be a breakdown product of the larger polypeptide, sequence data shows this not to be the case. We suggest that the ss ARS-T binding protein of Schmidt *et al.* (11), is also identical to PUB1p, since the purified protein was also around 60kDa and both poly dT and poly rU compete somewhat for ss ARS-T binding activity. However, the strong differences (apparently > 5-fold) in affinity reported by Schmidt *et al.* for point mutations in the ARS consensus were not reproducible in our hands with purified PUB1p, even using oligos kindly provided by G.Krauss. The third protein reported but not purified during our studies, migrates around 45kDa when crosslinked to the ssARS probe and shows even lower specificity for the ARS consensus. This may have been detected by others in bandshift assays (e.g. 12, 13).

Thus far the highest affinity ss ARS binding proteins that we have purified from either yeast nuclei or cytosolic fractions are RNP box containing proteins that also bind RNA. We did not purify RF-A by our approach since affinity chromatography was done in excess of random sequence ss DNA. Although we can not rule out that a low abundance nuclear protein that binds the ARS consensus has escaped our detection, it is clear that there was only one ss ARS binding activity associated with the nuclear scaffold (10), and that this is almost certainly identical to PUB1p.

ACKNOWLEDGEMENTS

We would like to thank M.Swanson and G.Dreyfuss for communicating preprints prior to publication, G.Krauss for ARS oligonucleotides, A.Sachs for samples of recombinant yPAB, J.Diffley for samples of yeast Replication Factor A, P.Legrain for anti-human PTB, J.Hofmann for anti-ACBP-67, and M.Swanson for the disrupted *PUB1* strain. We would like to thank Gian-carlo Alghisi for help in calculating dissociation constants, all members of the Gasser lab for helpful discussions and V. Simanis for carefully reading the manuscript. This research was supported by the Swiss National Science Foundation grant to SMG and in part by a Roche Research Fellowship granted to Moira Cockell.

REFERENCES

1. Campbell, J.L. and Newlon, C.S. (1993) In Broach, J.R., Pringle, J. and Jones, E. (eds.), The Molecular and Cellular Biology of the Yeast *S. cerevisiae*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 41–147.
2. Celniker, S.E., Sweder, K., Srienc, F., Bailey, J.E. and Campbell, J.L. (1984) *Mol. Cell. Biol.* **4**, 2455–2466.
3. Kearsey, S. (1984) *Cell*, 37, 299–307
4. Bouton, A.H. and Smith, M (1986) *Mol. Cell. Biol.* 6, 2354–2363
5. Palzkill, T.G. and Newlon, C.S. (1988) *Cell* **53**, 441–450.
6. Van Houten, J.V. and Newlon, C.S. (1990) *Mol. Cell. Biol.*, **10**, 3917–3925.
7. Deshpande, A.M. and Newlon, C.S. (1992) *Mol. Cell. Biol.* **12**, 4305–4313.
8. Bell, S. and Stillman, B. (1992) *Nature*, **357**, 128–134.
9. Diffley, J.F.X. and Cocker, J.H. (1992) *Nature*, **357**, 169–172.
10. Hofmann, J.F.X. and Gasser, S.M. (1991) *Cell*, **64**, 951–960
11. Schmidt, A.M.A., Heterich, S.U. and Krauss, G. (1991) *EMBO J.* **10**, 981–985.
12. Kuno, K., Kuno, S., Matsushima, K. and Murakami, S. (1991) *Mol. Gen. Genet.*, **230**, 45–48.
13. Kuno, K., Murakami, S. and Kuno, S. (1990) *Gene*, **95**, 73–77
14. Amati, B.A. and Gasser, S.M, (1988) *Cell*, **54**, 967–978.
15. Hofmann, J.F.X., Laroche, T., Brand, A.H. and Gasser, S.M. (1989) *Cell* **57**, 725–737
16. Hughes, G.J, Frutiger, S., Paquet, N. and Jaton, J.C. (1990) *Biochem. J.* **271**, 641–647
17. Laemmli, U.K. (1970) *Nature*, **227**, 680–685.
18. Sachs, A.B., Bond, M.W. and Kornberg, R.D. (1986) *Cell*, **45**, 827–835.
19. Anderson, J.T., Paddy, M.R. and Swanson, M.S. (1993) *Mol. Cell. Biol.* **13**, 6102–6113.
20. Sachs, A.B.and Davis, R.W (1989) *Cell*, **58**, 857–867.
21. Kim, C., Snyder, R.O. and Wold, M.S. (1992) *Mol. Cell. Biol.* **12**, 3050–3059.
22. Ripmaster, T.L. and Woolford, J.L., Jr (1993) *Nucleic Acids Res.*, **21**, 3211–3216.
23. Matunis, M.J., Matunis, E.L. and Dreyfuss, G. (1993) *Mol. Cell. Biol.* **13**, 6114–6123.
24. Chase, J.W. and Williams, K.R. (1986) *Annu. Rev. Biochem.*, **55**, 103–136.
25. Barrijal, S., Perros, M., Gu, Z., Avalosse, B.L., Belenguer, P., Amalric, F. and Rommelaere, J. (1992) *Nucleic Acids Res.*, **20**, 5053–5060.
26. Jong, AY.-S., Clark, M.W., Gilbert, M., Oehm, A. and Campbell, J.L. (1987) *Mol. Cell. Biol.* **7**, 2947–2955.
27. Gottleib, E. and Steitz, J.A. (1989) *EMBO J.* **8**, 841–850.
28. Gottleib, E. and Steitz, J.A. (1989) *EMBO J.* **8**, 851–861.
29. Hsu, T., King, D.L., LaBonne, C. and Kafatos, F.C. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 6488–6492.
30. Haynes, S, R. (1992) *New Biol.* **4**, 421–429
31. Kim, Y.J. and Baker, B.S. (1993) *Mol. Cell. Biol.*, **13**, 174–183
32. Scherly, D., Boelens, W., van Venrooij, W.J., Dathan, N.A., Hamm, J. and Mattaj, I.W. (1989) *EMBO J.*, **8**, 4163–4170.
33. Sachs, A.B., Davis, R.W. and Kornberg, R.D. (1987) *Mol. Cell. Biol.*, **7**, 3268–3276.
34. Query, C.C., Bentley, R.C. and Keene, J.D. (1989) *Cell* **57**, 89–101.
35. Ishikawa, F., Matunis, M.J., Dreyfuss, G. and Cech, T.R. (1993) *Mol. Cell. Biol.*, **13**, 4301–4310.
36. Santoro, I.M., Yi, T.-M., Walsh, K.(1991) *Mol. Cell. Biol.* **11**, 1944–1953
37. Davis, T.L., Firulli, A.B. and Kinniburgh, A.J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9682–9686
38. Yee, H.A., Wong, A.K.C., van de Sande, J.H. and Rattner, J.B. (1991) *Nucl. Acids Res.* **19**, 949–953.
39. Kolluri, R, Torrey, T.J A. and Kinniburgh, A.J. (1992) *Nucleic Acids Res.* **20**, 111–116
40. Brunel, F., Alzari, P.M., Ferrara, P. and Zakin, M.M. (1991) *Nucl. Acids Res.* **19**, 5237–5245.
41. Bergemann, A.D., Ma, Z. and Johnson, E.M. (1992) *Mol. Cell. Biol.* **12**, 5673–5682.
42. Tzfati, Y., Abeliovich, H., Kapeller, I. and Shlomai, J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6891–6895.
43. Carmichael, E.P., Roome, J.M. and Wahl, A.F. (1993) *Mol. Cell. Biol.* 13, 408–420.
44. Bell, L.R., Maine, E.M., Schedl, P. and Cline, T.W. (1988) *Cell* **55**, 1037–1046.
45. Amrein, H., Maniatis, T. and Nothiger, R. (1990) *EMBO J.*, **11**, 3619- 3629.
46. Tazi, J., Forne, T., Jeanteur, P., Cathala, G. and Brunel, C. (1993) *Mol. Cell. Biol.* **13**, 1641–1650.
47. Rinke, J. and Steitz, J.A. (1982) *Cell*, **29**, 149–159.
48. Shen, C.J. and Maniatis, T. (1982) *J. Mol. Appl. Genetics.* **1**, 343–360.
49. Chambers, J.C., Kurilla, M.G. and Keene, J.D. (1983) *J. Biol. Chem.* **258**, 11438–11441.
50. Lerner, M.R., Andrews, N.C., Miller, G. and Steitz, J.A. (1981)*Proc. Natl. Acad. Sci. USA* **78**, 805–809.
51. Terns, M.P., Lund, E. and Dahlberg, J.E. (1992) *Mol. Cell. Biol.*, **12**, 3032–3040.
52. Pollard, V.W., Harris, M.E. and Hajduk, S.L (1992) *EMBO J.* 11, 4429–2238.