PUF60: a novel U2AF65-related splicing activity

P. S. Page-McCaw, K. Amonlirdviman and P. A. Sharp

RNA 1999 5: 1548-1560

RNA (1999), *5*:1548–1560. Cambridge University Press. Printed in the USA.

PUF60: A novel U2AF65-related splicing activity

PATRICK SCHONLEBER PAGE-McCAW,[1,2] KEVIN AMONLIRDVIMAN,[1,3] and PHILLIP A. SHARP[1]

[1]Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

ABSTRACT

We have identified a new pyrimidine-tract binding factor, PUF, that is required, together with U2AF, for efficient reconstitution of RNA splicing in vitro. The activity has been purified and consists of two proteins, PUF60 and the previously described splicing factor p54. p54 and PUF60 form a stable complex in vitro when cotranslated in a reaction mixture. PUF activity, in conjunction with U2AF, facilitates the association of U2 snRNP with the pre-mRNA. This reaction is dependent upon the presence of the large subunit of U2AF, U2AF65, but not the small subunit U2AF35. PUF60 is homologous to both U2AF65 and the yeast splicing factor Mud2p. The C-terminal domain of PUF60, the PUMP domain, is distantly related to the RNA-recognition motif domain, and is probably important in protein–protein interactions.

Keywords: p54; PUF60; pump domain; pyrimidine tract; RRM; splicing factor; U2AF

INTRODUCTION

The branch sequence is recognized at least twice during spliceosome assembly, first by the branch sequence binding protein SF1/BBP and, subsequently, by U2 snRNP. The stable binding of U2 snRNP to the branch sequence is facilitated by U2 snRNP Auxiliary Factor (U2AF), which recognizes the pyrimidine tract (Ruskin et al., 1988; Zamore & Green, 1989; Roscigno et al., 1993). U2AF consists of two proteins of molecular weight 65 and 35 kDa (Zamore & Green, 1989) and it can be depleted from extracts by taking advantage of the larger subunit's very stable interaction with poly U RNA (Zamore & Green, 1989). The small subunit of U2AF enhances the activity of U2AF65 subunit in splicing (Zuo & Maniatis, 1996). U2AF is highly conserved in evolution both in sequence and function. The invertebrates *Drosophila melanogaster* (Kanaar et al., 1993) and *Caenorhabditis elegans* (Zorio et al., 1997) and the fission yeast *Schizosaccharomyces pombe* (Potashkin et al., 1993) have clear homologs of the large subunit of U2AF. The *Saccharomyces cerevisiae* gene MUD2 is a probable homolog of U2AF65; however, its sequence has diverged significantly from U2AF65 (Abovich et al., 1994). A homolog of U2AF35 has not been identified in *S. cerevisiae* but homologs are found in *D. melanogaster* (Rudner et al., 1996), *C. elegans* (Zorio & Blumenthal, 1999) and *S. pombe* (Wentz & Potashkin, 1996) as well as in vertebrates. The large and small subunits of U2AF are each known to be members of protein families. The hepatocarcinoma autoantigen HCC1 is homologous to U2AF65 (Imai et al., 1993), whereas the human protein Urp is homologous to U2AF35 (Tronchere et al., 1997)

U2AF and U1 snRNP, bound to the pyrimidine tract and 5′ splice site, respectively, together with several other proteins, including SR proteins, are part of the E, or commitment, complex that forms prior to the first ATP-dependent spliceosome assembly step (Michaud & Reed, 1991; Staknis & Reed, 1994). These complexes are precursors to the ATP-dependent assembly of U2 snRNP on the branch sequence, the A complex. An analog of the A complex, A3′, forms on 3′ half substrate pre-mRNA consisting of the branch sequence, pyrimidine tract, 3′ splice site, and 3′ exon (Barabino et al., 1990; Zamore & Green, 1991). An ATP-independent U2 snRNP-containing complex, A_{min}, forms on minimal branch sequence–pyrimidine tract RNAs and may represent an intermediate in A complex assembly (Query et al., 1997). The formation of the A_{min} complex is more sensitive to changes in substrate RNA se-

Reprint request to: Phillip A. Sharp, Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA; e-mail: sharppa@mit.edu.

[2]Present address: Department of Physiology, University of California, San Francisco, 513 Parnassus Avenue, Box 0444, San Francisco, California 94131-0444, USA.

[3]Present address: Department of Molecular & Cell Biology, University of California, Berkeley, 401 Barker Hall #3204, Berkeley, California 94720-3204, USA.

quences than is the ATP-dependent A complex, but these two complexes are quite similar in protein–RNA contacts as judged by photochemical cross-linking (Query et al., 1997).

Many splicing factors are concentrated in subregions of the nucleus known as "speckles." Repeats of the serine-arginine (SR) dipeptide sequence are an important signal for this subnuclear localization and the splicing factor p54 contains such repeats (Chaudhary et al., 1991; Zhang & Wu, 1996). p54 is conserved between vertebrates and the invertebrates *D. melanogaster* and *C. elegans* (Kennedy et al., 1998). p54 has a C-terminal domain rich in SR dipeptide repeats and contains an RNA-recognition motif (RRM) indicative of RNA-binding activity (Chaudhary et al., 1991). p54, like the general class of SR proteins, can activate S100 cytoplasmic extracts for splicing in vitro (Zhang & Wu, 1996). Unlike the general class of SR proteins, p54 associates with U2AF65, but does not associate with the 70-kDa protein bound to U1 snRNP, nor with the U2AF35 subunit (Zhang & Wu, 1996). The *Drosophila* homolog of p54 has been shown to bind C-rich pyrimidine tracts in short introns and promote the binding of U1 and U2 RNPs to the flanking splice sites (Kennedy et al., 1998).

We have identified a new pyrimidine-tract binding splicing factor, PUF, that is important, together with U2AF, for efficient reconstitution of a poly U-depleted nuclear extract. The activity has been purified to near homogeneity and consists of two proteins, PUF60 and the previously described splicing factor p54 (Zhang & Wu, 1996).

RESULTS

Depletion of nuclear extract for a novel pyrimidine-tract splicing factor

Splicing extracts can be depleted of the pyrimidine-tract binding factor U2 snRNP Auxiliary Factor (U2AF) by passing the extract over a poly U Sepharose column at high salt concentrations. Extracts depleted in this way (NEΔU2AF) are unable to catalyze the splicing reaction and are blocked at an early step in spliceosome assembly. Spliceosome assembly and splicing activity can be restored to these extracts through the addition of purified U2AF or recombinant U2AF65 (Zamore & Green, 1991). Modification of this depletion protocol (dialysis of the extract against 1.0 M KCl before absorption to the matrix) leads to the codepletion of a second pyrimidine-tract associated splicing activity. Nuclear extracts depleted using this method (NEΔU) were inactive for splicing in vitro (Fig. 1A, lane 2). Addition of saturating amounts of recombinant U2AF65 resulted in a partial restoration of splicing activity (Fig. 1A, compare lanes 1 and 4). Efficient restoration of splicing activity required the addition of a second activity (Fig. 1A, lane 5), which eluted from the poly U depletion column with 2.0 M KCl. This activity, poly U factor (PUF), was unable to restore splicing activity to NEΔU in the absence of U2AF65 (Fig. 1A, lane 3).

PUF activity is important for splicing of multiple introns. Two other pre-mRNAs in addition to the PIP85a intron were tested for PUF-dependent splicing activity. Both the PIPβG and AD10 (data not shown) chimeric substrates required both PUF activity and U2AF65 for efficient reconstitution of splicing activity (Fig. 1B, compare lanes 4 and 5).

U2AF consists of two subunits, U2AF65 and U2AF35. To determine if PUF compensated for the absence of U2AF35 in the U2AF65-reconstituted system, splicing was tested with U2AF purified from nuclear extract, containing both U2AF35 and U2AF65 (data not shown). For both the PIPβG and the PIP85a pre-mRNAs, addition of saturating levels of U2AF did not stimulate splicing more than U2AF65 either in the presence (Fig. 1B, compare lanes 5 to 7 and 12 to 14) or the absence (Fig. 1B, compare lanes 4 to 6 and 11 to 13) of the PUF activity. Further, addition of PUF to reactions complemented by U2AF resulted in a stimulation of splicing for both substrates. These results suggest that PUF activity does not simply substitute for U2AF35.

Identification of PUF60 and p54 as components of PUF activity

PUF activity was purified to near homogeneity using the reconstituted splicing reaction as an assay. Briefly, nuclear extract dialyzed against 1.0 M KCl was applied to a poly U Sepharose column and PUF activity eluted from the column at 2.0 M KCl, whereas U2AF activity eluted at 3.0 M KCl. The high salt eluate of the poly U Sepharose column was further fractionated on phosphocellulose, poly U Sepharose, and mono S Sepharose (see Materials and Methods). Chromatography of PUF activity on mono S Sepharose is shown in Figure 2A. The top panel shows the silver-stained gel of eluted fractions, whereas the bottom panel shows the activity of each eluted fraction. Polypeptides in the peak fraction (fraction 8) migrated with apparent molecular weights of 130 kDa, 60 kDa, and 48 kDa.

Application of the purified fraction 8 to a gel filtration column allowed further purification of the splicing activity and estimation of its size. The peak of activity eluted from the column with an estimated molecular weight of 400 kDa (Fig. 2B, bottom panel, lanes 5 and 6). The proteins present in the adjacent fractions included the 130-kDa band and the 60-kDa band (Fig. 2B, top panel, lanes 6, 7, and 8). Notably the 48-kDa protein did not copurify with PUF activity; the 48-kDa protein was found in fractions 20–22.

Polypeptides from the 130-kDa and 60-kDa bands were eluted, digested with trypsin, and sequenced. Three peptides were sequenced from the 60-kDa band

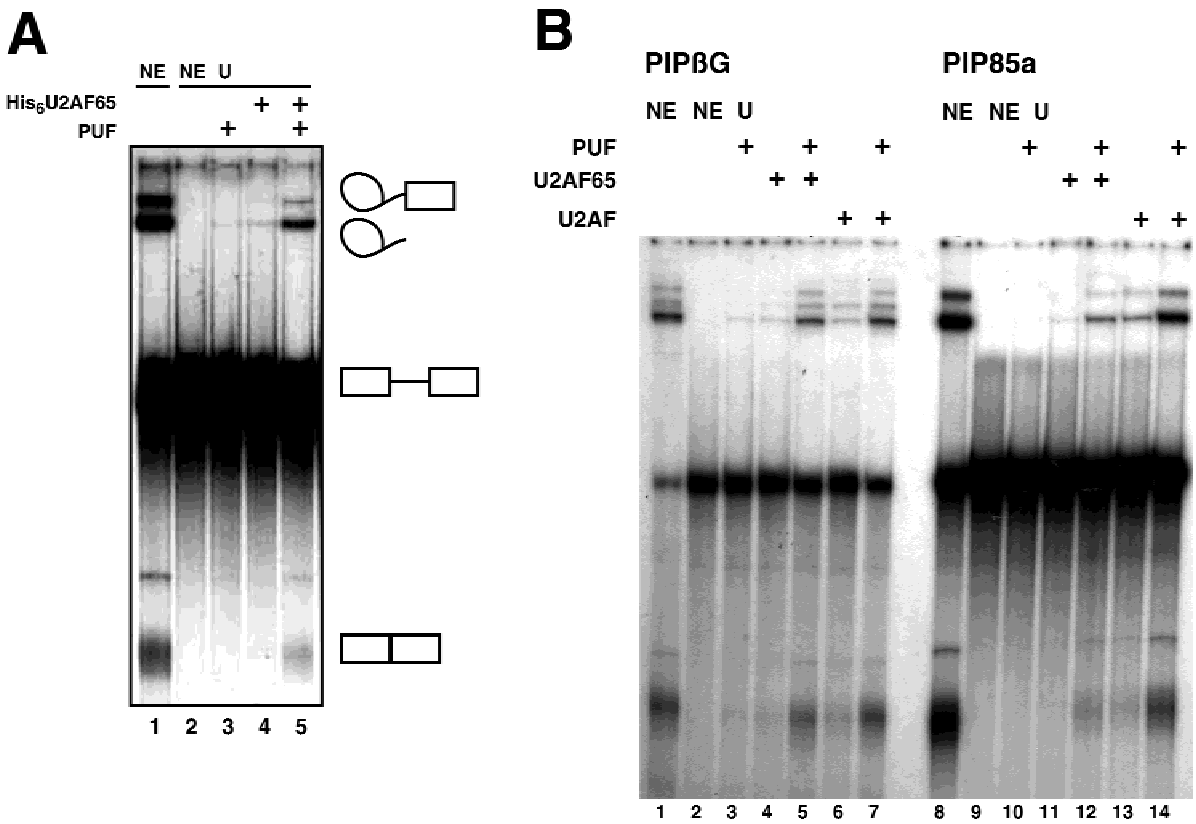

FIGURE 1. Poly U-depleted nuclear extract requires both U2AF and PUF. **A**: Splicing in vitro of the PIP85a pre-mRNA in poly U-depleted extract. Nuclear extract treated with high salt, but not depleted (NE) was compared to nuclear extract that had been depleted by passage over a poly U Sepharose column at high salt (NEΔU, lanes 1 and 2). Addition of recombinant U2AF65 to the depleted extract did not efficiently restore splicing activity to the depleted extract (lane 4). Addition of the 2.0-M KCl eluate fraction of the poly U depletion column (PUF) alone did not restore activity to NEΔU (lane 3), but addition of both recombinant U2AF65 and PUF restored splicing activity to the extract. The positions of the pre-mRNA, mRNA and lariat product and intermediate are indicated schematically to the right of the figure. **B**: PUF activity requirement for efficient splicing of pre-mRNAs and PUF activity requirement with either U2AF or U2AF65. U2AF purified from nuclear extract (the 3-M KCl eluate fraction of the poly U depletion column) was compared to U2AF65 in the reconstituted splicing system. Splicing in vitro of PIPβG (lanes 1–7) and PIP85a (lanes 8–14) with or without added purified PUF activity was compared in the presence of U2AF (lanes 6 and 7, 13 and 14) or recombinant U2AF65 (lanes 4 and 5, 11 and 12).

and the sequences of all three peptides matched that of the previously described splicing factor p54 (Chaudhary et al., 1991; Zhang & Wu, 1996). Thirteen peptides sequenced from the 130-kDa band matched a cDNA found in the EST database (Hillier et al., 1996). The longest EST identified, 1,989 bp, was sequenced. This human cDNA encoded all thirteen peptides identified from the 130-kDa band in a 559 amino acid open reading frame (ORF). The 559 amino acid open reading frame is predicted to encode a 60-kDa protein; the protein is referred to as PUF60 (poly U binding Factor–60 kDa). Poly A mRNAs from different human tissues were analyzed by Northern analysis with a PUF60 cDNA probe. The PUF60 transcript is ubiquitously expressed as an abundant mRNA of 2.0 kb (data not shown), consistent with the size of the sequenced cDNA. As will be discussed below, PUF60 forms SDS-resistant dimers with an apparent molecular weight of 130 kDa, explaining the aberrant mobility of this protein on SDS-polyacrylamide gels.

A search of the EST database identified 29 mouse cDNA sequences, allowing reconstruction of the murine ORF. Only 12 amino acids were found to differ between the human and reconstructed murine PUF60 homologs. These differences consisted of an insertion of five alanines in the N-terminal poly-alanine repeat, and the remaining seven differences were conservative substitutions (data not shown). A PUF60 homolog was identified in the *Drosophila* EST database and was sequenced. The *Drosophila* polypeptide (DPUF68) of 570 amino acids was 45.3% similar to PUF60 over its entire length (Fig. 3).

Domain structure of PUF60

Comparison of the hypothetical translation products of the human, mouse, and *Drosophila* homologs identified four domains in PUF60 (Fig. 3). The N-terminal domain is not well conserved between mammalian

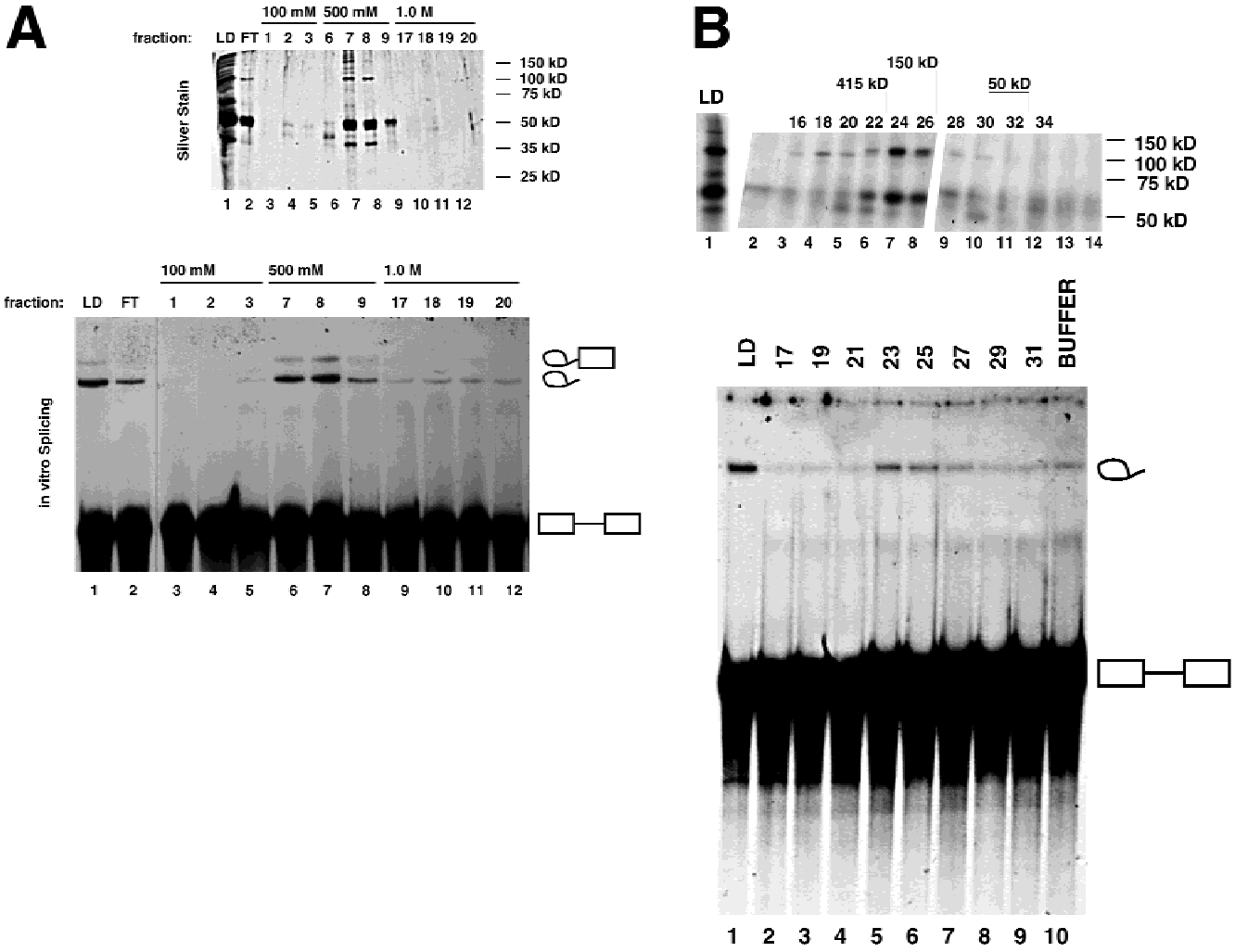

FIGURE 2. Purification of the PUF activity. **A**: Chromatography of the PUF activity on mono S Sepharose. Top panel: silver-stained protein gel of the fractionation. Bottom panel: reconstitution of in vitro splicing activity in reactions containing NEΔU supplemented with recombinant U2AF65, PIP85a substrate. Load on the column (LD) and flowthrough (FT) are shown in lanes 1 and 2. Equivalent quantities of each fraction were loaded in each lane. **B**: Purification of the mono S Sepharose purified material on Superose 6 gel filtration chromatography. Top panel: silver-stained gel of the even numbered column fractions, flanking peak fraction. Mobilities of gel filtration standards (BioRad; 415 kDa: fraction 23; 150 kDa: fraction 27; 50 kDa: fraction 33) are indicated above the gel. Bottom panel: reconstitution of in vitro splicing assay with odd numbered fractions. LD indicates load fraction. Equivalent quantities of each fraction were loaded in each lane.

PUF60 and DPUF68. Notably, DPUF68 contains four SR dipeptides in this region that are not present in the mammalian proteins. The conserved region begins with a segment of approximately 40 amino acids immediately upstream of the first RRM domain. The central domain consists of two RRM domains. Both RRM domains have good matches to the RNP1 and RNP2 consensus motifs; these domains are similar to the RRM domains of both U2AF65 and Mud2p. The C-terminal domain is preceded by a variable domain that, in *Drosophila*, contains a polyalanine motif similar to the polyalanine motif found near the N-terminus of the mammalian PUF60 homolog. The C-terminal domain has some similarity to an RRM-domain consensus, however it is unusual in having a poor match to the N-terminal RNP2 consensus motif, but a good match to the central RNP1 consensus motif. This C-terminal RRM-like domain is the most conserved part of the protein when U2AF65, Mud2p, and PUF60 are compared (Fig. 3B). Because this domain has both sequence features and biochemical activities that differ from the general class of RRM domains, it is referred to as the PUMP (PUF60, U2AF65, Mud2p protein–protein interaction) domain (see below).

PUF60 and p54 are depleted in NEΔU

Rabbit polyclonal antibodies raised to PUF60 were used to demonstrate that NEΔU is depleted of PUF60 as shown in Figure 4A. Equal volumes of different batches

A

```
PUF60     MAT---ATIALQVNGQQGGGSEPAAAAAVVAAG---DKWKP--POGTDSIK----MENGQ   47
          * :   *: : : : *:     **:  :   :*   *   *   *    :*    :: *
DPUF68    MGSNDRASRSPRSDDQREISDMPATKRTRSDSGKSTDSKIPYLSQPLYDLKQTGDVKFGP   60

PUF60     ST-AAKLG-----LPPLTPEQQEALQKAKKYAMEQSIKSVLVKQTIAHQQQQLTNLQMAA  102
           * :* **      ** *: **:: : ************ **:***:*******:
DPUF68    GTRSALLGLLGGALPKLSSEQHDLVSKAKKYAMEQSIKMVLMKQTLAHQQQQLA------  114

PUF60     VTMGFGDPLSPLQSMAAQRQRALAIMCRVYVGSIYYELGEDTIRQAFAPFGPIKSIDMSW  162
                    *    ***:***:********* :** ***** **:********:***
DPUF68    -----------TQRTQVQRQQALALMCRVYVGSISFELKEDTIRVAFTPFGPIKSINMSW  163

PUF60     DSVTMKHKGFAFVEYEVPEAAQLALEQMNSVMLGGRNIKVGRPSNIGQAQPIIDQLAEEA  222
          * :* ***********:** *********  ::************: *** :**:: ***
DPUF68    DPITQKHKGFAFVEYEIPEGAQLALEQMNGALMGGRNIKVGRPSNMPQAQQVIDEVQEEA  223

PUF60     RAFNRIYVASVHQDLSDDDIKSVFEAFGKIKSCTLARDPTTGKHKGYGFIEYEKAQSSQD  282
          ::*********:* ***::********** *  * **:  :   ********* : *: ::
DPUF68    KSFNRIYVASIHPDLSEEDIKSVFEAFGPILYCKLAQGTSLHTHKGYGFIEYANKQAMDE  283

PUF60     AVSSMNLFDLGGQYLRVGKAVTPPMPLLTPATPGGLPPAAAVAAAAATAKITAQEAVAGA  342
          *::********** ****:::***  *  *:*   :* ************* * :***
DPUF68    AIASMNLFDLGGQLLRVGRSITPPNALACPTTNSTMPTAAAVAAAAATAKIQALDAVASN  343

PUF60     AVLG-TLGTP-----------G---LVS---------PALTLAQPLGTLPQAVMAAQAP-  377
          **** :  **           *   :**         ***: * *   **  :: * :*
DPUF68    AVLGLSQNTPVMAAGAVVTKVGAMPVVSAATSAAALHPALAQAAP-ALLPPGIFQAPTPV  402

PUF60     -----GVITG------VTPARPP--------IPVTIPSVGV-VNPILAS----PPTLG--  411
           ** :*      * *: **         :*:*:  *** : * :*:      : *
DPUF68    APSLLGVPAGLQXLQAVVPTLPPPALLATPTLPMTVGGVGVGLVPTVATLAGAEASKGAA  462

PUF60     ------------------LLE--PKK-EKEEEELFPESERP---EMLSEQEHMSISGSSA  447
                          * *   *  **::***  :       : * :**:*** * **
DPUF68    AAAALSAAANNAAVTAANLSENIKKAHEKQQEELQKKLMDEGDVQTLQQQENMSIKGQSA  522

PUF60     RHMVMQKLLRKQESTVMVLRNMVDPKDIDDDLEGEVTEECGKFGAVNRVIIYQEKQGEEE  507
          *::***:*:*  :* *::***** *:*:*: *: *: *** ***:* ****::*** *:*
DPUF68    RQLVMQRLMRPVDSRVIILRNMVGPEDVDETLQEEIQEECSKFGTVSRVIIFNEKQTENE  582

PUF60     D---AEIIVKIFVEFSIASETHKAIQALNGRWFAGRKVVAEVYDQERFDNSDLSA       559
          *   ************  :*: :  :**:**:* **:****:***  **: ***
DPUF68    DDDEAEIIVKIFVEFSAGAEAMRGKEALDGRFFGGRRVVAELYDQGIFDQGDLSG       637
```

B

```
                  RNP 2                                                RNP 1
PUF60             VLRNMVDPKDIDD-----DLEGEVTEECG----KFGAVNRVIIYQEKQGEEED----AEIIVKIFVEFSIASETHKAIQALNGRWFAGRKVVAEVYDQERFDNSDLSA
DPUF68            ILRNMVGPEDVDE-----TLQEEIQEECS----KFGTVSRVIIFNEKQTENEDDDE-AEIIVKIFVEFSAGAEAMRGKEALDGRFFGGRRVVAELYDQGIFDQGDLSG

U2AF65            CLMNMVLPEELLDDEEYEEIVEDVRDECS----KYGLVKSIEIPRPVDG------VEVPGCGKIFVEFTSVFDCQKAMQGLTGRKFANRVVVTKYCDPDSYHRRDFW-
DU2AF50           CLLNMVTPDELRDEEEYEDILEDIKEDCT----KYGVVRSVEIPRPIEG------VEVPGCGKVFVEFNSVLDCQKAQQALTGRKFSDRVVVTSYFDPDKYIRREF--

Mud2p             LLLNCLDPLDLKDE----TFITEIKETLKY---SIAGADTIKICQPGVDYRLNFENLASGAGNIYIKFKTLEAAKHAMEELPGTQFNDRIVLCTYIDEDDFDMMEATQ

U2AF35            GLRCAVSDVEMQEH  YDEFFEEVFTEMEE   KYGEVEEMNVCDNLGD         HLVGNVYVKFRREEDAEKAVIDLNNRWFNGQPIHAELSPVTDFREACCRQ
DU2AF38           DLVANVSDEEMQEH--YDNFFEDVFVECED---KYGEIEEMNVCDNLGD---------HLVGNVYIKFRNEADAEKAANDLNNRWFGGRPVYSELSPVTDFREACCRQ
Urp               DYDPDASLEYSEEE--TYQQFLDFYEDVLPEFKNVGKVIQFKVSCNLEP---------HLRGNVYVQYQSEEECQAALSLFNGRWMAGRQLQCEFCPVTRWKMAICGL
RRM consensus     h hG h             h  h          G hh h h               G G h       h  hh  h G  h G  h h  h
```

FIGURE 3. Sequence of PUF60 and DPUF68. **A**: Alignment of PUF60 and DPUF68 with Clustal V; identical matches are indicated by * and similarities by :; underlined amino acids in the PUF60 sequence are peptides identified by amino acid sequencing of the 130-kDa protein in the purified PUF fraction. (GenBank accession numbers: human AF190744 and fly AF190745). **B**: Alignment of the PUMP domains of the U2AF65 family of proteins (PUF60, DPUF68, U2AF65, DU2AF50, and Mud2p) and of the U2AF35 family (U2AF35, DU2AF38, and Urp). Residues that are identical in seven of eight proteins are boxed and highlighted in grey, those that are similar in all proteins are highlighted in dark grey, and those that are similar in seven of eight proteins are highlighted in light grey. Alignment of the RRM consensus with the PUMP homology is shown in the bottom line (h indicates a conserved hydrophobic residue, G is a conserved glycine, + is a conserved positively charged amino acid; Birney et al. 1993). Position of the conserved motifs found in RRM domain, RNP1 and RNP2, is shown above the PUMP alignment. The homology of this region to the protein–protein interaction domain of the P-CIP2 has previously been noted (Alam et al., 1996). Similar amino acids are: A,I,L,M,V-C; F,Y,W; D,E,H,R,K,N,Q; S,T,A; and G,P.

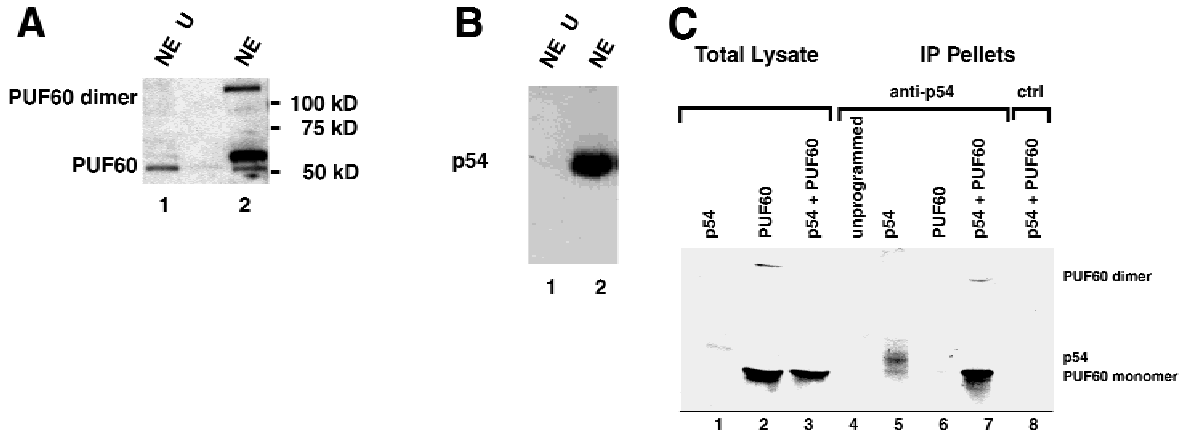

FIGURE 4. NEΔU is depleted of both PUF60 and p54. **A**: PUF60 immunoblotting analysis of NEΔU (lane 1) and the control NE (lane 2). The 55-kDa band present in the extracts is a cross-reacting band and is detected by preimmune serum. **B**: Immunoblotting analysis of NEΔU (lane 1) and NE (lane 2) detected with an antibody directed against a p54 peptide (Chaudhary et al., 1991). **C**: p54 forms complexes with PUF60 as shown by coimmunoprecipitation assays of in vitro translation reactions containing either p54, or PUF60, or both p54 and PUF60. Ten percent of each in vitro translation reaction is shown in lanes 1–3, p54 antibody precipitated pellets are shown in lanes 5–7, and p54 antibody precipitated pellets of an unprogrammed lysate is shown in lane 8. Control precipitates with no primary antibody are shown in lane 8.

of NEΔU (lanes 2, 4, and 5) and control extract, NE (lanes 3 and 7) were compared by immunoblotting. NEΔU was estimated to be depleted by at least 90% of PUF60 by this method. Antibodies raised against the splicing factor p54 demonstrate that p54 was also depleted from NEΔU to a similar extent (Fig. 4B, lanes 1 and 2).

The codepletion of PUF60 and p54 indicated that they were present as a complex. To determine if PUF60 and p54 could form a complex in vitro, cotranslation of PUF60 and p54 was performed. Antibody to p54 (Chaudhary et al., 1991) immunoprecipitated p54 (Fig. 4C, lane 6). Cotranslation reactions of p54 and PUF60 that were immunoprecipitated with p54 antibody resulted in coprecipitation of PUF60, whereas control immunoprecipitations with no p54 present did not result in precipitation of PUF60. This coprecipitation depended upon cotranslation of the two proteins, because mixture of the two reactions after completion of translation did not yield precipitation of PUF60. This coimmunoprecipitation strongly indicates that p54 and PUF60 form a complex in vitro.

PUF60 and formation of spliceosomal complexes

U2AF-depleted extract, NEΔU2AF, is blocked prior to the first step in splicing and does not form spliceosomal complexes (Zamore & Green, 1991). Both NEΔU and NEΔU reconstituted with U2AF were inefficient for the first step of splicing (Fig. 1A,B), indicating that PUF activity, like U2AF, acts early in spliceosome complex assembly. To determine if spliceosomal complex assembly was also blocked, formation of spliceosomal complexes was tested with a 3′ half pre-mRNA consisting of a branch sequence, pyrimidine tract, 3′ splice site, and 3′ exon. A time course of A3′ complex assembly is shown in Figure 5. While PUF activity was not required for formation of A3′ complex (Fig. 5, lanes 5 and 6 have about 30% of the A3′ found in lanes 11 and 12), it was required for efficient formation of this complex. This suggests that PUF activity facilitates formation of complexes containing U2 snRNP, but is not absolutely required for formation of this complex. This is in contrast to the stringent requirement of both PUF and U2AF activity for the formation of A_{min}, the minimal U2snRNP/branch sequence complex (Query et al., 1997).

The PUMP domain is a protein–protein interaction domain

Because PUF60 is encoded by an ORF of 559 amino acids, it would be expected to migrate in an SDS-polyacrylamide gel at 60 kDa. Instead, PUF60 was identified in the purified fraction as a 130-kDa product. The paradox of how an expected translation product of 60-kDa could result in a 130-kDa protein was resolved when PUF60 was produced in vitro by a translation reaction. The predominant in vitro translation product migrated as a 60-kDa polypeptide, but a small amount migrated as a 130-kDa polypeptide. This suggested that the 60-kDa primary translation product could form SDS-resistant dimers with the apparent mobility of a 130-kDa polypeptide. To test whether the 130-kDa form was a dimer of PUF60, full-length PUF60 and N-terminally truncated PUF60 (deleted from amino acids 1–76, PUF60ΔN) were cotranslated (Fig. 6A). Translation in

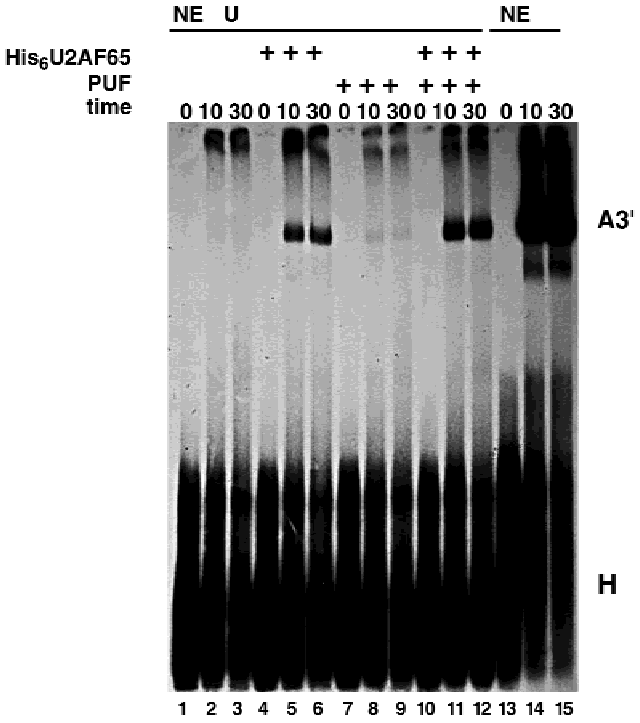

FIGURE 5. PUF activity is required for efficient A3′ spliceosome assembly. A time course (0, 10, and 30 min) of complex assembly in NEΔU alone (lanes 1–3), in the presence of PUF activity (lanes 4–6), in the presence of recombinant U2AF65 (lanes 7–9), or in the presence of both activities (lanes 10–12). The complexes formed in NE (lanes 13–15) are shown for comparison.

vitro of the full length PUF60 produced a predominant 60-kDa product and a small amount of a 130-kDa product. Translation in vitro of PUF60ΔN produced a predominant band of approximately 55 kDa and a small amount of a 120-kDa species. Cotranslation of both full-length and PUF60ΔN produced the expected 60-kDa and 55-kDa bands as well as bands of 130, 125, and 120 kDa. These latter bands were present in approximately the ratio of 1:2:1, consistent with the hypothesis that the 130-kDa and 125-kDa species were homodimers of the input forms of PUF60, whereas the 125-kDa species was a heterodimer of these two forms. Other experiments indicated that dimerization of PUF60 was stable in high concentrations of the reducing reagents dithiothreitol (DTT) (100 mM) or 2-mercaptoethanol (280 mM). PUF60 dimerization was also stable to the addition of 4-vinylpyridine, a thiol-alkylating reagent, when added in excess after reduction with 2-mercaptoethanol (data not shown). These results indicate that dimerization was not dependent on disulfide bond formation. Consistent with the hypothesis that PUF60 forms SDS-resistant dimers, Western blots of PUF60 protein in nuclear extracts detected bands migrating as concentration-dependent monomers (60 kDa) and dimers (130 kDa) (data not shown).

Bacterially expressed PUF60 also formed SDS-resistant dimers (Fig. 6B). Further, this dimerization was found to be concentration dependent; high concentrations of the protein formed dimers more efficiently than low concentrations. When the concentration of PUF60 protein, which was synthesized in bacteria as a His_6 fusion protein (His_6 PUF60), was increased by twofold increments from 60 nM to 2.0 μM, the fraction of protein migrating as a dimer increased dramatically. It was estimated that an approximate equal fraction of monomeric and dimeric forms of PUF60 formed at 1.0 μM, suggesting that the dissociation constant for the SDS-resistant form of PUF60 is 1 μM.

Formation of SDS-resistant dimers is a property of the PUMP domain alone. Different portions of PUF60 were expressed both in translation reactions in vitro and in bacterial expression systems. These proteins were individually tested for formation of SDS-resistant dimerization at high concentrations. For example, deletion of 44 amino acids from the C-terminus prevented dimerization of PUF60 both in translation in vitro reactions (data not shown) and in bacterially expressed protein (Fig. 6C, lanes 3 and 4). This deletion, His_6PUF60ΔC (deleted from amino acid 516), did not remove any cysteines from the protein, but did remove the RNP1 consensus motif of the PUMP domain. Further deletion and expression experiments showed that only proteins with an intact PUMP domain could form SDS-resistant dimers (Fig. 6D). Finally, SDS-resistant dimerization activity could be conferred on the GST protein by addition of a carboxyl-terminal 94 amino acid fusion of the PUMP domain (data not shown).

RNA-binding activity of PUF60

The PUMP domain is homologous to the RRM-domain family, which generally has RNA binding activity. To determine if the PUMP domain contributes significantly to RNA-binding activity, the affinities of bacterially expressed PUF60 and PUF60ΔC for a pyrimidine-tract RNA were determined. A series of different length pyrimidine-tract RNAs was synthesized using sequences from the PIP85a substrate, and RNA binding was evaluated using an electrophoretic mobility shift assay. Pyrimidine-tract RNAs of 14 (pyr14) and 23 (pyr23) nucleotides were bound to approximately the same extent by PUF60. In contrast, RNAs of 11 (pyr11) and 7 (pyr7) nt were not bound by PUF60. This suggested that optimal binding required sequences longer than 11 nt. The specificity of PUF60 binding to pyr23 was examined by competition with RNA homo- and heteropolymers (Fig. 7A). Pyr23 binding was competed only with RNA polymers containing uridine, that is, with poly[U], poly[C, U], and poly[G, U]. This binding could not be competed with homopolymers of poly[C], poly[G], poly[A], or poly[I].

When the binding activity of His_6 PUF60 was compared to that of His_6 PUF60ΔC for the substrate pyr23, the two proteins had indistinguishable affinities

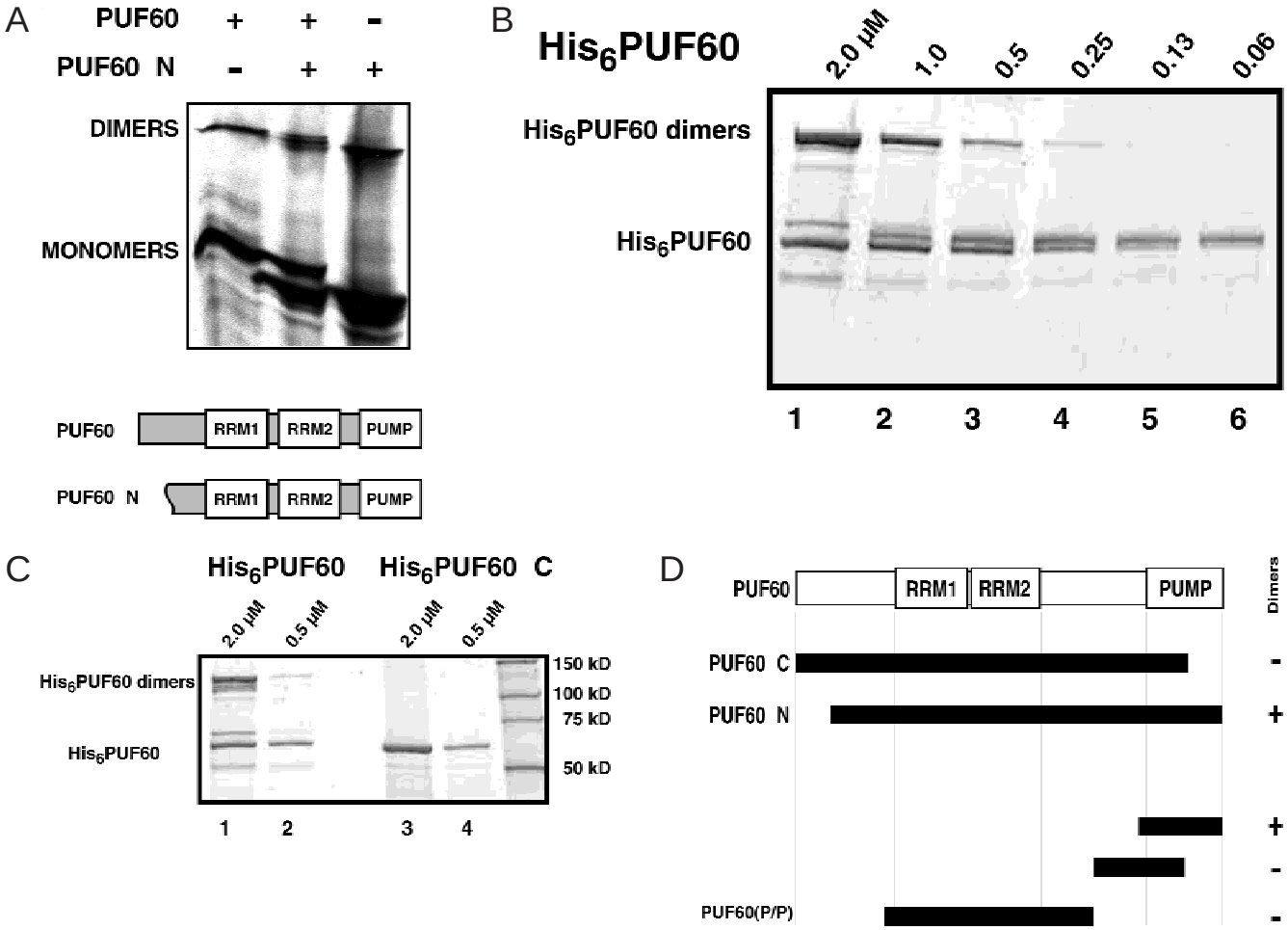

FIGURE 6. PUF60 forms SDS-resistant dimers. **A**: Translation in vitro of the full-length PUF60 yields a predominant band of 60 kDa and a band of 130 kDa (lane 1); an N-terminal truncation of HPUF60 yields a 55-kDa band and a 120-kDa band (lane 3); cotranslation of both yields the 60- and 55-kDa products as well as products of 120, 125, and 130 kDa (lane 2). **B**: Bacterially expressed His_6PUF60 also forms SDS-resistant dimers. Serial twofold dilutions of the protein (lanes 1–6) showed that monomeric His_6PUF60 could be shifted into dimeric form at high concentrations. **C**: Deletion of the carboxyl-terminal domain of PUF60 prevents SDS-resistant dimerization. Equivalent quantities (2.0 μM and 0.5 μM) of His, PUF60 (lanes 1 and 2) and His_6PUF60ΔC (lanes 3 and 4) boiled in SDS and resolved by SDS-polyacrylamide gel electrophoresis show that the carboxyl-terminal domain is required for formation of SDS-resistant dimers. **D**: The domain responsible for SDS-resistant dimerization was mapped by expression of PUF60 fragments in *Escherichia coli*. The expressed regions are shown as black bars beneath the PUF60 domain diagram and dimerization activity is indicated to the right.

(Fig. 7B; $K_{d\ obs}$ = 138 nM and 122 nM, respectively). The Hill coefficients, which indicate the degree of cooperativity during binding, were evaluated (Creighton, 1984). The coefficients were found to be 2.8 and 1.8, respectively. This difference in Hill coefficients suggests that the PUMP domain of PUF60 may mediate a slight cooperative pyrimidine tract binding activity. Under these same conditions, a His6 fusion of U2AF65 had a $K_{d\ obs}$ of 300 nM and a Hill coefficient of 1.1 (data not shown). Zamore et al. (1992) reported values of GSTU2AF65 binding ranging from 10 nM to 2 μM for pyrimidine tracts. In contrast to the slight effect on RNA binding of deletion of the PUMP domain from PUF60, deletion of this domain from U2AF65 resulted in a dramatic decrease in RNA binding affinity (Zamore et al., 1992).

DISCUSSION

We have identified a new pyrimidine-tract binding factor, PUF, that is required, together with U2AF, for the efficient reconstitution of polyU-depleted nuclear extracts. The activity consists of two proteins, PUF60 and the previously described splicing factor p54 (Zhang & Wu, 1996). PUF60 is a novel protein with striking similarity to U2AF65. PUF60 directly binds p54, which is an SR protein, and both proteins contain RRM domains indicative of RNA-binding activity. The PUF60/p54 complex chromatographs as a 400-kDa complex and thus may have an oligomeric structure; however, it is also possible that this complex contains other, yet to be identified, components.

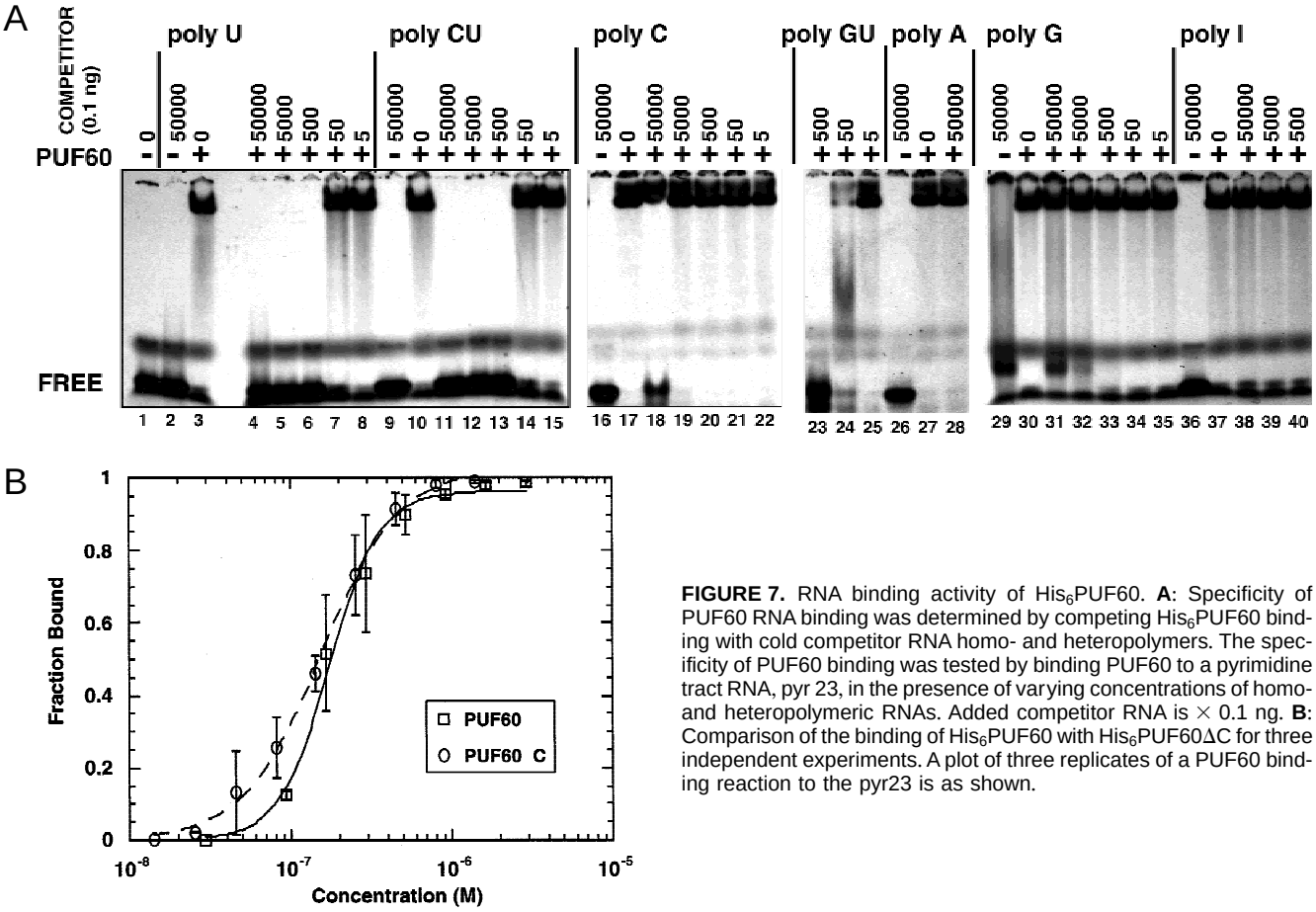

FIGURE 7. RNA binding activity of His_6PUF60. **A**: Specificity of PUF60 RNA binding was determined by competing His_6PUF60 binding with cold competitor RNA homo- and heteropolymers. The specificity of PUF60 binding was tested by binding PUF60 to a pyrimidine tract RNA, pyr 23, in the presence of varying concentrations of homo- and heteropolymeric RNAs. Added competitor RNA is × 0.1 ng. **B**: Comparison of the binding of His_6PUF60 with His_6PUF60ΔC for three independent experiments. A plot of three replicates of a PUF60 binding reaction to the pyr23 is as shown.

Two pyrimidine-tract RNA binding activities important for splicing

The PUF activity enhanced the efficiency of splicing of several introns and is probably best considered a general splicing factor. In contrast to the stimulatory nature of the PUF activity, U2AF was found to be an essential factor for splicing in these reconstituted reactions. Both factors contribute to the early steps in formation of a spliceosome. PUF activity was required for the efficient assembly of A3′ complex, the U2AF and ATP-dependent U2 snRNP-containing complex that forms on branch-sequence, pyrimidine-tract, 3′ exon RNAs. We have previously shown that PUF activity was more stringently required for the formation of the A_{min} complex (Query et al., 1997). This complex, a U2 snRNP-branch sequence complex that forms in the absence of ATP on a short oligonucleotide containing the branch sequence and pyrimidine tract, is more sensitive to alterations in the sequence and functional groups of the branch sequence than A or A3′ complexes. This enhanced sensitivity for PUF in formation of A_{min} complex may be due to the absence of binding sites for the SF3a factors on the pre-mRNA 5′ to the branch sequence (Chiara et al., 1994; Gozani et al., 1996). The activity of PUF in formation of the A_{min} complex is most consistent with it working in conjunction with U2AF in early steps in splicing.

PUF activity may augment or complement the activity of U2AF in the association of the U2 snRNP complex with the substrate. As it was not possible to reconstitute PUF activity with either bacterially expressed PUF60 or p54 protein nor with both in combination, the activity probably resides in some state of the heteromer. In that case, either protein could bind the substrate RNA and interact with either a protein that binds U2AF65 or some component of the U2 snRNP complex. For example, p54 contains SR subdomains that facilitate binding of the 17S U2 snRNP complex to the substrate (Kennedy et al., 1998). The p54 polypeptide has also been shown to interact directly with U2AF65 (Zhang & Wu, 1996). The specific role of U2AF65 in formation of the complex of U2 snRNP and substrate RNA is not completely clear. It is known that U2AF65 interacts through its carboxyl terminal domain with SF1/BBP, which recognizes sequences at the branch site (Berglund et al., 1997; Rain et al., 1998). PUF activity is not dependent upon interaction with the small subunit of U2AF, U2AF35, nor does it replace the activity of U2AF35 in splicing.

PUF60 has been identified by two other laboratories: as a protein, SIAH-BP1, that interacts in yeast two-hybrid assays with the mouse homolog SIAH of the *Drosophila* protein seven-in-absentia (*sina*; D. Bowtell, pers. comm.) and as RoBP1, which is part of an RNP complex containing the human auto-immune antigen Ro and the hY5 RNA (G. Boire and P. Bouffard, pers. comm.). The function of the Ro auto-antigen is not known, but Ro is found in both nuclear and cytoplasmic compartments (Peek et al., 1993). Nuclear Ro is associated with nuclear speckle domains (Wahren et al., 1996). The potential association of PUF60 with SIAH is also interesting. *sina*, the SIAH homolog in *D. melanogaster*, is known to target the product of the *tramtrack* (*ttk*) gene for degradation (Li et al., 1997; Dickson, 1998). *ttk* is a transcription factor that prevents cell-fate determination and its degradation is essential for generation of the retinal R7 cell. Strikingly, *ttk* protein exists as two alternatively spliced forms, one of which is degraded through the *sina* pathway (*ttk*88β; Li et al., 1997). It is interesting to speculate that DPUF68, the *Drosophila* homolog of PUF60, might interact with *sina* to regulate the *ttk* alternative splicing pathway as part of a regulatory loop.

The homology among PUF60 and U2AF65 extends across their entire length with the exception of the very N-terminal domain. Both proteins have a relatively nonconserved N-terminal domain, two central RRM domains, and a C-terminal RRM-like PUMP domain. Phylogenetic analysis shows that PUF60 and U2AF65 are approximately equally related to the yeast protein Mud2p. Mud2p is the only homolog of the U2AF65/PUF60 family in *S. cerevisiae* and, while not required for viability, is important in recognition of the 3′ splice site. Mud2p interacts genetically with U1 snRNA and is present in the commitment complex in vitro. Both Mud2p and U2AF65 interact with the branch site interacting protein SF1/BBP (Berglund et al., 1998). In fact, the PUMP domain of U2AF65 has recently been shown to interact with SF1/BBP1 (Berglund et al., 1998; Rain et al., 1998). Mud2p also interacts with the splicing factor Prp11, which adds to the spliceosome coincidentally with U2 snRNP (Abovich et al., 1994). These sequence comparisons strongly suggest that U2AF65, PUF60, and Mud2p are members of an extended family of factors that may have similar roles in the early steps of splicing.

The PUMP domain is a subset of the RRM domain family

The highest degree of conservation between PUF60, U2AF65, and Mud2p is found in the C-terminal PUMP domain. This domain was previously identified as an RRM-like domain (Birney et al., 1993) but was also recognized as being divergent from a consensus RRM domain. A consensus sequence formed of the PUF60, U2AF65, and Mud2p C-terminal domains was used to identify other possible homologs. Several such domains were identified; one of the most interesting is found in the small subunit of U2AF, U2AF35, and its homolog U2AFbp1/Urp. This domain has not previously been identified as having RRM homology, but showed significant homology to the PUMP domain.

The PUMP domains of U2AF35 and Urp have been implicated in binding U2AF65 (Zhang et al., 1992; Wentz & Potashkin, 1996; Tronchere et al., 1997). This domain has previously been described as the H2 homologous region of Urp and U2AF35 (Tronchere et al., 1997). The PUMP domain-containing region of U2AF35 recognizes a small region of U2AF65, N-terminal to the first RRM (Zhang et al., 1992). Thus, it is possible that PUMP domains are primarily important for protein–protein interactions and recognize short tracts of amino acid sequence.

The PUMP domain is responsible for the unusual dimerization of PUF60. Dimerization that is stable to the addition of SDS is unusual but not without precedent. For example, bacteriophage P22 tailspike Endorhamnosidase is resistant to SDS denaturation (Goldenberg et al., 1982). Like the RRM domain, and presumably the PUMP domain, the tailspike protein consists predominantly of β strands (Steinbacher et al., 1994). β-strand containing proteins can be remarkably resistant to denaturants. Extreme examples of denaturant-resistant proteins are the amyloid plaque and Prion proteins, which are insoluble under most conditions and these proteins are composed predominantly of β sheets (Meyer et al., 1986; Pan et al., 1993).

MATERIALS AND METHODS

Preparation of poly U-depleted nuclear extract

HeLa cell nuclear extract was prepared using standard protocols (Dignam et al., 1983). Briefly, 15–30 L HeLa cells grown in spinner culture were harvested by centrifugation, washed in ice cold PBS, and resuspended in five packed cell volumes of ice cold Buffer A (10 mM HEPES [from 1.0 M stock, pH 7.9, with KOH at room temperature], 1.5 mM $MgCl_2$, 10 mM KCl, 5 mM DTT), proteinase inhibitors were included in all buffers and included 50 μg/mL of PMSF and Leupeptin, Pepstatin A, and Aprotinin (Boehringer Mannheim, used at the concentration recommended by the manufacturer), and incubated for 10 min on ice. Cells were then spun down, resuspended in one swollen cell pellet volume of Buffer A, and dounced 10 strokes with a Hamilton B dounce to lyse the cells. Lysed cells were spun for 10 min at 1,200 $\times$ *g* and the supernatant was decanted off the loosely packed nuclei and discarded. Nuclei were packed at 25,000 $\times$ *g* for 20 min and the supernatant was discarded. Nuclei were resuspended in 1.25 vol Buffer C (20 mM HEPES, pH 7.9, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 5 mM DTT, 0.5 mM PMSF) per volume of packed cells with a Hamilton A dounce and incubated,

tumbling, at 4 °C for 45 min. Extracted nuclei were pelleted at 20,000 × *g* for 10 min and the pellet was discarded. Nuclear extract supernatant was twice dialyzed against 1.0 M KCl HENG10 (20 mM HEPES, pH 7.9, 0.2 mM EDTA, 0.05% Np40, 10% glycerol) for 1 h each on ice, until the conductivity of the extract matched that of the dialysis buffer. Control nuclear extract (NE) dialyzed to 1.0 M KCl was removed and set aside at 4 °C during the depletion procedure.

The high salt nuclear extract was depleted by application of the extract to a poly U Sepharose column, resuspended in ddH2O on ice, and washed several times in batch in ddH2O followed by equilibration in 1.0 M KCl HENG10. A volume of poly U Sepharose approximately equal to that of the high salt extract was used for the depletion. Extract was applied to the column at approximately one column volume per hour. The flow through of the column (NEΔU) was detected by Bradford, pooled, and dialyzed immediately against three changes of 0.1 M KCl HENG20 (HENG10 but with 20% glycerol) on ice for a total of 3 h. Extract was spun at 25,000 × *g* for 15 min to pellet insoluble material. The protein concentration of NEΔU consistently was two thirds to one half the protein concentration of the control extract (NE); loss was primarily due to precipitation during dialysis. Extract was frozen in small aliquots in liquid nitrogen and stored at −80 °C. The entire procedure was routinely performed in a single day for best results.

Purification

PUF activity eluted from the poly U Sepharose column with 2.0 M KCl HENG10. U2AF was eluted with 3.0 M KCl. Both factors eluted in a broad peak of 1.5- to 2-column volumes and were dialyzed against 0.1 M KCl HENG20. The 2.0-M KCl eluate was applied to a phosphocellulose (Whatmann P11) column according to manufacturer's instructions and washed at 0.1 M and 0.3 M and eluted in batch at 0.6 and 1.0 M KCl. The 0.6 and 1.0 M KCl eluates were pooled and dialyzed to 0.9 M KCl with HENG20 and applied to a 5-mL poly U Sepharose (Pharmacia) column at 0.25 mL/min and eluted with 1.0 M KCl. The poly U Sepharose eluate was dialyzed against 0.1 M KCl HENG20 and applied to a 5-mL S Sepharose column (Pharmacia). The S Sepharose column was step eluted at 0.1 M KCl, 0.55 M KCl, and 1.0 M KCl. Activity eluted at 0.55 M KCl.

Further purification of the activity was performed by gel filtration chromatography on Superose 6 (Pharmacia) by directly applying the S Sepharose eluate in 0.55 M KCl to the Superose column equilibrated at 0.55 M KCl at 0.25 mL/min KCl HENG20. Elution of the PUF activity was compared to that of gel filtration markers (Boehringer Mannheim Biochemicals HPLC markers, #1213 776).

RNAs used in this study

The PIP85a, PIPβG, and AD10 RNA substrates were transcribed from plasmids pPIP85a (Moore & Sharp, 1992), pPIPβG (Crispino et al., 1996), and pAD10 (Konarska & Sharp, 1986) using T7 RNA polymerase (United States Biochemicals) and α^{32}P UTP (New England Nuclear) under standard conditions (Query et al., 1996). The 3′ half substrate PIP85aΔRX was transcribed from pPIP85aΔRX, which is a deletion of PIP85a. pPIP85aΔRX was constructed by digesting pPIP85a with *Eco*RI and *Xho*I, blunting the resultant DNA, and ligating. Pyrimidine tract RNA, bs-ppt and ppt-bs RNAs, and ppt-AG were previously described RNA oligomers and were labeled at their 5′ ends (Query et al., 1997).

Splicing in vitro

In vitro splicing reactions were performed under standard conditions (Grabowski et al., 1984) and were 24% NE or NEΔU supplemented with recombinant U2AF65 or the 3 M KCl eluate of the poly U Sepharose column, the U2AF fraction (shown to contain both U2AF65 and U2AF35 by Western analysis), and the PUF fraction in 60 to 100 mM KCl, 20 mM HEPES, pH 7.9, 5 mM $MgCl_2$, 2 mM ATP, and 5 mM creatine phosphate for 1.5 h at 30 °C. The reaction was stopped with the addition of 250 μL of 2× Proteinase K buffer (Sambrook et al., 1989), 2 μg of glycogen (Boehringer Mannheim Biochemicals) and proteinase K. Reactions were digested for 30 min at 65 °C and precipitated with the addition of 4 vol of 95% ethanol, 1.5 M ammonium acetate, and 5 mM $MgCl_2$. 70% ethanol-washed pellets were resuspended in 5 μL 8 M urea, 1× TBE, heated to 100 °C for 3 min and resolved on 20% acrylamide (19:1)/50% urea (National Diagnostics) gels in 1× TBE.

Complex assembly assays

A3′ complex assembly was performed under standard splicing conditions (above) using an RNA transcribed from PIP85aΔRX (an *Eco*RI-to-*Xho*I deletion of PIP85a) and resolved on 4% acrylamide (60:1) Tris-glycine gels (Konarska & Sharp, 1986). Binding reactions were stopped with the addition of heparin to 5 mg/mL.

Gel shift

Binding of PUF60 to pyrimidine tract RNAs was carried out under splicing conditions at 80 mM KCl to ppt-bs RNA CQ58–19, previously described by Query et al. (1997), for 15 min at 30 °C. Binding reactions were stopped by addition of 0.5 mg/mL heparin and were incubated on ice until loading on 8% acrylamide (60:1) Tris-glycine gels and were resolved at 10 V/cm for 2.5 h at room temperature. Binding was quantified using ImageQuant (Molecular Dynamics) and data was fit to the Hill equation using Kaleidagraph.

Antibodies, immunoblotting, and immunoprecipitations and Gst pull-downs

PUF60 antibodies were generated in two rabbits by inoculation of bacterially expressed PUF60(P) (Covance). PUF60(P) is the *Pst*I internal fragment of PUF60 inserted into the *Pst*I site of pQE31 (Qiagen) and expressed and purified in 6 M urea on Ni^{+2}-NTA Agarose (Qiagen) according to the manufacturer's instructions. Protein was renatured on the Ni-NTA resin (Qiagen) according to manufacturer's instructions. PUF60(P) protein was further purified on mono S Sepharose.

The gel was run in the presence of 2 M urea to yield a sharper p54 band. All other gels were standard SDS 4–20%

polyacrylamide gels (BioRad). p54 antibody (anti pep C) was the generous gift of Nilabh Chaudhary (Chaudhary et al., 1991). p54 immunoblotting was detected using ^{125}I-protein A using standard techniques (Harlow & Lane, 1988). PUF60 immunoblot was detected using the ECL reagent (Amersham).

Immunoprecipitations were performed according to standard procedures (Harlow & Lane, 1988) using the p54 anti-pep C antiserum bound to protein A Sepharose (Pharmacia). In vitro translation reactions were performed according to manufacturer's instructions (Promega).

ACKNOWLEDGMENTS

We would like to thank Nilabh Chaudhary for generously contributing p54 antibodies and his substantial enthusiasm, Nadja Abovich for GstSF1/BBP and SF1/BBP antisera, and Helen Cargill, Margaret Beddall, Yubin Qiu, and Robbyn Issner for expert technical assistance. We would especially like to thank C. Query, A. MacMillan, A. Gil, D. Chasman, and J. Pomerantz for helpful discussion and support and M. Siafaca for wise advice. We thank D. Bowtell, R. Dickins, P. Bouffard, and G. Boire for sharing unpublished data. The protein sequencing of p54 and PUF60 was expertly performed by MIT Biopolymers. We thank the Berkeley *Drosophila* Genome Project for providing the *Drosophila* ESTs CK00112 and LD01934. This work was supported by U.S. Public Health Service MERIT award R37-GM34277 and RO1-AI32486 from the National Institutes of Health to P.A.S. and partially by a Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute.

Received August 24, 1999; returned for revision September 20, 1999; revised manuscript received October 6, 1999

REFERENCES

Abovich N, Liao XC, Rosbash M. 1994. The yeast MUD2 protein: An interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. *Genes & Dev 8*:843–854.

Alam MR, Caldwell BD, Johnson RC, Darlington DN, Mains RE, Eipper BA. 1996. Novel proteins that interact with the COOH-terminal cytosolic routing determinants of an integral membrane peptide-processing enzyme. *J Biol Chem 271*:28636–28640.

Barabino SML, Blencowe BJ, Ryder U, Sproat BS, Lamond AI. 1990. Targeted snRNP depletion reveals an additional role for mammalian U1 snRNP in spliceosome assembly. *Cell 63*:293–302.

Berglund JA, Abovich N, Rosbash M. 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes & Dev 12*:858–867.

Berglund JA, Chua K, Abovich N, Reed R, Rosbash M. 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell 89*:781–787.

Birney E, Kumar S, Krainer AR. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res 21*:5803–5816.

Chaudhary N, McMahon C, Blobel G. 1991. Primary structure of a human arginine-rich nuclear protein that colocalizes with spliceosome components. *Proc Natl Acad Sci USA 88*:8189–8193.

Chiara MD, Champion-Arnaud P, Buvoli M, Nadal-Ginard B, Reed R. 1994. Specific protein–protein interactions between the essential mammalian spliceosome-associated proteins SAP 61 and SAP 114. *Proc Natl Acad Sci USA 91*:6403–6407.

Creighton TE. 1984. *Proteins: Structures and molecular principles*. New York: WH Freeman and Company.

Crispino JD, Mermoud JE, Lamond AI, Sharp PA. 1996. *Cis*-acting elements distinct from the 5′ splice site promote U1-independent pre-mRNA splicing. *RNA 2*:664–673.

Dickson BJ. 1998. Photoreceptor development: Breaking down the barriers. *Curr Biol 8*:R90–R92.

Dignam JD, Lebovitz RM, Roeder RD. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res 11*:1475–1489.

Goldenberg DP, Berget PB, King J. 1982. Maturation of the tail spike endorhamnosidase of *Salmonella* phage P22. *J Biol Chem 257*: 7864–7871.

Gozani O, Feld R, Reed R. 1996. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes & Dev 10*:233–243.

Grabowski PJ, Padgett RA, Sharp PA. 1984. Messenger RNA splicing in vitro: An excised intervening sequence and a potential intermediate. *Cell 37*:415–427.

Harlow E, Lane D. 1988. *Antibodies: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Hillier LD, Lennon G, Becker M, Bonaldo MF, Chiapelli B, Chissoe S, Dietrich N, DuBuque T, Favello A, Gish W, Hawkins M, Hultman M, Kucaba T, Lacy M, Le M, Le N, Mardis E, Moore B, Morris M, Parsons J, Prange C, Rifkin L, Rohlfing T, Schellenberg K, Soares M, Tan F, Thierry-Meg J, Trevaskis E, Underwood K, Wohldman P, Waterston R, Wilson W, Marra M. 1996. Generation and analysis of 280,000 human expressed sequence tags. *Genome Res 6*:807–828.

Imai H, Chan EK, Kiyosawa K, Fu XD, Tan EM. 1993. Novel nuclear autoantigen with splicing factor motifs identified with antibody from hepatocellular carcinoma. *J Clin Invest 92*:2419–2426.

Kanaar R, Roche SE, Beall EL, Green MR, Rio DC. 1993. The conserved pre-mRNA splicing factor U2AF from *Drosophila*: Requirement for viability. *Science 262*:569–573.

Kennedy CF, Kramer A, Berget SM. 1998. A role for SRp54 during intron bridging of small introns with pyrimidine tracts upstream of the branch point. *Mol Cell Biol 18*:5425–5434.

Konarska MM, Sharp PA. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell 46*: 845–855.

Li S, Li Y, Carthew RW, Lai ZC. 1997. Photoreceptor cell differentiation requires regulated proteolysis of the transcriptional repressor Tramtrack. *Cell 90*:469–478.

Meyer RK, McKinley MP, Bowman KA, Braunfeld MB, Barry RA, Pruisner SB. 1986. Separation and properties of cellular and scrapie prion proteins. *Proc Natl Acad Sci USA 83*:2310–2314.

Michaud S, Reed R. 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes & Dev 5*:2534–2546.

Moore MJ, Sharp PA. 1992. Site-specific modification of pre-mRNA: The 2′ hydroxyl groups at the splice sites. *Science 256*:992–997.

Pan KM, Baldwin M, Nguyen J, Gasset M, Serban A, Groth D, Mehlorn I, Huang Z, Fletterick RJ, Cohen FE, Pruisner SB. 1993. Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc Natl Acad Sci USA 90*:10962–10966.

Peek R, Pruijn GJ, van der Kemp AJ, van Venrooij WJ. 1993. Subcellular distribution of Ro ribonucleoprotein complexes and their constituents. *J Cell Sci 106*:929–935.

Potashkin J, Naik K, Wentz HK. 1993. U2AF homolog required for splicing in vivo. *Science 262*:573–575.

Query CC, McCaw PS, Sharp PA. 1997. A minimal spliceosomal complex A recognizes the branch site and polypyrimidine tract. *Mol Cell Biol 17*:2944–2953.

Query CC, Strobel SA, Sharp PA. 1996. Three recognition events at the branch-site adenine. *EMBO J 15*:1392–1402.

Rain JC, Rafi Z, Rhani Z, Legrain P, Kramer A. 1998. Conservation of functional domains involved in RNA binding and protein–protein interactions in human and *Saccharomyces cerevisiae* pre-mRNA splicing factor SF1. *RNA 4*:551–565.

Roscigno RF, Weiner M, Garcia-Blanco MA. 1993. A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. *J Biol Chem 268*: 11222–11229.

Rudner DZ, Kanaar R, Breger KS, Rio DC. 1996. Mutations in the small subunit of the *Drosophila* U2AF splicing factor cause lethality and developmental defects. *Proc Natl Acad Sci USA 93*: 10333–10337.

Ruskin B, Zamore PD, Green MR. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell 52*: 207–219.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Staknis D, Reed R. 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol Cell Biol 14*:7670–7682.

Steinbacher S, Seckler R, Miller S, Steipe B, Huber R, Reinemer P. 1994. Crystal structure of P22 tailspike protein interdigitated subunits in a thermostable trimer. *Science 265*:383–386.

Tronchere H, Wang J, Fu XD. 1997. A protein related to splicing factor U2AF35 that interacts with U2AF65 and SR proteins in splicing of pre-mRNA. *Nature 388*:397–400.

Wahren M, Mellqvist E, Vene S, Ringertz NR, Pettersson I. 1996. Nuclear colocalization of the Ro 60-kDa autoantigen and a subset of U snRNP domains. *Eur J Cell Biol 70*:189–197.

Wentz HK, Potashkin J. 1996. The small subunit of the splicing factor U2AF is conserved in fission yeast. *Nucleic Acids Res 24*:1849–1854.

Zamore PD, Green MR. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc Natl Acad Sci USA 86*:9243–9247.

Zamore PD, Green MR. 1991. Biochemical characterization of U2 snRNP auxiliary factor: An essential pre-mRNA splicing factor with a novel intranuclear distribution. *EMBO J 10*:207–214.

Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature 355*: 609–614.

Zhang M, Zamore PD, Carmo-Fonseca M, Lamond AI, Green MR. 1992. Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc Natl Acad Sci USA 89*:8769–8773.

Zhang WJ, Wu JY. 1996. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Mol Cell Biol 16*:5400–5408.

Zorio DA, Blumenthal T. 1999. U2AF35 is encoded by an essential gene clustered in an operon with RRM/cyclophilin in *Caenorhabditis elegans*. *RNA 4*:487–494.

Zorio DA, Lea K, Blumenthal T. 1997. Cloning of *Caenorhabditis* U2AF65: An alternatively spliced RNA containing a novel exon. *Mol Cell Biol 17*:946–953.

Zuo P, Maniatis T. 1996. The splicing factor U2AF35 mediates critical protein–protein interactions in constitutive and enhancer-dependent splicing. *Genes & Dev 10*:1356–1368.