

Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit

(splicing/spliceosome/RNA binding/arginine, serine rich/coiled body)

MENG ZHANG*, PHILLIP D. ZAMORE*, MARIA CARMO-FONSECA†, ANGUS I. LAMOND†,
AND MICHAEL R. GREEN*

*Program in Molecular Medicine, University of Massachusetts Medical Center, 373 Plantation Street, Worcester, MA 01605; and †European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 102209, D-6900 Heidelberg, Federal Republic of Germany

Communicated by Mary Edmonds, June 19, 1992

ABSTRACT U2 small nuclear ribonucleoprotein auxiliary factor (U2AF), an essential mammalian splicing factor, is composed of two subunits: a 65-kDa protein (U2AF⁶⁵), which binds the pre-mRNA polypyrimidine tract and is required for *in vitro* splicing, and an associated 35-kDa protein (U2AF³⁵). Here we report the isolation of a cDNA encoding U2AF³⁵. U2AF³⁵ contains sequence motifs found in several mammalian pre-mRNA splicing factors. We show directly that U2AF⁶⁵ and U2AF³⁵ interact with each other and delineate the regions of both proteins that mediate this interaction. Using anti-peptide antibodies against U2AF³⁵, we show that the protein has the intracellular distribution characteristic of U2AF⁶⁵. Both U2AF⁶⁵ and U2AF³⁵ are concentrated in a small number of nuclear foci corresponding to coiled bodies, subnuclear organelles first identified by light microscopy in 1903.

In vitro experiments demonstrate that pre-mRNA splicing begins with the assembly of a multicomponent complex on the pre-mRNA (reviewed in refs. 1–4). This complex, the spliceosome, contains two types of factors: small nuclear ribonucleoprotein particles (snRNPs) and proteins. These components assemble on the RNA according to an ordered pathway. Several of these steps require ATP hydrolysis, the first of which is U2 snRNP binding.

Binding of U2 snRNP to the pre-mRNA branch site requires, in addition to U2 snRNP itself, U1 snRNP (5) and at least three protein factors: SF1, SF3 (6), and U2 snRNP auxiliary factor (U2AF; refs. 7 and 8). Of these protein factors, U2AF has been most extensively characterized.

Purified HeLa cell U2AF comprises two polypeptides, 65 and 35 kDa (8, 9). Our previous studies (8–10) have shown that the U2AF large subunit (U2AF⁶⁵) was essential for *in vitro* splicing of an adenovirus major late and a human β -globin pre-mRNA, but that the U2AF small subunit (U2AF³⁵) was dispensable. However, U2AF³⁵ may be required for splicing other pre-mRNAs (e.g., those that are inefficiently spliced), regulating selection of splice sites, or providing some intracellular function (e.g., subcellular localization). U2AF³⁵ is evolutionarily conserved (9) and copurifies with U2AF⁶⁵ (8). In this report, we describe the isolation of a cDNA clone encoding U2AF³⁵.[‡]

METHODS

Membrane-Bound Protein Binding Assay. Ten microliters of HeLa cell nuclear extract, U2AF-depleted nuclear extract, or purified U2AF was separated on a SDS/12.5% polyacrylamide gel and then transferred to poly(vinylidene difluoride) membranes (Millipore) by electroblotting. The membrane

was blocked for 1 h with 5% bovine serum albumin (BSA) in 20 mM Hepes-KOH, pH 7.9/3 mM MgCl₂/0.1 mM EDTA/0.1% Tween 20/10% (vol/vol) glycerol/1 mM dithiothreitol/100 mM to 1 M KCl. The membrane was incubated overnight with ³⁵S-labeled *in vitro* translated U2AF⁶⁵ or U2AF³⁵ (20 μ l) in 10 ml of blocking buffer and was then washed three times with the same buffer without BSA. The membrane was then dried and exposed to x-ray film.

Intracellular Localization Experiments. The anti-peptide antibody to U2AF³⁵ was described by Zamore and Green (9), where it was referred to as anti-pepC. The affinity-purified antibody (1.5 mg/ml) was injected into the nucleus of living HeLa cells by an AIS microinjection system (11). Cells were incubated for 1 h at 37°C, washed with phosphate-buffered saline (PBS), transferred to ice, extracted with 0.5% Triton X-100 in CSK buffer for 3 min, and fixed for 10 min with 3.7% paraformaldehyde in CSK buffer as described (12). After fixation, the cells were washed with 0.5% Tween 20 in PBS and then sequentially incubated with anti-p80 coilin patient antiserum (kindly provided by E. H. Tan, Scripps Clinic) diluted 1:200 in PBS and Texas Red-conjugated goat anti-human IgG and fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (Dianova, Hamburg, F.R.G.). Staining of fixed cells and confocal fluorescence microscopy were as described (12, 13).

RESULTS AND DISCUSSION

Isolation of a cDNA Clone for U2AF³⁵. To isolate a cDNA encoding the 35-kDa U2AF polypeptide, purified HeLa cell U2AF (8, 9) was separated by electrophoresis and transferred to a nitrocellulose membrane. The region of the membrane corresponding to the 35-kDa subunit was isolated and digested with trypsin. The resulting peptides were resolved by high-pressure liquid chromatography, and the sequences of three were determined (boxed). The sequences of two peptides permitted the design of degenerate primers that allowed us to isolate by PCR the nondegenerate DNA encoding each peptide. The two PCR products were cloned and sequenced. PCR primers were next designed to permit the isolation of the DNA between the two smaller DNA sequences. The product of this PCR was again cloned and sequenced. Finally, a portion of the sequence was used to screen $\approx 5 \times 10^6$ plaques of a λ gt10 human fetal brain cDNA library. Twenty-five positive clones were identified and purified and two of these were subcloned and sequenced. Both encoded all three peptides identified by protein microsequencing.

Fig. 1 shows the sequence of a 931-base-pair cDNA for U2AF³⁵. The cDNA contains a 240-amino acid open reading

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: snRNP, small nuclear ribonucleoprotein; U2AF, U2 snRNP auxiliary factor.

[‡]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M96982).

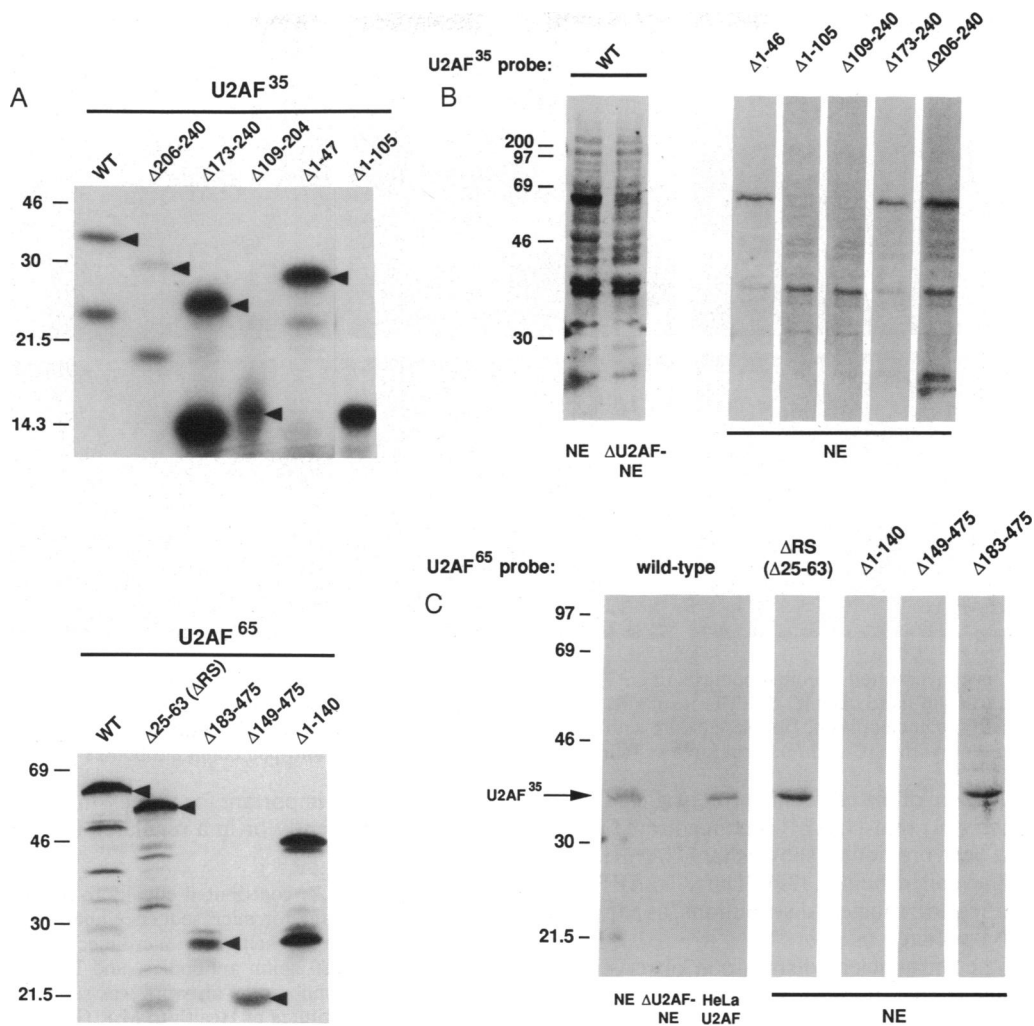


FIG. 3. Mapping the interaction surfaces of the two U2AF subunits. (A) Protein derivatives used in this study. The U2AF⁶⁵ and U2AF³⁵ coding sequences were fused to the human β -globin 5' untranslated sequence. RNA was transcribed from these plasmids with SP6 RNA polymerase and was translated in a wheat germ extract. Protein derivatives were created by linearizing the plasmid at appropriate restriction sites. The deletion mutant Δ RS has been described (10). The full-length *in vitro* translation products are indicated by arrows. Size standards (kDa) are indicated on the left. (B) U2AF³⁵ binds U2AF⁶⁵. The protein preparation used is indicated below and the ³⁵S-labeled *in vitro* translated derivative is indicated above. Size standards (kDa) are indicated on the left. The position of U2AF⁶⁵ is shown. (C) U2AF⁶⁵ binds U2AF³⁵. Other details are the same as in B. The position of U2AF³⁵ is shown. WT, wild type; NE, nuclear extract.

did not affect the ability of U2AF³⁵ to bind U2AF⁶⁵, but further deletion to residue 105 eliminated the interaction. C-terminal deletion of U2AF³⁵ to residue 173 had no effect, but further deletion to residue 109 abolished interaction with U2AF⁶⁵. Thus, a region between amino acids 47 and 172 of U2AF³⁵ is involved in binding to U2AF⁶⁵. This region does not include the RS motif or the glycine run.

A similar examination of U2AF⁶⁵ sequences (Fig. 3C) showed that neither its N-terminal RS motif (amino acids 25–63) nor an intact C-terminal sequence-specific RNA binding domain (amino acids 151–462) was required for interaction with U2AF³⁵. An in-frame deletion of the N-terminal RS motif (Δ RS) bound U2AF³⁵ normally, whereas deletion of the N-terminal 140 amino acids eliminated the interaction with U2AF³⁵. Deletion of amino acids C-terminal to residue 183 did not affect binding but further deletion to residue 149 disrupted the interaction. Thus, a region between amino acids 64 and 182 of U2AF⁶⁵ is involved in binding to U2AF³⁵.

Intracellular Localization of U2AF³⁵. The strong, specific *in vitro* interaction between the two U2AF subunits prompted us to ask whether the two proteins colocalize *in vivo*. We have previously reported that U2AF⁶⁵ is distributed throughout the nucleoplasm and concentrated in a small number of

nuclear foci (9). This same intranuclear distribution has been observed for U1 snRNA (13). U2, U4/6, and U5 snRNAs are highly concentrated in the foci (13) and can only be detected in the nucleoplasm upon long times of hybridization (28) or following extraction of fixed cells with SDS (29).

It has been recently shown that these snRNA-containing foci correspond to previously identified nuclear structures designated coiled bodies (29). Autoantibodies to the protein p80 coilin, a constituent of the coiled body (30), provide a convenient marker for these intranuclear structures. We used double-label immunofluorescence to determine the intranuclear distribution of U2AF³⁵ relative to that of p80 coilin. In these experiments, an affinity-purified, anti-peptide antibody against U2AF³⁵ (described in ref. 9) was either microinjected into the cell nucleus before fixation (Fig. 4 A–C) or was used to stain fixed cells (Fig. 4 D–F). Fig. 4 A and D shows that like U2AF⁶⁵ (9), U2AF³⁵ was concentrated in a small number of foci (arrows) and was diffusely distributed throughout the nucleoplasm excluding nucleoli. The staining observed with the anti-U2AF³⁵ antibody was eliminated by addition of the same U2AF³⁵ peptide used to derive the antibody (pepC; ref. 9) but not by an equal amount of an unrelated peptide (data not shown). These U2AF³⁵-containing foci correspond to the

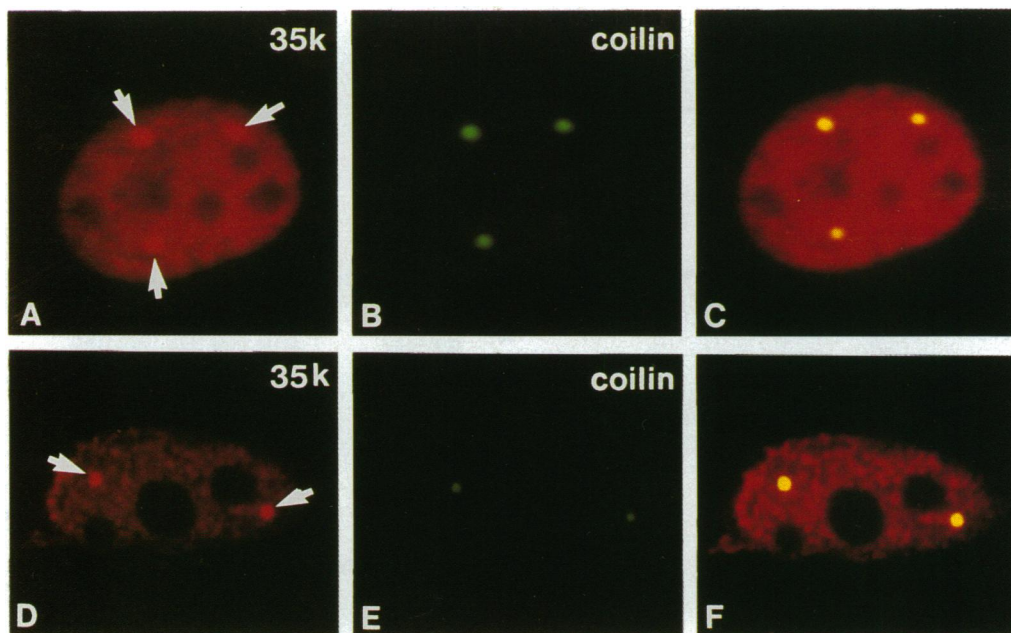


FIG. 4. U2AF³⁵ is concentrated in coiled bodies. An anti-peptide against U2AF³⁵ was either microinjected into the nucleus of living HeLa cells (A–C) or used to stain fixed cells (D–F). The results (A and D) show that U2AF³⁵ was concentrated in discrete foci (arrows) as well as distributed throughout the nucleoplasm. Double-labeling with anti-p80 coilin antibody (B, C, E, and F) demonstrates that these foci correspond to coiled bodies. In the overlay (C and F), structures staining with both anti-U2AF³⁵ and anti-p80 coilin antibodies appear yellow.

intracellular localization of the p80 coilin protein (Fig. 4 B, C, E, and F) demonstrating that U2AF³⁵ is concentrated in the coiled body. It has been previously shown that U2AF⁶⁵ is also concentrated in the coiled bodies (29). Thus, U2AF³⁵ and U2AF⁶⁵ have identical intracellular distributions, as expected if they are stably associated *in vivo*.

In contrast to the intracellular distribution observed for U2AF and the spliceosomal snRNAs, some snRNP structural proteins and the splicing factor SC-35 are concentrated in a larger number of discrete "speckles" (reviewed in ref. 1). Recently, it was proposed that the RS motif targets proteins to these speckled nuclear subregions (31). Our studies with U2AF demonstrate that the RS motif is not a dominant subnuclear targeting signal: both U2AF³⁵ (this report) and U2AF⁶⁵ (10) contain RS motifs, but neither has a speckled distribution. Instead, both U2AF³⁵ and U2AF⁶⁵ are distributed throughout the nucleoplasm and concentrated in coiled bodies. Our studies do not exclude, however, that the RS motif can affect nuclear localization by, for example, interacting with spliceosomal components, which themselves are distributed nonrandomly.

Several factors implicated in mammalian pre-mRNA splicing are ≈ 35 kDa, including U2AF³⁵, ASF/SF2 (20, 21), and SC-35 (32). The relationship among these three factors is unclear and has been a source of confusion. The sequence of U2AF³⁵ (Fig. 1) reveals that it is distinct from both ASF/SF2 (20, 21) and SC-35 (33, 34).

In light of the strong, specific interaction between the two U2AF subunits, what role might U2AF³⁵ play in pre-mRNA splicing? We have recently shown that the RS motif of U2AF⁶⁵ is required for splicing *in vitro* (10). U2AF⁶⁵ binds a specific sequence element, the polypyrimidine tract, thereby bringing the U2AF⁶⁵ RS motif to the pre-mRNA. Because U2AF³⁵ interacts with U2AF⁶⁵, its RS motif will also be brought to the pre-mRNA following binding of U2AF⁶⁵. We speculate that the RS motifs of U2AF³⁵ and U2AF⁶⁵ cooperate with one another to facilitate splicing by, for example, contacting the same component during spliceosome assembly. Such cooperative action of the two U2AF subunits may

be particularly important for processing pre-mRNAs that are spliced inefficiently or in a regulated fashion.

M.Z. and P.D.Z. contributed equally to the work reported here. We thank W. Lane for microsequence analysis, Drs. R. Peperkok and W. Ansorge for help with microinjection, Prof. Eng Tan for providing anti-p80 coilin antibodies, and L. Kunkel for the λ gt10 human fetal brain library. This work was supported by a grant from the National Institutes of Health to M.R.G.

- Green, M. R. (1991) *Annu. Rev. Cell Biol.* 7, 559–599.
- Maniatis, T. & Reed, R. (1987) *Nature (London)* 325, 673–678.
- Sharp, P. A. (1987) *Science* 235, 766–771.
- Guthrie, C. (1991) *Science* 253, 157–163.
- Barabino, S. M. L., Blencowe, B. J., Ryder, U., Sproat, B. S. & Lamond, A. I. (1990) *Cell* 63, 293–302.
- Kramer, A. & Utans, U. (1991) *EMBO J.* 10, 1505–1510.
- Ruskin, B., Zamore, P. D. & Green, M. R. (1988) *Cell* 52, 207–219.
- Zamore, P. D. & Green, M. R. (1989) *Proc. Natl. Acad. Sci. USA* 86, 9243–9247.
- Zamore, P. D. & Green, M. R. (1991) *EMBO J.* 10, 207–214.
- Zamore, P. D., Patton, J. G. & Green, M. R. (1992) *Nature (London)* 355, 609–614.
- Ansorge, W. & Peperkok, R. J. (1988) *Biochem. Biophys. Methods* 16, 283–292.
- Carmo-Fonseca, M., Peperkok, R., Sproat, B. S., Ansorge, W., Swanson, M. S. & Lamond, A. I. (1991) *EMBO J.* 10, 1863–1873.
- Carmo-Fonseca, M., Tollervey, D., Barabino, S. M. L., Merdes, A., Brunner, C., Zamore, P. D., Green, M. R., Hurt, E. & Lamond, A. I. (1991) *EMBO J.* 10, 195–270.
- Kozak, M. (1989) *J. Cell Biol.* 108, 229–241.
- Gardiner-Garden, M. & Fromer, M. J. (1987) *Mol. Biol.* 196, 261–282.
- Theissen, H., Etzerodt, M., Reuter, R., Schneider, C., Lottspeich, F., Argos, P., Luhrmann, R. & Philipson, L. (1986) *EMBO J.* 5, 3209–3217.
- Query, C. C. & Keene, J. D. *Cell* 51, 211–220.
- Etzerodt, M., Vignali, R., Ciliberto, G., Scherly, D., Mattaj, I. W. & Philipson, L. (1987) *EMBO J.* 7, 4311–4321.
- Mancebo, R., Lo, P. C. H. & Mount, S. M. (1990) *Mol. Cell. Biol.* 10, 2492–2502.

20. Ge, H., Zuo, P. & Manley, J. L. (1991) *Cell* **66**, 373–382.
21. Krainer, A. R., Mayeda, A., Kozak, D. & Binns, G. (1991) *Cell* **66**, 383–394.
22. Boggs, R. T., Gregor, P., Idriss, S., Belote, J. M. & McKeown, M. (1987) *Cell* **50**, 739–747.
23. Amrein, H., Gorman, M. & Nothiger, R. (1988) *Cell* **55**, 1025–1035.
24. Goralski, T. J., Edstrom, J.-E. & Baker, B. S. (1989) *Cell* **56**, 1011–1018.
25. Voelker, R. A., Gibson, W., Graves, J. P., Sterling, J. F. & Eisenberg, M. T. (1991) *Mol. Cell. Biol.* **11**, 894–905.
26. Chou, T.-B., Zachar, Z. & Bingham, P. M. (1987) *EMBO J.* **6**, 4095–4104.
27. Kenan, D. J., Query, C. C. & Keene, J. D. (1991) *Trends Biochem. Sci.* **16**, 214–220.
28. Huang, S. & Spector, D. L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 305–308.
29. Carmo-Fonseca, M., Peperkock, R., Carvalho, M. T. & Lamond, A. I. (1992) *J. Cell Biol.* **117**, 1–14.
30. Andrade, L. E. C., Chan, E. K. L., Raska, I., Peebles, C. L., Roos, G. & Tan, E. M. (1991) *J. Exp. Med.* **173**, 1407–1419.
31. Li, H. & Bingham, P. M. (1991) *Cell* **67**, 335–342.
32. Fu, X. D. & Maniatis, T. (1990) *Nature (London)* **343**, 437–441.
33. Vellard, M., Sureau, A., Soret, J., Martinerie, C. & Perbal, B. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2511–2515.
34. Fu, X. D. & Maniatis, T. (1992) *Science* **256**, 535–538.