RNA (2001), *7*:457–470. Cambridge University Press. Printed in the USA.

Purification of *Drosophila* snRNPs and characterization of two populations of functional U1 particles

EMMANUEL LABOURIER and DONALD C. RIO
Department of Molecular and Cell Biology, University of California, Berkeley, California 94720-3204, USA

ABSTRACT

U1 snRNP is required at an early stage during assembly of the spliceosome, the dynamic ribonucleoprotein (RNP) complex that performs nuclear pre-mRNA splicing. Here, we report the purification of U1 snRNP particles from *Drosophila* nuclear extracts and the characterization of their biochemical properties, polypeptide contents, and splicing activities. On the basis of their antigenicity, apparent molecular weight, and by peptide sequencing, the *Drosophila* 70K, SNF, B, U1-C, D1, D2, D3, E, F, and G proteins are shown to be integral components of these particles. Sequence database searches revealed that both the U1-specific and the Sm proteins are extensively conserved between human and *Drosophila* snRNPs. Furthermore, both species possess a conserved intrinsic U1-associated kinase activity with identical substrate specificity in vitro. Finally, our results demonstrate that a second type of functional U1 particle, completely lacking the U1/U2-specific protein SNF and the associated protein kinase activity, can be isolated from cultured Kc cell or Canton S embryonic nuclear extracts. This work describes the first characterization of a purified *Drosophila* snRNP particle and reinforces the view that their activity and composition, with the exception of the atypical bifunctional U1-A/U2-B″ SNF protein, are highly conserved in metazoans.

Keywords: pre-mRNA splicing; Sm proteins; SNF; U1-70K; U1-C; U1 snRNP kinase activity

INTRODUCTION

Splicing of nuclear introns occurs by a two-step pathway, conserved in eukaryotes, which takes place in a large ribonucleoprotein structure, termed the spliceosome (Sharp, 1994; Staley & Guthrie, 1998). The assembly of this complex requires the ordered interaction of numerous splicing factors and the five small nuclear ribonucleoproteins (snRNP) U1, U2, U5, and U4/U6. These spliceosomal snRNPs each contain a small RNA molecule, several snRNP-specific proteins and seven common proteins (B/B′, D1, D2, D3, E, F, and G; Will et al., 1993). The Sm core proteins are evolutionarily conserved (Hermann et al., 1995; Séraphin, 1995) and found in organisms as diverse as yeast, plants, fly, mouse, and man, with a molecular weight ranging from 9 to 29 kDa. They associate with the RNA polymerase II-transcribed U1, U2, U4, and U5 snRNAs in the cytoplasm on an uridylic acid-rich region flanked by two stem-loop structures, the so-called Sm binding site (Branlant et al., 1982). The formation of this Sm core RNP structure is essential for the hypermethylation of the snRNA 5′ cap structure to generate the 2,2,7-trimethylguanosine (m_3G) form (Mattaj, 1986). The m_3G cap, together with the Sm proteins, also constitutes the nuclear localization signal required for the nuclear targeting of the snRNP particles (Fischer & Lührmann, 1990; Hamm et al., 1990; Fischer et al., 1993). In contrast, the U6 snRNA, which is transcribed by RNA polymerase III, possesses a monomethylguanosine cap and associates in the nucleus with the U4 RNP to form the U4/U6 functional particle.

In addition to U1 snRNA and the common Sm proteins, the human U1 snRNP particle contains three U1-specific proteins denoted 70K, A, and C. Both protein–protein and protein–RNA interactions are required for the association of these proteins with U1 snRNP particles. For example, U1-70K and U1-C interact with Sm proteins during U1 snRNP assembly (Nelissen et al., 1994), whereas U1-70K and U1-A interact, through their highly conserved RNA binding domain (RBD or RNP-CS), with stem-loops I and II of U1 snRNA, respectively (Patton et al., 1989; Scherly et al., 1989). These specific proteins are primarily responsible for fulfilling the

Reprint requests to: Donald C. Rio, Department of Molecular and Cell Biology, University of California, 401 Barker Hall-3204, Berkeley, California 94720-3204, USA; e-mail: don_rio@uclink4.berkeley.edu.

function of U1 snRNP during the early events of spliceosome assembly, that is, the recognition of 5′ splice sites on the pre-mRNA. In mammalian nuclear extracts, this stage involves base pairing between the 5′ end of U1 snRNA with conserved sequences spanning the 5′ splice site (Zhuang & Weiner, 1986) and the formation of a stable U1 snRNP/pre-mRNA complex designated the early or E complex (Michaud & Reed, 1991). U1-C, but not U1-A, was shown to enhance the formation of E complexes (Will et al., 1996) as well as the interaction of U1 snRNP to the 5′ splice site (Heinrichs et al., 1990; Jamison et al., 1995).

It is also at this stage that intron definition takes place and the commitment to undergo splicing is made. Therefore, the early steps in spliceosome assembly are the primary target for regulatory factors that alter splice site choice or modulate splicing activity. Consistent with this idea, the splicing regulator SF2/ASF (Krainer et al., 1990) was shown to enhance the binding of U1 snRNP to the 5′ splice site by direct interaction with the U1-70K protein (Jamison et al., 1995; Kohtz et al., 1994). Such interactions are mediated by arginine–serine-rich (RS) domains, a cellular target for a number of phosphorylation/dephosphorylation events in vivo (Misteli, 1999). Interestingly, an extensive phosphorylation in vivo of the U1-70K protein and a U1 snRNP-associated kinase activity specific for RS domain-containing proteins in vitro have been reported (Woppmann et al., 1990, 1993). This activity may be crucial to regulate protein–protein interactions and/or the recruitment of splicing factors during the early recognition of 5′ splice sites.

Most of the functional information regarding the human snRNPs has come from studies on highly purified particles from HeLa cells. In particular, antibodies specific for the snRNAs m_3G cap allowed a single-step purification of snRNPs from nuclear extracts (Bringmann et al., 1983; Krainer, 1988). Despite their large size and very similar composition, the native particles recovered by desorption with an excess of m_3G or m^7G can subsequently be separated from each other by anion-exchange chromatography. It has also been possible to purify and study the function of U1 snRNP particles specifically lacking one or several of the U1-70K, U1-A, and U1-C proteins (Bach et al., 1990b).

The *Drosophila* U snRNAs closely resemble those of human, and a monoclonal anti-m_3G cap antibody was used to immunoaffinity purify a mixture of *Drosophila* snRNP particles from tissue culture cell nuclear extracts (Paterson et al., 1991). Although this study demonstrated that the major protein components of these particles appeared to be conserved between fly and vertebrates, their individual identity, composition, and functions in splicing are poorly defined. Only three *Drosophila* Sm proteins have been identified to date: the B (Brunet et al., 1993), D3 (Ivanov et al., 1998), and F proteins (Vincent et al., 1990). These proteins show an extensive conservation with their human counterparts and the same is true for the *Drosophila* U1-specific proteins 70K (Mancebo et al., 1990) and SNF (U1-A/U2-B″; Harper et al., 1992). However, the latter, encoded by the essential gene *sans-fille*, is an integral part of both U1 and U2 snRNPs in vivo and carries out the functions of U1-A and U2-B″ proteins (Flickinger & Salz, 1994; Polycarpou-Schwarz et al., 1996).

To initiate the characterization of *Drosophila* snRNPs, we have used an anti-m_3G immunochromatography followed by anion exchange microchromatography to individually purify the *Drosophila* snRNP particles from Kc cell nuclear extracts. The recovery of significant amounts of highly purified functional U1 snRNP particles allowed us to identify their protein components and activities, as well as to compare them with those of the well-characterized human U1 snRNP particle. The purification of functional *Drosophila* snRNPs that differ in their protein composition and/or activities, combined with *Drosophila* genetics, should facilitate the understanding of how snRNP particles are assembled and function in vivo.

RESULTS

Purification of *Drosophila* snRNP particles from Kc cell nuclear extracts

Drosophila nuclear extracts supporting in vitro pre-mRNA splicing were prepared from $\sim 3 \times 10^{10}$ exponentially growing Kc cells (Rio, 1988) and fractionated by immunoaffinity chromatography using the anti-m_3G monoclonal antibody K121 (Krainer, 1988). All the major U snRNPs bound to the affinity column and were recovered by successive elutions with free m^7G at low ionic strength (Fig. 1A, lanes 3 and 4). This snRNP preparation was highly pure at the RNA level and markedly enriched in the U1 snRNP particle. Based on the snRNA abundance in each fraction determined by RNA blot hybridization (data not shown), the binding of snRNPs to the column was not quantitative. Whereas 90–95% of the U1 snRNP was bound to the K121 column, only approximately 50% of the other snRNPs were bound (Fig. 1A, compare lanes 1 and 2).

As previously reported for *Drosophila* snRNP particles purified from S2 cells with the anti-m_3G H-20 antibody (Paterson et al., 1991), the purification yield was significantly lower than with HeLa cell extracts (0.5 mg per 10^{10} Kc cells). To resolve and recover individually purified snRNP particles in the low microgram or nanogram range, we decided to use an anion exchange micropreparative chromatography (see Materials and Methods). Figure 1B,C shows the elution profile of snRNPs from a 100 μL Mono Q column and the RNA composition of the respective RNP peaks. This profile closely resembles that of snRNPs purified from HeLa cells (Bach et al., 1990a). As observed with the human

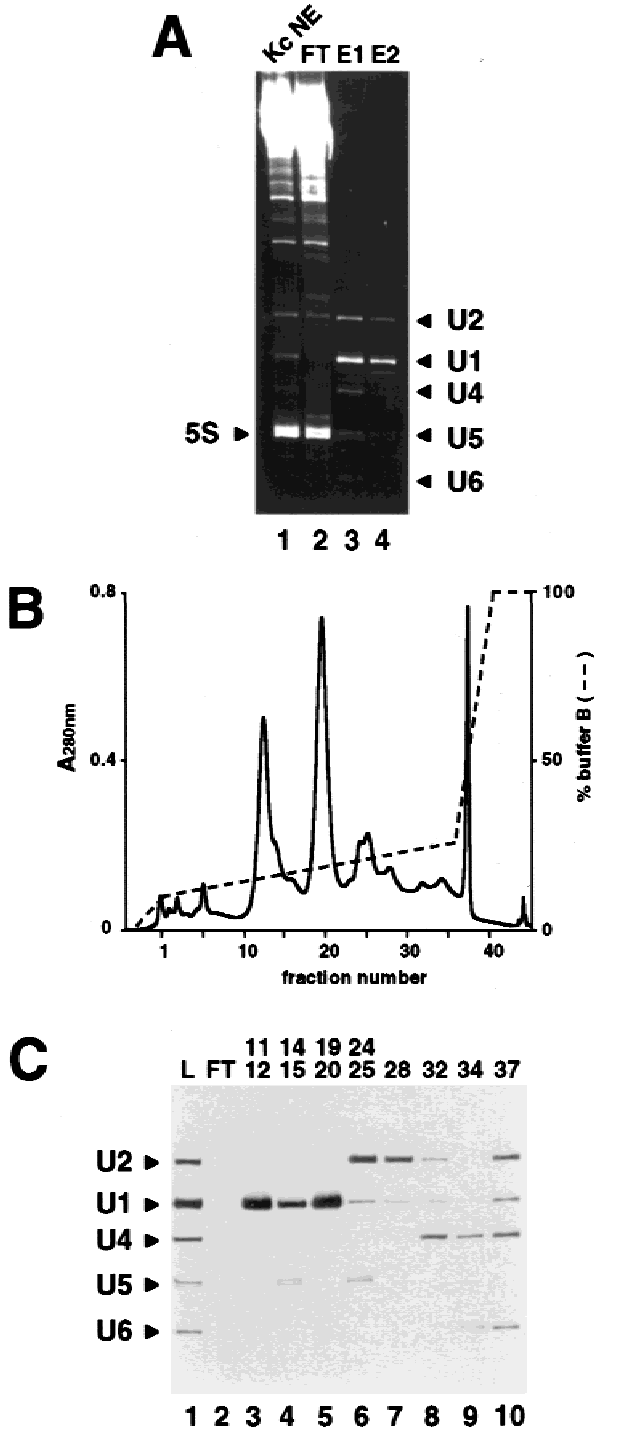

FIGURE 1. Purification of *Drosophila* snRNPs. **A**: Immunopurification of snRNPs using an affinity column. Kc cell nuclear extract loaded on the K121 column (Kc NE), the unbound flow through (FT) and the two successive elutions with 20 mM m^7G (E1, E2) were PK treated and RNAs were analyzed on a 10% polyacrylamide/urea gel stained with ethidium bromide. U snRNAs identity was confirmed by northern blot. The position of 5S rRNA is shown on the left of the panel. **B**: Elution profile of snRNPs from a 100-μL Mono Q column using the SMART system. Fractionation was performed with a linear KCl gradient at a flow rate of 0.1 mL/min and with 100-μL fractions being collected. Buffer B contained 2 M KCl. **C**: Silver staining of snRNAs purified on the Mono Q column and resolved on a 12% denaturing gel. The number of each fraction is indicated on the top of the panel. L: mixture of snRNPs loaded on the Mono Q column. FT: unbound flow through.

snRNP particles, U1 snRNP elutes at 370 mM KCl (fraction 11/12, Fig. 1C, lane 3), U2 together with U5 at 490 mM KCl (fractions 24/25, Fig. 1C, lane 6), U2 at 510 mM KCl (fraction 28, Fig. 1C, lane 7), and the fractions containing U4 and U6 at higher salt concentration: U4 with some residual U2 at 560 mM and U4/U6 at 585 mM (fractions 32 and 34, Fig. 1C, lanes 8 and 9, respectively). Furthermore, due to the small volume of the fractions collected (100 μL) and the sensitivity of the method we used, the fraction of total U5 snRNP that typically contaminates the HeLa U1 snRNP preparation (Bach et al., 1990a), elutes here at 390–400 mM KCl and is separated from the *Drosophila* U1 snRNP peak (fractions 14/15, Fig. 1C, lane 4). The last peak contains all five snRNAs (fraction 37, Fig. 1C, lane 10) and elutes at very high ionic strength ($\sim$1 M KCl). This fraction is lacking most of the U1 and U5 particles loaded on the column (Fig. 1C, compare lanes 1 and 10). Unexpectedly, the major difference from the elution profile observed with HeLa snRNPs was the presence of an additional strong U1 snRNP peak at 440 mM KCl (fraction 19/20, Fig. 1C, lane 5). The fractionation of five independent nuclear extracts showed a reproducible profile with an average 1:1 ratio between the two different *Drosophila* U1 snRNP peaks (data not shown). Each of them represented about 35% of the total U1 loaded on the column, the 30% remaining being mainly distributed between the fractions 14, 15, 24, 25, and 37 (Fig. 1B,C).

Two types of functional U1 snRNP particles are purified from Kc cell nuclear extract

The above results demonstrate that *Drosophila* U1 particles can be isolated directly from crude nuclear extracts. Therefore, the fractions corresponding to the two U1 peaks (dU1 370 and dU1 440) were pooled separately, dialyzed against the same buffer, and used at 50 μg/mL for further characterization. We first asked whether these snRNPs were active in splicing and able to complement a splicing extract that had been inactivated by oligonucleotide-directed RNase H treatment. As shown in Figure 2, incubation in a Kc cell nuclear extract of a short oligodeoxynucleotide complementary to the 5′ end of U1 snRNA results in a complete and specific cleavage of this snRNA (bottom panel, compare lanes 2 and 3). As expected, this treatment totally inhibits the splicing of a model *Tra* pre-mRNA substrate (Fig. 2, top panel). Because the inhibited extract was subsequently treated with DNase I, the splicing activity can be restored upon addition of purified U1 snRNPs. Importantly, both dU1 370 and dU1 440 (Fig. 2, lanes 4 and 5, respectively), but not other Mono Q fractions lacking U1 snRNP (data not shown), were functional in this assay to complement the inactivated extract.

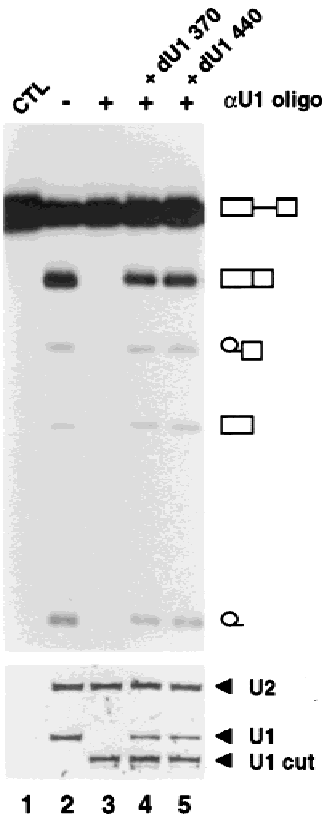

FIGURE 2. Complementation of a RNase H-inactivated splicing extract by the two types of purified *Drosophila* U1 snRNPs. In vitro splicing of Tra pre-mRNA was assayed in Kc cell nuclear extract treated with (+) or without (−) an oligonucleotide specific for U1 snRNA (αU1). After addition of purified snRNPs when indicated, ^{32}P-labeled Tra pre-mRNA was incubated at 22 °C under standard splicing conditions. The splicing products depicted on the right of the panel were analyzed on a 10% polyacrylamide/urea gel and revealed by autoradiography. The same gel was also silver stained (bottom panel) and the positions of U2, U1, and the targeted RNase H-cleaved U1 snRNAs are shown on the right of the panel. CTL: control unspliced Tra pre-mRNA.

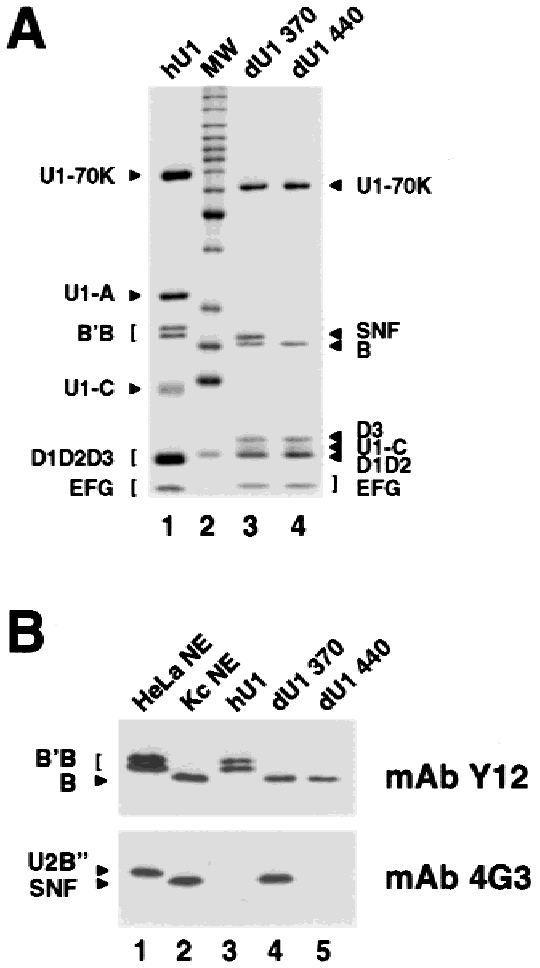

FIGURE 3. Polypeptide components of human and *Drosophila* U1 snRNPs. **A**: U1 particles purified from HeLa cells (hU1; ~1 μg) or Kc cells (dU1 370 and dU1 440; ~0.5 μg) were analyzed on a 12% SDS-PAGE and their proteins stained with Coomassie blue. The human proteins are indicated on the left of the panel and the *Drosophila* polypeptides identified in this article on the right. MW: molecular weight marker (220, 160, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, and 15 kDa from top to bottom). **B**: Western blot analysis. Proteins from the indicated purified U1 snRNPs and nuclear extracts were resolved on a 12% acrylamide/SDS gel, transferred to nitrocellulose, and probed with the monoclonal antibodies Y12 (top panel) or 4G3 (bottom panel).

One possible interpretation of these data is that both types of particles possess the determinants required for their function in splicing. Alternatively, one particle may lack one or several essential components, but a functional U1 snRNP is reconstituted in the nuclear extract. To distinguish between these possibilities, we compared the protein component of the purified particles by SDS-gel electrophoresis (Fig. 3A). The protein profiles were very similar; however, a polypeptide with an apparent molecular weight of 26 kDa was completely absent from the dU1 440 pool (Fig. 3A, compare lanes 3 and 4). Based on its size, this protein was likely the Sm B or SNF protein. To identify this protein, we tested several monoclonal antibodies against specific snRNP proteins. Immunoblot analysis with the anti-Sm monoclonal antibody Y12 indicated that the 25-kDa band present in both U1 particles was the *Drosophila* B protein (Fig. 3B, top panel, lanes 4 and 5). On the other hand, mAb 4G3 (Habets et al., 1989; Flickinger & Salz, 1994) specifically recognizes the human U2-B″ protein in HeLa nuclear extract (Fig. 3B, bottom panel, lane 1), as well as the SNF protein in both Kc cell nuclear extract and dU1 370 (Fig. 3B, bottom panel, lanes 2 and 4). Despite its presence in the dU1 370 fraction, in the anti-m_3G column supernatant, and in the dilute U1/U5 and U2-containing peaks (data not shown), SNF was not detected in the dU1 440 preparation (Fig. 3B, bottom panel, lane 5).

We next asked whether this SNF-free form of U1 particle was present in unfractionated Kc cell nuclear extracts prior to the K121 antibody and Mono Q ion exchange columns. The relative amounts of the SNF and B proteins in total U1 snRNP particle preparations,

purified from Kc cell nuclear extracts using a biotinylated 2′ *O*-methyl antisense U1 snRNA oligonucleotide, and in the dU1 370 fractions were compared by immunoblot analysis (Fig. 4A,C). The average SNF/Sm B ratio from three independent experiments was 2.3 for the dU1 370 fraction and 1.6 for total U1 snRNP particles purified from the same nuclear extracts (Fig. 4A, lanes 3 and 6; Fig. 4C). This significant difference implies that among the population of total U1 particles selected from the extracts, 25–35% were lacking the SNF protein. This observation is consistent with the results from the Mono Q fractionation, which showed that the SNF-free dU1 440 particles represented about 35% of the total U1 loaded on the column (Fig. 1B,C). Furthermore, the SNF/Sm B ratio was increased in the nuclear extract after complete depletion of U1 (from 1.6 to 2; Fig. 4A, lanes 1 and 2; Fig. 4C). Because SNF can only associate with U2 snRNP particles in the ΔU1-depleted extract, this variation further suggests the existence of an excess of snRNP-free SNF protein in Kc cell nuclear extracts. Finally, the same conclusions were drawn when U1 snRNP particles were purified from *Drosophila* Canton S embryonic nuclear extracts by 2′ *O*-methyl oligonucleotide affinity selection or by m_3G immunoaffinity chromatography followed by anion exchange chromatography. Two types of U1 particles were purified from these extracts and the SNF-free particle represented about 50% of the total U1 (Fig. 4B,C). Thus, we conclude that SNF is not necessarily associated with U1 snRNP particles in *Drosophila* nuclear extracts. Between 25 and 50% of the total U1 complexes completely lack SNF, and these species, which may possess a different activity or function in pre-mRNA processing, can be separated and individually isolated by biochemical fractionation.

To determine whether this SNF-free U1 snRNP can reassociate with free SNF in the Kc cell nuclear extract, we next performed a mixing experiment. After total depletion of all the forms of U1 snRNP from the Kc cell nuclear extract using a biotinylated 2′ *O*-methyl anti-

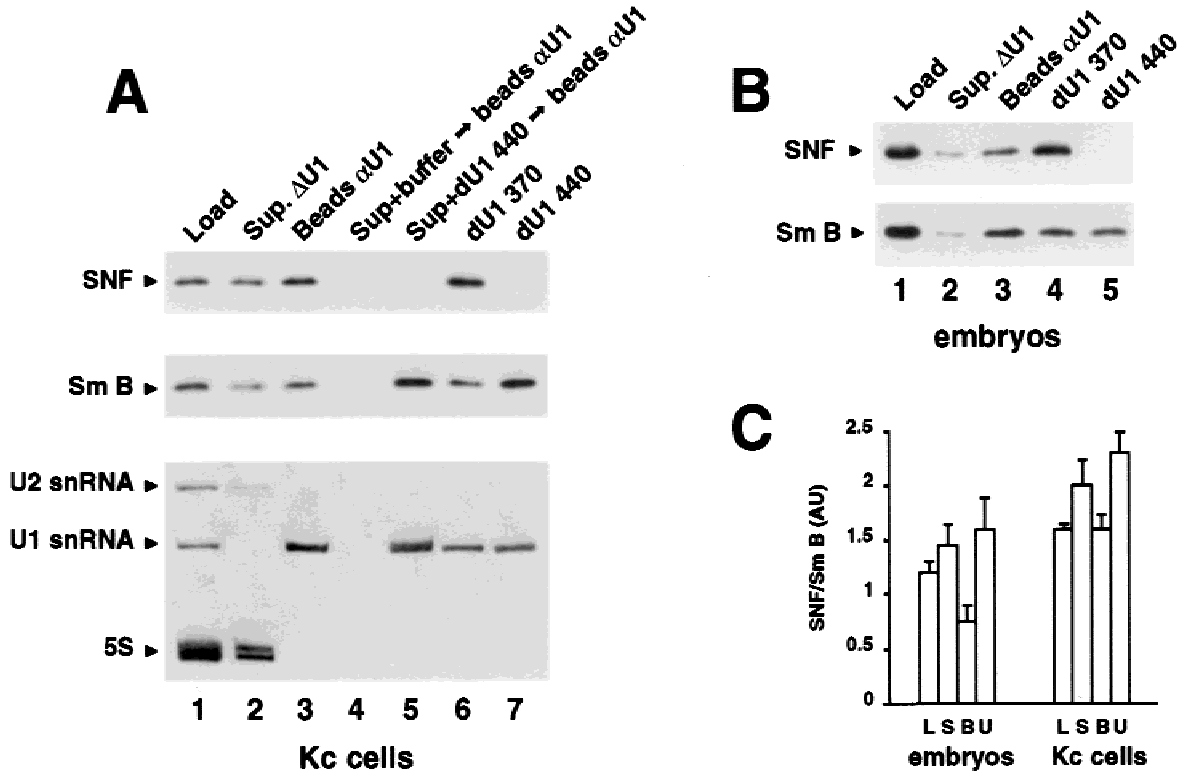

FIGURE 4. Comparison between the two populations of purified U1 particles and affinity selected total U1 snRNP. **A**: Kc cell nuclear extract (Load) was incubated under splicing conditions with a biotinylated 2′ *O*-methyl antisense U1 snRNA oligonucleotide. The total U1 snRNP affinity selected on streptavidin-agarose beads (Beads), the unbound fraction (Sup.), and the two types of U1 snRNPs purified from the Mono Q column (dU1 370 and 440) were analyzed by western blot using the 4G3 (top panel) or Y12 (middle panel) monoclonal antibodies. U1 snRNP was further affinity purified from the ΔU1 extract after incubation in presence of 20 μL of buffer or dU1 440 for 30 min at room temperature (lanes 4 and 5, respectively). The snRNAs from each fraction were analyzed on a 12% denaturing gel and silver stained (bottom panel). **B**: Comparison between U1 snRNP particles affinity selected using a biotinylated 2′ *O*-methyl antisense U1 snRNA oligonucleotide (Beads) or purified by immunoaffinity/Mono Q columns (dU1 370 and 440) from Canton S embryonic nuclear extracts. **C**: The relative SNF/Sm B abundance in nuclear extract (L), nuclear extract depleted in U1 (S), affinity selected total U1 snRNP (B) and purified dU1 370 (U) was calculated after quantification of western blots from two (embryos) or three (Kc cells) independent experiments using MacBAS, version 2.5.

sense U1 snRNA oligonucleotide, SNF was still abundant in the supernatant (Fig. 4A, top and bottom panels, compare lanes 1 and 2). However, incubation of dU1 440 in this ΔU1-depleted extract followed by a second round of U1 snRNP affinity selection did not revealed any detectable level of SNF protein (Fig. 4A, top panel, lane 5). This observation, together with the splicing complementation by dU1 440 observed in Figure 2, suggests that SNF is not required for U1 snRNP function in pre-mRNA splicing in vitro.

Both U1 snRNP-specific and Sm proteins are highly conserved between human and *Drosophila*

To determine the degree of similarity between the U1 snRNP proteins found in humans and flies, we next compared our dU1 preparations with the human U1 snRNP particle purified from HeLa cells by a similar two-step fractionation protocol (hU1). Eleven polypeptides are characteristic for hU1 (Fig. 3A, lane 1). With the exception of the Sm B/B′ proteins encoded by splice variants of the same gene, 10 corresponding *Drosophila* proteins were identified in the dU1 370 fraction: U1-70K (60 kDa), SNF (26 kDa), B (25 kDa), U1-C (15.5 kDa, see below), D3 (16.5 kDa), D2 (15 kDa), D1 (15 kDa), E (12.5 kDa), F (11 kDa), and G (10 kDa). These results are summarized in Figure 3A and Table 1.

Previous studies have shown that the *Drosophila* U1-70K, U1-A, B, D3, and F proteins possess significant sequence identities with their human counterparts (see Table 1; Fig. 5). Their presence in the purified dU1 snRNP fractions was confirmed using kinase assays (U1-70K; see below), immunoblotting with specific monoclonal antibodies (U1-A/SNF and Sm B; Figs. 3 and 4), and peptide sequencing (Sm D3 and F; Table 1; Fig. 5). Given the striking evolutionary conservation of the Sm motifs (Hermann et al., 1995; Séraphin, 1995), we next asked whether the small proteins detected in the dU1 preparations displayed substantial sequence homologies with the human D1, D2, E, and G proteins. Database searches with the human sequences revealed four equivalent *Drosophila* polypeptides denoted dD1, dD2, dE, and dG in Figure 5. Significantly, all of these proteins share a conspicuous conservation in their Sm motifs and an individual identity of 72 to 84% with their human counterparts (Table 1). As expected, microsequencing of the 15-, 12.5-, and 10-kDa protein bands only revealed peptide sequences that were specific for the Sm protein sequences retrieved from the database and corresponding to dD2, dE, and dG (Fig. 5). No peptide sequence was obtained for dD1, presumably because it comigrates with the 15-kDa dD2 protein band (Paterson et al., 1991). Furthermore, the calculated molecular weight for these *Drosophila* homologs of hD1, hD2, hE, and hG proteins correlates well with their observed mobilities in polyacrylamide-SDS gel (see Table 1).

TABLE 1. Comparison of human and *Drosophila* U1 snRNP polypeptides.

Human proteins		*Drosophila* proteins			
Name	Apparent MW (kDa)	Name	Apparent MW (kDa)	Calculated MW (kDa)	% Identity*
70K	70	70K	60	53	44 (68)
A	34	SNF	26	24.5	69 (67)
C	19	C	15.5	15.7	58 (93)
B′	27	—	—	—	—
B	26	B	25	21	63
D3	15	D3	16.5	15.6	75
D2	14.5	D2	15	13.5	84
D1	14	D1	15	13.8	78
E	12	E	12.5	11	72
F	11	F	11	9.4	78
G	9	G	10	8.5	72

*Percentage according to CLUSTALW program. In parentheses are shown the percentages of identity between the first 214 residues of the U1-70K proteins (68), between SNF and human U2-B″ (67), and between the first 62 amino acids of the U1-C proteins (93).

The lack of anti-U1-C antibodies cross-reacting with fly proteins (Paterson et al., 1991; our observations) did not allow us to directly detect this protein in the dU1 preparations. Nevertheless, the fact that U1-C is conserved and required for U1 snRNP function from yeast to human (Heinrichs et al., 1990; Tang et al., 1997), prompted us to microsequence the weak polypeptide species around 15.5 kDa (between D3 and the D1/D2 doublet; see Fig. 3A). We retrieved the peptide sequences KFYYQK and KWMEEQAQHLID, matching a unique polypeptide of 145 residues in the *Drosophila* genome sequence database (Adams et al., 2000). Significantly, this protein shares 58% identity with the human U1-C protein (Fig. 5; Table 1) and displays an extensive homology (93% identity) with its N-terminal domain required for dimerization and binding to U1 snRNP (Nelissen et al., 1991; Gunnewiek et al., 1995). Furthermore, the residues essential for hU1-C activity (cysteine 25, arginine 28, and lysine 29; Nelissen et al., 1991; Will et al., 1996) are conserved (Fig. 5). We conclude that the *Drosophila* U1 snRNP particle does contain a U1-C protein and is characterized by 3 U1-specific and 7 core proteins, each of them highly conserved with their human counterparts.

dU1 370, but not dU1 440, possesses an intrinsic kinase activity

Serine/threonine protein phosphorylation and dephosphorylation events play crucial roles during pre-mRNA splicing (Misteli, 1999) and a serine kinase activity specific for the RS domain of U1-70K and SF2/ASF proteins was shown to copurify with human U1 snRNP particles (Woppmann et al., 1993). Given the striking

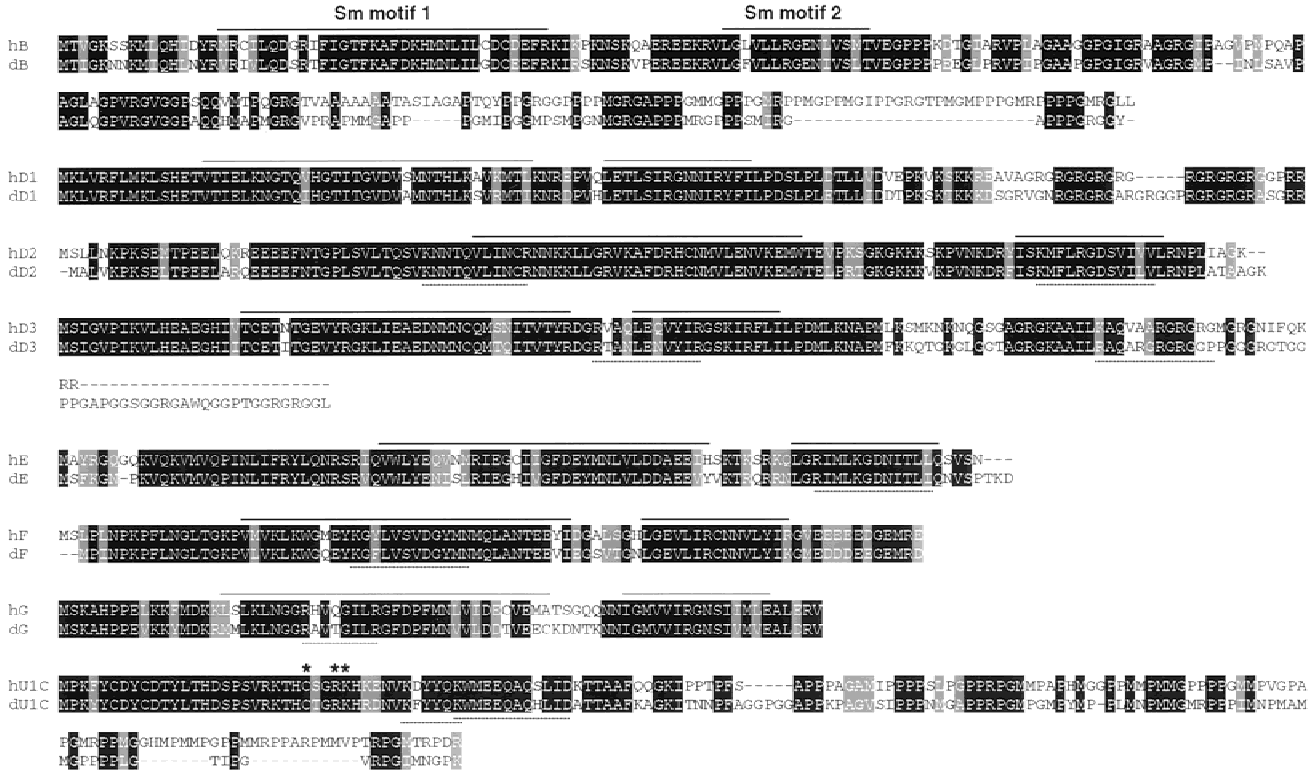

FIGURE 5. Sequence comparison of the human (h) and *Drosophila* (d) U1-C and Sm proteins. Alignments were performed using the CLUSTALW program. Identical and functionally related matches are boxed in black and gray, respectively. Sequences for the fly Sm proteins B, D3, and F were previously published (Brunet et al., 1993; Ivanov et al., 1998; Vincent et al., 1990). The other sequences were retrieved from the *Drosophila* genome database using the BLASTP program. GenBank accession numbers are: dD1: AAF49893; dD2: AAF54135; dE: AAF52576; dG: AAF48634; dU1-C: AAF55616. Percentages of identity are shown in Table 1. For each core protein alignment, the position of the conserved Sm motifs 1 and 2 (32 and 14 residues respectively; Hermann et al., 1995) is indicated by a solid line. ∗ shows the residues crucial for human U1-C activity (Nelissen et al., 1991; Will et al., 1996). The peptide sequences obtained by microsequencing of the 16.5- (dD3), 15.5- (dU1-C), 15- (dD2), 12.5- (dE), 11- (dF), and 10-kDa (dG) protein bands are underlined.

degree of similarity between the human and *Drosophila* splicing apparatus, one could expect that this activity is also conserved in flies. To determine whether the dU1 particles possess an associated kinase activity, purified U1 snRNPs were incubated with γ-^{32}P ATP and efficient ^{32}P incorporation was detected by autoradiography following SDS-gel electrophoresis. Specific phosphorylation of the respective U1-70K proteins in both hU1 and dU1 370 fractions is observed (Fig. 6, lanes 1 and 3). In contrast, no radiolabeled proteins are detected in the dU1 440 fraction (Fig.6, lane 5), suggesting that this particle lacks an associated kinase activity.

Alternatively, the 70K protein in this fraction may be in an hyperphosphorylated state, preventing any additional in vitro phosphorylation. This possibility was ruled out because a bacterially expressed, recombinant SF2/ASF protein, incubated with dU1 440 under the same conditions, was not phosphorylated (Fig. 6, lane 6). In contrast, this prototypical RS domain-containing splicing factor was an active substrate for the hU1 as well as for the dU1 370 kinases (Fig. 6, lanes 2 and 4). Moreover, kinase assays performed with other fractions purified from the Mono Q column showed only low levels of U1-70K or SF2/ASF phosphorylation (data not shown). Thus, the kinase activity preferentially associates with U1 snRNP purified from Kc cell nuclear extracts. However, the fraction of U1 depleted in SNF protein is also completely devoid of this protein kinase activity, implying that the two populations of U1 snRNPs may have different functions in pre-mRNA processing.

U1 snRNP-associated kinase substrate specificity is highly conserved between human and *Drosophila*

The above results suggest that the *Drosophila* U1 snRNP-associated kinase activity is specific for splicing factor containing arginine–serine-rich sequence motifs. To further characterize this activity, we chose to use the human prototypical SR protein SF2/ASF as a representative model recombinant substrate. Like U1-70K, this protein shows a modular structure with a C-terminal domain highly enriched in arginine-serine repeats (Fig. 7A). As shown in Figure 7B, the SF2/ASF RS domain is the specific target of dU1 370-associated

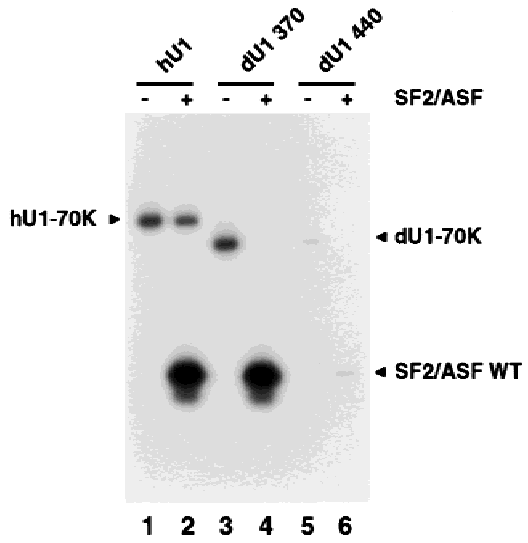

FIGURE 6. dU1 370 possesses an intrinsic kinase activity. U1 particles (250 ng) purified from HeLa cells (hU1) or Kc cells (dU1 370 and dU1 440) were incubated with (+) or without (–) an excess of recombinant SF2/ASF (200 ng; 6 pmol) purified from *E. coli* and γ-^{32}P ATP. Proteins were analyzed on a 12% SDS-PAGE and detected by autoradiography. The position and identity of the labeled proteins was determined by Coomassie blue staining of the same gel.

kinase activity. SF2/ASF RS mutant, which lacks the entire N-terminal 196 residues containing the RNA binding domain, was phosphorylated to the same extent as wild-type SF2/ASF (Fig. 7A, compare lanes 2 and 4). In contrast, no protein kinase activity was detected with a truncated protein deleted of its C-terminal region (ΔRS, Fig. 7A, lane 3). In vitro kinase assays with recombinant truncated *Drosophila* U1-70K proteins confirmed the requirement of an RS domain for an efficient phosphorylation (data not shown).

The RS domain of SF2/ASF contains 20 serine residues, most of them in a SR or RS dipeptide context. The use of progressive C-terminal deletion mutants of SF2/ASF revealed that these residues are not equivalent bona fide substrates (data not shown). As an example, a protein lacking eight repetitive RS dipeptides (SF2/ASF Δ207-224) is poorly phosphorylated by dU1 370 (Fig. 7B, lane 5), indicating that the majority of phosphorylation sites are located within this short domain. The low level of ^{32}P-labeling observed suggests that additional phosphorylation sites may be located in the most C-terminal 24 amino acids. This substrate specificity is exactly the same for the human U1-associated kinase activity (Woppmann et al., 1993; data not shown). Furthermore, replacement of all the RS repeats by either KS, GS, RG, or RT dipeptides within the SF2/ASF RS domain yielded mutant proteins that were not substrates for dU1 370-associated kinase (Fig. 7C). These results demonstrate that threonine cannot replace serine as substrate for the phosphotransferase activity and that phosphorylation of any of the

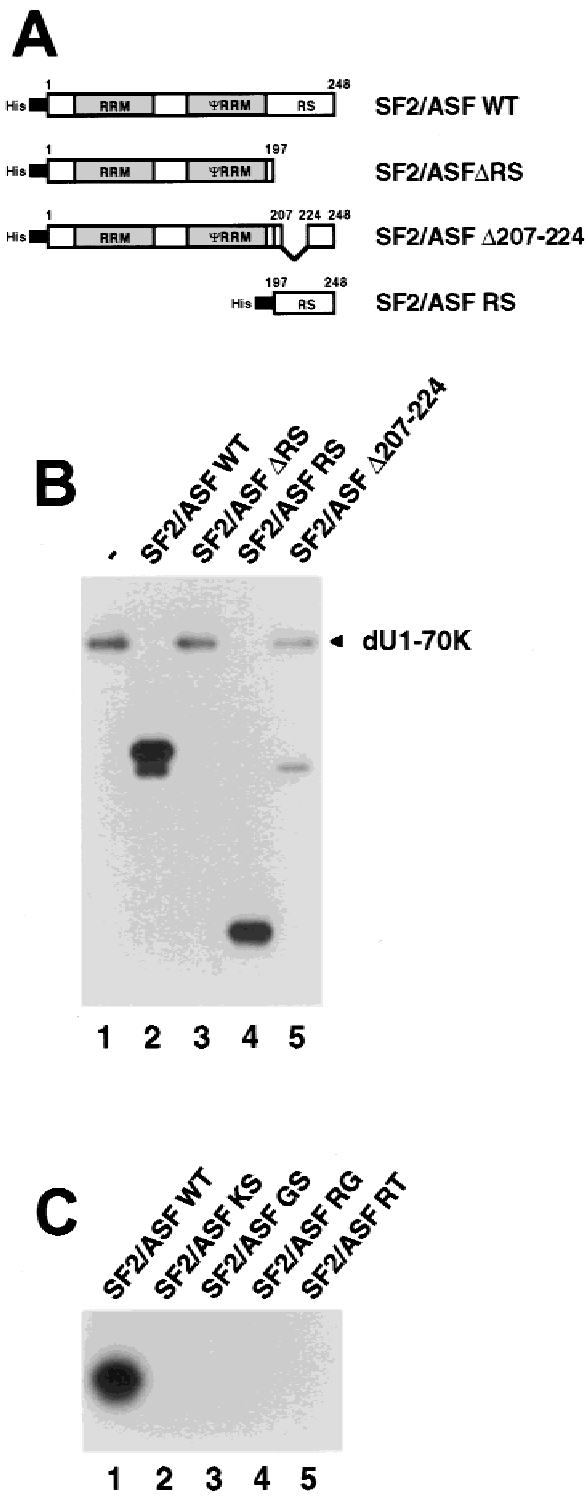

FIGURE 7. Substrate specificity of dU1 370-associated kinase activity. **A**: Schematic representation of the recombinant SF2/ASF proteins used in this study. RNA Recognition Motifs (gray), RS domains (white), and hexahistidine tags (black) are boxed. **B**: Kinase assays were performed as in Figure 6 with purified dU1 370 and 6 pmol of the indicated recombinant proteins. **C**: Same experiment with 6 pmol of recombinant mutant SF2/ASF proteins carrying specific substitution of their RS repeats: arginine to lysine (KS), arginine to glycine (GS), serine to glycine (RG) and serine to threonine (RT). Samples were analyzed on a 10% SDS-PAGE.

serine residues requires the presence of arginine in the RS dipeptides. Additionally, competition experiments with unlabeled nucleotides showed that the *Drosophila* U1 snRNP-associated kinase preferentially utilizes ATP as a phosphate donor (data not shown). We conclude that both human and fly U1 snRNP particles possess an intrinsic serine kinase activity specific for the RS domain of proteins like SF2/ASF.

We next asked whether other members of the broad family of serine–arginine-rich splicing factors were also in vitro substrates for the *Drosophila* kinase activity. Figure 8A shows that dU1 370 efficiently phosphorylates the recombinant *Drosophila* splicing regulators RSF1 and Rbp1 (Kim et al., 1992; Labourier et al., 1999) purified from *Escherichia coli* as GST fusion proteins (lanes 3 and 6). Importantly, the SR protein Rbp1 possesses a canonical RS domain with 12 RS or SR dipeptides and was phosphorylated to roughly the same extent as the human SR protein SF2/ASF (Fig. 8A, compare lanes 6 and 7). In contrast, RSF1 showed a weaker ^{32}P incorporation specific for its C-terminal domain (Fig. 8A, lanes 2–5). This domain is required for RSF1 to act as a functional SR protein antagonist (Labourier et al., 1999) and, like the SF2/ASF RS domain, contains 20 serine residues. However, only five of these residues are in a SR or RS context (16 for SF2/ASF), accounting for the low level of phosphorylation and further confirming the sharp substrate specificity of dU1 370 kinase activity.

Interestingly, the HeLa U1 snRNP particle possesses the exact same specificity, with an obvious correlation between the level of phosphorylation of the various substrates and the number of RS or SR repeats (Fig. 8, compare panels A and B). Both human and *Drosophila* recombinant SR proteins were better substrates than RSF1 for the human kinase activity. An efficient phosphorylation by hU1 of the smallest form of SF2/ASF, generated by proteolysis at its C-terminal end, was observed when the samples were resolved in a 12% acrylamide-SDS gel (see, for example, Fig. 6). Taken together, these data demonstrate that the *Drosophila* U1 snRNP-associated serine kinase activity has the same substrate specificity as the human U1 kinase. Furthermore, the fact that hU1 and dU1 370 phosphorylate both human and *Drosophila* splicing factors to the same extent suggests that the underlying mechanisms of pre-mRNA splicing regulation by phosphorylation events are extensively conserved between flies and vertebrates.

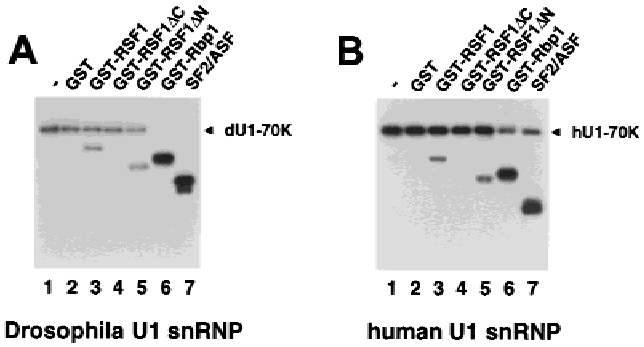

FIGURE 8. Conservation between the *Drosophila* and human U1 snRNP-associated kinase activities. **A**: Kinase assays were performed with purified dU1 370 (250 ng) and 6 pmol of the indicated hexahistidine or GST-tagged recombinant proteins. Samples were analyzed on a 12% acrylamide/SDS gel. **B**: Same experiment with U1 snRNP purified from HeLa cells (250 ng). Proteins were resolved on a 10% SDS-PAGE.

DISCUSSION

In this study, we demonstrate that two populations of functional *Drosophila* U1 snRNP particles can be individually isolated from cultured Kc cell or Canton S embryonic nuclear extracts. Although we show that the polypeptide components and activities of dU1 370 are highly conserved with the human U1 particle, our results indicate that *Drosophila* possesses a second type of U1 particle completely lacking the U1/U2-specific U1-A/U2-B″ protein SNF and the U1-associated kinase activity. A particle lacking SNF may represent a specialized form of U1 snRNP with different activities or functions in pre-mRNA processing.

Human snRNP particles are efficiently isolated from HeLa nuclear extracts by anion exchange chromatography of a mixture of anti-m_3G affinity purified snRNPs (Bach et al., 1990a). Here, we developed a purification of *Drosophila* snRNPs using a sensitive micropreparative biochemical fractionation of nuclear splicing extracts. While the *Drosophila* U snRNAs closely resemble those of human, the use of the H-20 (Paterson et al., 1991) or K121 (Krainer, 1988; this work) anti-m_3G monoclonal antibodies yielded only very low concentration snRNPs preparations. To overcome this problem, we optimized our method, and most of the U1 snRNP present in concentrated Kc cell nuclear extracts was recovered by lowering the glycerol and KCl concentrations during the batch adsorption and elution steps. Moreover, the use of a small (100 μL) Mono Q column combined with a microfraction collector improved the purification, both quantitatively and qualitatively. Surprisingly, the elution profile of the major U snRNPs from the Mono Q column was very similar to that observed with snRNPs purified from HeLa cells (Bach et al., 1990a). Several species eluted at the same salt concentration as the human snRNPs, suggesting similarities between the *Drosophila* and human particles.

Our data show that the *Drosophila* U1 snRNP particle, like the human one, contains seven Sm proteins (B, D1, D2, D3, E, F, G), three U1-specific proteins (70K, C, SNF), and an associated serine protein kinase activity specific for splicing factors containing arginine-serine repeats. After protein sequencing, we were able to retrieve from the *Drosophila* genome sequence data-

base unique protein sequences displaying high identity with the corresponding human U1-C, D1, D2, E, and G proteins. The extensive conservation within the functional domains of these proteins (Sm motifs 1 and 2 for the core proteins, dimerization and U1 snRNP binding domain for U1-C protein) and the correlation between the sizes observed in polyacrylamide/SDS gels and the theoretical molecular weights strongly suggest that these proteins identified by conceptual translation are the true *Drosophila* homologs of the human U1-C, D1, D2, E, and G proteins. Furthermore, cDNA clones derived from embryonic, larval, or pupal libraries and coding for the U1-C and D2 proteins were found by searching the BDGP EST database (clone ID LP03267, LD03002), implying that at least two out of the five new proteins identified in this article are efficiently expressed in flies.

In vitro kinase assays using various bacterially expressed recombinant splicing factors revealed that the dU1 370-associated kinase activity has the same substrate specify as the human U1 snRNP particle (Woppmann et al., 1993). We also noticed that the competition with an excess of recombinant SR protein does not completely abolish the phosphorylation of the human U1-70K protein (Fig. 6, lane 2; Fig. 8B, lanes 6 and 7). A careful examination of human U1-70K RS domain reveals 13 RS dipeptides and more identity with the RS domain of SR proteins, in comparison to that of the *Drosophila* U1-70K protein, which contains only 9 RS dipeptides. Thus, these observations may reflect a preferred specificity of both human and fly kinases for canonical RS domains. To our knowledge, the mammalian SR protein kinases SRPK1, associated with U1 snRNP, DNA topoisomerase I, and Clk/Sty all possess different substrate specificities (Woppmann et al., 1993; Colwill et al., 1996; Labourier et al., 1998). In flies, only the Doa SR protein kinase, which is required at multiple stages of development, has been reported (Yun et al., 1994). However, preliminary studies indicate that the *Drosophila* DNA topoisomerase I possesses a protein kinase activity (J. Tazi, pers. comm.) and three homologs of human SRPKs have been identified in the *Drosophila* genome (Mount & Salz, 2000). Taken together with the data presented in this article, these observations suggest that strong evolutionary constraints on the splicing machinery have led to a striking conservation of the fundamental phosphorylation/dephosphorylation mechanisms involved in pre-mRNA splicing regulation in both flies and vertebrates.

Among the snRNPs resolved on the Mono Q column, two fractions contained highly purified U1 snRNP particles. Immunoblot analysis of these fractions demonstrated that the SNF protein is totally depleted from the *Drosophila* U1 snRNP fraction eluting at 440 mM KCl. Although the SNF protein binds U1 and U2 snRNA in vitro and is an integral part of U1 and U2 snRNPs in vivo (Polycarpou-Schwarz et al., 1996), genetic analyses evidenced a snRNP-independent function for SNF in Sex-lethal splicing autoregulation (Cline et al., 1999). Consistent with the idea that there is extrinsic SNF protein, SNF was strongly detected in the anti-m_3G column supernatant lacking 95 and 50% of U1 and U2 snRNPs, respectively, as well as in a nuclear extract treated with a biotinylated 2′ *O*-methyl antisense U1 snRNA oligonucleotide. However, an integral U1 particle could not be reconstituted after incubation of dU1 440 in these depleted extracts, suggesting that a large part of the free SNF may, in fact, be held in other protein complexes. Such non-snRNP protein complexes containing the human U1-A protein have previously been reported (O'Connor et al., 1997).

Both dU1 370 and 440 are able to complement a splicing extract that had been inactivated by oligonucleotide-directed RNase H treatment. Consistent with the observations that U1-A is not required for the formation of early splicing complexes in human (Jamison et al., 1995; Will et al., 1996) or for yeast viability (Liao et al., 1993), this result suggests that SNF is not essential to restore in vitro splicing in *Drosophila* extracts. Unfortunately, all our attempts to complement with purified dU1 370 or 440 particles a nuclear extract depleted using an antisense U1 snRNA oligonucleotide failed. Control experiments with nonspecific biotinylated oligonucleotides revealed that both Kc cell and embryonic nuclear extract could not resume in vitro pre-mRNA splicing after incubation with the streptavidine beads, suggesting that other constitutive factor(s) are removed or inactivated during the treatment. Therefore, we cannot rule out the possibility that dU1 440 could complement the RNase H-inactivated splicing extract only after acquiring SNF by exchange with the oligonucleotide-cleaved U1 snRNP. Furthermore, only small amounts of functional U1 snRNP may suffice for the complete restoration of splicing in this type of assay. The development of assays directly reflecting the amount of U1 snRNP bound to the 5′ splice site (filter binding, gel mobility shift, or E complex assembly) may reveal qualitative differences between the two U1 particles described here and establish if SNF is required for the early steps of pre-mRNA splicing in *Drosophila*.

Approximately 25 to 50% of the total U1 snRNP purified from Kc cell or embryonic nuclear extracts was reproducibly found in the dU1 440 fraction, implying that these particles do not represent intermediate forms of U1 snRNP produced during the spliceosome cycle. The fractionation was performed at 4 °C and we did not detect other fractions containing U1 particles partially lacking one or several U1-specific proteins, as has been reported for human snRNPs isolated from a Mono Q column at room temperature (Bach et al., 1990b). Moreover, the quantification of the SNF and Sm B proteins relative abundance in the dU1 370 fractions as well as in the various unfractionated nuclear extracts and in

the total U1 snRNP particles affinity selected from these extracts confirmed that the isolation of two U1 snRNP populations was not an artifact of the purification procedure. These observations suggest that the dU1 440 fraction contains only stable U1 snRNP species that may possess a relevant biological function(s). For example, mammalian U1 snRNP was shown to coordinate the pre-mRNA splicing and polyadenylation process (Wassarman & Steitz, 1993) and to inhibit polyadenylation by direct interaction with the C-terminal domain of poly(A) polymerase (PAP; Gunderson et al., 1998). However, this domain appears to be poorly conserved within the putative PAP homolog identified by the *Drosophila* genome project (GenBank accession number AAF57528). Similarly, the fly U1-70K protein shares 68% identity with the RRM-type RNA binding domain of human U1-70K but only 21% with its C-terminal domain required for PAP inhibition (Gunderson et al., 1998). It is possible that the *Drosophila* splicing apparatus has evolved to produce a specialized U1 particle involved in the regulation of pre-mRNA 3′ end processing. Such a particle may not require the SNF protein and the associated kinase activity to specifically link the splicing and polyadenylation machineries.

Drosophila is an ideal model metazoan in which to study the mechanism and regulation of pre-mRNA processing. Genetic approaches should further facilitate the study of snRNP particles that may have other distinct functions in addition to their essential constitutive role in pre-mRNA splicing. The characterization of *Drosophila* snRNP components using biochemical studies, combined with the development of more sophisticated genetic tools, such as homologous gene targeting (Rong & Golic, 2000), will provide new insights into how snRNP particles are assembled and function in vivo.

MATERIALS AND METHODS

Cell growth, fly culture, and preparation of nuclear extracts

The *Drosophila* Kc cell line was grown in suspension at 25 °C in standard D22 medium, supplemented with 50 μg/mL streptomycin and penicillin (Sigma). Cells were split to $\sim 2.5 \times 10^6$ cells/mL daily and harvested at 4 to 5×10^6 cells/mL. Nuclear extracts were prepared essentially as described (Dignam et al., 1983) except that the packed nuclei were resuspended in 2.5 mL buffer C (Dignam et al., 1983) per 10^{10} cells. After dialysis against buffer D (Dignam et al., 1983), the KCl and protein concentrations ($\sim$0.11 M and $\sim$12 mg/mL, respectively) and the ability to support in vitro splicing were checked for each independent nuclear extract.

Canton S *Drosophila* were crossed on standard medium at 25 °C and embryos were collected for 12 h. Nuclear extracts were prepared according to Heberlein and Tjian (1988) and dialyzed against buffer D. The final protein concentration was $\sim$20 mg/mL.

snRNP purification

Affinity purification of 2,2,7-trimethylguanosine-capped snRNPs was performed using $\sim$5 mg of mAb K121 (Krainer, 1988) covalently bound to 1 mL of GammaBind Plus sepharose resin (Pharmacia Biotech) using dimethylpimelimidate (Harlow & Lane, 1988). Nuclear extracts from a 16-L Kc cells culture($\sim$19 mL) or from 40 g of embryos ($\sim$10 mL) were diluted to bring the final KCl and glycerol concentrations to 0.2 M and 5%, respectively, and incubated for 12 h at 4 °C with the beads already equilibrated in buffer A (20 mM HEPES-KOH, pH 7.6, 0.2 mM EDTA, 0.5 mM DTT, 0.2 M KCl, 5% glycerol). After extensive washes with buffer A at 0.3 M KCl, the antibody-bound snRNPs were desorbed from the beads by two successive elutions with 2 mL buffer A containing 0.1 M KCl and 20 mM 7-methylguanosine for 4 h at 4 °C. The eluates were pooled and 1 vol of buffer B (20 mM Tris-HCl, pH 7, 1.5 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM PMSF) was added to reduce the KCl concentration.

Using the SMART system (Pharmacia Biotech), the snRNP mixture was then applied to a 100-μL Mono Q PC 1.6/5 column (Pharmacia Biotech) equilibrated in buffer B containing 50 mM KCl. The column was washed with 10 column volumes of buffer B 50 mM KCl and the snRNPs were fractionated by elution with a linear KCl gradient in buffer B (0.05–2 M KCl; see Fig. 1B). All steps were carried out at 4 °C with a flow rate of 0.1 mL/min. The 100-μL fractions containing the species of interest were pooled separately, dialyzed using Spectro/Por CE membranes (5 kDa MWCO; Spectrum) against buffer D containing 5% glycerol and 1.5 mM $MgCl_2$, quickly frozen in liquid nitrogen, and stored at -80 °C.

Human U1 snRNP purified from HeLa cells (4 mg/mL) was generously provided by Prof. R. Lührmann, Germany.

RNA and protein analysis

RNA samples were treated with proteinase K in 1% SDS and extracted with phenol and phenol/chloroform. RNAs recovered by ethanol precipitation were fractionated through 10 or 12% polyacrylamide/urea gels and visualized by silver staining, ethidium bromide staining, or autoradiography (Fig. 2).

Proteins from nuclear extracts or purified snRNP particles were resolved on 10 or 12% polyacrylamide/SDS gels polymerized in the presence of very high TEMED concentration (0.07%). Proteins were detected by staining with Coomassie brilliant blue R-250 (and autoradiography when required) or by immunoblot with the monoclonal antibodies 4G3 (Habets et al., 1989; Flickinger & Salz, 1994) and Y12 (Lerner et al., 1981). Electroblotting was performed in 10 mM CAPS transfer buffer, pH 11, containing 10% methanol. Immunoreactive proteins were detected using horseradish peroxidase-conjugated secondary antibodies (Bio-Rad) and the ECL reagent (Amersham).

In vitro splicing and kinase assays

Radiolabeled Tra pre-mRNA was synthesized by in vitro transcription in the presence of 20 U of T7 RNA polymerase (Promega) and 5 μM α-^{32}P UTP (400 Ci/mmol; Amersham) in 25 μL reactions according to the manufacturer. Transcripts were purified by polyacrylamide/urea gels and quantitated by Cerenkov counting.

In vitro splicing reactions were performed at 22 °C for 90 min in a final volume of 20 μL containing 6 μL Kc cell nuclear extract, 50 fmol-labeled Tra pre-mRNA, 2% polyvinyl alcohol, 3 mM ATP, 20 mM creatine phosphate, 3.5 mM $MgCl_2$, 20 mM HEPES and 20 U RNase inhibitor (Promega). When required, samples were first incubated at 22 °C for 60 min with the oligodeoxynucleotide 5′-TACGCCAGGTAAGT-3′ (10 μM) complementary for bases 3–16 of *Drosophila* U1 snRNA. All the samples were treated for 30 min with 1 U of RQ1 RNase-free DNase I (Promega). Purified U1 snRNP particles (3 μL; 150 ng) or 3 μL of modified buffer D (5% glycerol, 1.5 mM $MgCl_2$) and the labeled pre-mRNA were then subsequently added.

In vitro kinase assays were carried out essentially as described (Labourier et al., 1998) with 250 ng of purified U1 snRNP (about 1 pmol) and 6 pmol of the various recombinant proteins purified from *E. coli* (Labourier et al., 1998, 1999). For these assays, human U1 snRNP was diluted in buffer D containing 5% glycerol and 1.5 mM $MgCl_2$ to bring the final protein concentration to 50 μg/mL.

Affinity selection of U1 snRNP

Streptavidin-agarose beads (Sigma) were blocked by rotating with an equal volume of 20 mM HEPES-KOH, pH 7.6, 0.3 M KCl, 0.01% NP40, 0.1 mg/mL glycogen, 1 mg/mL BSA, and 0.1 mg/mL yeast tRNA for 15 min at 4 °C. The beads were then washed three times with an equal volume of 20 mM HEPES-KOH, pH 7.6, 0.3 M KCl, 0.01% NP40.

Eighty microliters of Kc cell nuclear extract or 50 μL of embryonic nuclear extract were precleared by incubation with 50 pmol of a nonspecific, biotinylated, 2′ *O*-methyl RNA oligonucleotide (antisense Ftz 3′ exon 5′(bio-dT)$_4$-GUAGCGG GUGUACGUCUG) in a 100-μL reaction containing 20 mM HEPES-KOH, pH 7.6, 0.3 M KCl, 3 mM ATP, 5 mM creatine phosphate, and 4 mM $MgCl_2$. The reactions were incubated for 30 min at room temperature and spun at 13,000 × *g* for 3 min. The clarified supernatants were added to a 100 μL packed volume of blocked streptavidin-agarose beads and rotated for 30 min at 4 °C. The streptavidin-agarose beads were spun out and 50 pmol of a biotinylated, 2′ *O*-methyl antisense U1 snRNA oligonucleotide (5′GCCAGGUAAGUAU-(bio-dT)$_4$) was added to the precleared extracts. After incubation with the oligonucleotide for 1 h at room temperature, extracts were spun and incubated with 100 μL blocked streptavidin-agarose beads as described above. The supernatants were discarded and the beads were washed 3 times with 1 mL of 20 mM HEPES-KOH pH 7.6, 0.3 M KCl, and 0.1% NP40. Each wash was rotated for 5 min at 4 °C. During the final wash, the resin was divided in half and half was processed as described above for analysis of RNA. The other half was rotated in 100 μL of 8 M urea for 30 min at room temperature. Proteins in the supernatant were precipitated with 20% trichloroacetic acid, washed with cold acetone, resuspended in SDS sample buffer, and analyzed by immunoblotting.

Parallel control reactions were performed without the addition of the antisense U1 snRNA oligonucleotide. RNA synthesis and purification were carried out according to Lamond and Sproat (1994).

Peptide sequencing and biocomputing

Purified dU1 370 (~10 μg) was resolved on a 12.5% SDS-PAGE, transferred to nitrocellulose and each protein band was treated overnight with trypsin (Fernandez et al., 1992). The resulting peptides were separated by reversed phase HPLC (Applied Biosystem 172A) and sequenced by Edman degradation (Applied Biosystem 492; Fernandez et al., 1994).

Searches between a human query protein sequence and *Drosophila* protein databases were made using the BLASTP 2.0 program (National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov/blast). Sequence alignments were produced with CLUSTALW (European Bioinformatics Institute; http://www2.ebi.ac.uk/clustalw) and shaded using the BOXSHADE server (European Molecular Biology network; http://www.ch.embnet.org/software/BOX_form.html). Theoretical molecular weights were calculated with the ProtParam tool (Swiss Institute of Bioinformatics; http://expasy.cbr.nrc.ca/tools/protparam.html).

ACKNOWLEDGMENTS

We are grateful to R. Lührmann for providing the purified human U1 snRNP and to J. Tazi for the various recombinant splicing factors used in this study. We thank A. Krainer, J. Steitz, and T. Cline for the monoclonal antibodies K121, Y12, and 4G3, respectively. EL also acknowledges K. Collins and T. Cline for critical review of the manuscript as well as S. Zhou for help in peptide sequencing. This work was funded by a National Institutes of Health Grant R01HD28063 to DCR. EL was supported by a Philippe Foundation grant and a postdoctoral research fellowship from the Association pour la Recherche sur le Cancer (ARC).

Received July 12, 2000; returned for revision August 15, 2000; revised manuscript received December 21, 2000

REFERENCES

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor Miklos GL, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RD, Scheeler F, Shen H, Shue BC, Siden-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, Woodage T, Worley KC,

Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM, Venter JC. 2000. The genome sequence of *Drosophila melanogaster*. *Science 287*:2185–2195.

Bach M, Bringmann P, Lührmann R. 1990a. Purification of small nuclear ribonucleoprotein particles with antibodies against modified nucleosides of small nuclear RNAs. *Methods Enzymol 181*: 232–257.

Bach M, Krol A, Lührmann R. 1990b. Structure-probing of U1 snRNPs gradually depleted of the U1-specific proteins A, C and 70k. Evidence that A interacts differentially with developmentally regulated mouse U1 snRNA variants. *Nucleic Acids Res 18*:449–457.

Branlant C, Krol A, Ebel JP, Lazar E, Haendler B, Jacob M. 1982. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO J 1*:1259–1265.

Bringmann P, Rinke J, Appel B, Reuter R, Lührmann R. 1983. Purification of snRNPs U1, U2, U4, U5, and U6 with 2,2,7-trimethylguanosine-specific antibody and definition of their constituent proteins reacting with anti-Sm and anti-(U1)RNP antisera. *EMBO J 2*:1129–1135.

Brunet C, Quan T, Craft J. 1993. Comparison of the *Drosophila melanogaster*, human and murine Sm B cDNAs: Evolutionary conservation. *Gene 124*:269–273.

Cline TW, Rudner DZ, Barbash DA, Bell M, Vutien R. 1999. Functioning of the *Drosophila* integral U1/U2 protein Snf independent of U1 and U2 small nuclear ribonucleoprotein particles is revealed by snf(+) gene dose effects. *Proc Natl Acad Sci USA 96*:14451–14458.

Colwill K, Feng LL, Yeakley JM, Gish GD, Caceres JF, Pawson T, Fu XD. 1996. SRPK1 and Clk/Sty protein kinases show distinct substrate specificities for serine/arginine-rich splicing factors. *J Biol Chem 271*:24569–24575.

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

Fernandez J, Andrews L, Mische SM. 1994. An improved procedure for enzymatic digestion of polyvinylidene difluoride-bound proteins for internal sequence analysis. *Anal Biochem 218*:112–117.

Fernandez J, DeMott M, Atherton D, Mische SM. 1992. Internal protein sequence analysis: Enzymatic digestion for less than 10 micrograms of protein bound to polyvinylidene difluoride or nitrocellulose membranes. *Anal Biochem 201*:255–264.

Fischer U, Lührmann R. 1990. An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus. *Science 249*:786–790.

Fischer U, Sumpter V, Sekine M, Satoh T, Lührmann R. 1993. Nucleocytoplasmic transport of U snRNPs: Definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *EMBO J 12*:573–583.

Flickinger TW, Salz HK. 1994. The *Drosophila* sex determination gene *snf* encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein. *Genes & Dev 8*:914–925.

Gunderson SI, Polycarpou-Schwarz M, Mattaj IW. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell 1*:255–264.

Gunnewiek JM, van Aarssen Y, Wassenaar R, Legrain P, van Venrooij WJ, Nelissen RL. 1995. Homodimerization of the human U1 snRNP-specific protein C. *Nucleic Acids Res 23*:4864–4871.

Habets WJ, Hoet MH, De Jong BA, Van der Kemp A, Van Venrooij WJ. 1989. Mapping of B cell epitopes on small nuclear ribonucleoproteins that react with human autoantibodies as well as with experimentally-induced mouse monoclonal antibodies. *J Immunol 143*:2560–2566.

Hamm J, Darzynkiewicz E, Tahara SM, Mattaj IW. 1990. The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell 62*:569–577.

Harlow E, Lane DP. 1988. *Antibodies: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 522–523.

Harper DS, Fresco LD, Keene JD. 1992. RNA binding specificity of a *Drosophila* snRNP protein that shares sequence homology with mammalian U1-A and U2-B″ proteins. *Nucleic Acids Res 20*:3645–3650.

Heberlein U, Tjian R. 1988. Temporal pattern of alcohol dehydrogenase gene transcription reproduced by *Drosophila* stage-specific embryonic extracts. *Nature 331*:410–415.

Heinrichs V, Bach M, Winkelmann G, Lührmann R. 1990. U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5′ splice site. *Science 247*:69–72.

Hermann H, Fabrizio P, Raker VA, Foulaki K, Hornig H, Brahms H, Lührmann R. 1995. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein–protein interactions. *EMBO J 14*:2076–2088.

Ivanov IP, Simin K, Letsou A, Atkins JF, Gesteland RF. 1998. The *Drosophila* gene for antizyme requires ribosomal frameshifting for expression and contains an intronic gene for snRNP Sm D3 on the opposite strand. *Mol Cell Biol 18*:1553–1561.

Jamison SF, Pasman Z, Wang J, Will C, Lührmann R, Manley JL, Garcia-Blanco MA. 1995. U1 snRNP-ASF/SF2 interaction and 5′ splice site recognition: Characterization of required elements. *Nucleic Acids Res 23*:3260–3267.

Kim YJ, Zuo P, Manley JL, Baker BS. 1992. The *Drosophila* RNA-binding protein RBP1 is localized to transcriptionally active sites of chromosomes and shows a functional similarity to human splicing factor ASF/SF2. *Genes & Dev 6*:2569–2579.

Kohtz JD, Jamison SF, Will CL, Zuo P, Lührmann R, Garcia-Blanco MA, Manley JL. 1994. Protein–protein interactions and 5′-splice-site recognition in mammalian mRNA precursors. *Nature 368*: 119–124.

Krainer AR. 1988. Pre-mRNA splicing by complementation with purified human U1, U2, U4/U6 and U5 snRNPs. *Nucleic Acids Res 16*:9415–9429.

Krainer AR, Conway GC, Kozak D. 1990. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes & Dev 4*:1158–1171.

Labourier E, Bourbon HM, Gallouzi IE, Fostier M, Allemand E, Tazi J. 1999. Antagonism between RSF1 and SR proteins for both splice-site recognition in vitro and *Drosophila* development. *Genes & Dev 13*:740–753.

Labourier E, Rossi F, Gallouzi IE, Allemand E, Divita G, Tazi J. 1998. Interaction between the N-terminal domain of human DNA topoisomerase I and the arginine-serine domain of its substrate determines phosphorylation of SF2/ASF splicing factor. *Nucleic Acids Res 26*:2955–2962.

Lamond AI, Sproat BS. 1994. Isolation and characterization of ribonucleoprotein complexes. In: Higgins SJ, Hames BD, eds., *RNA processing: A practical approach*. New York: Oxford University Press. pp 103–140.

Lerner EA, Lerner MR, Janeway CA Jr, Steitz JA. 1981. Monoclonal antibodies to nucleic acid-containing cellular constituents: Probes for molecular biology and autoimmune disease. *Proc Natl Acad Sci USA 78*:2737–2741.

Liao XC, Tang J, Rosbash M. 1993. An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: Implications for snRNP protein function in pre-mRNA splicing. *Genes & Dev 7*:419–428.

Mancebo R, Lo PC, Mount SM. 1990. Structure and expression of the *Drosophila* melanogaster gene for the U1 small nuclear ribonucleoprotein particle 70K protein. *Mol Cell Biol 10*:2492–2502.

Mattaj IW. 1986. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell 46*:905–911.

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

Misteli T. 1999. RNA splicing: What has phosphorylation got to do with it? *Curr Biol 9*:198–200.

Mount SM, Salz HK. 2000. Pre-messenger RNA processing factors in the *Drosophila* genome. *J Cell Biol 150*:F37–44.

Nelissen RL, Heinrichs V, Habets WJ, Simons F, Lührmann R, van Venrooij WJ. 1991. Zinc finger-like structure in U1-specific protein C is essential for specific binding to U1 snRNP. *Nucleic Acids Res 19*:449–454.

Nelissen RL, Will CL, van Venrooij WJ, Lührmann R. 1994. The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. *EMBO J 13*:4113–4125.

O'Connor JP, Alwine JC, Lutz CS. 1997. Identification of a novel, non-snRNP protein complex containing U1A protein. *RNA 3*:1444–1455.

Paterson T, Beggs JD, Finnegan DJ, Lührmann R. 1991. Polypeptide components of *Drosophila* small nuclear ribonucleoprotein particles. *Nucleic Acids Res 19*:5877–5882.

Patton JR, Habets W, van Venrooij WJ, Pederson T. 1989. U1 small nuclear ribonucleoprotein particle-specific proteins interact with the first and second stem-loops of U1 RNA, with the A protein binding directly to the RNA independently of the 70K and Sm proteins. *Mol Cell Biol 9*:3360–3368.

Polycarpou-Schwarz M, Gunderson SI, Kandels-Lewis S, Séraphin B, Mattaj IW. 1996. *Drosophila* SNF/D25 combines the functions of the two snRNP proteins U1A and U2B′ that are encoded separately in human, potato, and yeast. *RNA 2*:11–23.

Rio DC. 1988. Accurate and efficient pre-mRNA splicing in *Drosophila* cell-free extracts. *Proc Natl Acad Sci USA 85*:2904–2908.

Rong YS, Golic KG. 2000. Gene targeting by homologous recombination in *Drosophila*. *Science 288*:2013–2018.

Scherly D, Boelens W, van Venrooij WJ, Dathan NA, Hamm J, Mattaj IW. 1989. Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. *EMBO J 8*:4163–4170.

Séraphin B. 1995. Sm and Sm-like proteins belong to a large family: Identification of proteins of the U6 as well as the U1, U2, U4, and U5 snRNPs. *EMBO J 14*:2089–2098.

Sharp PA. 1994. Split genes and RNA splicing. *Cell 77*:805–815.

Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell 92*:315–326.

Tang J, Abovich N, Fleming ML, Séraphin B, Rosbash M. 1997. Identification and characterization of a yeast homolog of U1 snRNP-specific protein C. *EMBO J 16*:4082–4091.

Vincent WS, Goldstein ES, Allen SA. 1990. Sequence and expression of two regulated transcription units during *Drosophila melanogaster* development: Deb-A and Deb-B. *Biochim Biophys Acta 1049*:59–68.

Wassarman KM, Steitz JA. 1993. Association with terminal exons in pre-mRNAs: A new role for the U1 snRNP? *Genes & Dev 7*:647–659.

Will CL, Behrens SE, Lührmann R. 1993. Protein composition of mammalian spliceosomal snRNPs. *Mol Biol Rep 18*:121–126.

Will CL, Rumpler S, Klein Gunnewiek J, van Venrooij WJ, Lührmann R. 1996. In vitro reconstitution of mammalian U1 snRNPs active in splicing: The U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic Acids Res 24*:4614–4623.

Woppmann A, Patschinsky T, Bringmann P, Godt F, Lührmann R. 1990. Characterisation of human and murine snRNP proteins by two-dimensional gel electrophoresis and phosphopeptide analysis of U1-specific 70K protein variants. *Nucleic Acids Res 18*:4427–4438.

Woppmann A, Will CL, Kornstadt U, Zuo P, Manley JL, Lührmann R. 1993. Identification of an snRNP-associated kinase activity that phosphorylates arginine/serine rich domains typical of splicing factors. *Nucleic Acids Res 21*:2815–2822.

Yun B, Farkas R, Lee K, Rabinow L. 1994. The Doa locus encodes a member of a new protein kinase family and is essential for eye and embryonic development in *Drosophila melanogaster*. *Genes & Dev 8*:1160–1173.

Zhuang Y, Weiner AM. 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell 46*:827–835.