Nucleic Acids Research, 2001, Vol. 29, No. 18 **3841–3847**

The *Drosophila* U2 snRNP protein U2A′ has an essential function that is SNF/U2B″ independent

Alexis A. Nagengast and Helen K. Salz*

Department of Genetics, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, OH 44106-4955, USA

Received June 15, 2001; Revised and Accepted July 24, 2001

ABSTRACT

Recruitment of the U2 snRNP to the pre-mRNA is an essential step in spliceosome assembly. Although the protein components of the U2 snRNP have been identified, their individual contributions to function are poorly defined. *In vitro* studies with the *Drosophila* and human proteins suggest that two of the U2 snRNP-specific proteins, U2A′ and U2B″, function exclusively as a dimer. In *Drosophila* the presence of the U2B″ counterpart, Sans-Fille (SNF), in the U2 snRNP is dispensable for viability, suggesting that SNF is not necessary for U2 snRNP function *in vivo*. With the identification of a single U2A′-like protein in the *Drosophila* genome, we can now investigate the relationship between SNF and its putative binding partner *in vivo*. Here we show that *Drosophila* U2A′ protein interacts with SNF *in vivo* and, like its human counterpart, is U2 snRNP specific. Unexpectedly, however, we find that loss of function causes lethality, suggesting that U2A′, but not SNF, is critical for U2 snRNP function. Moreover, although we demonstrate that several domains in the SNF protein are important for the interaction with the *Drosophila* U2A′ protein, including a redundant domain at the normally dispensable C-terminus, we find that U2A′ does not require heterodimer formation for either its vital function or U2 snRNP assembly. Thus together these data demonstrate that in *Drosophila* U2A′ has an essential function that is unrelated to its role as the partner protein of SNF/U2B″.

INTRODUCTION

RNA splicing is carried out by the spliceosome, a large catalytic RNA–protein machine that consists of smaller complexes called 'small nuclear ribonucleoprotein particles' (U1, U2, U4/U6 and U5 snRNPs) and many non-snRNP proteins (1,2). The spliceosome, which assembles *de novo* on each intron, both selects the intron/exon boundaries and catalyzes the splicing reaction. Identification of the intron/exon boundaries takes place during the earliest steps in spliceosome assembly, which include recruitment of the U1 and U2 snRNPs to the pre-mRNA. Each snRNP is made up of a unique snRNA backbone and a group of proteins, some of which are unique to each snRNP. Although the identities of many of the snRNP proteins are known, their individual contributions to spliceosome function remain poorly defined.

The recent completion of the *Drosophila* genome sequence has revealed a remarkable conservation between the *Drosophila* and human versions of the known snRNP proteins (3). One notable exception is that a single *Drosophila* protein, encoded by the *sans-fille* (*snf*) gene, is the counterpart of both the U1 snRNP protein U1A and the U2 snRNP protein U2B″ (4–7). Vertebrate U1A and U2B″ proteins have been the focus of many studies because they are nearly identical in sequence yet have different RNA-binding activities: U1A binds to stem–loop II of U1 snRNA, while U2B″, when bound to its partner protein U2A′, binds to stem–loop IV of U2 snRNA (see for example 8–18). The *Drosophila* Sans-Fille (SNF) protein behaves like U1A and binds U1, but not U2 snRNA (4,5). Unlike the human U1A protein, however, and like the U2B″ protein, SNF can be induced to bind U2 snRNA by the addition of either fly extract or the human U2A′ protein (5). Thus it is likely that a protein homologous to U2A′ is necessary for the U2-specific RNA-binding activity of SNF. In previous studies, we showed that *snf* is essential for viability; a deletion of the entire open reading frame results in an embryonic lethal phenotype (6). However, a number of mutations are viable, one of which, snf^{5MER}, is particularly noteworthy because it encodes a protein that is not stably associated with U2 snRNPs (7). The fact that animals carrying this mutation are viable suggests that SNF is not critical for U2 snRNP function. Although these studies suggest that the SNF/U2A′ heterodimer will also play a non-essential role in splicing, it is not known whether the two *Drosophila* proteins are dependent on each other for snRNP incorporation and/or function *in vivo*. With the identification of a single U2A′-like protein in the *Drosophila* genome (3), we can now investigate the relationship between these two proteins *in vivo*.

Here we report that the *Drosophila* U2A′ protein physically interacts with SNF *in vivo* and, like its human counterpart, is U2 snRNP specific. Surprisingly, we demonstrate that loss of function causes lethality. Thus our studies suggest that, contrary to the situation with SNF, the presence of U2A′ within the U2 snRNP is essential. Moreover, we find that although SNF and U2A′ do form heterodimers, U2A′ does not require heterodimer formation for either its vital function or U2 snRNP assembly. Together, these unexpected results

*To whom correspondence should be addressed. Tel: +1 216 368 2879; Fax: +1 216 368 3432; Email: hks@po.cwru.edu

demonstrate that in *Drosophila* U2A′ has at least one function in the cell that is unrelated to its role as the partner protein of SNF.

MATERIALS AND METHODS

U2A′ cloning

cDNAs corresponding to the *Drosophila* U2A′ protein were identified by TBLASTN searches (19) of the Berkeley *Drosophila* Genome EST database using the human U2A′ sequence (20). EST GM03681 (from clot 5902) was purchased from Research Genetics. This cDNA was sequenced on both strands by automated sequencing (Cleveland Genomics) and was found to encode a full-length protein with extensive sequence similarity to the human U2A′ protein.

Antibodies and co-immunoprecipitation experiments

The anti-SNF antibody mAb4G3 has been described previously (6,21). Polyclonal antibodies against dU2A′ were generated by standard methods. Briefly, glutathione *S*-transferase (GST) fusion proteins containing amino acids 1–265 of the *Drosophila* U2A′ protein were expressed and purified from *Escherichia coli* and then injected into guinea pigs by Convance Research Products International (PA). The resulting polyclonal antibodies were purified following standard procedures.

For co-immunoprecipitation studies, crude protein extracts from adult animals of the appropriate genotype were prepared in NETN buffer (150 mM NaCl, 50 mM Tris, pH 7.5, 5 mM EDTA, 0.5% NP-40) as described previously (7). Immunoprecipitation, western blotting, RNA isolation from the RNA–protein complexes and northern blot analysis were also carried out as previously described (7).

Construction of the genomic rescue construct and identification of *l(2)43Ee* as the gene encoding the *Drosophila* U2A′ protein

We have named the gene encoding the fly U2A′ protein *U2A*. *U2A* was mapped to the 43E region on the right arm of chromosome 2 by both *in situ* hybridization (data not shown) and by homology to one end of the fully sequenced P1 clone DS03503 (Berkeley *Drosophila* Genome Project, available at http://www.fruitfly.org). A comparison with the published genomic sequence reveals that *U2A* has no introns.

Based on this positional information and information provided by the Berkeley *Drosophila* Genome Project and FlyBase (http://flybase.bio.indiana.edu), we placed the *U2A* gene between the *Rpt1* and *torso* genes. In addition, four essential genes, *l(2)43Ed*, *l(2)43Ee*, *l(2)43Ef* and *l(2)43Eg*, are known to map near *torso* (22). We found that $l(2)43Ed^{1}$ fails to complement the three lethal P-element insertion alleles of *Rpt1* [$l(2)05643^{05643}$, $l(2)05643^{k11110}$ and *EP(2)2153*], indicating that *l(2)43Ed* is an allele of *Rpt1*.

We identified $l(2)43Ee^{1}$ as an allele of *U2A* by genetic rescue experiments with a genomic rescue construct. To make the genomic rescue construct, a 2.8 kb genomic region, which includes 1270 bp 5′ of the dU2A′ translation start site and ends 745 bp 3′ of the protein coding sequence, was amplified by PCR from the DS03503 P1 clone (Berkeley *Drosophila* Genome Project) into the pT-Adv vector (Clontech) using the primers 5′-GAATTCGGTGAAGACAGAAGGATAGAGG-3′ and 5′-GAATTCGGATCCTAATCCAATCCTTCGCATCTCC-3′. The resulting *Xba*I–*Kpn*I fragment was inserted into the pCaSpeR4 transformation vector. Germline transformants were obtained by standard methods.

To identify the mutation in $l(2)43Ee^{1}$ responsible for the mutant phenotype, DNA from $l(2)43Ee^{1}/CyO$ heterozygous adults was prepared by PCR using primers that cover the entire open reading frame (5′-CCTATTGGCAGCATTGGAGT-3′, 5′-GCCAGCGGTCTTAGTTGAAA-3′ and 5′-AGAATGATGGAGCCCAGGTT-3′). The resulting PCR products were sequenced on both strands by automated sequencing (Cleveland Genomics). The resulting sequence was analyzed with the CONSEDRUN program to identify likely heterozygous regions, using only quality scores ≥ 20 (23,24). This method identified several nucleotide substitutions that differed from our cDNA sequence information, only one of which resulted in a change in the amino acid sequence. To distinguish the change associated with the mutation from polymorphisms, we compared the sequence we obtained with the *U2A* sequence on the *CyO* chromosome [obtained from *Df(2)NCX8*/*CyO* animals, where *Df(2)NCX8* is a chromosome that carries a small deficiency that removes *U2A*] and the sequence from $l(2)43Ed^{1}/CyO$ animals, where $l(2)43Ed^{1}$ is a mutation isolated in the same genetic screen as $l(2)43Ee^{1}$ (22). Only a T→A change, which converts the conserved tyrosine at position 29 to a stop codon, was unique to the sequence obtained from $l(2)43Ee^{1}/CyO$ animals.

Yeast two-hybrid assays

The yeast two-hybrid assays used to detect interactions between the *Drosophila* U2A′ protein and different mutant SNF proteins were performed using standard methods (25). Briefly, the *U2A* cDNA was cloned into the pACT2 expression vector (Clontech) for expression as a GAL4 DNA-activation domain fusion protein. The *snf* cDNA and its mutant derivatives were cloned into the pAS2-1 expression vector (Clontech) as a GAL4 DNA-binding domain fusion protein. The two plasmids were co-transformed into the diploid yeast strain pJ69A/SL3004 and plated onto selective media to assay for GAL4-dependent transcription of the HIS3, ADE2 and LacZ reporter constructs. In control experiments, we found that substitution of either the *U2A* or *snf* fusion constructs by empty vectors does not support growth on selective media. For quantitative β-galactosidase assays, liquid cultures of each construct were grown to log phase in SD-Leu-Trp and tested for activity as specified by the Clontech Yeast Protocols Handbook.

snf alleles

The *snf* gene is located on the X chromosome and the null mutation used in this study, snf^{d210}, is a small deletion of the entire open reading frame described in Flickinger and Salz (6). The construction and characterization of the snf^{5MER} and snf^{E21D} transgenic alleles are described in Stitzinger *et al.* (7). The $snf^{\Delta Kpn}$ and $snf^{E21D+\Delta Kpn}$ mutant minigene constructs were generated by first cutting either wild-type or mutant *snf* cDNAs with *Kpn*I, then ligating the resulting truncated cDNA between 1.8 kb of upstream genomic sequences that contain the *snf* promoter and 700 bp of downstream genomic sequence and, finally, inserting the resulting minigene into a

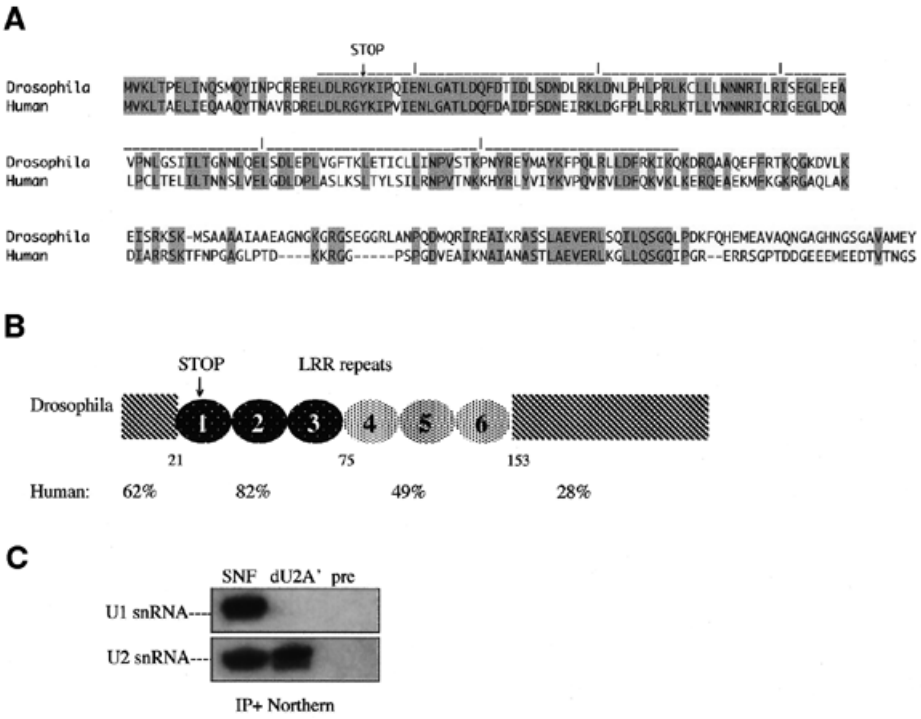

Figure 1. The *Drosophila* U2A′ ortholog is U2 snRNP specific. (**A**) Amino acid sequence alignment of the U2A′ protein from human and *Drosophila*. Identical residues are boxed in gray. The positions of the six LRRs are indicated by black lines above the sequence. The position of the *l*(2)*43Ee*^{1} mutation is indicated by an arrow. (**B**) Comparison of the human and *Drosophila* protein structures. The percent amino acid identity for the different domains is indicated. (**C**) dU2A′ is a U2 snRNP protein. snRNP incorporation was tested by immunoprecipitation of either SNF or dU2A′ from whole fly extracts followed by northern blotting to detect U1 and U2 snRNAs in the RNA extracted from the precipitated fractions.

CaSpeR transformation vector as described previously (7). Germline transformants were obtained by standard methods and each transgenic line was tested for its ability to rescue the lethality of males hemizygous for the *snf* null allele, *snf*$^{\delta 210}$, as described previously (7).

RESULTS

The *Drosophila* U2A′ protein is U2 snRNP specific

The *Drosophila* genome encodes a single protein that shares significant sequence similarity with the human U2A′ protein (3). This 265 amino acid protein has a structure that is nearly identical to the human protein, with sequence similarity extending over the entire length of the protein (Fig. 1A and B). U2A′ is composed primarily of six leucine-rich repeat motifs (LRRs); these motifs are most often found in tandem arrays and are believed to mediate protein–protein interactions (26,27). In addition to the LRR repeating units, U2A′ contains flanking N- and C-terminal non-LRR domains. While the overall conservation between the human and fly proteins is striking, it is interesting to note that the non-LRR C-terminal domain is the least conserved domain and it is the only portion of the human protein identified as dispensable (28,29). Based on this extensive sequence similarity and the functional data, presented below, we have named this *Drosophila* gene *U2A*. To avoid confusion with the human protein, we shall refer to the *Drosophila* protein as dU2A′.

As a first step in establishing that we have identified the functional counterpart of the human U2A′ protein, we raised antisera against the full-length protein (see Materials and Methods) and used this reagent to verify that dU2A′ is a snRNP protein (Fig. 1C). As in previous studies, U1 and U2 snRNAs were co-immunoprecipitated from whole cell extracts with antibodies directed against the SNF protein, demonstrating that SNF is incorporated into both U1 and U2 snRNPs. In contrast, the dU2A′-specific antibody precipitates U2 snRNA without precipitating significant amounts of U1 snRNA. Thus, we conclude that dU2A′ is a U2 snRNP-specific protein.

dU2A′ is essential for viability

To gain insight into the function of the dU2A′ protein, we sought to identify *U2A* mutations. *U2A* was first mapped to the 43E region of the second chromosome and then to a smaller region, which includes several essential genes previously described by Heitzler *et al.* (22). The embryonic lethality of one of these mutations, *l(2)43Ee*^{1}, was rescued by a 4 kb genomic transgene that contains only the *U2A* transcription unit, demonstrating that dU2A′ is encoded by the *l(2)43Ee* gene.

Sequencing of the *U2A* gene from the *l(2)43Ee*^{1} mutant chromosome revealed a single nucleotide substitution that results in conversion of Y29 to a stop codon within LRR 1 (Fig. 1A and B), indicating that *l(2)43Ee*^{1} is a complete loss-of-function mutation. Homozygous mutant animals complete embryogenesis but do not hatch (data not shown). Interestingly, a

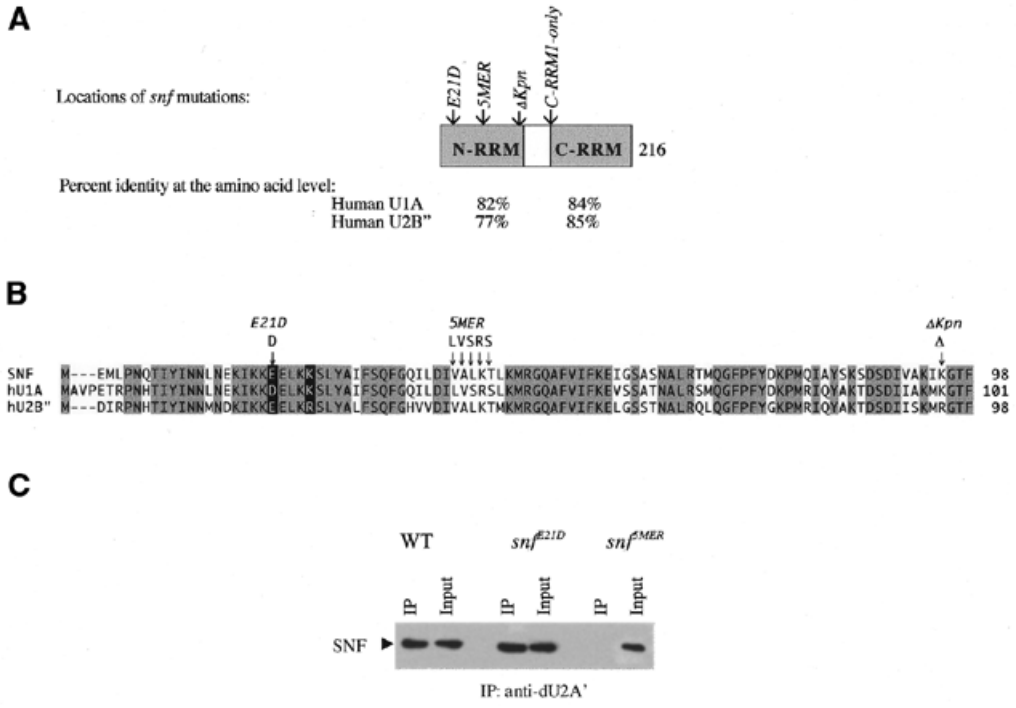

Figure 2. Impact of *snf* mutations on dU2A′/SNF heterodimer formation *in vivo*. (**A**) Schematic of the wild-type SNF protein showing the locations of mutations discussed in this study. SNF contains two RRM domains separated by a short linker region. As indicated, the sequence of the N- and C-terminal RRM motifs share significant sequence identity with both the human U1A and U2B″ proteins. The positions of the mutations used in this study are indicated by arrows. (**B**) Amino acid sequence alignment of the N-terminal RRM from SNF and human U1A and U2B″ proteins. Identical amino acids are boxed in gray. The critical residues for differential interactions with the human U2A′ protein are boxed in black. The relevant position of the point mutations and the position of the $snf^{\Delta Kpn}$ truncation are also indicated by arrows. (**C**) dU2A′ co-immunoprecipitates with wild-type SNF and SNF^{E21D}, but not SNF^{5MER}. dU2A′ was immunoprecipitated from protein extracts made from either wild-type or mutant adults followed by western blotting with an antibody against SNF protein. Mutant extracts were made from homozygous *y w* snf^{1210} animals which carried two autosomal copies of the appropriate *P*[w^+, *snf*] mutant *snf* minigene construct. Because snf^{1210} is a deletion of the entire open reading frame, all SNF protein is due to expression from the transgenic copy of *snf*. As a control for the amount of SNF protein expected, 20% of each protein extract used in the co-immunoprecipitation experiment is shown in the lane marked Input.

significant amount of maternally provided protein remains in homozygous mutant embryos as they are dying (data not shown), suggesting that by the end of embryogenesis the maternal stores of dU2A′ protein are no longer sufficient to support development. Similar late embryonic lethal phenotypes are caused by the loss of other splicing factors, including the orthlog of another U2 snRNP-specific protein, SF3a60 (30). Nevertheless, the finding that *U2A* mutations are lethal was unexpected because we previously showed that SNF, the presumed binding partner of dU2A′, is not essential for U2 snRNP function (7).

Several domains in SNF are important for the interaction with dU2A′

Although the human U2A′ protein associates with SNF *in vitro* (5), our finding that the$U2A^1$ mutation is lethal raised questions about the biological relevance of this interaction. To ascertain whether the dU2A′ protein associates with SNF *in vivo*, we tested for protein–protein interactions in *Drosophila* whole cell extracts and found that SNF and dU2A′ were co-immunoprecipitated by antibodies to either dU2A′ (Fig. 2) or SNF (data not shown). Hence, SNF and dU2A′, like their yeast and human counterparts, heterodimerize *in vivo*.

SNF, like its human counterparts, contains two RRM domains (for RNA recognition motif) separated by a short linker region. As illustrated in Figure 2, the sequence of the N-terminal RRM (amino acids 1–84) and the C-terminal RRM (amino acids 143–216) share significant similarity with both human proteins. Studies with the human proteins have shown that the N-terminal RRM of U2B″ is sufficient for heterodimer formation *in vitro* and have identified two critical U2B″-specific amino acids within this region (9,10). In agreement with the mutagenesis data, information from the crystal structure of the U2B″–U2A′–RNA ternary complex indicates that this motif contacts U2A′ directly. Interestingly, SNF has only one (E21) of these two critical amino acids; the other is identical to U1A (K25) (see Fig. 2A and B). In previous studies we tested the importance of this mixed motif by converting it to a U1A-like sequence with an E21D mutation and testing its function *in vivo*. Analysis of the SNF^{E21D} protein revealed that it retained its ability to be incorporated into both snRNPs, raising the question of whether the putative U2A′-binding site is necessary in *Drosophila* (7). Here, we extend these studies by directly asking whether the E21D mutation interferes with heterodimer formation. As illustrated in Figure 2C, dU2A′ and SNF^{E21D} could be co-immunoprecipitated from fly extracts in

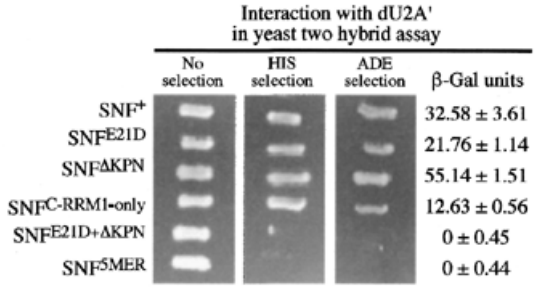

Figure 3. Several domains of SNF mediate the interaction with dU2A′ in yeast. dU2A′ was co-expressed in the yeast two-hybrid assay with wild-type SNF or different SNF mutant constructs. The positions of the mutants used in this study are indicated in Figure 2B. Positive interactions were tested by assaying the ability of the transformed yeast to grow on selective media after 3 days or by liquid culture assay for β-galactosidase activity.

which the sole source of SNF protein is the mutant *snf*^{E21D} transgene. Thus, the mixed motif in the U2A′-binding region is not essential for either heterodimer formation or U2 snRNP incorporation.

To better define the domain(s) of SNF responsible for the interaction with dU2A′, we used the yeast two-hybrid system (Fig. 3). Although both the full-length SNF^{+} and SNF^{E21D} constructs interacted with dU2A′ in yeast, we observed a decrease in reporter gene activity with the SNF^{E21D} construct relative to the wild-type construct. Thus, in the context of the full-length protein, the mixed motif in the U2A′-binding region influences but is not essential for heterodimer formation.

These studies, however, do not necessarily indicate a difference between the human and *Drosophila* proteins because the studies which identified E21 as a key residue for U2A′ binding were carried out using a truncated form of U2B″ (9,10). To explore the possibility that E21 is only essential for heterodimer formation in the absence of the C-terminal end of the protein, we designed both a 'wild-type' and a mutant version of a similarly truncated protein, called SNF$^{\Delta Kpn}$, and tested their ability to interact with dU2A′ in yeast (Fig. 3). Whereas the SNF$^{\Delta Kpn}$ construct activated the reporter constructs, an E21D mutation in this deletion construct abolished the interaction with dU2A′. Thus, we conclude that the mixed motif in the U2A′-binding region is only essential for heterodimer formation in the absence of the otherwise dispensable C-terminal end of the protein. Moreover, we find that the C-terminal RRM of SNF by itself mediates an interaction with dU2A′ in yeast (Fig. 3). Together these studies demonstrate that SNF has two redundant dU2A′ interaction domains: the previously identified motif in the N-terminal RRM and a second, previously unidentified, domain in the C-terminal RRM of the protein.

During the course of these studies, we also examined whether dU2A′ could associate with the SNF^{5MER} mutant protein, which we know from earlier studies is not stably associated with U2 snRNPs *in vivo* (7). *snf*^{5MER} has a five amino acid change in the RNA binding specificity domain of SNF that changes it from a U2B″-like to a U1A-like sequence (Fig. 2A and B). *In vivo* the *snf*^{5MER} mutation converts SNF from a bifunctional protein to a U1 snRNP-specific protein (7).

Table 1. Rescue of *snf*^{f210} lethal phenotype by representative transgenic lines

Transgene	Rescue in males	Rescue in females (n = expected no. of animals)
snf$^{E21D+\Delta KPN}$	78%	98% (n = 109)
snf$^{\Delta KPN}$	100%	93% (n = 80)

Because *snf* is an X-linked gene, only transgenic lines that carried autosomal insertions were tested. Each transgenic line was tested for its ability to rescue the lethal phenotype as follows: transgenic males were crossed to *y w snf*^{f210}*/FM7* females and the progeny scored for surviving *y w snf*^{f210} males carrying the transgene. Because *snf*^{f210} males are normally lethal, the recovery of *y w snf*^{f210} males demonstrates the ability of the transgene to rescue the lethal phenotype. The results presented here are from a second set of crosses designed to compare the ability of the transgenes to rescue males versus females. In these crosses *y w snf*^{f210} males carrying a single copy of the transgene were crossed to *y w snf*^{f210}*/FM7* females and the resulting progeny scored. The ability of the mutant transgene to rescue was assessed by comparing the number of *y w snf*^{f210}*/y w snf*^{f210} females and *y w snf*^{f210} males carrying the transgene to the number of 'expected' animals, as determined by the number of *y w snf*^{f210}*/FM7* females carrying the transgene recovered from the same cross.

Surprisingly, in co-immunoprecipitation experiments from *Drosophila* extracts in which the sole source of SNF protein is the *snf*^{5MER} mutant transgene we failed to detect an interaction between SNF^{5MER} and dU2A′ (Fig. 2). Furthermore, no interaction between SNF^{5MER} and dU2A′ was detectable in yeast (Fig. 3). Together, these studies demonstrate that the U2B″-like sequence within the 'RNA binding specificity domain' is critical for the interaction with dU2A′. Based on information from the crystal structure of the U2B″–U2A′–snRNA ternary complex, the region of the protein mutated in *snf*^{5MER} contacts the RNA without contacting the U2A′ protein (18). Thus, it is likely that the *snf*^{5MER} mutation interferes with dU2A′ binding indirectly.

snf mutations that do not associate with dU2A′ are viable

We have previously shown that animals carrying the *snf*^{5MER} mutation are viable (7). Because the *snf*^{5MER} mutation interferes with heterodimer formation, we conclude that SNF/dU2A′ heterodimer formation is not essential for viability.

Based on the results of the yeast two-hybrid experiments, the *snf*$^{E21D+\Delta KPN}$ complex mutation also eliminates heterodimer formation and therefore it is likely that animals carrying this mutation will also be viable. Using a *snf* minigene construct previously shown to replicate the high level expression of the endogenous locus (7), we generated transgenic animals bearing *snf*$^{E21D+\Delta KPN}$. Control *snf*$^{\Delta KPN}$ transgenic animals were also generated. Each line was then assayed for function by determining their ability to complement the lethal *snf* null allele *snf*^{f210}. In these assays, a single copy of the *snf*$^{E21D+\Delta KPN}$ mutant transgene (or the control *snf*$^{\Delta KPN}$ transgene) provided sufficient activity to complement the lethal phenotype of *snf*^{f210} (Table 1). Because the SNF-specific antibody recognizes only the C-terminal RRM (6), we were not able to confirm whether this mutant protein no longer interacts with dU2A′ *in vivo* or whether it is included in U2 snRNPs. However, the fact that a single copy of this mutant transgene fully rescues the *snf*^{f210} lethal phenotype indicates that elimination of both dU2A′ binding sites does not interfere with survival of the organism.

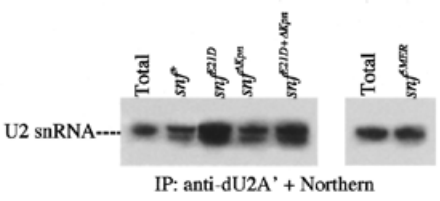

Figure 4. Assembly of dU2A′ into U2 snRNPs takes place in several different *snf* mutant backgrounds, including mutations that interfere with heterodimer formation. The ability of dU2A′ to incorporate into U2 snRNPs was tested by immunoprecipitation of dU2A′ from protein extracts made from either wild-type or homozygous mutant adults, followed by northern blotting to detect U2 snRNAs, as described in Materials and Methods. The positions of the mutants used in this study are indicated in Figure 2B. Mutant protein extracts were made from homozygous *y w* snf^{t210} animals which also carried two autosomal copies of the appropriate *P*[w^+, *snf*] mutant transgene.

Based on these studies, we conclude that dU2A′/SNF heterodimer formation is not essential for viability in *Drosophila*. The viable phenotype of these heterodimer-deficient animals therefore demonstrates that dU2A′ can accomplish its vital function without SNF.

Assembly of dU2A′ into U2 snRNPs is independent of SNF

Our finding that dU2A′ has an essential function that is independent of SNF led us to investigate whether incorporation of dU2A′ into the U2 snRNP is dependent on SNF. We therefore tested whether U2A′ could associate with U2 snRNAs in extracts made from animals homozygous for a number of different *snf* mutations, including $snf^{E21D+\Delta KPN}$ and snf^{5MER} (Fig. 4). Remarkably, dU2A′ and U2 snRNA can be co-immunoprecipitated from mutant fly extracts in which the sole source of SNF protein was provided by the $snf^{E21D+\Delta KPN}$ mutant transgene. Similarly, we find that dU2A′ associates with U2 snRNAs in snf^{5MER} mutant extracts. Together, these studies establish that dU2A′ incorporation into U2 snRNP does not depend on SNF/dU2A′ heterodimer formation. Furthermore, as an association between the mutant SNF^{5MER} protein and U2 snRNAs is not detectable under similar immunoprecipitation conditions, we conclude that dU2A′ can incorporate into U2 snRNPs in the absence of SNF.

DISCUSSION

Our *in vivo* analysis of the *Drosophila* U2A′ homolog (dU2A′) challenges the prevailing view that U2A′ is simply a cofactor of SNF/U2B″. Although SNF, like its human U2B″ counterpart, requires a U2A′-like protein to bind U2 snRNA *in vitro*, we find that *in vivo*, dU2A′ can assemble into U2 snRNP through a SNF-independent mechanism. How does dU2A′ assemble into U2 snRNP in the absence of SNF? dU2A′ is almost entirely composed of protein-interaction motifs and has no recognizable RNA-binding motifs. Nevertheless, *in vitro* studies with human U2A′ protein have shown that U2A′ is capable of binding stem–loop IV of U2 snRNA as an isolated protein, albeit weakly (28). Consistant with this possibility, the crystal structure of the U2B″–U2A′–RNA ternary complex shows that the U2A′ protein makes direct contact with U2 snRNA (18). Thus it is possible that U2A′ can bind U2 snRNA on its own and this activity is sufficient for snRNP assembly *in vivo*. However, studies with *Xenopus* oocytes have shown that stem–loop IV of U2 snRNA is itself dispensable for splicing (8,31). If, as our studies suggest, U2A′ is essential for splicing, then RNA binding to stem–loop IV is unlikely to be the only means by which this protein can assemble into U2 snRNPs. Further studies identifying the binding partners of the *Drosophila* U2A′ protein will provide insight into how dU2A′ is assembled into the appropriate snRNP.

Although dispensable for targeting dU2A′ to the appropriate snRNP, the ability of SNF and dU2A′ to form a heterodimer has been conserved in evolution. Given that several domains in the SNF protein are important for heterodimer formation, it is tempting to speculate that targeting SNF to U2 snRNP requires an interaction with dU2A′. Unfortunately, we have been unable to generate the tools necessary to test this idea directly, thus the biological role of heterodimer formation remains unknown.

The structure–function studies presented here have identified several motifs in SNF that are important for heterodimer formation, including a redundant domain in the normally dispensable C-terminal RRM. Importantly, our studies have shown that the C-terminal RRM can compensate for mutations in the N-terminal RRM dU2A′-binding motif previously identified by structure–function studies on the human proteins. The conclusion that the C-terminal RRM of SNF provides functions that are redundant to those provided by the N-terminal RRM was also reached by our earlier *in vivo* structure–function studies of SNF (7). In these studies, an N-terminal RRM mutant protein predicted to eliminate RNA binding still assembles into both snRNPs and is capable of rescuing the *snf* lethal phenotype only if the normally dispensable C-terminal RRM is present. Does the C-terminal RRM provide functions that are redundant to the N-terminal RRM in other eukaryotic species? This function is clearly absent in *Saccharomyces cerevisiae*, as the U2B″ ortholog has no C-terminal RRM (5,32). Although the human U2B″ protein does have a C-terminal RRM, its function has not been investigated. In fact, because the N-terminal RRM is both necessary and sufficient for both its RNA-binding activity and its ability to interact with U2A′, analysis of the human protein has focused almost entirely on the function of the N-terminal RRM as an isolated protein fragment (8–18). Whether the fact that the C-terminal RRM is 85% identical between the human and the fly proteins translates into conservation of function therefore remains to be determined.

A second striking feature of the studies presented here is our finding that dU2A′ is essential for viability. This result, coupled with the observation that heterodimer formation is not essential for viability in *Drosophila*, demonstrates that dU2A′ can accomplish its vital function without SNF. Is this SNF-independent function required for splicing? Evidence that U2A′ plays a key role in U2 snRNP function comes from studies of the *S.cerevisiae* U2A′ homolog, LEA1 (32). In these studies, a requirement for LEA1 in U2 snRNP function was suggested by the observation that in extracts made from strains lacking LEA1 spliceosome formation was blocked prior to addition of U2 snRNP. Interestingly, and in contrast to the situation in *Drosophila*, the assembly of LEA1 into U2 snRNP is dependent on the presence of the U2B″ counterpart, YIB9. This observation, coupled with the fact that dU2A′ is more

similar to its mammalian counterpart than to LEA1, suggests that the function of the *Drosophila* protein may not precisely parallel that of the yeast protein. It will be interesting to determine if the SNF-independent function of dU2A′ we have described here is conserved in metazoans or whether the *Drosophila* protein is unique in having acquired an independent function.

ACKNOWLEDGEMENTS

We thank the Umea and Bloomington Stock Centers for fly stocks, J. Cherry and S. Stitzinger for their excellent technical assistance and H. Lou, G. Matera and J. A. Wise for helpful comments on the manuscript. A.N. was supported by a Developmental Biology Training Grant from the NIH. This research was supported by a grant from the National Science Foundation.

REFERENCES

1. Burge,C.B., Tuschl,T. and Sharp,P.A. (1999) Splicing of precursors to mRNAs by the spliceosome. In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 525–560.
2. Reed,R. (2000) Mechanisms of fidelity during pre-mRNA splicing. *Curr. Opin. Cell Biol.*, **12**, 340–345.
3. Mount,S.M. and Salz,H.K. (2000) Pre-messenger RNA processing factors in the *Drosophila* genome. *J. Cell Biol.*, **150**, F37–F43.
4. Harper,D.S., Fresco,L.D. and Keene,J.D. (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.
5. Polycarpou-Schwarz,M., Gunderson,S.I., Kandels-Lewis,S., Seraphin,B. and Mattaj,I.W. (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.
6. Flickinger,T.W. and Salz,H.K. (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.
7. Stitzinger,S., Conrad,T., Zachlin,A. and Salz,H.K. (1999) Functional analysis of SNF, the *Drosophila* U1A/U2B″ homolog: identification of dispensable and indispensable motifs for both snRNP assembly and function. *RNA*, **5**, 1440–1450.
8. Scherly,D., Boelens,W., van Venrooij,W.J., Dathan,N.A., Hamm,J. and Mattaj,I.W. (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.
9. Scherly,D., Boelens,W., Dathan,N.A., van Venrooij,W.J. and Mattaj,I.W. (1990) Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B″ and their cognate RNAs. *Nature*, **345**, 502–506.
10. Scherly,D., Dathan,N.A., Boelens,W., van Venrooij,W.J. and Mattaj,I.W. (1990) The U2B″ RNP motif as a site of protein–protein interaction. *EMBO J.*, **9**, 3675–3681.
11. Scherly,D., Kambach,C., Boelens,W., van Venrooij,W.J. and Mattaj,I.W. (1991) Conserved amino acid residues within and outside of the amino-terminal ribonucleoprotein motif of U1A small nuclear ribonucleoprotein involved in U1 RNA binding. *J. Mol. Biol.*, **219**, 577–584.
12. Lutz-Freyermuth,C., Query,C.C. and Keene,J.D. (1990) Quantitative determination that one of two potential RNA-binding domains of the A protein component of the U1 small nuclear ribonucleoprotein complex binds with high affinity to stem-loop II of U1 RNA. *Proc. Natl Acad. Sci. USA*, **87**, 6393–6397.
13. Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P.R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515–520.
14. Bentley,R.C. and Keene,J.D. (1991) Recognition of U1 and U2 small nuclear RNAs can be altered by a 5-amino acid segment in the U2 small nuclear ribonucleoprotein particle (snRNP) B″ protein and through interactions with U2 snRNP-A′ protein. *Mol. Cell. Biol.*, **11**, 1829–1839.
15. Hall,K.B. and Stump,W.T. (1992) Interaction of N-terminal domain of U1A protein with an RNA stem–loop. *Nucleic Acids Res.*, **20**, 4283–4290.
16. Rimmele,M.E. and Belasco,J.G. (1998) Target discrimination by RNA-binding proteins: role of the ancillary protein U2A′ and a critical leucine residue in differentiating the RNA-binding specificity of spliceosomal proteins U1A and U2B″. *RNA*, **4**, 1386–1396.
17. Oubridge,C., Ito,N., Evans,P.R., Teo,C.-H. and Nagai,K. (1994) Crystal structure at the 1.92Å resolution of the RNA binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.
18. Price,S.R., Evans,P.R. and Nagai,K. (1998) Crystal structure of the spliceosomal U2B″-U2A′ protein complex bound to a fragment of U2 small nuclear RNA. *Nature*, **394**, 645–650.
19. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
20. Sillekens,P., Beijer,R., Habets,W. and van Venrooij,W. (1989) Molecular cloning of the cDNA for the human U2 snRNA-specific A′ protein. *Nucleic Acids Res.*, **17**, 1893–1906.
21. Habets,W.J., Hoet,M.H., de Jong,B.A.W., van der Kemp,A. and van Venrooij,W.J. (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.
22. Heitzler,P., Coulson,D., Saenz-Robles,M.T., Ashburner,M., Roote,J., Simpson,P. and Gubb,D. (1993) Genetic and cytogenetic analysis of the 43A-E region containing the segment polarity gene costa and the cellular polarity genes prickle and spiny-legs in *Drosophila melanogaster*. *Genetics*, **135**, 105–115.
23. Ewing,B. and Green,P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.*, **8**, 186–194.
24. Gordon,D., Abajian,C. and Green,P. (1998) Consed: a graphical tool for sequence finishing. *Genome Res.*, **8**, 195–202.
25. Bendixen,C., Gangloff,S. and Rothstein,R. (1994) A yeast mating-selection scheme for detection of protein–protein interactions. *Nucleic Acids Res.*, **22**, 1778–1779.
26. Buchanan,S. and Gay,N. (1996) Structural and functional diversity in the leucine-rich repeat family of proteins. *Prog. Biophys. Mol. Biol.*, **65**, 1–44.
27. Kajava,A.V. (1998) Structural diversity of leucine-rich repeat proteins. *J. Mol. Biol.*, **277**, 519–527.
28. Boelens,W., Scherly,D., Beijer,R.P., Jansen,E.J., Dathan,N.A., Mattaj,I.W. and van Venrooij,W.J. (1991) A weak interaction between the U2A′ protein and U2 snRNA helps to stabilize their complex with the U2B″ protein. *Nucleic Acids Res.*, **19**, 455–460.
29. Fresco,L.D., Harper,D.S. and Keene,J.D. (1991) Leucine periodicity of U2 small nuclear ribonucleoprotein particle (snRNP) A′ protein is implicated in snRNP assembly via protein-protein interactions. *Mol. Cell. Biol.*, **11**, 1578–1589.
30. Meyer,V., Oliver,B. and Pauli,D. (1998) Multiple developmental requirements of Noisette, the *Drosophila* homolog of the U2 snRNP-associated polypeptide SF3a^{60}. *Mol. Cell. Biol.*, **18**, 1835–1843.
31. Pan,Z.-Q. and Prives,C. (1989) U2 snRNA sequences that bind U2-specific proteins are dispensable for the function of U2 snRNP in splicing. *Genes Dev.*, **3**, 1887–1898.
32. Caspary,F. and Seraphin,B. (1998) The yeast U2A′/U2B″ complex is required for pre-spliceosome formation. *EMBO J.*, **17**, 6348–6358.