Polypeptide components of *Drosophila* small nuclear ribonucleoprotein particles

Trevor Paterson, Jean D.Beggs, David J.Finnegan and Reinhard Lührmann

Institute of Cell and Molecular Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR UK and 1Institut für Molekularbiologie und Tumorforschung, Klinikum der Philippus-Universität Marburg, Emil-Mannkopff-Strasse 2, D-3550 Marburg, FRG

Received August 20, 1991; Accepted October 14, 1991

ABSTRACT

In eukaryotes splicing of pre-mRNAs is mediated by the spliceosome, a dynamic complex of small nuclear ribonucleoprotein particles (snRNPs) that associate transiently during spliceosome assembly and the splicing reaction. We have purified snRNPs from nuclear extracts of *Drosophila* cells by affinity chromatography with an antibody specific for the trimethylguanosine (m$^3$G) cap structure of snRNAs U1-U5. The polypeptide components of *Drosophila* snRNPs have been characterized and shown to consist of a number of proteins shared by all the snRNPs, and some proteins which appear to be specific to individual snRNP particles. On the basis of their apparent molecular weight and antigenicity many of these common and particle specific *Drosophila* snRNP proteins are remarkably conserved between *Drosophila* and human spliceosomes. By probing western blots of the *Drosophila* snRNP polypeptides with a number of antisera raised against human snRNP proteins, *Drosophila* polypeptides equivalent to many of the HeLa snRNP-common proteins have been identified, as well as candidates for a number of U1, U2 and U5-specific proteins.

INTRODUCTION

In eukaryotes the excision of intervening sequences from pre-mRNA is catalyzed in a large ribonucleoprotein complex, the spliceosome. The spliceosome is composed of at least 30 distinct proteins and the five major small nuclear RNA (snRNA) molecules U1, U2, U4, U5 and U6. The U snRNAs can be recovered from nuclear extracts as four individual small nuclear ribonucleoprotein (snRNP) complexes, the monomeric U1, U2 and U5 snRNP particles and the dimeric U4/U6 snRNP in which the U4 and U6 snRNAs are found base paired (1,2).

The protein composition of the snRNP complexes has been studied in detail in human and rodent nuclear extracts (3,4). Eight distinct polypeptides, B, B', D1, D2, D3, E, F and G, with apparent molecular weights of between 9 and 29 (M$_r$ x 10$^3$), are common to all four snRNPs and form a common snRNP core. In addition, individual snRNPs have been shown to contain characteristic particle-specific proteins, denoted 70K, A and C for U1; A' and B" for U2; and 15, 40, 52, 100, 102, 116 and 200K for U5. The U5 snRNP has a particularly complex protein composition and sediments through glycerol gradients more rapidly than the other snRNPs (5). An even larger tripartite U4/U5/U6 complex can be isolated under low salt conditions and contains at least four further polypeptides (4,6—8).

Antisera from patients with a variety of connective tissue diseases recognize many of the snRNP polypeptides. Anti-RNP autoantibodies, from patients with mixed connective tissue disease (MCTD), react with the U1-specific proteins A, C, and 70K. Anti-Sm sera from patients with systemic lupus erythematosus (SLE) immunoreact with the core (Sm) proteins B', B and D (9,10), with variable reactivity to D1, D2 and D3 (11), and less frequently with E, F and G (12). These sera have been used to study the composition of snRNPs from a variety of species by immunoprecipitation and immunoblotting. These studies indicate that the protein composition of snRNPs is remarkably conserved throughout eukaryotes (reviewed 3) as indeed are the sequences of the U snRNAs (reviewed 13).

Purification of snRNPs has been facilitated by the use of the human autoantibodies, but more successfully by the use of monoclonal antisera raised against modified nucleosides peculiar to the snRNAs. In particular antibodies which recognize the highly methylated nucleoside 2,2,7-trimethylguanosine (m$^3$G), which is found as a 5'-cap on U1, A and C, and 70K. Anti-Sm sera from patients with systemic lupus erythematosus (SLE) immunoreact with the core (Sm) proteins B', B and D (9,10), with variable reactivity to D1, D2 and D3 (11), and less frequently with E, F and G (12). These sera have been used to study the composition of snRNPs from a variety of species by immunoprecipitation and immunoblotting. These studies indicate that the protein composition of snRNPs is remarkably conserved throughout eukaryotes (reviewed 3) as indeed are the sequences of the U snRNAs (reviewed 13).

Purification of snRNPs has been facilitated by the use of the human autoantibodies, but more successfully by the use of monoclonal antisera raised against modified nucleosides peculiar to the snRNAs. In particular antibodies which recognize the highly methylated nucleoside 2,2,7-trimethylguanosine (m$^3$G), which is found as a 5'-cap on U1, U2, U4 and U5 snRNAs, allow a single step purification of m$^3$G-capped snRNPs from nuclear extracts by immunooaffinity chromatography. Native snRNPs can be recovered by desorption with an excess of m$^3$G (or the cross-reactive nucleoside 7-methylguanosine, m$^7$G) or by elution with salt (14—17).

The *Drosophila* U snRNAs closely resemble those of vertebrates in size, structure and nucleotide sequence (18, reviewed 13). Anti-Sm and anti-RNP antisera have identified a number of low molecular weight potential snRNP polypeptides (19,20). In this study we have used a monoclonal antibody which recognizes the 5'-nucleotide cap of snRNAs (16) to immunooaffinity purify snRNPs from nuclear extracts of *Drosophila* cells in order to identify their protein components and to compare them to those of HeLa cells.
MATERIALS AND METHODS
Cell growth and preparation of nuclear extracts
The *Drosophila melanogaster* Schneider 2 cell line was propagated at 22 °C in suspension in Schneider’s (modified) *Drosophila* Medium (purchased from Imperial Laboratories, UK) supplemented with 5% foetal calf serum, 100U/ml penicillin and 100μg/ml streptomycin. Nuclear extracts were prepared by the method of Dignam *et al.*, (21) without homogenization of cells and with the addition of 0.5% v/v NP40 to disrupt the cytoplasmic membrane. HeLa cell nuclear extracts and purified snRNPs were kindly provided by Silke Börner.

Anti-m3G immunoaffinity chromatography of snRNPs
Affinity purification of m3G-capped snRNPs from nuclear extracts was performed by using mAb H-20 bound covalently to CNBr-activated Sepharose 4B essentially as described (16). Nuclear extract prepared as above was passed over a 5ml H-20-Sepharose column and bound RNA-protein complexes were desorbed with 20mM m7G.

Fig. 1. Immunoaffinity purification of *Drosophila* snRNPs. Analysis of fractions upon chromatography over H-20-Sepharose. (A) Coomassie stained proteins resolved on a 15% polyacrylamide/SDS gel compared with HeLa snRNP proteins (H) and size markers (S) (M, x 10^-3). L, nuclear extract loaded onto column; F, unbound flow through; W, initial wash after loading; E, eluate desorbed on addition of m7G; U, proteins in 6M urea wash to regenerate column. (B) Ethidium bromide stained RNA resolved on a 6% polyacrylamide/urea gel. Fractions as above except that W, wash fraction was sampled after extensive washing of the column prior to desorption. The identities of U1, 4, 5 and 6 RNAs were determined by hybridizing a northern blot of the gel with unique oligonucleotide probes. Positions of 5S rRNA and 4S tRNA are indicated.

Anti-sera
Rabbit: anti-FP8.1 raised against a portion of yeast U5 snRNP specific protein PRP8 fused to β-galactosidase and affinity-purified against the same portion of PRP8 fused to the trpE protein (22).

Monoclonal antibodies: Y12 anti-Sm type (9); 7.13 anti-HeLa D1 (23); H-20 anti-m3G (16); 4G3 anti-HeLa U2B' (24); H111 anti-HeLa U1 70K, H304 anti-HeLa U1 A and H386 anti-HeLa U1 70K and U5 100K (8,25).

Patient 7, E74 and Plasma D anti-Sm sera from patients with systemic lupus erythematosus were kindly provided by Prof. H.Peter, Freiburg.

Alkaline phosphatase-conjugated secondary antibodies (against rabbit, human or mouse IgG or mouse IgM antibodies) were purchased from Promega and Bio-Rad.

RNA and protein analysis
Samples were extracted with 1 volume of PCA (phenol, chloroform, isoamylalcohol 50:50:1) with 0.5% (w/v) SDS. Following phase separation by centrifugation RNAs were recovered from the aqueous phase by ethanol precipitation. Proteins were precipitated from the organic phase with 5 volumes of acetone.

RNAs were fractionated by electrophoresis through 10% polyacrylamide gels containing 7M urea, 10mM Tris-Borate pH8.2 and 2mM EDTA (TBE). RNA was visualized by staining with ethidium bromide (0.5/μg/ml) or with silver (26). RNA was transferred to nylon membrane (Amersham) by electrophoresis in 0.5X TBE for 3h at 45V and fixed by UV irradiation. U snRNA-specific oligonucleotides were 5'-end labelled using polynucleotide kinase and hybridized to pre-blocked filters in 6X SSC (3M NaCl, 3M trisodium citrate), 5xDenhardt's solution, 100/μg/ml denatured salmon sperm DNA for 2h at 4°C and washed briefly in 6X SSC at 20°C before autoradiography. Oligonucleotides were synthesized by OSWEL DNA Service, Edinburgh University, and were directed against: *Drosophila* U5

Fig. 2. Comparison of *Drosophila* (D) with HeLa (H) snRNP proteins resolved on a 12% acrylamide/SDS gel polymerized in the presence of high TEMED concentration (see Methods). The migration of size standards is shown (14-205 M, x 10^-3) and the major HeLa snRNP proteins are indicated.
bases 68–83 GACTCTAGAGTGTTC; Human U4 bases 58–75 GGAAAATTTTCAATTAG; Human U6 bases 78–95 GGCTTCACGAATTTGCGT.

Proteins were fractionated by SDS-gel electrophoresis (27). In order to resolve small proteins, in particular the HeLa D proteins, gels were polymerized where noted in the presence of high TEMED concentrations (0.04% compared to 0.01%, ref. 11). Protein bands were detected by staining with Coomassie brilliant blue G250 or were transferred to nitrocellulose by electrophoresis overnight at 25V in 25mM Tris, 19mM glycine (28). Blots were then probed with sera or monoclonal antibodies as described previously (25) and antibodies detected with alkaline-phosphatase-conjugated secondary antibodies and chromogenic substrates (Promega).

Glycerol gradient centrifugation
Affinity purified snRNP fraction (200μg protein, prepared as described above) was layered onto a linear 10—30% (v/v) glycerol gradient in a buffer containing 20mM Hepes-KOH pH8, 150mM KCl, 1.5mM MgCl₂, 0.5mM PMSF, 0.5mM DTT, and 2μg of leupeptin per ml. The gradient was centrifuged in a Beckman SW40 rotor at 29 krpm for 18h. 500μl fractions were collected and PCA extracted for analysis as described above.

RESULTS
Purification of Drosophila snRNPs from nuclear extracts by chromatography on H-20 affinity columns
Nuclear extract prepared from 0.5 x 10¹⁰ Drosophila cells was purified by H-20 immunoaffinity chromatography (Materials and Methods) and yielded 620μg of protein upon desorption of bound snRNPs (see Figure 1A). The yield was somewhat lower than routinely achieved with HeLa cell extracts. Analysis of the RNA eluted from the H-20 column revealed a marked enrichment of the U snRNAs from the nuclear extracts (Figure 1B).

Characterization of Drosophila snRNP proteins in relationship to their human counterparts
The protein component of purified Drosophila snRNPs was analysed by SDS-gel electrophoresis and compared with purified snRNPs from HeLa cells. A larger number of individual proteins were enriched from Drosophila in comparison with HeLa cells from which around 30 major proteins were selectively enriched (Figures 1A, 2). As a consequence, few direct correlations between the sizes of Drosophila and human proteins are immediately apparent.

Drosophila and HeLa snRNPs were transferred to nitrocellulose after gel electrophoresis and probed with a variety of antisera reactive to HeLa snRNP proteins in order to investigate the relationship between the affinity purified Drosophila and human proteins. When strips were probed with anti-Sm sera from three human SLE patients a number of Drosophila Sm proteins were identified (Figure 3A). On the basis of their apparent molecular weight these may include at least two D proteins, Du and DL (DL may in fact be a doublet, see lane 4 and also Figure 2A), at least one potential B protein, a potential A protein, several proteins with similar mobility to the HeLa Ul 70K protein and a doublet at about 100K (lane 4). The exact relationship of Du and DL to the HeLa Dl, 2 and 3 proteins is not clear but E 74 serum, which recognizes human Dl and D3, only recognizes the more slowly migrating Drosophila protein Du (lane 1) whilst sera recognizing all three HeLa D proteins also recognize DL (lanes 2 and 4).

Various monoclonal antibodies and rabbit antisera which react with specific snRNP proteins were also used as probes

Fig. 3. Immunological comparison of Drosophila (D) with HeLa (H) snRNP proteins. (A) Nitrocellulose strips of electroblotted snRNP proteins probed with various human anti-Sm antisera. 1, E 74; 2, Plasma D; 3, Patient 7 (1:200 dilution); 4, Patient 7 (1:50 dilution). (B) Strips probed with various monospecific rabbit and mouse antibodies: 1, Y12; 2, 7.13; 3, H111; 4, H386; 5, H304; 6, anti-FP8.1; 7, 4G3.
(Figure 3B). These antibodies confirmed the existence of at least two distinct D proteins D^U and D^L (lanes 1 and 2). Again the relationship of D^U and D^L to D1, D2 and D3 could not be unambiguously assigned since in this case antibodies raised against D1 (7.13) recognized D^L (lane 2). The mobility of the D proteins was clearly affected by the composition of the acrylamide gel used (Materials and Methods, 11). In a 12% acrylamide gel polymerized with high concentrations of TEMED, D^U and D^L ran with similar mobility to HeLa D proteins (Figure 3A lanes 1,2; Figure 3B lanes 1,2), but in a 14% acrylamide, low TEMED gel the Drosophila D proteins migrated more slowly than the HeLa counterparts (Figure 3A lanes 3,4).

As was observed with the Human Sm antisera only a single Drosophila B protein was detected by the anti-Sm monoclonal antibody Y12 (lane 1). The antibodies H111 and H386 (lanes 3,4) define a unique Drosophila band related to the HeLa Ul 70K protein whereas the human anti-Sm sera (Figure 3A) indicated that there may be multiple post-translationally modified variants of this protein as has been observed in HeLa cells. Antiserum raised against a portion of the yeast U5 snRNP specific protein PRP8 (22) recognized large 200—260kDa proteins in both HeLa and Drosophila snRNP particles (lane 6) as previously reported (29—31). Antisera raised against other regions of the yeast protein (22) also recognized both the HeLa and Drosophila protein (30, results not shown). A monoclonal (4G3) raised against the HeLa U2-specific B’ protein (24) recognized a similarly sized Drosophila protein (lane 7).

A monoclonal antibody specific to the HeLa A protein (H304, lane 5) and rabbit antipeptide antisera to the HeLa C and G proteins (results not shown) failed to recognize equivalent Drosophila proteins, perhaps reflecting lack of conservation of these epitopes between Drosophila and humans. Results from these and the later analyses are summarized in Table 1.

### Table 1. Comparison of human and Drosophila snRNP polypeptides

<table>
<thead>
<tr>
<th>HELA snRNP Proteins</th>
<th>Equivalent Drosophila proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designation</td>
<td>Designation</td>
</tr>
<tr>
<td>(a)</td>
<td>Apparent molecular weight (a,b)</td>
</tr>
<tr>
<td>Common proteins</td>
<td></td>
</tr>
<tr>
<td>B’</td>
<td>29</td>
</tr>
<tr>
<td>B</td>
<td>28</td>
</tr>
<tr>
<td>D3</td>
<td>18</td>
</tr>
<tr>
<td>D2</td>
<td>16.5</td>
</tr>
<tr>
<td>D1</td>
<td>16</td>
</tr>
<tr>
<td>E</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>11</td>
</tr>
<tr>
<td>G</td>
<td>9</td>
</tr>
<tr>
<td>U1 snRNP proteins</td>
<td></td>
</tr>
<tr>
<td>70K</td>
<td>70K(62)</td>
</tr>
<tr>
<td>A</td>
<td>34</td>
</tr>
<tr>
<td>C</td>
<td>22</td>
</tr>
<tr>
<td>U2 snRNP proteins</td>
<td></td>
</tr>
<tr>
<td>B’</td>
<td>28.5</td>
</tr>
<tr>
<td>U5 snRNP proteins</td>
<td></td>
</tr>
<tr>
<td>200K(+K)</td>
<td>&gt;200</td>
</tr>
<tr>
<td>116K</td>
<td>116</td>
</tr>
<tr>
<td>102K</td>
<td>102</td>
</tr>
<tr>
<td>100K</td>
<td>100</td>
</tr>
<tr>
<td>52K</td>
<td>52</td>
</tr>
<tr>
<td>40K</td>
<td>40</td>
</tr>
</tbody>
</table>

(a) From Lührmann et al. (4). (b) M×10^-3.

### Fig. 4. Sedimentation of Drosophila snRNP particles in a linear 10—30% glycerol gradient

The immunoaffinity purified Drosophila snRNP fraction was subject to sedimentation analysis in a linear 10—30% (v/v) glycerol gradient in order to investigate whether rapidly and slowly sedimenting particles could be resolved as has been demonstrated for HeLa snRNP particles (5). When the protein content of the glycerol gradient fractions was analysed by gel electrophoresis a pattern remarkably similar to that observed with HeLa extracts was found (Figure 4A). The snRNP particles appear to have sedimented in two different size ranges; the rapidly sedimenting particles (fractions 14—17) appear to correspond to the 20S U5 snRNP of HeLa (5) whilst the more slowly sedimenting particles (fractions 4—10) resemble in composition
the smaller snRNPs, including the U1 snRNP, with a value of around 10S.

When the RNA present in the gradient fractions was analysed by urea/acylamide electrophoresis and northern blot a less distinct pattern of separation was found. Whilst a peak of U5 snRNA indeed co-sedimented with the 20S particles (fractions 14 and 16, Figure 4B) all of the U snRNAs examined appeared to smear throughout the gradient, perhaps due to the loss of proteins from the snRNPs. However, the U1 and U6 snRNAs were more concentrated further up the gradient than U5 (fractions 6—12, Figure 4B). The protein composition of the two resolved particle types is shown in greater detail in Figure 4C. Common core proteins present in both particle types include two candidate B proteins, with apparent molecular weights marginally smaller than HeLa B and B' (the faster migrating species, B^L, corresponds to the B protein recognized by anti-Sm sera; Figure 3A lanes 1,2 and 4); the two D proteins D^U and D^L identified immunologically above, and two proteins with similar sizes to the HeLa core proteins E and F. Whilst a protein of similar size to Hela G is present in the total Drosophila snRNP fraction (arrowed in lane L), this band is not discernible in either of the glycerol gradient fractions.

The more slowly migrating complexes (fractions 4—10) appear to contain the U1 snRNP specific proteins A and 70K. Again a number of proteins with apparent size of about 70000 can be distinguished, perhaps indicating that multiple post-translationally modified variants of this protein are found in Drosophila as in HeLa. Alternatively this protein could have suffered proteolytic degradation during purification. Of the three candidate 70K proteins in Drosophila with M_r 54 000, 56 000 and 61 000 it appears to be the 56 000 species which is recognized by the anti-70K monoclonal antibodies. Whilst there is no protein of similar molecular weight to the HeLa C protein (M_r 22 000) a prominent polypeptide of M_r 14 500 was, however, resolved with reported molecular weights of 18 and 26 000 (19) and 14 and 26 000 (20).

Proteins which are specific to U1, U2 and U5 snRNPs are also conserved between humans and Drosophila. A counterpart to the HeLa U5-specific 200K protein which cross reacts with antibodies raised against the yeast PRP8 protein (29—31) is found in Drosophila U5-containing particles, as are at least two proteins of about M_r 100—120 000 and other proteins potentially corresponding to the less well characterized 40 and 52K HeLa U5 proteins.

Drosophila proteins corresponding to the HeLa U1 snRNP specific 70K and A proteins were recognizable on the basis of their size and antigenicity, but no protein of equivalent molecular weight to the HeLa C protein could be detected. An abundant protein of M_r 14 500 was, however, resolved by density centrifugation and may be the Drosophila counterpart to C (Figures 4A,C).

A protein of similar size to the HeLa U2-specific B' protein was recognized by an antibody raised against the human protein (24). This antibody also recognizes a protein of M_r 27 000 present in plant nuclear extracts and the gene encoding B' has recently been cloned from potato and shown to be highly conserved between the animal and plant kingdoms (43). A candidate gene encoding for either the U2-specific A' or B' proteins has recently been cloned from Drosophila (44).

As in HeLa extracts, a number of electrophoretically distinct species of around 70 000 were detected. The most abundant of these had an apparent molecular weight similar to that of the most abundant HeLa 70K protein (61—62 000 in this system, Figures 2; 4A). However, the species recognized by monoclonal antibodies raised against the HeLa 70K protein had an apparent molecular weight of 56—58 000 (Figure 3B). A further species of 52—54 000 was also detected (Figure 2) and all three or more bands, perhaps representing modified or partially degraded versions of 70K, were recognized by some anti-Sm sera (Figure 3A). The gene encoding the Drosophila 70K protein has been cloned and found to encode a protein of predicted M_r
53 000, the amino-terminal region of which is 68% identical to the human and Xenopus 70K proteins (45). Whilst there is no evidence in Drosophila that alternative RNA processing pathways give rise to 70K variants it is possible that some different forms of 70K may be generated by variable post translational modifications as has been reported for the mouse and human 70K proteins (46).

Until now the biochemical purification of snRNP particles for their characterization has largely been restricted to rodent and human cell types. Studies on a disparate range of species, primarily using autoimmune antibodies, have identified proteins corresponding to D, E, F and G in most systems (reviewed 3). This study shows that in Drosophila the conservation of the proteins involved in the mechanism of splicing extends much further. The striking degree of similarity between the snRNP proteins found in Drosophila and humans is perhaps not unexpected, given the evolutionary constraints on maintaining an accurate and efficient splicing apparatus.

ACKNOWLEDGEMENTS

This work was funded by project grants from the UK MRC (DJF) the Deutsche Forschungsgemeinschaft (RL) and an SERC postdoctoral research fellowship (TP). JDB is a Royal Society E.P.A. Cephalosporin Fund Senior Research Fellow. The western blot shown in Figure 3B lane 7 was performed by Helen Kulesza. TP would like to thank all the members of the Lührmann laboratory, especially Silke Börner, for their assistance and hospitality.

REFERENCES