The EMBO Journal vol.15 no.9 pp.2256–2269, 1996

The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*

Veronica A.Raker, Gabriele Plessel and Reinhard Lührmann[1]

Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Emil Mannkopff Straße 2, D-35037 Marburg, Germany

[1]Corresponding author

V.A.Raker and G.Plessel contributed equally to this work

Stable association of the eight common Sm proteins with U1, U2, U4 or U5 snRNA to produce a spliceosomal snRNP core structure is required for snRNP biogenesis, including cap hypermethylation and nuclear transport. Here, the assembly of snRNP core particles was investigated *in vitro* using both native HeLa and *in vitro* generated Sm proteins. Several RNA-free, heteromeric protein complexes were identified, including E·F·G, B/B′·D3 and D1·D2·E·F·G. While the E·F·G complex alone did not stably bind to U1 snRNA, these proteins together with D1 and D2 were necessary and sufficient to form a stable U1 snRNP subcore particle. The subcore could be chased into a core particle by the subsequent addition of the B/B′·D3 protein complex even in the presence of free competitor U1 snRNA. Trimethylation of U1 snRNA's 5′ cap, while not observed for the subcore, occurred in the stepwise-assembled U1 snRNP core particle, providing evidence for the involvement of the B/B′ and D3 proteins in the hypermethylation reaction. Taken together, these results suggest that the various protein heterooligomers, as well as the snRNP subcore particle, are functional intermediates in the snRNP core assembly pathway.

Keywords: cap hypermethylation/protein–protein interactions/RNP assembly/Sm proteins/snRNP biosynthesis

Introduction

The four major small nuclear ribonucleoprotein particles (snRNPs) U1, U2, U4/U6 and U5 are evolutionarily highly conserved RNA–protein complexes which participate in the splicing of pre-mRNA (Steitz *et al.*, 1988; Moore *et al.*, 1993). Each snRNP consists of one or two uridyl-rich snRNAs (U1, U2, U5 or U4/U6) and a set of snRNP proteins. The U1, U2, U4 and U5 snRNAs are transcribed by RNA polymerase II. They contain two common features, the so-called Sm site, which is a single-stranded, uridylic acid-rich region flanked by two stem–loop structures (Branlant *et al.*, 1982), and the 5′-terminal trimethylguanosine (m_3G) cap. U6 is exceptional in that it is transcribed by RNA polymerase III, has a γ-mono-methylphosphate cap and does not contain an Sm site (Reddy and Busch, 1988; Singh and Reddy, 1989).

The snRNP proteins fall into two classes: the specific proteins, which associate exclusively with one snRNP species, and the common proteins, which are present in each spliceosomal snRNP particle (Lührmann *et al.*, 1990). Due to their recognition by anti-Sm autoantibodies isolated from patients suffering from systemic lupus erythematosus (SLE), the snRNP core proteins are also called Sm proteins (Lerner and Steitz, 1979). To date, eight core proteins, denoted B′ (29 kDa), B (28 kDa), D3 (18 kDa), D2 (16.5 kDa), D1 (16 kDa), E (12 kDa), F (11 kDa) and G (9 kDa), have been identified in human HeLa cells. All of the cDNAs encoding the human core proteins have now been cloned (Hermann *et al.*, 1995, and references within). With the exception of B and B′, all core proteins are encoded by separate genes; the former are produced by alternative splicing of a single gene product and differ by only 11 amino acids at the C-terminus (Van Dam *et al.*, 1989; Chu and Elkon, 1991). Interestingly, two sequence motifs, denoted Sm motif 1 and 2, are found in all of the known Sm proteins, from organisms as diverse as yeast, nematodes, insects, plants and higher vertebrates (Cooper *et al.*, 1995; Hermann *et al.*, 1995; Séraphin, 1995). There is experimental evidence that the Sm motifs are important for Sm protein–protein interactions (Hermann *et al.*, 1995).

The Sm proteins bind to the Sm site of the U1, U2, U4 and U5 snRNAs to form a highly stable snRNP core structure, which is morphologically similar among all snRNPs (Kastner *et al.*, 1990). Little is known about the nature of protein–RNA interactions within the snRNP core structure. In the snRNP particle, the core proteins have been shown to protect a 15–25 nucleotide long region of the Sm site against hydrolysis by micrococcal nuclease (Liautard *et al.*, 1982). To date, the only known interaction between a core protein and U snRNA is that between the U1 snRNA Sm site and the G protein; this interaction was demonstrated by cross-linking analyses (Heinrichs *et al.*, 1992). Thus, protein–protein, rather than protein–RNA, interactions appear to dominate in determining the structure of the snRNP core.

The biogenesis of snRNPs is a complex process which takes place in both nuclear and cytoplasmic compartments. First, the RNA polymerase II-transcribed U snRNAs are synthesized in the nucleus as precursor molecules containing a m^7G cap structure and short 3′ extensions (Dahlberg and Lund, 1988; Mattaj, 1988). The precursors subsequently are exported to the cytoplasm, where large pools of the Sm proteins are stored (Zeller *et al.*, 1983; Zieve and Sauterer, 1990) and the latter then assemble onto the snRNA's Sm site. Association of the Sm proteins with the Sm site is essential for the hypermethylation of the snRNA cap structure to generate the m_3G cap (Mattaj, 1986); the core proteins probably provide a binding site for the *trans*-active snRNA-(guanosine-*N*2)-methyltransferase (Plessel *et al.*, 1994). In addition, the core proteins play

an important role in the formation of one part of the bipartite nuclear localization signal (NLS) of the snRNP core particle (Mattaj and De Robertis, 1985); the m_3G cap forms the second part of the NLS (Fischer and Lührmann, 1990; Hamm *et al.*, 1990). Since free Sm proteins do not enter the nucleus, the NLS of the snRNP core must be a masked property of the core proteins which is exposed as a result of proper snRNP core assembly. It has not yet been determined which of the core proteins contribute to the binding of the cap methyltransferase or to the NLS.

Despite the crucial role which the snRNP core particle plays in the biogenesis of snRNPs, information regarding the order of events leading to its formation is limited and has been provided mostly by experiments performed *in vivo*. Following pulse–chase and sucrose gradient fractionation of cytoplasmic proteins labeled *in vivo* with [^{35}S]methionine, immunoprecipitation with the anti-Sm monoclonal antibody Y12 indicated that E, F, G and one or more D protein co-sediment at 6S (Fisher *et al.*, 1985; Sauterer *et al.*, 1990). Thus, it was suggested that these proteins form an RNA-free, heteromeric complex that binds as such to the snRNA's Sm site, followed by the addition of the B/B′ proteins (Fisher *et al.*, 1985; Feeney *et al.*, 1989; Sauterer *et al.*, 1990). While these studies indicated for the first time that the assembly of the snRNP core may occur in multiple steps, important aspects of this assembly process remained unclear. Significantly, the complexity of the core snRNP protein composition was unknown at the time the aforementioned *in vivo* experiments were carried out, i.e. the three distinct D proteins, D1, D2 and D3, were considered to be a single protein. Thus, it was not clear whether all three D proteins were involved in an RNA-free complex together with the E, F and G proteins, or whether the individual D proteins exhibit differential protein association behavior. Our recent finding that D1 and D2 specifically interact with each other but not with D3 (Lehmeier *et al.*, 1994) supported the latter hypothesis and stimulated a detailed investigation of the core snRNP assembly pathway. We therefore analyzed the *in vitro* formation, not only of RNA-free core protein complexes, but also of core protein–UsnRNA complexes. Our results indicate that there is an ordered snRNP core assembly pathway that involves the initial formation of D1·D2, E·F·G, D1·D2·E·F·G and B/B′·D3 complexes, the subsequent association of U snRNA with D1·D2·E·F·G to form an snRNP subcore particle and, finally, the recruitment of the B/B′·D3 complex to generate an snRNP core particle functional in 5′ cap hypermethylation.

Results

Fractionation of native HeLa snRNP core proteins on sucrose density gradients reveals an E·F·G complex with an S value of 3.7

Understanding snRNP core assembly requires a precise knowledge of the heteromeric complexes formed by the core proteins in the absence of UsnRNA. We initially investigated whether protein heterooligomers could be identified in native HeLa snRNP core proteins, prepared from anti-m_3G immunoaffinity-purified snRNP particles. Interpretation of the results obtained by this approach relies on the assumption that strong protein–protein interactions

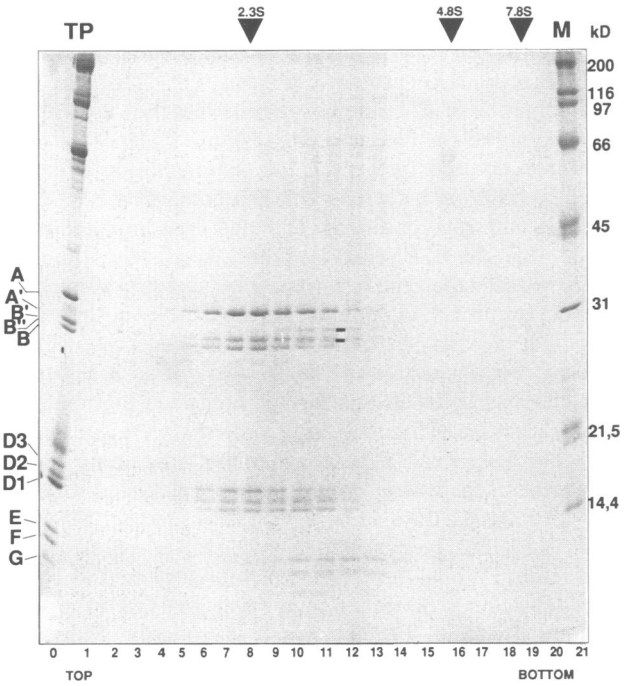

Fig. 1. Density gradient centrifugation of native, RNA-free, HeLa snRNP proteins. SnRNP proteins isolated from HeLa snRNP particles were fractionated on 6–20% sucrose gradients containing 300 mM KCl. After fractionation, proteins were separated on a high-TEMED, SDS–12.5% polyacrylamide gel and visualized by Coomassie blue staining. Lane 0, snRNP proteins; lanes 1–20, proteins from the gradient fractions, whereby the top of the gradient is at the left; and lane 21, marker proteins. The migration of the A′ and B″ bands observed is indicated with bars between lanes 11 and 12. The S values of various sedimentation standards (see Materials and methods) are indicated by triangles at the top.

once formed in RNA-free assembly intermediates may persist in the intact RNP particle and also survive RNP dissociation under mild conditions. Examples of this kind have been observed previously with ribosomes (see, for example, Dijk and Littlechild, 1979). The snRNP proteins were separated from the snRNAs by incubating the isolated snRNPs with the anion exchange resin DE53 in the presence of EDTA (Sumpter *et al.*, 1992). This method produces snRNP proteins which retain their native state, as demonstrated by their ability, under *in vitro* reconstitution conditions, to form snRNPs functional in nuclec-cytoplasmic transport (Fischer *et al.*, 1993) and splicing (Ségault *et al.*, 1995). Native snRNP protein preparations contain predominately the core proteins, in addition to some particle-specific proteins (e.g. the U1-specific A and C, and the U2-specific A′ and B″ proteins; see below).

To assay for the formation of snRNP core protein complexes, native snRNP proteins were fractionated on 6–20% sucrose gradients (Figure 1). Unexpectedly, the E, F and G polypeptides migrated faster than the bulk of core proteins, co-migrating in the 3.3–4S range with a peak at ~3.7S (lanes 10–14). This corresponds to the sedimentation of a globular protein of ~56 kDa (see Materials and methods). An internal marker which correlates with this rough estimation of 56 kDa was also provided by the U2-A′ and U2-B″ proteins (Figure 1; A′ and B″ are indicated with bars) which are known to form an RNA-free dimer of ~60 kDa (Scherly *et al.*, 1990).

U2-A′ and U2-B″ co-sedimented faster than the B, B′ and D proteins, but at a rate approximately equal to that of the E, F and G proteins (Figure 1). Thus, the co-sedimentation of E, F and G suggests that they are present as a complex on the gradient (Figure 1, lanes 10–14). This is supported further by the sedimentation behavior of E, F and G on a sucrose gradient containing 8 M urea: under denaturing conditions, E, F and G sedimented more slowly than the B, B′ and D proteins, consistent with their relative molecular masses (data not shown). Additional support for an E·F·G complex is provided by co-immunoprecipitation analyses of *in vitro* translated proteins (see below). Sedimentation of E, F and G at 3.7S under non-denaturing conditions further suggested that multiple copies of at least one of these proteins were present in the complex, since a heterooligomer containing single copies of each protein would have a calculated mol. wt of ~29 kDa (E, 10.8 kDa; F, 9.7 kDa; and G, 8.5 kDa) and would, thus, be expected to have a lower sedimentation coefficient. While the exact stoichiometry of the E·F·G complex cannot be determined here, analysis of snRNP proteins by SDS–PAGE reveals a roughly equivalent Coomassie blue staining intensity for E, F and G (see Figure 1), suggesting that the ratio of E:F:G within snRNP particles is 1:1:1 (the apparently lower staining intensity of G can be attributed to its migration as a doublet in high-TEMED gels; see also Hermann *et al.*, 1995).

In contrast to the U1-A protein, whose migration behavior corresponds well to its mol. wt of ~31 kDa (it peaks around 2.3S, Figure 1; Sillekens *et al.*, 1987), additional core proteins also appear to have an aberrant migration on sucrose gradients under native conditions (Figure 1). For example, the large S values of the D proteins do not correspond to their M_rs of 16–18 kDa, as they co-migrate on the gradient in the range of the B and B′ proteins (~29 kDa; Figure 1). This may be due either to unusual three-dimensional configurations or to their involvement in oligomeric complexes. It should be noted, however, that there was no evidence for the formation of a stable complex between E, F, G and one or more of the D proteins under these sedimentation conditions.

An E·F·G complex is formed with proteins translated in vitro

Protein–protein interactions were investigated more precisely by co-immunoprecipitation analyses. To determine if an E·F·G complex can indeed be formed, we initially tested *in vitro* translated, [^{35}S]methionine-labeled E, F and G polypeptides in immunoprecipitation assays with a polyclonal anti-F antibody, which was raised against recombinant F protein (Hermann *et al.*, 1995). While neither E nor G alone was precipitated significantly by the anti-F antibody (Figure 2A, lanes 4 and 6), both could be co-immunoprecipitated efficiently when F was present (lane 1). In addition, E, but not G, was co-precipitated efficiently in the presence of F (compare lane 2 with the E background precipitation in lane 4, and lane 3 with lane 6). This suggests that there is a stable interaction between E and F but not between F and G. Thus, since the anti-F antibody did not precipitate G when incubated with F alone (lane 3), the co-precipitation of E, F and G observed in lane 1 cannot be explained by the formation of E·F and G·F dimers, but rather demonstrates that an E·F·G complex

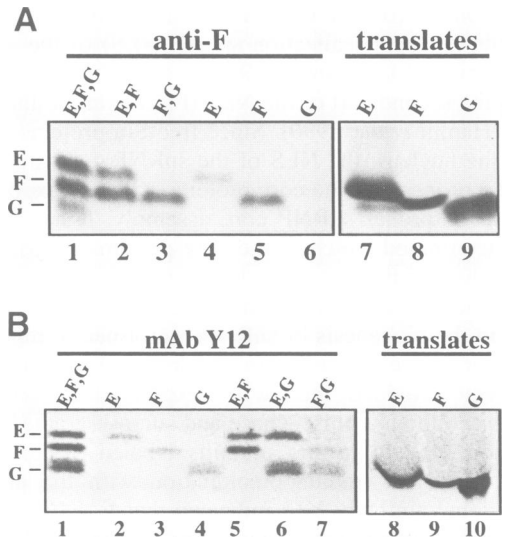

Fig. 2. Complex formation of E, F and G as shown by immunoprecipitation assays. Proteins were prepared by *in vitro* translation and labeled with [^{35}S]methionine. (**A**) Immunoprecipitation with an anti-F antibody, washed with buffer containing 150 mM KCl. Lanes 1–6, immunoprecipitation with E, F and G (lane 1), E and F (lane 2), F and G (lane 3), E (lane 4), F (lane 5) and G (lane 6); lanes 7–9, E, F and G proteins equivalent to 40% of the amount used in the immunoprecipitation assays. (**B**) Immunoprecipitation with mAb Y12, washed with buffer containing 450 mM KCl. Lanes 1–7, immunoprecipitation with E, F and G (lane 1), E (lane 2), F (lane 3), G (lane 4), E and F (lane 5), E and G (lane 6) and F and G (lane 7); lanes 8–10, proteins equivalent to 40% of the amount used for immunoprecipitation assays; the unusual migration of E, F and G is due to the high amount of globin present in the reticulocyte lysate. Fractionation was carried out by electrophoresis on a high-TEMED, SDS–12.5% polyacrylamide gel, and bands were visualized by fluorography.

was indeed formed. Note that the co-precipitation of E with F was reproducibly more efficient when G was also present (compare lanes 1 and 2). Co-precipitation of E and G with F by the anti-F antibody was observed not only under low salt conditions (i.e. 150 mM KCl as shown in Figure 2A), but also at high salt concentrations (e.g. 500 mM KCl, data not shown), indicating that the E·F·G complex is relatively stable.

The protein–protein interactions within the E·F·G complex were analyzed further by immunoprecipitation assays with the anti-Sm monoclonal antibody (mAb) Y12, which reacts with E, but not with F or G, on immunoblots (Lerner *et al.*, 1981). As shown in Figure 2B, the epitope recognition of Y12 in immunoprecipitation assays with *in vitro* translated proteins is unusual. That is, while the individually tested E, F and G proteins were precipitated only minimally (lanes 2–4), all were precipitated efficiently when mixed together prior to the addition of antibody (lane 1). This indicates that the Y12 epitope(s) is either hidden or not present on the individual proteins, and, thus, is first created or presented by conformational changes in the polypeptides which occur through their interactions with one another. Similarly, the combinations of either E and F, or E and G, in Y12 immunoprecipitation assays led to the efficient precipitation of both proteins (compare lanes 5 and 6 with lanes 2–4), indicative of protein–protein interactions between E and F as well as between E and G. No precipitation above background (i.e. the individually tested proteins) was observed when F and G

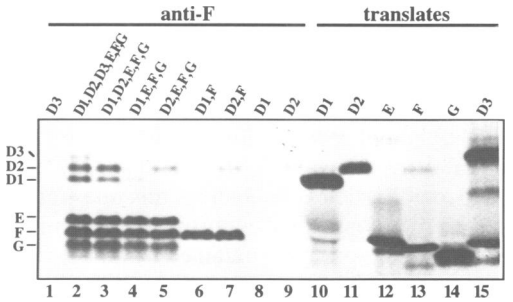

Fig. 3. Evidence that the D1, D2, E, F and G core proteins form an RNA-free complex *in vitro*. Proteins were prepared by *in vitro* translation and labeled with [^{35}S]methionine. Immunoprecipitation was performed with a polyclonal anti-F antibody. Lanes 1–9, immunoprecipitation of D3 (lane 1), D1, D2, D3, E, F and G (lane 2), D1, D2, E, F and G (lane 3), D1, E, F and G (lane 4), D2, E, F and G (lane 5), D1 and F (lane 6), D2 and F (lane 7), D1 (lane 8) or D2 (lane 9); lanes 10–15, proteins equivalent to 50% of the amount added to the assays. Protein fractionation was carried out by electrophoresis on a 12.5% high-TEMED SDS–polyacrylamide gel, and bands were visualized by fluorography.

were assayed together (lane 7), suggesting that these proteins do not interact. The results obtained with Y12 are in agreement with the anti-F immunoprecipitation data described above. Note that the immunoprecipitates were washed at a high ionic strength (450 mM KCl), indicating that the complexes which were formed are relatively stable. As the methionine content, and hence the ^{35}S-labeling, of each polypeptide is approximately the same (E contains five, and F and G, six methionines), the equal intensities of the co-immunoprecipitated E, F and G bands (Figure 2B, lane 1) strongly suggest that these polypeptides are present in stoichiometric amounts within an RNA-free, E·F·G complex.

RNA-free heteromeric complex formation with D1, D2, E, F and G

While fractionation of native snRNP proteins on sucrose gradients did not reveal a larger complex containing E, F, G and one or more of the D proteins (see Figure 1), evidence for the formation of such a complex was provided previously by pulse–chase experiments *in vivo* (see Introduction). We therefore investigated whether a larger heteromeric complex could be detected when *in vitro* translated E, F, G, D1, D2 and D3 proteins were mixed and then immunoprecipitated with the polyclonal anti-F antibody, which does not precipitate the individual D3, D1 or D2 proteins (Figure 3, lanes 1, 8 and 9, respectively) or the D1·D2 dimer (data not shown). Interestingly, D1 and D2 were co-precipitated efficiently with the E·F·G complex (Figure 3, lanes 2 and 3), while D3 was co-precipitated only slightly (but reproducibly) above background (lane 2). The same result was obtained when the reticulocyte lysate was treated with micrococcal nuclease after translation, but prior to addition to the immunoprecipitation assays (data not shown). Thus, D1 and D2 interact strongly with the E·F·G complex, while D3 has only a low affinity for this complex and is not stably associated. This result was also observed when identical immunoprecipitation assays were carried out with a polyclonal anti-D2 antibody (data not shown), demonstrating that the precipitation of a D1·D2·E·F·G complex is independent of the antibody used for precipitation.

The stable interaction of D1 and D2, but not D3, with the E·F·G complex is interesting in light of our previous finding that D1 and D2 specifically interact with each other, but not with D3 (Lehmeier *et al.*, 1994). We therefore investigated whether the interaction of D1 and D2 with E·F·G requires D1·D2 dimer formation, or if either D1 or D2 could interact individually with the E·F·G complex. Efficient co-precipitation of the D1 and D2 proteins with the E·F·G complex was observed only when both D proteins were present in the reaction mixture (Figure 3, compare lane 3 with lanes 4 and 5), suggesting that D1·D2 dimer formation is required for association with E·F·G. While neither D1 nor D2 were co-precipitated efficiently when incubated individually with F alone (lanes 6 and 7) or with the E·F·G complex (lanes 4 and 5), D2 was immunoprecipitated slightly above background when tested with the E·F·G complex (compare lanes 5 and 9). This might indicate that the interaction of D1·D2 with the E·F·G complex is mediated initially by D2. The D1·D2·E·F·G complex is less stable, however, than its constituent complexes, since it dissociates at ionic strengths (i.e. 500 mM KCl) at which the D1·D2 and E·F·G complexes remain intact (see above, and data not shown). By immunoprecipitation analyses with *in vitro* translated proteins we recently have provided evidence for an interaction between B′ or B and the D3 protein (Hermann *et al.*, 1995). This interaction was specific, since the D1 and D2 proteins interacted with B/B′ only weakly (for D1) or not at all (for D2). These observations are consistent with the results described above, i.e. that D1 and D2, but not D3, interact with the E·F·G heterooligomer.

Immunodepletion of the B/B′ and D3 proteins from total snRNP proteins

The existence of the various forms of RNA-free snRNP protein complexes described above raises the important question of whether these protein heterooligomers are functional intermediates in the assembly of snRNP core particles. One approach to demonstrate this would be to reconstitute snRNP particles *in vitro* with various subsets of the core proteins. Since the low protein concentration of polypeptides translated *in vitro* did not allow for efficient reconstitution, it was necessary to fractionate total snRNP proteins from isolated snRNPs so that appropriate subsets of core proteins, corresponding in their composition to the aforementioned heterooligomers, could be obtained. Subsets containing the E·F·G complex and E/F/G-depleted Sm proteins were isolated by sucrose gradient centrifugation (see Figure 1).

Further fractionation was necessary to obtain Sm protein subsets containing the D1·D2/E·F·G complex(es) or the B/B′·D3 complex. Removal of the D1·D2·E·F·G complex as such from the protein preparation was not possible due to its relative instability. Therefore, the fractionation of the total proteins was only possible by depletion of the B/B·D3 complex, rather than the D1·D2·E·F·G complex. As a prerequisite for this, the stability of the B/B′·D3 complex was first analyzed by immunoprecipitation with an anti-B/B′ mAb, KSm5, using ^{35}S-labeled polypeptides translated *in vitro* (Figure 4). In the presence of B′, D3 was co-precipitated even under stringent wash conditions (i.e. 400 mM KCl, lanes 5 and 6; 750 mM KCl, lanes 7 and 8); thus, the B/B′·D3 complex is highly stable. B′

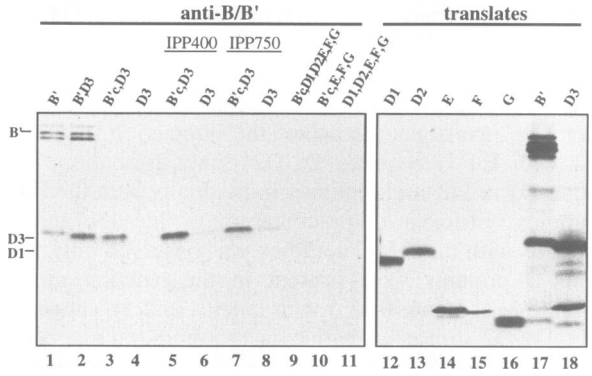

Fig. 4. Co-immunoprecipitation of D3 with B′. Proteins were prepared by *in vitro* translation and labeled with [^{35}S]methionine, with the exception of non-radiolabeled B′ polypeptide. Lanes 1–11, immunoprecipitation with the anti-B/B′ monoclonal antibody KSm5 with B′ (lane 1), B′ and D3 (lane 2), B′c (non-radiolabeled B′) and D3 (lane 3), D3 (lane 4), B′c and D3 (lane 5), D3 (lane 6), B′c and D3 (lane 7), D3 (lane 8), B′c, D1, D2, E, F and G (lane 9), B′c, E, F and G (lane 10) or D1, D2, E, F and G (lane 11), whereby immunocomplexes were washed either with 150 mM KCl (lanes 1–4 and 9–11), 400 mM KCl (lanes 5 and 6) or 750 mM KCl (lanes 7 and 8); lanes 12–18, proteins equivalent to 50% of the amount added to the assays. Note that the high molecular weight doublet observed for the B′ polypeptide is only observed under high-TEMED gel conditions and probably represents isoforms of B′. Fractionation was carried out by electrophoresis on a high-TEMED, SDS–12.5% polyacrylamide gel, and bands were visualized by fluorography.

does not interact stably with any of the other core proteins (lanes 9–10). Moreover, the mAb KSm5 reacts exclusively with B/B′, not with the other Sm proteins (D3, lane 4, and D1, D2, E, F and G, lane 11).

Given these results, we fractionated HeLa snRNP proteins by immunoaffinity chromatography with mAb KSm5; aliquots of each fraction from the immunoaffinity column were analyzed by SDS–PAGE (Figure 5A). The flow-through (lane 2) and wash fractions (lanes 3–8) of the KSm5 column contained the core proteins D1, D2, E, F and G, as well as the specific proteins A, C, A′ and B″, but no B/B′, and only minimal levels of D3. The eluate (lanes 9–12), on the other hand, contained exclusively B, B′ and D3; co-retention of D3 on the anti-B/B′ column is consistent with a strong B/B′·D3 interaction (see Figure 4). To distinguish between the multiple proteins migrating in the 28–30 kDa region (B, B′, A′ and B″) of the SDS gel, Western blotting was carried out with total snRNP proteins (Figure 5B, lanes 1–3), the flow-through fractions (lanes 4–6) and the eluate fractions (lanes 7–9), using either the monoclonal anti-B″ antibody (4G3), anti-B/B′ KSm5 or a polyclonal anti-D3 antibody raised against recombinant D3 protein. These results confirmed the absence of B and B′ in the flow-through fractions (lane 5) and of B″ in the eluate fractions (lane 7). While D3 was still present to a low degree in the flow-through (Figure 5B, lane 6), the majority of this protein was found in the eluate (lane 9) (the additional bands result from the low cross-reactivity of anti-D3 with denatured D1 and D2). Using anti-B/B′ affinity chromatography, we could thus separate B, B′ and most of D3 efficiently from the remaining core proteins.

E, F and G are necessary but not sufficient for the formation of a stable U1 RNP complex

An important question in the snRNP core assembly pathway is whether all core proteins interact concurrently during RNP complex formation or, alternatively, whether subsets of them can bind to create stable, intermediate RNP 'subcore' particles. To address this question, snRNP particles were reconstituted with the various mixtures of native snRNP core proteins, obtained by anti-B/B′ affinity chromatography and sucrose gradient centrifugation as described above, and 3′-end-labeled U1 snRNA that had been isolated from HeLa cells. Formation of stable U1 RNA–core protein complexes was monitored by co-precipitation of U1 snRNA with antibodies targeted against different core proteins (note that the mAb Y12 recognizes not only the E·F·G complex in immunoprecipitation, as mentioned above, but also the individual D1, D3, B and B′ proteins). The proteins fractionated by sucrose gradient centrifugation were tested initially in this RNP complex formation assay. As shown in Figure 6, U1 RNP formation was observed if reconstitution was performed with the complete set of core proteins; U1 was co-precipitated by both mAb Y12 (lane 3) and anti-F (lane 10). However, neither proteins depleted of E, F and G (lane 1) nor solely the E·F·G complex (lanes 4, 5, 11 and 12) were sufficient for stable RNP formation. U1 core formation was observed, however, when both the E·F·G complex and the E-, F- and G-depleted proteins (i.e. B′, B, D1, D2 and D3) were combined (lanes 2 and 9). In addition, the presence of sucrose was shown to have no effect on the reconstitution of core U1 snRNPs and their subsequent immunoprecipitation; the U1 snRNA was co-precipitated equally well by Y12 or anti-F after reconstitution with total HeLa snRNP proteins either in the presence or absence of sucrose (compare lanes 6 and 7, or lanes 13 and 14, respectively). Examination of the RNA in the immunoprecipitation supernatants demonstrated that the U1 snRNA remained intact throughout the assays (Figure 6, bottom panels; the low amount of U1 in the supernatant of lanes 6 and 7 is due to the nearly quantitative precipitation of the particle by Y12). We conclude from these results that E, F and G alone are not sufficient for the formation of a stable U1 RNP complex, although they must be present in order for RNP formation ultimately to occur.

Formation of a stable U1 RNP subcore particle with the D1, D2, E, F and G proteins

We next tested whether the proteins present in the RNA-free, D1·D2·E·F·G complex (Figure 3) were sufficient to form a U1 snRNP subcore particle. Core protein fractions containing either D1, D2, E, F and G (flow-through, FT) or B, B′ and D3 (eluate) were provided by anti-B/B′ immunoaffinity chromatography with the mAb KSm5 (see Figure 5). U1 snRNAs reconstituted with all of the core proteins (TP) could be co-precipitated by Y12 (Figure 7A, lane 1). As would be expected from the data presented above (see Figure 6), the B, B′ and D3 proteins alone did not bind U1 snRNA (eluate, lane 3). However, U1 snRNA was co-precipitated by Y12 after incubation with the proteins from the flow-through (D1, D2, E, F and G; lane 2) (note that an ~1000-fold molar excess of non-specific competitor tRNA was included in each reconstitution

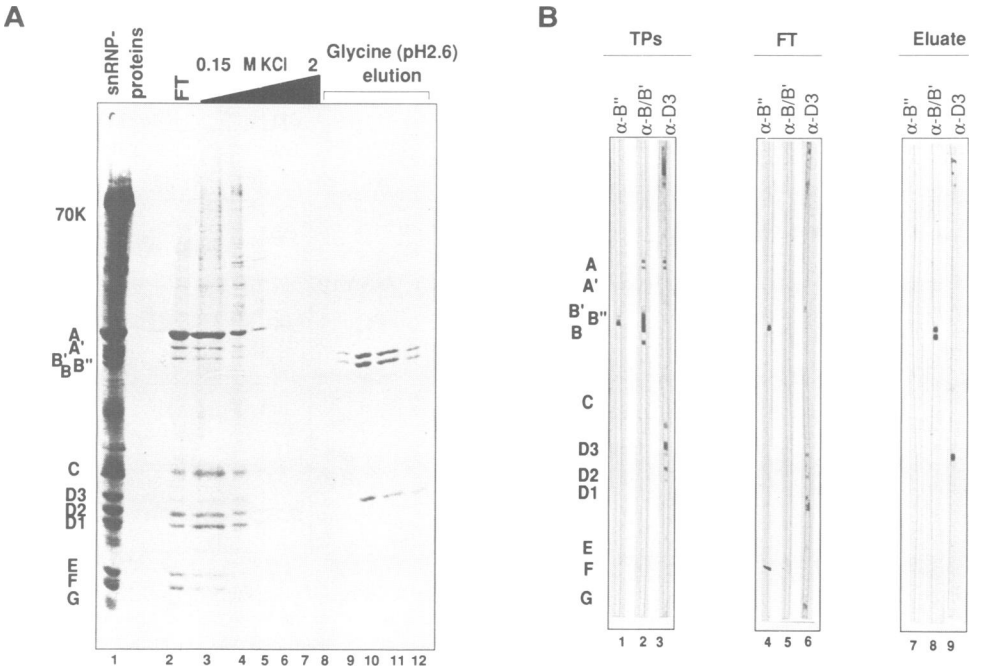

Fig. 5. Depletion of B/B′ and D3 proteins from total snRNP proteins by anti-B/B′ immunoaffinity chromatography. (**A**) SnRNP proteins separated by anti-B/B′ immunoaffinity chromatography with the mAb KSm5. Lane 1, snRNP proteins shown as a marker; lane 2, the column flow-through (FT); lanes 3–8, proteins washed from the column with increasing concentrations of KCl (0.15–2 M), as described in the Materials and methods; and lanes 9–12, proteins eluted from the column with 100 mM glycine (pH 2.6). One fifth (lanes 3–12) or 1/15 (lane 2) of each fraction was precipitated with acetone and the proteins analyzed on a high-TEMED, SDS–12.5% polyacrylamide gel and visualized with Coomassie blue staining. (**B**) Western blot analysis of the protein composition of the fractions obtained by immunoaffinity chromatography. Lanes 1–3, total snRNP proteins (TP), shown as a marker; lanes 4–6, column flow-through (FT); and lanes 7–9, column eluate. Antibodies used were mAb anti-B″ 4G3, lanes 1, 4 and 7; mAb anti-B/B′ KSm5, lanes 2, 5 and 8; and a polyclonal antibody against D3, lanes 3, 6 and 9.

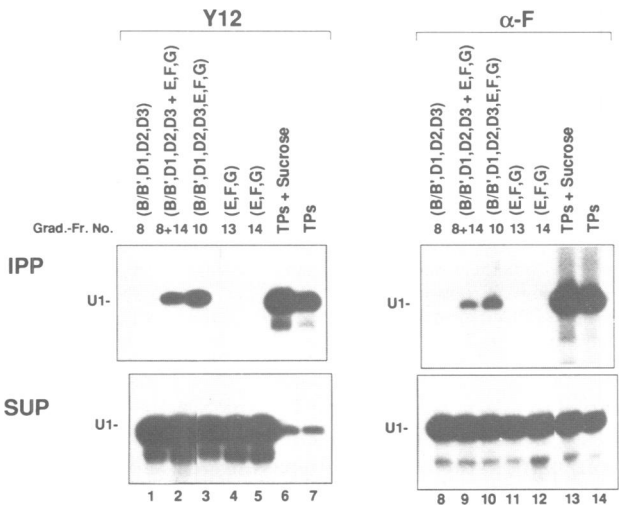

Fig. 6. The E, F and G core proteins are necessary but not sufficient for stable U1 snRNP complex formation *in vitro*. Reconstitution *in vitro* was carried out with [^{32}P]pCp-labeled U1 snRNA, isolated from HeLa cells, and HeLa snRNP proteins, previously separated on a 6–20% sucrose gradient (see Figure 1) as follows: gradient fraction 8 (B, B′, D1, D2 and D3; lanes 1 and 8); gradient fractions 8 and 14 (B, B′, D1, D2 and D3 plus E, F and G; lanes 2 and 9), gradient fraction 10 (all core proteins; lanes 3 and 10); or gradient fractions 13 or 14 (exclusively E, F and G; lanes 4, 5, 11 and 12). Total HeLa snRNP proteins (TP), which had not been separated on a sucrose gradient, were tested in reconstitution with (lanes 6 and 13) or without (lanes 7 and 14) the addition of 8% sucrose (final concentration). U1 RNP formation was analyzed by immunoprecipitation with either mAb Y12 (anti-Sm), lanes 1–7, or a polyclonal anti-F antibody, lanes 8–14. RNA was analyzed on a 10% polyacrylamide–7 M urea gel and visualized by autoradiography.

assay). This strongly suggests that, while the B/B′ and D3 proteins do not bind to the U1 snRNA on their own, the D1, D2, E, F and G core proteins are capable of assembling onto the U1 snRNA independently of B/B′ and D3. We have termed this putative Sm core RNP assembly intermediate a 'subcore' particle.

To verify its protein composition, the U1 subcore particle was analyzed further by immunoprecipitation with antibodies reactive against various core proteins. As tested by immunoprecipitation assays with ^{35}S-labeled, *in vitro* translated core proteins, the polyclonal anti-D2, anti-D3 and anti-F sera react specifically with D2, D3 or F, respectively (Figure 7B, lanes 8–24), whereas the mAb KSm5 recognizes specifically B and B′ (see Figure 4). All of these antibodies can precipitate isolated, native U snRNP particles efficiently, with the exception of the anti-D3 antiserum, for which the precipitation is very inefficient (data not shown). The subcore particle could be precipitated by anti-F (Figure 7A, lane 10) and anti-D2 (lane 14). However, no or minimal co-precipitation of the U1 snRNA was observed with the anti-B/B′ antibody (Figure 7A, lane 6) or anti-D3 (lane 18). This confirms that the particle formed from the FT fractions is indeed a 'subcore' particle that lacks B, B′ and D3 but contains D2 and F, and, presumably, the remaining core proteins, D1, E and G. The presence of the latter proteins could not be demonstrated directly by immunoprecipitation because antibodies specifically reacting with either D1, E or G were not available.

To investigate whether the subcore particle is a dead-end complex or a functional intermediate in the core particle assembly pathway, subsequent binding of B, B′

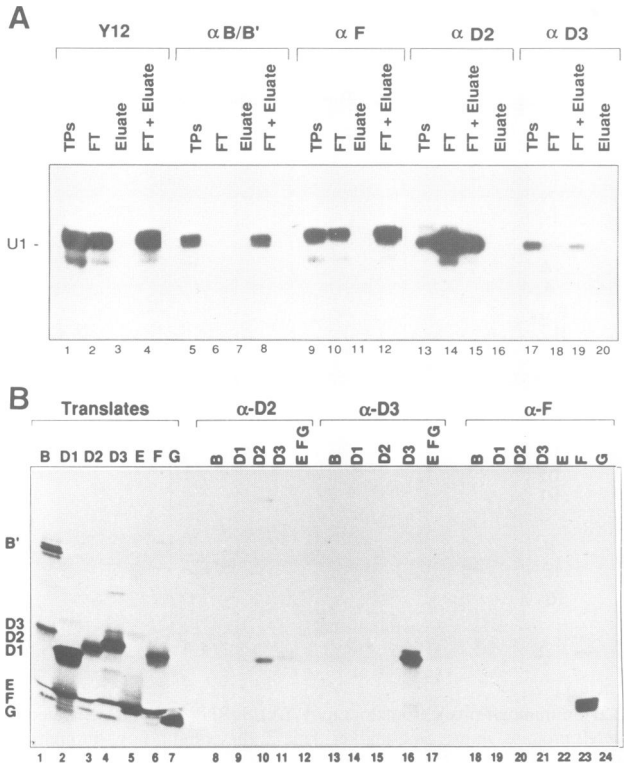

Fig. 7. Protein composition of U1 subcore snRNP particles, reconstituted *in vitro* from U1 snRNA and B/B′/D3-depleted snRNP proteins, as analyzed by immunoprecipitation. (**A**) [^{32}P]pCp-labeled U1snRNA was incubated with either total HeLa snRNP proteins (TP; lanes 1, 5, 9, 13 and 17) or proteins from an anti-B/B′ affinity column (shown in Figure 5) obtained from either the flow-through fraction (FT; lanes 2, 6, 10, 14 and 18), the eluate fraction (lanes 3, 7, 11, 16 and 20), or flow-through and eluate (FT + eluate; lanes 4, 8, 12, 15 and 19); the latter was performed in a two-step reconstitution as described in Materials and methods, with an initial incubation of the RNA with flow-through proteins, followed by the addition of the eluate proteins. The flow-through fraction contained the D1, D2, E, F and G core proteins, while the eluate contained the B, B′ and D3 core proteins. Immunoprecipitation was performed with the anti-Sm mAb Y12 (lanes 1–4), anti-B/B′ mAb KSm5 (lanes 5–8), or polyclonal anti-F (lanes 9–12), anti-D2 (lanes 13–16) and anti-D3 antibodies (lanes 17–20). Co-immunoprecipitated RNA was separated on a 10% polyacrylamide–7 M urea gel and visualized by autoradiography. (**B**) Antibody specificity was tested by immunoprecipitation with *in vitro* translated, ^{35}S-labeled proteins. Lanes 1–7, proteins equivalent to 50% added to the assays; lanes 8–24, immunoprecipitation with an anti-D2 polyclonal antibody (lanes 8–12), an anti-D3 polyclonal antibody (13–17) or an anti-F polyclonal antibody (lanes 18–24). *In vitro* translated proteins added to the immunoprecipitation assays are indicated at the top of each lane.

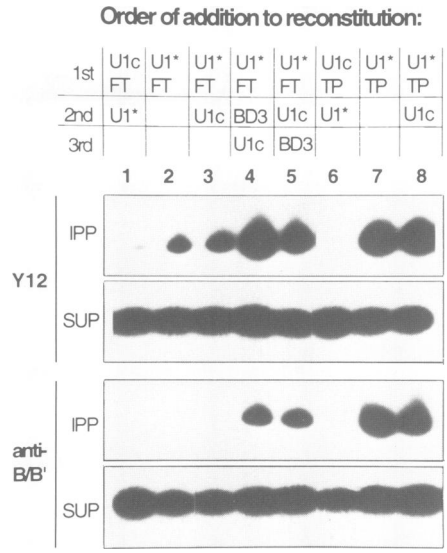

Fig. 8. Effect of competitor U1 snRNA on U1 subcore particle stability and on the stepwise assembly of core particles. An *in vitro* two-step reconstitution was performed with 0.1 pmol of [^{32}P]pCp-labeled U1 snRNA (U1*), 40 pmol of unlabeled competitor U1 snRNA (U1c), and HeLa snRNP proteins in a 10-fold molar excess over ^{32}P-labeled U1 snRNA. All mixtures were submitted to two consecutive 45 min reconstitution incubations. The order of addition is given in the flow chart at the top ('1st' and '2nd' indicate addition at the beginning of the first or second reconstitution incubations, while '3rd' indicates addition after a 5 min lapse after the beginning of the second incubation). In the initial reconstitution (1st), proteins from the anti-B/B′ immunoaffinity column flow-through fraction (FT; see Figure 5) or total proteins (TP) were added to either hot or cold U1 snRNA to reconstitute subcore (lanes 1–5) or core particles (lanes 6–8). Competitor U1 snRNA was added either to the first reconstitution (lanes 1 and 6) or to the second reconstitution (lanes 3–5 and 8). For the stepwise assembled core particles, proteins from the anti-B/B′ affinity column eluate fraction containing B/B′·D3 (denoted here 'BD3') were added to the subcore either 5 min prior to or subsequent to the addition of cold U1 snRNA (lanes 4 and 5, respectively). Each reconstitution assay was analyzed by immunoprecipitation with the mAb Y12 (top two panels; IPP: co-immunoprecipitated RNA, SUP: supernatant) and the mAb KSm5 (anti-B/B′; bottom two panels). Note that the increase in Y12 precipitation of core as compared with subcore particles (lanes 4 and 3, respectively) could reflect either an increase in particle stability or, alternatively, an improvement in Y12 recognition. Aliquots of the supernatants were analyzed to control for possible RNA degradation. RNA was separated on a 10% polyacrylamide–7 M urea gel and visualized by autoradiography.

and D3 to the subcore was assayed by immunoprecipitation after *in vitro*, two-step reconstitution. Upon addition of the B, B′ and D3 core proteins (eluate, Figure 5) to the U1 subcore, U1 snRNA was immunoprecipitated with both anti-B/B′ and anti-D3 antibodies, demonstrating that the B/B′ and D3 proteins were bound to the U1 snRNP (Figure 7A, lanes 8 and 19). In contrast, co-precipitation of U1 snRNA with anti-B/B′ or anti-D3 was not observed with the subcore (Figure 7A, lanes 6 and 18) or the eluate alone (lanes 7 and 20). Thus, a complete core particle could be formed from the subcore intermediate upon addition of B/B′ and D3.

The possibility remained that the subcore is not a stable assembly intermediate, but that the D1, D2, E, F and G proteins dissociate from the U1 snRNA in the presence of B, B′ and D3 and re-assemble on the RNA to form a core particle directly, thus bypassing the subcore in core assembly. To rule this out, we tested whether the pre-assembled subcore particle remained stable when mixed with competitor U1 snRNA in a reconstitution assay, either in the absence or presence of the B/B′·D3 complex (Figure 8). We first verified that a 400-fold molar excess of cold U1 snRNA over ^{32}P-labeled U1 snRNA was sufficient to titrate out the Sm proteins from either a mixture of total proteins (TP) or the D1/D2/E/F/G proteins (FT). When the Sm proteins were incubated first with competitor U1 snRNA, with the subsequent addition of radiolabeled U1 snRNA, no subcore or complete core U1 snRNP particles containing hot U1 snRNA could be detected by immunoprecipitation with the mAb Y12 (Figure 8, lanes 1 and 6, respectively). As expected, when hot U1 snRNA was first reconstituted with total Sm

proteins, the subsequent addition of competitor U1 snRNA did not dissociate the pre-formed core U1 snRNPs (lane 8; compare with lane 7, without competitor). Most importantly, the pre-formed, radiolabeled subcore particle likewise remained intact in the presence of competitor U1 snRNA (lane 3, compare with lane 2, without competitor; top panel). Moreover, the core particle assembled stepwise from the subcore and the B/B′·D3 complex was also stable in the presence of cold U1 snRNA, as shown by immunoprecipitation with the anti-B/B′ mAb KSm5 (lane 4, third panel). Finally, the addition of competitor U1 snRNA to pre-formed subcore particles prior to the addition of the B, B′ and D3 proteins did not prevent the formation of complete core particles (lane 5, third panel), demonstrating that the pre-formed subcore particles were chased into core particles despite the presence of free U1 snRNA. Taken together, our data strongly suggest that the subcore particle is indeed a genuine intermediate in snRNP core assembly *in vitro*. Additionally, the assembly of core from subcore particles in the presence of free U1 snRNA provides further evidence that the B/B′·D3 complex alone does not associate with U1 snRNA (see also Figure 6) but instead requires the pre-assembly of the snRNP subcore intermediate.

Finally, we investigated the sedimentation behavior of *in vitro* reconstituted subcore and core particles on sucrose gradients (5–20%). U1 core particles were reconstituted either in one step with total snRNP proteins (TP) or in two steps, adding first D1, D2, E, F and G (FT), and then B/B′ and D3 (eluate), to U1 snRNA. As a control, the sedimentation of ^{32}P-labeled U1 snRNA alone was determined. As shown in Figure 9A, naked U1 snRNA was found predominantly in fractions 8–10 (~5–6S). In contrast, both *in vitro* reconstituted U1 snRNP core particles (from either TP or FT + eluate), as well as the subcore particles, sedimented faster, peaking in fraction 12 (~8–9S). The identity of the RNP particles peaking in fraction 12 was confirmed subsequently by immunoprecipitation analyses. Immunoprecipitation with Y12 showed that RNP particles in the 8–9S range contain core proteins (Figure 9B, lanes 2 and 5). Immunoprecipitation with the anti-B/B′ mAb KSm5 verified that RNP particles in the subcore peak fractions did not contain B or B′ (Figure 9B, lane 6). In contrast, the core particles reconstituted in two steps with FT and eluate fractions were precipitated efficiently with mAb KSm5 (lane 9). These results provide further evidence for the relative stability of the subcore particle, since it withstands sucrose gradient centrifugation. That the partial and complete RNP particles co-sediment is not unique to the snRNP subcore and core particles but is also observed in other cases; for example, yeast SRP subparticles lacking either the Srp65 or Srp65/Srp54 proteins had sedimentation velocities comparable with that of the full SRP particle (Brown *et al.*, 1994).

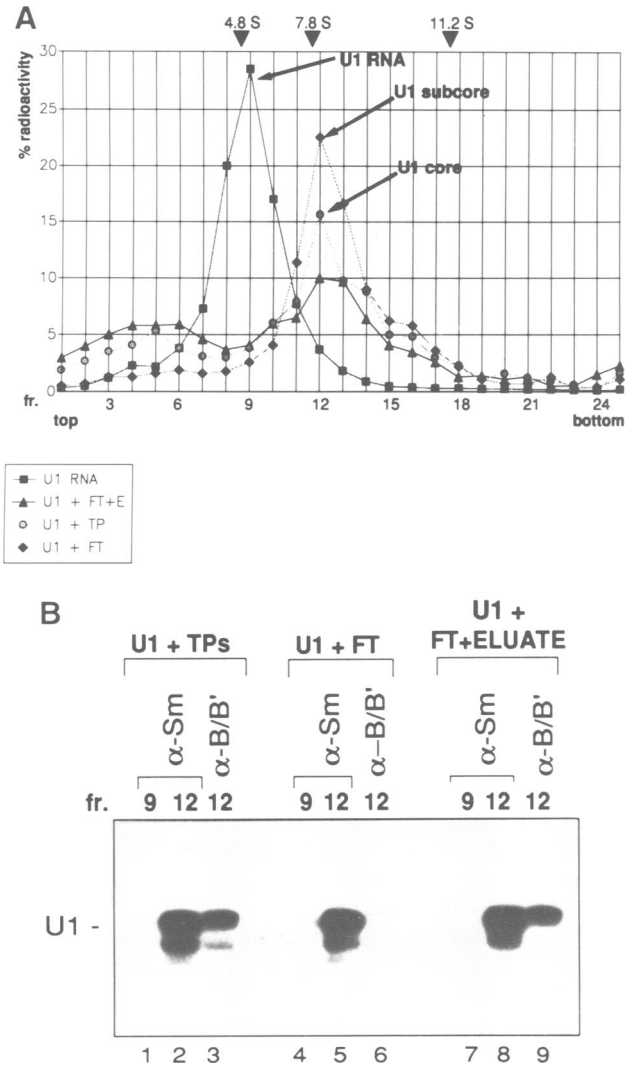

Fig. 9. Sedimentation analysis of *in vitro* reconstituted U1 snRNP subcore particles. (**A**) Sedimentation of U1 snRNA, U1 snRNP subcore particles and U1 snRNP core particles on a 5–20% sucrose gradient. Reconstitution *in vitro* was carried out with [^{32}P]pCp-labeled U1 snRNA and, for the core snRNP particles, either total HeLa snRNP proteins (TP), or flow-through (FT) and eluate, and for the subcore particle, only flow-through proteins. Flow-through and eluate fractions were obtained from anti-B/B′ affinity chromatography (shown in Figure 5) and contained D1, D2, E, F and G, or B, B′ and D3, respectively. The percentage of radioactivity in each fraction (in c.p.m.; left axis) is plotted against the fraction number (bottom axis), whereby the total radioactivity of each gradient was defined as 100% c.p.m. Sedimentation was from left to right. The sedimentation peaks of marker proteins fractionated on parallel gradients are indicated by triangles at the top. (**B**) Confirmation of the protein composition of the RNP particles fractionated on sucrose gradients by immunoprecipitation with the mAb Y12 (α-Sm; lanes 1, 2, 4, 5, 7 and 8), or with the mAb KSm5 (α-B/B′; lanes 3, 6 and 9). From each of the three RNP sucrose gradients (U1 + TP, U1 + FT and U1 + FT + eluate; see above), approximately equal amounts (c.p.m.) of either fraction 9, which contained free RNA, or fraction 12, which contained RNP particles (core RNPs, lanes 1–3 and 7–9; subcore RNPs, lanes 4–6), were analyzed.

Two-step reconstitution yields core U1 snRNP particles functional in a cap hypermethylation assay

In the last set of experiments, we investigated whether the core U1 snRNP reconstituted in two steps via the subcore was functionally active. As the concentration of core particles reconstituted under these conditions was too low to investigate their splicing activity, a recently established *in vitro* hypermethylation assay was employed to test the activity of the *in vitro* reconstituted particles (Plessel *et al.*, 1994). In this assay, 5′ cap hypermethylation (from m^7G to m_3G) can be induced when HeLa cytosolic extract, *S*-adenosyl methionine (SAM), and snRNP par-

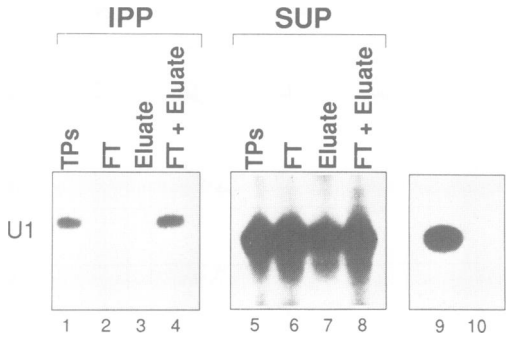

Fig. 10. Cap hypermethylation of *in vitro* reconstituted U1 snRNP particles. ^{32}P-Labeled, m^{7}G-capped U1 snRNA, transcribed *in vitro*, was incubated with either TPs (lane 1) or proteins in the flow-through fraction (lane 2), the eluate (lane 3) or the flow-through followed by the eluate fraction (lane 4), of the anti-B/B′ affinity column (shown in Figure 5). The reconstituted U1 RNP particles were then incubated in the presence of 100 µl of S100 extract (2 mg/ml) and 0.1 mM SAM at 37°C for 1 h. RNA was extracted and immunoprecipitated with the m$_3$G cap-specific antibody, R1131. The corresponding supernatants from each assay are shown in lanes 5–8. Antibody specificity was controlled by immunoprecipitation with m$_3$G- (lane 9) and m^{7}G-capped (lane 10) U1 snRNA. RNA was analyzed on a 10% polyacrylamide–urea gel and visualized by autoradiography.

ticles, reconstituted *in vitro* with m^{7}G-capped U1 snRNA, are incubated together. As the B/B′ proteins appear to be involved in hypermethylation (Plessel *et al.*, 1994), we were also curious as to whether the subcore, which lacks B, B′ and D3, could support cap hypermethylation and, if not, whether the addition of the B/B′ and D3 proteins would restore this activity.

For the functional analysis of *in vitro* reconstituted Sm core particles, m^{7}G-capped, ^{32}P-labeled U1 snRNA, transcribed *in vitro*, was incubated with either total snRNP proteins (TP), B/B′/D3-depleted proteins (see Figure 5), or B/B′·D3-depleted proteins followed by B/B′ and D3 (i.e. in two separate steps). Cap hypermethylation was then assayed by immunoprecipitation with the m$_3$G cap-specific rabbit antibody, R1131, which does not precipitate m^{7}G-capped RNA or RNP particles. As expected, snRNP core particles (U1 + TP) were hypermethylated efficiently (Figure 10, lane 1). Interestingly, the U1 subcore particle (U1 + FT) did not support hypermethylation, thereby demonstrating that it is indeed functionally distinct from the U1 core particle (lane 2). The B/B′ and D3 proteins alone, which are not capable of binding to U1 snRNA, also did not support U1 cap hypermethylation (lane 3). However, two-step reconstitution with the subcore plus B/B′ and D3 resulted in the formation of a core particle active in hypermethylation (lane 4). Thus, these results clearly demonstrate that the two-step assembly of the snRNP core particle leads to a functionally active particle. They further suggest that the *in vitro* Sm core assembly pathway described here is relevant to the assembly process *in vivo*, and that the B/B′ and D3 proteins contribute either directly or indirectly to the m$_3$G cap hypermethylation reaction.

Discussion

We have investigated in detail interactions among the various snRNP core proteins and their subsequent interaction with U1 RNA to form the core RNP structure of the U1 snRNP. We demonstrate that the assembly of the snRNP core is an ordered process which involves the initial formation of several heteromeric protein complexes and an intermediate snRNP 'subcore' particle.

Existence of multiple protein–protein interactions in the snRNP core

The structure and stability of the snRNP core has been proposed to be determined predominantly by protein–protein interactions (see Introduction). This idea is supported further by our finding that each of the eight Sm proteins is involved in one or more RNA-free heteromeric complexes. In addition to the previously reported D1·D2 (Lehmeier *et al.*, 1994), B·D3 and B′·D3 complexes (referred to here as B/B′·D3; Hermann *et al.*, 1995), we have identified, by immunoprecipitation analyses with *in vitro* translated Sm proteins, complexes of E·F, E·G, E·F·G and D1·D2·E·F·G. Each of these complexes remains intact at moderate salt concentrations (e.g. 250 mM KCl), which is consistent with the idea that they are formed not only *in vitro* but also *in vivo*. The E·F·G and B/B′·D3 complexes are extremely stable, withstanding buffers containing up to 750 mM KCl (see Figure 4 for B/B′·D3; data not shown for E·F·G). Since the E·F·G protein complex is much more stable than the E·F or E·G complex, the majority of the E, F and G proteins in the cytoplasm are presumably present in an E·F·G complex prior to their association with UsnRNA.

An E·F·G complex was also detected in RNA-free, native snRNP protein preparations which were generated by dissociation of snRNP particles in the presence of EDTA and the anion exchange resin, DE53. The existence of E, F and G as a multimeric complex in these preparations is supported by their sedimentation behavior on sucrose gradients; native E, F and G proteins co-sediment, peaking at 3.7S on the gradient (see Figure 1). Assuming that they form a globular complex, the estimated M_r of this E·F·G heterooligomer, based on its sedimentation coefficient, is ~56 kDa. Since the calculated M_r of an E·F·G trimer is 29 kDa, the sedimentation behavior of this complex suggests that it contains multiple copies of E, F and G. Analysis of the protein composition of the 3.7S complex by gel electrophoresis and silver staining (data not shown) indicated that equal molar amounts of each polypeptide are present; based on the relative radioactive intensity of its components, the E·F·G complex detected with *in vitro* translated proteins (see Figure 2B) also appears to contain stoichiometric amounts of E, F and G. Given its sedimentation coefficient and the apparent stoichiometric relationship of its constituents, we propose that the E·F·G multimer observed with native snRNP proteins contains two copies of each polypeptide and, thus, is a hexameric complex.

That the E·F·G multimer is a functional, core snRNP assembly intermediate is supported by several observations. The extreme stability of this complex *in vitro* suggests that it remains intact during assembly of the snRNP core. Furthermore, the 3.7S E·F·G complex obtained from sucrose gradient fractionation could be integrated into a U1 snRNP core particle by *in vitro* reconstitution with U1 snRNA and the remaining Sm proteins (see Figure 6). Additionally, electron microscopic analyses of the RNA-free E·F·G multimer revealed that its ultrastructure strongly resembles that of fully formed

snRNP core particles (Plessel *et al.*, in preparation). It is thus conceivable that the E·F·G multimer provides a structural platform for snRNP core assembly, and that the ultrastructural similarity of U1, U2, U4 and U5 snRNP cores observed by electron microscopy (Lührmann *et al.*, 1990) reflects the presence of the E·F·G multimer in each of these particles.

Previous *in vivo* pulse–chase studies identified an RNA-free complex containing E, F, G and one or more D protein, which sedimented as a 6S particle on sucrose gradients (Fisher *et al.*, 1985; Sauterer *et al.*, 1990). This complex is likely to be the D1·D2·E·F·G complex which we have detected by co-precipitation studies with *in vitro* translated proteins. While the D1·D2·E·F·G complex was not observed after sucrose gradient centrifugation of native HeLa snRNP proteins, the *in vivo* complex may have been stabilized by the presence of additional cellular factors which are lacking in our purified snRNP protein preparations. Determination of the sedimentation coefficient of the D1·D2·E·F·G complex formed with *in vitro* translated polypeptides was not possible, as the high amount of globin present in the reticulocyte lysate used for translation interfered with sedimentation.

The identification of numerous protein–protein interactions among the Sm proteins suggests that several of these proteins contain multiple protein interaction domains. Two evolutionarily conserved sequence motifs which are common to all of the Sm proteins have been shown to be involved in at least two of the Sm protein interactions. These so-called Sm motifs 1 and 2 are a hallmark of the Sm protein family (Cooper *et al.*, 1995; Hermann *et al.*, 1995; Séraphin, 1995). The Sm motifs, which contain few absolutely conserved residues, but rather related amino acids, have been proposed to mediate the proper folding necessary for interactions among the Sm proteins (Hermann *et al.*, 1995; for review, see Mattaj and Nagai, 1995). Indeed, truncation of the Sm motifs in either B′ or D3 (Hermann *et al.*, 1995) or amino acid substitution of a highly conserved residue in Sm motif 1 of the G polypeptide (M.Jahn, A.Fischer and R.Lührman, unpublished data) resulted in the loss of B′·D3 or E·F·G complex formation, respectively. Although it is likely that the Sm motifs play a role in other Sm protein interactions, the specificity of the Sm protein interactions suggests that they are mediated by less conserved residues within or adjacent to the Sm motifs or by the remaining regions of the Sm proteins. In the case of E, F and G, almost the entire protein is comprised of the Sm motifs, and thus protein–protein interactions involving these proteins most certainly are mediated by residues of the Sm motifs 1 and 2. Since E, F and G interact specifically with one another (e.g. F and G each interact with E, but not with each other), non-conserved residues in and around these conserved motifs must also play a decisive role.

Interestingly, the protein–protein interaction site(s) required for formation of a stable D1·D2·E·F·G complex appear to be generated only after the interaction of D1 with D2, and of E, F and G with each other. Specifically, the E·F·G complex interacts much more strongly with the D1·D2 complex than with either D1 or D2 alone (Figure 3). Similarly, D1·D2 interacts more efficiently with the E·F·G multimer than with any of its constituents alone (data not shown). The generation of functionally important intra- and/or intermolecular structural domains upon association of the E, F and G proteins is supported by the unexpected behavior of the prototypical anti-Sm antibody, Y12. While Y12 recognizes the individual E, F and G proteins only poorly, the E·F·G complex, on the other hand, is precipitated very efficiently by Y12; the extent of precipitation of the complex was much higher than that which could be accounted for by the collective precipitation of its individual components (see Figure 2B). Thus, the E·F·G complex, and possibly the D1·D2 complex, have intra- or intermolecular regions which are exposed or created during heterooligomerization and which are necessary for subsequent interactions with other Sm protein complexes, and ultimately with UsnRNA as well (discussed below).

The snRNP core assembly pathway

Given that the Sm proteins form several heteromeric complexes in the absence of RNA, we were interested in determining if these complexes represent actual RNP core assembly intermediates which associate as such with UsnRNA and, additionally, whether they interact concurrently or in a stepwise manner. For this purpose, HeLa snRNP proteins were fractionated by sucrose gradient centrifugation or immunoaffinity chromatography such that Sm protein subgroups which reflected the composition of the heteromeric complexes were obtained. Following *in vitro* reconstitution, the association of these defined Sm protein mixtures with U1 snRNA was assayed by immunoprecipitation. Interestingly, incubation of U1 RNA with E, F and G or with B/B′ and D3 alone did not lead to stable RNP formation (see Figures 6 and 7). Reconstitution of a stable RNP 'subcore' was observed first in the presence of D1, D2, E, F and G. Thus, the prior interaction of several Sm proteins appears to be a prerequisite for RNA binding. A similar phenomenon is observed with the mammalian signal recognition particle (SRP) proteins, SRP9 and SRP14, which bind to the SRP RNA only after forming a dimer (Strub and Walter, 1991). As in the case of SRP9 and SRP14, none of the Sm proteins contain well-defined RNA binding domains (for overview, see Hermann *et al.*, 1995). It thus seems highly likely that the RNA binding regions of the Sm proteins are intermolecular and, therefore, present only after heterooligomerization. In contrast to SRP9 and SRP14, which required only dimerization for RNA binding, the Sm proteins must form large multimeric complexes in order to associate stably with UsnRNA. At present, we cannot determine whether the D1·D2·E·F·G complex binds as such to the Sm site of the UsnRNA or whether the E·F·G multimer initially binds in a non-specific, non-stable manner that subsequently is stabilized by binding of D1·D2. In the case of the U2 snRNP-specific A′ and B″ proteins, for example, the specific and stable interaction of U2-B″ with the U2 snRNA requires its prior interaction with U2-A′ (Scherly *et al.*, 1990). In this context, it is interesting to note that addition of the 3.7S E·F·G multimer and the remaining Sm proteins to the U1 snRNP reconstitution assays without prior assembly of a D1·D2·E·F·G complex led to core particle assembly (see Figure 6). Thus, both interaction pathways are plausible and, in fact, both may be relevant *in vivo*. Cross-linking studies, which can detect less stable protein–RNA interactions than

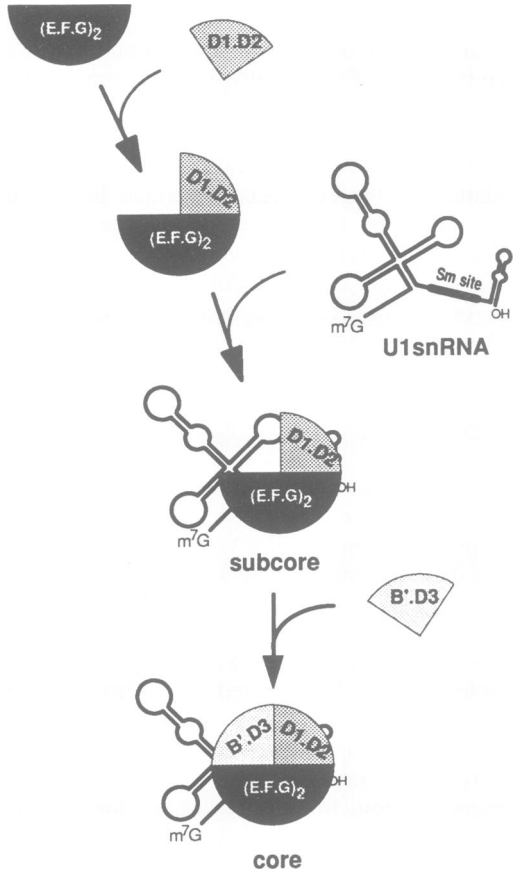

Fig. 11. Model of the probable assembly pathway of core snRNP particles. A schematic representation of core snRNP assembly based on *in vitro* observations is shown. A stable E·F·G hexamer binds to a D1·D2 dimer. This heteromeric complex is capable of interacting with U1 snRNA (as well as U2, U4 and U5 snRNA) to create the subcore particle, which is the first stable RNP particle formed along the snRNP core assembly pathway. A B′·D3 (or B·D3) complex binds the subcore to complete the assembly of the core particle. While the E·F·G complex appears to be present as a hexamer, our estimation of the number of polypeptides within the D1·D2 and B/B′·D3 complexes is speculative; the precise stoichiometry of the Sm proteins within the various protein complexes and snRNP particles remains to be elucidated.

immunoprecipitation analyses, should help to clarify the question of whether E·F·G interacts with the Sm site of U1 snRNA in the absence of other Sm proteins.

Association of the B, B′ and D3 proteins with U1 snRNA was not observed when they were incubated alone or in the presence of D1 and D2 (see Figure 7). Rather, the stable association of these proteins with the U1 snRNP required the prior formation of the D1·D2·E·F·G-containing subcore particle (see Figure 7). This observation is consistent with previous *in vivo* kinetic studies which suggested that B/B′ are the last proteins to associate during snRNP assembly (Fisher *et al.*, 1985; Feeney *et al.*, 1989). Our results thus demonstrate that the Sm protein heteromeric complexes associate in an ordered fashion during the *in vitro* assembly of the snRNP core. The most probable snRNP core assembly pathway, based on our data, is depicted schematically in Figure 11 (note that the stoichiometry of the various complexes has yet to be determined and is not intended to be represented in our cartoon). While it is conceivable that the B/B′·D3 complex also interacts with UsnRNA, in regions other than the Sm site, the general predominance of protein–protein interactions in the snRNP core suggests that the association of B/B′·D3 is mediated primarily by protein–protein contacts. The dependence of the B/B′·D3 interaction on the formation of a subcore particle indicates that the B/B′·D3 binding sites, whether protein–protein or protein–RNA in nature, are created by the prior association of the remaining Sm proteins. Several examples of cooperative binding, where the initial association of a particular protein results in conformational changes in either the RNA or protein component of an RNP particle such that new protein or RNA binding sites are created, are observed during the assembly of the SRP. For example, although the SRP19 and SRP54 proteins both interact directly with the SRP RNA, the SRP54–RNA interaction requires the prior association of SRP19 with the SRP RNA (Römisch *et al.*, 1989). Similarly, while the SRP68 and SRP72 proteins do not interact with each other in the absence of SRP RNA, an SRP68–SRP72 interaction is observed following the formation of an SRP72–RNA complex (Lütcke *et al.*, 1993). While the sequence of events during snRNP core assembly is now clear, the stoichiometry of the Sm proteins remains to be established. It is currently not clear, for example, whether the B/B′·D3 complex interacts with the UsnRNP subcore in the form of a dimer or multimer, containing either both the B and B′ proteins and/or multiple copies of D3.

Several lines of evidence suggest that the snRNP subcore particle which contains the D1, D2, E, F and G proteins is a bona fide functional intermediate in the *in vitro* snRNP core assembly pathway. Subcore formation was not restricted to U1 snRNPs, but was also observed with U2, U4 and U5 snRNA (data not shown). Based on its ability to withstand sucrose gradient centrifugation, immunoprecipitation and washes with buffers containing 300 mM NaCl, the U1 subcore appears to be a relatively stable RNP complex. Furthermore, the pre-formed U1 subcore particle could be chased into the core complex by the addition of B/B′·D3, even in the presence of a high concentration of competitor U1 snRNA (Figure 8). Finally, analyses of the U1 subcore snRNP in an *in vitro* 5′ cap hypermethylation assay demonstrated that the subcore particle does not support hypermethylation. However, consistent with it being a functional assembly intermediate, the subcore could be converted to a core U1 snRNP which is active in the cap hypermethylation assay, simply by adding the B/B′·D3 complex (Figure 10).

The ability to reconstitute the snRNP core stepwise with purified Sm proteins paves the way for a detailed investigation of the role of individual Sm proteins in the biogenesis and transport of snRNPs. The results of cap hypermethylation assays carried out with both U1 core and subcore particles (see Figure 10) demonstrate that the B, B′ and/or D3 proteins are required for the conversion of the U1 snRNA's m^7G cap to its hypermethylated m_3G form. This correlates well with previous studies, in which the B and B′ proteins appeared to be the target of iodoacetic acid modification which inhibited 5′ cap hypermethylation (Plessel *et al.*, 1994). These proteins may facilitate hypermethylation by providing a docking site for the cytoplasmic methyltransferase. It will now be possible to examine this process in more detail by initially reconstituting the

subcore particle and subsequently adding B/B′ and/or D3 mutants obtained either by *in vitro* translation or by recombinant methods. A similar strategy could also potentially be used to elucidate the NLS of the UsnRNPs, which is generated only after assembly of the UsnRNP core particle.

Materials and methods

Antibodies

Antisera specific for the D2, D3, G or F proteins were raised in rabbits according to standard immunization protocols (see also Hermann *et al.*, 1995). Additionally, the following monoclonal antibodies were used: anti-Sm Y12 (Lerner *et al.*, 1981), anti-B/B′ KSm5 (Williams *et al.*, 1986), R1131 (Lührmann *et al.*, 1982) and anti-B″ 4G3 (Habets *et al.*, 1989).

Preparation of snRNP proteins and anti-B/B′ affinity chromatography

The preparation of native snRNP proteins was carried out as described previously (Sumpter *et al.*, 1992). For immunoaffinity chromatography of the B, B′ and D3 proteins with mAb KSm5, 800 μl of ascites containing the mAb KSm5 were incubated overnight at 4°C with 1 ml of pre-swollen protein G–Sepharose (Pharmacia) in phosphate-buffered saline (PBS; 20 mM potassium phosphate and 130 mM NaCl, pH 8.0). The bound antibodies were cross-linked to protein G–Sepharose using dimethylpimelimidate (DMP) as described (Harlow and Lane, 1988). The anti-B/B′ affinity matrix was then equilibrated with PBS. For immunodepletion of the B, B′ and D3 proteins from snRNP proteins, ~300 μg of native snRNP proteins (TP, 0.3 mg/ml) were passed through the column at a rate of 1 ml/h and the column washed with 0.5 ml of PBS. Both the sample and wash fractions (~1.5 ml) were collected and denoted as flow-through (FT). The column subsequently was washed stepwise with 2 ml of wash buffer containing 50 mM HEPES–KOH, pH 7.9, 0.5 mM dithioerythrol (DTE) and increasing concentrations of KCl (i.e. 150 mM, 300 mM, 500 mM, 750 mM, 1 M or 2 M) at a rate of 4 ml/h. After a pre-elution wash with 10 ml of 10 mM KPO_4 buffer, pH 6.8, the bound proteins were eluted twice with 500 μl of 100 mM glycine, pH 2.6. The column was equilibrated subsequently with PBS. The eluates (500 μl) were collected and immediately neutralized with 1 M KPO_4 buffer, pH 8.0. Glycerol was added to a final concentration of ~5% (w/v) to all fractions before freezing in liquid nitrogen and storing at –80°C. For analysis on high-TEMED SDS–12.5% polyacrylamide gels (Lehmeier *et al.*, 1990), 1/15 of the flow-through and 1/5 of the other fractions were extracted with phenol–chloroform, precipitated by addition of 4–5 volumes of acetone and washed three times with 70% (v/v) ethanol.

For immunoblotting, ~15 μg of protein from either TP, FT or the eluted material were separated on a high-TEMED, SDS–12.5% polyacrylamide gel, transferred to nitrocellulose and immunostained as described (Lehmeier *et al.*, 1990). Incubation of nitrocellulose strips with antibodies was performed with a 1:1000 dilution of mAbs KSm5 (anti-B/B′) or 4G3 (anti-B″) and a 1:100 dilution of anti-D3 rabbit serum.

In vitro reconstitution of snRNPs

In vitro reconstitution of snRNP particles was performed essentially as described (Sumpter *et al.*, 1992). In brief, ~0.1 pmol of ^{32}P-labeled snRNA (isolated from native HeLa snRNP particles) was prepared as described (Plessel *et al.*, 1994) and then mixed with 10 μg of yeast tRNA and a 10-fold molar excess of either total snRNP proteins (0.3 mg/ml) or proteins isolated either by sucrose gradient centrifugation or anti-B/B′ immunoaffinity chromatography. The addition of equal amounts of ^{32}P-labeled U1 snRNA to each assay within an experiment was controlled by measuring the counts per min (c.p.m.). The reaction mixtures were incubated for 30 min at 30°C and then for 15 min at 37°C for reconstitution. The 'two-step' reconstitution of core from subcore particles consisted of two consecutive reconstitution incubations (each of 30 min at 30°C, 15 min at 37°C): snRNA was mixed with B/B′/D3-depleted core proteins for the first reconstitution, and the B/B′/D3 proteins were added to the second reconstitution. For the U1 competition experiments, unlabeled competitor U1 snRNA (isolated from native HeLa snRNP particles) was added in a 400-fold molar excess over ^{32}P-labeled U1 snRNA and added to the reconstitution at either 0, 45 or 50 min (described as 1st, 2nd and 3rd in Figure 8). Reconstitution mixtures were used directly for all subsequent assays.

In vitro transcription and translation

Plasmids containing the cDNAs encoding D1, D2, D3, E, F or G were linearized with *Sal*I, and the plasmid with the B′ cDNA was linearized with *Hin*dIII (restriction enzymes from New England Biolabs). The linearized plasmids were used as templates for transcription with T7 RNA polymerase. One microgram of each *in vitro* transcribed mRNA was translated with rabbit reticulocyte lysate in the presence of [^{35}S]-methionine (Amersham) in a total reaction volume of 75 μl, according to the manufacturer's instructions (Promega Biotech; see also Lehmeier *et al.*, 1994; Hermann *et al.*, 1995). A 3 μl aliquot was analyzed by high-TEMED, SDS–PAGE and the bands visualized by autoradiography.

Protein–protein RIPAs

For the protein–protein interaction studies (shown in Figures 2, 3 and 4), individual *in vitro* translated proteins (2–10 μl each) were mixed, incubated for 15 min at 30°C, followed by 30 min at 25°C, and then, in the presence of the appropriate antibody, for an additional 30 min at 25°C. The mixture was then added to 10 μl of protein A–Sepharose (Pharmacia), which had been pre-incubated for 2 h with PBS containing 1% bovine serum albumin (BSA). Incubation was continued in a total volume of 400 μl of PBS for 2 h at 4°C with constant mixing. The Sepharose-bound antibodies were pelleted and washed five times with 1 ml of IPP (10 mM Tris–HCl, pH 8.0, 0.1% NP-40, and 150 mM KCl, unless stated otherwise in the figure legend). Proteins were fractionated on high-TEMED, SDS–12.5% polyacrylamide gels, which were treated subsequently with Amplify (Amersham) and the bands visualized by fluorography, normally for 12 h.

Protein–RNA RIPAs

Antibodies used for immunoprecipitation were coupled initially to pre-swollen protein A–Sepharose (Pharmacia) in PBS for 12 h at 4°C. Particles reconstituted with [^{32}P]pCp-labeled snRNAs and proteins (described above) were incubated with antibody-coupled protein A–Sepharose in the presence of PBS containing 1% RNase-free BSA in a final volume of 400 μl. After a 2 h incubation at 4°C with constant mixing, the Sepharose-bound antibodies were pelleted and washed four times with IPP buffer containing 300 mM NaCl; in order to determine if the RNA remained intact during incubation, an aliquot from the supernatant was removed after the initial centrifugation. RNA was then extracted from the pellet and supernatant aliquot with phenol–chloroform, and precipitated with ethanol using sodium acetate as a carrier. RNA was separated on 10% polyacrylamide gels containing 7.0 M urea and visualized by autoradiography.

Sucrose gradient fractionation

Reconstituted RNP samples were mixed with reconstitution buffer (20 mM HEPES–KOH, pH 7.9, 50 mM KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTE) to a final volume of 200 μl before loading onto the sucrose gradients; native HeLa snRNP proteins (~35 μg) in 200 μl of reconstitution buffer were incubated at 30°C for 15 min and then applied directly to the sucrose gradient. Reconstituted RNP samples were fractionated on 5–20% (w/v) sucrose gradients prepared with PBS containing 0.01% (v/v) NP-40; snRNP proteins were fractionated on 6–20% (w/v) sucrose gradients prepared with buffer containing 20 mM KPO_4, pH 8.0 and 300 mM KCl. Centrifugation was carried out at 4°C for either 20 h at 39 000 r.p.m. in an SW40 rotor (for RNP particle gradients and their respective standards) or 26 h at 45 000 r.p.m. in a TLS55 rotor (for snRNP protein gradients and their respective standards). Gradients were fractionated manually from top to bottom, with a total of 20 fractions. Fractions containing ^{32}P-labeled RNA were analyzed by scintillation counting. Gradient fractions containing the sedimentation standards or the snRNP proteins were precipitated with 5 volumes of acetone and the protein analyzed by high-TEMED, SDS–PAGE. Sedimentation standards used were cytochrome c (2.3S), creatine kinase (4.8S), aldolase (7.8S) and catalase (11.2S). An approximate calculation of the M_r of the E·F·G complex from the 3.7S sedimentation coefficient was obtained by the formula $S = 0.00242 \times \text{mol. wt}^{0.67}$, which is based on the sedimentation of a standard globular protein (Eason, 1984) and is only intended to provide a rough estimate for our purposes (taking into account that the E·F·G complex is not likely to have a globular shape).

Cap hypermethylation

Analysis of cap hypermethylation of *in vitro* reconstituted particles in HeLa S100 extract was carried out essentially as described by Plessel *et al.* (1994).

Acknowledgements

We would like to thank J.Steitz, I.W.Mattaj, W.van Venrooij, D.Williams, T.Lehmeier and H.Hermann for their generous gifts of antibodies. We are grateful to M.Wicke, D.Meyer and S.Becker for antibody and HeLa snRNP protein preparations, C.L.Will for critical reading of the manuscript, and B.Kastner and H.Brahms for valuable discussions. This work was supported by grants from the Bundesministerium für Forschung und Technologie, the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References

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

Brown,J.D., Hann,B.C., Medzihradszhy,F., Miwa,M., Burlingame,A.L. and Walter,P. (1994) Subunits of the *Saccharomyces cerevisiae* signal recognition particle required for its functional expression. *EMBO J.*, **13**, 4390–4400.

Chu,J.-L. and Elkon,K.B. (1991) The small nuclear ribonucleoproteins, SmB and B′, are products of a single gene. *Gene*, **97**, 311–312.

Cooper,M., Johnston,L.H. and Beggs,J.D. (1995) Identification and characterisation of Uss1p (Sdb23p): a novel U6 snRNA-associated protein with significant similarity to core proteins of small nuclear ribonucleoproteins. *EMBO J.*, **14**, 2066–2075.

Dahlberg,J.E. and Lund,E. (1988) The genes and transcription of the major small nuclear RNAs. In Birnstiel,M.L. (ed.), *Structure and Function of the Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer Verlag, Berlin, pp. 38–70.

Dijk,J. and Littlechild,J. (1979) Purification of ribosomal proteins from *Escherichia coli* under nondenaturating conditions. *Methods Enzymol.*, **59**, 481–502.

Eason,R. (1984) Analytical ultracentrifugation. In Rickwood,D. (ed.), *Centrifugation—A Practical Approach*. IRL Press Ltd, Oxford, UK, pp. 251–287.

Feeney,R.J., Sauterer,R.A., Feeney,J.L. and Zieve,G.W. (1989) Cytoplasmic assembly and nuclear accumulation of mature small nuclear ribonucleoprotein particles. *J. Biol. Chem.*, **264**, 5776–5783.

Fischer,U., Sumpter,V., Sekine,M., Satoh,T. and Lührmann,R. (1993) Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m_3G cap. *EMBO J.*, **12**, 573–583.

Fischer,U. and Lührmann,R. (1990) An essential signaling role for the m_3G cap in the transport of U1 snRNP to the nucleus. *Science*, **249**, 786–790.

Fisher,D.E., Conner,G.E., Reeves,W.H., Wisniewolski,R. and Blobel,G. (1985) Small nuclear ribonucleoprotein particle assembly *in vivo*: demonstration of a 6S RNA-free core precursor and posttranslational modification. *Cell*, **42**, 751–758.

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.

Hamm,J., Datzynkiewicz,E., Tahara,S.M. and Mattaj,I.W. (1990) The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell*, **62**, 569–577.

Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Heinrichs,V., Hackl,W. and Lührmann,R. (1992) Direct binding of small nuclear ribonucleoprotein G to the Sm site of small nuclear RNA. Ultraviolet light cross-linking of protein G to the AAU stretch within the Sm site (AAUUUGUGG) of U1 small nuclear ribonucleoprotein reconstituted *in vitro*. *J. Mol. Biol.*, **227**, 15–28.

Hermann,H., Fabrizio,P., Raker,V.A., Foulaki,K., Hornig,H., Brahms,H. and Lührmann,R. (1995) SnRNP Sm proteins share two evolutionary conserved sequence motifs which are involved in Sm protein–protein interactions. *EMBO J.*, **14**, 2076–2088.

Kastner,B., Bach,M. and Lührmann,R. (1990) Electron microscopy of small nuclear ribonucleoprotein (snRNP) particles U2 and U5: evidence for a common structure-determining principle in the major U snRNP family. *Proc. Natl Acad. Sci. USA*, **87**, 1710–1714.

Lehmeier,T., Foulaki,K. and Lührmann,R. (1990) Evidence for three distinct D proteins, which react differentially with anti-Sm autoantibodies, in the cores of the major snRNPs U1, U2, U4/U6 and U5. *Nucleic Acids Res.*, **18**, 6475–6484.

Lehmeier,T., Raker,V.A., Hermann,H. and Lührmann,R. (1994) cDNA cloning of the Sm proteins D2 and D3 from human small nuclear ribonucleoproteins: evidence for a direct D1–D2 interaction. *Proc. Natl Acad. Sci. USA*, **91**, 12317–12321.

Lerner,M.R. and Steitz,J.A. (1979) Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc. Natl Acad. Sci. USA*, **76**, 5495–5499.

Lerner,E.A., Lerner,M.R., Hardin,J.A., Janeway,C.A. and Steitz,J.A. (1981) Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. *Proc. Natl Acad. Sci. USA*, **78**, 2737–2741.

Liautard,J.P., Sri-Widada,J., Brunel,C. and Jeanteur,P. (1982) Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. *J. Mol. Biol.*, **162**, 623–643.

Lührmann,R., Appel,B., Rinke,J., Reuter,R., Rothe,S. and Bald,R. (1982) Isolation and characterisation of rabbit anti-$m^{2,2,7}G$ antibodies. *Nucleic Acids Res.*, **10**, 7103–7113.

Lührmann,R., Kastner,B. and Bach,M. (1990) Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochim. Biophys. Acta*, **1087**, 265–292.

Lütcke,H., Prehn,S., Ashford,A.J., Remus,M., Frank,R. and Dobberstein,B. (1993) Assembly of the 68- and 72-kD proteins of signal recognition particle with 7S RNA. *J. Cell Biol.*, **121**, 977–985.

Mattaj,I. (1986) Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell*, **46**, 905–911.

Mattaj,I.W. (1988) UsnRNP assembly and transport. In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer Verlag, Berlin, pp. 100–114.

Mattaj,I.W. and De Robertis,E.M. (1985) Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. *Cell*, **40**, 111–118.

Mattaj,I.W. and Nagai,K. (1995) Recruiting proteins to the RNA world. *Nature Struct. Biol.*, **2**, 518–522.

Moore,M.J., Query,C.C. and Sharp,P.A. (1993) Splicing of precursors to messenger RNAs by the spliceosome. In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.

Plessel,G., Fischer,U. and Lührmann,R. (1994) m_3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) *in vitro*: evidence that the U1 small nuclear RNA-(guanosine-*N*2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. *Mol. Cell. Biol.*, **14**, 4160–4172.

Reddy,R. and Bush,H. (1988) Small nuclear RNAs: RNA sequences, structure, and modifications. In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer Verlag, Berlin, pp. 1–37.

Römisch,K., Webb,J., Herz,J., Prehn,S., Frank,R., Vingron,M. and Dobberstein,B. (1989) Homology of 54K protein of signal recognition particle, docking protein and two *E.coli* proteins with putative GTP-binding domains. *Nature*, **340**, 478–482.

Sauterer,R.A., Goyal,A. and Zieve,G.W. (1990) Cytoplasmic assembly of small nuclear ribonucleoprotein particles from 6S and 20S RNA-free intermediates in L929 mouse fibroblasts. *J. Biol. Chem.*, **265**, 1048–1058.

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.

Ségault,V., Will,C.L., Sproat,B.S. and Lührmann,R. (1995) *In vitro* reconstitution of mammalian U2 and U5 snRNPs active in splicing: Sm proteins are functionally interchangeable and are essential for the formation of functional U2 and U5 snRNPs. *EMBO J.*, **14**, 4010–4021.

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.

Sillekens,P.T.G., Habets,W.J., Beijer,R.P. and van Venrooij,W.J. (1987) cDNA cloning of the human U1 snRNA-associated A protein: extensive homology between U1 and U2 snRNP-specific proteins. *EMBO J.*, **6**, 3841–3848.

Singh,R. and Reddy,R. (1989) Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl Acad. Sci. USA*, **86**, 8280–8283.

Steitz,J.A., Black,D.L., Gerke,V., Parker,K.A., Krämer,A., Frendewey,D. and Keller,W. (1988) Functions of the abundant U-snRNPs. In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer Verlag, Berlin, pp. 115–154.

Strub,K. and Walter,P. (1990) Assembly of the Alu domain of the signal recognition particle (SRP): dimerization of the two protein components is required for efficient binding to SRP RNA. *Mol. Cell. Biol.*, **10**, 777–784.

Sumpter,V., Kahrs,A., Fischer,U., Kornstädt,U. and Lührmann,R. (1992) *In vitro* reconstitution of U1 and U2 snRNPs from isolated proteins and snRNA. *Mol. Biol. Rep.*, **16**, 229–240.

Van Dam,A., Winkel,I., Zijlstra-Baalbergen,J., Smeenk,R. and Cuypers,H.T. (1989) Cloned human snRNP proteins B and B′ differ only in their carboxy-terminal part. *EMBO J.*, **8**, 3853–3860.

Williams,D.G., Stocks,M.R., Smith,P.R. and Maini,R.N. (1986) Murine lupus monoclonal antibodies define five epitopes on two different Sm polypeptides. *Immunology*, **58**, 495–500.

Zeller,R., Nyffenegger,T. and De Robertis,E.M. (1983) Nucleocytoplasmic distribution of snRNPs and stockpiled snRNA-binding proteins during oogenesis and early development in *Xenopus laevis*. *Cell*, **32**, 425–434.

Zieve,G.W. and Sauterer,R.A. (1990) Cell biology of the snRNP particles. *CRC Crit. Rev. Biochem. Mol. Biol.*, **25**, 1–46.

Received on November 3, 1995; revised on January 16, 1996