4614–4623 *Nucleic Acids Research, 1996, Vol. 24, No. 23*

In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes

Cindy L. Will, Steffen Rümpler, Jacqueline Klein Gunnewiek[1], Walther J. van Venrooij[1] and Reinhard Lührmann*

Institut für Molekularbiologie und Tumorforschung, Philipps Universität Marburg, 35037 Marburg, Germany and [1]Department of Biochemistry, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands

Received September 13, 1996; Revised and Accepted October 16, 1996

ABSTRACT

We have established an *in vitro* reconstitution/splicing complementation system which has allowed the investigation of the role of mammalian U1 snRNP components both in splicing and at the early stages of spliceosome formation. U1 snRNPs reconstituted from purified, native snRNP proteins and either authentic or *in vitro* transcribed U1 snRNA restored both early (E) splicing complex formation and splicing activity to U1-depleted extracts. *In vitro* reconstituted U1 snRNPs possessing an m_3G or ApppG cap were equally active in splicing, demonstrating that a physiological cap structure is not absolutely required for U1 function. However, the presence of an m^7GpppG or GpppG cap was deleterious to splicing, most likely due to competition for the m^7G cap binding proteins. No significant reduction in splicing or E complex formation was detected with U1 snRNPs reconstituted from U1 snRNA lacking the RNA binding sites of the U1-70K or U1-A protein (i.e., stem–loop I and II, respectively). Complementation studies with purified HeLa U1 snRNPs lacking subsets of the U1-specific proteins demonstrated a role for the U1-C, but not U1-A, protein in the formation and/or stabilization of early splicing complexes. Studies with recombinant U1-C protein mutants indicated that the N-terminal domain of U1-C is necessary and sufficient for the stimulation of E complex formation.

INTRODUCTION

Nuclear pre-mRNA splicing requires the formation of a large ribonucleoprotein complex, the spliceosome, wherein the catalysis of the two sequential transesterification reactions responsible for intron removal and exon ligation occurs. Spliceosomes are formed by the ordered interaction of numerous splicing factors and the four small nuclear ribonucleoproteins (snRNPs), U1, U2, U5 and U4/U6, with conserved regions of the pre-mRNA (reviewed in 1). One of the initial contacts with the pre-mRNA in the spliceosome assembly pathway is the binding of U1 snRNP to the 5′ splice site. This interaction involves base pairing between the 5′ end of the U1 snRNA and conserved sequences spanning the 5′ splice site (2,3). A stable U1 snRNP/pre-mRNA complex, referred to as the commitment complex, was first identified in yeast (4,5). A similar complex, designated the early or E complex, was subsequently identified in mammalian splicing extracts by gel filtration (6,7). Subsequent to E complex assembly, the U2 snRNP interacts with the branch site of the intron, thereby forming the so-called pre-spliceosome. Mature spliceosomes are ultimately formed by the interaction of the U4/U6 and U5 snRNPs, in the form of a pre-assembled [U4/U6.U5] tri-snRNP complex (reviewed in 1).

The spliceosomal snRNPs consist of one snRNA molecule (or two in the case of U4/U6) complexed with eight so-called Sm or core proteins (B, B′, D1, D2, D3, E, F, G), which are present in all snRNP species, and a number of particle-specific proteins (reviewed in 8). The snRNP core proteins interact with an evolutionarily conserved region of the U1, U2, U4 and U5 snRNAs, the Sm site, which consists of a single-stranded, uridylic acid-rich region that is flanked by two stem–loop structures (9). The association of the snRNP Sm proteins results in the hypermethylation of the snRNA's monomethylguanosine (m^7G) 5′ cap structure to a 2,2,7-trimethylguanosine (m_3G) form (10). The m_3G cap, together with the snRNP Sm proteins, forms the karyophilic signal required for the nuclear import of the spliceosomal snRNPs (11,12).

The mammalian U1 snRNP contains, in addition to U1 snRNA and the common snRNP proteins, three U1-specific proteins denoted 70K, A and C. The U1-70K and U1-A proteins possess highly conserved RNA binding domains (RBDs) characteristic of a number of RNA binding proteins (13,14). These proteins are thus able to interact directly with the U1 snRNA; specifically, U1-70K binds stem–loop I of the U1 snRNA and U1-A interacts with stem–loop II (15–17). Protein–protein contacts also appear to contribute to the association of U1-70K and U1-A with the U1 particle. For example, a stable interaction between a U1-70K deletion mutant containing the N-terminal 97 amino acids, which do not bind U1 snRNA, and U1 snRNPs containing only the core snRNP proteins has been demonstrated (18). Recent studies with SNP1, the *Saccharomyces cerevisiae* U1-70K homolog, indicate

*To whom correspondence should be addressed. Tel: +49 6421 28 6240; Fax: +49 6421 28 7008; Email: luehrmann@imt.marburg.de

that contacts between 70K and other U1 proteins are sufficient for the formation of functional U1 snRNPs (19). The 70K protein also exhibits structural features, including regions rich in serine and arginine residues, characteristic of the SR family of essential splicing factors (reviewed in 20).

In contrast to U1-70K and U1-A, the U1-specific C protein does not contain an RBD and its association with the U1 snRNP appears to be mediated primarily by protein–protein interactions (21,22). Binding studies carried out with U1-C deletion mutants have shown that the N-terminal 45 amino acids suffice for its association with the U1 snRNP (22). Interestingly, this region of the U1-C protein contains a zinc finger-like motif which appears to be essential for binding, since point mutations in the cysteine and histidine residues of this putative zinc finger abolish U1-C binding (22). The association of U1-C with the U1 snRNP requires the presence of the N-terminal region of the 70K protein, as well as one or more Sm protein (18). More recent *in vitro* studies have demonstrated that the U1-C protein can form homodimers; U1-C dimerization also requires amino acids located in its N-terminal domain (23).

The U1 snRNP appears to function primarily during the early steps of splicing complex formation. Its main role is the recognition of the 5′ splice site which appears to be a nucleation event for spliceosome assembly. While U1 snRNP function relies heavily upon base pairing interactions of the U1 snRNA, protein components have also been shown to contribute to its activity. A general role for U1-specific proteins was initially suggested by splicing complementation studies in *Xenopus* oocytes using mutant U1 snRNAs; several mutants which did not support the stable association of U1-specific proteins were unable to restore splicing activity to oocytes that had been depleted of their endogeneous U1 snRNA (24). The first indication that a stable U1/5′ splice site interaction is mediated by U1 snRNP proteins was provided by studies demonstrating that mild proteolysis of the mammalian U1 particle inhibits its association with the 5′ splice site (25). Consistent with this observation, filter binding and gel mobility shift assays have provided evidence that the U1-C protein can augment the binding of the U1 snRNP to the 5′ splice site (26,27). The splicing factor ASF/SF2 has also been shown to enhance the interaction of U1 with the 5′ splice site; this enhancement appears to involve an interaction between the SR domain of ASF/SF2 and that of the U1-70K protein (28).

The bulk of functional information regarding the mammalian U1-specific proteins has been obtained by studies carried out with highly purified components in the absence of splicing extract. Here we have investigated the function of the mammalian U1-specific proteins in both splicing and early splicing complex formation, using U1-depleted splicing extracts and either *in vitro* reconstituted or biochemically purified U1 snRNPs. Complementation studies performed with purified U1 snRNPs lacking one or more of the U1-specific proteins demonstrated that the U1-C protein, but not U1-A, enhances the formation of early spliceosomal complexes. Mutagenesis experiments indicated that the N-terminal 60 amino acids of the U1-C protein are sufficient for this enhancement.

MATERIALS AND METHODS

Preparation of snRNPs and native snRNP proteins

Nuclear extracts were prepared from HeLa cells (Computer Cell Culture Center, Mons) as described (29). Native U1 snRNPs or U1 snRNPs specifically lacking either the U1-A (ΔA), U1-A and -C (ΔA,C) or U1-A -C, and -70K proteins (core U1 snRNPs) were isolated from HeLa nuclear extract by anti-m_3G immunoaffinity chromatography followed by Mono Q chromatography (30). Core U1, ΔA, and ΔA,C particles contained maximally 5% of each of the depleted U1-specific proteins. Native, RNA-free snRNP proteins (TPs) were isolated from a mixture of immunoaffinity purified U1, U2, U5 and U4/U6 snRNPs by dissociation in the presence of EDTA and the anion exchange resin DE53 (31). HeLa U1-A and U1-C proteins were isolated from native snRNP proteins by Mono S chromatography as previously described (18). Recombinant his-tagged U1-C proteins were constructed and purified as described previously (23). The substitution mutant (s28/29) was constructed in essentially the same manner.

Preparation of snRNAs and pre-mRNAs

HeLa U1 and U2 snRNAs were isolated from purified snRNPs as described previously (31). *In vitro* transcribed human U1 snRNA was prepared from *Pst*I linearized pHU1a (32). Wild-type and mutant *Xenopus* U1 snRNAs were prepared by *in vitro* transcription of *Bam*HI linearized plasmids with T7 polymerase (15). U1 snRNAs were transcribed in the presence of 1 mM chemically synthesized GpppG, ApppG, m^7GpppG or m_3GpppG (as indicated) and 0.1 mM GTP and were purified by denaturing polyacrylamide gel electrophoresis. Radiolabeled MINX (33) and pSP62Δi (34) pre-mRNA, with a specific activity of 2.5×10^6 c.p.m./pmol, were transcribed *in vitro* in the presence of GpppG essentially as described (35). 5′ and 3′ MINX pre-mRNAs, with a specific activity of 6.25×10^6 c.p.m./pmol, were transcribed from *Hin*dIII linearized pMINX and *Bam*HI linearized p3′MINX. p3′MINX was generated by isolating a *Hin*dIII–*Bam*HI fragment from pMINX and subcloning into pGEM4.

U1 snRNP depletion and splicing complementation

U1-depleted nuclear extract was prepared by affinity selection with a biotinylated 2′-*O*-methyl RNA oligonucleotide and streptavidin–agarose beads (36). An oligonucleotide complementary to U1 snRNA nucleotides 1–13 with the following sequence was used: 5′-GCCAGGUAAGUAUdC*dC*dC*dC*dT-3′ (where dC* denotes a biotinylated 2′-deoxycytidine and A, U, G and C represent 2′-*O*-methyl-ribonucleotides). Titration experiments and Western blotting indicated that >97.5% of the U1 snRNA, and 95% of each of the U1-specific proteins A, C and 70K, had been removed from the U1-depleted extract. Splicing reactions (12.5 µl) were incubated for 60 min at 30°C and typically contained 35% extract, 45 mM KCl, 2.5 mM $MgCl_2$, 2.0 mM ATP, 10 mM creatine phosphate and 12 fmol ^{32}P-labeled pSP62Δi pre-mRNA (3×10^4 c.p.m.). For complementation of U1-depleted extract, 0.6 pmol (200 ng) of Mono Q purified HeLa U1 snRNP were added directly to the reaction mixture. Complementation with *in vitro* reconstituted particles was accomplished by combining 1.8 pmol (100 ng) of authentic or *in vitro* transcribed U1 snRNA and 3.8 pmol (750 ng) of purified native snRNP proteins. This mixture was incubated for 30 min at 30°C in the presence of splicing reactions lacking pre-mRNA, and splicing was initiated by the addition of the pre-mRNA. No differences in complementation efficiency were observed when reconstitution was carried out either directly in splicing extract or by additionally pre-incubating in the absence of extract. Splicing intermediates and products were isolated by phenol/chloroform extraction followed by ethanol precipitation and analysed on 14.0% polyacrylamide–7 M urea gels.

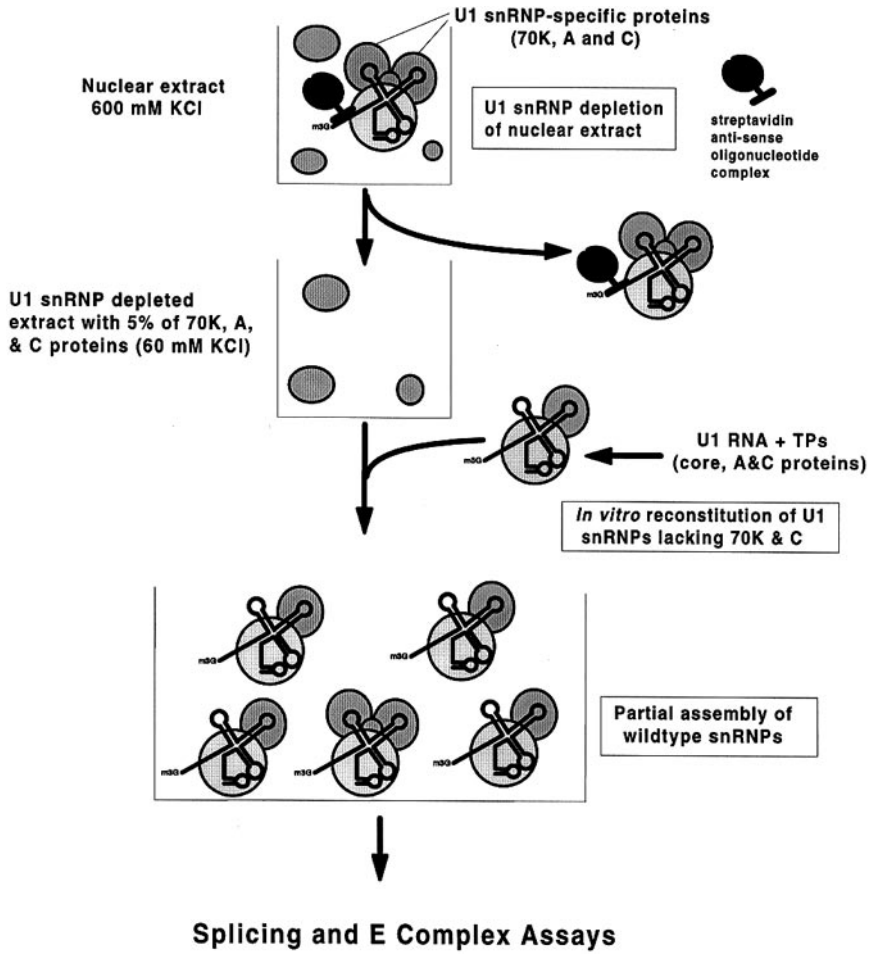

Figure 1. The U1 snRNP depletion/complementation assay. The depletion of U1 snRNPs from nuclear extract through a biotinylated 2′-*O*-methyloligoribonucleotide and streptavidin–agarose, and reconstitution of U1 snRNPs from U1 snRNA and purified snRNP proteins (TPs) are shown schematically. The association of stoichiometric amounts of U1-A, but only small amounts of U1-70K, with the U1 snRNA was confirmed by immunoprecipitation experiments with monoclonal antibodies specific for either the U1-A or U1-70K protein.

E complex assays

E complex formation was assayed by gel filtration on Sephacryl S-500 (6). Splicing extracts were initially depleted of ATP and magnesium by dialysis and subsequent incubation at room temperature for 30 min. Under these conditions no A complex formation was observed by native gel electrophoresis after a 30 min incubation at 30°C. Standard E complex reactions (75 µl) contained 30% splicing extract, 45 mM KCl and 0.18 pmol (4.5×10^5 c.p.m.) of ^{32}P-labeled MINX pre-mRNA and were incubated at 30°C for 25 min. Complexes were fractionated on a 60× 0.9 cm column at a flow rate of 10 ml/h, and the amount of radioactivity present in 350 µl fractions was determined by Cherenkov counting. Assays performed with 5′ or 3′ MINX pre-mRNA contained 72 fmoles (4.5×10^5 c.p.m.) of radiolabeled transcript. For complementation of U1-depleted extracts, 4 pmol of the indicated Mono Q purified U1 snRNPs were added either directly to the reaction mixture or after a 30 min incubation on ice with 20 pmol of Mono S purified or recombinant HeLa U1-A or U1-C protein, as indicated. For *in vitro* reconstitution studies, 600 ng of U1 snRNA and 3.75 µg native snRNP proteins (in a total of 3 µl) were incubated for 30 min at 30°C prior to incubation with extract.

RESULTS

In vitro reconstituted U1 snRNPs are active in splicing

As a potential means to investigate the function of the mammalian U1-specific proteins in pre-mRNA splicing, we established an *in vitro* reconstitution/splicing complementation system for HeLa U1 snRNPs (depicted schematically in Fig. 1). HeLa nuclear extracts specifically depleted of U1 snRNPs were prepared by affinity selection with a biotinylated 2′-*O*-methyl RNA oligonucleotide complementary to nucleotides 1–13 of the U1 snRNA. Mock-depleted extract was handled in an identical manner, except that oligonucleotide was omitted. Reconstitution of U1 snRNPs was carried out by incubating purified U1 snRNA and native snRNP proteins (TPs) in the presence of splicing extract. TPs, which are essentially free of any snRNA, consist of the snRNP Sm proteins, B, B′, D1, D2, D3, E, F and G, and the U1-specific proteins A and C, but only trace amounts of U1-70K (31,37). Since the association of U1-C is strictly dependent on the presence of the U1-70K protein (18), *in vitro* reconstituted U1 snRNPs thus consist predominantly of the snRNP Sm proteins and the U1-A protein. Due to low levels of the U1-70K protein

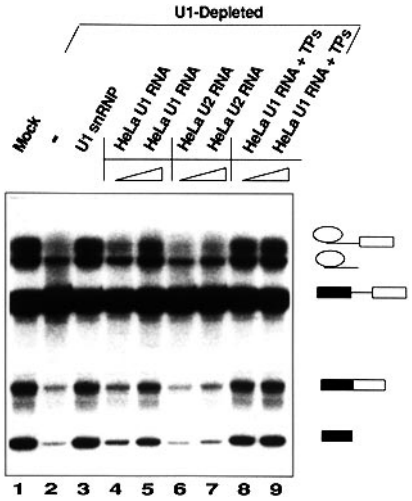

Figure 2. Complementation of U1-depleted extracts with U1 snRNPs reconstituted *in vitro*. Reconstitutions were performed in the presence of extract as described in the Materials and Methods. *In vitro* splicing reactions were performed with mock (lane 1), U1-depleted (lane 2) or U1-depleted extract plus the following: 0.8 pmol (200 ng) Mono Q purified U1 snRNPs (lane 3), 1.0 or 2.0 pmol each of purified HeLa U1 snRNA (lanes 4–5), HeLa U2 snRNA (lanes 6–7), or U1 snRNA plus 750 ng (~4.0 pmol) of purified snRNP proteins (TPs) (lanes 8–9). The addition of TPs alone did not restore splicing to a significant extent. Splicing was performed for 60 min with SP62Δi pre-mRNA as described in the Materials and Methods. Splicing intermediates and products as well as unspliced pre-mRNA (indicated at the right) were fractionated on a 14% polyacrylamide–7 M urea gel and visualized by autoradiography. All U1-depleted extracts typically exhibited a low level of residual splicing activity (~10–20% of mock extract activity).

in TPs and U1-depleted extract, a small amount of wild-type U1 snRNP is also reconstituted (data not shown).

Splicing of an adenovirus major late II pre-mRNA (pSP62Δi) was significantly reduced in U1-depleted extract (Fig. 2, lanes 1–2). However, the addition of a physiological amount (200 ng) of highly purified U1 snRNP restored splicing to mock extract levels (Fig. 2, lane 3), demonstrating that the observed reduction in activity is due specifically to the absence of U1 snRNPs. Splicing was also restored if 2 pmol of purified HeLa U1 snRNA (a 2-fold excess as compared to the mock extract level) was added (Fig. 2, lane 5). U2 snRNA, on the other hand, did not complement splicing (lanes 6–7), indicating that this effect is specific for U1 snRNA. Splicing complementation was, however, significantly enhanced, especially at the lower U1 snRNA concentration, if native snRNP proteins were added to the reconstitution mixture (compare lanes 4–5 with 8–9). TPs alone had little or no effect on the splicing activity of U1-depleted extract (Fig. 3, lane 23). Since the extent of U1 snRNA degradation was significantly reduced when TPs were present during reconstitution and *in vitro* splicing, the TP-induced enhancement of splicing complementation appeared to be primarily the result of increased U1 snRNA stability (data not shown). In sum, splicing could be restored by the addition of purified U1 snRNA alone, but even more efficiently by the combination of U1 snRNA and purified snRNP proteins.

U1 snRNPs reconstituted from *in vitro* transcribed snRNA can complement splicing

We next tested whether the splicing activity of U1-depleted extract could be restored by *in vitro* transcribed U1 snRNA. Concurrently, the role in splicing of the U1 snRNA 5′ cap was investigated by comparing *in vitro* transcribed RNAs possessing various cap structures. Synthetic HeLa U1 snRNA containing an m_3G or non-physiological ApppG cap (either alone or with TPs) restored splicing to U1-depleted extract as efficiently as U1 snRNA isolated from HeLa U1 snRNPs (compare lanes 13–14, 17–18 and 19–20). In contrast, U1 snRNAs containing either an m^7G or GpppG cap not only were unable to complement splicing (lanes 15–16 and 21–22), but also, in the absence of TPs, inhibited the activity of the mock extract as well as the residual splicing activity of the U1-depleted extract (compare lane 1 with lanes 3 and 9, and lane 12 with lanes 15 and 21). Since the stability of the m^7G- and GpppG-capped U1 snRNAs was similar to that of the m_3G- and ApppG-capped ones (data not shown), the decreased complementation efficiency of these RNAs cannot be attributed to an increase in their turnover. Rather, the latter result is consistent with previous data demonstrating that short RNAs possessing an m^7G or GpppG cap can inhibit splicing by competing for proteins (i.e., CBP20 and CBP80) which normally bind the m^7G cap of the pre-mRNA (38). Thus, although there is no absolute requirement for an m_3G cap, not all cap structures are compatible with U1 snRNP splicing activity *in vitro*.

Deletion of stem–loop I or II of the U1 snRNA does not affect the splicing activity of reconstituted U1 snRNPs

The ability to complement splicing with *in vitro* transcribed U1 snRNA allowed us to investigate the effect of U1 snRNA mutation on the activity of *in vitro* reconstituted U1 snRNPs. Since we were particularly interested in investigating the function of the U1-specific proteins, we tested the activity of *Xenopus* U1 snRNA mutants which lacked the RNA binding sites of the U1-70K and U1-A protein (mutants A and B, respectively). As shown in Figure 4, *in vitro* transcribed wild-type *Xenopus* U1 snRNA restored splicing activity to U1-depleted extract (lanes 6–7), albeit slightly less efficiently than U1 snRNA isolated from HeLa U1 snRNPs (lanes 4–5). As compared to wild-type, deletion of stem–loop I (mutant A) or stem–loop II (mutant B) had no significant effect on the complementation efficiency of *in vitro* reconstituted U1 snRNPs (Fig. 4, lanes 8–11). Thus, reconstitution of U1 snRNPs active in splicing does not require the presence of stem–loop I or II. This in turn suggests that either U1-70K and U1-A are dispensible for U1 snRNP function *in vitro* or, alternatively, that these proteins can functionally associate with the U1 snRNP by other means (e.g. by protein–protein interactions).

E complex formation in U1-depleted extracts can be complemented with purified or *in vitro* reconstituted U1 snRNPs

Since substoichiometric amounts of the U1 snRNP may suffice for the complete restoration of splicing activity, we reasoned that alterations in the U1 snRNP that affect its function might be more readily detected by an assay which directly measures the amount of functional snRNPs present. We thus analyzed the formation of the earliest detectable functional splicing complex, the E complex, by gel filtration (6). Due to inefficient splicing complex formation with SP62Δi transcripts (data not shown), E complex assays were performed with MINX pre-mRNA, which is also a derivative of the adenovirus major late II transcript. As compared to the mock-depleted extract, E complex formation was severely reduced in U1-depleted extract; predominantly pre-mRNA/

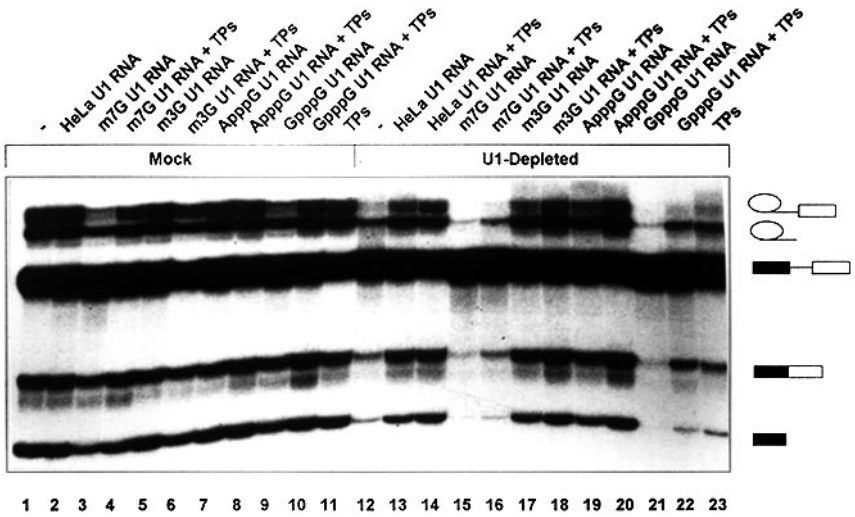

Figure 3. Complementation of U1-depleted extracts with *in vitro* transcribed U1 snRNA. U1 snRNA was transcribed *in vitro* with either an $m^{2,2,7}$GpppG (m_3G), m^7GpppG (m^7G), GpppG or ApppG cap as described in the Materials and Methods. *In vitro* splicing reactions were performed with mock (lanes 1–11) or U1-depleted extract (lanes 12–23) either alone (lanes 1 and 12) or after the addition of various U1 snRNAs/snRNPs. 100 ng of the following RNAs were added either alone or with 750 ng of purified snRNP proteins (TPs; as indicated above each lane): U1 snRNA purified from U1 snRNPs (HeLa U1 RNA; lanes 2, 13 and 14) or *in vitro* transcribed U1 containing an m^7G (lanes 3–4 and 15–16), m_3G (lanes 5–6 and 17–18), ApppG (lanes 7–8 and 19–20) or GpppG cap (lanes 9–10 and 21–22). In lanes 11 and 23, 750 ng of TPs alone were added to the reaction mixture. *In vitro* reconstitution and *in vitro* splicing assays were performed as described in the legend of Figure 2.

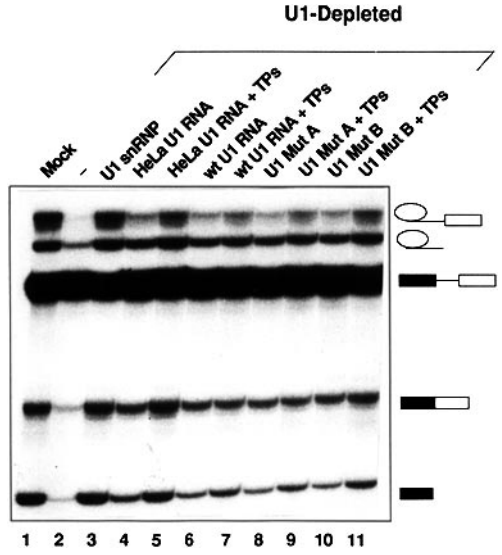

Figure 4. Complementation of U1-depleted extracts with U1 snRNPs reconstituted from U1 snRNAs lacking stem–loop I or II. *In vitro* splicing assays were performed with mock (lane 1), U1-depleted (lane 2) or U1-depleted extract plus 200 ng of HeLa U1 sRNP (lane 3) or 2.0 pmol of the following U1 snRNAs, either alone or with 750 ng TPs (as indicated above each lane): U1 snRNA isolated from HeLa U1 snRNPs (lanes 4 and 5), or *in vitro* transcribed, wild-type (lanes 6 and 7), mutant A (lanes 8 and 9) or mutant B (lanes 10 and 11) *Xenopus* U1 snRNA. Mutant A lacks nucleotides 17–47 which comprise stem–loop I of the U1 snRNA, whereas mutant B lacks nucleotides 50–91 which form stem–loop II (24). *In vitro* reconstitution and *in vitro* splicing assays were carried out as described in the legend of Figure 2.

hnRNP protein complexes (i.e., H complexes) were formed (Fig. 5A). The addition of physiological amounts of highly purified U1 snRNPs shifted the gel filtration profile of the U1-depleted extract to that of the mock extract (Fig. 5A), indicating that the reduction in early splicing complex formation is specifically due to the absence of U1 snRNPs. E complex formation could also be partially restored by the addition of U1 snRNPs reconstituted from TPs and either authentic HeLa or *in vitro* transcribed *Xenopus* U1 snRNA (Fig. 5B). The addition of *in vitro* transcribed *Xenopus* U1 snRNA or TPs alone had little or no effect on complex formation (data not shown). The ratio of E to H complex did not change significantly, as compared to wild-type *Xenopus* U1 snRNA, when complementation was performed with reconstituted U1 snRNPs lacking stem–loop I or II (mutants A and B; Fig. 5C). These results thus provide additional evidence that functional U1 snRNPs can be reconstituted even if the U1 snRNA binding sites for the U1-70K and U1-A protein are deleted. Deletion of the 5′ end of the U1 snRNA, on the other hand, led to a marked decrease in E complex formation (data not shown). The latter is consistent with the previous observation that deletion of the 5′ splice site leads to a dramatic reduction in E complex formation (7) and supports the idea that the early splicing complexes which we detect are indeed functional ones.

The U1-C, but not the U1-A, protein stimulates E complex formation in U1-depleted extracts

One drawback of studies performed with reconstituted particles is the presence of very large amounts of U1-A and U1-C, and smaller amounts of U1-70K, in the reconstitution mixture. Since conclusive results regarding the function of the U1-specific proteins could not be obtained with *in vitro* reconstituted particles, we tested the activity of equimolar amounts of biochemically purified HeLa U1 snRNPs lacking either the U1-A (ΔA U1 snRNPs), U1-A and U1-C (ΔA,C U1 snRNPs) or U1-A, U1-C and U1-70K proteins (core U1 snRNPs) (30). As shown in Figure 6A, wild-type U1 snRNPs and those lacking the U1-A protein complemented E complex formation to a similar extent. Particles lacking both the U1-A and U1-C protein only partially complemented E complex formation, whereas those lacking all three

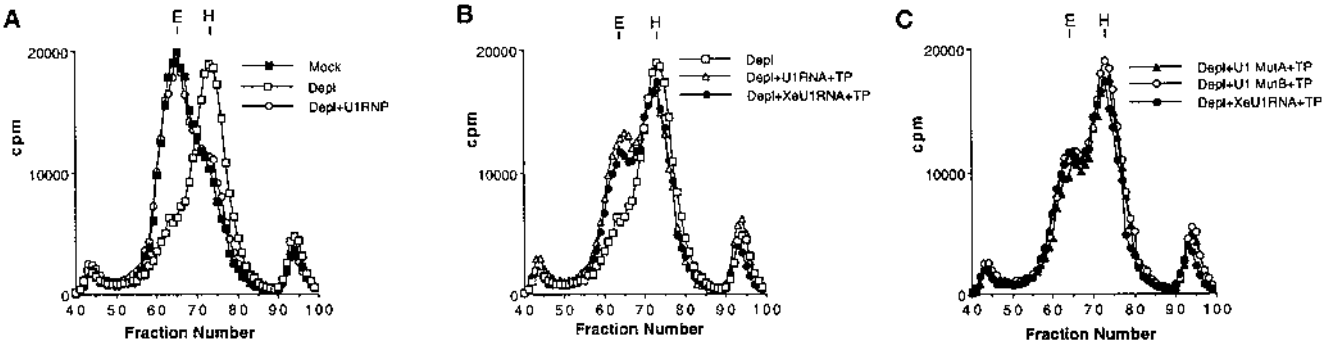

Figure 5. Complementation of E complex formation in U1-depleted extracts with *in vitro* reconstituted U1 snRNPs. Mock or U1-depleted extract (Depl), lacking ATP and magnesium, was incubated at 30°C for 25 min in the presence of 450 000 c.p.m. MINX pre-mRNA and then fractionated by gel filtration on Sephacryl S-500 as described in the Materials and Methods. The peaks containing E and H complexes are indicated above each column profile. The peak in fractions 40–50 is the void volume, whereas that in fractions 90–100 is degraded RNA. For complementation of U1-depleted extracts, 4 pmol of Mono Q purified U1 snRNP (**A**) or U1 snRNPs reconstituted from 600 ng of the indicated U1 snRNA and 4 µg of purified snRNP proteins (TPs) (**B** and **C**) were added. RNA designated U1 was isolated from HeLa U1 snRNPs whereas Xe or Mut A and B U1 snRNAs were transcribed *in vitro* and are from *Xenopus*.

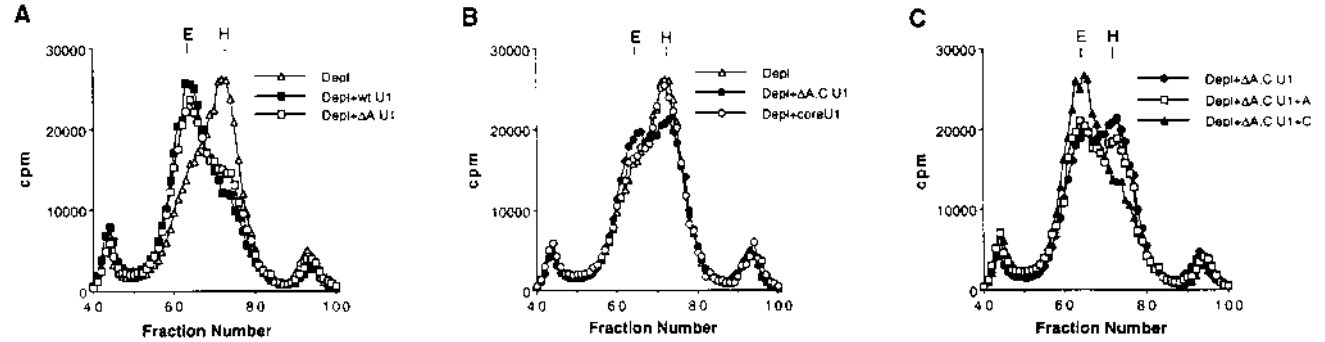

Figure 6. The U1-C protein stimulates E complex formation in U1-depleted extracts. E complex assays were performed as described in the legend to Figure 5. Complementation assays were performed in U1-depleted extract (Depl) with 4 pmol of Mono Q purified wild-type (**A**), ΔA (A), ΔA,C (**B**) or core (B) U1 snRNPs, or with ΔA,C U1 snRNPs that had been pre-incubated for 30 min on ice with 20 pmol of Mono S purified HeLa U1-A or U1-C protein (**C**). The peaks containing E and H complexes are indicated above each column profile.

U1-specific proteins were inactive (Fig. 6B). These results thus suggest that both U1-70K and U1-C, but not U1-A, contribute to the formation of early spliceosomal complexes. To directly determine whether E complex formation is enhanced by the U1-C protein, complementation was performed with ΔA,C U1 snRNPs that had been preincubated with biochemically purified U1-C or U1-A protein. While the complementation efficiency of ΔA,C particles was only slightly enhanced by the addition of purified U1-A protein, the addition of U1-C protein shifted the ratio of E to H complex to that of wild-type U1 snRNPs (Fig. 6C). The addition of the U1-C or U1-A protein alone had no effect on E complex formation (data not shown). These results thus clearly demonstrate that the U1-C protein, in the presence of the U1 snRNP, enhances the assembly of early splicing complexes.

The U1-C protein stimulates splicing complex assembly on the 5′ portion of the pre-mRNA

Previous studies have demonstrated that specific ATP-independent complexes, denoted E5′ and E3′, assemble on RNAs containing only the 5′ or 3′ portion of a pre-mRNA (7). E5′ complexes contain predominantly U1 snRNPs, whereas E3′ complexes are enriched in the splicing factor U2AF (7). To test whether the U1-C protein acts at the 5′ or 3′ splice site, RNAs consisting of the 5′ or 3′ half of the MINX pre-mRNA were prepared and their ability to form E5′ and E3′ complexes, respectively, was tested in mock and U1-depleted extract. Consistent with the known composition of E3′ complexes (i.e., they are for the most part devoid of U1 snRNPs), no significant difference in E3′ assembly was detected in mock versus U1-depleted extract (data not shown). In contrast, with the 5′ MINX transcript, an equal amount of E5′ and H complex was observed with mock-depleted extract, but essentially only H complex with U1-depleted extract (Fig. 7A). As observed with full-length MINX, the addition of physiological amounts of either wild-type or ΔA U1 snRNPs shifted the gel filtration profile of U1-depleted extract to that of the mock-depleted extract (Fig. 7A and B). Addition of ΔA,C U1 snRNPs, on the other hand, resulted in only a partial shift (Fig. 7B). However, pre-incubation of ΔA,C U1 snRNPs with the U1-C protein enhanced E5′ assembly such that the levels of E5′ and H complex were indistinguishable from those of the mock-depleted extract (Fig. 7C). In contrast, addition of the U1-A protein had no effect on the complementation activity of ΔA,C U1 snRNPs (Fig. 7C). Moreover, addition of the U1-C protein alone also had no significant effect on E5′ complex formation (data not shown). Thus, the U1-C protein stimulates

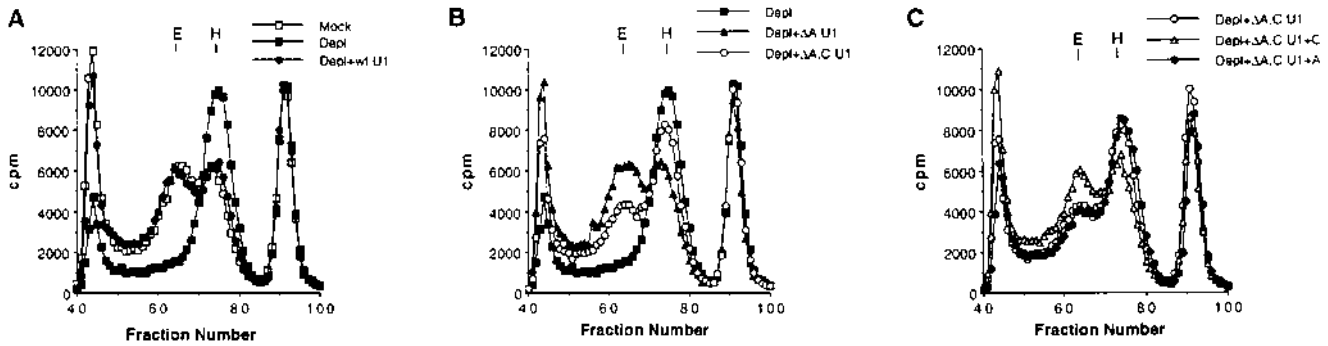

Figure 7. The U1-C protein stimulates splicing complex formation on an RNA containing the 5′ splice site. E complex assays were performed as described in the legend to Figure 5, except mock or U1-depleted extract (Depl) was incubated with the 5′ half of the MINX pre-mRNA. Complementations were performed with 4 pmol of Mono Q purified wild-type (**A**), ΔA (**B**), ΔA,C (B) U1 snRNPs or with ΔA,C U1 snRNPs that had been pre-incubated for 30 min on ice with 20 pmoles of Mono S purified HeLa U1-C or U1-A protein (**C**). The peaks containing E5′ or H complexes are indicated above each column profile.

interactions occurring on the 5′ half of the MINX pre-mRNA, namely those between the U1 snRNP and the 5′ splice site, even in the absence of a 3′ splice site.

The N-terminal domain of the U1-C protein is necessary and sufficient for its activity in splicing complex formation

To determine whether distinct regions of the U1-C protein are necessary and/or sufficient for the stimulation of E complex formation, we tested the activity of several HeLa U1-C mutants. To this end, histidine-tagged U1-C deletion and substitution mutants were constructed and over-expressed in *E.coli* (Fig. 8A). While the addition of ΔA,C U1 snRNPs alone only partially enhanced E complex formation in U1-depleted extract (Fig. 8B), preincubation with wild-type recombinant U1-C protein led to a significant increase in the ratio of E to H complex formed (Fig. 8B). Surprisingly, deletion of the C-terminal 99 amino acids of the U1-C protein did not significantly reduce complementation efficiency, indicating that the N-terminal 60 amino acids are sufficient for activity. In contrast, deletion of amino acids 1–29 abolished the ability of the U1-C protein to enhance E complex formation (Fig. 8B). These results thus demonstrate that residues within the first 29 amino acids of the U1-C protein are essential for its activity. We next tested the activity of two point mutants, (s25) and (s28/29), which exhibit both reduced binding and dimerization activity (23, data not shown). As shown in Figure 8C, substitution of the cysteine at position 25 with a serine led to only a slight decrease in the ratio of E to H complex, while substitution of arginine and lysine at positions 28 and 29 with glycine and serine, respectively, abolished the ability of U1-C to enhance E complex formation. These results thus indicate that the latter amino acids play an important role in the U1-C protein-mediated augmentation of E complex assembly.

DISCUSSION

We have established an *in vitro* reconstitution/splicing complementation system for HeLa U1 snRNPs which should facilitate future investigation of both structural and functional aspects of the U1 snRNP. Reconstitution/complementation systems have thus now been described for each of the mammalian spliceosomal snRNPs and a number of comparisons can be made among them (37,39,40). Restoration of the splicing activity of U1-depleted extracts could be achieved by the addition of an excess of purified U1 snRNA alone (Fig. 2), suggesting that sufficient amounts of those proteins required for the assembly of functional U1 snRNPs are present in U1-depleted extracts. While similar results were previously reported for U4-depleted extracts (39), U2- or U5-depleted extracts could not be complemented by the addition of U2 or U5 snRNA alone. Rather, reconstitution of functional U2 or U5 snRNPs required the addition of purified snRNP Sm proteins (37). In this respect, it is interesting to note that U1 and U4/U6 snRNPs, which contain only three and two particle-specific proteins, respectively, are relatively simple RNP complexes as compared to U2 and U5 snRNPs which contain eleven and nine particle-specific proteins, respectively. Functional reconstitution of the former may thus be more readily achieved in the presence of relatively low levels of Sm proteins or, alternatively, may not be strictly dependent on their presence (39).

U1 snRNPs reconstituted from *in vitro* transcribed U1 snRNA were as active in E complex formation and splicing as those reconstituted from U1 snRNA that had been isolated from U1 snRNPs (Figs 3 and 5). Since *in vitro* transcribed U1 is devoid of modified internal nucleotides, and it is unlikely that modification occurs during reconstitution and *in vitro* splicing, the three pseudouridines and two 2′-*O*-methylated nucleotides normally present at the 5′ end of U1 snRNA appear to be dispensible for U1 snRNP function in splicing. The ability to restore splicing to U1-depleted extract with synthetic U1 snRNA indicates that it may be possible to reconstitute functional U1 snRNPs containing photoactivateable nucleosides (e.g. 4-thiouridine). The *in vitro* reconstitution/splicing complementation system described here could thus potentially be used to study the interactions of the U1 snRNP with other components of the splicing reaction and thereby further our understanding of the three-dimensional architecture of the spliceosome.

As previously reported for the U5 snRNP (37), a modified 5′ cap structure (i.e., m_3G) was also dispensible for the activity of U1 snRNPs in *in vitro* splicing (Fig. 3). However, apparently due to competition for the m^7G cap-binding proteins CBP20 and CBP80, the presence of an m^7G or GpppG cap led to inhibition of the splicing reaction. Since we had previously reported that m^7G-capped U5 snRNPs could restore splicing activity to a U5-depleted extract (37), and the amount of m^7G-capped U5 snRNA used was only 2-fold less than the amount of m^7G-capped

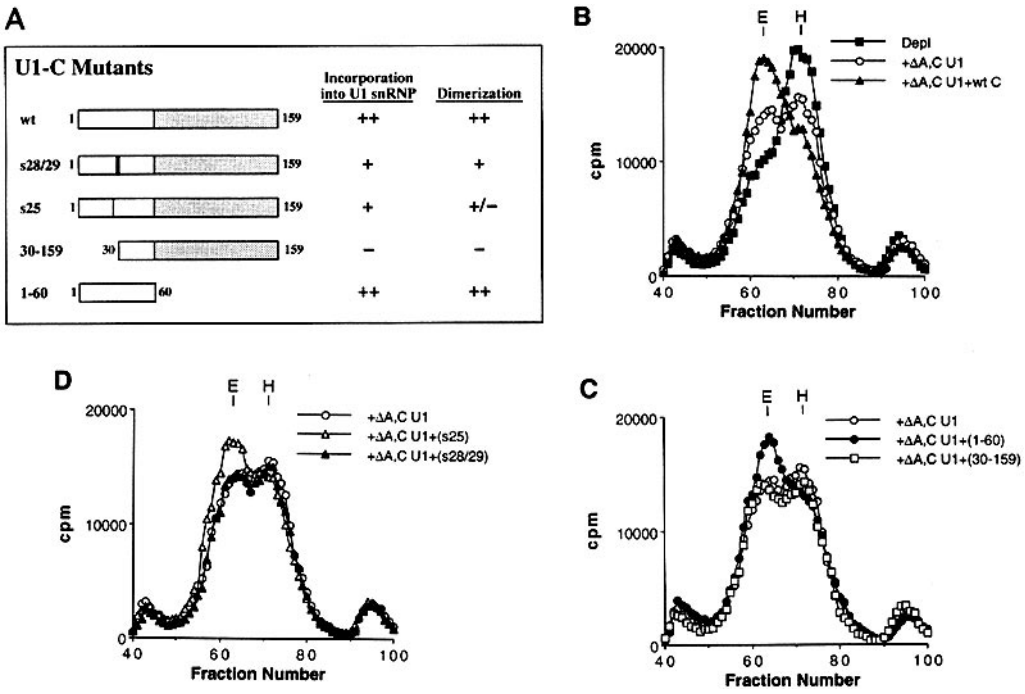

Figure 8. The N-terminal domain is necessary and sufficient for U1-C activity. The recombinant U1-C proteins used in E complex assays are shown schematically in (**A**). The shaded box represents the methionine/proline-rich C terminus which begins at position 61. In the substitution mutant (s28/29), the arginine and lysine at positions 28 and 29 have been substituted by glycine and serine, respectively, and in (s25), the cysteine at position 25 has been replaced by a serine. The ability of each of these proteins to associate with U1 snRNPs and to form dimers is summarized at the right (22,23). E complex assays (B–D) were performed as described in the legend to Figure 5. Complementation of U1-depleted extract (Depl) was performed with 4 pmol of Mono Q purified ΔA,C snRNPs (B) or with ΔA,C U1 snRNPs that had been pre-incubated for 30 min on ice with 20 pmol of the following recombinant U1-C proteins: wild-type (**B**), the deletion mutant (1–60) (**C**), the deletion mutant (30–159) (C), the substitution mutant (s28/29) (**D**), or the substitution mutant (s25) (D). The position of E and H complexes is indicated above each column profile.

U1 snRNA used here, we initially considered whether this inhibitory effect might be specific for the U1 snRNP. A direct comparison of the inhibitory effects of m^7G-capped U1 and U5 snRNPs on *in vitro* splicing indicated that both compete at a comparable level for the binding of the cap-binding proteins (data not shown). The detrimental effect of m^7G-capped snRNPs on splicing thus appears to be more pronounced in U1-depleted extracts. Since the cap-binding proteins have also been shown to stimulate E complex formation (41,42), it is perhaps not surprising that the effects of competition for these proteins are more readily detected in a system where E complex assembly is severely compromised due to the decreased level of U1 snRNPs. Alternatively, since CBP20 and CBP80 have been proposed to interact either directly or indirectly with components of the U1 snRNP (41), it is conceivable that they have been partially co-depleted with the U1 snRNP. The deleterious effect of m^7G-capped U1 or U5 snRNPs on splicing suggests that the requirement of an m_3G cap for the import of the spliceosomal snRNPs into the nucleus may have evolved, at least in part, to prevent the accumulation of m^7G-capped snRNPs in the nucleus.

Whereas *in vitro* splicing activity could be restored by U1 snRNA alone, complementation of E complex assembly required both U1 snRNA and TPs. This result could be explained by differences in the nature of these two assays. In contrast to *in vitro* splicing, the E complex assay is a binding assay and, therefore, should directly reflect the amount of functional U1 snRNP that is reconstituted. Since E complex formation may not be rate limiting for *in vitro* splicing, the reconstitution of small amounts of functional U1 snRNP, while having little impact on the overall level of E complex formation, may suffice for the complete restoration of splicing. In addition, the amount of E complex detected by gel filtration appears to be very dependent upon the stability of the complexes which are formed (7). Structural alterations in the U1 snRNP which reduce the amount of E complex detected by gel filtration (e.g., the removal of the U1-C protein), could thus primarily influence the stability, rather than the assembly, of these complexes. Thus, not only quantitative, but also qualitative differences in the U1 snRNP should be more readily apparent in the E complex assay. This could also explain why, in contrast to *in vitro* splicing, E complex formation was only partially restored by *in vitro* reconstituted U1 snRNPs (Fig. 5B). It is at present not clear, however, whether this partial complementation is due to the reconstitution of insufficient amounts of U1 snRNPs or, alternatively, to the formation of particles that are unable to support stable E complex formation, for example due to limiting amounts of the U1-70K protein which would limit binding of the U1-C protein.

Complementation studies with U1 snRNPs reconstituted from U1 snRNA deletion mutants indicated that functional U1 snRNPs can be formed even after extensive mutagenesis of the U1 snRNA, including deletion of stem–loop I or II (Figs 4 and 5). These results are somewhat surprising given the fact that these mutations significantly inhibited the ability of U1 snRNPs to complement splicing in *Xenopus* oocytes whose endogeneous U1 snRNA had been inactivated by oligonucleotide-directed RNase H cleavage (24). The basis for this difference is not clear, but may simply reflect differences in the assay systems employed. The ability of U1 snRNPs lacking stem–loop I or II to complement

splicing suggests, at first glance, that U1-A and U1-70K are dispensible for U1 snRNP function. However, there is a significant amount of evidence suggesting that these proteins can stably associate with the U1 snRNP solely via protein-protein interactions (16,18,24). In addition, recent studies carried out in *S.cerevisiae* indicated that contacts between the yeast 70K homologue and other U1 snRNP proteins are sufficient for the assembly of a functional U1 snRNP particle (19). Our results are thus consistent with the idea that, despite the removal of their primary RNA binding site, the U1-70K and U1-A proteins can stably and functionally associate with the U1 snRNP.

Evidence that the U1-C, but not U1-A, protein can enhance E complex formation and/or stability was provided by complementation studies with biochemically purified U1 snRNPs lacking one or more of the U1-specific proteins (Figs 6 and 7). These results are consistent with previous studies, performed in the absence of nuclear extract,which indicated that the U1-C protein enhances the interaction of the U1 snRNP with the 5′ splice site (26,27). The U1-C protein could enhance E complex formation by interacting directly with the 5′ splice site. Indeed, it has recently been shown that U1-C can be crosslinked to a short oligonucleotide containing a 5′ splice site (43). Alternatively, the binding and dimerization of the U1-C protein could indirectly enhance interactions with the pre-mRNA by affecting the overall structure of the U1 snRNP. The reduced activity of ΔA,C,70K (core) U1 snRNPs as compared to ΔA,C U1 snRNPs further suggested that the U1-70K protein also contributes to the formation and/or stability of E complexes. Consistent with this hypothesis, a stable, ASF/SF2-mediated interaction between U1 and the 5′ splice site appeared to involve an interaction between the U1-70K protein and ASF/SF2 (27,28). However, since U1-C protein binding is dependent on the U1-70K protein, the reduced activity of core U1 snRNPs as compared to ΔA,C U1 snRNPs could also potentially result from a decrease in the amount of U1-C association. Since neither U1 snRNPs specifically depleted of the U1-70K protein nor sufficient amounts of purified U1-70K protein were available, it was not possible to directly test the effect of this U1-specific protein on E complex formation.

Mutagenesis of the U1-C protein provided strong evidence that its N-terminal domain is necessary and sufficient for its function during the early stages of spliceosome assembly (Fig. 8). A functional role for the N-terminal but not C-terminal domain in E complex formation is consistent with recent studies demonstrating that solely amino acids in the N-terminal domain of the protein have been evolutionarily conserved from yeast to man (44). However, since we have only measured effects on E complex formation, we cannot rule out that the methionine/proline-rich C-terminal domain is involved in subsequent steps of the splicing process. The U1-C N-terminus contains residues essential for the interaction of U1-C with the U1 particle, as well as those required for dimerization (22,23). Due to the inability to overexpress sufficient amounts of point mutants which are solely deficient in U1 snRNP association, we cannot conclude at present whether a stable U1-C/U1 snRNP interaction is essential for the promotion of E complex assembly. However, given that the U1-C protein enhances the interaction of U1 with the 5′ splice site, it is reasonable to assume that the stabilization of this interaction requires the stable association of U1-C with the U1 snRNP. Complementation studies performed with point mutants demonstrated that the arginine and lysine present at positions 28 and 29, respectively, are essential for U1-C activity; substitution of the cysteine at position 25 with serine, on the other hand, only slightly reduced its ability to stimulate E complex formation. Since both mutations have been reported to diminish, but not abolish U1-C dimerization (23, data not shown), these results suggest that dimerization, while not absolutely essential, may still contribute to the activity of the U1-C protein during E complex formation. However, since these mutants also exhibit reduced binding (see Fig. 8), we cannot exclude that their phenotypes reflect, at least in part, their decreased ability to interact with the U1 particle. A more detailed mutational analysis of the U1-C protein should, in the future, allow us to more precisely define those amino acids essential to U1-C function. This system should thus allow a finer examination of those factors influencing the formation of the first functionally important complex in the spliceosome assembly pathway.

ACKNOWLEDGEMENTS

We thank Iain Mattaj for kindly providing the wild-type and mutant *Xenopus* U1 snRNA plasmids, as well as Edward Darzynkiewicz for providing m_3G cap. We also thank Silke Börner for excellent technical assistance, Christopher Marshallsay and Veronica Raker for constructive comments on the manuscript, and Véronique Ségault for helpful advice and discussions. We are grateful to Michael Krause for preparing biotinylated 2′-*O*-methyl oligonucleotides and Gabi Plessel for help in constructing p3′MINX. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 272/A3), the Netherlands Foundation for Chemical Reasearch (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO) and European Community grants ERBCHRXCT 930191 and 930176.

REFERENCES

1 Moore,M.J., Query,C.C. and Sharp,P.A. (1993) In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.
2 Zhuang,Y. and Weiner,A. (1986) *Cell* **46**, 827–835.
3 Siliciano,P.G. and Guthrie,C. (1988) *Genes Dev.* **2**, 1258–1267.
4 Ruby,S.W. and Abelson,J. (1988) *Science* **242**, 1028–1035.
5 Sraphin,B. and Rosbash,M. (1989) *Cell* **59**, 349–358.
6 Michaud,S. and Reed,R. (1991) *Genes Dev.* **5**, 2534–2546.
7 Michaud,S. and Reed,R. (1993) *Genes Dev.* **7**, 1008–1020.
8 Will,C.L., Behrens,S.-E. and Lührmann,R. (1993) *Mol. Biol. Rep.* **18**, 121–126.
9 Branlant,C., Krol,A., Ebel,J.P., Lazar,E., Haendler,B. and Jacob,M. (1982) *EMBO J.* **1**, 1259–1265.
10 Mattaj,I. (1986) *Cell* **46**, 905–911.
11 Fischer,U. and Lührmann,R. (1990) *Science* **249,** 786–790.
12 Hamm,J. and Mattaj,I.W. (1990) *Cell* **63**, 109–118.
13 Query,C.C., Bentley,R.C. and Keene,J.D. (1989) *Cell* **57**, 89–101.
14 Sillekens,P.T.G., Habets,W.J., Beijer,R.P. and van Venrooij,W.J. (1987) *EMBO J.* **6**, 3841–3848.
15 Hamm,J., Kazmaier,M. and Mattaj,I.W. (1987) *EMBO J.* **6**, 3479–3485.
16 Patton,J.R. and Pederson,T. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 747–751.
17 Scherly,D., Boelens,W., van Venrooij,W.J., Dathan,N.A., Hamm,J. and Mattaj,I.W. (1989) *EMBO J.* **8,** 4163–4170.
18 Nelissen,R.L.H., Will,C.L., van Venrooij,W.J. and Lührmann,R. (1994) *EMBO J.* **13,** 4113–4125.
19 Hilleren,P.J., Kao,H.Y. and Siliciano,P.G. (1995) *Mol. Cell. Biol.* **15**, 6341–6350
20 Fu,X.-D. (1995) *RNA* **1**, 663–680
21 Sillekens,P.T.G., Beijer,R.P., Habets,W.J. and van Venrooij,W.J. (1988) *Nucleic Acids Res.* **16**, 8307–8321.
22 Nelissen,R.L.H., Heinrichs,V., Habets,W.J., Simons,F., Lührmann,R. and van Venrooij,W.J. (1991) *Nucleic Acids Res.* **19**, 449–454.

23 Klein Gunnewiek,J.M.T., van Aarssen,Y., Wassenaar,R., Legrain,P., van Venrooij,W. and Nelissen,R.L.H. (1995) *Nucleic Acids Res.* **23**, 4864–4871.
24 Hamm,J., Dathan,N.A., Scherly,D. and Mattaj,I.W. (1990) *EMBO J.* **9**, 1237–1244.
25 Mount,S.M., Pettersson,I., Hinterberger,M., Karmas,A. and Steitz,J.A. (1983) *Cell* **33**, 509–518.
26 Heinrichs,V., Bach,M., Winkelmann,G. and Lührmann,R. (1990) *Science* **247**, 69–72.
27 Jamison,S.F., Pasman,Z., Wang,J., Will,C.L., Lührmann,R., Manley,J.L. and Garcia-Blanco,M.A. (1995) *Nucleic Acids Res.* **23**, 3260–3267.
28 Kohtz,J.D., Jamison,S.F., Will,C.L., Zuo,P., Lührmann,R., Garcia-Blanco,M.A. and Manley,J.L. (1994) *Nature* **368,** 119–124.
29 Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.* **11,** 1475–1489.
30 Bach,M., Bringmann,P. and Lührmann,R. (1990) *Methods Enzymol.* **181**, 232–257.
31 Sumpter,V., Kahrs,A., Fischer,U., Kornstädt,U. and Lührmann,R. (1992) *Mol. Biol. Rep.* **16**, 229–240.
32 Patton,J.R., Patterson,R.J. and Pederson,T. (1987) *Mol. Cell. Biol.* **7**, 4030–4037.
33 Zillmann,M., Zapp,M.L. and Berget,S.M. (1988) *Mol. Cell. Biol.* **8**, 814–821.
34 Frendeway,D. and Keller,W. (1985) *Cell* **42**, 355–367.
35 Melton,D.A., Krieg,P.A., Rebagliahi,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.* **12**, 7035–7056.
36 Lamond,A.I. and Sproat,B.S. (1994) In Higgins,S.J. and Hames,B.D. (eds), *RNA Processing: A Practical Approach.* IRL Press, Oxford, pp. 103–140.
37 Sgault,V., Will,C.L., Sproat,B.S. and Lührmann,R. (1995) *EMBO J.* **14**, 4010–4021.
38 Izaurralde,E., Lewis,J., McGuigan,C., Jankowska,M., Darzynkiewicz,E. and Mattaj,I.W. (1994) *Cell* **78**, 657–668.
39 Wersig,C. and Bindereif,A. (1992) *Mol. Cell. Biol.* **12**, 1460–1468.
40 Wolff,T. and Bindereif,A. (1992) *EMBO J.* **11**, 345–359.
41 Lewis,J.D., Izaurralde,E., Jarmolowski,A., McGuigan,C. and Mattaj,I.W. (1996) *Genes Dev.* **10**, 1683–1698.
42 Colot,H.V., Stutz,F. and Rosbash,M. (1996) *Genes Dev.* **10,** 1699–1708.
43 Rossi,F., Forne,T., Antoine,E., Tazi, J., Brunel,C. and Cathala,G. (1996) *J. Biol. Chem.*, **271**, 23985–23991.
44 Neubauer,G., Gottschalk,A., Fabrizio,P., Séraphin,B., Lührmann,R. and Mann,M. *Proc. Natl. Acad. Sci.*, in press.