The C-protein tetramer binds 230 to 240 nucleotides of pre-mRNA and nucleates the assembly of 40S heterogeneous nuclear ribonucleoprotein particles.

M Huang, J E Rech, S J Northington, P F Flicker, A Mayeda, A R Krainer and W M LeStourgeon

Mol. Cell. Biol. **1994, 14(1):518. DOI: 10.1128/MCB.14.1.518.**

MOLECULAR AND CELLULAR BIOLOGY, Jan. 1994, p. 518–533
0270-7306/94/$04.00+0

Vol. 14, No. 1

The C-Protein Tetramer Binds 230 to 240 Nucleotides of Pre-mRNA and Nucleates the Assembly of 40S Heterogeneous Nuclear Ribonucleoprotein Particles

M. HUANG,[1] J. E. RECH,[1] S. J. NORTHINGTON,[1] P. F. FLICKER,[1] A. MAYEDA,[2] A. R. KRAINER,[2] AND W. M. LESTOURGEON[1]*

Department of Molecular Biology, Vanderbilt University, Nashville, Tennessee 37235,[1] and Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724-2208[2]

Received 1 July 1993/Returned for modification 18 August 1993/Accepted 28 September 1993

A series of in vitro protein-RNA binding studies using purified native $(C1)_3C2$ and $(A2)_3B1$ tetramers, total soluble heterogeneous nuclear ribonucleoprotein (hnRNP), and pre-mRNA molecules differing in length and sequence have revealed that a single C-protein tetramer has an RNA site size of 230 to 240 nucleotides (nt). Two tetramers bind twice this RNA length, and three tetramers fold monoparticle lengths of RNA (700 nt) into a unique 19S triangular complex. In the absence of this unique structure, the basic A- and B-group proteins bind RNA to form several different artifactual structures which are not present in preparations of native hnRNP and which do not function in hnRNP assembly. Three $(A2)_3B1$ tetramers bind the 19S complex to form a 35S assembly intermediate. Following UV irradiation to immobilize the C proteins on the packaged RNA, the 19S triangular complex is recovered as a remnant structure from both native and reconstituted hnRNP particles. C protein-RNA complexes composed of three, six, or nine tetramers (one, two, or three triangular complexes) nucleate the stoichiometric assembly of monomer, dimer, and trimer hnRNP particles. The binding of C-protein tetramers to RNAs longer than 230 nt is through a self-cooperative combinatorial mode. RNA packaged in the 19S complex and in 40S hnRNP particles is efficiently spliced in vitro. These findings demonstrate that formation of the triangular C protein-RNA complex is an obligate first event in the in vitro and probably the in vivo assembly the 40S hnRNP core particle, and they provide insight into the mechanism through which the core proteins package 700-nt increments of RNA. These findings also demonstrate that unless excluded by other factors, the C proteins are likely to be located along the length of nascent transcripts.

Whether released from isolated nuclei by sonic disruption or by low-salt extraction, the majority of the pre-mRNA molecules remain dispersed in solution following chromatin removal by brief centrifugation. Under conditions of minimal nuclease activity, 70 to 95% of this RNA is recovered in large ribonucleoprotein (RNP) complexes which sediment from 30S to more than 200S (26, 39, 52). Electron micrographs of the faster-sedimenting complexes reveal 20- to 25-nm particles arranged either as an array of polyparticles (29, 36, 42, 43, 52, 53, 61) or as clusters of particles when nuclease activity is aggressively inhibited (56). Upon mild nuclease activity the polyparticle complexes are lost and the cleaved RNA (mostly 500- to 1,000-nucleotide [nt] fragments) is recovered in 20- to 25-nm 30S-40S monoparticles (heterogeneous nuclear RNP [hnRNP] particles or ribonucleosomes) (3, 10, 18, 56, 64). Monoparticles purified from HeLa nuclei via glycerol gradients are primarily composed of six abundant nuclear proteins (the core particle proteins) (10, 28, 54, 65), which exist as three heterotypic tetramers, $(A1)_3B2$, $(A2)_3B1$, and $(C1)_3C2$ (4, 7, 39). The $(A2)_3B1$ and $(C1)_3C2$ tetramers have been isolated and partially characterized (4, 7). The $(A1)_3B2$ tetramer has not been isolated, but its existence is inferred from chemical cross-linking studies which reveal that A1 exists in monoparticles as homotrimers (34, 41) and in a 3:1 ratio with B2. Like the histones (reviewed in reference 63), the core particle proteins are transcribed from multigene families (13, 20, 50), and they reveal charge and nonallelic variants (14, 15, 20) which can be resolved in various two-parameter electrophoretic systems (11, 17, 35, 38, 46, 65).

Isolated 40S monoparticles completely dissociate upon RNA digestion with nuclease (22, 64). Spontaneous reassembly occurs in vitro upon the addition of 700 ± 20 nt of exogenous RNA or single-stranded DNA (22). Multiples of this length support the spontaneous in vitro assembly of dimers, trimers, and polyparticle complexes (22, 39). Reconstituted particles possess the same sedimentation coefficient, protein stoichiometry, chemical and UV cross-linking properties, pattern of salt-induced protein dissociation, nuclease sensitivity, and ultrastructural morphology as native hnRNP (22, 27, 64). Most of the noncore proteins present in initial hnRNP preparations do not quantitatively reconstitute with the core particle proteins (22, 64), a finding which indicates that they are not primarily involved in the RNA packaging mechanism (64). The in vitro packaging activity of the core proteins, the existence of nascent transcripts in polyparticle complexes, and the nuclease-induced conversion of polyparticle complexes to 30S-40S monoparticles suggest that the core proteins function in vivo to package approximately 700-nt increments of pre-mRNA into regular structures. If this is true, then unless excluded, phased, or dissociated by other factors, the core proteins should generally be distributed along the length of nascent transcripts. Such a distribution has been observed for the A- and B-protein homologs of *Drosophila melanogaster* (2), and UV light-induced protein-RNA cross-linking studies have shown that the C proteins exists over the entire length of adenovirus late transcripts in vivo (62). This distribution is consistent with the

* Corresponding author. Phone: (615) 322-2588. Fax: (615) 343-6707. Electronic mail address: LESTOUWM@CTRVAX.VANDERBILT.EDU.

findings described here which indicate that purified C-protein tetramers bind RNA in vitro in a highly self-cooperative manner and are located along the entire length of packaged RNA. Further evidence for this binding mode in vivo is seen in chemical cross-linking studies which reveal that the C-protein tetramers exist in oligomeric arrays in isolated monoparticles (41) and in the hnRNP complexes present in splicing-competent nuclear extracts (33, 34). Cooperative RNA binding has been shown for purified $(A2)_3B1$ tetramers (7) and for purified core protein A1 (21). Prior to this report, the mechanism through which the core proteins measure 700-nt increments of RNA for packaging in 40S monoparticles has not been directly addressed.

Efforts to understand the topology of RNA and protein in monoparticles are dependent on a knowledge of the exact number of each tetramer per particle. On the basis of the mass of gradient-purified 40S monoparticles (about 1.5 million) (39), the individual protein masses, and the mass of a 700-nt packaged substrate, each monoparticle should contain three copies of each tetramer (39). Definitive evidence for this stoichiometry can be found in several findings described here. Specifically, only three C-protein tetramers will bind particle-length RNA (700 nt). Upon binding of the third tetramer, a folding event results in the formation of a 19S triad structure with apparent C_3 symmetry. Multiples of this RNA length support the assembly of polytriad complexes. The triangular 19S C protein-RNA complex functions as an obligate first intermediate in the in vitro assembly of 40S hnRNP particles, and it is recovered from native and reconstituted monoparticles as a remnant structure following dissociation of the A- and B-group proteins.

MATERIALS AND METHODS

Cell culture, labeling, and nuclear isolation. HeLa S3 cells (ATCC CCL 2.2) were cultured in 2-liter spinner flasks in modified Eagle's minimum essential medium (Earl's minimal essential medium suspension powder, with L-glutamine and without sodium bicarbonate; GIBCO Laboratories, Grand Island, N.Y.) supplemented with 5% bovine calf serum (HyClone, Logan, Utah), 0.09% pluronic F 68, penicillin G (6,000 U/liter), and streptomycin sulfate (60 μg/ml). To ensure a generation time of 24 to 28 h, the cells were harvested daily at a density of 5×10^5 to 6×10^5 cells per ml. Cells were stored at −70°C in 50% glycerol. For protein labeling, cells were cultured 24 h after addition of [^{35}S]methionine to a final concentration of 1.0 μCi/ml. Nuclei free of cytoplasmic tags and other cellular material (as judged by phase-contrast microscopy at a magnification of ×1,000) were isolated through a rapid two-step procedure described in detail elsewhere (6).

Isolation of 40S hnRNP particles and protein purification. 40S hnRNP particles were freed from isolated nuclei by brief sonic disruption and purified by sedimentation in either 15 to 30% (wt/vol) sucrose or 15 to 30% (vol/vol) glycerol gradients as described in detail elsewhere (6). The C-protein tetramers $(C1)_3C2$ were purified by a method modified from that of Barnett et al. (6). Gradient-purified 40S hnRNP particles were loaded on a Mono Q HR5/5 anion-exchange column (Pharmacia LKB Laboratories) equilibrated with buffer B (20 mM Tris-HCl [pH 8.0], 1 mM dithiothreitol, 200 mM NaCl). The column was eluted with a gradient of 200 to 600 mM NaCl in buffer B at a flow rate of 0.5 ml/min. Protein elution was monitored at 214 nm. The C protein tetramers elute at 460 mM NaCl. If an Econo-Pac Q column (Bio-Rad Laboratories, Richmond, Calif.) is used, the tetramers elute at about 350 mM NaCl. In some cases, the eluted protein is subjected to a second purification using the anion-exchange column. Small aliquots of the pooled fractions containing the purified protein are taken for sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and for quantification via A_{214}. Purification of the $(A2)_3B1$ tetramers was done essentially as described by Barnett et al. (7). Nuclear sonic extracts are incubated at 37°C for 10 min after addition of RNase A (5 μg/10^8 nuclei). The 43S rearrangement complex which spontaneously forms through the reassociation of $(A2)_3B1$ tetramers following RNA cleavage was then purified through 15 to 30% glycerol gradients. The 43S complex was then dialyzed for 4 h against pH 8.0 STM buffer (90 mM NaCl, 1.0 mM $MgCl_2$, 20 mM Tris) and adjusted to an optical density at 280 nm (OD_{280}) of 0.3 with pH 8.0 STM. After addition $CaCl_2$ and dithiothreitol to a concentration of 1 mM, the endogenous RNA remaining was digested with micrococcal nuclease (Boehringer Mannheim) for 1 h on ice at a concentration of 300 U/ml. This procedure yields soluble $(A2)_3B1$ tetramers which at temperatures above 10°C form insoluble helical fibers (6, 39, 41). The extent of fiber formation can be seen in a heterodisperse pattern of sedimentation for proteins A2 and B1 in glycerol gradients (for examples, see Fig. 7A and 9). Because 40S hnRNP assembly results from highly regular protein-protein and protein-RNA interactions, the amount of fiber formation (A2 and B1 loss) determines the yield of reconstituted 40S hnRNP. This explains why, upon fiber formation, increased amounts of the other core proteins remain at the top of glycerol gradients. Fibers of $(A2)_3B1$ do not form in high-salt-dissociated hnRNP preparations, and particle reconstitution is quantitative upon dialysis to conditions of low ionic strength (41).

Plasmid constructions and in vitro RNA synthesis. Plasmid pHBG709 was used as the template for the in vitro synthesis of three of the eight RNA substrates (192, 709, and 2,087 nt) used in these studies. The plasmid contains the entire human β-globin gene and was produced by the insertion of a 281-nt *Alu*I fragment from pBR322 via blunt-end ligation into the *Hin*dIII site of pHBG. Cleavage of pHBG709 with *Bam*HI and transcription with T7 polymerase yields the 709-nt transcript containing the pBR322 sequences, exon 1, intron 1, and part of exon 2. This transcript is spliced in vitro and functions in the efficient in vitro assembly of 40S hnRNP particles. The 192-nt transcript contains only pBR322 sequences and is obtained after cleavage with *Tth*111I. The 2,087-nt RNA, produced after cleavage with *Eco*RI, contains the three exons and two introns of the human β-globin gene. The 685-nt transcript used in some studies was obtained by digesting plasmid pGL737b (provided by G. Mosig) with *Bst*EII. This transcript, produced by T7 polymerase, is of the gene 22/23 region of bacteriophage T4 and contains no eukaryotic RNA processing signals. The 230- and 456-nt RNA substrates were synthesized by plasmid pMBG2020 as the template after truncation with *Mbo*II and *Bgl*I, respectively. The 230-nt RNA contains the first exon and part of the first intron of mouse β^{maj}-globin gene. The 456-nt transcript contains the first exon, the first intron, and part of the second exon. The 1,452-nt RNA was transcribed from plasmid pAD2MLT-1967 after truncation with *Bgl*I. It contains the first and second leaders and the first intron of the adenovirus type 2 major late transcript region (22). The 962-nt RNA was transcribed from plasmid pAD2MLT-1487/IVS after cleavage with *Bgl*I. The intron of this construct is shortened by 491 nt. The procedures for in vitro transcription and RNA purification are described elsewhere (22). Prior to their use as substrates in RNP assembly reactions,

all transcription products were examined for correct length and homogeneity by electrophoresis of 2.0 μg of RNA in 8 M urea sequencing-type gels and staining in ethidium bromide. In addition to these pre-mRNA transcripts, 9.7S homoribopolymers of uridine (average length, 280 nt) (Pharmacia) were used as a binding substrate.

In vitro RNP assembly. Purified C-protein tetramers were dialyzed against STE buffer (20 mM Tris-HCl [pH 8.0], 10 mM EDTA, 90 mM NaCl) for 4 h at 4°C in the presence of 0.001% phenylmethylsulfonyl fluoride. In most cases, C protein purified from a 5-ml preparation of hnRNP particles (adjusted to an OD_{260} of 1.0) was mixed with in vitro-synthesized pre-mRNAs of differing length. The final protein concentration in the reaction mix was approximately 10.5 μg/ml, and the RNA concentration was 50 μg/ml. For the assembly of the 35S complex, purified $(A2)_3B1$ tetramers were present at different concentrations prior to addition of gradient-purified 19S C protein-RNA complexes. In these experiments, the amount of added 19S complex was determined by OD_{280} measurements. The 35S complex was also assembled in additional experiments by mixing purified $(A2)_3B1$ and $(C1)_3C2$ tetramers prior to addition of the 709-nt transcript. The protein-RNA mixtures were held on ice for 1 h and then loaded on 15 to 30% (vol/vol) glycerol gradients in STE buffer. It is convenient but not necessary to hold the samples on ice for 1 h. The in vitro assembly of 40S hnRNP particles was essentially as described previously (6). In experiments in which preformed C protein-RNA complexes were used to nucleate hnRNP assembly, the C protein-RNA complexes were allowed to assemble under conditions of excess protein and then purified through sedimentation in glycerol gradients. Through this procedure, no unbound RNA was present upon addition of these complexes to preparations of dissociated total hnRNP or to purified $(A2)_3B1$ tetramers. In these experiments, 1.0 ml of the various C protein-RNA complexes (with an OD_{260} of 0.2 to 0.3) was mixed with 1.0 ml of soluble 40S hnRNP monoparticles with an OD_{260} of 1.0.

In all of the in vitro RNA binding studies so far conducted, purified C protein binds RNA in the presence or absence of ATP and/or divalent cation whether or not eukaryotic RNA processing signals are present. For example, the T4 transcripts and the GC-rich 192-nt pBR transcripts contain no introns or eukaryotic RNA processing signals. These findings are in disagreement with the report of Swanson and Dreyfuss (59) that C protein does not bind in vitro to bacterial RNA or to transcripts of the β-globin thalassemia mutant lacking the polypyrimidine sequence. In addition, we find that the binding of C protein to homoribopolymers of uridine (average length of 268 nt) is fully reversible at 0.5 M salt. This is in disagreement with the finding of Swanson and Dreyfuss (58) that C protein-poly(U) complexes cannot be dissociated in 2.0 M salt but require boiling in SDS-mercaptoethanol preparations to effect dissociation.

Isopycnic density determination of C protein-RNA complexes. C protein-RNA complexes assembled on transcripts of 192, 230, 456, 512, and 709 nt were purified by sedimentation through glycerol gradients as described above. The C proteins were covalently cross-linked to the bound RNA by 15 to 20 min of UV irradiation. The samples were placed on a rotating platform 9 cm below a 15-W Minerallight lamp with a maximum spectral output at 254 nm (Ultraviolet Products, Inc.) and irradiated at 2,000 mW/cm^3. Following dialysis against STE (pH 8.0) and concentration in Centriprep-30 membrane concentrators (Amicon, Lexington, Mass.), the complexes were layered onto preformed CsCl gradients. The CsCl concentration in the preformed gradient was 49 to 55.2%, corresponding to a density of 1.48 to 1.63 g/cm^3 from top to bottom. Gradients were centrifuged at 40,000 rpm at 22°C for 70 h. The gradients were fractionated, and aliquots of each fraction were taken for density determinations and for SDS-PAGE analysis after dialysis and precipitation with alcohol to confirm that the distribution of C protein was coincident with RNA as monitored by counts per minute of [^{3}H]uridine. Prior to electrophoresis, the RNA was digested with RNase A (300 U of enzyme for each 1 ml of sample with an OD_{260} of 1.0). The samples were digested for 40 min at 25°C. This procedure was also used to confirm that UV-induced C protein-RNA cross-linking was quantitative and that CsCl did not cause C-protein dissociation from the bound RNA.

In Vitro RNA splicing. To determine whether RNP-packaged transcripts are splicing competent in vitro, the 19S C protein-RNA complex was assembled on two [^{32}P]UTP-labeled transcripts of the human β-globin gene. The 709-nt transcript contains 281 nt of pBR322 sequence preceding exon 1 and terminates at the *Bam*HI site in exon 2. The 799-nt transcript, which does not contain the 5′ pBR322 sequences, contains a hybrid exon 2 and terminates in exon 3. Both transcripts were also used as substrates for the in vitro assembly of 40S hnRNP particles. The 709-nt transcript was also used for the assembly of the 35S complex which contains three copies of the $(C1)_3C2$ and $(A2)_3B1$ tetramers. In these experiments, the various RNP complexes were assembled under conditions of protein excess and purified through 15 to 30% glycerol gradients. This procedure ensures that no unbound RNA was present in the in vitro-assembled RNP complexes. The in vitro splicing assay was performed with nuclear extracts and S100 cytoplasmic extracts with added splicing factor SF2/ASF as described previously (44), and both protein-bound and protein-free RNAs were used as substrates.

Electron microscopy and gel filtration. Preparations of the C-protein tetramer were taken for electron microscopy directly after purification from the Mono Q column and after sedimentation of this material through 5 to 15% glycerol gradients. In some cases, samples were stored at −70°C. Generally the best negative staining results were obtained with freshly purified protein. This observation was true for both the purified tetramer and the various C protein-RNA complexes examined in this study. All C protein-RNA complexes were purified through 15 to 30% glycerol gradients. Both the tetramer and tetramer-RNA complexes were fixed for 10 min on ice with 1% glutaraldehyde at pH 7.0. Samples were applied to Formvar-carbon-coated grids that were glow discharged within 30 min of sample application. The grids were stained with filtered 2% uranyl acetate blotted and dried over a beaker of distilled water. In some experiments, ferritin was added for use as an internal size standard. Sample preparation and platinum shadowing were performed as described in detail elsewhere (66). In some experiments, tobacco mosaic virus or myosin was used as an internal standard at a final concentration of 20 to 25 μg/ml. Electron microscopy was performed on an Hitachi H-800 transmission electron microscope at 100 kV or on a Philips CM-12 transmission electron microscope at 120 kV. Micrographs were recorded at nominal magnifications of ×40,000 and ×70,000. Microscope calibrations were performed by photographing tropomyosin paracrystals at appropriate magnification and then measuring the 395-Å (39.5-nm) internal repeats. Measurements of the C-protein tetramer and of the various C protein-RNA complexes were made from nega-

tives by using a Nikon profile projector. On average, 150 measurements were made for each complex.

Our previous gel filtration experiments on the purified C-protein tetramer yielded an approximate Stokes radius of 6.2 nm. Upon observing an average tetramer diameter of 9 nm in negatively stained electron micrographs, we performed a second series of gel filtration experiments using Sephacryl S-300 (Pharmacia LKB Laboratories) instead of Bio-Gel A1.5 (Bio-Rad Laboratories) or Superose 6 (Pharmacia LKB Laboratories). In the first set of experiments, the columns were eluted with a Tris-HCl buffer (pH 8.0) containing 1.0 M NaCl. The internal standards used for Stokes radius determinations were thyroglobulin (8.1 nm), immunoglobulin G (5.2 nm), and bovine serum albumin (3.5 nm) (1). In recent experiments, the Sephacryl S-300 columns were eluted at several salt concentrations from 200 to 900 mM, and bovine liver catalase (232 kDa; Stokes radius of 5.2 nm) and horse spleen ferritin (481 kDa; Stokes radius of 6.71 nm) (23) were used as standards. Calculations to approximate the molecular weight and Stokes radius were as described elsewhere (16, 23). In these experiments, as in our previous studies, the C-protein tetramer elutes from Sephacryl S-300 columns just before ferritin. This material was examined in negatively stained electron micrographs, and its sedimentation in 15 to 30% glycerol gradients was compared with that of the C-protein tetramers purified via ion-exchange chromatography.

RESULTS

Morphology and size of the C-protein tetramer. The RNA-free C-protein tetramers $(C1)_3C2$, purified through ion-exchange chromatography, represent more than 98% of the Coomassie blue-stained material which can be detected through SDS-PAGE at a protein load of 50 μg per lane (5, 6). Samples of this material reveal spherical or slightly ellipsoidal structures in negatively stained electron micrographs which range in size from 8.5 to 11 nm, with a two-dimension average diameter of 9 nm. Sixty-five to 70% of the tetramers observed in micrographs do not clearly reveal their subunit morphology. The remaining tetramers appear more dispersed on the grids (ranging from 10 to 11 nm), and these more clearly reveal a four-subunit morphology (Fig. 1A). One of the four subunits is often poorly resolved in electron micrographs. This may reflect an altered orientation of one of the four polypeptides or the orientation of the tetramer on the grid. The C2 polypeptide, with its 13-amino-acid insert (12, 45), confers compositional polarity to the tetramer.

The dimensions of the C-protein tetramer in negatively stained electron micrographs are smaller than the overall size of the tetramer in solution. The C-protein tetramer (in 90 to 500 mM salt) elutes from Sephacryl S-300 columns as a symmetrical peak just before ferritin (4, 5) with an approximate Stokes radius of 6.8 nm (23). This corresponds to an overall hydrated dimension of 13.6 nm. For comparison, the diameter of the three doughnut-like ferritin molecules shown in Fig. 1A is 12 nm (31), and their hydrated size is 13.4 nm (23). The tetramer's anomalously large Stokes radius in solution may be due in part to the presence of 41 acidic and 21 basic residues in the 130-residue carboxy terminus of C1 and C2 (12) and to the hydration shell associated with these four highly charged domains. The C-protein tetramer does not dissociate into free polypeptides in 10 mM to 2.0 M salt-containing buffers. It does dissociate in 6.0 M urea and in SDS-containing solutions in the absence of reducing reagent.

The C-protein tetramer has an RNA binding site size of 230 to 240 nt, and three tetramers fold 700 nt of RNA into a triangular 19S hnRNP assembly intermediate. When purified C-protein tetramers bind 230-nt transcripts of the mouse β^{maj}-globin gene, C protein-RNA complexes which appear as slightly ellipsoidal structures in electron micrographs with an average diameter of 9.7 nm spontaneously form (Fig. 1B). These complexes sediment in 15 to 30% glycerol gradients at approximately 11.5S (Fig. 2). Their density in CsCl is 1.51 g/cm^3 (not shown). When two tetramers bind 456-nt transcripts of the mouse β^{maj}-globin gene, two morphologically similar but unique complexes are formed. At stoichiometric protein/RNA ratios, dimers of the complexes seen in Fig. 1B are observed. In the presence of excess protein, the dimers reveal a characteristic angular orientation of the two RNA-bound tetramers, and an additional small structure (thought to be a protein-free tetramer) is located centrally between the two RNA-bound tetramers (Fig. 1C). These structures are shown here because they are morphologically consistent with the structures which form when three tetramers bind RNA and because they may provide insight into the mechanism of tetramer association (see Discussion). Like the one-tetramer C protein-RNA complex, the two-tetramer complex sediments at approximately 11.5S (Fig. 2A). Also like the one-tetramer complex, the two-tetramer complex has a density of 1.51 g/cm^3 in CsCl (not shown). Three C-protein tetramers bind 709-nt transcripts of the human β-globin gene to form a unique triangular complex (Fig. 1D) that sediments at 19S in glycerol gradients (Fig. 2). While the three-tetramer complex sediments appreciably faster than the one- and two-tetramer complexes, its density (1.50 g/cm^3) is essentially the same. About 80% of the structures in a typical field appear as rather tightly packed symmetrical complexes (Fig. 1D). The remaining complexes are more dispersed, and three globular subunits can be counted. Rotary-shadowed preparations of the three-tetramer 19S C protein-RNA complexes clearly reveal their triangular topology (Fig. 3). In both negatively stained and rotary-shadowed preparations, one side of the triangle is slightly shorter (arranged as the base in Fig. 1D and 3). The longer sides of the triangle complexes measure on average 2.3 nm in negatively stained preparations and 2.2 nm in rotary-shadowed preparations.

The common densities of the one-, two-, and three-tetramer C protein-RNA complexes which assemble on 230-, 456-, and 709-nt transcripts together with their characteristic one-, two-, and three-subunit morphology argues that a single tetramer binds each 230- to 240-nt increment of RNA. This is confirmed by the finding that two- and three-triangular complexes assemble on transcripts of 1,452 and 2,087 nt (described below) and through velocity sedimentation analysis which yields a mass of 600 kDa for the 19S complex (Fig. 2B). More specifically, from their amino acid sequence (60), the mass of three C protein tetramers is 387.2 kDa and the mass of the 709-nt RNA substrate is 235 kDa. The difference between the experimental (600 kDa) and expected (622 kDa) masses of the 19S complex (which contains 9C1 and 3C2 polypeptides) is less than the mass of a single C1 polypeptide (31,931 Da).

In the presence of a 2.3-fold molar RNA excess, the C-protein tetramers bind 709-nt transcripts to form only the 19S triangular RNP complex. In these experiments, no unbound protein can be detected at 5.8S (the position of native tetramer in gradients), as a pellet in the gradient tubes, and no intermediate-size C protein-RNA complexes (possessing one or two tetramers per transcript) are formed.

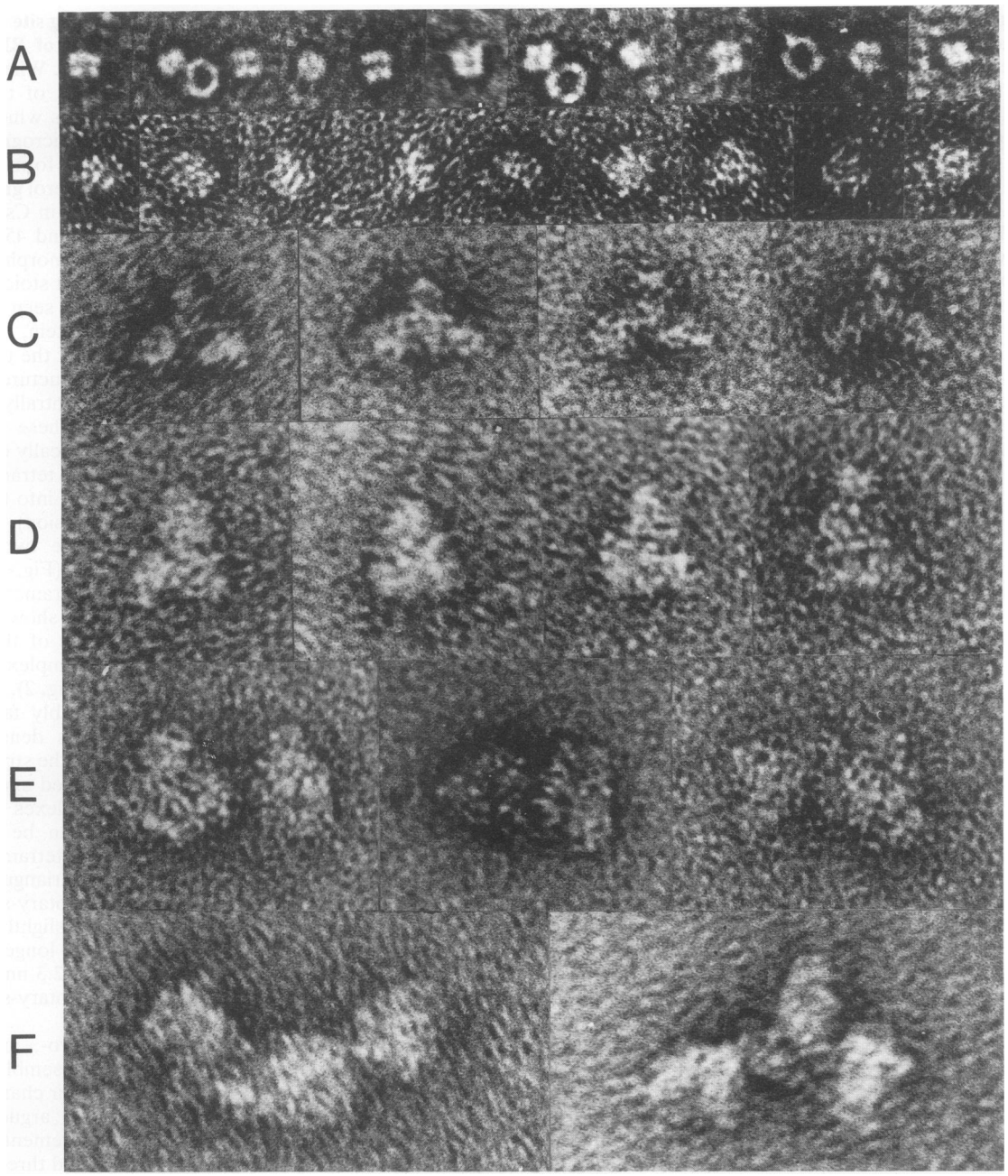

FIG. 1. Electron micrographs of negatively stained purified RNA-free C-protein tetramers (A). The holes in the doughnut-like horse spleen ferritin molecules added as an internal size standard are about 8 nm (23). (B to F) Structures which form when purified C-protein tetramers bind pre-mRNAs of increasing length. (B) A collection of the slightly ellipsoidal structures typical of those observed when a single C-protein tetramer binds a 230-nt pre-mRNA transcript of the mouse β^{maj}-globin gene. (C) A collection of structures typical of those formed when two tetramers bind the 456-nt transcript of the same gene. Each of the two subunits are similar in size and morphology to the complexes shown in panel A. Note the angular association of the two subunits and the putative RNA-free tetramer located centrally and above the subunits. (D) The triangular 19S C protein-RNA complexes which form when three C-protein tetramers bind a 709-nt transcript of the human β-globin gene. In panel D, the triangular 19S complexes are arranged with the shorter side as the base (measuring on average 18.2 nm). The isosceles sides of the triangles measure 23.0 nm. (E and F) The two (E)- and three (F)-triangle structures which form when six and nine C-protein tetramers bind 1,452-nt adenovirus transcripts and 2,087-nt transcripts of the human β-globin gene. Note that the two and three contiguous triangular complexes are of the same general size and morphology as the triangular 19S C protein-RNA complexes which form on the 709-nt transcripts. The structures shown in panel B are representative of more than 90% of the complexes seen in electron micrographs. The remaining complexes are apparent aggregates of these complexes. The structures shown in panel C represent about half of the complexes observed in micrographs. The remaining complexes appear as apparent aggregates of these structures. The structures shown in panel D are typical of about 80% of the structures seen in micrographs. The remaining structures appear more dispersed, and the three subunits are more readily observed. The structures shown in panels E and F represent about 60% of the complexes observed in micrographs. Most of the remaining material appears either as smaller globular masses presumably due to nonspread complexes. As described in the text, regardless of the RNA substrate used in these in vitro assembly experiments, almost all of the protein-bound RNA sediments as structurally homogeneous material. Bar = 30 nm for panel A and 19 nm for panels B to F.

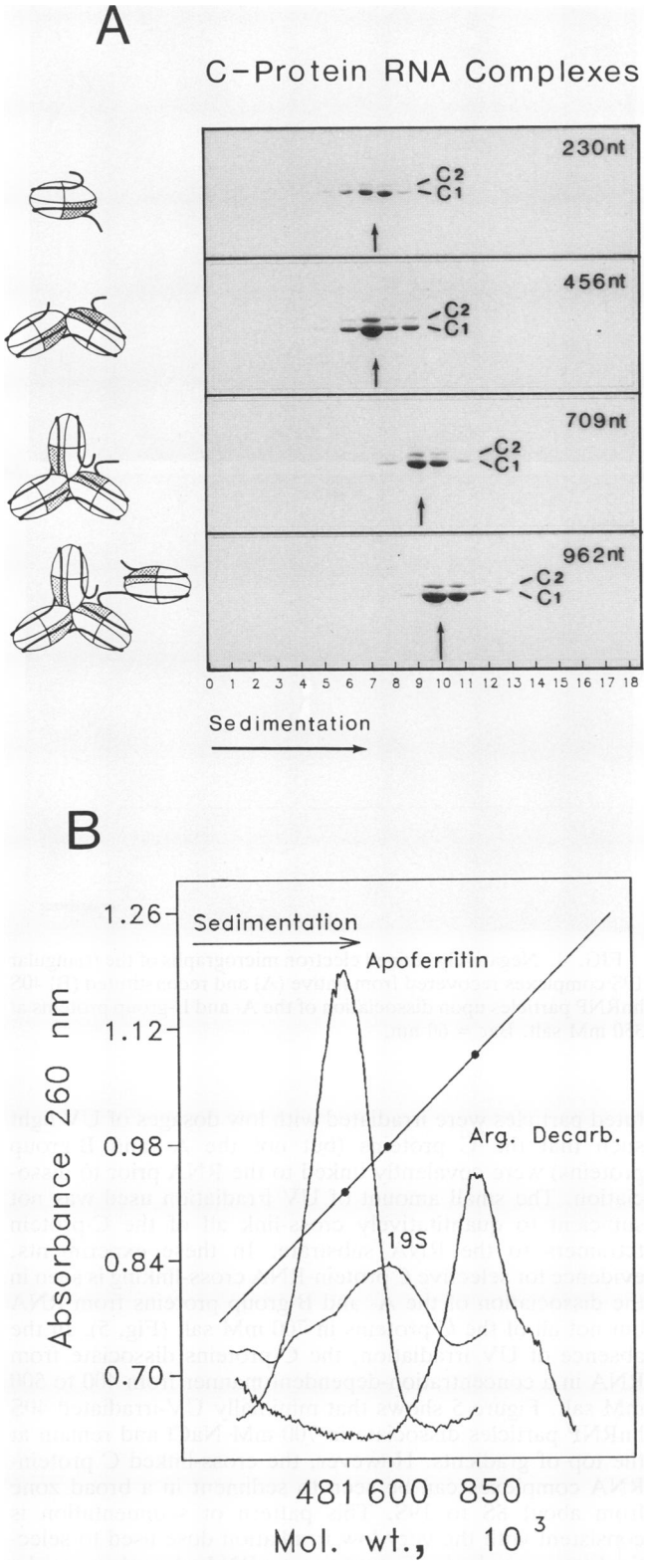

FIG. 2. (A) Four Coomassie blue-stained SDS-polyacrylamide gels showing the distribution in 15 to 30% linear glycerol gradients of the C protein-RNA complexes which spontaneously assemble on increasing lengths of RNA. The arrows denote respective sedimentation coefficients of 11.5S, 11.5S, 19S, and 24S for the C protein-RNA complexes which form on transcripts of 230, 456, 709, and 962 nt. Although the arrangement of C-protein tetramers to the left of each gel is consistent with the information presented here, it is intended only to denote tetramer number and not the absolute arrangement of tetramers or bound RNA in each RNP complex. The shaded subunit denotes the C2 polypeptide. (B) The sedimentation of apoferritin (481.2 kDa), arginine decarboxylase (Arg. Decarb.; 850 kDa), and the triangular three-tetramer C protein-RNA complex in parallel linear 5 to 15% glycerol gradients. The triangular three-tetramer complex sediments at 19S between the two standards with a relative mass of 600 kDa. The position of the 19S complex in the gradients was determined by monitoring OD_{260} distribution and confirmed by SDS-PAGE. The slowly sedimenting shoulder of the 19S peak is a small amount of free 709-nt RNA, as in these experiments the complex was assembled under conditions of excess RNA and was not gradient purified prior to sedimentation. The position of the standards was determined by OD_{280} measurements. The apoferritin peak corresponded to an OD_{280} of 0.3 (the absorbance scale measured at 280 nm is not shown).

These findings demonstrate that the great majority, if not all, of the purified C protein used in these in vitro binding studies is competent to bind RNA in vitro and that the three-tetramer triangular complex is a thermodynamically favored structure. The absence of intermediate complexes possessing one or two tetramers per transcript (under conditions of RNA excess) indicates and that the tetramers bind RNA in a self-cooperative mode. The C proteins bound in the 19S triangular complex dissociate from RNA at 400 to 500 mM salt as do the C proteins of native 40S monoparticles.

Evidence that a distinct folding event actually occurs when the third tetramer binds RNA can also be seen in the sedimentation properties of the C protein-RNA complexes which form on increasing lengths of RNA. Figure 2A shows the sedimentation of one-, two-, three-, and four-tetramer complexes in parallel 15 to 30% glycerol gradients. The one- and two-tetramer complexes have very similar sedimentation properties. This is consistent with a tetramer arrangement depicted for the 456-nt RNP complex in Fig. 2A and with the actual arrangement seen in electron micrographs of these structures (Fig. 1C and D). The increase in RNA and protein mass is associated with a concomitant increase in the frictional coefficient of the complex. However, a significant increase in sedimentation rate, consistent with a reduction in frictional coefficient and increased mass, is associated with the binding of the third tetramer to 709-nt transcripts.

C-protein tetramers package long lengths of RNA into oligomers of triangular complexes. Having observed that a folding event follows the binding of a third C-protein tetramer to 700-nt lengths of RNA, we wished to determine whether successive groups of three tetramers associate in a combinatorial mode to fold longer RNA substrates into a contiguous array of triangle complexes. This event would provide mechanistic insight into the finding that when long RNAs are added to preparations of dissociated total hnRNP, a contiguous array of 40S hnRNP particles spontaneously assemble in vitro (22). Figures 1E and F show that purified C-protein tetramers spontaneously package two- and three-triangle-length RNAs (1,452 and 2,087 nt, respectively) into two- and three-triangle complexes. In all electron micrographs examined to date, no protein-free regions of RNA are seen to separate the triangle complexes.

Triangular 19S complexes are recovered from native and reconstituted 40S hnRNP particles. As shown above, three C-protein tetramers possess the intrinsic ability to fold approximately 700-nt increments of RNA into 19S triangular complexes in vitro, and previous studies have shown that

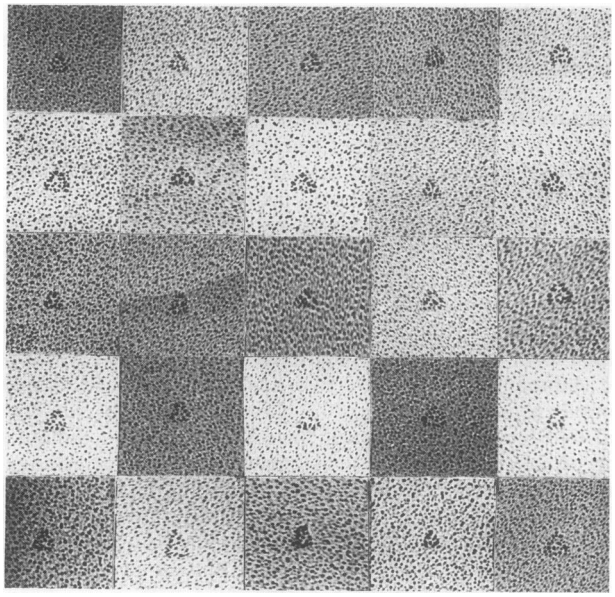

FIG. 3. A collection of single triangular structures from four electron micrographs of rotary-shadowed 19S C protein-RNA complexes. In these micrographs, the average length of an isosceles side is 22.5 nm. Protein-free lengths of RNA were not detected in association with these structures.

700 ± 20-nt increments of RNA are packaged in each 40S monoparticle (22). If the C-protein-driven folding event exists as the mechanism for measuring monoparticle lengths of RNA and is a bona fide step in the assembly pathway leading to 40S monoparticles, then it may be possible to recover these structures from native hnRNP complexes. To determine whether the 19S complex exists in hnRNP particles, a series of salt-induced dissociation studies were performed. Previous studies have shown that the A- and B-group proteins dissociate from native and reconstituted 40S hnRNP particles at salt concentrations between 150 and 250 mM and that the C proteins are the only core proteins bound to RNA at 350 mM salt (10). In the previous study, the sedimentation coefficient of the C protein-RNA complex which remains in high salt was not determined, and these structures were not examined with the electron microscope. In the present characterization, it was observed that the salt-resistant C protein-RNA complex cosediments with the triangular 19S complex, and electron micrographs of material taken from these gradient regions reveal the characteristic triangular 19S complex (Fig. 4). Electron micrographs reveal that more and better-defined triangles are consistently recovered from the reconstituted monoparticles (Fig. 4B) than from native gradient-isolated 40S hnRNP particles (Fig. 4A). This is consistent with the fact that the RNA substrate in reconstituted 40S hnRNP is homogeneous in length (i.e., 685, 709, or 726 nt; see Materials and Methods), while the RNA recovered from native hnRNP is nicked and heterogeneous in length. Most of the RNA recovered from native 40S monoparticles ranges from 600 to about 1,000 nt in length (10, 22).

The recovery of triangular 19S C protein-RNA complexes from native and reconstituted hnRNP particles may be dependent on salt-induced protein rearrangements on the RNA substrate following dissociation of the A- and B-group proteins. To address this question, both native and reconstituted particles were irradiated with low dosages of UV light such that the C proteins (but not the A- and B-group proteins) were covalently linked to the RNA prior to dissociation. The small amount of UV irradiation used was not sufficient to quantitatively cross-link all of the C-protein tetramers to the RNA substrate. In these experiments, evidence for selective C protein-RNA cross-linking is seen in the dissociation of the A- and B-group proteins from RNA but not all of the C proteins in 700 mM salt (Fig. 5). In the absence of UV irradiation, the C proteins dissociate from RNA in a concentration-dependent manner from 400 to 500 mM salt. Figure 5 shows that minimally UV-irradiated 40S hnRNP particles dissociate in 700 mM NaCl and remain at the top of gradients. However, the cross-linked C protein-RNA complexes can be seen to sediment in a broad zone from about 8S to 19S. This pattern of sedimentation is consistent with the very low irradiation dose used to selectively cross-link the C proteins to RNA. In other words, some of the RNA substrates possess one, two, and three cross-linked tetramers. Electron micrographs of samples taken from the 19S region of the gradient reveal typical triangle-shaped 19S C protein-RNA complexes (Fig. 6).

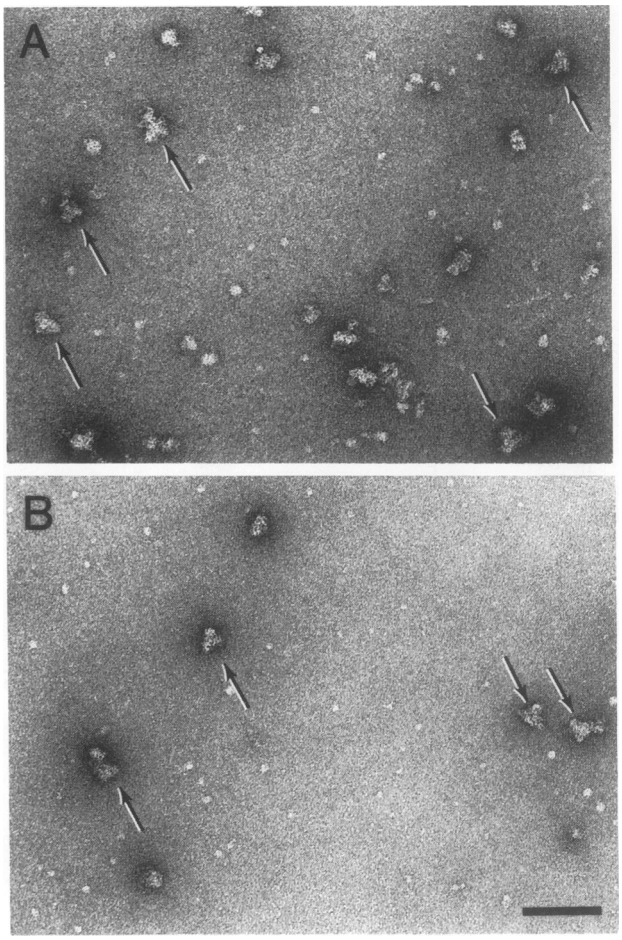

FIG. 4. Negatively stained electron micrographs of the triangular 19S complexes recovered from native (A) and reconstituted (B) 40S hnRNP particles upon dissociation of the A- and B-group proteins at 350 mM salt. Bar = 60 nm.

The triangular 19S C protein-RNA complex nucleates the correct assembly of 40S hnRNP particles in vitro. The recov-

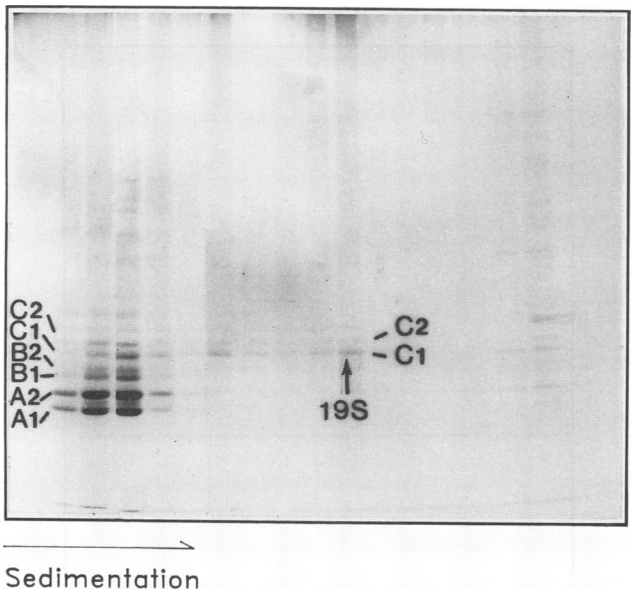

FIG. 5. A Coomassie blue-stained gel showing the sedimentation of low-dose UV-irradiated 40S hnRNP particles in a 15 to 30% glycerol gradient containing 700 mM salt. Following low-dose irradiation, it can be seen that the A- and B-group proteins dissociate at 700 mM salt and remain in the top gradient fractions. The faster-sedimenting material (near the middle of the gradient) contains C protein-RNA complexes. The C protein-RNA complexes which sediment in a disperse manner from 5.8S (near the top) to 19S contain one or two cross-linked C-protein tetramers per transcript. See Fig. 6 for electron micrographs of the 19S material.

ery of triangular C protein-RNA complexes from native hnRNP complexes indicates that purified C-protein tetramers bind RNA in vivo in a physiologically relevant manner. Additional evidence that the 19S complex is a physiologically relevant structure is seen in its ability to nucleate the

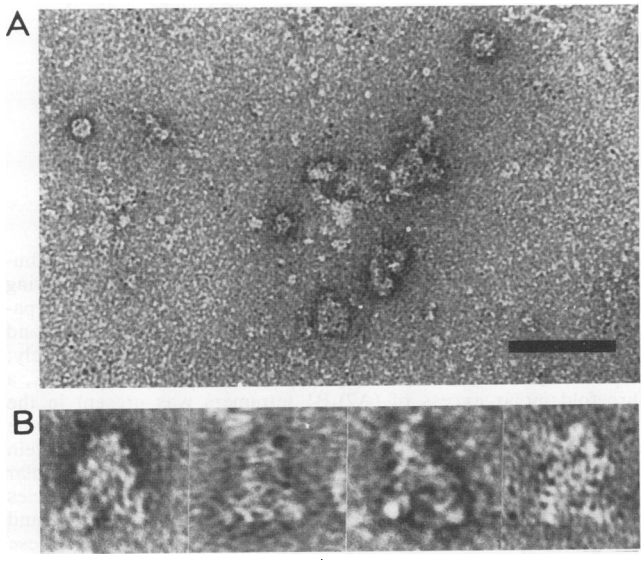

FIG. 6. Electron micrographs of negatively stained C protein-RNA complexes which sediment at 19S following UV irradiation and salt dissociation of 40S hnRNP particles. (A) A typical field containing three 19S complexes (bar = 60 nm); (B) a twofold enlargement of these complexes arranged with the shorter side as the base.

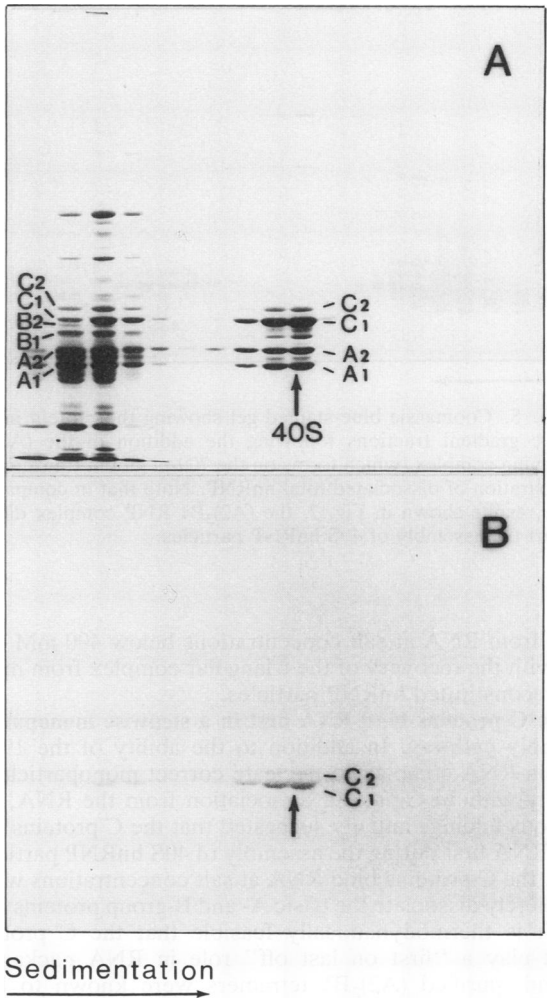

FIG. 7. (A) Coomassie blue-stained gel of the proteins in successive 15 to 30% glycerol gradient fractions following the sedimentation of the 40S hnRNP particles which spontaneously assemble when [^{35}S]methionine-labeled gradient-purified 19S C protein-RNA complexes are used as the assembly substrate; (B) autoradiogram of the gel shown in panel A. The autoradiogram shows the distribution of the radiolabeled C protein.

stoichiometric assembly of 40S hnRNP particles in vitro when added to preparations of soluble core particle proteins. In Fig. 7 are shown the results of a typical experiment in which purified 19S C protein-RNA complexes (instead of protein-free RNA) were added to a crude preparation of nuclease-dissociated hnRNP particles. The 40S monoparticles that assemble on the pre-formed 19S complex are seen in Fig. 7A to sediment in 15 to 30% glycerol gradients as do intact 40S hnRNP particles and to possess typical core particle stoichiometry. In this experiment, the C protein used to form the 19S complex was purified from cells grown for 24 h in [^{35}S]methionine. The autoradiogram (Fig. 7B) demonstrates that almost all of the prebound C protein sediments with the in vitro-assembled monoparticles. This finding demonstrates that during monoparticle assembly, the prebound C proteins do not exchange with the nonlabeled C protein present in the preparation of total dissociated hnRNP. This finding is consistent with the stability of the 19S complex, with the inability to detect C protein dissoci-

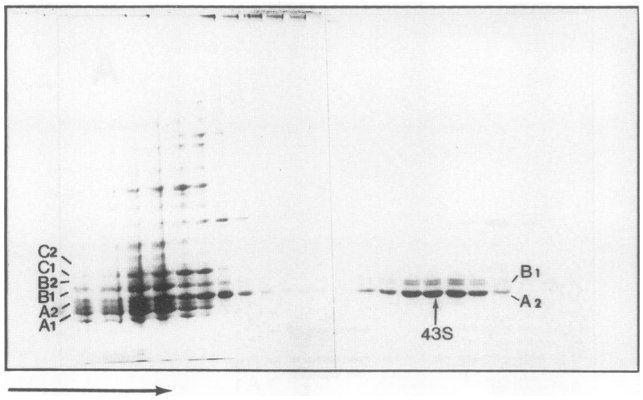

FIG. 8. Coomassie blue-stained gel showing the protein in successive gradient fractions following the addition of the $(A2)_3B1$ containing complex (which forms on the 709-nt globin transcript) to a preparation of dissociated total hnRNP. Note that in comparison to the results shown in Fig. 7, the $(A2)_3B1$ RNP complex did not support the assembly of 40S hnRNP particles.

ation from RNA at salt concentrations below 400 mM (10), and with the recovery of the triangular complex from native and reconstituted hnRNP particles.

The C proteins bind RNA first in a stepwise monoparticle assembly pathway. In addition to the ability of the 19S C protein-RNA complex to nucleate correct monoparticle assembly without C-protein dissociation from the RNA, two previous findings initially suggested that the C proteins may bind RNA first during the assembly of 40S hnRNP particles. First, the C proteins bind RNA at salt concentrations which completely dissociate the basic A- and B-group proteins (10), and it is thermodynamically feasible that the C proteins could play a "first on last off" role in RNA packaging. Second, purified $(A2)_3B1$ tetramers were known to bind RNA in vitro to form artifactual 43S complexes which are not present in native monoparticle preparations (41). Described here are three experimental observations which indicate that the C-protein tetramers bind RNA first during 40S hnRNP assembly in vitro and direct the stoichiometric association of the basic A- and B-group proteins.

(i) Unlike the C proteins, RNA prebound by purified $(A2)_3B1$ tetramers does not support the in vitro assembly of hnRNP particles. In these experiments, purified native $(A2)_3B1$ tetramers were allowed to bind monoparticle-length transcripts, and the resulting 43S $(A2)_3B1$-RNA complex was added to a preparation of total dissociated 40S hnRNP (as described above for the 19S C protein-RNA complex). The results shown in Fig. 8 demonstrate that RNA prebound by proteins A2 and B1 does not function as a substrate for the binding of the other core particle proteins.

(ii) When purified 19S C protein-RNA complexes are added to purified $(A2)_3B1$ tetramers in vitro such that equal molar amounts of the two tetramer types are present, a 35S RNP complex spontaneously forms which, like native hnRNP, contains a 1:1 molar ratio of the $(A2)_3B1$ and $(C1)_3C2$ tetramers (Fig. 9B). The protein composition and stoichiometry of the 35S RNP complex are like those of 40S monoparticles that are missing the major core protein A1. As in the case of native and reconstituted 40S hnRNP particles, the 19S triangular complex can be recovered from the 35S complex following salt-induced dissociation of the $(A2)_3B1$ tetramers (not shown). If a twofold molar excess of $(A2)_3B1$

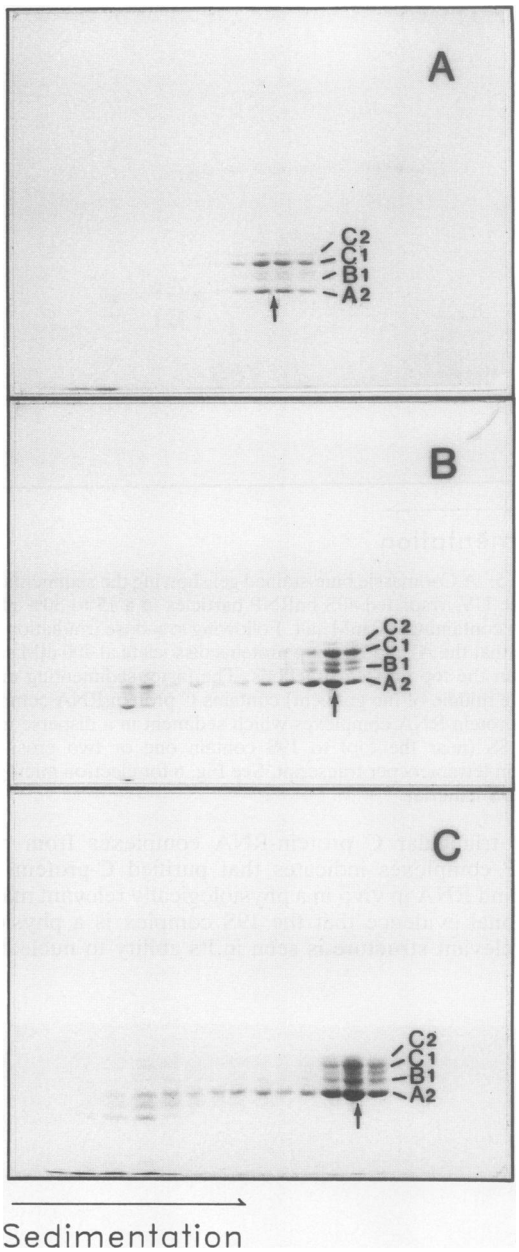

FIG. 9. Three Coomassie blue-stained gels showing the distribution of RNP complexes in successive gradient fractions following the addition of preformed 19S C protein-RNA complexes to preparations of purified $(A2)_3B1$ tetramers. In panel A, the $(C1)_3C2$ and $(A2)_3B1$ tetramers were present at a molar ratio of 2:1, respectively; in panel B, the molar ratios were approximately 1:1; in panel C, a threefold molar excess of $(A2)_3B1$ tetramers was present in the assembly mix. For panels A, B, and C, the vertical arrows denote the 23S, 35S, and 40S positions of these gradients. The protein composition of the 35S RNP complex which formed at a 1:1 molar ratio of the two tetramer types (B) looks like 40S hnRNP particles except for the absence of proteins A1 and B2. The protein A2- and B1-rich complex which sediments at 40S (C) may contain these proteins at sites normally occupied by the closely related $(A1)_3B2$ tetramers. Although in significant molar excess, the $(A2)_3B1$ tetramers do not displace the prebound C proteins from the RNA substrate.

tetramers exists prior to addition of the triangular complex, then a particle is formed that contains twice the normal complement of $(A2)_3B1$ tetramers (Fig. 9C). This complex sediments at 40S but lacks the basic proteins A1 and B2. This suggests that $(A2)_3B1$ tetramers may bind at sites normally bound by proteins A1 and B2.

(iii) The stoichiometric in vitro assembly of 40S hnRNP particles occurs only when the 19S C protein-RNA complex is used as the assembly substrate. In these experiments, C protein-RNA complexes composed of one to four tetramers (formed on transcripts of 230, 456, 709, and 962 nt) were added to aliquots of dissociated hnRNP to determine whether stoichiometric 40S monoparticle assembly is actually dependent on the presence of the three-tetramer 19S complex. In Fig. 10 it can be seen that single-tetramer complexes nucleate the assembly of fast-sedimenting A2/B1-rich complexes. This is an expected result because proteins A2 and B1 alone bind short lengths of RNA to form artifactual 43S complexes (22). The two-tetramer complex appears to nucleate the assembly of two structures. The more slowly sedimenting structure (immediately to the left of the arrow in Fig. 10B) appears to possess almost correct core protein stoichiometry. Because this structure assembles on a substrate that is two-thirds the length of RNA in 40S monoparticles and sediments more slowly than monoparticles, it may represent a two-thirds particle. The faster-sedimenting material is an artifactual C protein-rich complex. As in Fig. 7, it can be seen in Fig. 10 that the three-tetramer 19S C protein-RNA complex nucleates the correct assembly of stoichiometric 40S hnRNP. When a four-tetramer C protein-RNA complex is used as the assembly substrate, a C-protein-rich complex forms, which sediments slightly faster than 40S monoparticles (Fig. 10D). This is consistent with the formation of a 40S monoparticle that contains a fourth C-protein tetramer bound to the extra 262 nt of RNA. In these and other in vitro assembly experiments, various amounts of protein A2 and B1 can be seen to sediment in a heterodisperse manner between the pool of soluble protein near the top of gradients to regions near the bottom of the gradient. This is due to the artifactual assembly of RNA-free fibers composed entirely of $(A2)_3B1$ tetramers (6, 39, 41). Fibers spontaneously assemble when intact monoparticles are dissociated by digesting the RNA substrate with nuclease. Fiber formation can be retarded but not completely inhibited by conducting the assembly reactions at 0°C (see Materials and Methods).

RNA packaged in the 19S C protein-RNA complex, in the 35S complex, and in 40S hnRNP particles is efficiently spliced in vitro. The findings described above indicate that the C proteins initially package RNA through a length-dependent self-cooperative binding mode and that folding interactions between three contiguous tetramers define the architecture of a repeating array of regular particles. The findings presented above also demonstrate that the triangular C protein-RNA complex is highly stable in solution and that once it is formed, little exchange between bound and unbound protein occurs. Therefore, to determine whether the RNA packaged in the 19S C protein-RNA complex, in the 35S assembly intermediate, and in reconstituted 40S hnRNP particles is accessible to *trans*-acting factors which function in RNA splicing, these complexes were tested for their ability to support in vitro RNA splicing.

The results shown in Fig. 11 reveal that the 709-nt transcript of the human β-globin gene is efficiently spliced in nuclear extracts and in cytoplasmic extracts (containing added splicing factor SF2/ASF). In these experiments, the in vitro-assembled RNP complexes were allowed to form under conditions of protein excess and then purified via sedimentation in 15 to 30% glycerol gradients to ensure that no protein-free RNA was present. Because of the large amount of RNA necessary for preparative gradient isolation of RNP complexes, the pre-mRNA was not capped. Capping is not required for splicing in this system, although uncapped RNA is more susceptible to 5′ and 3′ exonucleolytic degradation (37). Equal amounts of RNP-packaged and protein-free RNA were assayed in parallel, and product yield was monitored at 15, 30, and 60 min. Figure 11 reveals that the intermediate intron-exon 2 product either is formed more efficiently in the S100 extracts or is more slowly degraded than in nuclear extracts. The amount of spliced product may be slightly higher for RNAs spliced in S100 preparations than in nuclear extracts, and more spliced product remains after 60 min in the case of RNP-packaged RNA compared with protein-free RNA. The 252-nt RNA is generated by cleavage at a site in intervening sequence 1 which contains an AG dinucleotide (37). This site resembles the consensus for 3′ splice junctions. The 252-nt fragment is a major aberrant product observed in splicing reactions utilizing noncapped substrate (37). This product and exon 1 either form more efficiently or are more stable in nuclear extracts than in S100 preparations. The amplification of this phenomenon in RNP-packaged transcripts (in comparison with protein-free RNA) may indicate enhanced stability in nuclear extracts rather than an enhanced rate of formation. Somewhat more of the intermediate intron-exon 2 product appears after 30 min for RNP-packaged transcripts in nuclear extracts than for protein-free RNA. In other experiments, an 802-nt transcript of the human β-globin gene (substrate DUP-171 described in reference 24, which contains exon 1, intron 1, a hybrid exon 2, through part of exon 3) was efficiently spliced when packaged in 40S hnRNP particles, and a 497-nt transcript of human β-globin containing the first exon through part of exon 2 was spliced when packaged by the C proteins into a two-tetramer C-protein complex but not more efficiently than protein-free RNA. This was also true for the 709-nt transcript packaged into the 35S complex which contains three copies each of the $(C1)_3C2$ and $(A2)_3B1$ tetramers (not shown). Because ^{35}S-labeled C-protein tetramers do not spontaneously dissociate from the RNA substrate in solution or exchange with unlabeled C protein during monoparticle assembly, it is assumed (but not proven) that the same is true in these experiments.

DISCUSSION

The association of a single C-protein tetramer with 230 to 240 nt of RNA is demonstrated here through different but mutually supportive experimental approaches. First, the density and thus the protein/RNA ratio are constant for C protein-RNA complexes which assemble on approximate multiples of 235 nt. Second, the experimentally determined mass of the three-tetramer triangular 19S C protein-RNA complex (600 kDa) is very close to its expected mass (622 kDa, from the known mass of the RNA substrate and three C-protein tetramers). Finally, electron micrographs of the C protein-RNA complexes which form on RNA substrates ranging in length from 230 to 2,087 nt demonstrate that a single C-protein tetramer binds 230 to 240 nt of RNA and that a single three-tetramer triangular complex assembles on each 700 ± 20-nt increment of RNA. More specifically, if 2,087 nt, the length of RNA bound by nine tetramers (three triangle complexes), is divided by tetramer number, then

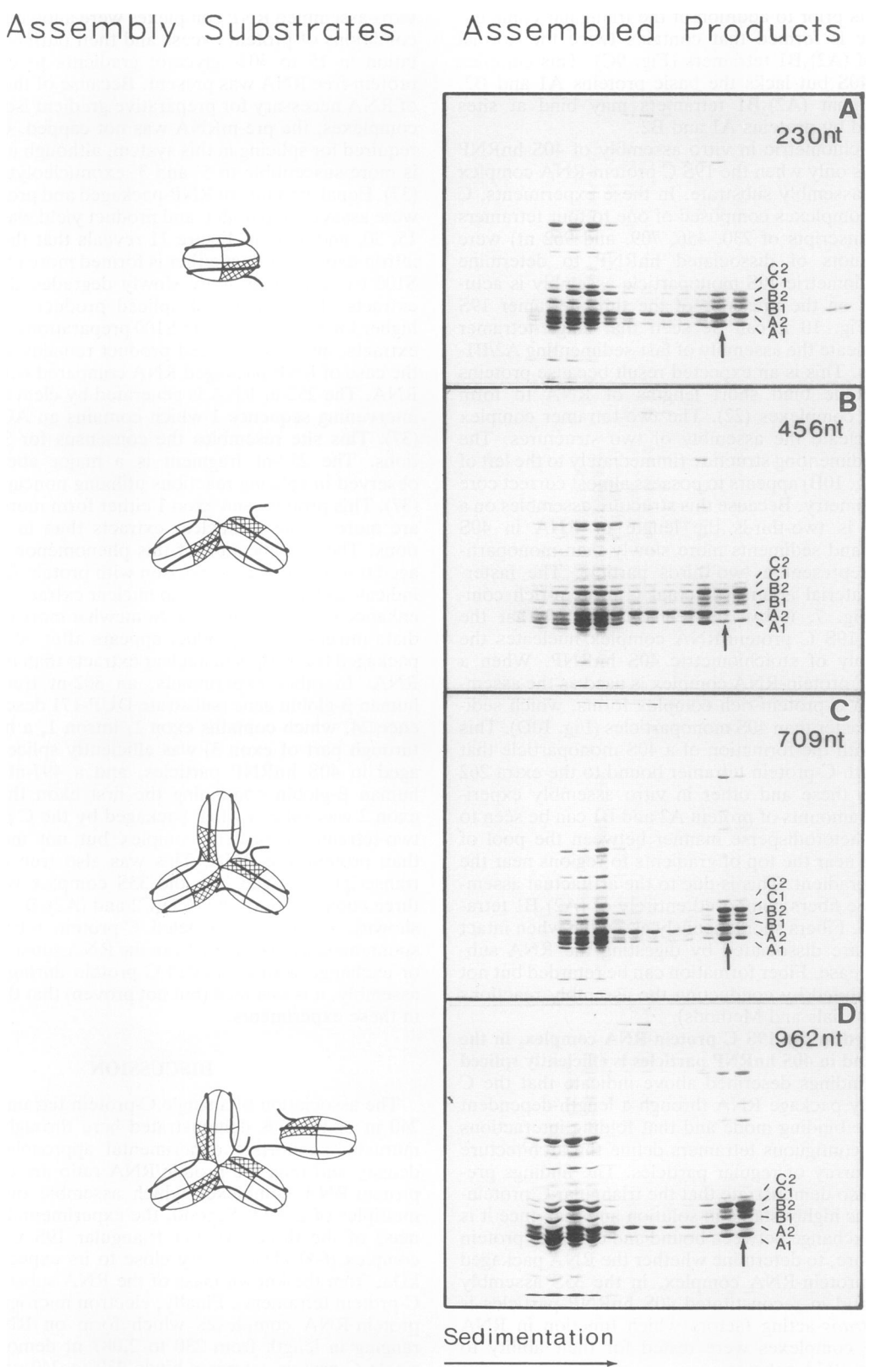

FIG. 10. Four Coomassie blue-stained gels showing the protein in successive 15 to 30% glycerol gradient fractions after four different C protein-RNA complexes were used to nucleate hnRNP assembly. In these experiments, a single preparation of hnRNP particles was dissociated via RNA digestion and divided into four equal aliquots. After addition of the four preformed C protein-RNA complexes (which assemble on transcripts of 230, 456, 709, and 962 nt) to the dissociated hnRNP preparations, the samples were subjected to sedimentation

each C-protein tetramer is associated with 232 nt of RNA. The association of six tetramers with 1,452 nt yields a single tetramer length of 242 nt, and each tetramer in the 709-nt complex could associate with 236 nt of RNA. Estimates of the single-tetramer RNA binding site size through this approach are not precise because electron micrographs of the two- and three-triangle complexes may not resolve an additional tetramer if present, and a terminal tetramer might force bind (due to tetramer-tetramer interactions) to short lengths of nonoccluded RNA. The approximate site size reported here is not likely to be in significant error since only one tetramer binds 230 nt, since two tetramers bind twice this length, and since a folding event occurs when the third tetramer binds monoparticle-length RNA to form the triangular complex. The ability of the C-protein tetramer to measure and fold the same length of RNA packaged in intact particles is not without precedent. Namely, the $(H3\text{-}H4)_2$ histone tetramer (which is considerably smaller than the C-protein tetramer) folds 146 bp of DNA (corresponding to the length wrapped on the core particle) (63), and it must bind DNA first during core particle assembly (25, 32). As described here, it is the triad structure rather than the presence of C-protein-bound RNA that functions in the correct nucleation of monoparticle assembly.

Crystallographic studies on the C-protein-like RNA binding domain of small nuclear RNP protein U1A (49), nuclear magnetic resonance studies on the 93-residue amino terminus of C protein which contains the RNA binding domain of C1 and C2 (30), and fluorescence spectroscopy of protein-RNA interactions (48) conducted on hnRNP core protein A1 (which also contains the conserved RNA binding domain) indicate that the actual occluded RNA contact length is probably between 8 and 29 nt. This suggests that much of the RNA associated with each tetramer is available for binding by the other core proteins during monoparticle assembly. Our inability to detect protected lengths of RNA upon nuclease digestion of intact 40S hnRNP particles or the various C-protein complexes described here is consistent with this interpretation. The fact that hnRNP-packaged RNA is efficiently spliced in vitro demonstrates that splicing factors can access to the RNA substrate.

The association of purified C-protein tetramers with relatively long lengths of RNA in vitro is consistent with the tetramers' relatively large Stokes radius (6.8 nm). It also suggests that the four amino-terminal RNA binding domains probably bind RNA independently to regularly spaced sites along the 230-nt RNA substrate. This would be expected if the RNA loops around each polypeptide in the tetramer or if the RNA wraps upon the tetramer's surface. The four RNA binding domains could be spatially positioned in close proximity within the tetramer and bind RNA such that loops could exist around each of the four polypeptides. The 130-residue carboxy terminus of proteins C1 and C2 contains 41 acidic and 21 basic residues and has an estimated pI of 4.1. The negatively charged character of the carboxy terminus could contribute to the presence of loosely associated loops or lengths of RNA in the C protein-RNA complex. As

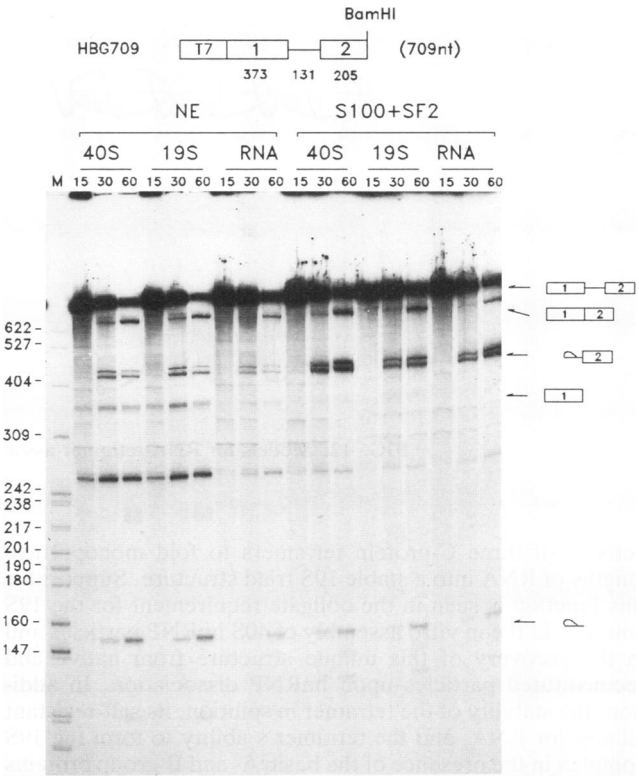

FIG. 11. In vitro splicing of packaged and nonpackaged 709-nt transcripts of the human β-globin gene. The 19S C protein-RNA complexes and the 40S hnRNP particles used in these assays were assembled on [^{32}P]UTP-labeled 709-nt transcripts under protein excess and purified through sedimentation on 15 to 30% glycerol gradients. The splicing assays were conducted with both nuclear extracts (NE) and S100 extracts containing added SF2 as described previously (44). The substrates, intermediates, and products of the reaction are indicated schematically above and at the right of the PhosphImager print. For each input substrate, aliquots were taken after 15, 30, and 60 min. The lanes under the heading RNA contained the same amount of OD_{260} material as used for the 19S and hnRNP complexes and was from the same in vitro RNA synthesis as used for RNP complex formation and was in the same buffer.

reported previously (4) and observed again in this study, the $(C1)_3C2$ tetramer does not dissociate in solution, and it is the tetramer and not soluble C1 and C2 which actually binds RNA. That this occurs in vivo is seen in the presence of the triad complex in native hnRNP particles, in the spontaneously assembly of the triad when three tetramers bind monoparticle lengths of RNA, and in the obligate requirement for this structure in 40S hnRNP assembly.

The mechanism through which approximately 700-nt increments of pre-mRNA are packaged into a repeating array of hnRNP particles in vivo is explained by the intrinsic

through 15 to 30% glycerol gradients. In these in vitro assembly studies, the four C protein-RNA substrates were adjusted by dilution to approximate equal molar amounts of RNA based on OD_{260} measurements before being added to the respective aliquots of dissociated crude hnRNP. The vertical arrows denote the 40S position of each gradient. The structures shown at the left are intended only to show tetramer number and that the assembly substrates were preformed C protein-RNA complexes. The shaded subunit of each tetramer is intended to represent protein C2. Note that the three-tetramer C protein-RNA complex was the only preformed complex to direct correct 40S hnRNP assembly. See text for further information.

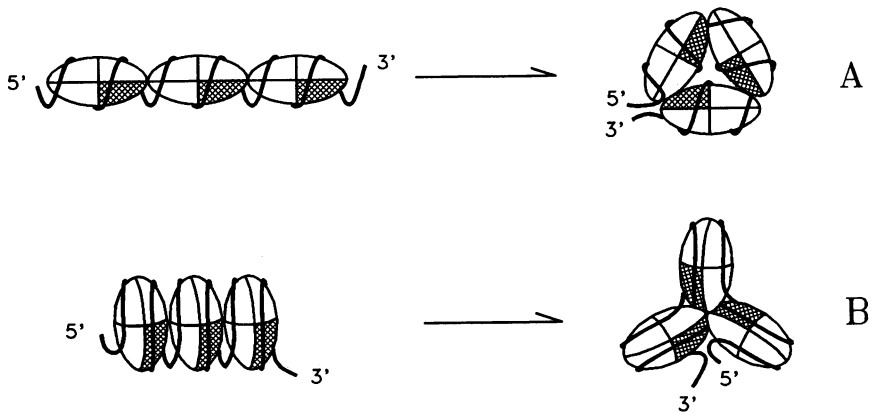

FIG. 12. Models for RNA-tetramer association and formation of the triangular complex.

activity of three C-protein tetramers to fold monoparticle lengths of RNA into a stable 19S triad structure. Support for this function is seen in the obligate requirement for the 19S complex in the in vitro assembly of 40S hnRNP particles and in the recovery of this unique structure from native and reconstituted particles upon hnRNP dissociation. In addition, the stability of the tetramer in solution, its salt-resistant affinity for RNA, and the tetramer's ability to form the 19S complex in the presence of the basic A- and B-group proteins indicate that the C-protein tetramer is well suited for this packaging function. Further evidence that the C proteins function to direct correct monoparticle assembly is seen in the ability of the 19S C protein-RNA complex to direct the binding of 3 $(A2)_3B1$ tetramers to form a 35S hnRNP intermediate complex, in contrast to the artifactual complexes which form when the A- and B-group proteins bind RNA either in the absence of or before C protein binding. The spontaneous assembly of the 35S complex which occurs when RNA is added to a mixture of $(A2)_3B1$ and $(C1)_3C2$ tetramers or when purified $(A2)_3B1$ tetramers are added to the 19S C protein-RNA complex suggests the existence of a stepwise pathway toward 40S hnRNP assembly. Namely, three C-protein tetramers bind first to form the 19S complex, and after $(A2)_3B1$ binding, proteins A1 and B2 complete monoparticle assembly. In the absence of purified native A1 and B2, the findings described here argue only that the C proteins are obligately involved in the first event of monoparticle assembly. The ability of $(A2)_3B1$ tetramers to bind at sites normally bound by $(A1)_3B2$ tetramers may explain in part why erythroleukemia cells lacking the genes for protein A1 are viable (8). It further suggests that the intranuclear concentration of these and other proteins may ultimately determine hnRNP composition, stability, and function. Our initial interest in this possibility came with the finding that only in rapidly dividing or metabolically active cells does one see equal molar ratios of A1 and A2 in 40S monoparticle preparations (13, 17, 40, 47, 51). More recently, it has been observed that the mix of A1 isoforms differs with the state of cellular differentiation (46), and it has been shown that protein A1 can influence alternative splice site selection in vitro in a concentration-dependent manner (44).

Several findings described here indicate that purified C-protein tetramers bind RNA in a self-cooperative manner. In the presence of excess of RNA (long enough for three or nine tetramers to bind), complexes possessing fewer than three or nine tetramers, respectively, are not observed either in gradients or in electron micrographs. Rather, the products are protein-free and protein-saturated RNA. When RNA substrates that are too short to support cooperative binding (230 nt) are used to nucleate 40S hnRNP assembly, nonstoichiometric A-B-rich complexes are formed. This suggests that the intrinsic RNA-binding affinity of a single C-protein tetramer is not alone sufficient to preclude the artifactual binding of the basic A- and B-group core proteins to short lengths of RNA. Finally, the folding event which leads to the formation of the triangular 19S complex could not occur if tetramer-tetramer interactions were absent. In this context, it appears that the C-protein tetramers must be free to slide or otherwise relocate from an initial contact site since the RNA substrates (whether 456, 709, 962, or 2087 nt) are always maximally bound by protein (one tetramer per 230 to 240 nt).

The tetramers' morphology in negatively stained electron micrographs seems to exclude a linear association of the four polypeptides, but it does not distinguish between a tetrahedral and a planar arrangement. When purified tetramers associate with 230-nt RNA substrates, there is an increase in observed size from about 9 to 10.2 nm. Precisely how RNA associates with each tetramer or how three RNA bound tetramers fold to form the triangular complex is unknown, although two possibilities are immediately apparent, as diagrammed in Fig. 12. The tetramers may associate along the length of the RNA in a head-to-tail fashion (the 13-amino-acid insert in C2 confers compositional polarity to the tetramer) and wrap up such that the head of the third tetramer associates with the tail of the first (as in model A). Alternatively, the tetramers could coalesce through some type of head-to-head-to-head interaction (as in model B). The complex which forms when two tetramers bind 456-nt substrates reveals an angular association of the two RNA bound tetramers and a centrally located structure which may be a protein-free tetramer. If this interpretation is correct, then the folding mechanism shown in model B is favored.

The association of three C-protein tetramers with monoparticle lengths of RNA (700 nt) must occur in such a manner that the RNA is available for the binding of three $(A2)_3B1$ and three $(A1)_3B2$ tetramers to complete 40S hnRNP assembly. The additional tetramers could associate above and below a planar 19S complex. If the 19S complex exists as an open triad complex (a three-blade propeller), then the other tetramers may interdigitate in an ordered manner. Most reports place the diameter of 30S-40S hnRNP particles

between 20 and 25 nm (39). In negatively stained preparations, the isosceles sides of the 19S C protein-RNA complex average 23 nm and the height averages 20 nm. Dimensions taken from rotary-shadowed preparations are only slightly smaller. Thus, whether the $(A2)_3B1$ and $(A1)_3B2$ tetramers associate on the faces of a planar 19S complex or interdigitate within an open structure, the similarity in dimensions between intact hnRNP particles and the 19S complex further indicates that the 19S complex functions to establish the basic topology of protein and RNA in 40S hnRNP particles. A packaging mechanism based on an open arrangement of tetramers in the 19S complex (and thus in monoparticles) is favored by the lability of the substrate to nuclease and by the fact that RNA is available for the binding of splicing factors. The inhibition of RNA splicing by antibodies against the A-, B-, and C-group proteins (19, 55) could result from the inability of *trans*-acting factors to access the RNA in antibody-bound hnRNP complexes.

If the C proteins associate with RNA in vivo as in vitro, then unless excluded by other factors, pre-mRNA molecules are likely to be bound along their entire length by these highly abundant nuclear proteins. This might be as expected since many, if not most, transcripts possess hundreds of nucleotides 5′ to the first splice site and there are many genes with intronic sequences measuring in the thousands of nucleotides. Stated differently, the short sequences which lie at intron/exon junctions must represent a very small percentage of total nuclear premRNA sequences. The finding that the C proteins are cross-linked by UV irradiation to all regions of adenovirus transcripts (corresponding to 16.2 to 91.5 map units) (62) and the recovery of oligomeric arrays of C-protein tetramers after cross-linking of splicing-competent native hnRNP complexes (33, 34) with dimethyl 3,3′-dithiobis propionimidate support this binding mechanism. The obligate C-protein-directed stepwise assembly of 40S hnRNP particles demonstrates that the core proteins interact together to package RNA into an array of regular ribonucleosomes. As shown here, this does not preclude the association of splicing factors with the packaged substrate and is consistent with the finding that a unique combination of RNA-binding proteins associate with various transcripts (108 to 547 nt) upon their addition to splicing extracts (9). For at least two reasons, it is not likely that transcripts long enough to be packaged into a single 40S hnRNP particle (about 700 nt) would possess normal core protein stoichiometry after their recovery from splicing extracts. Splicing extracts, prepared by the extraction of nuclei with 0.4 M salt (followed by low-salt dialysis and protein precipitation) may not contain the proper relative concentrations of hnRNP core proteins. More importantly, it is likely that snRNP complexes and other splicing factors immediately compete for their relevant binding sites. In this situation, information regarding the dissociation constants and the cooperativity parameters of the various factors involved in RNA processing would be necessary to predict the complement of proteins present on a given length of RNA after its addition to splicing extracts. The observation that hnRNP particles are located along the length of polyomavirus late transcripts but that intervening sequences are under represented in the pool of RNase T_1 fragments following hnRNP digestion (57) also seems consistent with a perturbation of RNA packaging near splice sites.

ACKNOWLEDGMENTS

This work was supported by NSF grant MCB 8819051 and by NIH grant GM48567 to W.M.L.

REFERENCES

1. **Ackers, G. K.** 1964. Molecular exclusion and restricted diffusion processes in molecular sieve chromatography. Biochemistry **3**:723–730.
2. **Amero, S. A., G. Raychaudhuri, C. L. Cass, W. J. vanVenrooij, W. J. Habets, A. R. Krainer, and A. L. Beyer.** 1992. Independent deposition of heterogeneous nuclear ribonucleoproteins and small nuclear ribonucleoprotein particles at sites of transcription. Proc. Natl. Acad. Sci. USA **89**:8409–8413.
3. **Ast, G., D. Goldblatt, D. Offen, J. Sperling, and R. Sperling.** 1991. A novel splicing factor is an integral component of 200S large nuclear ribonucleoprotein (lnRNP) particles. EMBO J. **10**:425–432.
4. **Barnett, S. F., D. L. Friedman, and W. M. LeStourgeon.** 1989. The C proteins of HeLa 40S nuclear ribonucleoprotein particles exist as anisotropic tetramers of $(C1)_3C2$. Mol. Cell. Biol. **9**:492–498.
5. **Barnett, S. F., W. M. LeStourgeon, and D. L. Friedman.** 1988. Rapid purification of native C protein from nuclear ribonucleoprotein particles. J. Biochem. Biophys. Methods **16**:87–98.
6. **Barnett, S. F., S. J. Northington, and W. M. LeStourgeon.** 1990. Isolation and in vitro assembly of nuclear ribonucleoprotein particles and purification of core particle proteins. Methods Enzymol. **181**:293–307.
7. **Barnett, S. F., T. A. Theiry, and W. M. LeStourgeon.** 1991. The core proteins A2 and B1 exist as $(A2)_3B1$ tetramers in 40S nuclear ribonucleoprotein particles. Mol. Cell. Biol. **11**:864–871.
8. **Ben-David, Y., M. R. Bani, B. Chabot, A. D. Koven, and A. Bernstein.** 1992. Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth. Mol. Cell. Biol. **12**:4449–4455.
9. **Bennett, M., S. Pinol-Roma, D. Staknis, G. Dreyfuss, and R. Reed.** 1992. Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly in vitro. Mol. Cell. Biol. **12**:3165–3175.
10. **Beyer, A. L., M. E. Christensen, B. W. Walker, and W. M. LeStourgeon.** 1977. Identification and characterization of the packaging proteins of core 40S hnRNP particles. Cell **11**:127–138.
11. **Brunel, C., and M. Lelay.** 1979. Two-dimensional analysis of proteins associated with heterogenous nuclear RNA in various animal cell lines. Eur. J. Biochem. **99**:273–283.
12. **Burd, C. G., M. S. Swanson, M. Gorlach, and G. Dreyfuss.** 1989. Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. Proc. Natl. Acad. Sci. USA **86**:9788–9792.
13. **Buvoli, M., G. Biamonti, P. Tsoulfas, M. T. Bassi, A. Ghetti, S. Riva, and C. Morandi.** 1988. cDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms. Nucleic Acids Res. **16**:3751–3770.
14. **Buvoli, M., F. Cobianchi, M. Bestagno, M. T. Bassi, G. Biamonti, and S. Riva.** 1990. A second A1-type protein is encoded by the human hnRNP A1 gene. Mol. Biol. Rep. **14**:83–84.
15. **Buvoli, M., F. Cobianchi, M. G. Bestagno, A. Mangiarotti, M. T. Bassi, G. Biamonti, and S. Riva.** 1990. Alternative splicing in the human gene for the core protein A1 generates another hnRNP protein. EMBO J. **9**:1229–1235.
16. **Cantor, C. R., and P. R. Schimmel.** 1980. Biophysical chemistry. Part II. Techniques for the study of biological structures and function. W. H. Freeman & Co., New York.
17. **Celis, J. E., R. Bravo, H. P. Arenstorf, and W. M. LeStourgeon.** 1986. Identification of proliferation-sensitive human proteins amongst components of the 40 S hnRNP particles. Identity of hnRNP core proteins in the HeLa protein catalogue. FEBS Lett. **194**:101–109.
18. **Choi, Y. D., and G. Dreyfuss.** 1984. Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly. Proc. Natl. Acad. Sci. USA **81**:7471–7475.
19. **Choi, Y. D., P. J. Grabowski, P. A. Sharp, and G. Dreyfuss.**

1986. Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. Science **231**:1534–1539.
20. **Cobianchi, F., G. Biamonti, M. T. Bassi, M. Buvoli, and S. Riva.** 1990. The protein A1 of mammalian hnRNP complex: a new insight into protein complexity and gene family structure, p. 561–577. *In* P. R. Strauss and S. H. Wilson (ed.), The eukaryotic nucleus: molecular biochemistry and macromolecular assemblies, vol. 2. The Telford Press, Caldwell, N.J.
21. **Cobianchi, F., D. N. SenGupta, B. Z. Zmudzta, and S. H. Wilson.** 1986. Structure of rodent helix-destabilizing protein revealed by cDNA cloning. J. Biol. Chem. **261**:3536–3543.
22. **Conway, G., J. Wooley, T. Bibring, and W. M. LeStourgeon.** 1988. Ribonucleoproteins package 700 nucleotides of pre-mRNA into a repeating array of regular particles. Mol. Cell. Biol. **8**:2884–2895.
23. **deHaen, C.** 1987. Molecular weight standards for calibration of gel filtration and sodium dodecyl sulfate-polyacrylamide gel electrophoresis: ferritin and apoterritin. Anal. Biochem. **166**: 235–245.
24. **Dominski, Z., and R. Kole.** 1991. Selection of splice sites in pre-mRNAs with short internal exons. Mol. Cell. Biol. **11**:6075–6083.
25. **Dong, F., and K. E. vanHolde.** 1991. Nucleosome positioning is determined by the (H3-H4)2 tetramer. Proc. Natl. Acad. Sci. USA **88**:10596–10600.
26. **Dreyfuss, G.** 1986. Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. Annu. Rev. Cell Biol. **2**:457–495.
27. **Economidis, I. V., and T. Pederson.** 1982. Assembly of nuclear ribonucleoprotein particles during in vitro transcription. Proc. Natl. Acad. Sci. USA **79**:1469–1473.
28. **Economidis, I. V., and T. Pederson.** 1983. Structure of nuclear ribonucleoprotein: heterogeneous nuclear RNA is complexed with a major sextet of proteins in vivo. Proc. Natl. Acad. Sci. USA **80**:1599–1602.
29. **Faiferman, I., M. G. Hamilton, and A. O. Pogo.** 1970. Nucleoplasmic ribonucleoprotein particles of rat liver. I. Selective degradation by nuclear nucleases. Biochim. Biophys. Acta **204**:550–563.
30. **Gorlach, M., M. Wittekind, R. A. Beckman, L. Mueller, and G. Dreyfuss.** 1992. Interaction of the RNA-binding domain of the hnRNP C proteins with RNA. EMBO J. **11**:3289–3295.
31. **Hahn, M., and W. Baumeister.** 1974. High resolution negative staining of ferritin molecules on vermiculite single crystal supports. Biochim. Biophys. Acta **371**:267–274.
32. **Hansen, J. C., K. E. vanHolde, and D. Lohr.** 1991. The mechanism of nucleosome assembly onto oligomers of the sea urchin 5S DNA positioning sequence. J. Biol. Chem. **266**:4276–4282.
33. **Harris, S. G., S. O. Hoch, and H. C. Smith.** 1988. Chemical crosslinking of Sm and RNP antigenic proteins. Biochemistry **27**:4595–4600.
34. **Harris, S. G., T. E. Martin, and H. C. Smith.** 1988. Reversible chemical cross-linking and ribonuclease digestion analysis of the organization of proteins in ribonucleoprotein particles. Mol. Cell. Biochem. **84**:17–28.
35. **Karn, J., G. Vidali, L. Boffa, and V. Allfrey.** 1977. Characterization of the non-histone nuclear proteins associated with rapidly labeled heterogeneous nuclear RNA. J. Biol. Chem. **252**:969–976.
36. **Kish, V. M., and T. Pederson.** 1978. Isolation and characterization of ribonucleoprotein particles containing heterogeneous nuclear RNA. Methods Cell Biol. **17**:377–399.
37. **Krainer, A. R., T. Maniatis, B. Ruskin, and M. R. Green.** 1984. Normal and mutant human beta-globin pre-mRNAs are faithfully and efficiently spliced in vitro. Cell **36**:993–1005.
38. **Leser, G. P., and T. E. Martin.** 1987. Changes in heterogeneous nuclear RNP core polypeptide complements during the cell cycle. J. Cell Biol. **105**:2083–2094.
39. **LeStourgeon, W. M., S. F. Barnett, and S. J. Northington.** 1990. Tetramers of the core proteins of 40S nuclear ribonucleoprotein particles assemble to package nascent transcripts into a repeating array of regular particles, p. 477–502. *In* P. R. Strauss and S. H. Wilson (ed.), The Eukaryotic nucleus: molecular biochemistry and macromolecular assemblies, vol. 2. The Telford Press, Caldwell, N.J.
40. **LeStourgeon, W. M., A. L. Beyer, M. E. Christensen, B. W. Walker, S. M. Poupore, and L. P. Daniels.** 1977. The packaging proteins of core hnRNP particles and the maintenance of proliferative cell states. Cold Spring Harbor Symp. Quant. Biol. **42**:885–898.
41. **Lothstein, L., H. P. Arenstorf, S. Chung, B. W. Walker, J. C. Wooley, and W. M. LeStourgeon.** 1985. General organization of protein in HeLa 40S nuclear ribonucleoprotein particles. J. Cell Biol. **100**:1570–1581.
42. **Lukanidin, E. M., E. S. Zalmanson, L. Komaromi, O. P. Samarina, and G. P. Georgiev.** 1972. Structure and function of informofer. Nature (London) New Biol. **238**:193–196.
43. **Malcolm, D. B., and J. Sommerville.** 1977. The structure of nuclear ribonucleoprotein of amphibian oocytes. J. Cell Sci. **24**:143–165.
44. **Mayeda, A., and A. R. Krainer.** 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell **68**:365–375.
45. **Merrill, B. M., S. F. Barnett, W. M. LeStourgeon, and K. R. Williams.** 1989. Primary structure differences between proteins C1 and C2 of HeLa 40S nuclear ribonucleoprotein particles. Nucleic Acids Res. **17**:8441–8449.
46. **Minoo, P., T. E. Martin, and R. M. Riehl.** 1991. Nucleic acid binding characteristics of group A/B hnRNP proteins. Biochem. Biophys. Res. Commun. **176**:747–755.
47. **Minoo, P., W. Sullivan, L. R. Solomon, T. E. Martin, D. O. Toft, and R. E. Scott.** 1989. Loss of proliferative potential during terminal differentiation coincides with the decreased abundance of a subset of heterogeneous ribonuclear proteins. J. Cell Biol. **109**:1937–1946.
48. **Nadler, S. G., B. M. Merrill, W. J. Roberts, K. M. Keating, M. J. Lisbin, S. F. Barnett, S. H. Wilson, and K. R. Williams.** 1991. Interactions of the A1 heterogeneous nuclear ribonucleoprotein and its proteolytic derivative, UP1, with RNA and DNA: evidence for multiple RNA binding domains and salt-dependent binding mode transitions. Biochemistry **30**:2968–2976.
49. **Nagai, K., C. Oubridge, T. H. Jessen, J. Li, and P. R. Evans.** 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. Nature (London) **348**:515–520.
50. **Nakagawa, T. Y., M. S. Swanson, B. J. Wold, and G. Dreyfuss.** 1986. Molecular cloning of cDNA for the nuclear ribonucleoprotein particle C proteins: a conserved gene family. Proc. Natl. Acad. Sci. USA **83**:2007–2011.
51. **Planck, S. R., M. D. Listerud, and S. D. Buckley.** 1988. Modulation of hnRNP A1 protein gene expression by epidermal growth factor in Rat-1 cells. Nucleic Acids Res. **16**:11663–11673.
52. **Samarina, O., and A. Krichevskaya.** 1981. Nuclear 30S RNP particles. Cell Nucleus **9**:1–48.
53. **Samarina, O. P., E. M. Lukanidin, J. Molnar, and G. P. Georgiev.** 1968. Structural organization of nuclear complexes containing DNA-like RNA. J. Mol. Biol. **33**:251–263.
54. **Schenkel, J., C. E. Sekeris, A. Alonso, and E. K. Bautz.** 1988. RNA-binding properties of hnRNP proteins. Eur. J. Biochem. **171**:565–569.
55. **Sierakowska, H., W. Szer, P. J. Furdon, and R. Kole.** 1986. Antibodies to hnRNP core proteins inhibit in vitro splicing of human β-globin pre-mRNA. Nucleic Acids Res. **14**:4475–4489.
56. **Spann, P., M. Feinerman, J. Sperling, and R. Sperling.** 1989. Isolation and visualization of large compact ribonucleoprotein particles of specific nuclear RNAs. Proc. Natl. Acad. Sci. USA **86**:466–470.
57. **Steitz, J. A., and R. Kamen.** 1981. Arrangement of 30S heterogeneous nuclear ribonucleoprotein on polyoma virus late nuclear transcripts. Mol. Cell. Biol. **1**:21–34.
58. **Swanson, M. S., and G. Dreyfuss.** 1988. RNA binding specificity of hnRNP proteins: a subset bind to the 3′ end of introns. EMBO J. **7**:3519–3529.
59. **Swanson, M. S., and G. Dreyfuss.** 1988. Classification and

purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. Mol. Cell. Biol. **8**:2237–2241.

60. **Swanson, M. S., T. Y. Nakagawa, K. LeVan, and G. Dreyfuss.** 1987. Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-RNA binding proteins. Mol. Cell. Biol. **7**:1731–1739.
61. **Tsanev, R. G., and L. P. Djondurov.** 1982. Ultrastructure of free ribonucleoprotein complexes in spread mammalian nuclei. J. Cell Biol. **94**:662–666.
62. **van Eekelen, C., R. Ohlsson, L. Philipson, E. Mariman, R. van Beek, and W. van Venrooij.** 1982. Sequence dependent interaction of hnRNP proteins with late adenoviral transcripts. Nucleic Acids Res. **10**:7115–7131.
63. **van Holde, K. E.** 1989. The proteins of chromatin. I. Histones, p. 69–180. *In* A. Rich (ed.), Chromatin. Springer-Verlag, New York.
64. **Wilk, H. E., G. Angeli, and K. P. Schafer.** 1983. In vitro reconstitution of 35S ribonucleoprotein complexes. Biochemistry **22**:4592–4600.
65. **Wilk, H. E., H. Werr, D. Friedrich, H. H. Kiltz, and K. P. Schaefer.** 1985. The core proteins of 35S hnRNP complexes: characterization of 9 different species. Eur. J. Biochm. **146**:71–81.
66. **Willison, J. H., and A. J. Rowe.** 1980. Shadowing, p. 59–93. *In* A. M. Glauert (ed.), Practical methods in electron microscopy, vol. 8. North-Holland Publishing Co., New York.