Structure of Nuclear Ribonucleoprotein:
Identification of Proteins in Contact with Poly(A)⁺
Heterogeneous Nuclear RNA in Living HeLa Cells

SANDRA MAYRAND, BUDDY SETYONO, JAY R. GREENBERG, and THORU PEDERSON
Cell Biology Group, Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts 01545

ABSTRACT The processing of heterogeneous nuclear RNA into messenger RNA takes place in special nuclear ribonucleoprotein particles known as hnRNP. We report here the identification of proteins tightly complexed with poly(A)⁺ hnRNA in intact HeLa cells, as revealed by a novel in situ RNA-protein cross-linking technique. The set of cross-linked proteins includes the A, B, and C "core" hnRNP proteins, as well as the >42,000 mol wt species previously identified in noncross-linked hnRNP. These proteins are shown to be cross-linked by virtue of remaining bound to the poly(A)⁺ hnRNA in the presence of 0.5% sodium dodecyl sulfate, 0.5 M NaCl, and 60% formamide, during subsequent oligo(dT)-cellulose chromatography, and in isopycnic banding in Cs₂SO₄ density gradients. These results establish that poly(A)⁺ hnRNA is in direct contact with a moderately complex set of nuclear proteins in vivo. This not only eliminates earlier models of hnRNP structure that were based upon the concept of a single protein component but also suggests that these proteins actively participate in modulating hnRNA structure and processing in the cell.

Eukaryotic messenger RNA is derived from precursors collectively known as heterogeneous nuclear RNA (hnRNA). The conversion of hnRNA to mRNA involves a number of well-defined steps, including the addition of poly(A) and 5'-caps, and the removal of intervening DNA sequence transcripts through a series of coupled cleavage-ligation reactions known as splicing. These mRNA processing steps take place while hnRNA is tightly complexed with nuclear proteins in ribonucleoprotein particles known as hnRNP (reviewed in reference 1). hnRNP particles were first seen as nascent RNP fibrils on the lateral loops of amphibian oocyte lambrush chromosomes (2, 3) and later by electron microscopy in a variety of eukaryotic cells (4–14). hnRNP particles have been isolated (15, 16), and their constituent RNA and proteins have now been characterized in some detail (15–32).

It is likely that the structure of hnRNP is functionally related to mRNA processing (1). We have been studying the nucleoprotein structure of defined RNA sequences in hnRNP using nucleases and other probes. For example, the use of nucleases specific for either single- or double-stranded RNA has revealed that intramolecular double-stranded regions of hnRNP are essentially free of associated protein, in contrast to single-stranded hnRNA which is extensively complexed with protein (27). Recently, we have found that messenger RNA-homologous sequences in hnRNP are extensively complexed with protein as revealed by nuclease protection experiments (31, 33). The recent demonstration that hnRNP from mouse Friend erythroleukemia cells contains the pre-spliced 15S precursor of β-globin mRNA (30) supports the idea that hnRNP particles are the sites of mRNA processing in the cell.

In this investigation we used a novel methodology (32) to address another important aspect of hnRNP structure: the number of different protein species that are complexed with hnRNA. Because the number of different proteins in hnRNP is fairly large (16–19, 24, 25, 30), it becomes important to identify which of these make direct contact with the hnRNA. To examine this issue, we exploited RNA-protein cross-linking by 254-nm light. We have showed previously that 254-nm irradiation of intact HeLa cells leads to the photochemical formation of covalent cross-links between hnRNA and hnRNP proteins (32). This not only conclusively establishes that hnRNA is indeed complexed with nuclear proteins in vivo but also provides a new approach for differentiating hnRNP particle proteins that are in direct contact with the hnRNA from those that are not. Only the former will remain bound to the RNA when cross-linked hnRNP is exposed to protein-dissociating agents such as detergent or high concentrations of salt. In this investigation we used this powerful approach to identify
hnRNP particle proteins that are in direct contact with HeLa cell poly(A)-containing hnRNA in vivo.

MATERIALS AND METHODS

Cells, Labeling, and Cross-linking

HeLa cells were grown in suspension culture as previously detailed (34, 35). Cultures in medium containing one-half the usual l-methionine concentration were grown for 40 h with 2.5 µCi of [35S]methionine/ml (1,000–1,200 Ci/m mole, New England Nuclear, Boston, Mass.). Cells were harvested, resuspended at 4 x 10^6 cells/ml in fresh, prewarmed medium, incubated for 15 min with 0.04 µg antibiotic, followed by 20 min with 5-10 µCi of [3H]uridine/ml to label hnRNA. The cells were then washed, resuspended at 10^7 cells/ml in 0.1 M NaCl. 0.01 M Tris-HCl, pH 7.2, and irradiated, with stirring, in Petri dishes at 4°C, all as previously described in detail (32).

Cell Fractionation and hnRNP Isolation

After irradiation, the cells were centrifuged and fractionated by Dounce homogenization in reticulocyte standard buffer (RBSB: 0.01 M NaCl, 1.5 mM MgCl2, 0.01 M Tris-HCl, pH 7.2), and the hnRNP particles were isolated from nuclear envelope as previously described (16). Aliquots were removed at the 4,500 g 0.30% sucrose interface step and monitored for hnRNA:protein croslinking by the phenol:water phase partitioning assay previously described (32). After separation on 15-30% sucrose gradients as previously described (16), the purified hnRNP sedimenting at 40-250S was pelleted at 45,000 rpm in a Spinco 60Ti rotor (Beckman Instruments, Inc., Spinco Div., Palo Alto, Calif.) for 12-14 h (4°C).

Oligo(dT)-cellulose Chromatography, Electrophoresis, and Fluorography

The pellet of hnRNP was resuspended in 2 ml of 1 mM EDTA, 50 mM Tris-HCl, pH 7.6, 0.5% sodium dodecyl sulfate (SDS). When dissolved (5-10 min at 20-23°C with occasional agitation), the hnRNP were made 0.5 M in NaCl. Deionized formamide was added to a final concentration of 60%, and the samples were heated at 60°C for 3 min. This step, referred to hereafter as "SDS-high salt-formamide dissociation," is essential for the complete separation of noncross-linked proteins from the hnRNA which, in turn, is required to permit a clear separation of noncross-linked proteins by 254-nm light has been extensively described in 10% polyacrylamide slab gels and fluorographed, as previously detailed (36). After application of the sample, the column was washed with 5 ml of elution buffer (1 mM EDTA, 14 mM β-mercaptoethanol, 0.01 M Tris-HCl, pH 7.6, 0.5% SDS). After application of the sample, the column was washed (flow rate = 0.5 ml/min) with 10 volumes of binding buffer, after which additional washes did not release further 35S radioactivity. The column was then washed with 5 ml of elution buffer (1 mM EDTA, 0.01 M Tris-HCl, pH 7.2, 0.5% SDS) to release the poly(A)+ hnRNA and any cross-linked protein. The eluted material was then made 0.3 M in NaCl and rechromatographed and eluted once again. The oligo(dT)-bound and nonbound column fractions were ethanol-precipitated and centrifuged as described above, dissolved in 0.15 ml of digestion buffer (10% glycerol, 1 mM dithiothreitol, 0.5% Sarkosyl, 0.0625 M Tris-HCl, pH 6.5), and digested with 12.5 U of ribonuclease T1 for 60 min at 37°C. The samples were then made 2% in SDS and 0.01% in bromophenol blue and electrophoresed in 10% polyacrylamide slab gels and fluorographed, as previously detailed (36).

RESULTS AND DISCUSSION

Strategy

The photocatalyzed cross-linking of RNA to tightly associated proteins by 254-nm light has been extensively described in studies on RNA and proteins in solution (37), as well as in native ribonucleoprotein complexes such as ribosomes (38), transfer RNA:aminoacyl-tRNA synthetases (39), viruses (40), and eukaryotic messenger RNP particles (41, 42). Unlike chemical reagents, the cross-linking radius of 254-nm light is essentially "zero-Å", so that only proteins in direct contact with RNA are candidates for cross-linking (e.g. 38). Moreover, by the use of suitable UV doses, the technique can be extended to probe ribonucleoprotein structure in intact cells, as we have recently done for hnRNP particles in HeLa cells (32). In this study we wished to determine which of the different HeLa hnRNP protein species, defined by SDS polyacrylamide gel electrophoresis (16, 24, 25), can be cross-linked to the hnRNA in vivo. This required a method for separating noncross-linked proteins from those that are cross-linked to the hnRNA. We have found that this can be done expeditiously by a combination of 0.5% SDS, 0.5 M NaCl, and 60%, formamide, which completely strips noncross-linked protein, followed by oligo(dT)-cellulose chromatography, which retrieves poly(A)-containing hnRNA and its cross-linked proteins. A minor disadvantage of this approach is that it permits direct analysis only of that fraction of the hnRNA that is polyadenylated. The compelling advantage, as will be shown, is that all traces of noncross-linked protein can be removed from the hnRNP before the specifically cross-linked proteins are analyzed.

Analysis of Proteins Cross-linked to Poly(A)+ hnRNA

HeLa cells were labeled with [35S]methionine and [3H]uridine and irradiated in vivo (32). hnRNP particles were purified, stripped in SDS/high salt/formamide, and prepared for oligo(dT)-cellulose chromatography, as detailed in Materials and Methods. The 35S-labeled proteins in the total, oligo(dT)-nonbound and oligo(dT)-bound fractions were analyzed by electrophoresis and fluorography. Fig. 1 illustrates the results typical of several independent experiments. In lane b are displayed the total proteins of hnRNP from nonirradiated HeLa cells, as a control. As previously reported (16, 24, 25), HeLa cell hnRNP proteins are a moderately complex set of polypeptides ranging from 33,000 to 125,000 daltons. The
33,000–42,000 mol wt components are known as the A (32,000–34,000), B (36,000–37,000) and C (42,000–44,000) "core" proteins (25) and are major components of a basic 30–40S hnRNP subparticle that can be generated from intact hnRNP by endogenous or exogenous nuclease digestion (16, 24, 25). The almost complete absence of proteins <33,000 is noteworthy and attests to the absence from the hnRNP particle preparation of both chromatin and nucleolar preribosomal RNP particles (43, 44). Lanes c and d in Fig. 1 show the proteins present in the oligo(dT)-nonbound and -bound fractions from nonirradiated cells. Analysis of the distribution of 3H radioactivity revealed that 6% of the hnRNA bound to oligo(dT) in this experiment. (In other experiments, values of 6–15% were obtained. These values are somewhat less than the poly(A) fraction observed with HeLa hnRNP isolated directly from nuclei by phenol extraction (15–20%), which probably reflects the partial loss of poly(A)-distal regions due to occasional nicking of the molecules during the lengthy hnRNP isolation and subsequent preparation for analysis.) Despite the binding of poly(A)+ hnRNA to the column, it can be seen in lane d that the bound fraction is completely devoid of 35S-labeled proteins. This shows that the SDS/high salt/formamide treatment (see Materials and Methods) is effective in stripping noncross-linked protein from hnRNP, thereby providing a low background against which to assess cross-linking. We wish to emphasize the importance of formamide. If noncross-linked hnRNP is treated with 0.5 M NaCl and 0.5% SDS alone and then run on oligo(dT)-cellulose, we find that some protein still remains associated with the poly(A)+ hnRNA, constituting a sufficient background to complicate the analysis of cross-linking. When formamide was used, no significant amounts of 35S-labeled proteins were observed in gels of the oligo(dT)-bound fractions of nonirradiated hnRNP in any of several experiments.

Lanes e, f, and g of Fig. 1 show the total, oligo(dT)-nonbound and -bound fractions from cross-linked hnRNP. The total proteins (lane e) are indistinguishable from those present in noncross-linked particles (lane b). This demonstrates that the in vivo irradiation does not somehow "capture" in hnRNP additional species of nuclear proteins beyond those which are normally seen in the particles. More importantly, however, is the fact that a substantial amount of protein is now recovered in the oligo(dT)-bound fraction (lane g). Rather than consisting of a small, selected subset of the total hnRNP proteins, these appear to represent the full analytical complexity of the particle proteins (compare g with e). Of particular interest is the presence in the cross-linked hnRNP of the A, B, and C core proteins (33,000–44,000 daltons) in apparently normal stoichiometry. Densitometry (not shown) reveals that the relative concentrations of individual proteins in the oligo(dT)-bound fraction are very similar to those in total hnRNP. A possible exception is a doublet at ~92,000 daltons, which was slightly but reproducibly depleted in the oligo(dT)-bound fraction relative to the other proteins. These proteins may be restricted to the poly(A)+ hnRNP or, alternatively, may be organized in the poly(A)+ hnRNP in such a way that their cross-linking efficiency is very low.

The results illustrated in Fig. 1, lanes e, f, and g, were obtained with hnRNP isolated from cells irradiated at a dose of 6 × 10⁸ ergs/mm². This cross-links ~50% of the hnRNA molecules to protein, as defined by the conversion of the hnRNA from aqueous to phenol solubility (see reference 32 for details). When cells were irradiated at a UV dose 60 times higher (3.6 × 10⁸ ergs/mm²), which cross-links 100% of the hnRNA to protein (32), the pattern of proteins cross-linked to poly(A)+ hnRNA was identical to that in Fig. 1, lane g (data not shown), although the total amount of cross-linked protein was somewhat higher per unit of RNA, as might be expected due to the higher dose. From analyses of trichloroacetic acid-precipitable [35S]methionine and [3H]uridine radioactivity, it can be calculated that the mass of protein remaining cross-linked to poly(A)+ hnRNA after the SDS/high salt/formamide stripping and oligo(dT)-cellulose chromatography steps represents, on a per unit RNA basis, ~10% of the initial hnRNP particle protein content. This is confirmed by isopycnic banding of the cross-linked, oligo(dT)-bound hnRNA:protein complexes in Cs₂SO₄. As shown in Fig. 2, the majority of these cross-linked complexes band at an average density of 1.60 g/cm³. In these gradients, completely protein-free hnRNA is 1.66 g/cm³, and intact, cross-linked hnRNP (not stripped with SDS/high salt/formamide) is 1.31 g/cm³ (32). By use of the equation of Perry and Kelley (45) for the relationship between observed buoyant density and RNA:protein ratio, and taking 1.66 and 1.18 as the densities of RNA and protein in Cs₂SO₄, respectively (32, 46), the peak at 1.60 g/cm³ in Fig. 2 is calculated to contain 91% RNA and 9% protein. Because intact hnRNP is ~80% protein (16), it follows that the average protein content of the 1.60 g/cm³ complexes (9%) is ~11% that of the total hnRNP particle protein. This is in excellent agreement with the estimate (10%) based on 35S:3H ratios in the oligo(dT)-bound fractions.

![Figure 2: Cs₂SO₄ isopycnic banding of cross-linked poly(A)+ hnRNP complexes after SDS/high salt/formamide stripping and elution from oligo(dT)-cellulose. The eluted hnRNP, in 0.01 M NaCl, 0.01 M Tris-HCl, pH 7.2, 0.5% SDS, was made 0.1 M in NaCl and precipitated in 2.5 vol of ethanol (−20°C). The hnRNP was redissolved in 0.3 ml of RSB containing 0.5 M NaCl and banded in a Cs₂SO₄ gradient (initial density 1.50 g/cm³). Centrifugation was in a Beckman Spinco SW50.1 rotor at 34,000 rpm for 60 h (20°C). Density and trichloroacetic acid-precipitable [3H] radioactivity of the gradient fractions were determined as previously described (27, 36). The arrow at 1.66 g/cm³ indicates the position of naked hnRNA run in a parallel gradient.](https://example.com/figure2.png)
The fact that the cross-linked complexes band at 1.60 g/cm³ also provides additional, independent evidence that the proteins analyzed by electrophoresis (Fig. 1, lane g) were indeed physically linked to the hnRNA.

We attempted to determine whether there is an upper limit to the fraction of hnRNP particle protein (mass) that can be cross-linked to hnRNA in vivo, by increasing the dosage of UV irradiation further. At 7.2 × 10⁵ ergs/mm², the amount of protein cross-linked to poly(A)⁺⁻ hnRNA was increased slightly, but electrophoresis revealed no new or missing protein species as compared to those seen in Fig. 1, lane g (data not shown). At still higher UV doses, such as 1.5–3.0 × 10⁵ ergs/mm², we find that cell fractionation and hnRNP isolation become problematic (S. Mayrand, unpublished results), and we have therefore not explored these high doses further. We conclude that there is at least an experimental limit on the fraction of hnRNP protein mass than can be cross-linked to hnRNA in vivo, ~10% of the total. This may be due to (a) the optical density of the intact cell for 254-nm light or (b) the topography of hnRNP proteins on different hnRNA molecules. Possibility (a) is supported by the fact that the UV dose vs. cross-linking profile for hnRNP irradiated as isolated particles is shifted to tenfold lower doses relative to the cross-linking profile for hnRNP in intact cells (32). Of course, the two possibilities stated above are not mutually exclusive.

hnRNP Structure and Function

We have used 254-nm light-catalyzed RNA-protein cross-linking in vivo to establish that hnRNP particles exist in the cell before fractionation (32), and the present results confirm this important point. In addition, we have now shown that the poly(A)⁺⁻ fraction of HeLa cell hnRNA is complexed with a set of proteins that are indistinguishable from those of total hnRNP (Fig. 1), at least within the limits of resolution of one-dimensional gel electrophoresis. These include not only the A, B, and C groups of “core” proteins (25) but also the >42,000 mol wt species previously identified in total, noncross-linked hnRNP (16). The initial studies of hnRNP by Georgiev et al. (reviewed in reference 16) suggested that these particles contained only a single protein component of ~40,000 mol wt. Although the notion of a single hnRNP protein might have been attractive, more detailed analyses (16, 24, 25) have now convincingly established that the hnRNP proteins are in fact rather complex. However, it still remained possible that some of these additional proteins are derived from nonRNP nuclear structures copurifying with hnRNP or, alternatively, represent soluble proteins that bind adventitiously to the hnRNA during isolation, although reconstruction experiments argue against this latter possibility (16, 19, 30). The present results establish unequivocally that the complete set of protein species previously identified in hnRNP are actually in direct contact with poly(A)⁺⁻ hnRNA in vivo before cell fractionation. Thus, all of these proteins can now be regarded as authentic components of hnRNP particles and, as such, are candidates for roles in modulating hnRNA structure and facilitating mRNA processing (splicing). Whether they interact with hnRNA as individual protein molecules or, perhaps more likely, as parts of multimeric homotypic or heterotypic particles awaits analyses with protein:protein cross-linking reagents.

The majority of the hnRNP analyzed in these experiments does not contain poly(A) of sufficient length to bind oligo(dT)-cellulose. The oligo(dT) nonbound fraction of cross-linked hnRNP thus consists of the hnRNP proteins bound to, or stripped from, poly(A)⁺⁻ hnRNA, plus those stripped from the poly(A)⁺⁻ hnRNA. Therefore, our experiments do not directly identify the hnRNP proteins that are cross-linked to poly(A)⁺⁻ hnRNA. However, because ~85% of the hnRNP-RNA fractionates as poly(A)⁺⁻, it follows that the proteins observed in the nonbound fraction are largely derived from poly(A)⁺⁻ hnRNP. These proteins are identical to those cross-linked to poly(A)⁺⁻ hnRNA (Fig. 1, lanes f vs. g). The fact that a large fraction of the hnRNA in mammalian cells lacks poly(A) (47, 48) has never been satisfactorily explained. About 10% of the mRNA sequences in HeLa cell polyribosomes lack poly(A) (49), and some of the poly(A)⁺⁻ hnRNA might be precursors to this poly(A)⁻⁻ mRNA. A second possibility is that poly(A)⁺⁻ hnRNA simply reflects defective polyadenylation of transcripts which, in other cases, become poly(A)⁺⁻ hnRNA. Studies of the primary transcripts of SV 40, adenovirus and β-globin genes reveal that transcription frequently proceeds well beyond the poly(A) site (50–52), and it is possible that many of these molecules remain poly(A)⁺⁻ thereafter. Irrespective of the actual basis for the existence of a large fraction of nonpolyadenylated hnRNA, our results indicate that the RNP forms of poly(A)⁺⁻ and poly(A)⁻⁻ hnRNA contain an essentially identical set of hnRNP proteins. We therefore consider it unlikely that the major hnRNP proteins play a role in the early posttranscriptional decision (53) as to whether or not a given transcript will be polyadenylated. Of course, this does not exclude such a role for minor hnRNP proteins which are not resolved in our analyses.

A final issue is the relationship of hnRNP to polyribosomal messenger RNP. We previously showed that the proteins of HeLa cell hnRNP are a more complex set than those associated with polyribosomal mRNP (20). The present cross-linking results, taken together with recent data on the proteins cross-linked to mouse L-cell polyribosomal mRNA (42), confirm and extend the original observation (20) that hnRNP and mRNP proteins are different. This suggests that a substantial change occurs in the RNP structure of hnRNA as it is processed and exported to the cytoplasm as mRNA, a conclusion which is also supported by recent nuclease protection experiments (33), as well as by analyses of the different proteins cross-linked to nuclear vs. polyribosomal poly(A) sequences (54). Further investigation of how and why these changes in RNP structure occur during mRNA processing may be facilitated by the photochemical cross-linking approach we have introduced.

Supported by National Science Foundation grant PCM77–19394 to J. R. Greenberg and National Institutes of Health grants GM21595, GM23914, GM28274, and P30 12708 to T. Pederson. Central instrumentation facilities were supported in part by the Mimi Aaron Greenberg Fund.

This is paper 16 in a series from the laboratory of T. Pederson entitled “Ribonucleoprotein Organization of Eukaryotic RNA.” The 15th paper in the series is reference 33.

Received for publication 9 March 1981.

REFERENCES
