MOLECULAR AND CELLULAR BIOLOGY, May 1988, p. 2237–2241
0270-7306/88/052237-05$02.00/0

Vol. 8, No. 5

Classification and Purification of Proteins of Heterogeneous Nuclear Ribonucleoprotein Particles by RNA-Binding Specificities

MAURICE S. SWANSON AND GIDEON DREYFUSS*

Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, Illinois 60208

Received 21 December 1987/Accepted 8 February 1988

Several proteins of heterogeneous nuclear ribonucleoprotein (hnRNP) particles display very high binding affinities for different ribonucleotide homopolymers. The specificity of some of these proteins at high salt concentrations and in the presence of heparin allows for their rapid one-step purification from HeLa nucleoplasm. We show that the hnRNP C proteins are poly(U)-binding proteins and compare their specificity to that of the previously described cytoplasmic poly(A)-binding protein. These findings provide a useful tool for the classification and purification of hnRNP proteins from various tissues and organisms and indicate that different hnRNP proteins have different RNA-binding specificities.

The nuclear precursors of mRNAs in eucaryotes are bound with specific proteins with which they form complexes termed heterogeneous nuclear ribonucleoprotein (hnRNP) particles (3, 13–16, 19–22, 27). These particles are major constituents of the nucleus and are important elements in the posttranscriptional pathway of the expression of genetic information. The composition and structure of hnRNP particles have therefore been the focus of much research (for a review, see reference 10). The purification of hnRNP particles with monoclonal antibodies to individual hnRNP proteins revealed that hnRNP particles are composed of heterogeneous nuclear RNA (hnRNA) and at least 20 polypeptides (7, 28). Antibodies against some of these proteins have been obtained and have proven to be extremely valuable tools; they have made it possible to study the role of hnRNP proteins in pre-mRNA processing (9, 30), to investigate their cellular localization (8, 12, 14, 18), and to isolate cDNA clones (24, 32). But little is known so far about most of the hnRNP proteins, and methods for their purification are needed so that their properties can be studied and antibodies against them can be prepared.

During studies on the binding of hnRNP proteins to pre-mRNA in nuclear extracts, we noticed that several of these proteins bind with very high affinity to certain RNA sequences and can be selectively eluted with different RNA homopolymers (M. S. Swanson and G. Dreyfuss, manuscript in preparation). We have therefore investigated systematically whether hnRNP proteins possess RNA-binding specificity towards RNA homopolymers and whether this characteristic can be used to purify these proteins. To do so, we examined the binding of HeLa cell proteins to the ribonucleotide homopolymers poly(G), poly(A), poly(U), and poly(C) at various salt concentrations. The use of homopolymer RNA-cellulose chromatography to isolate bacterial and bacteriophage proteins was first described by Carmichael (6) and was used in this study for the isolation of ribosomal protein S1 and termination factor ρ on poly(U) cellulose. This technique has also been extensively used to isolate poly(A)-binding proteins (PABPs) (2, 4). Here we show that several of the hnRNP proteins bind to specific RNA homopolymers at high salt concentrations and that this property can be used to identify and purify these proteins.

Nucleoplasm from [^{35}S]methionine-labeled HeLa cells was first digested with micrococcal nuclease, and the Ca^{2+}-dependent nuclease digestion was stopped by the addition of excess EGTA [ethylene glycol-bis(β-aminoethylether)-*N*,*N*,*N*′,*N*′-tetraacetic acid]. The treatment of the nucleoplasm with micrococcal nuclease was done to free hnRNP proteins from endogenous hnRNA and, more important, to prevent spurious binding of proteins to the ribopolymers, such as intact polyadenylated hnRNPs binding to poly(U). The RNA-binding experiments were carried out by incubating nucleoplasm, which had been adjusted to the indicated salt concentration (from 0.1 to 2 M), with Sepharose-bound polyribonucleotides for 10 min at 4°C (the binding of most of the observed protein occurred very rapidly and was essentially complete after 1 min). Identical aliquots of nucleoplasm corresponding to material from an identical number of cells were used for each incubation with each of the polyribonucleotides. Nonspecifically bound proteins were removed by a wash in the same buffer at the indicated salt concentration; the first wash also contained 2 mg of heparin per ml as an RNA competitor (28) and RNase inhibitor, although elimination of the heparin wash primarily affected the number of polypeptides bound at 0.1 to 0.5 M NaCl. The entire purification procedure required less than 30 min to complete and was performed at 4°C to minimize possible proteolysis and nuclease degradation of the ribopolymers. The bound proteins were eluted by being boiled in sodium dodecyl sulfate (SDS)-containing buffer and analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 1). For reference, total nucleoplasm, hnRNP particles immunopurified with the anti-hnRNP C protein monoclonal antibody 4F4 (lane 4F4), and C proteins immunopurified from micrococcal nuclease-digested nucleoplasm (lane 4F4MN) were included (Fig. 1).

At 0.1 M NaCl (Fig. 1), the hnRNP proteins A1 and A2 and/or the E proteins (two of which comigrate with A1 and A2 [see Fig. 3]) bound to all four ribonucleotide homopolymers, but for many proteins there were already clear binding differences at this salt concentration. The binding of several proteins (e.g., the 120,000-molecular-weight U protein [120K U protein]) to poly(G) and poly(U) was higher at 0.5 M NaCl than at 0.1 M NaCl, whereas the binding of other proteins was lower at 0.5 M NaCl, and the binding of still other proteins was the same at either salt concentration. The most significant differences in binding specificity were observed at NaCl concentrations of 1 M or higher. At these high salt

* Corresponding author.

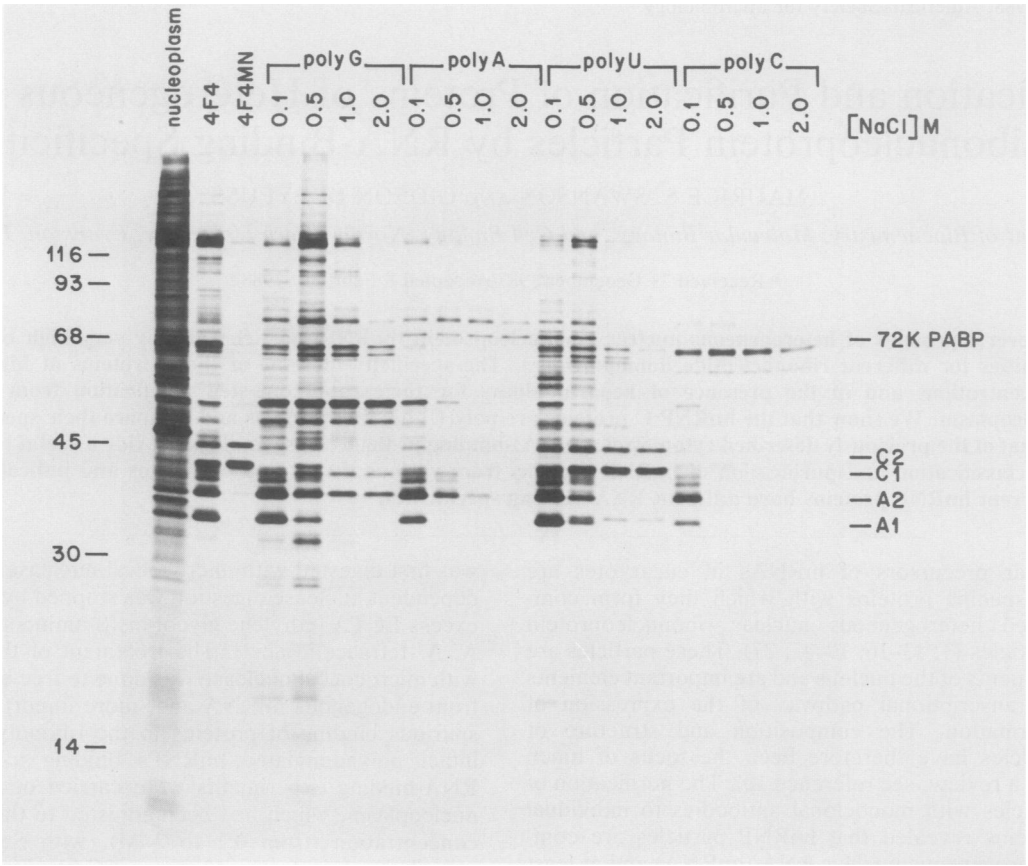

FIG. 1. Binding of nuclear proteins to ribonucleotide homopolymers. HeLa S3 cells were labeled overnight with [^{35}S]methionine (20 μCi/ml), and nucleoplasm was isolated as previously described (8, 28); 500 μl of nucleoplasm fraction was produced from one tissue culture plate (10-cm diameter) of subconfluent S3 cells. Samples (250 μl) of nucleoplasm (lane nucleoplasm) were immunoprecipitated either directly (lane 4F4) with the 4F4 monoclonal antibody, which had been previously bound to protein A-Sepharose in 10 mM Tris hydrochloride (pH 7.4)–100 mM NaCl–2.5 mM $MgCl_2$–0.5% Triton X-100–1 μg each of leupeptin and pepstatin per ml, and 0.5% aprotinin as described elsewhere (8), or following treatment with 100 U of micrococcal nuclease per ml (lane 4F4MN) for 10 min at 30°C in the presence of 1 mM $CaCl_2$, after which EGTA was added to 5 mM to inactivate the nuclease. Samples (250 μl) of this nuclease-treated nucleoplasm were then either left in 100 mM NaCl or had their NaCl concentrations increased to 0.5, 1, or 2 M by the addition of solid NaCl, and Triton X-100 was added to a final concentration of 0.5%. These samples were then incubated with 25 μl of poly(G), poly(A), poly(U), or poly(C), each of which had been bound to Sepharose beads (type 6, Pharmacia; 0.8 to 1.5 mg of polyribonucleotide per ml of agarose) and washed in the buffer described above, for 10 min at 4°C on a rocking platform. The beads were then spun down briefly in a Microfuge, incubated for another 10 min at 4°C in the buffer described above in the indicated salt concentrations plus 2 mg of heparin (porcine intestinal mucosa, grade I; Sigma Chemical Co.) per ml, and then washed four times in the same buffer without heparin. The drained beads were then boiled for 3 min in 50 μl of SDS-polyacrylamide gel electrophoresis loading buffer, and 10 μl of the sample was applied to 12.5% acrylamide SDS-polyacrylamide gels (11). After electrophoresis, the gels were dried and fluorographed (17). The positions of the mRNP 72K PABP and the hnRNP proteins A1, A2, C1, and C2 are indicated.

concentrations, proteins specific for each of the four ribonucleotide homopolymers were seen. The binding at such high salt concentrations reflects very high affinity. Two proteins, 41K and 43K, identified as the C1 and C2 hnRNP proteins, respectively, showed a clear binding preference for poly(U) (Fig. 1, lane polyU, 2 M NaCl). This identification was confirmed by immunoblotting with the specific anti-C protein monoclonal antibody, 4F4 (8) (Fig. 2A). Binding of the C proteins to poly(G) was also detected, but only at 0.5 M NaCl. However, in contrast to binding of many of the other poly(G)-binding proteins, C-protein binding to poly(G) was not resistant to 1 M NaCl. The binding of nucleoplasm proteins to poly(G) reproducibly increased at 0.5 M NaCl, possibly due to an overall increase in solubility at salt concentrations higher than 0.1 M or to a change in the structure of poly(G). For reference, we examined binding of the PABP, the only previously described eucaryotic ribonucleotide homopolymer-binding protein, to the four ribonucleotide homopolymers under the same conditions. In this experiment, we used yeast lysate instead of HeLa cell material to examine the cytoplasmic PABP, because this protein from yeast is the only PABP for which antibodies are available. We have previously described the production and specificity of this antibody (1). It is apparent that the PABP, at salt concentrations of 0.5 M or less, binds not only to poly(A) but also to poly(U) and poly(G). Only at a salt concentration of 1 M NaCl does the PABP bind only to poly(A) (Fig. 2B). The PABP is, therefore, not an absolutely base-specific binding protein, and its maximal specificity occurs at 1 M NaCl. The RNA-binding specificity of RNP proteins is thus a relative term and can be adequately determined by comparing the binding of the protein to all four ribonucleotide homopolymers at various salt concentrations. Using this criterion as an operational definition, we concluded that the C proteins are poly(U)-binding proteins and that there are also poly(C)- and poly(G)-binding proteins

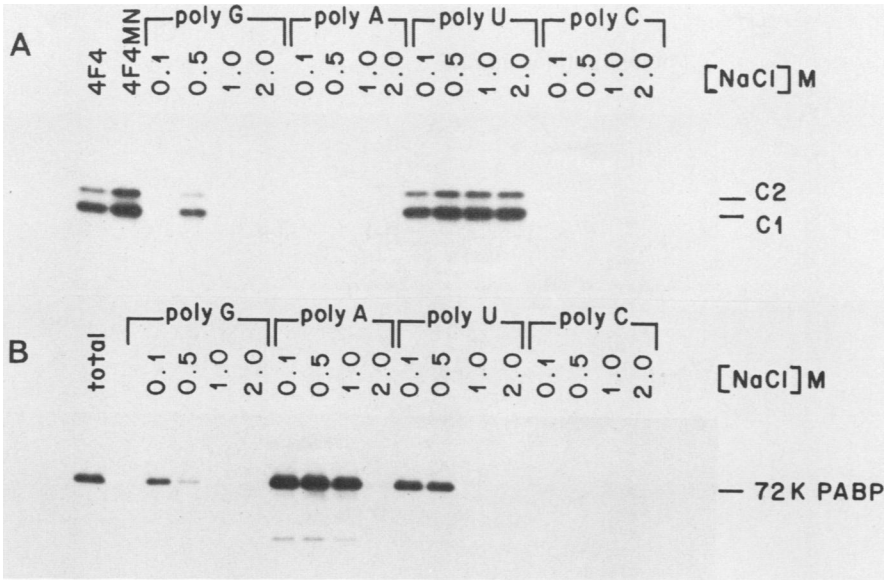

FIG. 2. Immunoblot analysis of nuclear proteins bound to ribonucleotide homopolymers. Samples prepared as described in the legend to Fig. 1 were electrophoresed on 12.5% polyacrylamide gels, and the proteins were transferred to nitrocellulose. Blots were probed with either the anti-human C protein monoclonal 4F4 (A) (8, 32) or a rabbit polyclonal antiserum to the yeast PABP (B) prepared as described elsewhere (1). The positions of the hnRNP C proteins C1 and C2 (A) and the yeast 72K PABP (B) are indicated.

in cells (Fig. 1, lanes polyC and polyG, 1 and 2.0 M NaCl). In an attempt to determine whether any of these proteins are hnRNP proteins, we have examined their mobilities by two-dimensional gel electrophoresis. By comparing the positions of known hnRNP proteins obtained by immunopurification (28) with the positions of the ribonucleotide homopolymer-binding proteins in nucleoplasm, Fig. 3 shows that these latter proteins are also components of hnRNP particles. The poly(C)-binding proteins are the hnRNP proteins K and J. The two faster-migrating E proteins, H, the slower-migrating F, and the M proteins are all poly(G)-binding proteins. The hnRNP P proteins are PABPs. Of these, the faster-migrating one is of 72 kilodaltons, and it migrates to a position which would be expected, from comparison with the yeast PABP (1), for the mRNA PABP. If this assignment of the 72K PABP is correct, then the binding of the mammalian mRNA 72K PABP to poly(A) is more resistant to higher salt concentrations (cf. Fig. 1, lane polyA, 2.0 M NaCl, and Fig. 2B, lane polyA, 2.0 M NaCl) than is the yeast protein. The slower-migrating P protein appears to be a polypurine-binding protein, rather than a specific PABP, because it binds well to poly(G) and poly(A). The M proteins also bind to poly(U), but over 90% of the material bound to poly(U) is C proteins.

It is striking that the binding of the C proteins to poly(U) at 2 M NaCl is as specific as the binding of C proteins to the anti-C protein monoclonal antibody itself (Fig. 1, lane 4F4MN). This feature provides the basis for a general, one-step affinity chromatography method for the purification of these proteins from any biological material. This, of course, also applies to the other high-affinity ribonucleotide homopolymer binding proteins. The method can be readily scaled up and used in a column, and the bound proteins can be eluted with 8 M urea or 6 M guanidine hydrochloride instead of SDS. Using this method, we have recently isolated the nuclear poly(U)- and poly(C)-binding proteins (C1 and C2 and K and J, respectively) from several organisms and used them as immunogens (M. Matunis, R. Bandziulis, and G. Dreyfuss, unpublished results). Poly(U)-Sepharose chromatography has also been previously used as a step in the purification of the human lupus antigen La protein. This protein is not likely to contaminate a C-protein preparation, because it elutes from the column at between 1.5 and 2.5 M NaCl. The purification of several hnRNP proteins by affinity chromatography on ribonucleotide homopolymers complements the fractionation of these proteins by chromatography on single-stranded-DNA columns as described recently (26, 28). In the latter, additional steps, such as ion-exchange chromatography, are required to obtain proteins of comparable purity to that obtained by a single-step purification on RNA-Sepharose columns.

The findings described here, in addition to their practical application for the purification of hnRNP proteins, provide a useful means of classifying these proteins. For example, once hnRNP proteins can be purified from organisms for which no such information is presently available, such a criterion as RNA-binding specificity may be essential for relating these proteins to the HeLa proteins. Furthermore, these findings provide a new view of hnRNP proteins. These proteins have long been considered general RNA-binding proteins which have no sequence specificity, akin to histones in chromatin. On the basis of the observations described here, it appears that several of the hnRNP proteins do, in fact, have RNA-binding specificities. Of course, these observations have only limited usefulness in predicting where on a particular hnRNA the hnRNP proteins will bind, except in the cases of several specific RNA segments. For example, the 3′ poly(A) tail is the longest, most common, and best-characterized ribohomopolymer segment in mRNAs and pre-mRNAs, and it is bound by the PABP which was first described in 1973 (5). In addition, long homopolynucleotide segments in both hnRNA and mRNA have been described. Large hnRNAs also contain 30-nucleotide-long poly(U) segments (23), and long stretches of U residues are common at the 3′ end of eucaryotic genes—for example, the 3′ noncoding region of the yeast PABP contains 30 consecutive pyrimidines, 22 of which are U residues (1). We also note that the polypyrimidine stretch at the 3′ end of introns,

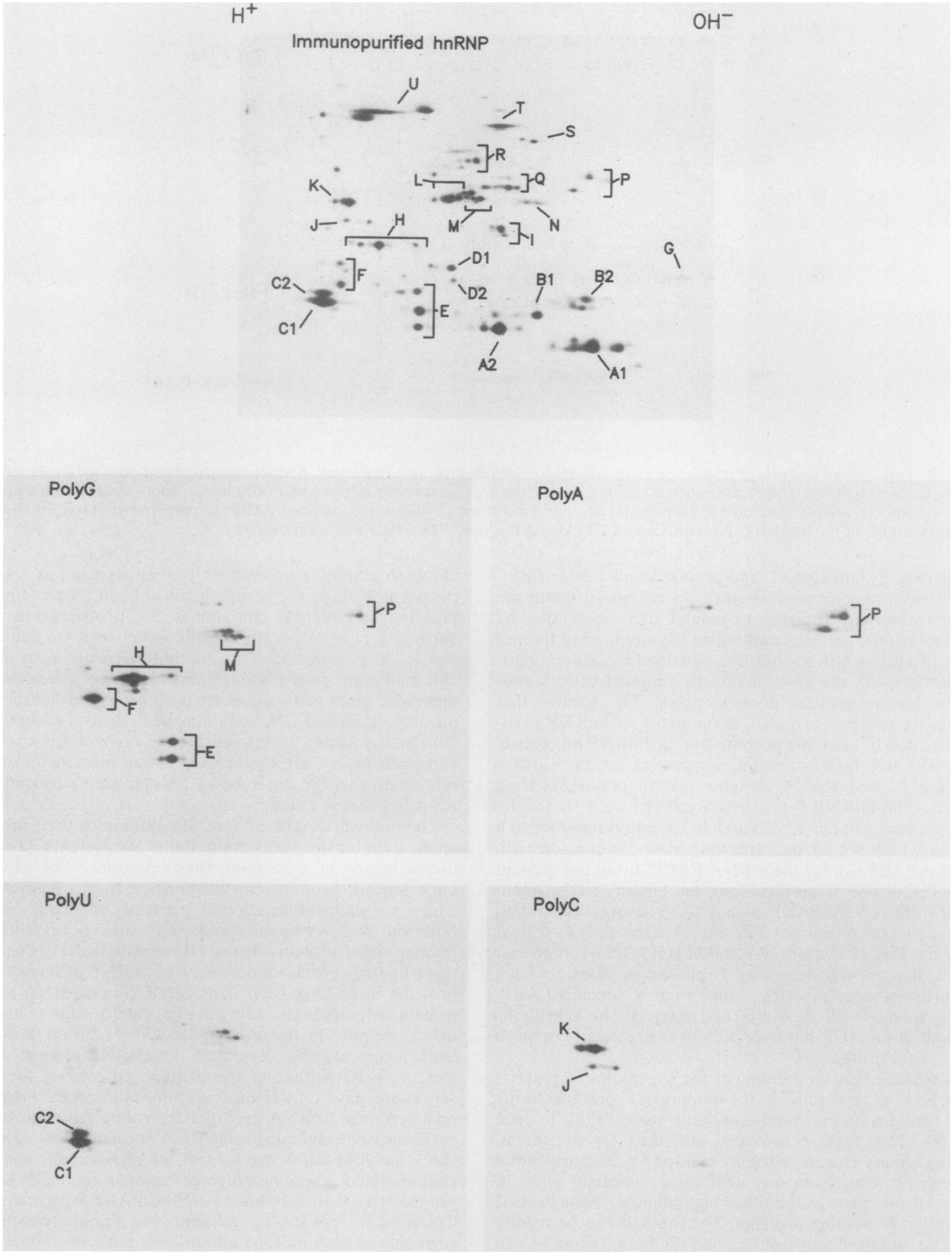

FIG. 3. Two-dimensional nonequilibrium gel electrophoresis of specifically bound hnRNP proteins. Immunopurified hnRNP particles and nucleoplasm proteins bound to Sepharose beads bearing poly(G), poly(A), poly(U), or poly(C) at 2 M NaCl were electrophoresed in the first dimension on nonequilibrium isoelectric-focusing gels (25) and in the second dimension on 12.5% polyacrylamide gels. The designation of the major hnRNP polypeptides is according to Piñol-Roma et al. (28).

sequences within the 3′ untranslated regions of pre-mRNAs, and sequences downstream of the AAUAAA poly(A) recognition element are often very uridine rich (29). Direct experiments are in progress to determine where hnRNP proteins bind on specific pre-mRNAs.

We thank Michael J. Chamberlin for helpful discussion and Raymond Bandziulis and Serafin Piñol-Roma for comments on the manuscript.

This research was supported by Public Health Service grants GM-31888 and GM-37125 from the National Institutes of Health. G.D. is an Established Investigator of the American Heart Association.

LITERATURE CITED

1. **Adam, S. A., T. Nakagawa, M. S. Swanson, T. K. Woodruff, and G. Dreyfuss.** 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. Mol. Cell. Biol. **6**:2932–2943.
2. **Baer, B. W., and R. D. Kornberg.** 1983. The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. J. Cell Biol. **96**:717–721.
3. **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.
4. **Blanchard, J. M., C. Brissac, and P. Jeanteur.** 1974. Characterization of a protein species isolated from HeLa cell cytoplasm by affinity chromatography on polyadenylate-sepharose. Proc. Natl. Acad. Sci. USA **71**:1882–1886.
5. **Blobel, G.** 1973. A protein of molecular weight 78,000 bound to the polyadenylate region of eukaryotic messenger RNA. Proc. Natl. Acad. Sci. USA **70**:924–928.
6. **Carmichael, G. G.** 1975. Isolation of bacterial and phage proteins by homopolymer RNA-cellulose chromatography. J. Biol. Chem. **250**:6160–6167.
7. **Choi, Y. D., and G. Dreyfuss.** 1984. Isolation of the heterogeneous nuclear ribonucleoprotein complex (hnRNP): a unique supramolecular assembly. Proc. Natl. Acad. Sci. USA **81**: 7471–7475.
8. **Choi, Y. D., and G. Dreyfuss.** 1984. Monoclonal antibody characterization of the C proteins of heterogeneous nuclear ribonucleoprotein complexes in vertebrate cells. J. Cell Biol. **99**:1997–2004.
9. **Choi, Y. D., P. J. Grabowski, P. A. Sharp, and G. Dreyfuss.** 1986. Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. Science **231**:1534–1539.
10. **Dreyfuss, G.** 1986. Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. Annu. Rev. Cell Biol. **2**:459–498.
11. **Dreyfuss, G., S. A. Adam, and Y. D. Choi.** 1984. Physical change in cytoplasmic messenger ribonucleoproteins in cells treated with inhibitors of mRNA transcription. Mol. Cell. Biol. **4**: 415–423.
12. **Dreyfuss, G., Y. D. Choi, and S. A. Adam.** 1984. Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. Mol. Cell. Biol. **4**:1104–1114.
13. **Economides, 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.
14. **Fakan, S., G. Leser, and T. E. Martin.** 1984. Ultrastructural distribution of nuclear ribonucleoprotein as visualized by immunochemistry on thin sections. J. Cell Biol. **98**:358–363.
15. **Jacob, M., G. Devilliers, J. P. Fuchs, H. Gallinaro, R. Gattoni, C. Judes, and J. Stévenin.** 1981. Isolation and structure of the ribonucleoprotein fibrils containing heterogeneous nuclear RNA, p. 194–259. *In* H. Busch, (ed.), The cell nucleus, vol. 8. Academic Press, Inc., New York.
16. **Karn, J., G. Vidali, L. C. Boffa, and V. G. Allfrey.** 1977. Characterization of the non-histone nuclear proteins associated with rapidly labeled heterogeneous nuclear RNA. J. Biol. Chem. **252**:7307–7322.
17. **Laskey, R. A., and A. D. Mills.** 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. Eur. J. Biochem. **56**:335–341.
18. **Leser, G. P., J. Escara-Wilke, and T. E. Martin.** 1984. Monoclonal antibodies to heterogeneous nuclear RNA-protein complexes. J. Biol. Chem. **259**:1827–1833.
19. **Lothstein, L., H. P. Arenstorf, J. C. Wooley, S. Y. Chung, B. W. Walker, and W. M. LeStourgeon.** 1985. The general organization of protein and RNA in HeLa 40S ribonucleoprotein particles. J. Cell Biol. **100**:1570–1581.
20. **Martin, T. E., P. B. Billings, J. M. Pullman, B. J. Stevens, and A. J. Kinniburgh.** 1978. Substructures of nuclear ribonucleoprotein complexes. Cold Spring Harbor Symp. Quant. Biol. **42**: 899–909.
21. **Mayrand, S., B. Setyono, J. R. Greenberg, and T. Pederson.** 1981. Structure of nuclear ribonucleoprotein: identification of proteins in contact with poly(A)$^{+}$ heterogeneous nuclear RNA in living HeLa cells. J. Cell Biol. **90**:380–384.
22. **Miller, O. L., and A. H. Bakken.** 1972. Morphological studies of transcription. Karolinska Symp. Res. Methods Reprod. Endocrinol. **5**:155–167.
23. **Molloy, G. R., W. Jelinek, M. Salditt, and J. E. Darnell.** 1974. Arrangement of specific oligonucleotides within poly(A) terminated hnRNA. Cell **1**:43–53.
24. **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.
25. **O'Farrell, P. Z., H. M. Goodman, and P. H. O'Farrell.** 1977. High resolution two-dimensional electrophoresis of basic as well as acidic proteins. Cell **12**:1133–1142.
26. **Pandolfo, M. O., O. Valentini, G. Biamonti, P. Rossi, and S. Riva.** 1987. Large-scale purification of hnRNP proteins from HeLa cells by affinity chromatography on ssDNA-cellulose. Eur. J. Biochem. **162**:213–220.
27. **Pederson, T.** 1974. Proteins associated with heterogeneous nuclear RNA in eukaryotic cells. J. Mol. Biol. **83**:163–183.
28. **Piñol-Roma, S., Y. D. Choi, M. M. Matunis, and G. Dreyfuss.** 1988. Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA binding proteins. Genes Dev. **2**:215–227.
29. **Shapiro, M. B., and P. Senapathy.** 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. **15**:7155–7174.
30. **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**:5241–5254.
31. **Stefano, J. E.** 1984. Purified lupus antigen La recognizes an oligouridylate stretch common to the 3′ termini of RNA polymerase III transcripts. Cell **36**:145–154.
32. **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-rRNA-binding proteins. Mol. Cell. Biol. **7**:1731–1739.