The EMBO Journal vol.13 no.1 pp.213-221, 1994

RNA annealing activities in HeLa nuclei

Douglas S.Portman and Gideon Dreyfuss[1]

Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6148, USA
[1]Corresponding author

Communicated by R.Lührmann

RNA-RNA base pairing plays a critical role in the interactions between pre-mRNAs and *trans*-acting factors during the processing of pre-mRNAs (hnRNAs) into mRNAs, and it is likely that specific factors are required to promote the annealing of RNAs. To identify particular nuclear components that have such activity, we fractionated HeLa nucleoplasm and assayed for activity which promoted the hybridization of a pre-mRNA with an antisense RNA probe complementary to 60 nucleotides (nt) encompassing the 3′ splice site. At least nine major RNA annealing activities were identified and, surprisingly, eight of these copurified partially or to homogeneity with known hnRNP proteins. The activities of three of these proteins, hnRNP A1, C1 and U, were confirmed using purified recombinant proteins. Moreover, we found that the RNA binding domain alone of hnRNP C1/C2 had significant activity, indicating that this RNA annealing may result, at least partly, from chaperone activity: a direct modulation of RNA conformation by hnRNP proteins. The finding that hnRNP proteins have strong RNA annealing activity indicates that they can profoundly affect the interactions of pre-mRNAs with *trans*-acting factors and suggests this to be an important function of hnRNP proteins in the processing of pre-mRNAs.

Key words: hnRNP/pre-mRNA/RNA binding proteins/RNA chaperones/RNA processing

Introduction

Nascent RNA polymerase II transcripts (pre-mRNAs, or hnRNAs) mature into processed, translatable mRNAs by a complex and highly orchestrated sequence of events whereby the 5′ cap structure is generated, introns are excised and the poly(A) tail is added. Inherent in these processes, particularly for pre-mRNA splicing, are rapid, precise and ordered interactions between the pre-mRNA and *trans*-acting factors (Green, 1991; Guthrie, 1991; Steitz, 1992; Moore *et al.*, 1993). Of these, the best characterized are the base pairing interactions between the pre-mRNA and snRNPs, which form the core of the spliceosome. It has recently been proposed (Steitz, 1992) that these associations, which involve multiple RNA-RNA interactions (between U2 snRNA and the pre-mRNA branch point, between U1 snRNA and 5′ and 3′ intron sequences and between U5 snRNA and 5′ and 3′ exon sequences), result in a dynamic Holliday-like structure within which the splicing reaction occurs. Direct pre-mRNA-protein interactions are also important for constitutive and regulated splicing (reviewed in Lamm and Lamond, 1993) and also for 3′ end formation (Takagaki *et al.*, 1992). In addition to these associations *in trans*, specific *cis*-acting interactions within the pre-mRNA are also thought to be important for splicing in certain cases (Clouet d'Orval *et al.*, 1991; Libri *et al.*, 1991; Goguel and Rosbash, 1993).

It is not yet understood how these specific interactions can occur with the speed and high fidelity required *in vivo*. Pre-mRNA introns average over 1 kb in length (Hawkins, 1988) and some exceed 100 kb (e.g. Miyatani *et al.*, 1992), presenting a considerable background of sequence information within which short, specific regions must be precisely located. This problem of recognition seems likely to be exacerbated by the potential for formation of incidental higher order structure in the pre-mRNA, which could obscure the *cis*-acting signals required by the processing apparatus. Because pre-mRNA processing maintains accuracy and speed despite these complications, it seems likely that factors have evolved which modulate the structure of the pre-mRNA and its associations with *trans*-acting factors.

As an assay to identify factors which may facilitate interactions involving pre-mRNAs, we studied the annealing of a complementary RNA probe to a pre-mRNA molecule *in vitro*. Some of the factors which stimulate this reaction are likely to act in a manner which is specific to RNA-RNA base pairing; however, this assay should also identify proteins which are able to promote specific interactions between *trans*-acting factors in general and pre-mRNAs. An RNA annealing assay is therefore useful for identifying candidates for proteins which help to overcome some of the complications inherent in pre-mRNA processing.

The capacity of HeLa nuclear extract to facilitate RNA annealing has been noted previously in studies of *in vitro* splicing reactions (Konarska *et al.*, 1985; Munroe, 1988). Subsequently, hnRNP A1, a major protein of the human hnRNP complex (reviewed in Dreyfuss *et al.*, 1993; Görlach *et al.*, 1993), has been shown to stimulate strongly DNA-DNA as well as RNA-RNA annealing *in vitro* (Kumar and Wilson, 1990; Pontius and Berg, 1990; Munroe and Dong, 1992). Other nuclear HeLa factors previously reported to have RNA annealing activity include the pre-mRNA binding proteins SF2/ASF (Krainer *et al.*, 1990) and U2AF^{65} (Lee *et al.*, 1993) and the tumor-suppressor protein p53 (Oberosler *et al.*, 1993).

Here we describe a systematic fractionation of RNA annealing activities from HeLa cell nucleoplasm. We found at least nine distinct activities after four chromatographic steps. Surprisingly, eight of these copurified with known proteins of the human hnRNP complex. Moreover, a fragment containing only the RNP-consensus sequence (RNP-CS) RNA binding domain of the human hnRNP

C1/C2 proteins has strong RNA annealing activity. These data demonstrate that hnRNP proteins can profoundly affect the character of pre-mRNAs through RNA–protein interactions. Furthermore, they suggest that certain hnRNP proteins function as important components of the nuclear RNA processing pathway, probably by modulating the conformation of the pre-mRNA and promoting associations between pre-mRNAs and *trans*-acting factors.

Results

RNA annealing assay

The assay we used to measure RNA annealing activity is based on base pairing between two *in vitro* synthesized RNAs, as shown in Figure 1A. The 'sense' RNA is a 454 nt long capped pre-mRNA derived from the adenovirus-2 first leader intron (Konarska *et al.*, 1984). The 'probe' RNA is a 93 nt long capped, internally ^{32}P-labeled RNA, consisting of 60 nt of sequence completely complementary to the region surrounding the 3′ splice site of the sense RNA, flanked by 17 and 16 nt at the 5′ and 3′ ends, respectively, of non-complementary sequence. Upon duplex formation these single-stranded overhangs remain susceptible to RNase T1 digestion, leaving a 66 nt protected fragment. Any unannealed probe RNA is digested to small RNA fragments. Therefore, the amount of the 66 nt fragment protected upon RNase treatment provides a direct measurement of the extent of the formation of dsRNA. The use of this assay is shown in Figure 1B. Lanes 2 and 4 show the RNase digestion pattern of the probe RNA in its single-stranded and annealed forms, respectively, and lanes 5–11 show that no RNA annealing takes place in the absence of protein or in the presence of bovine serum albumin, *Escherichia coli* single-strand binding protein or cytochrome *c*. Lane 12 shows that HeLa nucleoplasm contains RNA annealing activity as previously reported (Konarska *et al.*, 1985; Munroe, 1988) and lane 13 shows that the assay is suitable for detecting the activity of a known RNA annealing protein, hnRNP A1 (Kumar and Wilson, 1990; Pontius and Berg, 1990; Munroe and Dong, 1992).

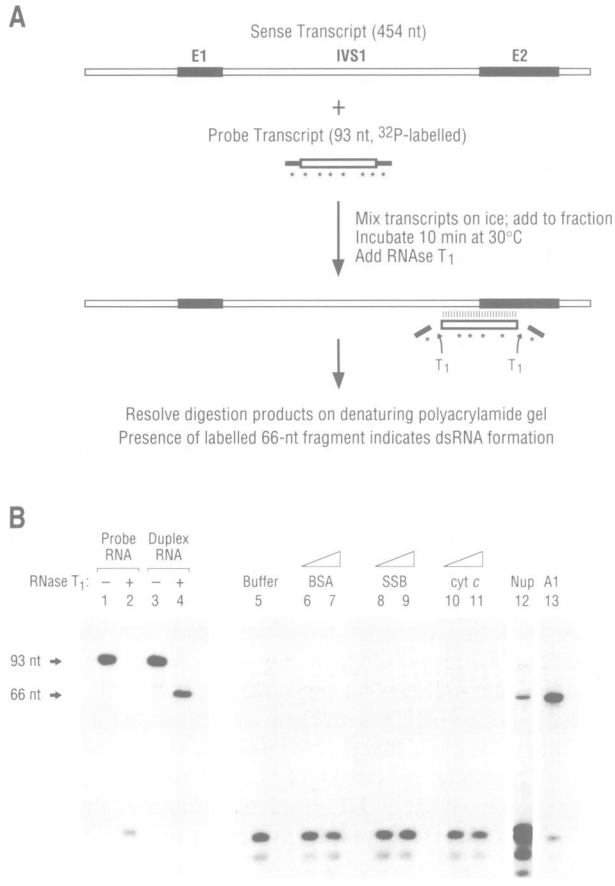

Fig. 1. RNA annealing assay. (**A**) RNA annealing activity was assayed using two complementary RNAs: the 454 nt sense RNA and a 93 nt, ^{32}P-labeled probe RNA, 60 nt of which were perfectly complementary to the sense RNA. Upon digestion by RNase T1, a 66 nt fragment of the probe RNA will be protected if the probe is annealed to the sense RNA. (**B**) RNA standards and control RNA annealing assay. Lanes 1 and 2, probe RNA without and with RNase T1 digestion, respectively. Lanes 3 and 4, probe RNA which had been annealed to sense RNA (by preincubation at 65°C in 1 M KCl) without and with RNase T1 digestion, respectively. Lane 5, RNA annealing assay in the absence of protein. Lanes 6–11, control RNA annealing assays containing 10 nM and 1 μM bovine serum albumin, *E.coli* single-strand binding protein and cytochrome *c*, respectively. Lanes 12 and 13, RNA annealing assays of total HeLa nucleoplasm and 200 nM purified, bacterially expressed hnRNP A1.

Fractionation of RNA annealing activities

As an initial step in identifying factor(s) in HeLa cell nuclei which have RNA annealing activity, we fractionated nucleoplasm on a ssDNA–cellulose column and assayed fractions for RNA annealing activity. The ssDNA

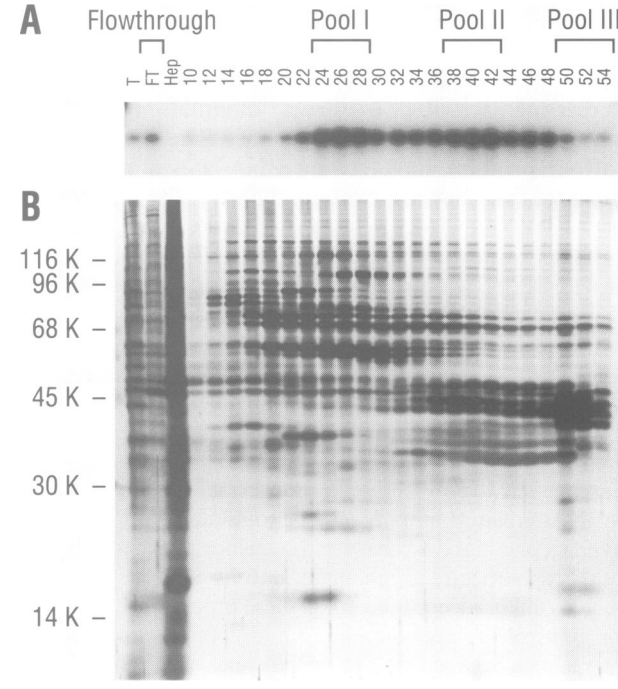

Fig. 2. ssDNA chromatography of RNA annealing activities. Nucleoplasmic proteins were fractioned on a ssDNA–cellulose column; fractions were collected and assayed for RNA annealing activity. (**A**) RNA annealing assay of the ssDNA–cellulose column fractions. For clarity, only the region of the gel containing the 66 nt protected fragment is shown. The four pools selected for further chromatography are indicated (see text). (**B**) SDS–PAGE of the corresponding column fractions. Proteins were visualized by silver staining.

chromatography step significantly enriches the bound fractions for RNA binding proteins (Piñol-Roma *et al.*, 1988) that are good candidates for proteins which can modulate the interactions of RNAs. The results of this fractionation are illustrated in Figure 2, which shows the protein composition of the ssDNA column fractions (Figure 2B) and the RNA annealing activities of these fractions (Figure 2A). Notably, multiple distinct RNA annealing activities can be observed in this set of fractions. To further purify these factors four pools of activity were retained for further chromatography, as indicated in Figure 2. Pool I (fractions 24–28) and Pool II (fractions 38–42) have high levels of activity and appear to have different protein compositions. Pool III, though it appears to have a lower level of activity, has a protein composition distinct from that of Pool II, since it contained fractions which eluted from the column upon a step from 900 mM to 2 M KCl; it therefore seemed possible that Pool III might contain unique active factors. Also analyzed was the flowthrough from the ssDNA column, in order to determine whether the activities in the ssDNA column fractions represented the majority of RNA annealing activities present in nucleoplasm.

Each of these four pools was then further fractionated using a TSK–SP cation exchange column and the fractions were again assayed for RNA annealing activity. The results from these four fractionations are shown in Figure 3. The protein compositions of the fractions, as visualized by silver staining, are shown (Figure 3B, E, H and K), as well as their RNA annealing activities (Figure 3A, D, G and J) and the results of Western blotting experiments, to confirm the identities of particular proteins (see below; Figure 3C, F and I).

In the fractionation of Pool I (Figure 3A and B), three peaks of RNA annealing activity were apparent. (We named the fractionated activities 'RA' for RNA annealing factor, followed by a number.) RA1 activity, strongest in fraction 24, correlated well with the presence of a protein of ~75 kDa. While this activity is rather weak, it appeared consistently during other similar fractionations. To attempt to identify the 75 kDa protein we compared its mobility in a two-dimensional (2-D) NEPHGE/SDS–PAGE system with that of known nucleoplasmic ssDNA binding proteins (data not shown). This analysis showed that the 75 kDa protein exactly comigrated with the P protein of the hnRNP complex, which is known to elute from ssDNA–cellulose at approximately the same salt concentration as that which eluted Pool I fractions (Piñol-Roma *et al.*, 1988). The RNA annealing activity in fractions 27 and 28, called RA2, was probably a result of the nearly pure 33 kDa protein present in these fractions. Analysis of fraction 27 on 2-D gels (data not shown) showed that it contained a very basic protein which did not comigrate with any known nucleoplasmic ssDNA binding protein. It is possible that this protein is the splicing factor SF2/ASF, a basic 33 kDa pre-mRNA binding protein which has been previously shown to have RNA annealing activity (Krainer *et al.*, 1990). RA3 activity, which elutes at the end of the gradient fractionation of Pool I, copurified with the nearly pure 120 kDa protein in fractions 33–36. This factor was identified unambiguously as hnRNP U (Kiledjian and Dreyfuss, 1992) by its reactivity in a Western blot with the monoclonal antibody 3G6 (Dreyfuss *et al.*, 1984b) (Figure 3C). Verification that hnRNP U has RNA annealing activity is shown in Figure 5. Interestingly, this fractionation revealed that there are nucleoplasmic ssDNA binding proteins which do not have RNA annealing activity (fractions 12–22), indicating that RNA annealing activity is not a general property of single-stranded nucleic acid binding proteins.

Fractionation of Pool II (Figure 3D and E) also revealed at least three activities. The specific protein(s) responsible for RA4 activity (fractions 13–18) and RA5 activity (fractions 20 and 21) could not be definitively identified because multiple proteins were present in these fractions. However, at least four known pre-mRNA binding proteins were present in the RA4 fractions: hnRNP A2/B1 (which differs only by a small peptide insert; Burd *et al.*, 1989), hnRNP I (also known as the polypyrimidine tract binding protein, or PTB; Brunel *et al.*, 1991; Gil *et al.*, 1991; Patton *et al.*, 1991; Ghetti *et al.*, 1992) and hnRNP K (Matunis *et al.*, 1992). RA5 activity is probably due at least partly to the known RNA annealing protein hnRNP A1 which is present in these fractions, as judged by 2-D gel mobility (data not shown). By the same criteria, the other major proteins present in these fractions were hnRNP B2 and two isoforms of the hnRNP D proteins (Piñol-Roma *et al.*, 1988), both of which may contribute to the observed activity. Fractions 24–28, which have RA6 activity, contain at least two proteins, both of which were visualized very poorly upon silver staining. One factor, not visible in the figure, was seen in the stained gel as a 'halo' around a negatively staining band of ~43 kDa (indicated by the dotted box). Because this protein comigrated on 2-D gels (data not shown) with the hnRNP G protein (Piñol-Roma *et al.*, 1988), it was tested in a Western blot with a canine autoimmune antiserum which recognizes the human G protein (Soulard *et al.*, 1991, 1993) which confirmed its identity as hnRNP G (Figure 3F). Also present in fractions 26 and 27 was a band of ~30 kDa. It is not certain which of these proteins is responsible for RA6 activity.

Three additional RNA annealing activities were found upon the fractionation of Pool III, shown in Figure 3G and H. Of these three, only RA9 (fractions 29–31) could be definitively identified. Based upon their high salt elution from ssDNA–cellulose as well as molecular weight, we suspected that the two bands which copurified with RA9 represented the hnRNP C1 and C2 proteins (Piñol-Roma *et al.*, 1988). To test this, we probed a Western blot of these fractions with the monoclonal antibody 4F4 (Choi and Dreyfuss, 1984). The results, shown in Figure 3I, indicate that RA9 activity correlated extremely well with the presence of the C proteins. The RNA annealing activity of hnRNP C1 is shown below (Figure 5). Both the RA7 and RA8 activities, in fractions 12–14 and 18–22, respectively, contained at least two proteins, preventing the assignment of particular species to these activities. RA7 activity is probably a result of one of the two major proteins present in these fractions, hnRNP K and a member of the hnRNP E family. Further fractionation of the factor(s) equivalent to RA8 will be required for their unambiguous identification; candidate proteins present in these fractions are hnRNP A1 and members of the hnRNP D and E families. The identification of candidate factors for both RA7 and RA8 was based on 2-D gel electrophoresis (data not shown).

Fractionation of the ssDNA column flowthrough in a similar fashion is shown in Figure 3D. While the majority of nuclear proteins are present in these fractions, very little annealing activity was evident. Though we cannot rule out

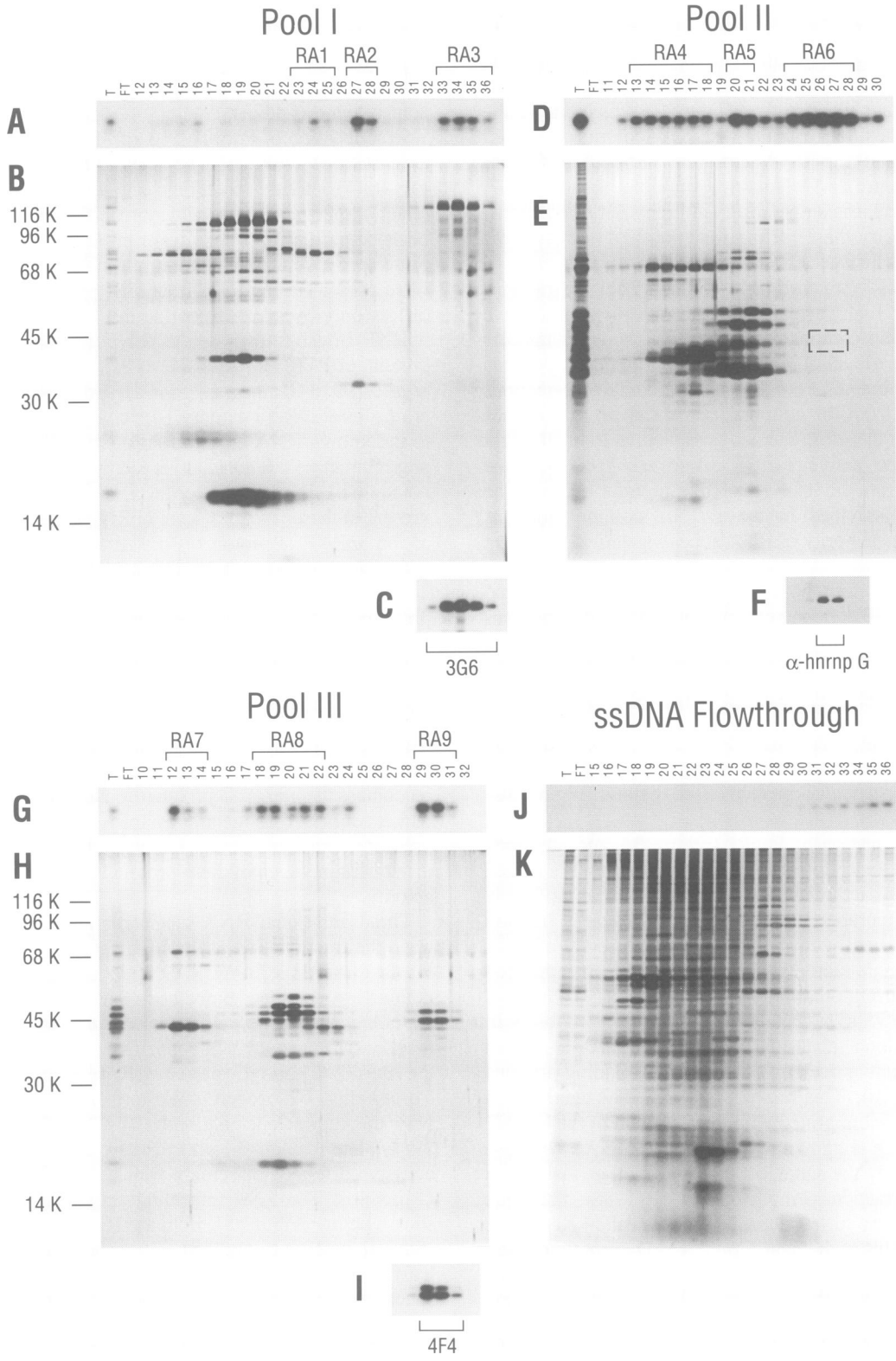

Fig. 3. Cation-exchange chromatography of ssDNA column fractions. (**A, D, G** and **J**) RNA annealing assays of the SP column fractions from ssDNA Pools I, II, III and flowthrough, respectively. The nine identified RNA annealing activities are indicated. For clarity, only the region of the gel containing the 66 nt protected fragment is shown. (**B, E, H** and **K**) Protein compositions of the corresponding column fractions, resolved by SDS−PAGE and visualized by silver staining. The location of the poorly staining 43 kDa protein in fractions 26 and 27 of Pool II is indicated by the dashed box (see text). (**C**) Western blot of fractions 32−36 of Pool I probed with the monoclonal antibody 3G6, which recognizes the human hnRNP U protein. (**F**) Western blot of fractions 25−28 of Pool II probed with a canine autoimmune antiserum which recognizes human hnRNP G (see Materials and methods). (**I**) Western blot of fractions 28−31 of Pool III probed with the monoclonal antibody 4F4, which recognizes the human hnRNP C1/C2 proteins.

the possibility of the presence of an inhibitor of RNA annealing, this experiment suggests that the three above fractionations were able to account for the majority of RNA annealing activities in HeLa nucleoplasm.

The purification scheme and the results described above are summarized in Figure 4. The nine RNA annealing activities resolved after the ssDNA column and four cation exchange gradients are shown, along with the identity of the active factor or candidate factors for each.

RNA annealing activity of recombinant hnRNP proteins

Because the data shown in Figure 3 identified several hnRNP proteins as having or being likely to have RNA annealing activity, we used purified, bacterially expressed proteins to test directly their activity in our assay. Figure 5A shows the results of an experiment in which serial dilutions of purified, bacterially expressed hnRNP A1, C1 and U proteins were tested for RNA annealing activity. A1 had maximal activity at 128–512 nM (corresponding to a protein:nucleotide ratio of between 1:1 and 1:4), causing the conversion of 60–70% of the probe RNA into the double-stranded form after 10 min. Activity was detectable at protein concentrations as low as 16 nM, a protein:nucleotide ratio of 1:31. [A previous study of the RNA annealing activity of A1 has reported maximal activity at a protein:nucleotide ratio of 1:11–1:19 (Munroe and Dong, 1992); such differences may be due to the length, concentration or sequence of RNA used in our assay.] The factor likely to be RA9, hnRNP C1, was also tested for activity. [HnRNP C2, also in the RA9 fractions, is closely related to C1, differing only by an insert of 13 amino acids (Burd *et al.*, 1989).] Purified, bacterially expressed C1 had strong RNA annealing activity *in vitro*, reaching a peak of activity at 64 nM (a protein:nucleotide ratio of 1:8), in which 45% of the probe RNA was protected from RNase T1 digestion. The activity of C1 was consistently observed to decrease quickly at above optimal concentrations (see Discussion). HnRNP U, fractionated as RA3, was also assayed as a purified bacterially expressed protein. The U protein had especially strong activity in this assay, stimulating the annealing of 100% of the input probe RNA. Moreover, hnRNP U had activity at significantly lower concentrations than A1 or C1, reaching its maximum at 32 nM (a protein:nucleotide ratio of 1:16), and having detectable activity at protein concentrations as low as 4 nM (a protein:nucleotide ratio of 1:125). These results confirm that hnRNP C1 and U are responsible for RA9 and RA3 activity, respectively.

The RNA binding domain of hnRNP C1/C2 has RNA annealing activity

It was of interest that hnRNP C1 exhibited activity comparable to that of A1, especially in light of the domain structures of these proteins: C1 does not contain a glycine-rich domain, which has been shown in the case of A1 to have strong RNA annealing activity (see Discussion). We tested the RNA annealing activity of the 94 amino acid RNA binding domain of C1 (designated 'K94') whose structure has recently been determined by multi-dimensional NMR (Görlach *et al.*, 1992). K94 contains only the core RNP-CS RNA binding domain of hnRNP C1 (Görlach *et al.*, submitted), it is a monomer in solution, and is unlikely to participate in protein–protein interactions (Wittekind *et al.*, 1993). Surprisingly, K94 had RNA annealing activity nearly as strong as that of full length hnRNP C1 (Figure 5B). At a protein concentration of 128–256 nM (a protein:nucleotide ratio of between 1:2 and 1:4), K94 caused 40–45% conversion of the single-stranded probe RNA into the duplex form. These data suggest that it is the RNA binding domain of hnRNP C1, rather than its acidic auxiliary domain, which mediates its RNA annealing activity.

Discussion

We have reported here a systematic detection and fractionation of nucleoplasmic HeLa cell RNA annealing activities. This approach has detected nine chromatographically distinct RNA annealing activities, which together are likely to

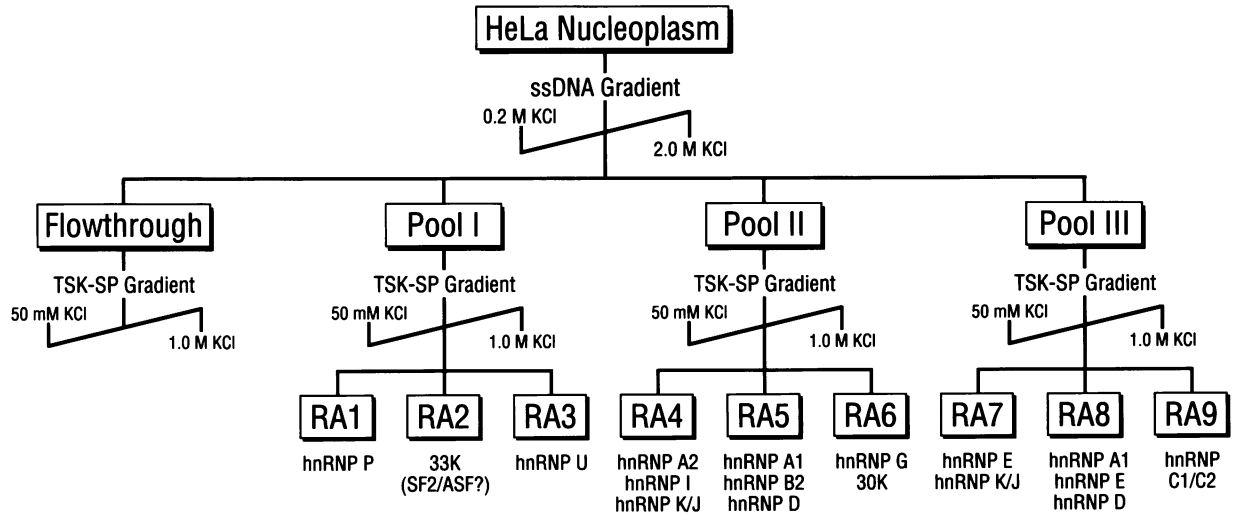

Fig. 4. Purification of RNA annealing activities. The fractionation scheme, as described in the text, used to identify nine RNA annealing activities present in HeLa nucleoplasm, RA1–RA9. Known factors copurifying with the RNA annealing activities, as identified by Western blots and/or 2-D gel electrophoresis, are listed under each observed activity.

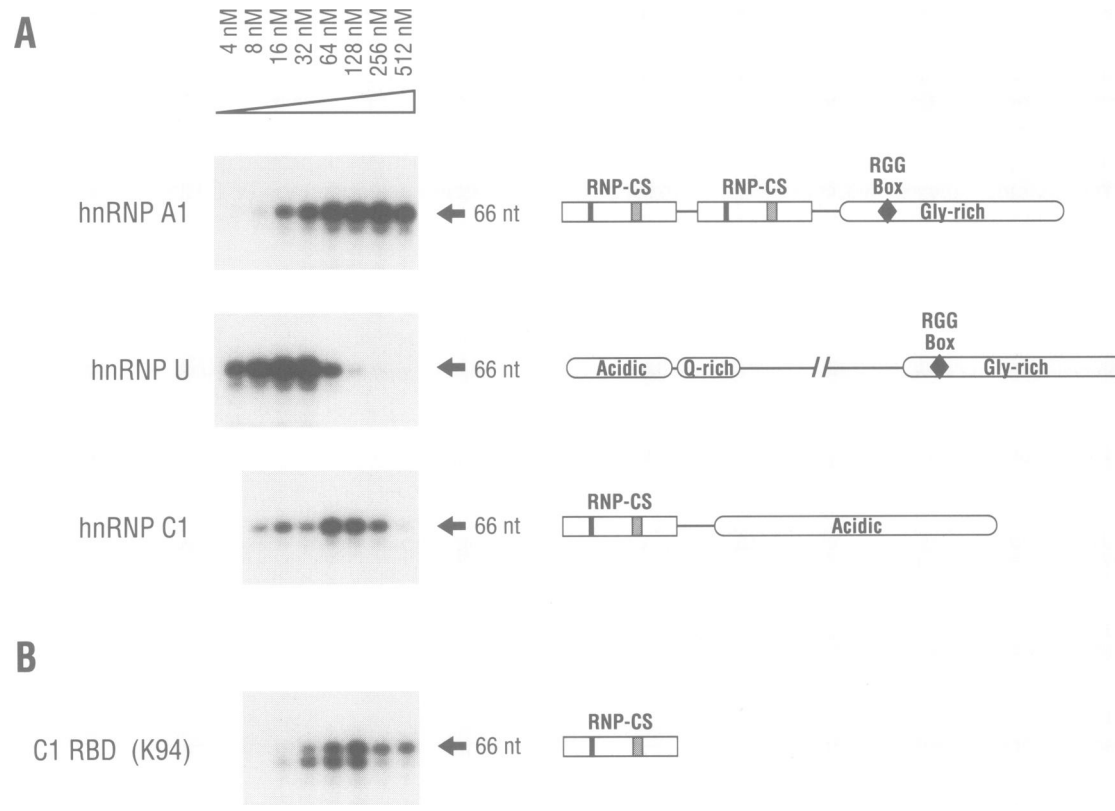

Fig. 5. RNA annealing activities of recombinant human hnRNP proteins. (**A**) RNA annealing activities of the human hnRNP A1, C1 and U proteins, overexpressed in *E.coli* and purified to near homogeneity. Serial dilutions of each protein were assayed, as indicated. For clarity, only the region of the gel containing the 66 nt protected fragment is shown. The domain structures of these proteins are illustrated on the right. (**B**) RNA annealing activity of the 94 amino acid fragment of the human hnRNP C proteins, which contains only their RNP-CS-type RNA binding domain.

represent the majority of the total nuclear RNA annealing activity in HeLa nuclei; moreover, we have shown that eight of them copurify with previously known hnRNP proteins. There are ~20 abundant human hnRNP proteins, denoted hnRNP A1 through hnRNP U (Piñol-Roma *et al.*, 1988), that associate with nascent RNA polymerase II transcripts (for reviews see Dreyfuss *et al.*, 1993; Görlach *et al.*, 1993). The finding that multiple hnRNP proteins have RNA annealing activity is particularly interesting, since the recognition of splice sites appears to occur when the pre-mRNA is associated with hnRNP proteins (Beyer and Osheim, 1988; Bennett *et al.*, 1992; Matunis *et al.*, 1993). While the specific contributions of these pre-mRNA binding proteins to the process of RNA metabolism is still not fully understood, our findings suggest that one function of hnRNP proteins is to promote or modulate the interactions of pre-mRNAs with *trans*-acting factors, making the pre-mRNAs suitable substrates for RNA processing events.

While the main point we wish to emphasize is that hnRNP proteins can strongly influence the interactions of pre-mRNAs, the mechanisms by which they do this are also of interest. There are two extreme but not mutually exclusive general models, shown in Figure 6, whereby proteins can facilitate RNA annealing. We refer to the active factors in these two models as 'matchmakers' and 'chaperones', respectively, after the nomenclature of Racker (1992) in his discussion of protein chaperones, which have essential roles as modulators of the conformations and associations of nascent polypeptides *in vivo* (reviewed in Gething and Sambrook, 1992). We propose here that hnRNP proteins can act on pre-mRNA molecules in functionally analogous ways.

In the first model (Figure 6A), hnRNP proteins act as matchmakers, bringing together two ribonucleoprotein complexes through protein−protein interaction. This scenario has recently been proposed for the DNA renaturation activity of hnRNP A1 (Pontius and Berg, 1992; Pontius, 1993), whereby flexible, repeating domains (such as the glycine-rich domain of A1) can stabilize the transient complex formed when two A1-bound strands collide. Prolonging the lifetime of this complex through protein−protein interactions increases the opportunity for the formation of stable base pairing interactions, making it significantly more likely that each encounter of complementary strands will lead to annealing. In support of this model, the carboxy-terminal glycine-rich domain of A1, which has strong RNA annealing activity (Munroe and Dong, 1992), confers co-operative binding of A1 to RNA (Cobianchi *et al.*, 1988), a property which may be a result of homotypic protein−protein interactions. Additionally, there are indications that some hnRNP proteins are capable of oligomerization under certain conditions (Barnett *et al.*, 1991), consistent with a potential for protein−protein interaction.

In the second model (Figure 6B), we propose that some hnRNP proteins can act as chaperones of RNA, thereby modulating RNA conformation. This activity would be functionally analogous to that of the hsp70 family of protein chaperones in maintaining unstructured conformations of nascent polypeptides to prevent misfolding and inappropriate

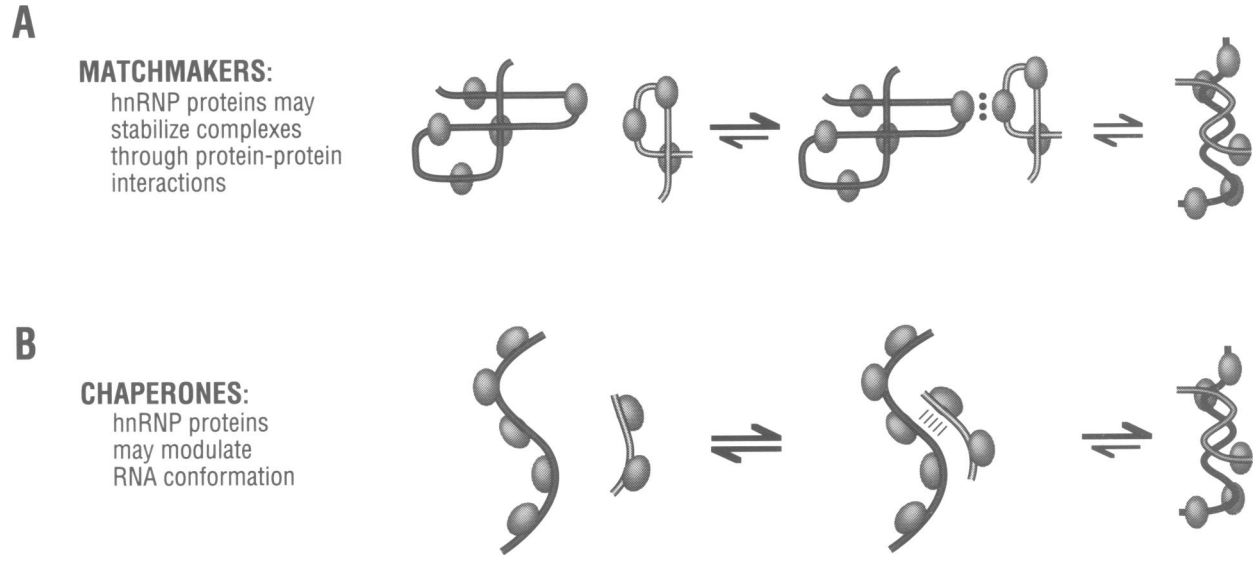

Fig. 6. Mechanisms for RNA annealing (see text). (**A**) Matchmakers serve to facilitate RNA annealing by first bringing two ribonucleoprotein complexes together through protein–protein interaction. (**B**) Chaperones modulate the conformation of the RNA, increasing the accessibility of the nucleotides for interactions *in trans*.

interactions (Craig, 1993). In the case of chaperones of RNA, the proteins may influence RNA conformation upon binding, resulting in the stabilization of a relatively open, extended RNA structure in which the nucleotides are exposed and available for interaction. This effect could stimulate RNA annealing both by increasing the likelihood that collisions between complementary molecules will be productive, since regions of complementarity are not concealed, and by reducing the extent of structure which must be undone during the zippering process (Munroe and Dong, 1992). The binding of these proteins could also cause a shielding of the charge of the phosphate backbone, which could contribute to RNA annealing activity by reducing intermolecular repulsion.

Several lines of evidence support a role for hnRNP proteins as chaperones of RNA which can modulate RNA–RNA interactions. An amino-terminal fragment of hnRNP A1 consisting of its two RNP-CS domains, called UP1, has been shown to facilitate the renaturation of tRNA and 5S RNA molecules from biologically inactive conformations (Karpel *et al.*, 1982). Though UP1 has only weak RNA annealing activity (Munroe and Dong, 1992), its renaturation activity is consistent with an ability of RNP-CS domains to modulate RNA structures and promote stable interactions (see below). More recently, the glycine-rich domain of nucleolin, which is similar to that of A1, has been shown to have the ability to unstack bases in RNA molecules (Ghisolfi *et al.*, 1992), suggesting that A1 may also be able to affect RNA conformation in this manner; interestingly, nucleolin also has DNA annealing activity (Sapp *et al.*, 1986).

Furthermore, the chaperone model is strongly supported by the demonstration that K94, the 94 amino acid RNP-CS RNA binding domain of hnRNP C1/C2, has significant RNA annealing activity. Because K94 is thought not to be involved in protein–protein interactions, whether free or when bound to RNA (Wittekind *et al.*, 1993), the matchmaker model is ruled out in this case. Since this isolated RNP-CS domain is able to stimulate RNA annealing, it seems quite likely that the activity of hnRNP C1 is through a direct effect on the structure of the pre-mRNA. Though not all RNP-CS domains have strong RNA annealing activity [e.g., those of hnRNP A1 (Munroe and Dong, 1992)], we propose that RNA annealing may be a general activity of many proteins containing this domain. This idea is supported by the recently solved structures of K94 (Görlach *et al.*, 1992) and an RNP-CS domain of the U1 snRNP A protein (Nagai *et al.*, 1990; Hoffmann *et al.*, 1991), which suggest that RNP-CS domains can act as platforms, binding substrates in a manner in which they remain accessible to other factors (Görlach *et al.*, 1992).

Implications for functions of hnRNP proteins

The data presented here strongly support a model (Dreyfuss *et al.*, 1993; Görlach *et al.*, 1993) in which hnRNP proteins play an important role in the processing of pre-mRNA, by acting as both chaperones and matchmakers. Upon binding to nascent transcripts, hnRNP proteins would cause the pre-mRNA to adopt a relatively extended conformation in which critical sites on the molecule, such as the 5′ and 3′ splice sites, the branch point, the poly-pyrimidine stretch and polyadenylation signals are accessible to specific factors which recognize them *in trans*. Intriguingly, several studies have demonstrated that many predicted pre-mRNA secondary structures which exist *in vitro* do not form *in vivo* (Solnick and Lee, 1987; Eperon *et al.*, 1988; Senecoff and Meagher, 1992). This effect is consistent with a role for hnRNP proteins as chaperones of RNA. As matchmakers, hnRNP proteins may also recruit particular factors to the pre-mRNA, probably through protein–protein interactions involving their auxiliary domains (Dreyfuss *et al.*, 1993).

Furthermore, since hnRNP proteins exhibit sequence specificity (Swanson and Dreyfuss, 1988; C.Burd *et al.*, submitted) and bind pre-mRNAs in a transcript specific fashion (Bennett *et al.*, 1992; Matunis *et al.*, 1993), it is possible that these proteins can act preferentially on particular sites in pre-mRNAs to modulate their accessibilities and influence their interactions. In this vein, it is particularly interesting that the ratio of hnRNP A1 to the splicing factor SF2/ASF has been shown to determine 5′ splice site choice

in vitro (Mayeda and Krainer, 1992). The models presented here suggest that these and other splicing factors could influence regulated splicing both by modulating the accessibilities of competing splice sites and by recruiting (or hindering the interaction of) other factors. Recently, it has been suggested that the RNA annealing activity of A1 may promote the interaction between the pre-mRNA branch point and U2 snRNA (Buvoli *et al.*, 1992) and that *in vitro* phosphorylation of hnRNP A1 can abolish its annealing activity (Cobianchi *et al.*, 1993), suggesting that this activity may be regulated *in vivo*.

An interesting characteristic of the RNA annealing activity of hnRNP proteins *in vitro* is that their activity is reduced or abolished at high protein:RNA ratios (Figure 5A); this effect has previously been noted for hnRNP A1 (Munroe and Dong, 1992). Although the reasons for this are unclear, it is possible that excess protein stabilizes the single strands, sterically blocks RNA base pairing or 'squelches' sites required for protein–protein interactions. One important consequence of this window of activity is that some RNA annealing activities may have been missed during the fractionation because of their concentrations. To circumvent this problem, each assay was repeated at least once using different concentrations of the fractions (data not shown). We feel that the conditions shown were suitable to detect the majority of the RNA annealing activities present. Another interesting implication of this phenomenon is that *in vitro* studies examining the role of hnRNP proteins in pre-mRNA processing reactions should be performed over a range of protein concentrations.

The stimulation of RNA annealing by hnRNP proteins represents a profound effect on the interactions of RNAs with each other and provides interesting insights into the functional consequences of the binding of these proteins to nascent pre-mRNAs *in vivo*. We suggest that hnRNP proteins can act to overcome some of the problems intrinsic to pre-mRNA processing by bringing about changes in the structure of pre-mRNAs and in the association of pre-mRNAs with *trans*-acting factors. Clearly, further studies will be required to understand both the ability of hnRNP proteins to promote RNA–RNA interactions, and the specific roles of these proteins during the process of mRNA formation.

Materials and methods

Preparation of RNA transcripts

RNA transcripts were synthesized by SP6- and T7-directed *in vitro* transcription reactions in the presence of [α-^{32}P]UTP (for the probe transcript only) and ^{7}mGpppG and were gel purified as described (Choi *et al.*, 1986). *Bgl*I-linearized pRSP-1-ΔIVS plasmid DNA (Konarska *et al.*, 1984) was used as the template for the 'sense' RNA. *Sma*I-linearized pAd(60)-3′ss+ DNA was used as the template for the probe transcription. This plasmid contains the sequence from 46 nt upstream to 14 nt downstream of the pRSP-1-ΔIVS encoded 3′ splice site, chemically synthesized as two complementary 76mer oligonucleotides, cloned into the *Eco*RI site of pGEM-1 (Promega). Transcript yields were quantitated optically at 260 nm and by scintillation counting.

RNA annealing assay

For RNA annealing assays the protein or fraction of interest was brought to 5 μl in 2 × RNA annealing buffer (40 mM HEPES pH 7.6, 200 mM KCl, 1 mM magnesium acetate and 1 mM dithiothreitol) and preincubated for 10 min at 30°C. For assays on purified proteins 0.1 mg/ml RNase-free BSA (Pharmacia) was included. An equal volume of a solution containing 2 nM sense RNA and 1 nM probe RNA was then added to the protein and the annealing reaction was allowed to take place for 10 min at 30°C. (The RNA solution was prepared freshly before the start of each assay by heating 4 nM and 2 nM solutions of sense and probe RNA, respectively, to 65°C for 5 min, quickly chilling on ice and mixing equal volumes of each.) RNase T1 (Boehringer-Mannheim) was then added to a final concentration of 1 U/μl and the reaction was incubated for 15 min at 37°C. An equal volume of 2 × stop solution (0.2% SDS, 1 mg/ml proteinase K, 2.5 mg/ml *E.coli* tRNA and 600 mM sodium acetate pH 5.2) was then added and the reaction was again incubated for 15 min at 37°C. Finally, the mixture was extracted once with an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol and precipitated with 2.5 vol of ethanol. RNA pellets were recovered by centrifugation, washed with 70% ethanol and resuspended in 99% formamide loading buffer. RNA fragments were resolved electrophoretically on 12% (29:1) acrylamide:bis/8 M urea gels which were then dried and exposed to film. Quantitation of the relative amounts of ssRNA and dsRNA was performed by analysis of the dried gel with a Molecular Dynamics PhosphorImager.

Fractionation of RNA annealing activities from HeLa nucleoplasm

Micrococcal nuclease-digested HeLa nucleoplasm was prepared as described (Pederson, 1974) from 8 l of HeLa S3 cells. Nucleoplasm was cleared by centrifugation at 15 000 *g* and loaded onto a 10 ml ssDNA–cellulose column (United States Biochemical) at 0.5 ml/min. The column was washed with column buffer (50 mM HEPES pH 7.6, 1 mM EDTA and 10% glycerol) at 100 mM KCl and then with 1 mg/ml heparin at 100 mM KCl in column buffer. Bound material was eluted with a linear gradient from 200 mM to 900 mM KCl and then a step to 2M KCl, in column buffer. Fractions collected were 1 ml each; 1 μl of each fraction was used for an annealing assay and 25 μl were TCA precipitated and used for SDS–PAGE analysis on 12.5% acrylamide gels as described (Dreyfuss *et al.*, 1984a), followed by silver staining (Morrissey, 1981).

Four pools from the ssDNA column, as described in the text, were then individually brought to 50 mM KCl by the addition of column buffer and loaded onto a TSK–SP column at 1.0 ml/min. The column was washed with 40 mM KCl in column buffer at 0.5 ml/min and bound proteins were eluted with a 128 min linear gradient from 40 mM to 1000 mM KCl in column buffer. Fractions (1 ml each) were collected. Half of each fraction was desalted and concentrated using Microcon-10 tubes (Amicon) for 2 h as suggested by the manufacturer. Concentrated fractions were recovered by the addition of 20 μl of column buffer containing 400 mM KCl and 0.5 mM DTT. Generally, 0.5 μl of each concentrated fraction was used for an annealing assay and 2 μl for SDS–PAGE analysis, followed by silver staining. Western blotting with the monoclonal antibodies 3G6 (against the hnRNP U protein) and 4F4 (against the hnRNP C proteins) was performed as described (Choi and Dreyfuss, 1984). Western blotting with the canine anti-hnRNP G antiserum (Soulard *et al.*, 1991; generously provided by C.-J.Larsen) was performed at 1:800 dilution. Bound antibody was detected with peroxidase-conjugated protein A followed by visualization using ECL (Amersham).

Expression and purification of hnRNP proteins

The hnRNP A1 cDNA (from pBS01, kindly provided by S.Riva) was cloned in between the *Nde*I and *Bam*HI sites of pET-11a (Studier *et al.*, 1990) to generate pET11a-A1. Two liters of transformed *E.coli* BL21(DE3) were induced as described (Studier *et al.*, 1990). After 3 h of growth cells were pelleted and frozen at −70°C. The cell pellet was thawed in lysis buffer [50 mM HEPES pH 7.6, 1 mM EDTA, 10% glycerol, 500 mM KCl, 1 mM DTT, 1 mM PMSF, 10 mM $Na_2S_2O_5$ and 1 mg/ml each of leupeptin and pepstatin A (Sigma)] and sonicated. Whole cell extract was obtained by centrifugation at 20 000 *g* for 20 min at 4°C. The cleared lysate was diluted 1:1 with column buffer and loaded onto an 80 ml DEAE–sephacel column; the flowthrough from this column, which contained the recombinant hnRNP A1, was then immediately loaded onto a 10 ml ssDNA–cellulose column. The column was washed with column buffer at 100 mM KCl followed by 1 mg/ml heparin in column buffer at 100 mM KCl. Bound proteins were then eluted with a linear gradient to 1 M KCl. Fractions containing A1, as judged by OD_{280}, were pooled, diluted to 100 mM KCl with column buffer and loaded onto an SP–TSK column at 2 ml/min. After washing with column buffer at 100 mM KCl at 0.5 ml/min, bound proteins were eluted with a linear gradient to 500 mM KCl. Fractions containing A1, as judged by SDS–PAGE (data not shown), were pooled, brought to 700 mM KCl by the addition of column buffer containing 2 M KCl, aliquoted and stored at −70°C. The final preparation was judged to be virtually homogeneous by SDS–PAGE (data not shown). A1 protein concentration was determined using the Bio-Rad protein assay reagent according to the manufacturer's instructions. From 2 l of induced *E.coli* ~11 mg of purified A1 were obtained. Preparations of the hnRNP U and C1 proteins were generously provided by M.Kiledjian and M.Görlach, respectively (M.Kiledjian, unpublished data; M.Görlach *et al.*, submitted).

Acknowledgements

We are grateful to M.Görlach and M.Kiledjian for providing purified hnRNP C1 and U, respectively; to S.Wilson and A.Kumar for a sample of A1 used in the preliminary stages of this work; to C.Larsen and M.Soulard for the anti-hnRNP G antiserum; to V.Apkarian for oligonucleotide synthesis; and to G.Daly for help with the manuscript. We thank members of our laboratory, especially S.Piñol-Roma, C.Burd, M.Görlach, M.Michael, H.Siomi and M.Matunis, for insightful discussion and critical reading of the manuscript, and S.Munroe for helpful discussions. This work was supported by the Howard Hughes Medical Institute and by grants from the National Institutes of Health.

References

Barnett,S.F., Thiery,T.A. and LeStourgeon,W.M. (1991) *Mol. Cell. Biol.*, **11**, 864–871.
Bennett,M., Piñol-Roma,S., Staknis,D., Dreyfuss,G. and Reed,R. (1992) *Mol. Cell. Biol.*, **12**, 3165–3175.
Beyer,A.L. and Osheim,Y.N. (1988) *Genes Dev.*, **2**, 754–765.
Brunel,F., Alzari,P.M., Ferrara,P. and Zakin,M.M. (1991) *Nucleic Acids Res.*, **19**, 5237–5245.
Burd,C.G., Swanson,M.S., Görlach,M. and Dreyfuss,G. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9788–9792.
Buvoli,M., Cobianchi,F. and Riva,S. (1992) *Nucleic Acids Res.*, **20**, 5017–5025.
Choi,Y.D. and Dreyfuss,G. (1984) *J. Cell Biol.*, **99**, 1997–2004.
Choi,Y.D., Grabowski,P.J., Sharp,P.A. and Dreyfuss,G. (1986) *Science*, **231**, 1534–1539.
Clouet D'Orval,B., D'Aubenton Carafa,Y., Sirand-Pugnet,P., Gallego,M., Brody,E. and Marie,J. (1991) *Science*, **252**, 1823–1828.
Cobianchi,F., Karpel,R.L., Williams,K.R., Notario,V. and Wilson,S.H. (1988) *J. Biol. Chem.*, **263**, 1063–1071.
Cobianchi,F., Calvio,C., Buvoli,M. and Riva,S. (1993) *Nucleic Acids Res.*, **21**, 949–955.
Craig,E.A. (1993) *Science*, **260**, 1902–1903.
Dreyfuss,G., Adam,S.A. and Choi,Y.D. (1984a) *Mol. Cell. Biol.*, **4**, 415–423.
Dreyfuss,G., Choi,Y.D. and Adam,S.A. (1984b) *Mol. Cell. Biol.*, **4**, 1104–1114.
Dreyfuss,G., Matunis,M.J., Piñol-Roma,S. and Burd,C.G. (1993) *Annu. Rev. Biochem.*, **62**, 289–321.
Eperon,L.P., Graham,I.R., Griffiths,A.D. and Eperon,I.C. (1988) *Cell*, **54**, 393–401.
Gething,M.-J. and Sambrook,J. (1992) *Nature*, **355**, 33–45.
Ghetti,A., Piñol-Roma,S., Michael,W.M., Morandi,C. and Dreyfuss,G. (1992) *Nucleic Acids Res.*, **20**, 3671–3678.
Ghisolfi,L., Joseph,G., Amalric,F. and Erard,M. (1992) *J. Biol. Chem.*, **267**, 2955–2959.
Gil,A., Sharp,P.A., Jamison,S.F. and Garcia-Blanco,M.A. (1991) *Genes Dev.*, **5**, 1224–1236.
Goguel,V. and Rosbash,M. (1993) *Cell*, **72**, 893–901.
Görlach,M., Wittekind,M., Beckman,R.A., Mueller,L. and Dreyfuss,G. (1992) *EMBO J.*, **11**, 3289–3295.
Görlach,M., Burd,C.G., Portman,D.S. and Dreyfuss,G. (1993) *Mol. Biol. Rep.*, **18**, 73–78.
Green,M.R. (1991) *Annu. Rev. Cell Biol.*, **7**, 559–599.
Guthrie,C. (1991) *Science*, **253**, 157–163.
Hawkins,J.D. (1988) *Nucleic Acids Res.*, **16**, 9893–9908.
Hoffman,D.W., Query,C.C., Golden,B.L., White,S.W. and Keene,J.D. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2495–2499.
Karpel,R.L., Miller,N.S. and Fresco,J.R. (1982) *Biochemistry*, **21**, 2102–2108.
Konarska,M.M., Padgett,R.A. and Sharp,P.A. (1984) *Cell*, **38**, 731–736.
Konarska,M.M., Padgett,R.A. and Sharp,P.A. (1985) *Cell*, **42**, 165–171.
Krainer,A.R., Conway,G.C. and Kozak,D. (1990) *Genes Dev.*, **4**, 1158–1171.
Kumar,A. and Wilson,S.H. (1990) *Biochemistry*, **29**, 10717–10722.
Lamm,G.M. and Lamond,A.I. (1993) *Biochim. Biophys. Acta*, **1173**, 247–265.
Lee,C.-G., Zamore,P.D., Green,M.R. and Hurwitz,J. (1993) *J. Biol. Chem.*, **268**, 13472–13478.
Libri,D., Piseri,A. and Fiszman,M.Y. (1991) *Science*, **252**, 1842–1845.
Matunis,M.J., Michael,W.M. and Dreyfuss,G. (1992) *Mol. Cell. Biol.*, **12**, 164–171.
Matunis,E.L., Matunis,M.J. and Dreyfuss,G. (1993) *J. Cell Biol.*, **121**, 219–228.
Mayeda,A. and Krainer,A.R. (1992) *Cell*, **68**, 365–375.
Miyatani,S., Copeland,N.G., Gilbert,D.J., Jenkins,N.A. and Takeichi,M. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 8443–8447.
Moore,M.J., Query,C.C. and Sharp,P.A. (1993) In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.
Morrissey,J.H. (1981) *Anal. Biochem.*, **117**, 307–310.
Munroe,S.H. (1988) *EMBO J.*, **7**, 2523–2532.
Munroe,S.H. and Dong,X. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 895–899.
Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P.R. (1990) *Nature*, **348**, 515–520.
Oberosler,P., Hloch,P., Ramsperger,U. and Stahl,H. (1993) *EMBO J.*, **12**, 2389–2396.
Patton,J.G., Mayer,S.A., Tempst,P. and Nadal-Ginard,B. (1991) *Genes Dev.*, **5**, 1237–1251.
Pederson,T. (1974) *J. Mol. Biol.*, **33**, 251–263.
Piñol-Roma,S., Choi,Y.D., Matunis,M.J. and Dreyfuss,G. (1988) *Genes Dev.*, **2**, 215–227.
Pontius,B.W. (1993) *Trends Biochem. Sci.*, **18**, 181–186.
Pontius,B.W. and Berg,P. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 8403–8407.
Pontius,B.W. and Berg,P. (1992) *J. Biol. Chem.*, **267**, 13815–13818.
Racker,E. (1992) *Curr. Topics Cell. Regul.*, **33**, 127–143.
Sapp,M., Knippers,R. and Richter,A. (1986) *Nucleic Acids Res.*, **14**, 6803–6820.
Senecoff,J.F. and Meagher,R.B. (1992) *Plant Mol. Biol.*, **18**, 219–234.
Solnick,D. and Lee,S.I. (1987) *Mol. Cell. Biol.*, **7**, 3194–3198.
Soulard,M., Barque,J.P., Della Valle,V., Hernandez-Verdun,D., Masson,C., Danon,F. and Larsen,C.-J. (1991) *Exp. Cell Res.*, **193**, 59–71.
Soulard,M., Della Valle,V., Siomi,M.C., Piñol-Roma,S., Codogno,P., Bauvy,C., Belli,M., LaCroix,J.-C., Dreyfuss,G. and Larsen,C.-J. (1993) *Nucleic Acids Res.*, **21**, 4210–4217.
Steitz,J.A. (1992) *Science*, **257**, 888–889.
Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
Swanson,M.S. and Dreyfuss,G. (1988) *EMBO J.*, **7**, 3519–3529.
Takagaki,Y., MacDonald,C.C., Shenk,T. and Manley,J.L. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 1403–1407.
Wittekind,M., Görlach,M., Friedrichs,M., Dreyfuss,G. and Mueller,L. (1992) *Biochemistry*, **31**, 6254–6265.

Received on September 16, 1993