The EMBO Journal vol.9 no.4 pp.1229–1235, 1990

Alternative splicing in the human gene for the core protein A1 generates another hnRNP protein

Massimo Buvoli, Fabio Cobianchi, Marco G.Bestagno, Angelo Mangiarotti, Maria Teresa Bassi, Giuseppe Biamonti and Silvano Riva

Istituto di Genetica Biochimica ed Evoluzionistica CNR, Via Abbiategrasso 207, 27100 Pavia, Italy

Communicated by A.Falaschi

The human hnRNP core protein A1 (34 kd) is encoded by a 4.6 kb gene split into 10 exons. Here we show that the A1 gene can be differentially spliced by the addition of an extra exon. The new transcript encodes a minor protein of the hnRNP complex, here defined $A1^B$ protein, with a calculated mol. wt of 38 kd, that coincides with a protein previously designated as B2 by some authors. *In vitro* translation of the mRNAs selected by hybridization with A1 cDNA produced two proteins of 34 and 38 kd; Northern blot analysis of poly(A)$^+$ RNA from HeLa cells revealed that the abundance of the $A1^B$ mRNA was ~5% that of A1. The $A1^B$ protein was detected by Western blotting with an anti-A1 monoclonal antibody both in enriched preparations of basic hnRNP proteins and in 40S hnRNP particles. The $A1^B$ protein exhibits a significantly higher affinity than A1 for ssDNA. The recombinant $A1^B$ protein, expressed in *Escherichia coli*, shows the same electrophoretic mobility and charge as the cellular one.

Key words: alternative splicing/hnRNP proteins

Introduction

Heterogeneous nuclear RNAs (hnRNAs) are associated in the nucleus with specific proteins to form ribonucleoprotein complexes commonly called hnRNP particles (Dreyfuss, 1986). These proteins are attracting increasing attention after some recent experiments that suggest their involvement in hnRNA metabolism and particularly in the splicing process (Choi *et al.*, 1986; Sierakowska *et al.*, 1986; Swanson and Dreyfuss, 1988).

One of the most debated questions in the study of hnRNP complexes centres around their protein composition, which still remains poorly defined. The development of selective UV cross-linking techniques and the use of specific monoclonal antibodies has gradually allowed a more stringent definition of genuine hnRNP proteins, i.e. of the proteins that are bound to RNA in the nucleus (Mayrand *et al.*, 1981; Dreyfuss, 1986). By such criteria >20 polypeptides appear now to be 'bona fide' hnRNP components (Piñol-Roma *et al.*, 1988). Among these, six polypeptides commonly defined as hnRNP 'core' proteins (A1, A2, B1, B2, C1 and C2; mol. wt 34–42 kd) have been more thoroughly studied by several groups in the last 10 years as reviewed by Dreyfuss (1986). Early studies on the physicochemical properties and on the amino acid composition of these proteins led to their subdivision into sets of closely related polypeptides. Other experimental evidence indicated that the most basic proteins (A and B) are structurally and antigenically related (Leser *et al.*, 1984) and that the same is true for the C1/C2 doublet (Dreyfuss *et al.*, 1984).

These results prompted us to investigate the gene structure of this family of proteins to elucidate the genetic basis of such structural relationships. To date, only the cDNAs of human proteins A1 and C1 have been isolated and sequenced (Swanson *et al.*, 1987; Buvoli *et al.*, 1988). More recently the isolation in our laboratory of an active gene for protein A1 has led to the discovery of a differential splicing event in the A1 pre-mRNA (Biamonti *et al.*, 1989). In this paper we show that the product of alternative splicing is a minor component of the hnRNP complex, previously described as protein B2 by some authors (Wilk *et al.*, 1985). On the basis of this correlation the new protein is here named $A1^B$. Our results support the hypothesis that the relatedness between and among hnRNP proteins derives at least in part from alternative splicing of common pre-mRNAs.

Results

Alternative splicing in the A1 gene

The occurrence of alternative splicing in an active gene for the human hnRNP protein A1 was reported in a previous paper (Biamonti *et al.*, 1989). Indeed, we found that in a Northern type hybridization to poly(A)$^+$ RNA from HeLa cells an intron probe recognized two prominent mRNA species in addition to the expected pre-mRNA (data not shown). The two mRNA species were 100–200 nt longer than the two previously reported A1 mRNAs (Buvoli *et al.*, 1988), suggesting the presence of A1-specific mRNAs with an additional exon. In order to confirm the occurrence of differentially spliced A1-specific mRNAs and to define the borders of the putative additional exon, we screened a HeLa cell cDNA library in λgt11 with a probe corresponding to the *Sph*I–*Bam*HI fragment of intron no. 7 of the A1 gene (Biamonti *et al.*, 1989). Screening of 2×10^5 plaques yielded four positive clones (pRP16–pRP19), whose cDNA inserts were subcloned in pUC19 and sequenced on both strands.

Alignment with the previously published A1 cDNA sequence (pRP15; Buvoli *et al.*, 1988; EMBL data base accession no. X06747) revealed in all four cases the same insertion of 156 nt at position 836, whereas the overlapping regions had strictly identical sequences. In Figure 1A the longest (nearly full length) cDNA (pRP18) is shown aligned to pRP15: it is 1441 nt long, starting at nt 94 (inside the third codon) and ending at nt 1378 of pRP15, and contains an extra exon corresponding to nt 3078–3228 in the 7th intron of the A1 gene. It should also be noticed that the extra exon (here named exon 7 bis) is bordered by the canonical

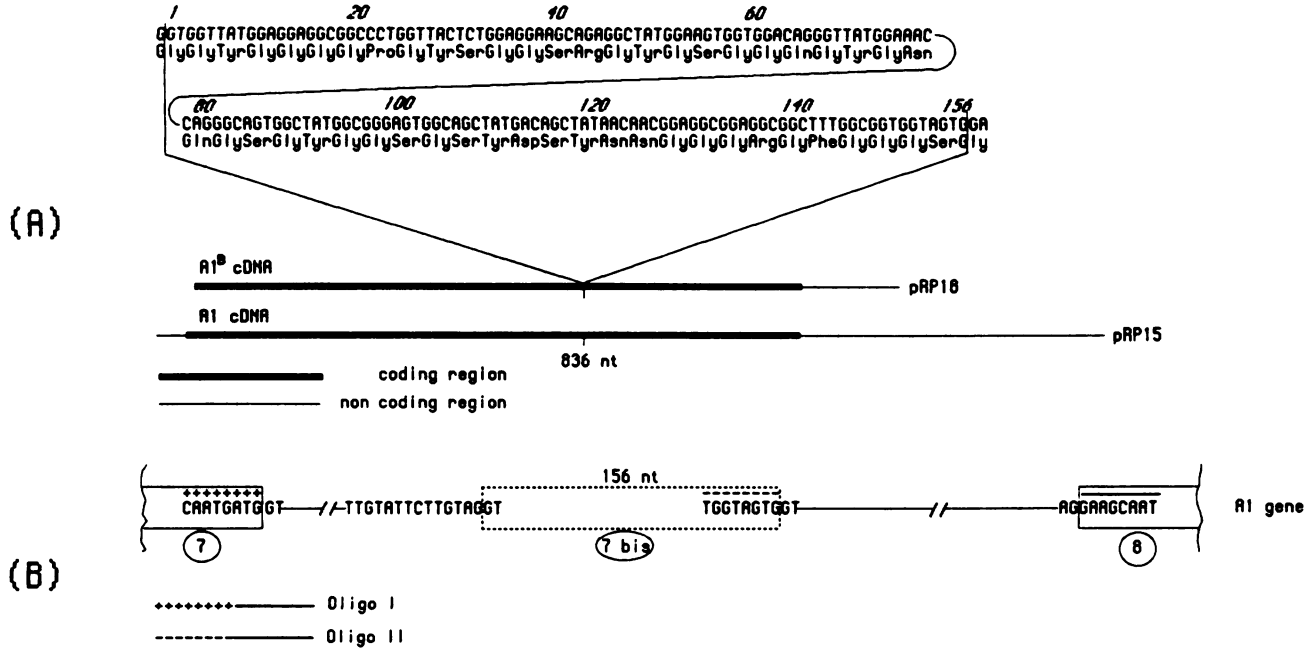

Fig. 1. (**A**) Schematic representation of the $A1^B$ cDNA (pRP18) in comparison with A1 cDNA (pRP15). The sequence of the extra exon (exon 7 bis, 156 nucleotides, 52 amino acids) is outlined. (**B**) Localization of exon 7 bis in A1 gene (Biamonti *et al.*, 1989). Exon–intron junction and splicing signals are shown. The two 'split-sequence' oligonucleotides (16 mers) used in Figure 2A, complementary to exon junction 7–8 and 7 bis–8 respectively, are depicted.

exon/intron junctions and splicing signals (see Figure 1B). When translated into protein, addition of exon 7 bis leads to the insertion of 52 amino acids (27 of which are glycines) at residue 251 in the glycine rich C-terminal domain of protein A1 (Buvoli *et al.*, 1988), generating a new protein hereinafter named $A1^B$.

Quantification of the differential splicing event

When the full length A1 cDNA (pRP15) was used to probe a Northern blot of poly(A)$^+$ RNA from HeLa cells, two strong bands at 1900 and 1500 nt were detected, that were shown to be derived from the same pre-mRNA by differential polyadenylation (Buvoli *et al.*, 1988). However, this experiment was not suitable for detecting hybridization to the alternative splicing products described in the previous section since their bands are weak and very close in length to the two main A1 mRNA species. Therefore, in order to estimate the ratio between the two splicing products in HeLa cells, an experiment was devised whereby each of the two splicing products is separately detected by a specific oligonucleotide probe. Two 'split-sequence' oligonucleotides (16 mers) were synthesized complementary to the junctions (8+8 nt) between exons 7 and 8, or 7 bis and 8 respectively (see scheme in Figure 1B). The two oligos (I and II) were labelled with ^{32}P at the same specific activity and used to probe the same Northern blot of poly(A)$^+$ mRNA from HeLa cells under stringency conditions that ensure specific hybridization to the cognate RNAs (see Materials and methods). As shown in Figure 2A both oligonucleotide probes detected two mRNA species probably due to the previously described differential polyadenylation. As expected from the presence of the additional exon, the mRNAs recognized by oligo II are slightly longer than those recognized by oligo I, although much weaker. The relative intensity of the two doublets of bands after equivalent times

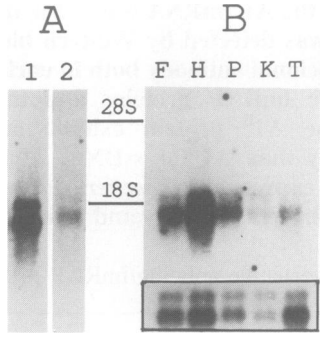

Fig. 2. Quantification of the alternative splicing in different human tissues. (**A**) Relative abundance of A1 and $A1^B$ mRNAs in HeLa cells as determined by Northern blot hybridization to poly(A)$^+$ RNA with the 'split-sequence' oligonucleotide probes shown in Figure 1: oligo I (**lane 1**); oligo II (**lane 2**). (**B**) Quantification in the $A1^B$ mRNAs in comparison with total A1 mRNA in different human tissues as determined by Northern blot hybridization to poly(A)$^+$ RNA with an exon 7 bis specific probe and with an A1 cDNA (pRP15) probe (boxed panel). F, SV40 transformed fibroblasts; H, HeLa cells; P, placenta; K, kidney; T, thymus. The positions of rRNAs (18S and 28S) are indicated.

of exposure allows an approximate estimation of the relative abundance of the splicing products (A1 versus $A1^B$) in the order of 20:1. In a previous paper (Buvoli *et al.*, 1988) we estimated by Northern type hybridization the amount of A1-specific mRNA in five human tissues: SV40 transformed fibroblasts, HeLa cells, placenta, kidney and thymus. In order to quantify the differential splicing event in these five tissues a Northern blot of poly(A)$^+$ RNAs was hybridized first with the probe specific for exon 7 bis and, after washing, with A1 cDNA. The results shown in Figure 2B demonstrate that differentially spliced mRNAs are present in all five tissues; the estimation of the A1 versus $A1^B$ ratio, based on

exposure times, is ~20:1 in all tissues, with the noticeable exception of thymus, where the ratio is much lower. This finding raises the interesting possibility of a cellular specificity in the regulation of the splicing pattern.

Hybrid selected A1-specific poly(A)$^+$ RNA is translated into two proteins

In order to assess the biological relevance of alternative splicing in the A1 gene we addressed the question as to whether the A1^B mRNA can be translated. In a preliminary experiment, *in vitro* translation was performed on the two differentially spliced A1-specific mRNAs transcribed *in vitro* with T7 polymerase from the corresponding cDNAs. In order to produce full length mRNA from pRP18 cDNA (that lacks the 5′ end non-translated region and the first three codons) this cDNA was cut at the *Bal*I site and ligated to the *Eco*RI−*Bal*I fragment of pRP15 so as to supply the 5′ missing portion of pRP18.

As depicted in Figure 3 (lanes 2 and 3), *in vitro* translation of the two RNAs yields two polypeptides with the size expected for the A1 and and A1^B proteins (34 and 38 kd respectively). Both proteins can be separately immunoprecipitated by an anti-A1 monoclonal antibody (4B10, see Materials and methods) thus indicating that the A1^B protein also contains the reactive epitope (not shown). The two translation products, mixed and immunoprecipitated, are show in Figure 3, lane 4.

To demonstrate that the cellular A1^B mRNA is translatable, a hybrid selected translation reaction was performed. The full length A1 cDNA (pRP15) was adsorbed to a nitrocellulose membrane and used to hybrid select HeLa poly(A)$^+$ RNA as described in Materials and methods. After elution from the membrane, the RNA was translated *in vitro* in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine. The ^{35}S-labelled polypeptides were then immunoprecipitated with the anti-A1 monoclonal antibody (4B10) and separated by SDS−PAGE. As shown in Figure 3, lane 5, two ^{35}S-labelled polypeptides were detected with exactly the same size as the two proteins described above. The relative abundance of the two proteins is ~20:1 in accordance with the mRNA estimations reported in the previous section. The largest protein (38 kd) has the size expected from the insertion of 52 amino acids into protein A1 (34 kd).

These data further confirm that HeLa cells contain two translatable A1-specific mRNAs produced by differential splicing of the same gene.

Detection of the A1^B protein in HeLa cells by Western blotting

The results of the hybrid selected translation strongly indicate that the A1^B protein is in effect present in HeLa cells, while at the same time raising the question of its identification among the described hnRNP proteins.

Probably due to its low abundance the A1^B protein cannot be detected in crude nuclear extracts with current Western blotting techniques either with an anti-A1 serum or with the monoclonal antibody specific for protein A1 (4B10) (data not shown). In a more thorough identification effort we decided to look for this protein (which is assumed to copurify with A1) in an enriched preparation of hnRNP basic proteins, prepared as described in Materials and methods.

The protein preparation was separated by SDS−PAGE, blotted onto nitrocellulose and probed with the 4B10 monoclonal antibody, as described in Materials and methods.

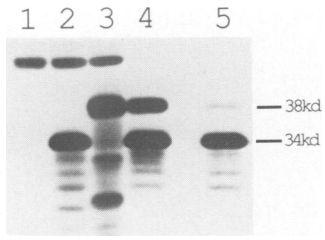

Fig. 3. *In vitro* translation and hybrid selection experiments. **Lanes 1−4**: *in vitro* translation of the A1 and A1^B T7 transcripts in a rabbit reticulocyte lysate. **Lane 1**, −RNA; **lane 2**, +A1 mRNA; **lane 3**, +A1^B mRNA; **lane 4**, the ^{35}S-labelled translation products of lanes 2 and 3 mixed and immunoprecipitated with the anti-A1 monoclonal antibody 4B10. **Lane 5**, hybrid selection with A1 cDNA and *in vitro* translation. HeLa cells poly(A)$^+$ RNA was hybridized to filter bound pRP15 cDNA and translated *in vitro* in a rabbit reticulocyte lysate after elution. The ^{35}S-labelled translation products were immunoprecipitated with 4B10 antibody. Bands below 34 kd in lanes 4 and 5 are produced by limited proteolysis.

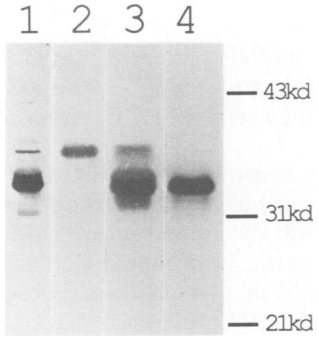

Fig. 4. Detection of the A1^B protein by SDS−PAGE and Western blotting with the anti-A1 monoclonal antibody 4B10. **Lane 1**, 20 μg of an enriched preparation of basic hnRNP proteins (see text); **lane 3**, 30 μg of purified 40S hnRNP particles; **lanes 2** and **4**, 2 and 3 μg of recombinant A1^B and A1 proteins respectively (see text).

As shown in Figure 4, lane 1, antibody staining reveals a band at the mol. wt expected for the A1^B protein (38 kd) in addition to the A1 band. A faint band migrating in the position expected for the A1^B protein is detectable also in SDS−PAGE of 40S hnRNP particles prepared as previously described (Pandolfo *et al.*, 1987), provided the amount of particles examined is sufficiently large (>25 μg of protein) (see Figure 4, lane 3). These experiments also demonstrate that the size of the cellular A1^B protein and its abundance relative to A1 are in agreement with those obtained in the *in vitro* translation experiment described above.

The purified preparation of basic hnRNP was also fractionated by two-dimensional (2D) non-equilibrium pH gradient polyacrylamide gel electrophoresis (NEPHGE) (Figure 5). The pattern of spots obtained is similar to that described by Wilk *et al.* (1985) and Piñol-Roma *et al.* (1988). Staining of the 2D blots with 4B10 monoclonal antibody revealed that, besides the A1 protein (34 kd; pI 9.5), an additional protein at 38 kd and pI 9.2 contains a reactive epitope (see insert in Figure 5), thus allowing its identification with the A1^B protein. These findings pose the question of the identity of A1^B protein with some of the already known hnRNP proteins. In fact a survey of the literature data revealed that the A1^B protein most likely

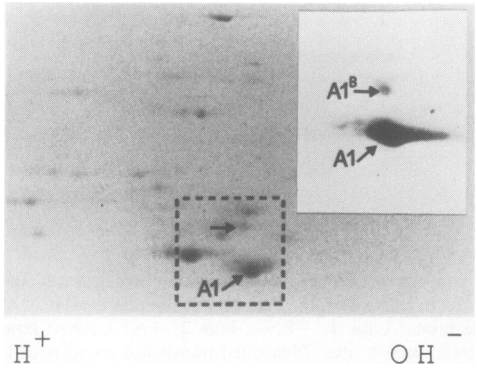

Fig. 5. Detection of the $A1^B$ protein in 2D gel electrophoresis (NEPHGE) of purified basic hnRNP proteins. 25 μg of proteins were fractionated, blotted and stained as described in Materials and methods. The arrows indicate the spots of A1 and $A1^B$ proteins after Coomassie staining of the gel. The insert in the top right corner shows the squared region of the blot after staining with the anti-A1 monoclonal antibody 4B10.

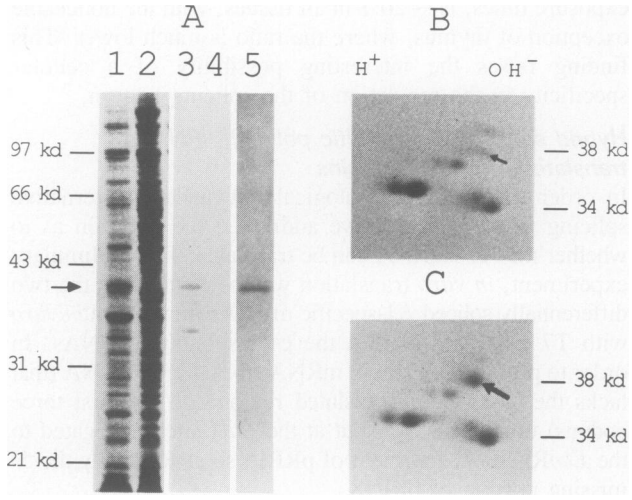

Fig. 6. Overexpression of the recombinant $A1^B$ protein in *E.coli*. Plasmid construction, cell growth and induction were as described in Materials and methods. (**A**) **Lanes 1** and **2**, Coomassie stained SDS−PAGE of induced (**lane 1**) and non induced (**lane 2**) bacterial extracts. The arrow on the left side indicates the position of the induced protein band. **Lanes 3** and **4**, Western blot of the gels in lanes 1 and 2 stained with 4B10 monoclonal antibody. **Lane 5**, SDS−PAGE and Coomassie staining of the recombinant protein $A1^B$ (2 μg) purified as described in Materials and methods. (**B**) and (**C**) Comigration of the recombinant protein $A1^B$ with the cellular one in 2D gel electrophoresis. (**B**) Coomassie stained 2D gel of hnRNP proteins as in Figure 5. (**C**) Same experiment as in (B) after addition of 4 μg of purified recombinant $A1^B$ protein. The reinforced spot produced by the recombinant protein is indicated by the thick arrow.

coincides with hnRNP protein B2 from HeLa cells as described by Wilk *et al.* (1985), corresponding to protein no. 306 of the HeLa cell NEPHGE protein catalogue (Celis *et al.*, 1988).

Moreover, an unnamed spot attributable to $A1^B$ can be seen in 2D gels of hnRNP particles immunopurified with a monoclonal antibody directed against hnRNP core protein C (Piñol-Roma *et al.*, 1988). It can therefore be concluded that the $A1^B$ protein was previously observed by other authors among hnRNP proteins.

Overexpression of the $A1^B$ cDNA in E.coli

For an initial investigation of the structure−function relationship of the $A1^B$ protein in comparison with A1, we undertook its expression in *Escherichia coli* in order to obtain appreciable amounts of unfused protein. Cloning and expression of the $A1^B$ cDNA was performed essentially as previously described for the case of protein A1 (Cobianchi *et al.*, 1988). The $A1^B$ protein was overexpressed from the open reading frame of the pRP18 cDNA, arranged as described in Materials and methods, using the P_L promoter based bacterial expression system pRC23 (Crowl *et al.*, 1985). The resulting plasmid construct maintains unaltered the open reading frame at the 5′ end and positions the ribosome binding site of the vector 8 bp upstream of the initiation codon in the cDNA as required for optimal expression in *E.coli*.

After induction with temperature, cells transformed with the recombinant expression plasmid produced large amounts of a new 38 kd protein. Immunoblotting experiments with anti-A1 monoclonal antibody (4B10) revealed that the 38 kd protein contained a reactive epitope (see Figure 6A). Using the purification procedure previously described for protein A1 (Cobianchi *et al.*, 1988), the $A1^B$ protein could be purified easily to homogeneity from induced bacterial extracts (~5 mg protein from 1 g of pelleted cells).

Physico-chemical properties of the $A1^B$ protein

As shown in Figure 4 the recombinant $A1^B$ protein has exactly the same mobility in SDS−PAGE as the cellular one. Moreover, as demonstrated in Figure 6 (panels B and C), it comigrates with the cellular protein also in NEPHGE (mol. wt 38 kd, pI 9.2). These results indicate that the two proteins have a similar structure and argue against the existence of major perturbing chemical modifications in the cellular species. In the course of these experiments we could also verify that the same is true for A1 protein (data not shown).

As deduced from its cDNA sequence, the $A1^B$ protein (372 amino acids) contains 52 additional amino acids in the glycine rich C-terminal domain. The overall amino acid composition of the extra exon closely matches that of the surrounding regions, both in the very high content of glycine residues and in the presence of a sequence motif repeated through the whole C-terminal portion, consisting of aromatic residues (Phe or Tyr) separated by five to eight residues, mostly serines and glycines (Wilson *et al.*, 1987; Cobianchi *et al.*, 1988).

It was proposed that such a periodicity could provide the basis for an interaction of the C-terminal domain with nucleic acids (Wilson *et al.*, 1987). In fact, although the binding of A1 protein to RNA is primarily due to an 'RNA binding consensus' present in the N-terminal portion of the protein (Adam *et al.*, 1986; Merrill *et al.*, 1988) a role of C-terminal domain in nucleic acid binding has recently gained some experimental support (Cobianchi *et al.*, 1988, 1989).

On the basis of these considerations we addressed the question of possible differences between proteins A1 and $A1^B$ in their nucleic acid binding properties. To test this possibility a purified preparation of the basic hnRNP proteins (see Materials and methods) was applied onto a ssDNA−cellulose column and eluted with a NaCl concentration gradient. As shown in Figure 7, the $A1^B$ protein elutes from the column at a significantly higher salt concentration with respect to A1, indicating a higher affinity for ssDNA.

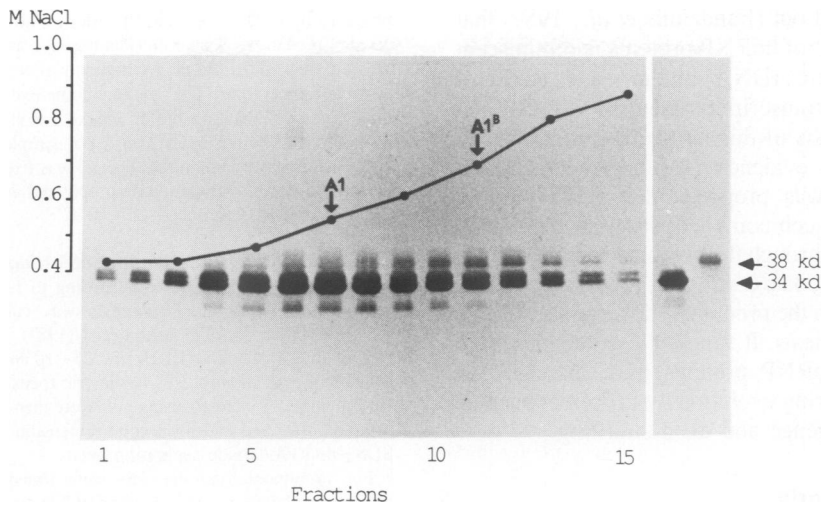

Fig. 7. Differential elution of A1 and $A1^B$ proteins from ssDNA−cellulose. 1 mg of basic hnRNP proteins was absorbed onto a ssDNA−cellulose column (5 ml) and eluted with a NaCl concentration gradient (0.4−1.0 M). 100 μl aliquots of each fraction (2.5 ml), corresponding to 1.7−8.0 μg of total proteins, were submitted to SDS−PAGE, blotted onto nitrocellulose and stained with 4B10 monoclonal antibody. The molarity of peak elution of the two proteins is indicated. The last two lanes contain 4 and 2 μg of purified recombinant A1 and $A1^B$ respectively. The band below 34 kd is a limited proteolysis product.

This result supports the hypothesis of a role of the glycine rich C-terminal domain of the protein in the binding to nucleic acid.

Discussion

In this paper we demonstrate that a gene encoding the human hnRNP protein A1 is differentially spliced to produce a second protein of the hnRNP complex. Such a protein, here defined as $A1^B$, contains an extra exon (156 nt; 52 amino acids) in the C-terminal portion of the molecule. $A1^B$ protein is a minor component of the complex (~5% of A1 protein in HeLa cells) and comigrates in 2D gels with the protein B2 as defined by Wilk *et al.* (1985). $A1^B$ can also be identified with protein no. 306 of the HeLa cell NEPHGE protein catalogue (Celis *et al.*, 1988). A spot in the position of $A1^B$ protein is also visible in 2D gels of other authors (Piñol-Roma *et al.*, 1988) which, however, do not classify this protein as the authors mentioned above. Such a discrepancy evidences once again the need for a better classification of hnRNP proteins. If, as it seems likely, other cases of hnRNP proteins related through alternative splicing were to be found, a more rational nomenclature could be adopted for hnRNP proteins.

The splicing event that gives rise to the $A1^B$ protein seems to occur with similar frequency in a number of different human tissues. Only in thymus is such frequency significantly lower, suggesting the possibility of a cell type dependent regulation; however, such a hypothesis still remains to be demonstrated.

Preliminary experiments (data not shown) indicated that this alternative splicing event occurs also in the A1 gene of rodents, pointing to a relevant role for $A1^B$ protein in nuclear RNA metabolism. This finding, together with the discovery of $A1^B$-specific pseudogenes in the human genome (paper in preparation), suggests that alternative splicing in the A1 gene is a rather ancient event in evolution.

In addressing the question of the role of the $A1^B$ protein in the hnRNP structure and/or function two points should be considered. The first concerns the structure of the $A1^B$ protein in comparison to that of protein A1; the second has to do with their relative abundance. The $A1^B$ protein contains an extra exon that renders the glycine rich C-terminal domain significantly larger (177 versus 125 amino acids) than the corresponding domain in protein A1. The result shown in Figure 7 strongly indicates that alterations in the C-terminal domain of the protein change its affinity for ssDNA and probably also for RNA since it was shown both for A1 and for other proteins that binding to ssDNA is a faithful model for RNA binding (Merrill *et al.*, 1988; Michel and Zinder, 1989). Several authors have proposed an involvement of the C-terminal domain in protein−protein interactions on the basis of experimental evidence (Cobianchi *et al.*, 1988; Bandziulis *et al.*, 1989) and in analogy with other classes of proteins (Steinert *et al.*, 1985). If this were to be confirmed, it is conceivable that the $A1^B$ protein might differ from A1 also with regard to this parameter.

It is tempting, therefore, to speculate that the $A1^B$ protein is different from A1 in its capacity to interact both with RNA and with the other proteins of the hnRNP particles. As a consequence of such differences their participation in particle assembly could generate variability (polymorphism) of hnRNP structure. In effect, when one considers the stoichiometry of the 'core proteins' in hnRNP particles (three copies of A1 per particle; Lothstein *et al.*, 1985) and the A1 versus $A1^B$ ratio, it is evident that the latter protein can be present only in a subset of particles. Moreover it should be noted that polymorphism of hnRNP could also be generated by other mechanisms, as indicated by the observation that the relative content of the A1 protein in 40S particles varies with the state of cellular proliferation (LeStourgeon *et al.*, 1978; Celis *et al.*, 1986). It is important to outline that polymorphism in hnRNP structure could have a wider biological significance in that it could reflect a functional specialization in hnRNA metabolism. This possibility is particularly appealing in the light of some recent studies that indicate a binding preference of many hnRNP proteins for specific pre-mRNA sequences and a possible regulatory role of hnRNP complexes in pre-mRNA processing (Swanson and Dreyfuss, 1988). Bearing this in

mind it has been pointed out (Bandziulis *et al.*, 1989) that the two-domain structure of hnRNP proteins is reminiscent of the modular structure (DNA binding and auxiliary domains) of eukaryotic transcription factors (Mitchell and Tjian, 1989). On the basis of this structural similarity and of some recent genetic evidence (Amrein *et al.*, 1988; Bell *et al.*, 1988), it was proposed that RNA binding proteins might regulate such complex processes like intron recognition and splicing through the combined action of their variable domains. In this context our finding of a mechanism that generates diversity in the protein A1 C-terminal domain is in line with this hypothesis. It is possible in fact that, like transcription factors, hnRNP proteins have increased in number and diversity during evolution by processes such as gene duplication, divergence and exon shuffling.

Materials and methods

Isolation and characterization of A1^B cDNA clones

A human HeLa cell cDNA library in λgt11 (HL 1022, Clontech, Palo Alto, CA) was screened by standard procedures with a probe corresponding to the 550 bp of *Sph*I–*Bam*HI fragment in the seventh intron of A1 gene (Biamonti *et al.*, 1989). Hybridization conditions were 42°C in 4 × SSC, 4 × Denhardt's solution, 0.2% SDS, 100 μg/ml yeast tRNA, 25 mM NaPPi, 50% formamide, followed by washing at 68°C with 0.2 × SSC, 0.1% SDS for 30 min. DNA was ^{32}P-labelled by standard procedures. The inserts of four positive plaque purified clones were subcloned in pUC19 (Vieira and Messing, 1982) and sequenced in both orientations by the dideoxy method (Sanger *et al.*, 1977). Of the four clones obtained (pRP16–pRP19), only pRP18 contained a nearly full length A1^B cDNA (see text).

Oligonucleotide synthesis

Oligonucleotide probes were synthesized on a Beckman system 1 plus DNA synthesizer, purified by polyacrylamide gel electrophoresis and ^{32}P end-labelled with T4 polynucleotide kinase.

RNA preparation and Northern blot hybridization

Total and poly(A)$^+$ RNA preparations and Northern blot analysis were performed as described (Buvoli *et al.*, 1988).

Hybridization with 'split-sequence' oligonucleotides was carried out at 38°C with oligo I and at 40°C with oligo II (see Figure 2A). High stringency washing conditions were 44°C and 46°C for 2 min in 6 × SSC respectively.

In vitro transcription and translation

The full length A1 cDNA (pRP15) was previously described (Buvoli *et al.*, 1988). The full length A1^B cDNA was constructed according to standard procedures from overlapping clones by replacing the *Bal*I–*Hin*dIII fragment of pRP15 with the same fragment of the A1^B cDNA (pRP18) in Bluescribe vector. 1 μg of each template was linearized with *Hin*dIII and transcribed with T7 polymerase according to Nielsen and Shapiro (1986) in the presence of 500 μM each of ATP, CTP and UTP, 50 μM GTP and 500 μM 5′ me-GPPPG 3′. The mRNA was phenol extracted, ethanol precipitated and resuspended in water. *In vitro* translation was performed using treated rabbit reticulocyte system and [^{35}S]methionine according to the protocols provided by the supplier (Promega Biotech).

Hybrid selection, translation and immunoprecipitation

Hybrid selection was performed according to Maniatis *et al.* (1982). 20 μg of A1 cDNA (pRP15) were immobilized on 0.4 × 0.4 cm squares of nitrocellulose after NaOH denaturation and neutralization. The membranes were then incubated at 50°C for 3 h in 60% formamide, 20 mM PIPES, pH 6.4, 0.4 M NaCl, 0.2% SDS, 0.1 mg/ml tRNA in the presence of 400 μg/ml of poly(A)$^+$ RNA preheated in the same solution at 70°C for 10 min. After hybridization, the membranes were washed at 65°C 10 times with 10 mM Tris–HCl, pH 7.6, 0.15 M NaCl, 1 mM EDTA and 0.5% SDS and twice with the same solution without SDS. The RNA was eluted from the hybrid by boiling for 1 min in 0.3 ml of 5 mM KCl, 2 mM EDTA and 0.1 mg/ml tRNA, then snap-frozen in a dry ice/ethanol bath and thawed. The RNA was extracted with phenol–chloroform, precipitated with ethanol, resuspended and used for *in vitro* translation performed as described.

The products of translation were diluted with four volumes of 0.1 M NaCl, 30 mM Tris–HCl, pH 7.5, 2 mM EDTA, 0.5% Triton X-100 and incubated for 2 h with 1 μl of anti-A1 monoclonal antibody 4B10 (kindly provided by G.Dreyfuss, Northwestern University, Evanston, IL, USA). 30 μl of protein A–Sepharose (Pharmacia; corresponding to 30 μg of protein A) were then added and the incubation was continued for an additional hour at room temperature. The immune complexes were pelleted, and washed three times with the same buffer without Triton X-100. The proteins released by boiling for 5 min in 25 μl of SDS sample buffer were analysed on a 10% SDS polyacrylamide gel. The gel was fixed in 30% methanol and 10% acetic acid treated with autoradiography enhancer (Enlightning, Du Pont), dried and autoradiographed at −70°C.

Gel electrophoresis and immunoblotting

SDS–PAGE was performed according to Laemmli (1970).

Two-dimensional electrophoresis was carried out according to the NEPHGE procedure of O'Farrell *et al.* (1977). The pH gradient was obtained by blending Ampholine (LKB) pH 3.5–10 and 9–11 at a 3:2 ratio to give a final concentration of 2%; isoelectric focusing was towards the cathode for 2 h at 600 V. The focusing gels were then incubated for 30 min in SDS sample buffer and laid for second dimension on the stacking gels of 10% SDS–polyacrylamide separating gels.

For immunoblotting the gels were transferred onto a nitrocellulose membrane according to the method of Towbin *et al.* (1979) and processed as described (Bestagno *et al.*, 1987). The membrane was subsequently incubated with monoclonal antibody 4B10, diluted 1:2000, and with alkaline phosphatase conjugated anti-mouse immunoglobulin antibodies (Promega Biotech). The substrate solution for alkaline phosphatase was as previously described (Bestagno *et al.*, 1987).

Purification of hnRNP basic proteins on ssDNA–cellulose

The previously described procedure (Pandolfo *et al.*, 1987; Cobianchi *et al.*, 1988) was followed with minor modifications. 3 g of frozen HeLa cells were resuspended in 15 ml of 0.35 M potassium phosphate buffer (KPB), pH 7.5, and sonicated six times for 10 s each with a Braun Labsonic 1510 sonicator set at 100 W. All buffers used contained 1 mM dithiothreitol (DTT), 1 mM phenylmethylsulphonyl fluoride (PMSF), 10 mM $Na_2S_2O_5$ and 1 μg/ml Pepstatin A (Sigma).

The cellular extract was centrifuged and the supernatant loaded onto a 20 ml DEAE–cellulose column, connected in series with a 5 ml ssDNA–cellulose column. The ssDNA–cellulose column was then disconnected and washed with five volumes of 0.35 M KPB and subsequently with five volumes of 1 M KPB. The column was then equilibrated with 10 volumes of 20 mM KPB and washed with five volumes of 20 mM KPB and 0.4 M NaCl. The hnRNP proteins were eluted with two volumes of 20 mM KPB and 1.2 M NaCl and the peak fractions pooled (1.5 mg total protein).

Large scale overproduction of A1^B protein in Escherichia coli

Expression of A1^B cDNA in *E.coli* was accomplished by inserting the A1^B cDNA into the expression vector pRC23 previously used for overexpression of the A1 protein (Cobianchi *et al.*, 1988). An intermediate plasmid construct was devised to obtain the complete coding region of A1^B cDNA by the addition of the first three codons of A1 cDNA (pRP15) to pRP18 cDNA. Cell growth, induction and purification of A1^B protein were as described for the recombinant A1 protein (Cobianchi *et al.*, 1988).

Acknowledgements

The authors are grateful to Drs G.Dreyfuss, J.Celis and C.Morandi for helpful discussions. The technical assistance of D.Arena is acknowledged. M.G.B. and M.T.B. are fellows of the Fondazione A.Buzzati Traverso. The work was supported in part by the Progetto Finalizzato CNR: 'Biotecnologie e Biostrumentazione'.

References

Adam,S.A., Nakagawa,T., Swanson,M.S., Woodruff,T.K. and Dreyfuss,G. (1986) *Mol. Cell. Biol.*, **6**, 2932–2943.

Amrein,H., Gorman,M. and Nothiger,R. (1988) *Cell*, **55**, 1025–1035.

Bandziulis,R.J., Swanson,M.S. and Dreyfuss,G. (1989) *Genes Dev.*, **3**, 431–437.

Bell,L.R., Maine,E.M., Schedl,P. and Kline,T.W. (1988) *Cell*, **55**, 1037–1046.

Bestagno,M., Cerino,A., Riva,S. and Astaldi-Ricotti,G.C.B. (1987) *Biochem. Biophys. Res. Commun.*, **146**, 1509–1514.

Biamonti,G., Buvoli,M., Bassi,M.T., Morandi,C., Cobianchi,F. and Riva,S. (1989) *J. Mol. Biol.*, **207**, 491–503.

Buvoli,M., Biamonti,G., Tsoulfas,P., Bassi,M.T., Ghetti,A., Riva,S. and

Morandi,C. (1988) *Nucleic Acids Res.*, **16**, 3751–3770.
Celis,J.E., Bravo,R., Arenstorf,H.P. and LeStourgeon,W.M. (1986) *FEBS Lett.*, **194**, 101–109.
Celis,J.E., Ratz,G.P., Celis,A., Madsen,P., Gesser,B., Kwee,S., Madsen,P.S., Nielsen,H.V., Yde,H., Lauridsen,J.B. and Basse,B. (1988) *Leukemia*, **2**, 561–602.
Choi,Y.D., Grabowski,P.J., Sharp,P.A. and Dreyfuss,G. (1986) *Science*, **231**, 1534–1539.
Cobianchi,F., Karpel,R.L., Williams,K.R., Notario,V. and Wilson,S.H. (1988) *J. Biol. Chem.*, **263**, 1063–1071.
Cobianchi,F., Biamonti,G., Bassi,M.T., Buvoli,M. and Riva,S. (1989) In Strauss,P. and Wilson,S.H. (eds), *The Eukaryotic Nucleus; Molecular Structure and Macromolecular Assemblies*. The Telford Press, Caldwell, NJ, in press.
Crowl,R., Seamans,G., Lomedico,P. and McAndrew,S. (1985) *Gene*, **38**, 31–38.
Dreyfuss,G. (1986) *Annu. Rev. Cell. Biol.*, **2**, 459–498.
Dreyfuss,G., Choi,Y.D. and Adam,S.A. (1984) *Mol. Cell. Biol.*, **4**, 1104–1114.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Leser,G.P., Escara-Wilke,J. and Martin,T.E. (1984) *J. Biol. Chem.*, **259**, 1827–1833.
LeStourgeon,W.M., Beyer,A.L., Christensen,M.E., Walker,B.W. and Poupore,S.M. (1978) *Cold Spring Harbor Symp. Quant. Biol.*, **42**, 885–898.
Lothstein,L., Arenstorf,H.P., Chung,S., Walker,B.W., Wooley,J.C. and LeStourgeon,W.M. (1985) *J. Cell Biol.*, **100**, 1570–1581.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Mayrand,S., Setyono,B., Greenberg,J.R. and Pederson,T. (1981) *J. Cell Biol.*, **90**, 380–384.
Merrill,B.M., Stone,K.L., Cobianchi,F., Wilson,S.H. and Williams,K.R. (1988) *J. Biol. Chem.*, **263**, 3307–3313.
Michel,B. and Zinder,N.D. (1989) *Nucleic Acids Res.*, **17**, 7333–7344.
Mitchell,P.J. and Tjian,R. (1989) *Science*, **245**, 371–378.
Nielsen,D.A. and Shapiro,D.J. (1986) *Nucleic Acids Res.*, **14**, 5936.
O'Farrell,P.Z., Goodman,H.M. and O'Farrell,P.H. (1977) *Cell*, **12**, 1133–1142.
Pandolfo,M., Valentini,O., Biamonti,G., Rossi,P. and Riva,S. (1987) *Eur. J. Biochem.*, **162**, 213–220.
Piñol-Roma,S., Choi,Y.D., Matunis,M.J. and Dreyfuss,G. (1988) *Genes Dev.*, **2**, 215–227.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Sierakowska,H., Szer,W., Furdon,P.J. and Kole,R. (1986) *Nucleic Acids Res.*, **14**, 5241–5254.
Steinert,P.M., Steven,A.C. and Roop,D.R. (1985) *Cell*, **42**, 411–419.
Swanson,M.S. and Dreyfuss,G. (1988) *EMBO J.*, **7**, 3519–3529.
Swanson,M.S., Nakagawa,T.K., LeVan,K. and Dreyfuss,G. (1987) *Mol. Cell. Biol.*, **7**, 1731–1739.
Towbin,NH., Staehelin,T. and Gordon,J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350–4354.
Vieira,J. and Messing,J. (1982) *Gene*, **19**, 259–268.
Wilk,H.E., Werr,H, Friedrich,D., Kiltz,H.H. and Schäfer,K.P. (1985) *Eur. J. Biochem.*, **146**, 71–81.
Wilson,S.H., Cobianchi,F. and Guy,H.R. (1987) In Thompson,E.B. and Papaconstantinou,J. (eds), *DNA:Protein Interactions and Gene Regulation*. University of Texas Press, pp. 129–146.

Received on November 18, 1989; revised on January 18, 1990