1996–2002 Nucleic Acids Research, 1994, Vol. 22, No. 11

Two homologous genes, originated by duplication, encode the human hnRNP proteins A2 and A1

Giuseppe Biamonti*, Matteo Ruggiu, Salvatore Saccone[1], Giuliano Della Valle[1] and Silvano Riva
Istituto di Genetica Biochimica ed Evoluzionistica del CNR, Via Abbiategrasso 207, 27100 Pavia and [1]Dipartimento di Genetica e Microbiologia 'A. Buzzati Traverso', Università di Pavia, Via Abbiategrasso 207, 27100 Pavia, Italy

Received April 5, 1994; Revised and Accepted May 11, 1994 GenBank accessions nos U09120–U09123 (incl.)

ABSTRACT

Heterogeneous nuclear ribonucleoprotein (hnRNP) A2 belongs, with A1, B1 and B2, to the basic protein subset of the hnRNP complex in mammalian cells. All these proteins share a modular structure consisting of two conserved RNA binding domains linked to less conserved Gly-rich domains (2xRBD-Gly). In the framework of our studies on the genetic basis of hnRNP proteins structure and diversity we have isolated and sequenced the A2 gene and compared it to the previously described A1 gene. The A2 gene, which exists in a single copy on Ch. 7 band p15, is split in 12 exons including an alternatively spliced 36 nt mini exon specific for the human hnRNP protein B1. In this work we show that the intron/exon organisation of the A2 gene is identical to that of the A1 gene over the entire length, indicating a common origin by gene duplication. Moreover the comparison of corresponding exons evidences significant conservation also in the apparently divergent Gly-rich domains that could define previously unenvisaged structural and/or functional motifs. The A2 gene promoter is also analysed in comparison to that of the A1 gene.

INTRODUCTION

The splicing of pre-mRNA into the mature product occurs in a multi-component complex constituted by small nuclear ribonucleoprotein particles (snRNP) and proteins such as splicing factors and hnRNP proteins. The hnRNP proteins are a family of at least 20 polypeptides whose precise role in RNA processing is still largely undefined. Many evidences indicate that hnRNPs might package hnRNA in a transcript specific manner so that the other components of splicing (splicing factors and snRNPs) recognise each transcript as a unique entity (1, 2). Moreover other experiments suggest a more specific role in splicing for some hnRNPs (hnRNP C and A1 [3–5]). Although the understanding of the mode of action of hnRNPs is of considerable interest, progress in this field is hampered by the lack of knowledge on the functional determinants of these proteins.

In the past, we concentrated our efforts on the hnRNP protein A1 by isolating its cDNA and gene (6, 7) and by performing a number of *in vitro* studies with the recombinant protein (8–10). In this work we analyse the structure and sequence of the human hnRNP protein A2 gene. hnRNP protein A2 is similar to A1 and belongs (with A1, $A1^B$, B1 and B2) to the subset of basic hnRNPs which are related in sequence (11) and characterised by a modular structure consisting of two tandemly arranged RNA binding domains (RBD) linked to a Gly-rich region at the C-terminal (2xRBD-Gly) (12). The RBDs are in all cases highly conserved while the Gly-rich domains can be very different (for example between A1 and A2). Such differences, which are conserved in mammals, suggest a role of the Gly-rich domain in the functional specialisation of each member of the group. That these domains harbour important functional determinants is also suggested by a number of experiments performed in our (8–10) and in other laboratories (13). Yet the identification of the specificity determinants is hampered by the lack of suitable assays. In fact, so far, the most successful way of identifying such determinants was through the alignment of aminoacid sequences of different RNA binding proteins looking for meaningful conservations. In this way the RBD motif (14) and the RNP-CS sequences (12) were discovered. However, with this approach no obvious conservation could be seen in the Gly-rich domains of A1 and A2 (11). In this paper we show that significant similarities in the two domains do in fact emerge when corresponding genomic sequences are compared. From previous studies (6, 11, 15) it was reasonable to assume that the basic hnRNP proteins of the A/B type are encoded by only two genes, one specific for $A1/A1^B$ (15) and the other for A2/B1 and possibly B2 (11). We reasoned that, if the two genes were originated by gene duplication, possible conserved motifs could be better revealed by the comparison of corresponding exons. In effect, the comparison of the two genes revealed that they did in fact arise by the duplication of a common ancestral gene, and not from an independent assembly of domains, since the intron/exon organisation has remained essentially the same both in the RBD and in the Gly-rich region. Interestingly, by comparing the corresponding Gly-rich exons, an overall conservation emerges greater than that detected at cDNA or

*To whom correspondence should be addressed

protein level (11) along with some important differences. These observations provide a useful guide-line for identifying important structural \ functional determinants in the Gly-rich domain and for starting to unravel the complexity of hnRNP function and specificity.

MATERIALS AND METHODS

Plasmids

pHCA2 (11), containing the full length cDNA encoding the human hnRNP protein A2, was a generous gift of G.Dreyfuss.

pEE26, pEE55 and pEE65 were obtained by subcloning, into the *Eco*RI site of pUC19 vector, the three *Eco*RI fragments contained in the λC53 phage isolated from the screening of a human liver DNA library (16).

Plasmid pA2 CAT contains a DNA insert of 500 nt corresponding to the region upstream of the ATG in the A2 gene, cloned into the *Sma*I site of pEMBL8CAT0 vector (17). The insert was obtained by the amplification of pEE26 clone with oligonucleotides A (5′ CTGCGGATTAAGCCCCTC 3′) and B (5′ CGCGGACTCAGTCGCTTC 3′) as primers. These oligonucleotides are complementary to nt −338 to −320 and to +145 to +162 respectively (+1 refers to the position of the transcription initiation site). *Taq* polymerase (Perkin-Elmer Cetus) was used according to the instructions of the manufacturer. The sequence was verified by the dideoxy-chain termination method (Sequenase kit; US Biochemical Corp.). pSH CAT and pSV2 CAT plasmids were previously described (17).

Isolation and analysis of genomic clones

The A2 specific genomic clones were isolated from λCh4A genomic library (16) of human liver DNA partially digested with *Eco*RI. A total of 5×10^5 plaques were screened with the pHCA2 cDNA probe by standard procedures (18). Hybridisation conditions: 4×SSC, 4×Denhardt's solution, 0.2% SDS, 250 μg yeast tRNA/ml, 50% formamide, ^{32}P-labelled pHCA2 cDNA (10^6 cts/min per ml) at 42°C for 18 h. Washing conditions: twice, for 30 min, at room temperature in 1% SDS, 0.2×SSC; twice, for 30 min, at 60°C in 0.2% SDS, 0.2×SSC; twice, for 30 min, at 68°C in 1% SDS, 0.2×SSC. A2 positive phages were replated and rescreened several times to ensure single plaques purity. Analytical amounts of recombinant bacteriophage were prepared by a minilysate procedure (19).

Primer extension

A 19-mer synthetic oligonucleotide (oligo C) complementary to nucleotide −83 to −101 with respect to the ATG (5′ GTCCTGGCGCTGTAGTGAG 3′) was 5′-end-labelled with polynucleotide kinase. The oligonucleotide (10^6 cts/min) was coprecipitated with 4 μg of poly(A)$^+$ RNA, prepared from HeLa cells, redissolved in 400 mM NaCl, 10 mM Pipes (pH 6.4) heated at 80°C and hybridised at 54°C for 3 h. After hybridisation, the sample was precipitated with ethanol and resuspended in 30 μl of 50 mM Tris−HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 40 mM dithiothreitol, 100 μg bovine serum albumin/ml, 500 μM NTPs, 1 unit RNasin/ml (Promega Co.) and 25 units of Moloney murine leukamia virus (M-MLV) reverse transcriptase (Boehringer Mannheim) and incubated at 44°C for 1 h. The reaction was stopped by the addition of 25 mM EDTA final concentration and precipitated with ethanol and carrier tRNA. The primer extension products were denatured and fractionated on 8% polyacrylamide/urea sequencing gel.

RNase mapping

A DNA fragment corresponding to position −67 to −358 with respect to the ATG was PCR amplified with the *Taq* polymerase (Perkin Elmer Cetus) using oligonucleotides D (5′ CATTTACGGAAGTAACGTCGG 3′) and E (5′ GTCCCTAAATTAATACGACTCACTAT*AGGGAGA*AACACGAACCGGACTCGTCC 3′) as primers. The E primer contains a promoter for T7 RNA polymerase (underlined) followed by 7 more bases present in all the class III promoters (italics). The transcript start site is bolded (20). The PCR product was then used as a template in the T7 RNA polymerase (Boehringer Mannheim) reactions in presence of α^{32}P-CTP (400 Ci/mmole) to produce the labelled antisense transcripts. The full length riboprobe was then purified on polyacrylamide gels as described (19). Hybridisation to total RNA and digestion of hybrids with RNase A and T_1 was performed as described (21). After precipitation with ethanol, total HeLa cell RNA (30 μg) was dissolved in 30 μl of hybridisation buffer (80% formamide, 40 mM Pipes (pH 6.4), 400 mM NaCl, 1 mM EDTA) containing 10^6 cts/min of probe, heated at 85°C for 10 min, and incubated at 45°C for 10 h. Following hybridisation, the sample was diluted 10 fold with 10 mM Tris−HCl (pH 7.5), 5 mM EDTA, 800 mM NaCl containing 2.5 units RNase A/ml, 20 units T_1/ml and incubated for 60 min at 30°C. The reaction, stopped by addition of 3 μl of proteinase K (10mg/ml) and 10 μl of 20% SDS, was incubated for 15 min at 37°C, extracted with phenol, and precipitated with ethanol and carrier tRNA. The sample was then dissolved, denatured and fractionated on 8% polyacrylamide/urea sequencing gel.

DNA transfections and CATase assay

HeLa cells were transfected with recombinant plasmid by the calcium phosphate precipitation technique of Graham and van der Eb (22). 5×10^5 cells/dish were plated in Dulbecco's modified Eagle's medium (DMEM, Gibco-BRL), 10% fetal calf serum (Gibco-BRL), 80 μg gentamicin/ml. After 24 h, 10 μg of recombinant plasmid and 10 μg of high molecular weight calf thymus DNA carrier (Boehringer Mannheim) were added. At 48 h after transfection cells were harvested and extracts were prepared and heated at 60°C for 10 min (23). Extracts were subjected to the CATase assay as described (24) Reactions were carried out at 37°C for 15 min with 0.5 μCi of ^{14}C chloramphenicol (Amersham) substrate.

In situ hybridisation

The genomic clone λ53C DNA was labelled by nick translation with biotin-16-dUTP (Boehringer Mannheim) and hybridised at a concentration of 5μg/ml to metaphase chromosome spreads of normal males. To suppress hybridisation of repeated sequences possibly present in the probe, unlabelled sonicated human placental DNA was added to the hybridisation mixture at a final concentration of 2.5 mg/ml. After DNA denaturation, preannealing was performed at 37°C for 30 min. Hybridisation and detection with FITC-conjugated avidin were performed as previously described (25). Chromosomes were counter-stained with propidium iodide in antifade medium and photographed. Subsequently, the chromosomes were G-banded with Wright's stain according to Chandler and Yunis (26) and rephotographed.

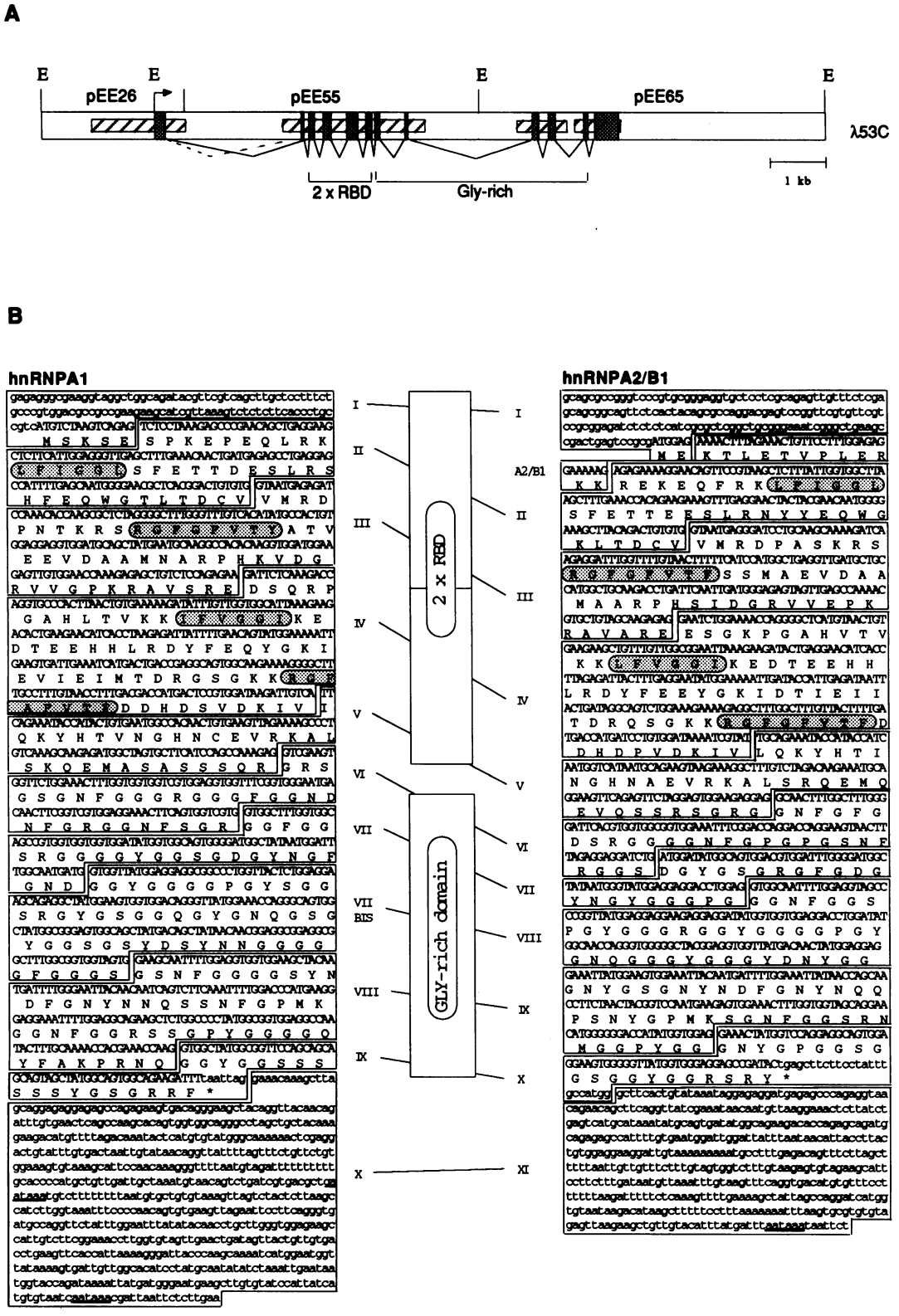

Figure 1. A: schematic representation of the λ53C clone and of the pEE subclones (E: *Eco*RI sites). Exons are indicated by grey boxes except for the alternatively spliced B1 specific exon, drawn in black. The two main protein domains (2xRBD and Gly-rich) are bracketed. Sequenced regions are indicated by the internal slashed boxed (GenBank accession numbers: U09120, U09121, and U09122, U09123 for pEE26, pEE55 and pEE65 respectively). The arrow indicates the transcription start site. **B:** comparison of A1 and A2/B1 exon structure. Exons are boxed and numbered as indicated on margin. Coding regions are in capital letters; aminoacids are in one letter code. RNP-CS 1 and 2 are evidenced in grey ovals. Corresponding exons, connected by lines, are matched to the protein domains as depicted in the centre of the figure.

RESULTS

Isolation of the human gene for the hnRNP core proteins A2/B1

On the basis of previous observations (11) it could be expected that, unlike the hnRNP A1 gene (6), the A2/B1 gene has few or no pseudogenes, since an A2/B1 cDNA probe recognised only a few restriction fragments in Southern hybridisation to HeLa cell DNA (11). Thus, we undertook the isolation of this gene by screening a λCh4A genomic library of human liver DNA, partially digested with *Eco*RI, with a cDNA probe without counterselection for pseudogenes (6). All the independently isolated clones contained the same three *Eco*RI fragments of 2.6, 5.5 and 6.5 kb. The insert of one phage (λ53C) was subcloned into the *Eco*RI site of pUC19 to produce pEE26, pEE55 and pEE65 (Fig. 1, A).

The strategy to determine the exon/intron structure of the A2/B1 gene was based on the assumption that the A1 and A2/B1 genes could share the same overall organisation, being probably originated by gene duplication. Such assumption is supported by the observation that both proteins have the same 2xRBD-Gly structure (12) with the RBD portions about 80% identical (11). Therefore, by analogy with the A1 gene structure (6) we expected that: 1) the first A2 exon encodes only few aminoacids downstream of the ATG; 2) the 2xRBD region of the gene is split into four exons; 3) one intron exactly separates the RBDs from the Gly-rich region; 4) the penultimate exon contains the stop codon as observed in A1 and in other hnRNP genes (6, 27). Furthermore, the 12 aminoacid insert near the amino terminus, found only in the B1 protein, is expected to be encoded by a small differentially spliced exon (11). In contrast, no assumption could be made on the exon/intron structure of the part of gene coding for the Gly-rich domain that was reported to be only 29% identical in the two proteins (11). On the basis of these consideration we synthesised oligonucleotides targeted to the following A2/B1 cDNA regions: 1) upstream of the ATG in the expected first exon; 2) to the A2/B1 small alternative exon; 3) to the most conserved sequence elements in the RBDs (RNP-CS 1 and 2 [12]), and 4) to the last untranslated exon, a few nucleotides downstream of the stop codon. These oligonucleotides were assigned to pEE subclones by hybridisation and used as primers in sequencing reactions. The resulting genomic sequences were matched to the A2/B1 cDNA sequence and exon/intron boundaries were determined. Intron sequences were further extended with suitable additional primers and, overall, the entire 2xRBD region of the gene was sequenced except for a gap in the first intron (see Fig. 1, A). The results confirmed that in this region the A2/B1 gene is indeed organised like the A1 gene. In detail (Fig. 1, B), 5 exons are envisaged, the first one codes for only two aminoacids and, alike in the A1 gene, the conserved sequence elements (RNP-CS 1 and 2) in the first RBD are coded by two separate exons. Vice versa, the same elements in the second RBD are encoded in both genes by a single exon. It is interesting to note that in the RBD region the exon boundaries are remarkably conserved at both the nucleotide and aminoacid

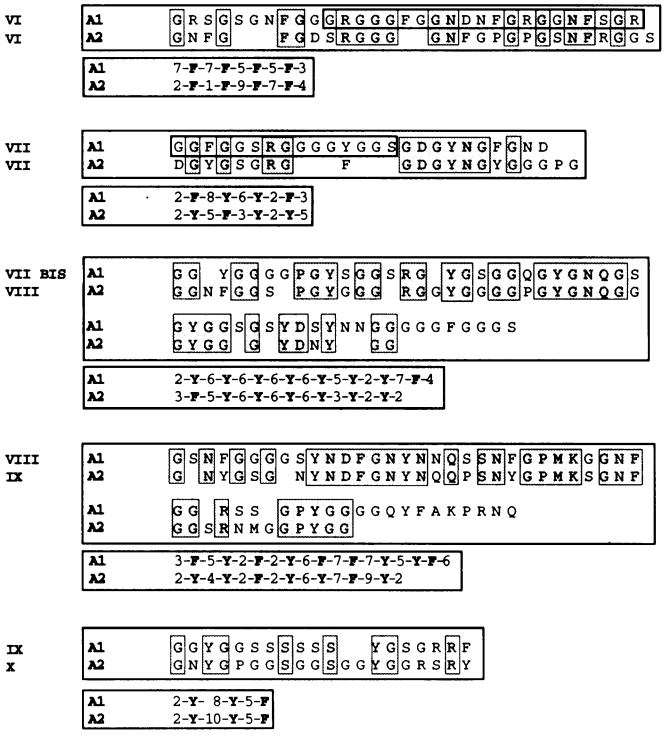

Figure 2. Alignment of hnRNP A1 and A2/B1 exons coding for the Gly-rich domain. Corresponding exons are compared and conserved aminoacids are evidenced by grey boxes. The 'RGG box' in exons VI and VII of A1 (see text) is framed. The conservation of the spacing of aromatic residues in corresponding exons is shown underneath each exon pair.

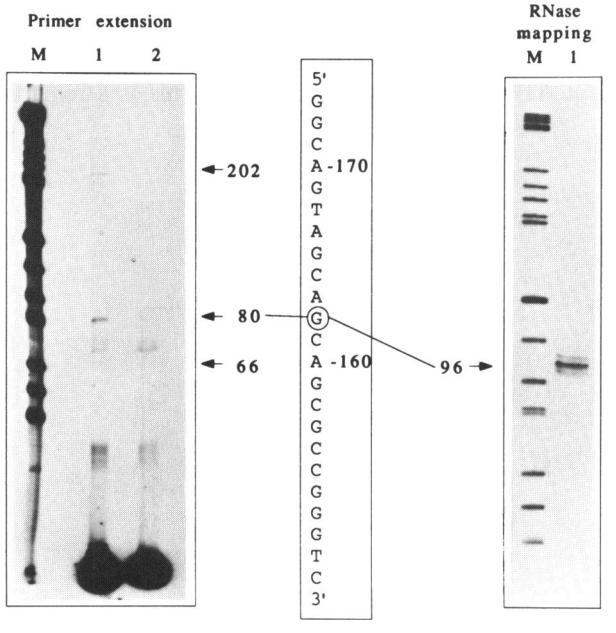

Figure 3. Mapping of transcription start-site by primer extension and RNase protection experiments. For primer extension (left panel) ^{32}P-labelled primer was annealed to HeLa cell poly(A)$^+$ RNA (lane 1) or to yeast tRNA (lane 2). RNase mapping (right panel, lane 1) was performed on total HeLa cell RNA. For details see Materials and Methods. The bands resulting from extended products or protected fragments are indicated (arrows) and the deduced transcription start-site is matched to the corresponding nucleotide. Negative numbering refers to the distance from the first ATG. Lanes M are molecular weight markers (pBR322 DNA digested with *Hae*III).

level. As expected, in both genes most of the 3′ untranslated sequences are contained in a single exon and are preceded by a short exon containing the translational stop codon. This unusual organisation is found in other hnRNP genes in species as distant as human and Drosophila (27) suggesting that it may be important for regulation. Finally, as expected from published information (11), a small exon that by alternative splicing gives rise to B1 mRNA, can be envisaged in the first intron.

The portion of the gene encoding the Gly-rich domain was sequenced using as start points the 5′-end of the first expected exon in this region (sixth exon) and the 3′-end of the penultimate exon as deduced from the previous analysis and from the cDNA sequence. In turn, the end points of these exons determined the borders of the neighbouring ones, that were sequenced, along with most of the flanking introns, with the same strategy as described above. Five exons were found to code for the Gly-rich domain of A2/B1.

In summary, as depicted in Fig. 1, the A2/B1 gene is divided into 12 exons, one of which unique to the B1 mRNA. Exons are separated by introns bounded by the canonical GT...AG dinucleotides and presenting a Py-rich tract at the 3′ end. All determined sequences can be found in the GenBank database (accession numbers: U09120, U09121, U09122, and U09123).

Exon/intron structure of the Gly-rich domains

As shown in the previous section, besides sequence similarity the A1 and A2/B1 genes share the same exon/intron organisation in the 2xRBD domains. This finding is in accord with the initial hypothesis of a common origin by gene duplication and suggests that also the structure of the remaining portion of the genes could be maintained. In fact, as shown in Fig.1 (panel B), the C-terminal portions of the two proteins are encoded by the same number of exons if the additional exon (VII bis) present in the variant $A1^B$ protein (15) is considered. This is a remarkable finding if one considers that the two proteins seem almost unrelated in this region (only 29% identity) and suggests that further similarities could emerge from the comparison of corresponding exons, in line with the generally accepted contention that exons encode protein building blocks. The result of this analysis (Fig. 2) demonstrates that this is indeed the case. Several homology islands can be envisaged in all exons and an overall sequence relatedness, not simply attributable to the abundance of Glycines, is present even in the less conserved exons if small insertions and deletions are allowed. For example, exon VIII shows a remarkable homology to the exon VII bis of A1. Both exons are characterised by aromatic residues (Y and F) regularly interspersed in a G and S rich milieu with only one Arginine residue present (see Fig. 2). Exon IX of A2 and exon VIII of A1 have the highest number of identities. Interestingly, while the periodicity of aromatic residues is maintained, their aminoacid composition is significantly different from that of the preceding exons, being much less Gly-rich. The remaining exons are less conserved, but the spacing of aromatic residues is maintained (Fig. 2).

Transcriptional initiation site and 5′-flanking sequences

The start site of the A2/B1 mRNA was mapped by RNase protection and primer extension experiments.

The result of primer extension is shown in Fig. 3 (left panel). Three extended products are detectable, suggesting start sites at position −148, −162 and −284 with respect to the ATG. To perfect this analysis, we performed an RNase protection experiment with an antisense transcript spanning the region containing all the previously detected start sites. The result (Fig. 3, right panel) evidences only one start site corresponding to the most intense, primer extension band.

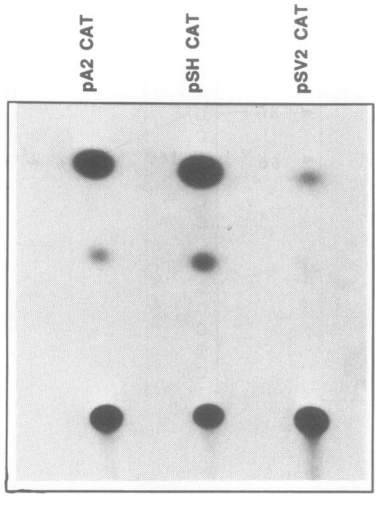

Figure 4. CATase assay of the 5′ flanking sequence of the A2/B1 gene in transiently transfected HeLa cells. The 500 nt upstream of the ATG were cloned in pEMBL8CAT0 and assayed as described in Materials and Methods in parallel with the same vector harbouring the A1 promoter (pSH CAT) and the SV40 promoter (pSV2 CAT). The autoradiography of the chromatography is shown. The experiment was repeated three times with reproducible results.

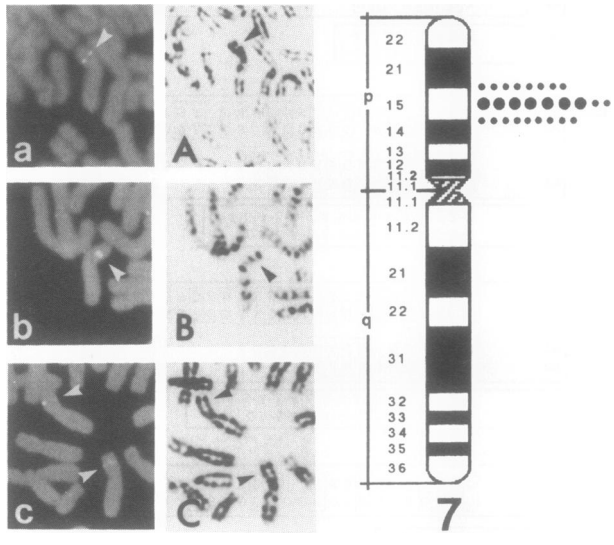

Figure 5. Localisation of the human A2/B1 gene to chromosome 7p15. Biotinylated λ53C DNA was hybridised to metaphase preparations and visualised by FIT-C-conjugated avidin. (**a**, **b** and **c**): partial metaphases showing hybridisation signals on chromosome 7 (arrow heads). (**A**, **B** and C): the same partial metaphases G-banded with Wright's stain. The ideogram of the G-banded human chromosome 7 illustrates the distribution of the hybridisation signals. Each large dot represent ten fluorescent spots.

A total of 1348 nt upstream of transcription start site were sequenced (Fig. 1, panel A and GenBank accession number U09120). As in the A1 promoter, no TATA Box is detectable. On the other hand the 5'-flanking region of A2 gene presents little similarity with that of A1 either in sequence or in putative promoter elements. However, in spite of such divergence, the A2 upstream sequence seems to have a promoter activity comparable to that of A1, as expected from previous evidence (17). 500 nt upstream of the ATG were PCR amplified, inserted into the pEMBL8CAT0 vector and assayed for CATase activity in transfected HeLa cells in parallel with the A1 and the SV40 promoters (see Materials and Methods) (17). The results (Fig. 4) clearly show that the upstream 500 nt of A2 are as active as the A1 promoter, both being much stronger than the SV40 promoter. Interestingly, as indicated by our previous results, both promoters must also be subject to yet unidentified control mechanism that ensure a coordinate expression in different cell types and tissues (17).

Chromosomal localisation of hnRNP A2/B1 gene

The chromosomal location of hnRNP A2/B1 gene was determined by *in situ* hybridisation of the biotinylated λ53C DNA to human metaphase spreads as described in Materials and Methods.

Analysis of 26 metaphases revealed a total of 109 fluorescent signals on chromosomes. 89 fluorescent spots were localised on chromosome 7, of which 72 on the p15 band and 17 at the border between p15 and the two adjacent bands, as identified from the G-banding pattern (Fig. 5). Most chromosomes 7 showed specific signals on both chromatids. This result permits to unequivocally assign hnRNP A2/B1 gene to the G-negative band p15 of human chromosome 7.

DISCUSSION

In this paper we report the isolation of the human gene encoding the hnRNP proteins A2/B1 and we compare its structure to that of the gene for the closely related hnRNP protein A1 (6). While the A1 gene maps on Ch. 12 band q13.1 and has pseudogenes scattered on many human chromosomes (25), a single A2 locus is present on the G-negative band p15 of Ch. 7. In spite of this, we show that the two genes are highly related in sequence and in the overall intron/exon organisation suggesting an origin by duplication of an ancestral gene.

It is worthy of note that, while the A1 and A2 proteins are highly similar in the 2xRBD domains, their Gly-rich domains are apparently unrelated (29% aminoacid identity) (11). However, our results show that conserved elements could be found also in the Gly-rich domain. In effect, when corresponding exons are compared in this domain, conserved blocks of aminoacids and a regular spacing of aromatic residues can envisaged. This observation, reinforces the hypothesis of a gene duplication event and provides a guideline for the interpretation of structure/function relations within the Gly-rich domain. For instance, significative sequence similarity exists between exon VII-bis of A1 and exon VIII of A2. While exon VII-bis is included only in the variant protein $A1^B$ (5% of A1 in HeLa cells) and is subject to a tissue specific regulation (15), a similar exon is constitutively present in the A2 protein. This suggests an important structural and/or functional role for the sequences encoded by these exons. The comparison of exon VIII and IX of A1 and A2 respectively, reveals an even greater similarity which is remarkable in light of the fact that these exons have the lowest Glycine content. It is interesting to observe that exon VIII of A1 encodes for sequences previously implicated in the cooperativity of binding to RNA and thus, probably, in protein−protein interactions (28). It is therefore likely that similar functions could be performed by sequences encoded by exon IX of A2.

In addition to conserved blocks of aminoacids, another type of conservation is clearly observable in this region, i.e. a regular spacing of aromatic residues (Tyr and Phe) (see Fig. 2).

On the other hand, some important divergences can be envisaged. For example exons VI and VII of A1, encoding the putative RNA binding motif designated 'RGG box' (29), differ significantly from the corresponding A2 exons in which no such motif is recognisable.

Altogether, the similarities between A1 and A2 suggest an overlapping of functions between the two proteins. This hypothesis gained some support from the reported isolation of a viable A1-less cell line (30), which indicates that the A1 function can be substituted by another protein such as, for example, the highly related A2. However, the transformed phenotype of this cell line might indicate that the substitute protein is unable to fulfil all the A1 functions. The question then arises of the possible functional specialisation of these proteins and of the structure determinants involved, some of which might reside in the less conserved Gly-rich domains. To this regard, a number of in vitro experiments revealed a role of the Gly-rich domain of A1 in RNA binding and cooperativity (28), in RNA annealing (10, 31), and probably in protein−protein interactions (28, 32). Although no such data are available for the corresponding A2 domain, it is plausible that sequence conservations might underlie common functions while differences could account for functional specialisation.

As an example of functional differentiation, it is worth mentioning that while the A1 protein was reported to modulate 5' splice site choice *in vitro*, both $A1^B$ and A2 seem to be inactive in this assay (4) suggesting a specific role of sequences encoded by exon VII bis and exon VIII respectively.

Therefore the possibility that the Gly-rich domains of A1 and A2 are constituted by an assembly of specialised modules should be considered and tested experimentally. The results will shed new light on many unexplored aspects of RNA−protein interactions.

Finally, it should be noted that, although the 5'-flanking sequences of the two genes bear no apparent similarity, they seem to harbour housekeeping promoters of comparable strength (see Fig. 3). Moreover, since our previous experiments suggest that the two genes are coordinately expressed *in vivo* (17), the promoters must share regulatory pathways that are presently under investigation.

ACKNOWLEDGEMENTS

This work was supported by grants from the Consiglio Nazionale delle Ricerche, Progetti Finalizzati 'Ingegneria Genetica' and 'Biotecnologie e Biostrumentazione'. We thank D.Arena for technical assistance.

REFERENCES

1. Bennett M., Piñol-Roma S., Staknis D., Dreyfuss G., and Reed R. (1992) *Mol. Cell. Biol.* 12: 3165−3175.
2. Matunis E.L., Matunis M.J., and Dreyfuss G. (1993) *J. Cell Biol.* 121: 219−228.

3. Choi Y.D., Grabowski P.J., Sharp P.A., and Dreyfuss G. (1986) *Science* 231: 1534–1539
4. Mayeda A., and Krainer A.R. (1992) *Cell* 68: 365–375.
5. Mayeda A., Helfman D.M., and Krainer A. R. (1993) *Mol. Cell. Biol.* 13: 2993–3001
6. Biamonti G., Buvoli M. Bassi M.T. Morandi C., Cobianchi F., and Riva S. (1989) *J. Mol. Biol.* 207: 491–503.
7. Buvoli M., Biamonti G., Tsoulfas P., Bassi M.T., Ghetti A., Riva S., and Morandi C. (1988) *NucleicAcids Res.* 16: 3751–3770.
8. Buvoli M., Cobianchi F., Biamonti G., and Riva S. (1990) *Nucleic. Acids Res.* 18: 6595–6600.
9. Buvoli M., Cobianchi F., and Riva S. (1992) *Nucleic. Acids Res.* 20: 5017–5025.
10. Cobianchi F., Calvio C., Stoppini M., Buvoli M., and Riva S. (1993) *Nucleic. Acids Res.* 21: 949–955.
11. Burd C.G., Swanson M.S., Görlach M., and Dreyfuss G. (1989) *Proc. Natl. Acad. Sci. USA* 86: 9788–9792.
12. Dreyfuss G., Matunis M.J., Piñol-Roma S., and Burd C.G. (1993) *Annu. Rev. Biochem.* 62: 289–321.
13. Kumar A., Casas-Finet J.R., Luneau C.J., Karpel R.L., Merrill B.M., Williams K.R., Wilson S.H. (1990) *J. Biol. Chem.* 265: 17094–17100.
14. Adam S.A., Nakagawa T., Swanson M.S., Woodruff T., and Dreyfuss G. (1986) *Mol. Cell. Biol.* 6: 2932–2943.
15. Buvoli M., Cobianchi F., Bestagno M., Mangiarotti A., Bassi M.T., Biamonti G., and Riva S. (1990) *EMBO J.* 9: 1229–1235.
16. Mottes M., Tsai Lai S.A., Montoya J., and Attardi G. (1984) *Gene* 27: 109–113.
17. Biamonti G., Bassi M.T., Cartegni L., Metcha F., Buvoli M., Cobianchi F., and Riva S. (1993) *J. Mol. Biol.* 230: 77–89.
18. Benton W. D., and Davis R.W. (1977) *Science* 196: 180–182.
19. Sambrook J., Fritsch E.F., and Maniatis T. (1989) *Molecular Cloning. A Laboratory Manual.* 2nd Edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
20. Dunn J.J. and Studier F.W. (1983) *J. Mol. Biol.* 166: 477–535
21. Melton D.A., Krieg P.A., Rebagliati M.R., Maniatis T., Zinn K., and Green M.R. (1984) *Nucl. Acids Res.* 12: 7035–7056.
22. Graham F.L., and van der Eb A.J. (1973) *Virology* 52: 456–467.
23. Mercola M., Goverman J., Mirell C, and Calame K. (1985) *Science* 227: 266–270.
24. Gorman C.M., Moffat L.F., and Howard B.H. (1982) *Mol. Cell. Biol.* 2: 1044–1051.
25. Saccone S., Biamonti G., Maugeri S., Bassi M.T., Bunone G., Riva S., and Della Valle G. (1992) *Genomics* 12: 171–174.
26. Chandler M.E., Yunis J.J. (1978) *Cytogenetic Cell Genet.* 22: 352–356.
27. Haynes S.R., Johnson D., Raychaudhuri G., and Beyer A.L. (1990) *Nucleic. Acids Res.* 19: 25–31.
28. Nadler S.G., Merrill B.M. Roberts W.J., Keating K.M., Lisbin M.J., Barnett S.F., Wilson S.H., and Williams K.R. (1991) *Biochemistry* 30: 2968–2976.
29. Kiledjian M., and Dreyfuss G. (1992) *EMBO J.* 11: 2655–2664.
30. Ben-David Y., Bani M.R., Chabot B., De Koven A., and Bernstein A. (1992) *Mol. Cell. Biol.* 12: 4449–4455.
31. Kumar A., and Wilson S.H. (1990) *Biochemistry* 29: 10717–10722.
32. Casas-Finet J.R., Smith J.D., Kumar A., Kim J.G., Wilson S.H., and Karpel R. L. (1993) *J. Mol. Biol.* 229: 873–889.