
The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA

A.S.Krayev, D.A.Kramerov, K.G.Skryabin, A.P.Ryskov, A.A.Bayev and G.P.Georgiev

Institute of Molecular Biology, USSR Academy of Sciences, Vavilov street 32, Moscow B-34, USSR

Received 22 February 1980

ABSTRACT

Three copies of a highly repetitive DNA sequence B1 which is complementary to the most abundant class of mouse fold-back RNA have been cloned in pBR322 plasmid and sequenced by the method of Maxam and Gilbert. All the three have a length of about 130 base pairs and are very similar in their base sequence. The deviation from the average sequence is equal to 4% and the overall mismatch between each two is not higher than 8%. One of the recombinant clones used contained two copies of B1 oriented in the same direction. All of the B1 copies are flanked with sequences which possess nonidentical but very similar structure. They consist of a number of $A_m C_n$ blocks (where m varies from 2 to 8 and n equals 1-2). These peculiar sequences in all cases are separated from B1 by non-homologous DNA stretches of 2-8 residues. In one case, a long polypurine stretch is located next to such a block. It consists of 74 residues most of which represent a reiteration of the basic sequence AAAAG. We have found two regions within the B1 sequence which are homologous to the intron-exon junctions, especially to those present in the large intron of the mouse β -globin gene. It may indicate the involvement of the B1 sequence in pre-mRNA splicing.

INTRODUCTION

Nuclear pre-mRNA (hnRNA) contains long double-stranded regions /1-3/. These sequences designated as double-stranded RNA (dsRNA) can be isolated from nuclear RNA preparations by exhaustive digestion with a mixture of RNAases T1 and A followed by chromatography of a Sephadex G-75 column. Long dsRNA which remains undigested is eluted in the void volume and can be further fractionated by gel electrophoresis into two fractions: (1) dsRNA-A 300-800 base pairs long incapable of folding back upon pre-mRNA melting, and (2) dsRNA-B 100-200 base pairs long capable of folding back after pre-mRNA melting /4/.

DsRNA-A is probably the product of intermolecular interaction of certain nuclear RNAs while dsRNA-B seems to originate from the self-complementary regions present in the same pre-mRNA molecules, i.e. from hairpin-like structures /4/. DsRNA-B is transcribed from reiterated DNA sequences and represented by a few number of different sequences as follows from the studies on reassociation kinetics /4,5/. The existence of some homology between dsRNA and polysomal poly(A)⁺RNA suggested the involvement of dsRNA-B in pre-mRNA processing /4,5/.

Recently, we have reported the cloning, in E.coli, of mouse DNA fragments containing sequences hybridizing to dsRNA-B /6/. About a quarter of all the clones obtained in a shotgun experiment contained such sequences. Two main families of the sequences were isolated, B1 and B2, which bound 30% and 20% of all dsRNA-B, respectively, and which did not cross-hybridize. The most abundant was a sequence designated B1 which was really ubiquitous in mouse genome. 40,000-80,000 copies of the B1 sequence were shown to be scattered throughout the whole genome. Some of them form palindromes, but the majority of the B1 sequences do not. Therefore, only about a quarter of the B1 sequences have the double-stranded configuration in pre-mRNA.

The cloning of the B1 unit made it possible to determine its nucleotide sequence. We have sequenced three different members of the B1 family and found that the size of the B1 sequence is equal to about 130 base pairs and the degree of divergency does not exceed 8%. In all the cases, a block consisting of short repeating units of the A_mC_n type (m = 2-8; n = -1-2) was found on one and the same side of the B1 sequence. In one case, a huge polypurine block 74 base pairs long is located next to such a block. Besides this, we could also detect the regions of homology between the sequence B1 and the intron-exon junctions in eukaryotic pre-mRNA.

MATERIALS AND METHODS

Recombinant clones containing mouse DNA fragments (designated Mm31 and Mm35) have been described in the previous paper /6/.

Recombinant bacteria were grown in the M9 medium supplemented with 2% casamino acids in 10 l vessels in a New Brunswick MF-14 laboratory fermentor. Plasmid DNA was isolated as described /7/.

Isolation of mouse DNA inserts. Inserted mouse DNA fragments were isolated by cleaving plasmid DNA with EcoRI or with EcoRI+HindIII and separated from the residual plasmid DNA molecule by sucrose gradient centrifugation (Mm31) or by electrophoresis in 2.6% polyacrylamide gel. The isolated fragments were further cleaved by endonucleases TaqI, BspI, HpaII isolated as described /8/. Restriction digests were separated by electrophoresis in 3% agarose gels, blotted onto nitrocellulose filters and hybridized in probes as described by Southern /24/.

DsRNA-B to serve as a hybridization probe was isolated from mouse Ehrlich ascite carcinoma cells and was labeled with ^{125}I as described previously /3,4,6/. In some cases, DNA fragments labeled with ^{32}P by nick-translation were used as probes

DNA sequencing procedures. DNA sequencing was done essentially as described by Maxam and Gilbert /9/. To obtain labeled fragments of interest, DNA was cleaved by BspI, TaqI, HpaII, AluI or their mixtures, labeled with ^{32}P at the 5'-ends (T4 polynucleotide kinase and a high specific activity γ - ^{32}P ATP, from Amersham, England) and purified from the excess of unreacted labeled ATP by passing through a small column of Sephadex G-50 (Superfine). The fragments were separated by electrophoresis in 5-8% polyacrylamide gels, eluted and either cleaved with another restriction enzyme or strand-separated according to the method of Szalay *et al.* /10/.

Fragments carrying the label at a single end were subjected to several chemical degradation reactions which included:

for G - "alternate G" reaction /9/;

for A+G - depurination either in 60% formic acid for 10 min at 20°C, followed by ethanol precipitation and piperidine cleavage /11/, or in 0.01% formic acid for 1 min at 90°C, followed

by a two-fold dilution with 1 M piperidine and incubation for further 30 min at 90° (the total volume of both reaction mixtures was 25 ul); for T+C and C - hydrazinolysis as described /9/.

In some cases, duplicate experiments were performed both with the standard chemistry (outlined above) and with new chemical reactions employing diethylpyrocarbonate (to locate G's and A's; to be described in detail elsewhere).

Thin gels according to /12/ were used to fractionate chemical cleavage products. Sequencing gels were exposed to pre-flashed medical X-ray films at -70° /13/.

RESULTS

1. Mapping of B1 sequences within DNA fragments of recombinant clones Mm31 and Mm35

Mouse DNA inserts in the recombinant clones Mm31 (1.0 kb) and Mm35 (2.3 kb) are framed with two HindIII or HindIII and EcoRI restriction sites, respectively. Detailed maps with endonucleases BspI (isoschizomer of HaeIII), TaqI and HpaII were constructed for the both of them (Fig. 1).

The location of B1 sequences was established by hybridization of ¹²⁵I-labeled dsRNA-B to the Southern filters containing the Mm31 and Mm35 restriction fragments.

One can see from Fig. 2 that dsRNA-B hybridized only to the Bsp A and Bsp C fragments of Mm31 DNA and to the Bsp A, B and C fragments of Mm35 mouse DNA. Similar results were obtained in cross-hybridization experiments performed in two variants. First, Mm31 DNA was immobilized on a nitrocellulose filter and hybridized to ¹²⁵I-labeled dsRNA-B. The bound RNA was eluted from the filter and rehybridized to a blot with the Bsp-fragments of Mm35 mouse DNA. Second, the ³²P-labeled Mm31 mouse DNA insert was hybridized to the blotted Bsp-fragments of Mm35 DNA. In both cases, the same label was bound to the same subfragments as in hybridization experiments with total dsRNA-B (data not shown). It means that both clones analysed do really contain closely related B1 copies and lack any additional regions of sequence homology. Finally, we found

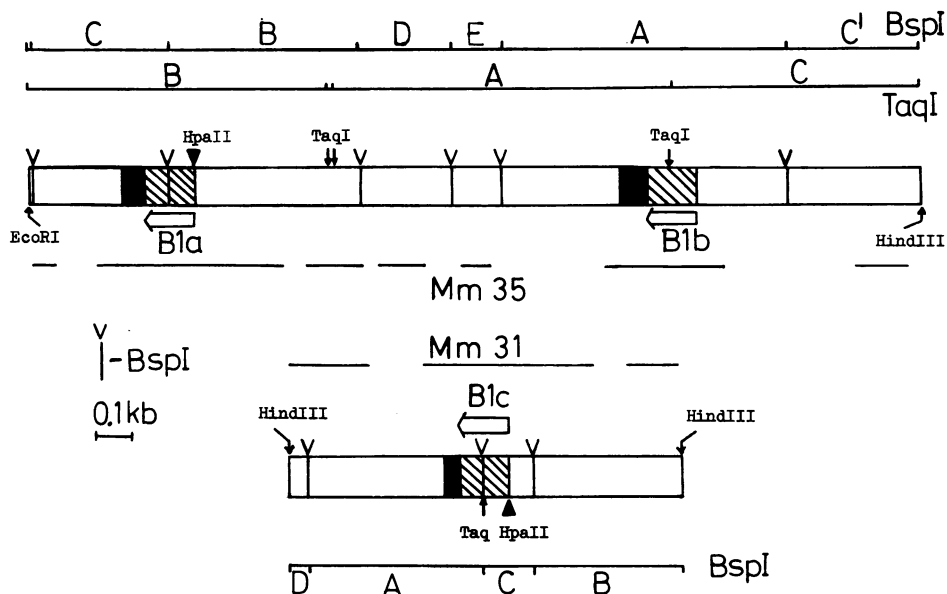


Fig. 1. Restriction map of mouse DNA inserts from the clones Mm31 and Mm35. B1-coding sequences are shown as shaded boxes. Solid boxes show the position of a region consisting of A_mC_n blocks (see text).

Broken lines indicate areas which have been actually sequenced.

that ¹²⁵I-dsRNA-B hybridized to both Taq-derived parts of the BspA-fragment from the clone Mm35 insert (Fig. 2).

Summarizing these data, one may conclude that the Mm31 fragment contains one, and the Mm35 fragment - two separate copies of the B1 sequence. Cross-hybridization experiments have also revealed that each B1 is an asymmetric sequence and that two B1 copies should have the same orientation within clone Mm35 DNA.

2. Sequencing of cross-hybridizing regions in Mm31 and Mm35 inserts

Three regions hybridizing to dsRNA-B and cross-hybridizing to each other (see the previous section) were sequenced. The regions of homology thus obtained are shown in Fig. 3.

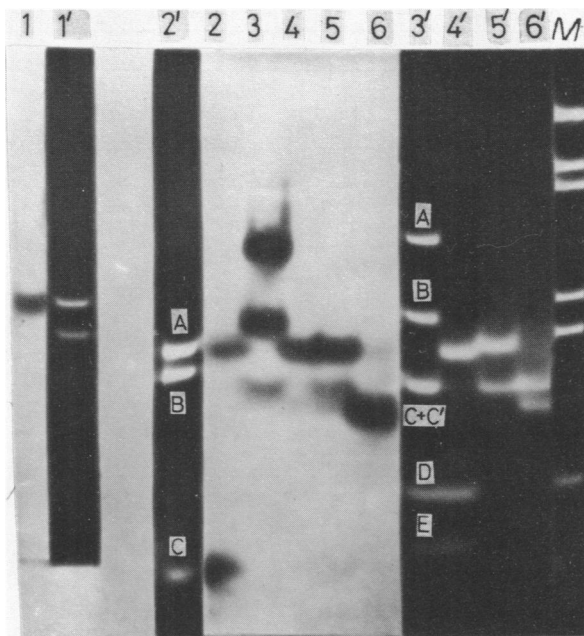


Fig. 2. Southern hybridization of Mm35 and Mm31 DNAs cleaved by restriction enzymes with ¹²⁵I-labeled dsRNA-B as a probe. 1'-6' denote electrophoretic patterns stained with EtBr, 1-6 designate autoradiographs of the corresponding Southern blots.

- 1 and 1' - HpaII fragments of Mm31.
- 2 and 2' - Bsp -"- -"-
- 3 and 3' - Bsp fragments of Mm35.
- 4 and 4' - Taq A fragment of Mm35, cleaved with BspI.
- 5 and 5' - Taq B -"- -"- -"- -"-
- 6 and 6' - Taq C -"- -"- -"- -"-
- M - size markers (HindIII digest of SV40 DNA).

An average sequence, or a "consensus copy", was constructed from three copies determined in such a way that when a difference in the nucleotide sequence between real copies occurred we chose the base present in two copies of the three.

The consensus copy differs from each of the three really existing (see lines 2-4, Fig. 3) by no more than from 5 to 6 substitutions. In one case, the insertion of a nucleotide was observed (B1c). The total length of the homologous region is

```

consensus
sequence   CCGGGCA TG GTGGTGCATG CCTTTAATCC CAGCACTCGG GAGGCAGAGG

      B1c           G A                               C
      B1b   T     -           C                               T
      B1a           -   A   C

consensus
sequence   CAGCGCGATT TCTGAGTTCG AGGCCAGCCT GGTCTTCAGA GTGAGTTCCA

      B1c
      B1b                               T
      B1a           G           A T                               A

consensus
sequence   GGACACCAGG GCTACAGAGA AACCCCTGTCT

      B1c           CA
      B1b                               A
      B1a
  
```

Fig. 3. Sequences of the three B1 copies. The "consensus" sequence is given above (see text). Below, only differences between the three real copies and the consensus one are indicated with letters. Sequences which show homology with several exon-intron junctions are underlined (see Fig. 6).

equal to 129-130 b.p. The boundaries of the homologous regions are clear-cut: beyond the limits of the 130 b.p. homologous stretch, the three sequences deviate remarkably from each other (see below).

The mismatching between any two B1 sequences analysed is 8%. The divergency is the same in copies present in the same cloned DNA fragment and in distantly located copies (Fig. 3). The substitutions are scattered more or less evenly within the unit: no hot-spots of preferential mutability can be observed. The overall GC content of an average B1 sequence is 59%. The sequence itself does not contain long true palindromes or any other long stretches of internal homology. We could only find frequent repetition of the tetranucleotide AGGC which occurs four times in one strand and two times in the other.

The absence of symmetry of the B1 sequence allows one to orient it. Two of them present in the clone Mm35 are oriented in the same direction.

3. Flanking sequences

Fig. 4 shows sequences which flank each of the three B1 units. One can easily see differences between them. At the 5'-end of DNA strands shown in Fig. 3, the flanking sequences completely differ from each other and lack any noticeable homology (Fig. 4B). The only characteristic feature is a rather frequent occurrence of termination signals in both strands and at different frames. On the other hand, the flanking sequences located at the 3'-ends of all the three B1 units are very peculiar and similar to each other although they cannot be called true homologs (Fig. 4A). They represent stretches

A

B1c AAAAAACCAAACCAAACCAAACAAAAACCGAGACAACAATCAAATGGTGTA...

B1b AAAAAACAAAAAACAAAAAACAAAAAACCAACCAACAAACAAACAAACAAAC
 AAAAAACAACAAAAAAGAAAACCTCTCTCTGCAACAGGTGGAGACCATTTAAAAA
 AAAAAA...

B1a AAAAAACAAACAAAAACAAACAAACAAACAAACAAACACCAAAAAAAAAAAGAAAGG
 AAAG(AAAAG)₁₀AA...

B

...TAACTAATACAATTTAATTACTGAGTTYGGCCGTGGTAATCCTTGCAGCAAAATATGACATATA
 ATAATGAGCAGTTAAGAGACAACAATCAGAG (B1c)

...TGATATCCCTAAGGTATAATAGTGGCAGGTATACCTTGGTGGTAACAACAGCTCTCTAATTGGAC
 TATTGAAGACTTGTTCACAAGAGGAAATCAAGC (B1b)

...AAGAGTGGAACTGTGTCAACAGACTGTGTCAACAGACAAAACCTGGGAAGGCTGAGTGGACCCAGA
 AACCTCCCTAAATGCTCATGTCTGGCA (B1a)

Fig. 4. Sequences which flank the B1 unit to the right (B) and to the left (A) as drawn in Fig. 1.

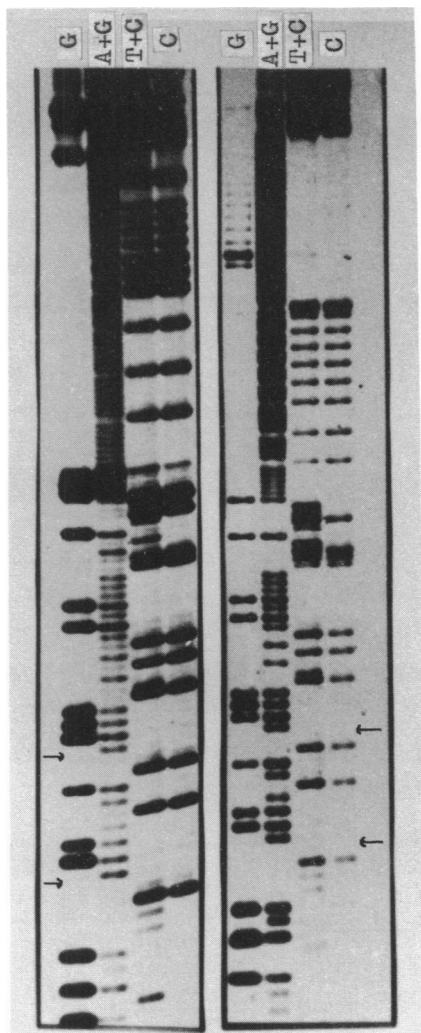


Fig. 5. Sequencing gels which show DNA regions of two B1 units from the clone Mm35. B1 sequences can be read off from the bottom to the approximately middle part of each gel. At the top, A_mC_n sequences and a poly-purine block are also visible though rather poorly resolved. The gels are 8% acrylamide - 7 M urea - 100 mM tris-borate, pH 8.3, 0.4 mm thick, run at 1600 V for 2.5-3 hrs. The specificity of chemical cleavages is denoted with letters at the top. Missing bands marked by arrows correspond to methylated cytosine residues /19/.

from 31 to 69 base pairs long consisting of a number of tandemly repeated $A_m C_n$ blocks where m varies from 2 to 8 and n equals 1 or 2. Such blocks are clearly visible on representative sequencing gels (Fig. 5). Each of $A_m C_n$ motives is separated from the end of the 131 bp sequence by a few nonhomologous base pairs.

It should be stressed that $A_m C_n$ motives are located at the same flank of the B1 unit. In one case, the strand asymmetry is further expanded into another peculiar sequence which appears as a huge polypurine block in the same strand where one reads $A_m C_n$ blocks. The major part of the latter is a reiteration of the AAAAG sequence. This bizarre block is seen unresolved at the top of the sequencing gel in Fig. 5.

The flanks of the two other B1 sequences were not sequenced to such a long distance. However, in the case of B1b, a sequence of 12 consecutive A residues was observed after the $A_m C_n$ motive. It may be considered as a homolog of the above mentioned huge polypurine. Beyond these regions, the sequences show a more or less random distribution of bases and deviate considerably from each other.

Summarizing these data, we may conclude that each B1 is followed with a strand-asymmetric sequence.

DISCUSSION

General properties of B1 sequences

In this work, the representatives of a highly repeated sequence family from the eukaryotic genome were first sequenced. The sequence in question is responsible for transcription of the most abundant class of mouse fold-back RNA. It seems to be scattered throughout the genome and to occur mostly in the form of solitaire copies or direct repeats and less frequently as inverted repeats. In our clones, we have found a solitaire copy and a direct repeat which represents a more probable situation. The sequencing confirmed the results of hybridization experiments showing close similarity between different copies of the B1 family. However, some results require additional explanations.

The sequencing data show that about one half of the 130 b.p. homologous unit is located in each of the fragments Bsp B and Bsp C from the clone Mm35 insert. In view of this fact, it is not quite clear why Bsp C shows a markedly weaker hybridization with ^{125}I -dsRNA-B comparing to that of Bsp B (see blot in Fig. 2). One might suggest that dsRNA-B is not a full copy of the 130 b.p. unit. However, the same pattern is observed in the Southern blots where the Bsp fragments of the Mm35 insert have been hybridized to ^{32}P -DNA of Mm31 used as a probe. Therefore, one is bound to conclude that, due to a yet unknown reason, the Bsp C fragment hybridizes to both probes with an abnormally low efficiency. This feature may be attributed to a peculiar nucleotide sequence close to B1. For example, a huge polypurine block containing the reiterated basic sequence AAAAG is located in the Bsp C fragment. Probably, this simple sequence could have folded back even at room temperature during the DNA transfer step, thus markedly reducing the binding of Bsp C to a nitrocellulose filter. Similarly, A_mC_n sequences might have interfered with the normal hybridization. The latter consideration is supported by the finding (also visible in Fig. 2) that the Bsp A fragment of the Mm31 insert hybridizes to ^{125}I -dsRNA-B weaker than does the Bsp C fragment of the same clone. These observations strongly indicate that one should be cautious in interpreting the data based on the band intensities in Southern hybridization if only short homologous sequences with an unusual composition are to hybridize.

Hence, basing on the data of restriction mapping, Southern blot hybridization with RNA probes, cross-hybridization between clones and quantitative hybridization /6/ we believe that a strict correspondence exists between the 130 b.p. units having been sequenced and the dsRNA-B1 fraction of mouse fold-back RNA.

The degree of deviation from the consensus copy is 4% and the mismatching between any two members of the B1 family is 8%. The latter figure is somewhat higher than that deduced from hybridization experiments /6/. It can be explained by the fact that the Mm35 clone was taken for hybridization as

a standard for perfect hybrid formation. However, as was shown in the present paper, Mm35 DNA contains two different copies of B1 and therefore half of hybrids formed were already mismatched. The same fact accounts for overestimating the length of the B1 sequence in the previous report /6/. In fact, its size is only 130 base pairs.

The sequencing has revealed the existence of very peculiar A-rich sequences in close proximity to each of the B1 units. A question arises whether these sequences should be considered as part of B1. However, these regions can be regarded as "similar motives" and not as true homologous parts though we cannot rule out the possibility of their contribution to hybridization. Recently, Tokarskaya *et al.* /14/ have shown that DNA segments hybridizing to B1 sequences are frequently located closely to the poly(T) hybridizing regions. Various oligo(A) sequences seem to be most probable candidates to explain these observations.

Ubiquitous DNA units, analogous to mouse B1 and B2, have been recently found in DNA of several mammals including man. Their structure, however, remains unknown.

The search for homology with known eukaryotic DNA sequences

The determination of the structure of B1 sequences allows one to begin a systematic search for homologies between them and other eukaryotic sequences studied. Such an analysis may elucidate possible functions of repetitive DNA in the eukaryotic genome which still remain completely obscure.

The most striking similarity was found while comparing the B1 sequence with sequences at the intron-exon junctions (Fig. 6).

Different introns in the eukaryotic genome possess a similarity which allowed Breathnach *et al.* /15/ and Lerner *et al.* /16/ to write a consensus sequence for the beginning and the end of introns. The beginning is better defined because variations of bases are much more prominent at the end of an intron. We have found that the B1 sequence contains two regions separated with 46 nucleotides from each other which are homologous to some extent to the consensus sequence at the exon-intron and intron-exon boundaries.

	Sequence	Source	References
I	...CTTCAGAGTGAGT...	BI	
	<u>CAG</u> <u>GT</u> <u>AAGT</u>	consensus sequence at exon-intron junction	[16]
	CTTCAG <u>GGT</u> GAGT	β -globin	[20] [21]
	<u>AGG</u> TGAGT	kappa chain of Ig	[22]
	<u>TAG</u> <u>GT</u> GAGT	gamma 1 chain of Ig	[23]
	<u>CAG</u> <u>GT</u> <u>AAGT</u>	gamma 2 chain of Ig	[23]
	<u>exon</u> <u>intron</u>		
II	...CCTTTAATCCCAGCACTC	BI	
	YYYYY <u>YYXCAGG</u>	consensus sequence at exon-intron junction	[16]
	TCCTC <u>CCACAGC</u> CTC	β -globin	[20] [21]
	TATTC <u>CCACAGC</u>	kappa chain of Ig	[22]
	<u>intron</u> <u>exon</u>		

Fig. 6. The existence of homologies between two regions of the B1 sequence and intron-exon junctions of eukaryotic genes. The averaged sequence taken from Lerner *et al.* /16/ was slightly modified: we used Y's instead of T in the sequence TYPYYTXCAGG as there was no preferential occurrence of any particular pyrimidine in these positions of the rodent intron-exon junctions - see the data collected in /16/. Bases corresponding to deletion or substitution are underlined, except those in a variable polypyrimidine sequence of the intron end where the underlining was uninformative. Y - pyrimidines.

A better matching can be obtained if one makes a comparison with some individual introns instead of the average one. An especially good correlation was obtained with the large intron of the mouse β -globin gene. In this case, 12 among 13 bases of B1 and of the first exon-intron junction and 8 among 10 bases of B1 and of the second intron-exon junction coincide. Such a degree of homology seems to be quite significant. The homology with B1 expands also to the exon sequence. Therefore, some homology can be observed while comparing the B1 sequence with the sequences in mature β -globin mRNA.

It was shown previously that dsRNA-B could form hybrids of variable length with cytoplasmic and polysomal poly(A)⁺ RNA, in particular, with preparations of globin mRNA /4, 5, 17/. These hybrids are not very stable. It is not clear whether their presence may be explained by the above mentioned homology but such a possibility cannot be excluded.

It should be pointed out that the 10-nucleotide sequence of B1 unit homologous to intron-exon junction exactly coincides with one of the large oligonucleotides observed by Jelinek /18/ in T1 hydrolyzates of human dsRNA (UAAUCCCAGC). This indicates the evolutionary stability of certain parts of the dsRNA coding sequences.

Lerner *et al.* /16/ reported the existence of complementarity between 18 bases of the 5'-end of small nuclear RNA (snRNA) and average eukaryotic introns at the points of their junction with exons.

The significance of all these observations remains tentative. It is possible that the both repetitive sequences, i.e. B1 and snRNAs, are parts of a complex system controlling splicing and other types of pre-mRNA processing. The ubiquity of the B1 sequence in the mouse genome makes its participation in such a general process as splicing not too unlikely.

It is interesting that dsRNAs-B were detected not only in isolated pre-mRNA but also in nuclear RNP particles. Recently, we found that a significant fraction of such hairpin-like structures present in native RNP particles belonged to B1 sequences (Samarina & Kramerov, unpublished).

The mechanisms of splicing in eukaryotes are unknown and it is too early to speculate on the possible ways of the B1 sequence involvement in the process. Further studies of the transcribed repetitive DNA sequences in eukaryotes may, however, provide important information for understanding such a fundamental step in the eukaryotic gene expression as the processing of pre-mRNA is.

REFERENCES

- 1 Ryskov, A.P., Farashyan, V.R., Georgiev, G.P. (1972) *Biochim. Biophys. Acta* 262, 568-572.

- 2 Jelinek, W. and Darnell, J.E. (1972) Proc. Natl. Acad. Sci. USA 69, 2537-2541.
- 3 Ryskov, A.P., Saunders, G.F., Farashyan, V.R., and Georgiev, G.P. (1973) Biochim. Biophys. Acta 312, 152-164.
- 4 Kramerov, D.A., Ryskov, A.P., Georgiev, G.P. (1977) Biochim. Biophys. Acta 475, 461-475.
- 5 Ryskov, A.P., Kramerov, D.A., Georgiev, G.P. (1976) Biochim. Biophys. Acta 447, 214-229.
- 6 Kramerov, D.A., Grigoryan, A.A., Ryskov, A.P., Georgiev, G.P. (1979) Nucleic Acids Res. 6, 697-713.
- 7 Bolivare, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Grosa, J.M. and Falkow, S. (1977) Gene 2, 95-113.
- 8 Bickle, T.A., Pirotta, V., Imber, R. (1977) Nucleic Acids Res. 4, 2561-2572.
- 9 Maxam, A.M., Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
- 10 Szalay, T., Grohman, K., Sinsheimer, R.L. (1977) Nucleic Acids Res. 4, 1569-1578.
- 11 Skryabin, K.G., Zaharyev, V.M., Bayev, A.A. (1978) Doklady AN SSSR, v.241, 488-492.
- 12 Sanger, F., Coulson, A.R. (1978) FEBS Lett. 82, 314-317.
- 13 Laskey, R.A., Mills, A.D. (1977) FEBS Lett. 82, 314-317.
- 14 Tokarskaya, O.N., Tchurikov, N.A., Ivanov, P.L., Kramerov, D.A., Ryskov, A.P. (1980) Nucleic Acids Res. 8, N3 (in press).
- 15 Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., Chambon, P. (1978) Proc. Natl. Acad. Sci. USA 75, 4853-4857.
- 16 Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., Steiz, J.A. (1980) Nature, 283, 220-223.
- 17 Ryskov, A.P., Tokarskaya, O.N., Georgiev, G.P., Coutelle, Ch. and Thiele, B. (1976) Nucleic Acids Res. 3, 1487-1498.
- 18 Jelinek, W.R. (1977) J. Mol. Biol. 115, 591-601.
- 19 Ohmori, H., Tomizawa, J., Maxam, A.M. (1978) Nucleic Acids Res. 5, 1479-1485.
- 20 van den Berg, J., van Ooyen, A., Mantei, N., Schamböck, A., Grosveld, G., Flavell, R.A., Weissmann, C. (1978) Nature 276, 37-44.
- 21 Konkel, D.A., Maizel, J.V., Leder, P. (1979) Cell 18, 865-873.
- 22 Max, E.E., Seidman, J.G., Leder, P. (1979) Proc. Natl. Acad. Sci. USA 76, 3450-3454.
- 23 Sakano, H., Rogers, J.H., Hüppi, K., Brack, C., Trautnecher, A., Maki, R., Wall, R., Tonegawa, S. (1979) Nature 277, 627-633.
- 24 Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.