Nucleotide Sequence of Bovine 1.723 Satellite DNA

G. Pluciennicak, J. Skowronski, A. Pluciennicak, and J. Jaworski

Department of Biochemistry, Institute of Physiology and Biochemistry, Medical School, ul. Lindleya 6, 90-131 Lodz, Poland

Z. Naturforsch. **40 c**, 242–246 (1985); received April 24/November 27, 1984

1.723 bovine satellite DNA, Satellite DNA evolution

The nucleotide sequence of the bovine 1.723 satellite DNA repeated unit was determined. The 680 bp long period of this satellite DNA does not show any significant sequence similarities with the known bovine satellite DNAs. Short repetitive sequences which are parts of 680 bp long repeated units do not form any orderly periodical structure. It seems, however, that the basic repeated unit of the 1.723 bovine satellite DNA has been formed by successive duplications of two, about 100 bp long sequences. The sequence divergence between different copies of the 680 bp repeated unit was also analyzed.

Introduction

The nucleotide sequence of 5 out of 8 bovine satellite DNAs have been described [1–7]. The basic repeated units of all sequenced bovine satellite DNAs are internally repetitive. 1.706, 1.720b and 1.711a bovine satellite DNAs [1–3] have internal 23 bp long repetitive sequences with pronounced sequence homology. This is strong evidence for the common origin of these three satellite DNAs. The 1399 bp basic repeated unit of the 1.715 bovine satellite DNA reveals 31 bp internal periodicity and the 31 bp average sequence is homologous to the 23 bp sequences coming from the 1.706, 1711a and 1.720b bovine satellite DNAs [4]. The 1.711b bovine satellite DNA basic repeated unit consists of one or more 1.4 kb DNA fragments highly homologous to the 1399 bp period of the 1.715 satellite DNA [7, 8]. Therefore, all the above mentioned bovine satellite DNAs form a family of related sequences. This close relationship can be also observed when comparing the nucleotide sequences of unperiodical segments inserted into 31 bp periods of 1.711 b and 23 bp periods of 1.711a bovine satellite DNAs. These insertions have also pronounced sequence homology [3, 4, 7].

The aim of this work was to analyze the sequence organization of the basic repeated unit of the 1.723 bovine satellite DNA and its possible relation to other bovine satellite DNAs.

0341-0382/85/0300-0242 $ 01.30/0

Materials and Methods

Total bovine DNA was isolated from kidney [9], thymus [9], sperm [10] and leucocytes [11]. For isolation of 1.723 satellite DNA, preparations enriched in this satellite DNA by means of precipitation with histon H1 [12] have been centrifuged in the Cs_2SO_4/Ag^+ preparative gradients as described by Macaya *et al.* [13].

The 680 bp repeated unit produced by digestion of 1.723 satellite DNA with *Bcl*I was ligated to BamHI digested plasmid pUR222. Construction of recombinant plasmids and transformation was performed according to Ruther *et al.* [14]. Recombinant plasmids were isolated by a rapid alkaline extraction method [15] and used for the isolation of the inserted DNA fragments.

Restriction endonucleases

Restriction endonucleases *Eco*RII, *Sau*3AI, *Alu*I, *Sau*96I, *Hae*III, *Msp*I and *Sst*I were obtained and used as described earlier [16–18]. *Hha*I, *Hpa*II were from BRL and *Bcl*I was purchased from Sigma. Digestion conditions for *Sau*96 were as in [19] and for *Bcl*I as in [20]. For all other enzymes the incubations were done in the medium of 6 mM $MgCl_2$, 6 mM mercaptoethanol and 6 mM Tris-HCl, pH = 7.6.

DNA sequence analysis

Sequence analysis of different restriction fragments derived from the 680 bp long repeated unit was carried out according to Maxam and Gilbert [21]. Five specific reactions for cytosine, adenine

```
                                        C
GATCATGGGC  TGGTTCCCCT  GCGCTGCGGC  TGTGCTAGGG  AGCCCAAACG      50
                                        *

                                            C   C
CCCAGGGGAA  GCCCGCACAG  CACATCCACC  TGCCTCTAGG  CGGGAGGACA     100
                                            *   *

GCTCCAGGGA  AGGCCCGGTC  TGGCAGGCGC  TTTCTGGGCA  CAAACGGAGC     150

AACACTCTGG  CTGTGGAGCT  GCCTCCCCAA  GCCAGGGGCA  CCGGGACCCT     200
                                               ⌐

GATTCCCTCC  GTGCGGCGCT  AGTCTGGGCG  GGGAAGCACC  AGCTTGGGAA     250

   G
TGCAGCCTGG  GGCGCTTCCT  TTCCCTCGGA  AAGAGCTCTC  CCCGCGGGAG     300
 *  ⌐

         C  C
GCTTTGGCAG  TGTGCACACC  CCCAGCCCCT  AAGCCCTCTG  CAAGGACGCC     350
        *  * ⌐

TGGCTCCTAG  GTCCGAAGGA  AGGCGAGAGC  TGAGCTGAGG  CACCAGTCAG     400

              A
GCACCACTCA  CGGCTCCCAC  ACTCCCGCAG  CCTGCTCCTT  CCCGGGCGCG     450
              *

                 C
AACGCAGAAG  CTGGGCAGGG  CTTGGGGAGG  CTCGCCCAGC  ACAGAGCAGC     500
                 *

    C        C        G       G                    A  G
CTCTTGGCCG  GGGCTCTGTT  CTATCCCCTG  GGCGCAGAGC  TGGAGCTCCC     550
 *          *        *       *                    *  ✤

             G     A                                 G
CGGTGTGGAA  GCTCACTCGG  TCCTTCCCAG  TACAGGCTTG  CAAGCACCCA     600
             *     *                                 *

                                    G       A         A
CCGGCAGCCA  GCCTCTTCTC  TCCCGCGTTC  CGGGCATTGG  GCTCCAAACC     650
                                    *       *         *

TGGGACCAAA  GTAACTGCCC  GGAGCTCACT                        680
```

Fig. 1. Nucleotide sequence of the 680 bp *Bcl*I repeated unit of 1.723 bovine satellite DNA. Marks represent the base changes (*), deletions (−) or insertions (+) in the sequence of the 680 bp repeated unit.

```
AG CAACACTC      /148-157/

AGGCACCAGTC      /388-398/

AGGCACCACTC      /399-409/

GGAAGCTCACT      /557-567/

GG AGCTCACT      /671-680/
```

Fig. 2. Regions of significant homology scattered on the 680 bp unit and their positions.

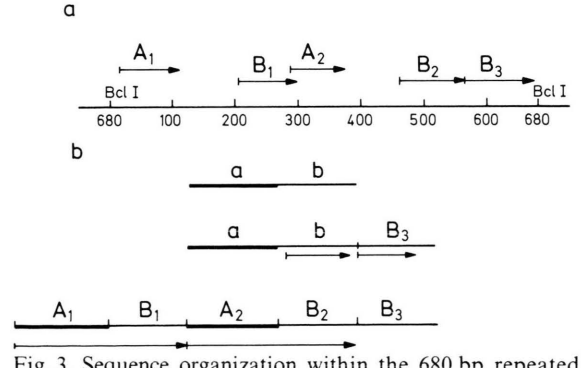

Fig. 3. Sequence organization within the 680 bp repeated unit of 1.723 bovine satellite DNA; a) localization of segments of A and B type along 680 bp unit; b) hypothetical scheme of evolution of the 1.723 bovine satellite DNA.

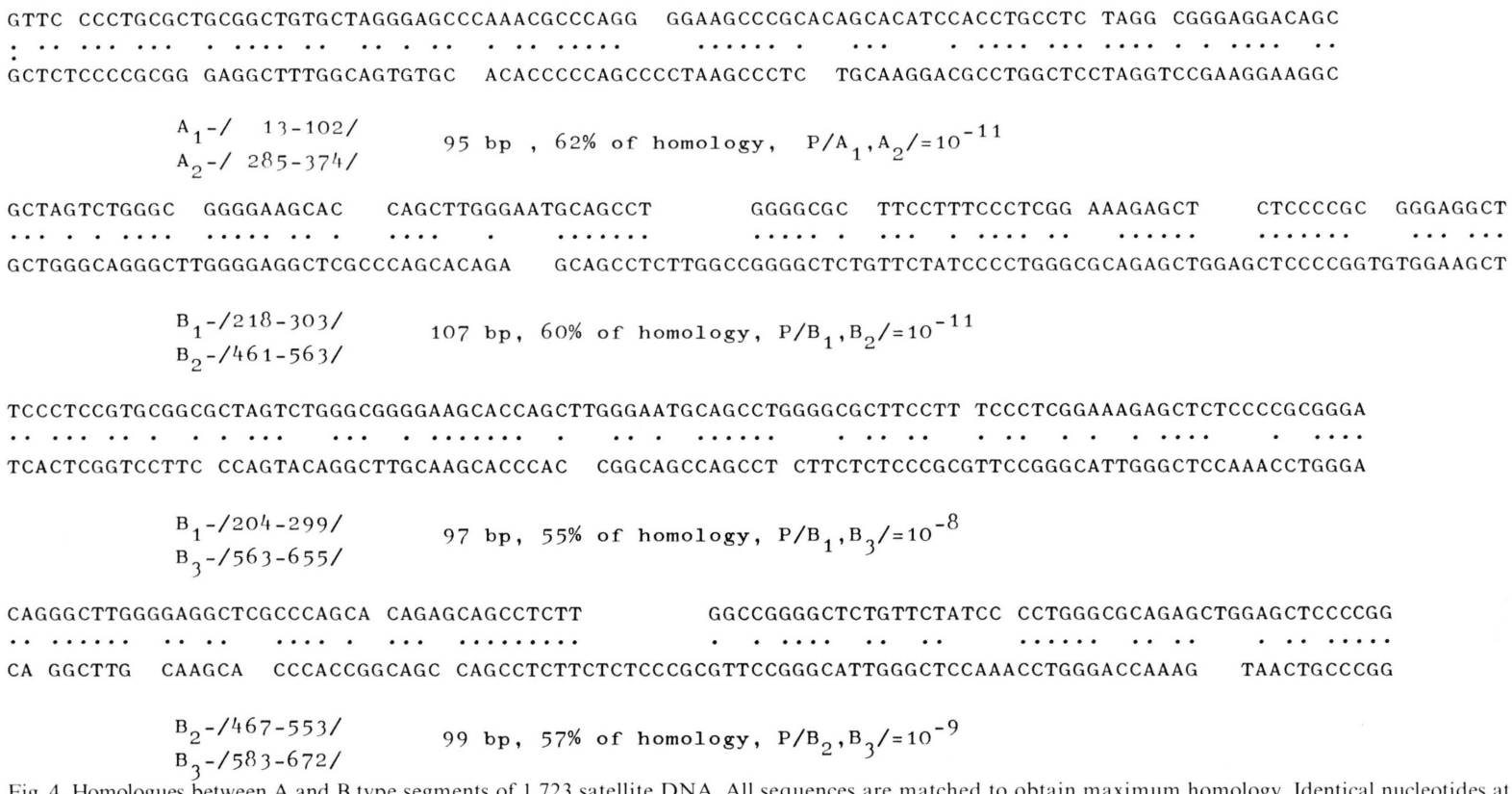

Fig. 4. Homologues between A and B type segments of 1.723 satellite DNA. All sequences are matched to obtain maximum homology. Identical nucleotides at particular positions are marked with dots. Data concerning positions, lengths, percentage of homology and double matching probability are shown under the compared sequences.

plus cytosine, pyrimidines, purines and guanine were performed.

The nucleotide sequence was analyzed on computers MERA 305 and ODRA 1305 using programs developed in this laboratory [4]. The probability that the observed sequence homologies between different regions of sequence occur by chance was computed according to double matching probability method of McLachlan [22] and Staden [23] and was corrected for the sequence length.

Results and Discussion

For sequencing the 1.723 bovine satellite DNA, we used total (uncloned) fragments. Despite this, most of the sequencing gels showed no ambiguity. The G+C content of the bovine 1.723 satellite DNA amounts to 66.4%. The complete nucleotide sequence of the 680 bp repeated unit is shown in Fig. 1. The strand presented has the following base composition (in mol%): A, 17%, T, 31%, G, 31% and C, 35%. The buoyant density of the bovine 1.723 satellite DNA, calculated from its base composition, is 1.725 g/cm^3. This value is 0.002 g/cm^3 higher than that estimated in analytical ultracentrifugation experiments [1, 2]. Similarly to the 1.715 bovine satellite DNA [4], this difference may be due to the presence of methylated cytosine residues.

The computer analysis of the 680 bp sequence performed according to McLachan [22], Staden [23] and the study of distribution of distances between identical di- or trinucleotides have not revealed the presence of any periodicity spread over the 680 bp repeated unit. On the other hand, the frequency of certain tetranucleotides of 1.723 satellite DNA deviates from that expected for DNA sequence of a given base composition. For example, tetranucleotides AGCT, CTCT, TCTG and TTCC occur, respectively, 5.5, 4.2, 3.7 and 3.3 times more frequently than expected. We have also found several regions of significant homology of 11 bp scattered on the 680 bp unit (Fig. 2). Further computer analysis based on construction of a simplified matrix of homology revealed two groups of segments A and B (Fig. 3a). There are two A type segments A1 and A2 and three B type segments B1, B2 and B3. The percentage of sequence homologies between these segments and the probabilities of their appearance by chance in a 680 bp long random sequence of base composition as in 1.723 satellite DNA are given in Fig. 4.

The localization of segments of the A and B type along the 680 bp unit (Fig. 3a) enabled us to consider this unit as composed of imperfect direct repeats formed by A1B1 and A2B2 segments combined with the B segment. Such an arrangement of the segments taken together with the data presented in Fig. 4 suggests the possible sequence of events during the generation of the basic repeated unit of the bovine 1.723 satellite DNA. The highest values of sequence homology observed for A1 and A2, B1 and B2 (Fig. 4) suggest that duplication of a b (Fig. 3b) ancestral segment was the last step in the formation of the repeated unit. The B3 segment which reveals much lower but significant homology with B1 and B2 segments probably originated earlier from the ancestral b segment, possibly by duplication. In summary, it seems possible that the ancestral segment a b or its parts after two consecutive duplications gave rise to the ancestral 680 bp sequence.

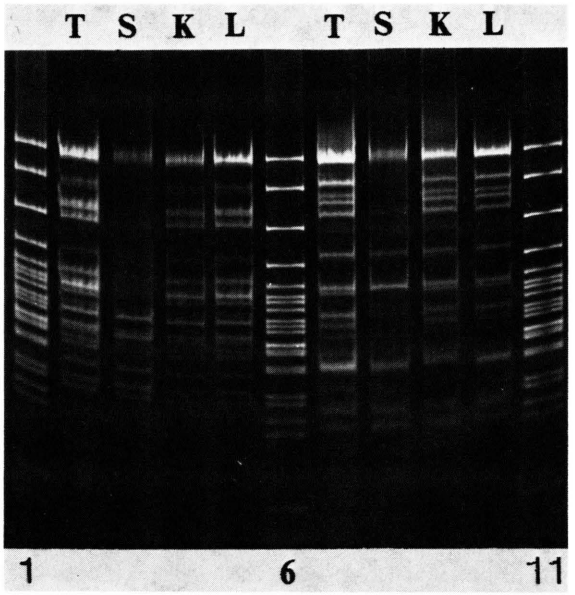

Fig. 5. Methylation pattern of the 1.723 bovine satellite DNA as revealed by digestion with *Hha*I and *Hpa*II restriction nucleases. The 1.723 satellite DNA preparations isolated from thymus – T, sperm – S, kidney – K and leucocytes – L after digestion with *Bcl*I restriction nuclease were subjected to hydrolysis with *Hha*I (channels 2–6) and *Hpa*II (channels 7–10) restriction nucleases. Channels 1, 6 and 11 – 1.5 μg pBR322 plasmid DNA digested with *Msp*I restriction nuclease. Electrophoresis has been carried in 4% polyacrylamide gel (bis : acrylamide = 1 : 19).

The sequence heterogeneity within the 680 bp repeated unit was investigated. The restriction analysis of total and cloned 680 bp repeated units reveals the sequence heterogeneity between different 680 bp DNA segments (Fig. 1). The data from the restriction analysis were confirmed by sequence determination of 3 cloned 680 bp segments of 1.723 bovine satellite DNA. Only point mutations were observed. The distribution of these mutations along the 680 bp repeated unit indicates that the changes of the sequence are grouped in a particular region, at positions 500–650.

The methylation pattern of cytosine residues in 33 CpG dinucleotides occurring in the basic repeated unit as revealed by digestion with HhaI and HpaII restriction nucleases is similar for DNA isolated from thymus, kidney and leucocytes. In a part of the 680 bp repeated units all the cleavage sites for *Hha*I and *Hpa*II are blocked but in most of them only partial methylation is observed (Fig. 5). We have also noticed that the 1.723 satellite DNA is undermethylated in sperm. Similar results has been reported for other satellite DNAs [1–5, 25].

Acknowledgements

We thank Mrs. K. Siewierska for her skillful technical assistance. This work was supported by the Polish Academy of Sciences within the project 09.7.2.3.2.

[1] M. Pech, R. E. Streeck, and H. G. Zachau, Cell **18,** 883–893 (1979).
[2] E. Poschl and R. E. Streeck, J. Mol. Biol. **143,** 147–154 (1980).
[3] R. E. Streeck, Science **213,** 443–445 (1981).
[4] A. Pluciennicak, J. Skowronski, and J. Jaworski, J. Mol. Biol. **158,** 293–304 (1982).
[5] C. Gaillard, J. Doly, J. Cortadas, and G. Bernardi, Nucleic Acids Res. **10,** 6069–6082 (1981).
[6] E. J. Taparowsky and S. A. Gerbi, Nucleic Acids Res. **10,** 1271–1281 (1982).
[7] R. E. Streeck, Nature **298,** 767–769 (1982).
[8] E. J. Taparowsky and S. A. Gerbi, Nucleic Acids Res. **10,** 5503–5515 (1982).
[9] A. Pluciennicak, J. Bartkowiak, A. Krzywiec, and H. Panusz, Biochem. Biophys. Res. Comm. **56,** 799–806 (1974).
[10] E. Borenfreund, E. Fitt, and H. Bendich, Nature **4796,** 1375–1377 (1961).
[11] M. Poncz, D. Solowiejczyk, B. Harpel, Y. Mory, E. Schwartz, and S. Surrey, Hemoglobin **6,** 27–36 (1982).
[12] A. Pluciennicak, J. Bartkowiak, A. Krzywiec, and H. Panusz, Nucleic Acids Res. **1,** 1675–1690 (1974).
[13] G. Macaya, J. Cortadas, and G. Bernardi, Eur. J. Biochem. **84,** 179–188 (1978).
[14] U. Ruther, M. Koenen, K. Otto, and B. Miller-Hill, Nucleic Acids Res. **9,** 4087–4098 (1981).
[15] H. C. Birnboim and J. Doly, Nucleic Acids Res. **7,** 1513–1524 (1979).
[16] T. A. Bickle, V. Pirrotta, and R. Imber, Nucleic Acids Res. **4,** 2561–2572 (1977).
[17] O. H. Smith and Wilcox, J. Mol. Biol. **51,** 379–391 (1970).
[18] D. I. Smith, F. Blattner, and J. Davies, Nucleic Acids Res. **3,** 343–353 (1976).
[19] J. S. Sussenbach, P. H. Steenberg, J. A. Rost, W. J. van Leeuven, and J. D. V. van Embden, Nucleic Acids Res. **5,** 1153–1163 (1978).
[20] A. H. A. Bingham, T. Atkinson, D. Sciaky, and R. J. Roberts, Nucleic Acids Res. **5,** 3457–3471 (1978).
[21] A. M. Maxam and W. Gilbert, Methods Enzymol. **65,** 499–560 (1980).
[22] A. D. McLachlan, J. Mol. Biol. **61,** 409–424 (1971).
[23] R. Staden, Nucleic Acids Res. **10,** 2951–2961 (1982).
[24] H. Ohmori, J. Tomizawa, and A. M. Maxam, Nucleic Acids Res. **5,** 1479–1485 (1978).
[25] H. Sano and R. Sager, Proc. Natl. Acad. Sci., USA **79,** 3584–3588 (1982).