Biochimie (1992) 74, 187–194

On the tertiary structure of satellite DNA

T Beridze[1], N Tsirekidze[1], MA Roytberg[2]

[1]*Institute of Plant Biochemistry, Georgian Academy of Sciences, Tbilisi, 380031;*
[2]*Computer Research Centre, Academy of Sciences of USSR, Pushchino, Moscow Region, 142292, USSR*

(Received 10 October 1990; accepted 25 May 1991)

Summary — The primary structure of the *Citrus ichangensis* satellite DNA repeating unit has been estimated. The repeat is 181 bp long and contains four pentanucleotides of adenine residues. Oligomer forms of the stDNA repeating unit were detected by a partial hydrolysis of the *C ichangensis* stDNA by *Bsp*I restriction endonuclease. Experiments on comparative mobility of oligomers in agarose and polyacrylamide gels evidenced a certain retardation of those in polyacrylamide gel indicating to a slight bend in the repeating unit. The BEN computer program [9] was employed to calculate the spatial positions of monomer and oligomer axes of the satellite DNA repeating unit of *Citrus ichangensis*, mouse and African green monkey, and to plot their two-dimensional projections. The bends in the monomer for higher oligomer form proved to result in a hypothetical solenoid-like structure, termed coiled double helix (CDH).

key words

Introduction

An eukaryotic genome containing unique and moderately repeating sequences also contains tandemly arranged multiply repeating ones termed satellite DNAs (see for review [2]). In some cases stDNA are detectable when a nuclear DNA is subject to equilibrium ultracentrifugation in the CsCl neutral density gradient, wherein it occupies a separate band from major DNA mass. Whenever the stDNA and major DNA densities coincide, stDNAs can be isolated in the CsCl gradient in the presence of some antibiotics or other DNA-binding low-molecular ligands, and in the Cs_2SO_4 density gradient in the presence of mercury or silver ions. Besides, stDNAs can be isolated using restriction endonucleases, which are to be chosen so as to digest the major DNA while leaving the stDNA intact. The stDNA content in the eukaryotic genome is quite high, reaching a half of the whole nuclear DNA in some species. The length of stDNA repeat varies greatly, from 2 (for the genus Cancer) to several hundreds. The stDNA GC-content is also quite variable, both from species to species and within species. Despite the numerous studies devoted to stDNAs, their functional role still remains obscure. They are known to be located in the constitutive heterochromatin areas of the chromosomes defining in a way their compact structure.

Study of the stDNA-containing chromatin has shown that the arrangement of nucleosomes along the stDNA chain is not random and all the studied cases revealed one major and several minor nucleosome positions in the stDNA. The amount of DNA per volume is 1.5 times as much as euchromatin in the constitutive heterochromatin. It still remains obscure why DNA is so compact in heterochromatin, the compact packing seeming to be the basis of heterochromatin visualization in a nucleus.

The recent studies have shown clearly that the linear DNA molecules can contain tertiary structure elements which are associated with the presence of bends. Bends in the DNA molecules were discussed by Trifonov and Sussman [21]. The first experimental proof was obtained by Marini *et al* in the study of the *Leishmania tarentolae* kinetoplast DNA [12]. Local bends in the DNA molecules were detected in the A-tract locations (n = 2–9). When the A-tracts repeated multipliable in phase with the helix repeat, the total bend grew to a considerable size and could be recorded by various physical methods [20, 00]. Programs were compiled to build tertiary structure models on the basis of the DNA primary structure [19].

This work deals with: the problem of identifying the primary structure of the *C ichangensis* stDNA repeating unit; the mobility of the stDNA repeating unit oligomers in agarose and polyacrylamide gels; and gives a hypothetical model of the tertiary structure of

Abbreviations: bp, base pair; stDNA, satellite DNA; AGM, African green monkey; CDH, coiled double helix.

the repeating unit oligomers of the *C ichangensis*, mouse and AGM stDNA, built on the basis of the BEN program by Eckdahl and Anderson [9].

Materials and methods

stDNA were isolated from *Citrus ichangensis* leaves by the procedure described earlier [1].

The *C ichangensis* stDNA was hydrolyzed by restriction endonucleases *Bsp*I or *Hae*III in a low salt buffer solution during the time interval shown in the figures.

The stDNA digestion products were subjected to electrophoretic analysis in polyacrylamide gel in an LKB apparatus (16 x 18 x 0.2 cm) in the 0.05 M Tris-borate buffer solution, pH 8, 3 and at 8 V/cm. After that the gel was placed into ethidium bromide solution (1 μg/ml) for 10 min and then photographed with a transilluminator Chroma 43 (Helling).

Preparative separation of the *Bsp*I-fragments of the *C ichangensis* stDNA was carried out on a plate (20 x 40 x 0.3 cm) in an 8.0% polyacrylamide gel at 4 V/cm during 16 h. DNA fragments were extracted from the gel by a solution containing 0.5 M CH_3COONH_4, 0.01 M $(CH_3COO)_2$ mg, 1 mM of EDTA, 1% Na dodecylsulphate for 10 h at 37°C, and precipitated by ethanol. Later they were reprecipitated by ethanol, with 1/2 volume of 7.5 M CH_3COONH_4 added previously.

Electrophoretic analysis was also performed in an agarose horizontal block (10 x 20 x 0.3 cm) in the Tris-borate buffer solution at 5 V/cm. Ethidium bromide was introduced into the agarose solution in 1 μg/ml concentration.

Cloning and sequencing of the *C ichangensis* stDNA *Bsp*I-fragments were performed in the M13mp8 bacteriophage according to Amersham or Pharmacia prescription. The replicative form of M13mp8 was digested by the *Sma*I restriction endonuclease and then ligated with the analysed fragments. α-^{32}P-dATP was used as a radioactive label.

The primary structure of the cloned fragments was analysed with a computer using the program of multiple comparison of sequencesd, LINNEUS, contained in the SAMSON packet [23].

The models of the tertiary structure were built in the Walters PC AT computer using the BEN program by Eckdahl and Anderson [9].

Reagents

Cloning and sequencing were performed using the kits of Amersham and Pharmacia; the restriction endonucleases used in the work were *Sma*I and *Hae*III of Amersham, *Bsp*I of Reanal; α-^{32}P-dATP of Radiopreparat (Tashkent, USSR).

Results

Determination of the primary structure of the C ichangensis stDNA repeating unit

The *C ichangensis* stDNA is fully cleaved into small fragments by *Bsp*I or its isoschizomer *Hae*III (fig 1). Eleven bands were detected altogether in analytical gel. Some additional bends were found in the preparative gel. They are numbered in figure 1. In order to define the primary structure of the *C ichangensis*

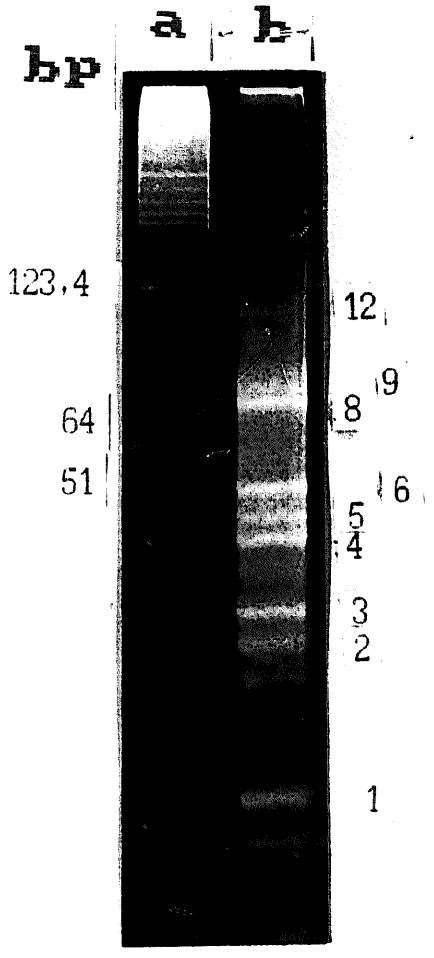


Fig 1. Description of products of the *C ichangensis* stDNA hydrolysis by the *Bsp*I restriction endonuclease in 16% polyacrylamide gel. The hydrolysis condition: 1 unit of enzyme/μg of DNA, 20 h, 37°C. The isolated and sequenced bands are numbered. The lacking bands 7, 10, 11 are seen in the preparative gel.

stDNA repeating unit, DNA was isolated preparatively from 12 bands, cloned in the M13mp8 replicative form and sequenced by a Sanger's dideoxy method. Forty clones of the total length of about 2400 bp were sequenced.

The method of computer analysis of primary structures consisted of defining 'arguable' sites and subsequent mapping of the sequences according to the detected sites. The analysed sequences turned out to be correlated with a 181-bp long period. This period was 'covered' by sequences 12–1 and 10–9 so that the 12–1 end was homologous to the 10–9 beginning. Clone 3–1 is a two-fold repeat of the consensus site. The other sequences (excluding 1–2, 1–3, 1–4, 2–2,

3–1, 5–3) were homologous to the fragments of the discovered consensus. The sequences shown in the brackets were 'glued up' of two consensus areas so that each part is either homologous to a certain area of the consensus, or complementary (the end part of fragments 1–3, 1–4, 5–3) to a certain consensus area.

A summary of all the studied sequences is showed schematically in figure 2. One direct repeat is observable in the consensus:

```
38                                          65
GGCCGGGGC**GCCAAGTTCGTCCAGCGGA
——GT————TG———C————GC———
102                                         131
```

A certain homology exists also between the end and beginning of the defined consensus:

```
154                                         181
CGAGGCGCCTCCGTCTGCCAAAAATAGG
          AA          C
–CGC–G———G——*–*————————CC
1                                           29
```

It is therefore quite possible that the first 153 symbols are the period of consensus, the remaining symbols being a repeat of the beginning of the next period. It should be, however, mentioned that the similarity between the consensus end and its beginning is lower than the average level of similarity between the initial fragments and consensus. Thus, a more accurate definition of the consensus length calls for more information.

Detection of the oligomer forms of the C ichangensis stDNA repeating unit

Analysis of products of a partial hydrolysis provided an accurate estimation of the consensus length in the *C ichangensis* stDNA repeating unit. Figure 3 shows distribution of fragments obtained after partial hydrolysis of the stDNA in the 2% agarose gel. The figure shows clear bands of the monomer and subsequent oligomers. Band length calculations showed that the monomer was about 180 bp long, and the oligomer lengths were multipliable to it, *ie* 360, 540, and so on. It should be noted that there were four faint bands between the oligomers (they are better seen in fig 4). While looking through the consensus of a repeating unit, one would discover four *Hae*III-sites within a sequence (positions 38, 70, 111, 159). It might thus be presumed that the faint bands in the electrophoregram correspond to fragments formed during the DNA cleaving in these four sites. This presumption is supported by the arrangement of the sequenced clones in figure 2.

The beginning of these sequenced clones coincides as a rule with the consensus *Hae*III-sites. Thus the four intermediate fragments between a monomer and a dimer can be fragments with one end corresponding to the beginning of the consensus, while the other end corresponds to the *Hae*III site in the next monomer. The fact that oligomer forms are predominant in a partial hydrolysis and none of the sequenced clones overlaps the end and beginning of the consensus indicates that the enzyme cleaves more readily the sites located in the beginning and end of the consensus. So, the consensus length may be of 181 bp rather than 153 bp.

Oligomer forms of the C ichangensis stDNA repeating unit have a reduced mobility in polyacrylamide gel

As mentioned above, the latest studies have proved existence of bends in DNA molecules due to the A-tracts. One of the experimental proofs of existence of bends in DNA molecules is a decreased mobility of DNA fragments in polyacrylamide gel, while the same molecules move in accordance with their length in agarose gel. We have compared the mobilities of oligomers of the *C ichangensis* stDNA repeating unit in polyacrylamide and agarose gels. The dimer and trimer moved in the 2% agarose gel like 360- and 540-bp long fragments; meanwhile, they showed a slower mobility in the 3.5% polyacrylamide gel. For instance, the trimer moved in the 3.5% polyacrylamide gel as slow as the 587-bp long fragment of pBR322. The retardation became even more obvious with longer fragments (fig 4).

Models of stDNA tertiary structure

We have calculated the coordinates of monomer, tetramer and octamer axes of the repeating unit in various stDNAs, and built their two-dimensional projections using the BEN program by Eckdahl and Anderson (1987). The program is based on Ulanovski and Trifonov's 'wedge' model which suggests that the dinucleotide ApA causes a 8.4° bend in the roll and a 2.3° one in the tilt directions [22]. The dinucleotide TpT causes a reverse effect only in the tilt, but not in the roll direction. Besides, the program calculates the ENDS ratio showing the relation of the contour length of a given molecule axis to the distance between the ends.

The two-dimensional projection of the monomer and octamer axes of the *C ichangensis* stDNA is shown in figures 5 and 6. It follows from the figures that the slight bends observed in the monomer induce formation of a coiled form in the stDNA hypothetical model during transition to a higher oligomer form. The ENDS ratio shows the periodic maxima to reach the 1.1 value with about 200-bp intervals. As for the AT-content, its amount along DNA chain varies in narrow range, remaining below 40%M (fig 7a).

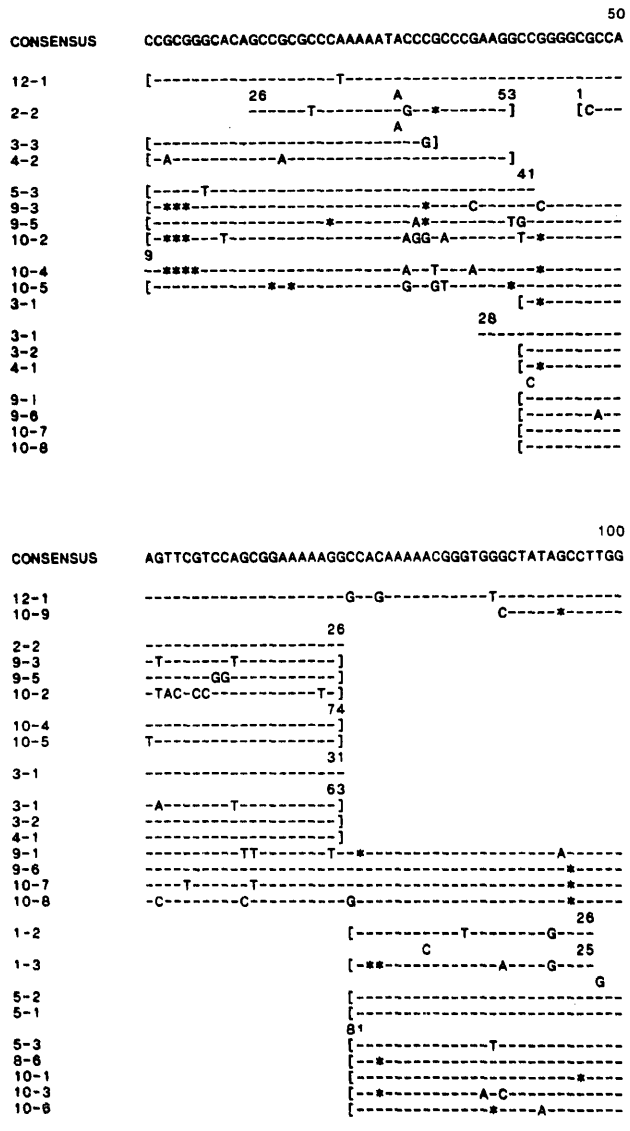


Fig 2. The sequence of the *C ichangensis* stDNA repeating unit. The consensus has been obtained from analysis of 40 clones. The first figure in the clone designation corresponds to the band number on the electrophoregram, from where the DNA was isolated. The second figure stands for the clone number from the given band. The letter above the line means that the given nucleotide follows the symbol above which it is. The asterisk means deletion. Brackets show the beginning and end of clones. If a clone consists of two consensus fragments standing far apart, no brackets are put at the separation points. A figure designing the nucleotide position stands above the beginning and end of each fragment of the 'divided' clone.

Models were built for other stDNAs whose structure had been studied well. These were the mouse and AGM stDNAs. Figures 5 and 6 show the two-dimensional projections of monomer and octamer images of these stDNAs, and their ENDS ratio and average AT-composition (fig 7). These data indicate that the form of the models built for the mouse and AGM stDNA octamers resembles that of the *C ichangensis* stDNA. Differences are observable only in the solenoid parameters and identity period.

The solenoidal form is particularly distinct in the octamer of the repeating unit of the mouse stDNA. The 1872-bp long octamer axis coils twice. The ENDS ratio shows clear maxima exceeding the value 1.2 with about a 150-bp long interval.

Discussion

stDNA of citric plants are very convenient to study. First, their high content in a genome, occasionally reaching 20% or more, and second, their high level of

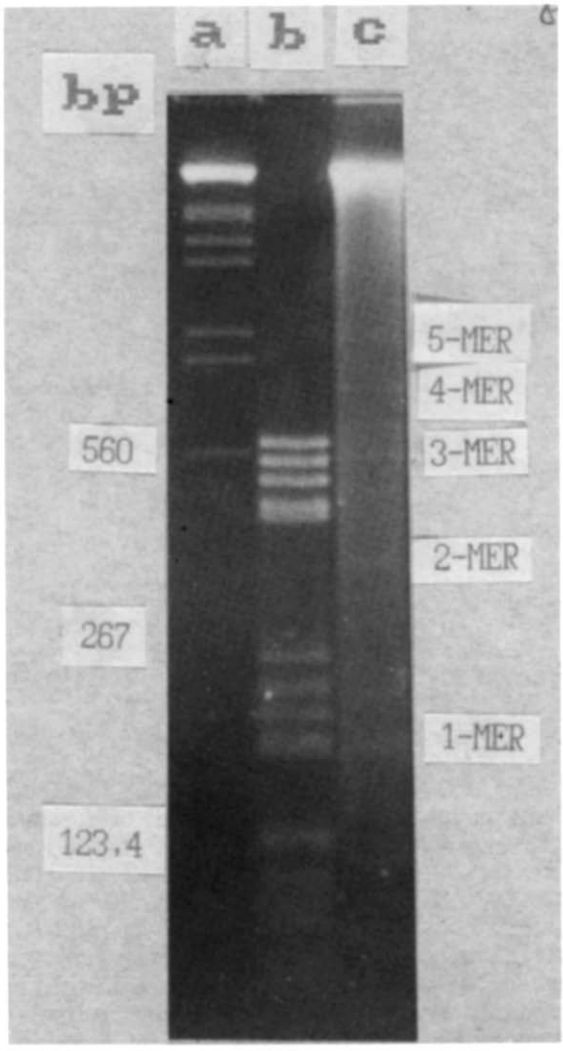


Fig 3. Gel-electrophoresis of products of the *C ichangensis* stDNA partial hydrolysis in a 2% agarose gel. The hydrolysis condition: 1 unit of *Bsp*I/μg of stDNA, 10 min, 37°C. Lane a: (*Hind*III + *Eco*RI)-digest products of bacteriophage λ. Lane b: *Hae*III digest of pBR322. Lane c: *C ichangensis* stDNA *Bsp*I digest products. The monomer, dimer etc of the repeating unit are designated respectively.

GC-content, 65–70%, is the highest among all the studied plants. Citrus stDNA were discovered by Ingle *et al* [11]. They succeeded in isolating a pure stDNA of *C sinensis* and indicated to its ability of a rapid reversible renaturation. Later on detailed studies of citrus stDNA were published [1, 3, 18].

In this work we attempted to define the primary structure of the *C ichangensis* stDNA repeating unit. The 181-bp long repeating unit revealed four pentanucleotides of adenine residues at positions 21–25, 65–69, 77–81 and 173–177.

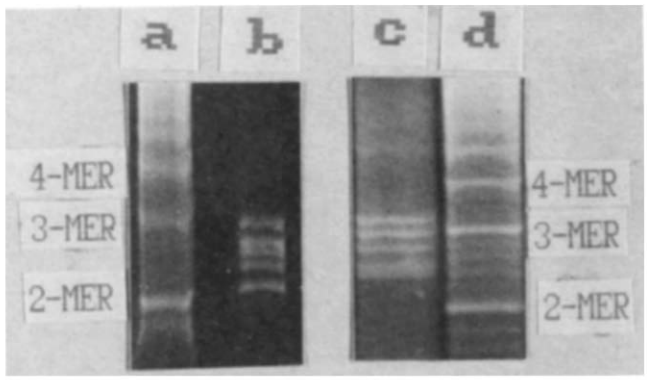


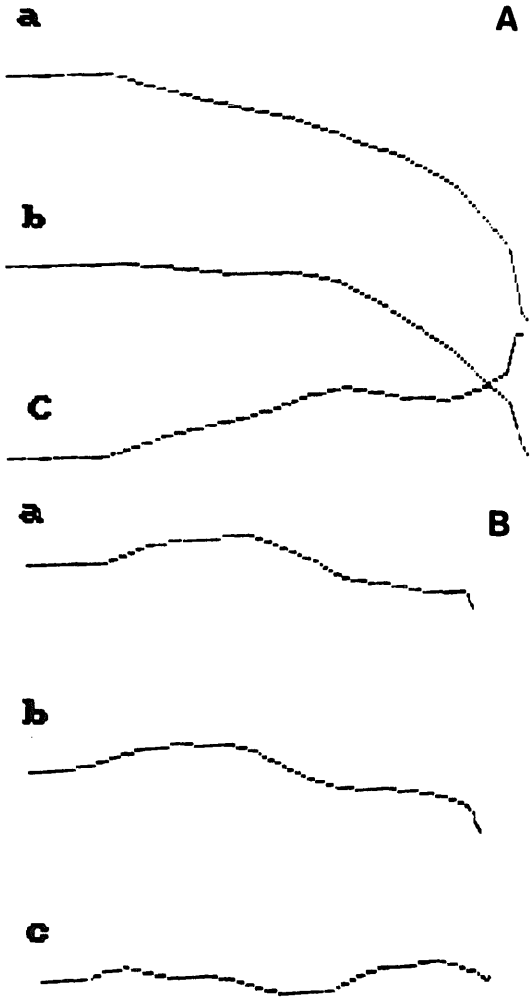
Fig 4. Gel-electrophoresis of oligomers of the *C ichangensis* stDNA. A: comparison of mobilities of the stDNA oligomers (a) and *Hae*III-digest of pBR322 (b) in 3.5% polyacrylamide gel. B: Comparison of mobilities of the oligomers (d) and the *Hae*III digest of pBR322 (c) in a 2% agarose gel. The oligomers are numbered. The length of the pBR322 *Hae*III restriction fragments is 434, 458, 504, 540 and 587 bp.

An examination of cloned fragment structure has revealed that the *C ichangensis* stDNA is diverged to a considerable degree, just as many other studied stDNAs. Deletions and insertions are observed as well as transitions and transversions. Some of the clones are made up of separated consensus fragments, indicating recombinant processes in the evolution of these molecules.

A partial hydrolysis of the *C ichangensis* stDNA, using *Bsp*I, allowed to detect the oligomer forms of the stDNA repeating unit. Experiments on the comparative mobility of oligomers in agarose and polyacrylamide gels demonstrated a certain deceleration of mobility in polyacrylamide gel, thus proving the existence of a slight bend in the repeating unit.

The stDNA problems are closely intertwined with the constitutive heterochromatin problem. The compact state of the constitutive heterochromatin seems to be due to the structure of the stDNAs, since no proteins characteristic only of heterochromatin and absent in euchromatin, have been detected [147, 24]. On the other hand, Strauss and Varshavsky isolated HMG-like stDNA-binding protein form crude extracts of AGM cells. Experiments on stDNA-containing nucleosome reconstitution, proceeding from a pure DNA and core histons have proved that the nucleosome position characteristic of the stDNA-containing chromatin is due to the independent contribution of many different DNA-histone contacts in an additive feature [15].

A hypothesis set forth in 1986 and described in greater detail in a later publication, suggested that the compact state of heterochromatin must be due to the



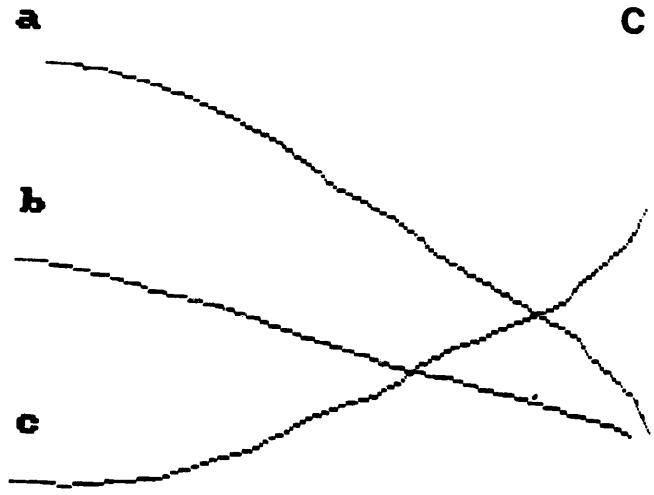


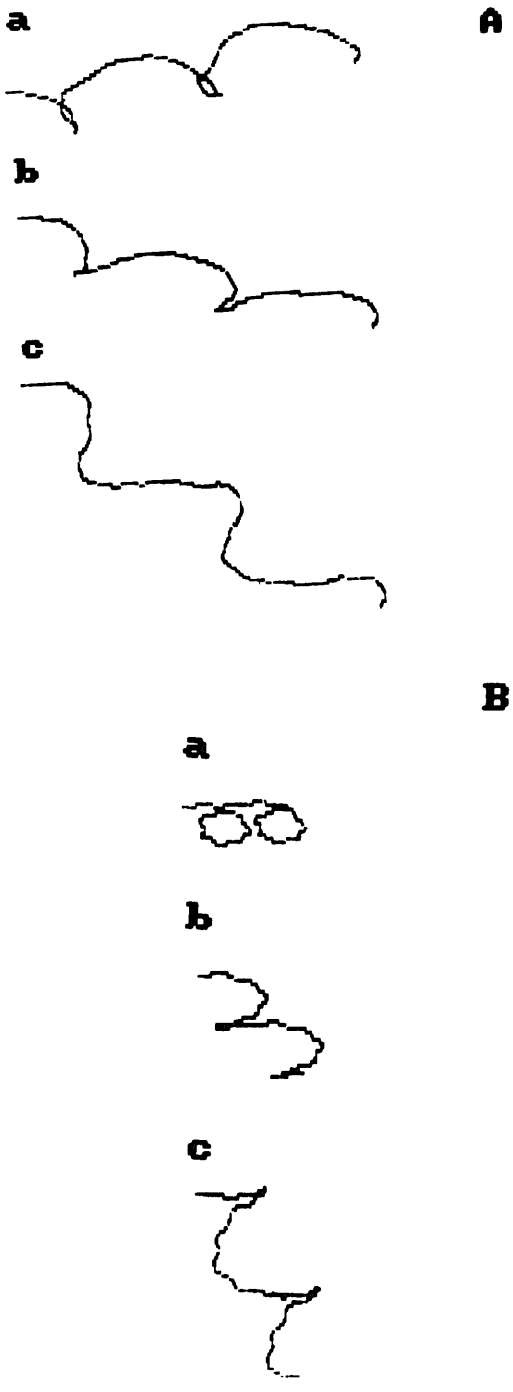
Fig 5. A two-dimensional projection of the stDNA monomer axis calculated by the BEN program. The projections are obtained by rotating the molecule axis by 0° (a), 72° (b) and 144° (c). **A**: *Citrus ichangensis*; **B**: mouse; **C**: AGM.

tertiary structure of stDNAs which might be caused by certain oligonucleotides specifically arranged along the stDNA chain [2, 3]. The tertiary structure must have been caused by the repeating arrangement of the stDNA. If the repeating unit has a bend making the repeat ends deviated from the straight line, then any DNA molecule consisting of such tandemly arranged units ought to form a solenoid-like structure which we will term as 'coiled double helix' (CDH).

The latest experimental results on DNA bendability supported that hypothesis and encouraged further experiments to check it.

In order to confirm the hypothesis concerning the tertiary structure, three different stDNAs were used: *C ichangensis* stDNA with 65% of GC-content and two AT-rich stDNAs of the mouse and AGM. As the bending in DNA molecules increases with the AT-content we considered it basically important to present material on GC-rich DNA. As for the *C ichangensis* stDNA, its oligomer mobility in polyacrylamide gel was estimated; while for the mouse and green monkey stDNAs models were built on the basis of the primary structure known from the literature [10, 17]. A bend in the mouse stDNA monomer was reported [16], being expressed through a slower mobility in polyacrylamide gel. In a recent publication of Martinez-Balbas *et al* [13] the presence of a slight bend in mouse and AGM stDNA monomers was definitely proved, using the electron microscopy technique. The form of the monomers is in good agreement with the theoretically calculated one in this work.

We have built tertiary structure models for three stDNAs. Although the monomer forms were different, the tetramers and octamers revealed a likeness, *viz*, formation of a solenoid-like structure, with quite different parameters, though. There seem to be no conclusive data that could ensure building of an absolutely precise tertiary structure for DNA molecules on the basis of the primary structure. The initial conditions on which the BEN program is based may undergo changes, even quite serious ones. Thus, Calladine *et al* [5] mentioned that the angular parameters employed in Ulanovski and Trifonov's model and being basical in the BEN program, contradicted the work on $CGCA_6GCG$. However, any stDNA molecule made up of bent monomers must have the form of a coiled double helix. It is up to future investigations, including electron microscopy of the repeating unit oligomers, to define the real parameters of



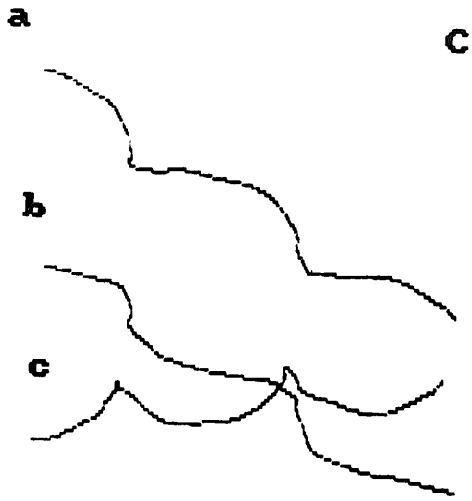


Fig 6. A two-dimensional projection of the stDNA octamer axis calculated by the BEN program. The projections are obtained by rotating the molecule axis. **A**: *Citrus ichangensis*; **B**: mouse; **C**: AGM. Rotation angles: **A**: 0° (a), 54° (b), 108° (c); **B**: 18° (a), 72° (b), 144° (c); **C**: 0° (a), 72° (b), 144° (c).

those molecules and confirm their agreement or disagreement with the presented models.

Besides, the theoretical computer analysis of DNA spatial structure is not a perfect one, as it is based on a static model, whereas the bend angle coincides in the order of value the bend angle fluctuation because of the thermal motion. Therefore, the stDNA has to contain an ensemble of different structures. An exceptionally solenoidal form of stDNAs has to be valid only for an absolute zero. With this in view the calculations in this paper have basically an illustrative character.

Should the existence of a specific tertiary structure in stDNAs be proven, the functionl role of those molecules will require a new approach. The role of stDNA may consist only in compactization of heterochromatin. A number of important functions are ascribed to heterochromatin as a macroscopic structure by cytogenetics. Heterochromatin can fulfill these functions only when it is in a compact structure typical of itself, while a compact structure of heterochromatin may be ensured by the stDNA tertiary structure. The existence of the CDH form will also facilitate a new approach to the study of stDNA-containing nucleosomes, their phasing, proceeding from the fact that DNA in those elements has a structure different from linear.

An indirect proof of the stDNA tertiary structure is decondensation of the constitutive heterochromatin during distamycin A treatment of animal cells, the distamycin A being known to straighten the bends in DNA molecule [8, 16].

The primary structure of the repeating unit has been identified in many stDNAs. Their analysis revealed A-tracts in most of them, therefore, all of them can

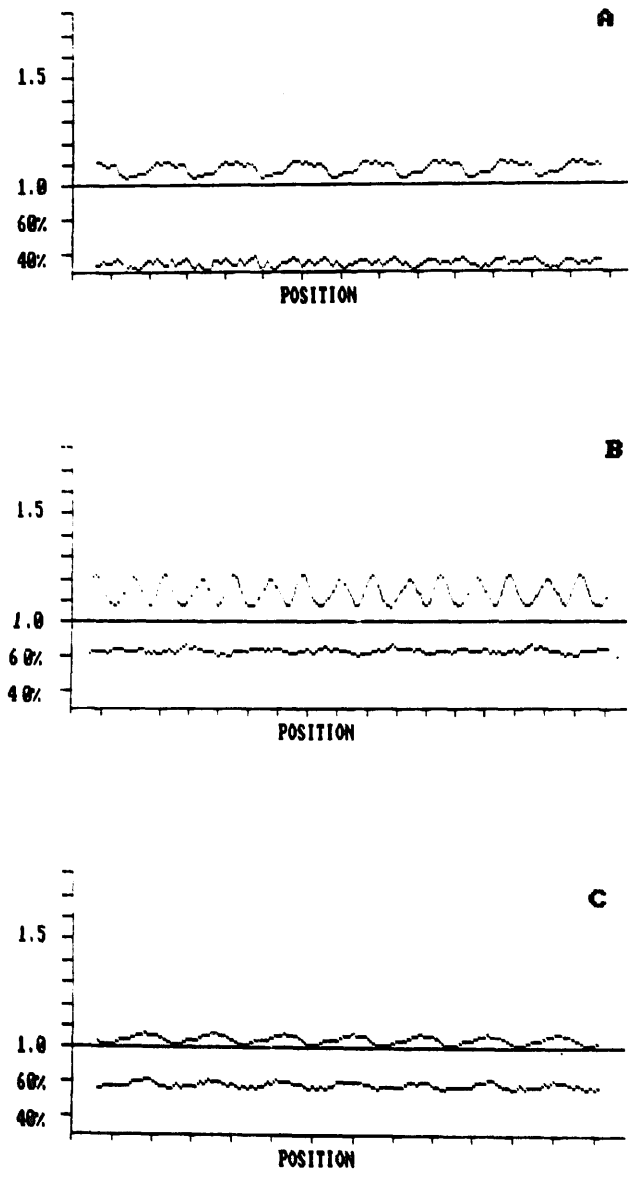


Fig 7. Plots of the ENDS ratio (upper part) and AT-content (lower part) as a function of length of stDNAs. **A**: *Citrus ichangensis*; **B**: mouse; **C**: AGM.

create the CDH-form, except some GC-rich molecules wherein A-tracts were not observed. It was, however, presumed that bends can be produced by other combinations of nucleotides different from the A-tracts [7]. Therefore, when the repeating unit contains a bend, those molecules ought to have a coiled form owing to the tandem arrangement of repeats in the satellite DNA.

Acknowledgments

The authors extend their gratitude to JN Anderson for the BEN program, to M Masiukov for his help in building the tertiary structure models, to L Gabodze and N Kvashilava for technical assistance and D Gvishiani for the English rendering of the manuscript of the paper.

References

1 Beridze T (1980) *Plant Sci Lett* 19, 325–338
2 Beridze T (1986) Satellite DNA, Springer-Verlag, 1
3 Beridze TG, Shengelia NI (1986) *Bioorg Khim* 12, 112–115
4 Beridze TG, Turishcheva MS, Kokonashvili GN, Masiukov MV (1988) *Abstracts of 14-th International Congress of Biochemistry*, FR 136
5 Calladine CR, Drew HR, McCall MJ (1988) *J Mol Biol* 201, 127–137
6 Diekmann S (1989) *In: Nucleic Acids a. Mol Biology vol 1* (F Eckstein, DMJ Lilley, eds) Springer, 138–156
7 Diekmann S, von Kitzing E (1988) *In: Structure and Expression, vol 3. DNA bending and curvature*, Adenine Press, 57–67
8 Dolfini SF (1987) *Chromosoma* 95, 57–62
9 Eckdahl TT, Anderson JN (1987) *Nucleic Acids Res* 15, 8531–8545
10 Hörz W, Altenburger W (1981) *Nucl Acids Res* 9, 683–696
11 Ingle J, Pearson GG, Sinclair J (1973) *Nature* 242, 193–197
12 Marini JC, Levene SD, Crothers DM, Englund PT (1982) *Proc Natl Acad Sci USA* 79, 7664–7668
13 Martinez-Balbas A, Rodriguez-Campos A, Garcia-Ramirez M, Sainz J, Carrera P, Aymami J, Azorin F (1990) *Biochemistry* 29, 2342–2348
14 Mazrimas JA, Balhorn R, Hatch FT (1979) *Nucleic Acids Res* 7, 935–946
15 Neubauer B, Linxweiler W, Horz W (1986) *J Mol Biol* 190, 639–645
16 Radic MZ, Lundgren K, Hamkalo B (1987) *Cell* 50, 1101–1108
17 Rosenberg H, Singer M, Rosenberg M (1978) *Science* 200, 394–402
18 Shengelia NI, Japiashvili MV, Kokonashvili GN, Mazin AL, Beridze TG (1986) *Mol Biol* 20, 798–810
19 Tan RKZ, Harvey SC (1987) *J Biol Struct Dynamics* 5, 497–512
20 Trifonov EN (1985) *CRC Crit Rev Biochem* 19, 89–106
21 Trifonov EN, Sussman JL (1980) *Proc Natl Acad Sci USA* 77, 3816–3820
22 Ulanovsky LE, Trifonov EN (1987) *Nature* 326, 720–722
23 Wernoslov SE, Kondrashov AS, Roytberg MA, Shabalina SA, Yureva OW, Nazipova NN (1990) *Mol Biol* 24, 524–529
24 Zhang XY, Horz W (1982) *Nucleic Acids Res* 10, 1481–1494