Sequence-dependent DNA curvature: conformational signal present in the main regulatory region of the rat mitochondrial genome

G.Pepe, G.Gadaleta, G.Palazzo and C.Saccone

Centro SMME-CNR and Dipartimento di Biochimica e Biologia Molecolare, Università di Bari, Italy

Received August 8, 1989; Revised and Accepted September 28, 1989

EMBL accession no. X14848

ABSTRACT

Theoretical analysis and experimental approaches by gel electrophoresis in retarding conditions allowed us to identify the presence of an intrinsic bending in the D-loop containing region of the rat mitochondrial genome. The curvature was located in the right domain of the sequence analyzed, between the origin of replication of the heavy strand and its promoter. The preliminary evidence of a specific recognition of the bent DNA with mitochondrial matrix proteins suggests a probable role of this DNA conformation in the duplication and/or expression of the mammalian mitochondrial genome.

INTRODUCTION

In the last few years new interest has been focused on the study of DNA conformation and the fascinating involvement of higher order structures in the regulation of molecular processes have come to light. Most interestingly, in several genomes of both prokaryotic and eukaryotic organisms, the presence and the biological importance of the sequence-directed DNA bending has clearly emerged. The exact role of this conformation is still to be completely understood but its involvement in some direct or even indirect protein interactions during recombination, replication and transcription processes is universally accepted (1-5).

For many years our group has been engaged in the study of the genetic and structural organization of the rat mitochondrial DNA (6).

Our attention was focused in particular on the main non coding region, called D-loop containing region since the nascent strand displaces the parental H-strand, thus creating a triple stranded loop structure. This portion of the genome was sequenced from different species of rats and accurately compared with its counterpart from other vertebrates (7,8). It ranges

from about 900-1000bp in mammals to about 2000 in amphibia (9) and contains the origin of the heavy-strand replication (Ori-H) and the promoters (HSP and LSP) for the transcription of the two strands (9-11).

In spite of its generally high variability, this part of the animal mitochondrial DNA contains small conserved sequences, CSBs (Conserved Sequence Blocks) and TASS (Termination Associated Sequences), that are thought to play a role in the regulation of the replication process. Another segment of about 200 bp, in the middle part of the region, revealed homology in all the analyzed systems. Its potential for a small ORF and for an antisense RNA is at present under investigation by another group from our laboratory (7,12-14).

Moreover we have found that in rat, as well as in all other vertebrates, defined tracts of the right and left D-loop domains may well be folded up in tRNA-like secondary structures which, owing to their high conservation, should also play a role in the regulation of the functions that take place there (7,8). Thus this segment of the animal mitochondrial genome, which does not possess direct information for known products, appears to be rich in signals at level of both primary and secondary structures.

Furthermore, as concerning the base content, the D-loop containing region can be divided in three domains. The central domain is the mitochondrial DNA tract with the highest G-C content. The left and right domains are instead particularly rich in As and Ts(8).

The analysis reported here reveals that the rat D-loop containing region displays in the right domain an (An-Tn) motif, which, according to actual views (1-4), is consistent with a sequence-dependent DNA curvature, i.e. with a tertiary structure of the DNA. The rest of the region contains An and Tn stretches, but not in periodical phase.

In addition, this paper reports the experimental proof of a DNA curvature in the right domain of this region, giving the first evidence for such type of conformation in the mitochondrial DNA from an animal.

Preliminary evidence for a specific interaction of the bend

DNA with mitochondrial matrix proteins is also shown.

MATERIALS AND METHODS

Source of DNA

The region we are interested in is contained within the smaller BamHI subfragment (2863bp long), obtained from the EcoRI-A fragment of *Rattus norvegicus* mitochondrial DNA (mtDNA). This part of the genome, that we called Bam2, was cloned in pUC8 and used either as a recombinant molecule or as a pure mitochondrial fragment, after excision from the plasmid. In all our experiments useful restriction enzymes were used and the products of the digestions were subjected to electrophoresis in retarding condition.

Mono and two-dimensional polyacrylamide gel electrophoresis

Electrophoresis was performed, according to Mizuno (5) with some modification, in 6% PAA (30:0.5, monomer/bis) in 40mM Tris acetate, 5mM Na acetate, 1mM EDTA, pH7.4. The slab gels (20x25 cm x 2mm thick) were run at 200V in the same buffer maintained at 60°C or at 4°C in a temperature controlled hand-made apparatus. After gel staining, the patterns obtained at different temperatures were compared. In other experiments, the slot containing the samples run at 60°C was directly cut from the gel, without staining (first dimension), and laid on another gel with the same composition, to be run at 4°C (second dimension).

Construction of head-to-tail dimer

The construction of the plasmid containing a tandem dimer, described in Fig.5, was carried out as follows. A Blue Scribe recombinant plasmid, containing the AccI-BglI fragment (250bp) cloned in the SmaI site was used as recipient (it was kindly provided by Dr. Shisa'). After digestion with BamHI and filling in, the recombinant was cut with AccI (both enzymes have the site in the polylinker flanking the insert). It was then ligated with the purified AccI-BglI fragment extracted from the same recombinant by the double digestion EcoRI (followed by filling in)x AccI. In this way the fragment was inserted in the desired head-to-tail orientation, as verified by sequencing.

Fig.1 Organization of the cloned Bam2 fragment.

a) Physical and genetic organization.
b) Primary structure depicted on a cylindrical projection (10bp/turn) with arrows indicating the direction of the projection. Only the L strand of the DNA helix is shown for clarity. The An and the Tn stretches ($n \geq 3$) are plotted with filled bars and empty bars, respectively.

Protein-DNA binding assay and gel retardation

Proteins from different mitochondrial compartments were prepared as described by Greenawalt (15). DNA-binding assays were performed under the following conditions. 5ng of $\gamma^{32}\text{P}$ 5'end labeled AccI-BglI fragment were incubated with 12 μg of each protein fraction and 0.5 μg of pUC8 or other plasmid DNA for 20' at 28° C in 10mM Tris-HCl pH7.9, 10mM MgCl₂, 40mM KCl, 5% glycerol and 0.5 mg/ml bovine serum albumin. The samples were immediately loaded in a vertical 4% PAA gel in 90mM Tris, 90mM boric acid, 2.5mM EDTA and electrophoresed at 4°C, at 220V, for 4h. Gels were dried and autoradiographed.

RESULTS

Analysis of sequence periodicity in the Bam2 fragment

The entire sequence of Bam2 fragment was carefully checked in order to identify the presence of a characteristic motif consistent with the DNA curvature. In this regard we considered both An and Tn tracts ($n \geq 3$) since, according to the studies of many authors, the periodic phasing of the mixture of these two tracts can produce intrinsic bending (16,17).

Fig.1 shows a schematic representation of the Bam2 region, cloned in pUC8, with its gene organization and the restriction sites used in our experiments (a). Searching of the An-Tn periodicity reveals the presence of three major stretches with a medium phasing of 10.5bp in the right portion of the D-loop containing region. The position of these blocks is shown at the bottom of the figure (b), where the sequence is depicted as a cylindrical projection having 10bp per turn.

Fig.2 presents the position of the structural elements already known within the nucleotide sequence of the D-loop containing region. The identified A-T rich blocks are located close to each other in a small portion between Ori-H and HSP site. In this region we have noted the presence of a perfect mirror symmetry CATAAATAT[>]TTATAAATAC between the positions 633 and 652, which is well conserved among vertebrates. Such symmetric structures are considered as functionally meaningful sequence elements often associated with regulatory regions, and their probable signal role in interactions with other macromolecules has been suggested (18).

We have also analyzed the entire Bam2 fragment using the algorithm of Plaskon and Wartell which assigns a numerical curvature score based on the phasing of (An-Tn) tracts to a DNA sequence, following a set of rules derived from gel mobility studies (19). The authors find that the high transcription rate promoters tend to have high curvature score in their upstream region (about at -150bp) and suggest that curved DNA structure is a characteristic feature of regulative regions and may influence the promoter strength.

Our analysis on the entire Bam2 sequence showed high curvature score particularly in the A-T rich portion of the D-loop containing region depicted in Fig.1b and in Fig.2. It has a curvature score=6.75, quite close to the values reported by Plaskon for regions upstream highly active promoters. Indeed this sequence, upstream the HSP, seems to be a good candidate for a DNA curvature.

Furthermore our D-loop containing region was analysed by the De Santis et al. computational model (20). The result supported

the presence of a curvature in this area, with a maximum around the position 600 of the sequence in Fig 2 (data not shown).

Experimental detection of curved-DNA

An anomalously low electrophoretic migration can reveal the presence of a curved DNA structure. Bending is increased at lower temperatures, whereas high temperatures tend to reduce the abnormal migration (16,17,21).

Based on this property, we performed a series of experiments in which the isolated Bam2 fragment or the entire pUC8-Bam2 recombinant were digested with different enzymes and analyzed on 6% PAA monodimensional gels at 4°C and at 60°C. In another series of experiments the digestion products were run through a two-dimensional electrophoresis, as described in Materials and Methods.

The strategy used to identify the sequence responsible for the bending and to place its locus more accurately was the following. We obtained smaller subfragments, in which the region of our interest (An-Tn blocks) was progressively better defined, by double digestions performed by HinfI and, as the secondary enzyme, HapII, Sau3A, BglI, TaqI, HincII and AccI, respectively (see the physical map in Fig.1). A short sequence of 250bp, which includes the blocks I and II, where we theoretically identified the curvature, derived in a separate experiment by BglIAccI digestion.

The one-dimension electrophoresis at different temperatures and particularly the two-dimension gels clearly showed the fragments with more or less evident retardation (Fig.3). The sequence that theoretically might have the bend structure was obviously contained in various partially overlapped subfragments, depending on the digestion, which exhibited different degrees of electrophoretic anomaly. The phenomenon of the retardation appeared to a major extent in the fragments which covered the right domain of the D-loop containing

Fig.2 Nucleotide sequence and structural organization of the D-loop containing region. Within the sequence, the structural elements conserved in vertebrates, CSBs and TASSs, are defined by the brackets and the promoters are located. The An and Tn blocks are underlined. On the bottom the tRNA-like structures present at the 5' and 3' ends of the region are shown.

Fig.3 Two dimensional gel electrophoresis in 6% PAA at 4°C of the restriction products of pUC8Bam2. From A to D: recombinant digested by HapIIxHinfI; Sau3AxHinfI; BglIIxHinfI and BglIIxAccI, respectively. Arrows label the fragments within the D-loop containing region, which show abnormal mobility. Asterisks mark the fragments corresponding to part of the 12S rRNA. Circles mark the fragments corresponding to part of cyt b.

region, between the 5' terminus of the H-strand replication and the HSP and particularly in the BglI-AccI fragment.

A measure of the retardation was also obtained by the R factor, i.e. the ratio of the apparent size (at 4°C) versus the actual size (from the sequence analysis). Table I summarizes the data obtained by the gel electrophoresis experiments, reporting the real size and the retardation,

TABLE I. SCHEMATIC DATA FROM GEL ELECTROPHORESIS

SUBFRAGMENTS	POSITIONS	REAL DIMENSIONS	APPARENT DIMENSIONS	INCREASE (%)	R
<i>AccI-HinfI</i>	782 945	163	162 $\pm$ 10	0	1,00
<i>TaqI-TaqI</i>	433 602	170	170 $\pm$ 10	0	1,00
<i>BglI-AccI</i> $\leftarrow$	532 781	249	310 $\pm$ 10	24,5	1,24
<i>HincII-HinfI</i>	665 945	280	280 $\pm$ 10	0	1,00
<i>TaqI-HinfI</i>	603 945	343	380 $\pm$ 10	11	1,10
<i>BglI-HinfI</i>	532 945	414	510 $\pm$ 10	23,5	1,23
<i>Sau3A-HinfI</i>	461 945	485	600 $\pm$ 10	24	1,23
<i>HapII-HinfI</i>	319 945	627	750 $\pm$ 10	19	1,20
<i>HincII-HinfI</i>	-88 664	710	700 $\pm$ 10	0	1,00

expressed as R ratio and percentage of the increase in size. The R factors in our experiments are comparable with the values generally reported for natural DNAs. Fig.4 shows, in diagrammatic form, the electrophoretic behaviour of all the subfragments of the right domain of the D-loop containing region, which potentially contain the bend. This allowed us to locate the center of the curvature in the position marked with the asterisk.

Both the theoretical and experimental analysis described above suggest that also other portions of the mtDNA Bam2 fragment might be curved or might contain a particular conformation able to confer electrophoretic retardation.

The fragment *HinfI-HinfI* of 1020bp (real size), containing part of the 12S rRNA gene, shows anomalously low migration (Fig.3). Also this sequence has, at some point, in phase An-Tn tracts, which could be responsible for a bend in DNA and possess a high curvature score value.

A second case we want to mention is that of another fragment *HinfI-HinfI* and its subsequence *Sau3A-HinfI*, where the portion of the cyt b contained in Bam2 is located (see Fig.1). These fragments show electrophoretic retardation, as appears in Fig.3, but they do not possess the characteristic An-Tn motif and in fact they have a very low curvature score (see Discussion).

Another approach to a more accurate identification of the

Fig.4 Diagrammatic form of the electrophoretic behaviours. All the subfragments, having Hinfi at one end, potentially containing a bend, are plotted. The presumptive center of the curvature is indicated by the asterisk.

bending locus is the circular permutation analysis described by Wu and Crothers (22). Since, in the experiments above described, the presence of curved DNA was found to be limited to the BglI-AccI fragment (250bp long), we focused our attention on it. Unfortunately, this fragment has an unsatisfactory number of restriction sites, so that a complete permutation analysis was impossible. Nevertheless we constructed the head-to-tail tandem dimer of this fragment, as described in Materials and Methods and in Fig.5, and we used it in the following comparative experiment. Fig.6 shows the electrophoretic behaviour (in the retarding conditions) of the cloned monomer,

Fig.5 Construction strategy of the head-to-tail tandem dimer of the AccI-BglI fragment, cloned in Blue Scribe digested by *SmaI*.

after digestion by *EcoRI* × *BamHI* and of the cloned dimer, after digestion by *HincII* (slot A and C respectively). The first digestion produces a monomeric form which corresponds to the BglI-AccI fragment of the D-loop containing region (from the position 532 to the position 781 of Fig.2), plus some bases of the vector (270bp total size). Also in this case, as in the gels shown in Fig.3, this sequence appeared clearly retarded. On the other hand the retardation disappears in the

Fig.6 Gel electrophoresis at 4°C of fragments with the same size but with the presumptive center of bending at different positions.
A: monomer, derived from the Blue Scribe recombinant digested by EcoRI/BamHI (corresponding to BglI-AccI fragment of Bam 2).
C: monomer, derived from the tandem dimer, digested by HincII.
B-D: pBR 322 digested by AluI and by HaeIII, as markers, respectively.

monomer resulting from the HincII digestion, which has the same size. According to Wu and Crothers, this indicates that the center of the curvature is located on the middle of the BglI-AccI fragment, very close to the HincII site: digestion at this site destroys the bending and the HincII-monomer loses its retardation.

All these experiments allowed us to define quite precisely the locus of the bending between block I and block II of Fig.1.
Study of the biological function of curved DNA

In order to investigate whether our DNA bending represents a possible signal for protein binding, we have carried out the

Fig.7 Protein-DNA binding assay by gel retardation experiments. Results with different compartment proteins are shown in slots from A to D: Inner membrane; External membrane; Intermembrane; Matrix. F) indicates the position of the free BglI-AccI fragment. C) indicates the retarded bands, corresponding to the protein-DNA complexes, which appear only in the lane of the matrix.

experiment shown in Fig.7, which demonstrates that the fragment BglI-AccI is recognized by the matrix proteins and not by other mitochondrial protein fractions. The interaction clearly produces complexes which show different dimension, probably because of their different degree of specificity.

DISCUSSION

Bending is a phenomenon which has been widely studied in the trypanosome kinetoplast (22-24). Recently it has been found also in the yeast mitochondrial genome, although, in this case, it seems to be induced by the RNA polymerase-promoter binding (25).

Nobody, however, at our knowledge, has described a sequence-dependent curvature in animal mitochondrial DNA.

It is well known that this mitochondrial genome has a very compact gene organization, which allows the presence of only a main non coding region, called in vertebrates D-loop containing region. Since the promoters of both strands and the origin for the synthesis of the H-strand are located here, this region

Fig.8 Possible secondary structures in the cyt b containing fragment of Bam2. Asterisk marks the position of the point mutation which can be responsible of the modification in the nucleic acids and protein structure.

appears of great interest for studying the regulatory mechanisms operating in the organelle replication and expression, which should be under direct nuclear control. In fact the enzymes responsible for mtDNA replication and transcription, as well as the ribosomal proteins and the factors involved in genome expression, are all coded for by nuclear genes and translated at level of cytoplasmic ribosomes.

In this paper we report the first experimental proof of the presence of an intrinsic bending in the mtDNA from a mammal. The strategy used to determine the position of the curvature allowed us to locate it quite precisely within the BglI-AccI fragment of the D-loop containing region, near the HincII restriction site. With respect to the other primary and secondary structural features previously detected in this area, the bending appears limited to a sequence upstream the HSP site and closely linked to the CSB2. The genome tract spanning between the Ori-H and the two promoters is very likely involved in recognition of proteins responsible for mtDNA synthesis and transcription.

The presence of a bend structure in this particular segment of the main regulative region of the molecule suggests its possible functional role during the transcription and/or

duplication process. For example, it has been proposed that such a particular conformation, which is often present in the promoter areas of *E. coli*, could facilitate the initiation step of the transcription (19). It might increase the RNA polymerase-DNA contacts and induce the transition to the open complex (26-28), or it might "per se" provide an additional protein-DNA interaction, required to confer the specificity or the right polarity to the transcription (like the TF1 factor found by Clayton in human mitochondria (29,30)). In mitochondrial genome, transcription and duplication are strictly connected, so that a regulatory signal could influence both processes.

Our gel retardation experiment seems to support the idea of a specific interaction between curved DNA and mitochondrial proteins. Since the latter belong to the matrix compartment, where all proteins are coded for by the nucleus, such interaction should represent a possible communication pathway between the two genetic systems. Further experiments are necessary to characterize the nature of the bound proteins and the features of this protein-DNA interaction.

We want to make some short comment about the other fragments of Bam2, which showed electrophoretic retardation, such as the 12S rRNA and cyt b containing sequences. Since static bending in DNA has been normally found in regulative regions, it is difficult to understand its presence inside a sequence coding for an rRNA. We cannot "a priori" exclude a possible regulatory site present in this region of the genome. Anyway we want to remember the importance of specific interactions of rRNA with proteins in ribosome assembly and the roles played by rRNAs for the correct contacts with the mRNA during translation. These functions might be mediated also by particular RNA conformations. As concerning the other fragment, we suggest that another particular structure, different from bending, might be the cause of its retardation. The fragment codes for the hydrophobic domain VI of the cyt b, a region highly conserved in all eukaryotes, involved in the binding of one of the heme groups. By a careful analysis of the sequence of this region in rat we found two contiguous inverted repeats able to form stable

hairpin-loop structures (Fig.8), which, in our opinion, should explain the migration retardation of the DNA fragment. Similar explanation for anomalous DNA behaviour was also suggested by Howell et al. (31). The authors report that, in mouse L-cells mutants, the increased resistance to some cyt b inhibitors depends on a single base alteration that, causing a Gly->Asp change, gives the protein a new charged residue that would distort its secondary structure. The authors suggest that the base substitution can create two alternative structures in the wild type and in mutant form, responsible for the different migration of the corresponding DNA fragment.

In conclusion the results presented in this paper suggest that in the D-loop containing region of rat mtDNA messages written in different language must co-exist, to give co-ordinate regulatory signals. Thus a tertiary structure, as that created by bending DNA conformation, together with the primary and secondary structures above mentioned, may be strictly involved in a fine regulation of the mammalian mitochondrial genome.

During the revision of this manuscript, the paper of Welter et al. appeared in the volume 17 of August on NAR, pp. 6077-6086. The authors presented evidence of a curvature in the Ori-L of human mtDNA and its specific protein binding, thus strongly supporting our study.

ACKNOWLEDGMENTS

We are grateful to Drs. P.De Santis and M.Savino, who have analyzed our sequence with their method. We thank Dr. S.Liuni for helping in computational work, Mr. N.Cataldo for photography and the student C.De Virgilio for her help in some of the experiments reported. We gratefully acknowledge Prof Trifonov for his helpful comments and suggestions during the revision of the manuscript.

This work was supported by M.P.I. grants.

REFERENCES

1. Trifonov, E.N. (1985) CRC Crit. Rev. Biochem. 19, 89-106.
2. Lilley, D.M.J. (1986) Nature 320, 487-488.
3. Eisenberg, H. (1986) TRENDS Biochem. Sci. 11, 350-351.
4. Diekmann, S. (1987) EMBO J. 6, 4213-4218.
5. Mizuno, T. (1987) Nucl. Acids Res. 15, 327-341.
6. Gadaleta, G., Pepe, G., De Candia, G., Quagliariello, C., Sbisa, E. and Saccone, C. (1989) J. Mol. Evol. 28, 497-515.
7. Brown, G.G., Gadaleta, G., Pepe, G., Saccone, C. and Sbisa, E. (1986) J. Mol. Biol. 192, 503-511.
8. Saccone, C., Attimonelli, M. and Sbisa, E. (1987) J. Mol. Evol. 26, 205-211.

-
9. Saccone, C. and Sbisa, E. (1989) In Tager, J.M., Guerrieri, F., Azzi, A. and Papa, S. (eds.), *Organelles of Eukaryotic Cells: Molecular Structure and Interactions*, Plenum Pub. Co., New York, in press.
 10. Clayton, D.A. (1982) *Cell* 28, 693-705.
 11. Clayton, D.A. (1984) *Annu. Rev. Biochem.* 53, 573-594.
 12. Walberg, M.W. and Clayton, D.A. (1981) *Nucl. Acids Res.* 9, 5411-5421.
 13. Doda, J.N., Wright, C.T. and Clayton, D.A. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6116-6120.
 14. Sbisa, E., Nardelli, M. and Saccone, C. (1988) *Gene* 72, 309-310.
 15. Greenawalt, J.W. (1974) *Methods in Enzymology* 31, 310-323.
 16. Diekmann, S. (1986) *FEBS Letters* 195, 53-56.
 17. Koo, H.-S., Wu, H.-M. and Crothers, D.M. (1986) *Nature* 320, 501-506.
 18. Trifonov, E.N. (1988) In Bock, H.H. (ed.), *Classification and Related Methods of Data Analysis*, Elsevier Science Publishers, North Holland, pp.57-64.
 19. Plaskon, R.R. and Wartell, R.M. (1987) *Nucl. Acids Res.* 15, 785-796.
 20. De Santis, P., Palleschi, A., Savino, M. and Scipioni, A. (1988) *Biophys. Chem.* 32, 305-329.
 21. Hagermann, P.J. (1985) *Biochemistry* 24, 7033-7037.
 22. Wu, H.-M. and Crothers, D.M. (1984) *Nature* 308, 509-513.
 23. Marini, J.C., Levene, S.D., Crothers, D.M. and Englund, P.T. (1982) *Proc. Natl. Acad. Sci. USA* 79, 7664-7668.
 24. Diekmann, S. and Wang, J.C. (1985) *J. Mol. Biol.* 186, 1-11.
 25. Schinkel, A.H., Groot Koerkamp, M.J.A., Teunissen, A.W.R.H. and Tabak, H.F. (1988) *Nucl. Acids Res.* 16, 9147-9163.
 26. Bossi, L. and Smith, D.M. (1984) *Cell*, 39, 643-652.
 27. Kadesch, T., Williams, R. and Chamberlin, M. (1980) *J. Mol. Biol.* 136, 65-78.
 28. Travers, A.A., Lamond, A.I., Mace, H.A.F. and Berman, M.L. (1983) *Cell* 35, 265-273.
 29. Fisher, R.P. and Clayton, D.A. (1985) *J. Biol. Chem.* 260, 11330-11338.
 30. Fisher, R.P., Topper, J.N. and Clayton, D.A. (1987) *Cell* 50, 247-258.
 31. Howell, N., Appel, J., Cook, J.P. and Hauswirth, W.W. (1987) *J. Biol. Chem.* 262, 2411-2414.
-