Isolation of a 25-kDa Protein Binding to a Curved DNA Upstream the Origin of the L Strand Replication in the Rat Mitochondrial Genome*

(Received for publication, January 7, 1996, and in revised form, February 28, 1996)

Gemma Gadaleta, Domenica D’Elia, Lara Capaccio, Cecilia Saccone, and Gabriella Pepe

From the Centro di Studio sui Mitocondri e Metabolismo Energetico-CNR and Department of Biochemistry and Molecular Biology, University of Bari, via Orabona 4A, 70126 Bari, Italy

The presence of a curved DNA sequence in the gene for the NADH-dehydrogenase subunit 2 of rat mitochondrial genome, upstream from the origin of the light strand replication, have been demonstrated through theoretical analysis and experimental approaches. Gel retardation assays showed that this structure makes a complex with a protein component extracted from the mitochondrial matrix. The isolation and purification of this protein is reported. With a Sepharose CL-6B and magnetic DNA affinity chromatography a polypeptide was purified to homogeneity having 25-kDa mass as shown by gel electrophoresis. To functionally characterize this protein, its capability to bind to other sequences shown by gel electrophoresis; PCR, polymerase chain reaction; EMSA, electrophoretic mobility shift assay; bp, basepair(s); ND2, NADH dehydrogenase subunit 2.

Mammalian mtDNAs have two separate and distinct replication origins. The origin of the heavy strand replication (Ori-H), inside the so-called D-loop containing region which also contains the promoters for the transcription of both strands (H and L strand promoters); and the origin of the L strand synthesis (Ori-L), nestled within a cluster of five tRNA genes at two-thirds of the molecule, clockwise with respect to the Ori-H (for review, see Ref. 1). Initiation of the L strand synthesis only occurs after Ori-L is exposed as a single-stranded template. The outline of mtDNA replication in mammals has been elucidated (2) and the γ-DNA polymerase has been identified and partially purified in humans (3). Several other proteins take part in this process with different roles. An RNA polymerase and an RNase-MRP have a role in the initiation of RNA priming, the very first step of the new H strand synthesis; a DNA primase is involved in priming the L strand synthesis; unwinding factors have helicase or topoisomerase functions. Other protein candidates for process regulation by specific targeting to mtDNA sequences associated with replication have been identified from different eukaryotes, including mammals (for review, see Ref. 4). For many years we have deeply studied the genetic and structural organization of the mt genome in mammals, especially in rat (5, 6).

Moreover, we analyzed the rat mtDNA sequence searching for primary or higher order structures that could be recognized by functional proteins. Namely, we checked for sequence-dependent DNA curvatures, frequently present in the regulatory regions of a genome, involved in protein binding (7–11). Then we concentrated our efforts on the isolation, purification, and characterization of the proteins able to interact with these structures. We have already reported the presence of a curved DNA in the D-loop containing region, located close to Ori-H. This structure acts as a binding site for a nucleoeyrly 67-kDa polypeptide isolated from the matrix and we suggested that it is a regulatory factor promoting the first step of the H strand replication through the RNA → DNA transition (12, 13). Thereafter we focused our attention on the region of the genome containing the other replication origin, the Ori-L.

In this paper we report the localization of a curved DNA sequence inside the gene for the NADH dehydrogenase subunit 2 (ND2) upstream Ori-L. We also present the purification to homogeneity and characterization of a 25-kDa protein unknown so far, able to bind this DNA structure specifically. This is one of the few proteins, nuclearly coded, purified until now from higher eukaryotes which should participate in the biogenesis of mitochondria and for which we suggest a possible physiological role.

MATERIALS AND METHODS

Source of DNA—Several recombinants were constructed by cloning different portions of the Rattus norvegicus mtDNA around the Ori-L. The organization of this mtDNA region, its restriction map, and the position of the cloned fragments are shown in Fig. 1. The larger recombinant, called BS-C, was a Blue Scribe plasmid containing the entire 2982-bp long EcoRI–C restriction fragment as the insert. From BS-C, another recombinant was derived inserting the EcoRI–HindI subfragment, called C1, in the pUC8 vector. This third recombinant, called BS-A, was a Smal-digested Blue Scribe containing the 520-bp-long EcoRI–TaqI subfragment, inserted in the direction T3 → T7 of the promoters contained in the vector.

Curvature Analysis—The theoretical analysis was performed on the entire EcoRI–C sequence according to Plaskon and Wartell (14). For the experimental approach the above mentioned recombinants were digested with suitable enzymes and the restriction products underwent gel electrophoresis in retarding conditions, as described previously (12).

Electrophoretic Mobility Shift Assay (EMSA)—The DNA probe was the 520-bp insert excised by suitable digestion from the BS-A recombinant. Alternatively a smaller probe was obtained by PCR amplification of BS-A using a mitochondrial upstream primer (about 200 bp from the EcoRI site) and the T7 downstream primer. This probe was 350 bp long, out of which 320 bp were mtDNA. EMSA analysis was carried out as in Gadaleta et al. (13), by incubating the 32P-end-labeled probe with proteins prepared according to Greenawalt (15) and purified as described.

* This work was supported by Ministry of University and Scientific and Technological Research 40% and 60% grants and by CNR-Progetto Finalizzato Ingegneria Genetica. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, University of Bari, Via Orabona 4A, 70126 Bari, Italy. Tel.: 39-80-5443471; Fax: 39-80-5443317.

§ These authors contributed equally to this article.

1 The abbreviations used are: mt, mitochondrial; L strand, light strand; H strand, heavy strand; Ori-L, origin of the L strand replication; Ori-H, origin of the H strand replication; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; EMSA, electrophoretic mobility shift assay; bp, basepair(s); ND2, NADH dehydrogenase subunit 2.
RESULTS

Detection of the Curved DNA—The entire EcoRI-C sequence was carefully checked for the characteristic An-Tn stretches consistent with a curved DNA (7–12). This property was indeed present in several regions along the fragment. Plaskon and Wartell algorithm was applied which assigns a curvature number based on the phasing of the An-Tn tracts and allows prediction of the presence of a curved DNA in fragments with higher values. Plaskon and Wartell also suggested that the curved DNA is a characteristic feature of regulatory regions in a genome (14). Fig. 1 shows a schematic representation of the EcoRI-C fragment. The analysis showed a curvature score of 6.25, significantly high according to Plaskon evaluation, only in the second half of the ND2 gene upstream Ori-L. Thus, the region seemed to be a good candidate for containing a DNA curvature in rat mt genome.

The region was experimentally tested by gel electrophoresis. It is known that an anomalously low migration of a fragment can reveal the presence of a curved DNA and the retardation due to the bending phenomenon is increased at lower temperatures and reduced at higher temperatures (20–22). Thus, a number of experiments was performed in which the recombinant DNAs described under “Materials and Methods” were digested with different enzymes as shown in Fig. 1 and analyzed on a monodimensional gel electrophoresis (at 4 °C and at 60 °C) or on a two-dimensional gel (12). BS-C was double digested by EcoRI-HindI giving three subfragments of 1913, 849, and 220 bp, according to the sequence data (6). However, on polyacrylamide at 4 °C the 849-bp fragment migrated as a 913-bp long one with R = 1.1, where R is the ratio of the apparent size, obtained at 4 °C, versus the actual size (results not shown). This region (C1) was cloned in pUC8 and the derived recombinant was carefully analyzed to localize the curvature. It was double digested by TaqI-HindI giving a 976-bp subfragment with the apparent size of 1291 bp (R = 1.33). Fig. 2a shows the two-dimensional gel that highlights this retarded fragment consisting of 520 bp from the recombinant and a vector portion. To verify that the anomalous migration was due to a curvature present in the mtDNA sequence, the pUC8-C1 recombinant underwent EcoRI-TaqI digestion and the new products were analyzed by two-dimensional gel electrophoresis as shown in Fig. 2b. This experiment definitely localized the curvature in the 520-bp sequence of rat mtDNA inside ND2 gene from nucleotides 4252 to 4772. The fact that this curved DNA did not correspond to that found in human mt genome by Welter et al. (23) is discussed below. Finally, the sequence containing the curvature was cloned in the Blue Scribe vector giving the BS-A recombinant that was thereafter used in most of our experiments. Fig. 2c shows gel retardation performed with the shorter 350-bp sequence obtained by PCR and a vector portion. The DNA solution for the coupling to the labeled fragment. A reusable magnetic solid-phase was formed with a specific affinity for the target protein (10 μg of specific DNA fragment per mg of beads). The beads were then washed with TGED buffer, 50 mM NaCl (20 mM Tris-HCl, pH 8, 10% glycerol, 1 mM dithiothreitol, 0.01% Triton X-100, 50 mM NaCl) and mixed with the pooled fractions from the heparin-Sepharose active in the EMSA. The beads were overloaded so that the limiting number of available DNA sites were certainly saturated only by the specific proteins. The contaminating unbound proteins remained in the solution after the magnetic separation and were discarded. Finally, the bound protein component was eluted by TGED buffer, 1 M NaCl and characterized by 11% SDS-PAGE as in Laemmli (19).

849-bp fragment migrated as a 913-bp long one with R = 1.1, where R is the ratio of the apparent size, obtained at 4 °C, versus the actual size (results not shown). This region (C1) was cloned in pUC8 and the derived recombinant was carefully analyzed to localize the curvature. It was double digested by TaqI-HindI giving a 976-bp subfragment with the apparent size of 1291 bp (R = 1.33). Fig. 2a shows the two-dimensional gel that highlights this retarded fragment consisting of 520 bp from rat mtDNA and a vector portion. To verify that the anomalous migration was due to a curvature present in the mtDNA sequence, the pUC8-C1 recombinant underwent EcoRI-TaqI digestion and the new products were analyzed by two-dimensional gel electrophoresis as shown in Fig. 2b. This experiment definitely localized the curvature in the 520-bp sequence of rat mtDNA inside ND2 gene from nucleotides 4252 to 4772. The fact that this curved DNA did not correspond to that found in human mt genome by Welter et al. (23) is discussed below. Finally, the sequence containing the curvature was cloned in the Blue Scribe vector giving the BS-A recombinant that was thereafter used in most of our experiments. Fig. 2c shows gel retardation performed with the shorter 350-bp sequence obtained by PCR as described under “Materials and Methods.” This sequence had almost normal migration with respect to the marker fragments. We demonstrated by EMSA that both the 520-bp and the 350-bp fragments bound a nuclearily coded protein component extracted from the organelle matrix. These results are consistent with the presence of the curvature roughly in the middle of the 520-bp and near the 5’ terminus of the 350-bp fragment. Thus, both these probes were used in EMSA and
Protein-DNA Interaction Analysis — The functional role of the identified curved structure was studied. EMSA experiments were performed by incubating the selected DNA probe with the mt proteins extracted from the matrix and partially purified by heparin-Sepharose CL-6B. Fig. 3a shows the experiment with different amounts of proteins: with 10 μg the probe was completely saturated in the complex. When the same proteins were pretreated with proteinase K the complex disappeared (Fig. 3b) indicating the real protein nature of the component bound to the DNA. To evaluate the stability of the protein-DNA interaction, the dissociation rate of the complex was monitored by EMSA as described under "Materials and Methods" (Fig. 4). The in vitro estimated lifetime was about 3.5 h, when the complex appeared destabilized. Binding assays were also carried out in the presence of a competing ribonucleotide synthesized by using the 520-bp sequence of BS-A as the template and corresponding to the ND2-RNA or the ND2-anti-RNA. The competition causes the decreasing or disappearing of the complex and the increasing of the free probe. Fig. 5 shows the inhibition of the protein-DNA binding activity by both riboprobes, although the anti-RNA is less efficient. In Figs. 4 (lane 9) and 5 (lanes 3–5) intermediate complexes appeared, suggesting the interaction of more protein subunits at the DNA binding site.

Protein Purification — The result of the affinity purification performed as described under "Materials and Methods" is presented in the SDS-PAGE of Fig. 6. The polypeptide eluted by the activated beads is essentially pure, with a 25-kDa mass as measured by the marker. A purification of 26.3-fold with respect to the heparin chromatography was calculated by gel densitometry of the lanes loaded with the proteins before and after magnetic separation.

Binding Properties of the 25-kDa Polypeptide — As shown in Fig. 7a, the purified factor still retained a strong activity in EMSA, whereas the unbound contaminant proteins lost this property. In order to investigate if the 25-kDa factor recognizes any DNA curvature, the binding experiment was performed with the other curved region present in the rat mt genome upstream Ori-H (12). Indeed, this probe formed a complex, although it was a much less efficient target for this protein (Fig. 7b).

The binding capability of the 25-kDa factor was also tested on the two corresponding regions from the human mtDNA. In both cases a much lower intensity complex was observed (data not shown).

**DISCUSSION**

As far as the mtDNA replication and transcription is concerned the knowledge of regulatory factors is still poor and only a few protein factors have been purified (see Refs. 1 and 2, for review). Thus it is of great interest to determine the require-
Purification and Characterization of a Rat mtDNA Binding Protein

Here we provide evidence for another structural element at the level of Ori-L beside the stem-loop, namely a curved DNA which recognizes a nuclearly coded protein extracted from the rat mt matrix. We purified this protein to homogeneity and studied its possible functional role in the machinery of the L strand synthesis.

It has been reported by in vitro studies on human cells that the synthesis of the L strand is accomplished by the enzymes DNA primase and polymerase co-assisted by protein factors that interact with the stem-loop DNA structure of Ori-L. The evolutionary conservation of this secondary structure suggests it may play a functional role in vertebrate mt genomes (4). Furthermore, over the last years the importance of DNA higher-order structures and of specifically binding proteins in the regulation of molecular processes has emerged clearly. DNA regions with peculiar conformations, such as the curved DNA, located in proximity of the origins of replication, could play a role in this process, via the recognition and binding of proteins that facilitate the positioning of the multienzyme replicative complex.

In a previous study we demonstrated the presence in rat mt genome of a curved DNA inside the main regulatory region, the D-loop containing region. This structure, localized close to Ori-H, was a conformational signal able to bind a specific 67-kDa protein with a presumptive regulatory role (13).

In human mtDNA a curvature was localized inside the ND2 gene, 310 bp upstream of Ori-L, which bound a protein factor of about 100 kDa isolated from mitochondria and the authors proposed that this complex could promote the initiation of replication at the correct site (23).

In the present paper we demonstrate in rat mtDNA the existence of a curvature in a different sequence of ND2 gene, over 600 bp upstream Ori-L. It corresponds to the 520-bp fragment cloned in BS-A and near the 5' end of its 350-bp subfragment, obtained by PCR.

The binding protein factor was isolated to an excellent degree of purification by magnetic affinity and appeared in the SDS-PAGE as a single polypeptide of 25 kDa. However, Figs. 4 and 5, that showed the intermediate complexes, suggest that more protein subunits could be involved in the curved DNA interaction. The existence of different co-migrating polypeptides is excluded by preliminary sequencing data which reveal only one NH₂ terminus.

G. Gadaleta, D. D’Elia, L. Capaccio, C. Saccone, and G. Pepe, unpublished data.
The data discussed above strongly support the need of a curvature near Ori-L to bind a protein factor required in the L strand replication mechanism. The position of this structure seems not to be constrained strictly and it is likely to have different locations, in different species, not too far upstream of Ori-L.

In order to functionally characterize the 25-kDa protein, we performed a series of binding assays. The dissociation rate experiment indicated a strong stability of the complex in vitro, compatible with its functional role. Other experiments using different probes showed that the protein recognizes the rat curved DNA close to Ori-H, and the human curved DNA within ND2 gene. The low efficiency of these binding can be explained by a general capability of the protein to recognize a curved structure.

Furthermore, the two riboprobes synthesized by the curvature-containing DNA as the template, corresponding to the ND2-RNA and its antisense, were able to compete with the DNA probe in forming the complex with the protein component. The competition was complete when the higher amount of the ND2-RNA was used; the antisense only decreased the complex. The incomplete competition gave faster complexes (Fig. 5), again suggesting that a multimeric protein factor interacts with the curved DNA. The anti-RNA seems to interfere more generally with the complex through its secondary structure; whereas the binding of the sense-RNA to the protein seems to reflect a more specific recognition between the two molecules. Further studies will be necessary to clarify the functional role of these interactions.

Therefore, we can hypothesize that the replication of the L strand and the correct positioning of the DNA primase at the initiation site within Ori-L need an accessory protein, probably a multimeric one. A possible function for the 25-kDa protein in DNA replication is suggested by the proximity of its binding site to the L strand origin. By analogy with other genetic systems (25-27), this protein, when bound, could stall an advancing replication complex involving the newly synthesized H strand, by antagonizing the action of an associated DNA helicase (whose activity in mammalian mitochondria has been demonstrated so far in bovine (28)). This could facilitate the formation, on the displaced strand, of the priming structure for L strand initiation. Consequently, the local presence of DNA polymerase and other components of the replication machinery initiates efficiently the L strand synthesis.

In conclusion, the reported data are relevant in several respects. The importance of the structural elements in regulatory regions of mt genomes is stressed. The 25-kDa protein we isolated from rat mt matrix is the first purified factor that interacts at the level of Ori-L, with a presumptive regulatory function in the initiation of L strand replication. Finally, the different position of the curved structure in rat and human mtDNA and the isolation from the two systems of two proteins with different molecular weight offer new interesting topics to investigate mitochondrial and nuclear genome co-evolution in the biogenesis of mammalian organelles.

Acknowledgments—We thank Dr. E. Sbisa’s research group for kindly providing the riboprobes and human recombinant clones. We are grateful to Dr. D. Catalano for help in performing some experiments and D. Munno for his skilful technical assistance. Thanks are due to Dr. M. Lonigro for English revision of this manuscript.

REFERENCES
1. Saccone, C., and Sbisa, E. (1993) in Principles of Medical Biology (Bittar, E. E., ed) Vol. 1B, pp. 37–70, J. a Press Inc., Greenwich, CT
2. Wong, T. W., and Clayton, D. A. (1985) Cell 42, 951–958
3. Gray, H., and Wong, T. W. (1992) J. Biol. Chem. 267, 5835–5841
4. Clayton, D. A. (1992) Int. Rev. Cytol. 141, 217–232
5. Brown, G. G., Gadaleta, G., Pepe, G., Saccone, C., and Sbisa, E. (1986) J. Mol. Biol. 192, 503–511
6. Gadaleta, G., Pepe, G., De Candia, G., Quagliariello, C., Sbisa, E., and Saccone, C. (1989) J. Mol. Biol. 28, 497–516
7. Trifonov, E. N. (1985) CRC Crit. Rev. Biochem. 19, 89–106
8. Lilley, D. (1986) Nature 320, 487–488
9. Eisenberg, H. (1986) Trends Biochem. Sci. 11, 350–351
10. Diekmann, S. (1987) EMBO J. 6, 4213–4218
11. Mizuno, T. (1987) Nucleic Acids Res. 15, 327–341
12. Pepe, G., Gadaleta, G., Palazzo, G., and Saccone, C. (1989) Nucleic Acids Res. 17, 8803–8819
13. Gadaleta, G., D’Elia, D., Saccone, D., and Pepe, G. (1995) Gene (Amst.) 160, 229–234
14. Plaskon, R. R., and Wartell, R. M. (1987) Nucleic Acids Res. 15, 785–796
15. Greenawalt, J. W. (1974) Methods Enzymol. 31, 310–323
16. Carey, J. (1991) Methods Enzymol. 208, 103–117
17. Gabrielsen, O. S., Hornes, E., Korsnes, L., Ruet, A., and Øyen, T. B. (1989) Nucleic Acids Res. 17, 6253–6267
18. Gabrielsen, O. S., and Huet, J. (1993) Methods Enzymol. 218, 508–525
19. Laemmli, U. K. (1970) Nature 227, 680–685
20. Diekmann, S., and Wang, J. C. (1985) J. Mol. Biol. 186, 1–11
21. Koo, H. S., Wu, H. M., and Crothers, D. M. (1986) Nature 320, 501–506
22. Hagermann, P. J. (1985) Biochemistry 24, 7033–7037
23. Welter, C., Dooley, S., Zang, K. D., and Blin, N. (1989) Nucleic Acids Res. 17, 6077–6096
24. Waddell, W. J., and Hill, C. (1956) J. Lab. Clin. Med. 48, 311–314
25. Lee, E. H., Kornberg, A., Hidaka, M., Kobayashi, T., and Hozuchi, T. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9104–9108
26. Mayhook, A. G., Rinaldi, A. M., and Jacobs, H. T. (1992) Proc. Roy. Soc. Lond. B. 248, 85–94
27. Qureshi, S., and Jacobs, M. T. (1993) Nucleic Acids Res. 21, 2801–2808
28. Hehman, G. L., and Hauswirth, W. W. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 8562–8566