Mapping and positioning DNA-binding proteins along genomic DNA. Structure of *D.melanogaster* ribosomal 'Alu-repeats' and 1.688 satellite chromatin

Sergei V.Belikov⁺, Alexander I.Belgovsky, Marina P.Partolina, Vadim L.Karpov and Andrei D.Mirzabekov*

W.A.Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 117984, Russia

Received June 24, 1993; Revised and Accepted September 9, 1993

ABSTRACT

Chromatin structure of so-called 'Alu-repeat' in *D.melanogaster* ribosomal non-transcribed spacer that contains sequences homologous to the promoter of ribosomal genes has been studied. Using the 'protein image' hybridization assay based on UV-light-induced DNA-protein crosslinking and 2-D gel retardation electrophoresis, two proteins of the molecular mass of 50 kD (rABP50) and 70 kD (rABP70), associated with 'Alu-repeat' DNA have been found. Exo III mapping of crosslinking sites and DNase I footprinting have provided a detailed map of H1, rABP50 and rABP70 contacts within the 'Alu-repeat' and H1 and a non-histone protein contacts on satellite DNA. These data indicate precise positioning of non-histone proteins, histone H1 and nucleosomes within genomic regions studied and account for the presence of unusual 240 bp long nucleosomal particles in 'Alu-repeats'. The same approach can be adapted for successive mapping and positioning proteins on genomic DNA.

INTRODUCTION

Regulation of gene activity results from intricate interactions of several proteins with regulatory gene elements and is governed by such factors as chromatin structure, protein–DNA, protein–protein interactions and proteins and DNA modifications (1–4)

Ribosomal genes provide a convenient and well-studied model for addressing such questions as gene regulation and the structure of active and inactive chromatin. There are about 250 copies of ribosomal RNA genes in *D.melanogaster* organized in tandem repeats in which the transcription units are separated from each other by so-called 'non-transcribed' spacers (NTS) (for review see (4). Each NTS contains an about 300 bp-long promoter

sequence and 6 to 10 tandemly repeated 239 bp-long elements, flanked by duplicated Alu I cleavage sites (for convenience called 'Alu-repeats') (4–8) (Figure 1A). They contain sequences similar to the transcription initiation region (9): 52 base pairs of the 'Alu-repeat' match the promoter region, situated from –24 to +28 with respect to the transcription initiation site of the ribosomal RNA precursor (7,8,10) (Figure 1B) and seem to play an essential role in the transcription of ribosomal genes (11–13). There is almost a linear correlation between the number of 'Alu-repeats' and rRNA transcription rate (12). An intriguing feature of the ribosomal 'Alu-repeats' is the same 240 bp length in *D.virilis* and *D.hydei* despite their extensive sequence divergency (7). The 240 bp-long DNA repeat in *D.melanogaster* ribosomal NTS corresponds to the unusual 240 bp-long chromatin repeat of this region as revealed by micrococcal and DNase I digestion (14,15).

Another type of sequence we used as hybridization probe was 'satellite' DNA. This highly repeated DNA is arranged largely in long tandem arrays and is found mostly within centromeric and telomeric chromosomal regions, where it comprises the bulk of constitutive heterochromatin. One of the four major *D.melanogaster* satellite DNAs, the 1.688 g/cm³ satellite, consists of tandemly repeated copies of a 359 bp sequence; it is AT-rich and constitutes up to 4% of the total *D.melanogaster* DNA in nonpolytene tissues (16). Correspondence between the DNA repeat length and the length of chromatin repeat suggest positioning nucleosomes. Phased arrangement of nucleosomes on a tandemly repetitive DNA can be enforced by a phasing protein, such as α -protein (HMG1(Y) firstly described by Strauss and Varshavsky (17).

We have applied UV-induced DNA-protein crosslinking for whole *Drosophila* cells together with the 'protein image hybridization' method (18–21) and Exo III trimming for identification of proteins and localization of their contact sites along the 'Alu-repeat' of rDNA and 1.688 g/cm³ satellite DNA.

* To whom correspondence should be addressed

⁺ Present address: Karolinska Institutet, Department of Cell and Molecular Biology, Medical Nobel Institute, S-17177, Stockholm, Sweden

MATERIALS AND METHODS

Mapping of DNA-protein interactions

Covalent crosslinking of proteins to DNA by UV-light was performed as follows. 0.5 cm-thick suspension of *D. melanogaster* tissue-culture cells, subline K_c, in Echallier medium (22) (about $5-10 \times 10^6$ cells/ml) was irradiated for 10 min at 4°C in a jacketed dish. The source of UV-light was an inverted transilluminator (UV-Products, USA) without filter, placed 10 cm above the cell suspension. Nuclei were isolated from irradiated cell as described (23), then lysed in 0.15 M NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.1% NP-40, 2% lauryl sarcosinate, sonicated for 3 to 5 sec at 22 kHz, and the solution was centrifuged at $6,000 \times g$ for 5 min. To remove uncrosslinked proteins the supernatant was loaded onto a CsCl gradient containing 0.5% lauryl sarcosinate, 10 mM Tris-HCl, pH 8.0 and 1 mM EDTA (23). Centrifugation was carried out at 16°C for 48 hr at $200,000 \times g$ in a fixed-angle rotor (Ti 50, Beckman). The protein-containing top layer was discarded and the DNA-containing fractions were collected and dialysed three times against 0.25% NP-40, 10 mM Tris-HCl, pH 8.0. MgCl₂ was added to 5 mM and DNA and DNA-protein complexes were digested in the buffer with restriction endonuclease Alu I (2 units of enzyme per 1 µg of DNA) at 37°C for 16 hr in presence of 0.5 mM of protease inhibitor diisopropylfluorophosphate (DFP).

For precise mapping of crosslinking sites along DNA, digestion of DNA-protein complexes with Exo III (5 units per 1 µg of DNA) was performed at 37°C for 0.5 hour in presence of 0.5 mM DFP.

To remove most of the uncrosslinked DNA, the solution was mixed with an equal volume of phenol-chloroform (4:1, the phenol was equilibrated with 50 mM Tris-HCl, pH 8.0). The mixture was shaken, and after centrifugation the aqueous layer containing uncrosslinked DNA was discarded. The organic phase and interphase containing DNA-protein complexes were extracted five times with equal volumes of 1 M Tris-HCl, pH 8.0. tRNA-carrier (20 µg) was added to the organic layer and the material was precipitated by ethanol. The pellet was washed with 70% ethanol, dissolved in electrophoresis buffer and loaded onto the gel. Two-dimensional diagonal polyacrylamide gel electrophoresis of the DNA-protein complexes was carried out as described (20) to obtain 'protein images' diagonals. The DNA fragments were transferred onto Hybond nylon membrane (Amersham) in 10 mM Tris-acetate, pH 7.8, 5 mM CH₃COONa, 0.5 mM EDTA, at 10°C, 25 V, for 12 hr. The conditions for DNA labeling, hybridization and washing of the filter were described earlier (24-26).

The (+) and (-) strands of 'Alu-repeat' DNA fragment (239 bp) of ribosomal genes and satellite repeat DNA subcloned in ss M13 phage were used as strand-specific hybridization probes, the labeling was carried out using the 2nd-strand building method (25).

Nucleases protection experiments

D. melanogaster tissue-culture cells, subline K_c (22), were grown as described earlier (20). Nuclei were isolated (27) in a buffer containing 15 mM NaCl, 60 mM KCl, 15 mM Tris-HCl, pH 7.4, 3 mM MgCl₂, 1 mM DTT, 0.25 M sucrose and suspended in the buffer at a concentration of 20 A₂₆₀ U/ml. Nuclei were digested either with DNase I (10-25 units/ml) for 3-10 min at 25°C or by micrococcal nuclease (10 units/mg of DNA) at 25°C for 15 min in the same buffer in the presence

of 1 mM CaCl₂ and reaction was stopped by addition of EDTA to 10 mM. Nuclei were lysed by adding SDS to 1% and DNA was isolated as described (27). DNA after digestion with micrococcal nuclease was electrophoresed in 2% agarose gel. DNA with size of mono- and dinucleosome was eluted from gel also described as (Maniatis *et al.*, 1982). DNA (20-30 µg) was digested with Alu I or Dde I, subjected to sequencing reactions as previously described (28) and separated in 6% sequencing gel. After electrotransferring onto Hybond nylon membrane (Amersham) DNA was hybridized with ³²P-labeled single-stranded probe and autoradiographed.

Analysis of DNA-protein complexes on polyacrylamide gels

Analysis of DNA-protein complexes on polyacrylamide gels. The DNA oligonucleotides used in this work are shown in Figure 1B. The fragment F69 was labeled with Klenow polymerase. The initial F69 was digested with Alu I to yield F31 and F38 and the 'halves' obtained were separated electrophoretically in polyacrylamide gel and isolated. Labeled DNAs were mixed with the nuclear protein extract from *D. melanogaster* tissue culture cells subline K, incubated and then analysed by polyacrylamide gel electrophoresis (29).

RESULTS

Mapping and positioning DNA-protein contacts along genomic DNA

The experimental design is shown in Figure 2. UV-crosslinking and 'protein imaging' (Figure 2, path A) were done as described in Materials and Methods (see also (19)).

To localize precisely the crosslinked proteins along DNA, the enriched DNA-protein complexes are treated with exonuclease III (Exo III) (Figure 2, path B). Trimming of both 3'-ends of

Figure 1. A. A map of *D. melanogaster* ribosomal repeat. The thick lines indicate the regions corresponding to ribosomal 18S and 28S RNAs. The start and direction of transcription are indicated with an arrow. Type I and II ribosomal insertions are shown separately. Alu I restriction sites in 239 bp 'Alu-repeats' are designated as A. **B.** Internal organisation of ribosomal 'Alu-repeat'. Two adjacent repeats are shown. Alu I and Dde I—restriction sites in 239 bp 'Alu-repeats'. (+1)—a region homologous to the transcription initiation site. Grey boxes—the minimal (72 bp-long) fragment capable of activating rRNA transcription (Grimaldi *et al.*, 1990). Shaded boxes—the region highly homologous to RNA-polymerase I promoter (Coen and Dover, 1982). Double broken lines—region corresponding to hybridization probes used in DNase I footprinting experiments (24 and 47—size of the probes in nucleotides). F69, F31 and F38—regions of 'Alu-repeat' which correspond to oligonucleotides used in gel retardation assay.

double-stranded DNA fragment by Exo III stops at the DNA-protein crosslinking sites (see also 30). Thus, the sizes of crosslinked 5'-terminal single-stranded DNA fragments show precisely position of protein binding sites on a DNA fragment as the distance from its 5'-termini. After Exo III digestion, the protein-crosslinked complexes are fractionated in the two-dimensional diagonal gel electrophoresis system (as described above). A multiplicity of protein binding sites will result in several crosslinked 5'-terminal DNA fragments. Hybridization with ^{32}P -labeled M13 single-stranded probes for upper (+) and lower (-) strands of genomic DNA (Figure 2) enables one to localize the protein crosslinking sites on complementary DNA strands.

Two non-histone proteins and histone H1 are detected by crosslinking on the 'Alu-repeat' of rDNA

The 'Alu-repeat' regions of chromatin were shown earlier (15) to produce upon micrococcal nuclease digestion unusual 240-bp chromatin particles along with 'canonic' ones. In order to identify proteins in these particles, crosslinked DNA-protein complexes were subjected to 'protein image' analysis.

The results of such a hybridization with the 'Alu-repeat' probe are shown in Figure 3A. One can see the spots that correspond to histone H1 and two non-histone proteins. Core histones are not crosslinked by UV-light in significant amounts. To calibrate the gel and estimate molecular masses of these proteins, we carried out model experiments on chemical crosslinking to DNA (see Methods) of known proteins (*E. coli* RNA polymerase σ subunits (90 kD), bovine serum albumin (67 kD), topoisomerase

II from corn seedlings (135 kD) and histone H1 (37 kD)) with subsequent separation of DNA-protein complexes in the same two-dimensional gel (not shown). From the calibration experiments the molecular masses of these two proteins associated with ribosomal 'Alu-repeats' were found to be about 50 kD and 70 kD (accuracy of our method was $\pm 10\%$ for the proteins with molecular masses less than 40 kD and $\pm 15\%$ for the larger proteins). They were designated as rABP50 and rABP70 (ribosomal 'Alu-repeat' binding protein). The extreme left spot ('X' in Figure 3A) may correspond to DNA crosslinked to eukaryotic RNA polymerase I subunits which does not seem to enter the gel due to their large size.

Positioning of non-histone proteins and histone H1 on ribosomal 'Alu-repeats'

To localize crosslinked proteins on DNA, Alu-restricted DNA-protein complexes were digested with exonuclease III and

Figure 2. Experimental design for mapping and positioning of crosslinked proteins on genomic DNA by 'protein image' hybridization (see the text).

Figure 3. Mapping and positioning proteins on 'Alu-repeat' of ribosomal genes. A. Mapping by 'protein image' hybridization with 'Alu-repeat'. Diagonal lines under the autoradiographs indicate positions of 'protein image' diagonals for histone H1 (H1), for two proteins of apparent molecular masses of 50 kD (rABP50) and 70 kD (rABP70), and for an unknown protein with high molecular mass (X). Bottom diagonal lines indicate positions of 'marker' proteins: 37 kD (H1 histone), 67 kD (bovine serum albumin), 90 kD (σ -subunit of RNA polymerase from *E. coli*) and 135 kD (topoisomerase II). M — *Drosophila* DNA digested with Alu I restriction endonuclease (DNA-'marker') run as marker in the 2nd dimension; the size of DNA in the spot corresponds to 239 b.p. B, C. Positioning of proteins on *D. melanogaster* 'Alu-repeat' by Exo III digestion. (B)—hybridization with 'Alu-repeat' probe (upper strand, (+)). (C)—hybridization with 'Alu-repeat' probe (lower strand, (-)). The figures on the photo indicate the size in nucleotides of 5'-crosslinked DNA fragments. Numbers 1–4 above and below designate diagonals of uncrosslinked DNA (1), DNA crosslinked to histone H1 (2), rABP50 (3) and rABP70 (4). The arrow in Figure 3B indicates a nonspecific spot.

Figure 4. (A) Arrangement of rABP50, rABP70, histone H1 and nucleosomes within 'Alu-repeat' as revealed by DNA-protein crosslinking and nuclease footprinting assay. The arrow indicates the boundary of micrococcal nuclease protection region. R.P.—proposed RNA-polymerase binding region. (B) Arrangement of nucleosomes, histone H1 and pX within *D. melanogaster* (1.688-g cm⁻³) satellite repeat. The figures on both strands indicate the distance in bases from the 5'-termini of DNA fragments. Crosslinking sites are shown by the boxes (black boxes—strong contacts, dotted boxes—weak contacts). (+) and (—)—plus and minus strands.

then subjected to 'protein image' hybridization method. The filter was successively hybridized with probes for two complementary strands of 'Alu-repeat' DNA (Figure 3B,C) and 1.688 satellite DNA (see below, Figure 7).

Discrete size of ExoIII-treated crosslinked DNA fragments indicates unique positioning of rABP50, rABP70 and histone H1 on the 'Alu-repeat', as shown schematically in Figure 4A. Both rABP50 and rABP70 show a similar pattern of crosslinking. The size of DNA strongly crosslinked to rABP50 and rABP70 corresponds to their binding to the terminal regions of the 'Alu-repeat' at the positions distanced by 215 nucleotides from the 'Alu-repeat' 5'-end on the (+)strand and by 210 nucleotides on the (—)strand. H1 binds to the terminal regions (210–225 and 195 nucleotides from the 5'-end on the (+)strand and 160, 200–215 nucleotides from the 5'-end on the (—)strand) and also to the central region of 'Alu-repeat' (85, 105, 140–145 nucleotides from the 'Alu-repeat' 5'-end on the (+)strand and 115–125 nucleotides from 5'-end of 'Alu-repeat' on the (—)strand). DNA fragments shorter than 60 nucleotides were not well fixed on the filter and therefore were not identified. However, testing both DNA complementary strands made up in part for this failure.

Positioning proteins and nucleosomes by genomic nuclease footprinting

The ability of micrococcal nuclease to split DNA in chromatin preferentially between nucleosomes was used to localize nucleosomes on the 'Alu-repeat' by genomic footprinting assay (25). The same procedure was also used to localize on DNA the sites protected by proteins and nucleosomes against DNase I digestion. Isolated nuclei were treated either by micrococcal nuclease or by DNase I, and DNA was purified and subjected to Alu I restriction. After electrophoretic separation and electroblotting on the filter, DNA was hybridized with a short synthetic oligonucleotide (24 nucleotides from the 5'-end of (—) strand) as a strand-specific hybridization probe ('Alu 24 probe',

Figure 5. Micrococcal nuclease and DNase I footprinting in *D. melanogaster* 'Alu-repeat'. The figures (on A and B) indicate the distance in nucleotides counting from left Alu I restriction site on (+)strand (see Figure 1B). A. Hybridization with Alu 24 probe (see Figure 1) after Alu I digest. 1. Nucleosomes were prepared by digestion of nuclei with micrococcal nuclease. DNA isolated from mono- and dinucleosomes was digested with Alu I. The arrow indicates the boundary of micrococcal nuclease protection region. 2. G+A genomic sequencing marker. 3. Purified DNA digested first with DNase I and then with Alu I. 4. Nuclei were digested with DNase I and isolated DNA was then treated with Alu I. B. Hybridization with Dde I-Alu I probe (see Figure 1) after Dde I digest. 1. Purified DNA digested with DNase I and then with Dde I. 2. DNA from nuclei digested with DNase I and then treated with Dde I.

see Figure 1). The data obtained show that micrococcal nuclease makes cuts at positions around the 70–75th nucleotide on the (+)strand thus indicating the approximate location of the exposed nucleosomal end (Figure 5A, track 1). The precise positioning of the nucleosome on the 'Alu-repeat' is also suggested by the ten-nucleotide periodicity in the pattern of DNase I digestion extending from the 70th to 220th nucleotide (Figure 5A, lane 4). A somewhat smeared pattern of digestion in Figure 5A, lane 4 could be due to a close location of two Alu I restriction sites which are spaced by only four nucleotides and therefore can be digested alternatively at one or another site (Figure 1). Indeed, restriction of DNA with Dde I that gives rise to one 'Alu-repeat' fragment rather than two fragments for Alu I; Dde I produces therefore a more distinct ten-nucleotide periodicity pattern (Figure 5B).

Two strongest regions in 'Alu-repeat' chromatin protected against DNase I digestion are located in the area 43–63 and 210–220 nucleotides. The second protected region corresponds quite well to strong crosslinking sites for rABP50, rABP70 and histone H1. No crosslinking of rABPs or H1 was found within 43–63 region. In some experiments we observed crosslinking of some large proteins to 'Alu-repeat' (Figure 3A, spot X) which may correspond to RNA polymerase I large subunits. We speculate that this region could interact with RNA polymerase

Figure 6. Gel retardation assay for fragments of 'Alu-repeat' DNA. Autoradiograph of a 5% polyacrylamide gel shows protein-DNA complexes between 1 ng of 'Alu-repeat' fragments (see Figure 1) F69 (b), F31 (d), F38 (f) and extract from *D.melanogaster* nuclei in presence of 400 ng of poly dA-dT, as a competitor. (a,c,e)—the same oligonucleotides, respectively, but without extract.

Figure 7. Alignment of Alu I restriction site along *D.melanogaster* (1.688-g/cm³) satellite. Double broken line—186 bp-long DNA fragment was used as a hybridization probe B,C. Exo III mapping experiments in *D.melanogaster* (1.688-g/cm³) satellite. Numbers 1–3 designate diagonals of uncrosslinked DNA, unknown protein (pX) and histone H1, respectively. The figures indicate the distance in bases from the 5'-termini of DNA. Hybridization with B—Sat/Alu 186, upper (+) strand, C—Sat/Alu 186, lower (–) strand.

I (see Discussion). The data on localization of protein crosslinking sites and a nucleosome on 'Alu-repeat' are summarized in Figure 4A.

A specific protein binds to 'Alu-repeat' as revealed by gel retardation

Specific crosslinking of a protein to 'Alu-repeat' DNA was corroborated by gel retardation assay (29) (shown in Figure 6). A synthetic double-stranded 69 bp long DNA fragment (F69) and two subfragments produced by its digestion with Alu I (F31 and F38) (see Figure 1,B) were used in the experiments. All three

Figure 8. A model for the arrangement of DNA in ribosomal 'Alu-repeat' chromatin. The open boxes correspond to histone octamer positioning. Grey boxes correspond to rABPs contacts, black boxes to supposed RNA-polymerase positioning. (+1)—points to transcription initiation site. The 80 bp spaced histone H1 contacts on the adjacent turns of chromatin superhelix are designated with the white vertical arrows. MN points to the micrococcal nuclease cutting site. The figures along the helix correspond to distance in base pairs from the left Alu I restriction site in every 'Alu-repeat'.

fragments bound specifically one protein in this assay. F69 covers two rABP crosslinking sites: 215 on the (+) strand and 200–215 on the (–) strand whereas F31 and F38 include each of them, separately.

Positioning of proteins within *D.melanogaster* satellite DNA (1.688-g/cm³)

Organization of *D.melanogaster* satellite DNA (repeat unit 359 bp) is shown in Figure 7A (30). The same method was applied for mapping DNA–protein interactions within this genomic region. We used cloned sequences (186 b.p.) between Alu I sites designated as Sat/Alu 186 (see Figure 7A). The data are presented in Figure 7B,C. DNA discrete size of ExoIII-treated crosslinked DNA fragments indicates unique positioning of histone H1 and unknown protein (pX) of the apparent molecular mass of 15 kD on the satellite repeat. Crosslinking sites, where ExoIII-nuclease stops, schematically shown in Figure 4B. H1 histone strong crosslinking sites correspond to the positions of 135th, 165th and 185th nucleotide from the left Alu-restriction site on the (+) strand and 115th, 140th and 175th nucleotides from right Alu-restriction site on the (–) strand. The strong sites of crosslinking for pX correspond to 186th and 135th nucleotides from 5' end of 186 bp Alu-Alu fragment for (+) strand and 175 nucleotides from the 5' end of 186 bp Alu-Alu fragment for (–) strand.

DISCUSSION

Structure of 'Alu-repeat' chromatin

Nucleosomes were mapped on the 'Alu-repeat' region by micrococcal nuclease and DNase I digestion studies (14). Crosslinking experiments revealed a lower extent of crosslinking for H1 and core histones on this DNA as compared with silent chromatin (15, 31 see also 32) which suggested that interaction of histones with DNA was somehow changed.

The DNase I digestion pattern shows nucleosome-like 10-nucleotide periodicity in the 60–220 nucleotide region of the digested 'Alu-repeat' DNA (Figure 5). Micrococcal nuclease specifically cuts out nucleosomes from chromatin and splits the repeat DNA at the site of the 70–75th nucleotide on the (+) DNA strand. These data suggest that the nucleosome is positioned on the 'Alu-repeat' DNA beginning from the 70–75th nucleotide from the 5'-end of 'Alu-repeat'.

On early stages of digestion micrococcal nuclease cuts the 'Alu-repeat' region with a periodicity equal to the DNA periodicity, i.e. about 240 bp (14,15). at the site 20–30 bp downstream from the Alu I restriction site. We localized this site roughly at the same place (Figure 5A, lane 1) however it covers much extended area. Most probably this is due to using in our experiments later stages of micrococcal nuclease digestion.

It is essential to stress here that the nucleosome positioning revealed by micrococcal nuclease digestion was precisely predicted on the basis of the theoretical model of sequence-dependent curvature in DNA fragments (33).

For histone H1, strong crosslinking sites are expected to be located on the terminal regions of nucleosomal DNA. There are also some H1 contacts within nucleosomal DNA, in particular within its middle part at the sites (+)140 and (–)120 on complementary strands (see Figure 4). The binding of H1 to internal regions of nucleosomal DNA was observed by chemical crosslinking (34) and at the center of nucleosomal DNA as shown by nuclease protection experiments (35).

Non-histone proteins rABP50/70 are positioned outside the nucleosome although there is a weak binding site within the nucleosome. It is worth mentioning here that non-histone proteins HMG-14/17 occupy similar regions of DNA within core particle or a chromatosome (36). The binding sites for rABPs and histone H1 are partly overlapped. The most straightforward conclusion is that H1 and rABPs may compete with each other for DNA binding. Indeed, in the case of 5S RNA gene, histone H1 interferes with protein regulatory factor TFIIIA binding to regulatory DNA sequence located inside nucleosome due to presumably closing up compact nucleosome structure (37). However, crosslinking data do not exclude a possibility of their simultaneous interaction with the same DNA sites (see below). Gel retardation assay demonstrated that two DNA fragments 0–31th nucleotide (F31) and 201–239th nucleotide (F38) (see Figure 1B) bind specifically (separately or within one fragment) only one protein from nuclei extract. F31 and F38 contain crosslinking sites for rABPs- +185, +215, +239 and –239, –210, respectively. It seems that the bound protein is rABP50 or/and rABP70. Taken together with crosslinking rABP50 and rABP70 to the same DNA sequences, this suggests that rABP50 is either a product of partial degradation of rABP70 or these two proteins are very similar and run with the same mobility in native gel upon electrophoresis but with different mobilities in denaturing SDS-gel.

The presence of 'Alu-repeat' activates rRNA transcription in vitro, and its minimal activating fragment is 72 bp-long extending from 211 to 44 bp (38). This precisely coincides with crosslinking sites for rABP70/50.

Another interesting feature of 'Alu-repeat' is the fact that the region of 10–60th bp on 'Alu-repeat' has the same sequence as for transcription initiation site of the ribosomal RNA gene (7,8,10) and is located close by and downstream from the rABP70/50 binding site. We have not established directly the crosslinking of RNA polymerase to this site (probably for technical reasons, see Results) however this homology taken together with essential protection of the region against DNase I digestion strongly suggest that this site is occupied by RNA polymerase I. Significantly, Kohorn & Rae have shown that the 'Alu-repeat' can serve as templates for in vitro transcription and that initiation occurs roughly at the same position as in *in vivo* (8).

Crosslinking, nuclease protection and digestion studies revealed a highly organized structure of 'Alu-repeat' where a nucleosome, H1 histone, rABP70/50 and probably RNA polymerase are precisely positioned on its DNA. It is not surprising therefore that the 'Alu-repeat' chromatin seems to be folded into a higher order structure. It was proposed that the existence of a variable peculiarity in three-dimensional structure of this chromatin results in difference in the kinetics of micrococcal nuclease cutting between different 'Alu-repeats'. Relevant to this is the electron microscopy observation of a highly compacted structure of non-transcribed spacer in *Xenopus* rDNA that contains 'Alu-repeat'-like elements, while transcribed rDNA chromatin regions were found to be in an extended state (39). What could be a three-dimensional organization of this chromatin?

We propose therefore that chromatin DNA of 'Alu-repeat' is organized as continuous superhelix with 80 bp per turn making the 10nm chromatin fiber that in some cases may be further folded into continuous solenoid structure. The model of superhelical structure of 'Alu-repeat' chromatin is shown in Figure 8. The arrangement in H1 crosslinking sites has 80 bp periodicity; it appears that histone H1 interacts simultaneously with adjacent regions on the DNA superhelix thereby stabilizing the latter (see also (40)). rABP70/50 and proposed RNA polymerase binding sites are always located on the same side of the superhelical surface and may be permanently exposed in folded solenoidal structure. It is worth mentioning here that 'Alu-repeats' seem to play a role of a depot to accumulate RNA polymerase in front of the transcription initiation site. If this is the case, the compact solenoidal structure of the 'Alu-repeat' exposes polymerase molecules on the surface of the chromatin superhelix, like bullets in a revolver, to make polymerases ready for fast initiation of RNA synthesis when transcription begins.

D.melanogaster satellite (1.688-g/cm³) chromatin

The same methods have been used for investigation of *D.melanogaster* satellite chromatin with the repetitive unit of 359 bp. Exo III mapping experiments have shown precise positioning of H1 and pX protein contacts within the sequences analyzed (Figures 7). It is interesting that in both cases precise positioning of nonhistone proteins and histone H1 on DNA correlates with periodicity of DNA sequence.

In the vicinity of the diagonal relevant to the core histones we found regions of intensive contacts. According to our data (32) in intact cells UV-crosslinking of core histones to DNA proceeds rather poorly and we did not observe core histone crosslinking

by hybridization of the the same filter with 'Alu-repeat' DNA. It appears that we observed contacts of some nonhistone protein with molecular mass of approximately 10–15 kD. A non-histone protein which was found to bind preferentially to α -satellite DNA was isolated from African green monkey cells (17). One can suggest for this kind of protein a role of a phasing signal, which determines precise nucleosome positioning. This suggestion is supported by the data showing that the periodicity in localization of pX contacts along DNA is equal to the nucleosomal repeat length (180 bp) of *D.melanogaster* satellite (unpublished data). Another contact of this protein is distanced from the main one by 40 bp and can correspond to alternative phasing of nucleosomes on satellite DNA.

The periodicity of histone H1 contacts along DNA differ from that shown for 'Alu-repeat' chromatin. Whilst in 'Alu-repeat' the contacts of histone H1 were observed within nucleosomal DNA region, in satellite chromatin we determined a high density of contacts probably at the ends of nucleosome and at the internucleosomal spacer. It appears that for highly compact satellite chromatin and activated ribosomal 'Alu-repeat' different modes of histone H1-nucleosome binding do exist. It was demonstrated, as an example, that the contacts of Thr-1 and His-15 in histone H5 and similar residues in histone H1 are effected by chromatin condensation (41).

ACKNOWLEDGEMENTS

We thank Miss E.I.Novikova for help in manuscript preparation. This study was supported by grant No. gg562 from the Russian Human Genome Program and the Russian State Program 'Frontiers in Genetics'.

REFERENCES

- Wasylyk, B. (1988) *Biochimica et Biophysica Acta*, **951**, 17–35.
- Bach, R., Allet, B. and Crippa, M. (1981) *Nucl.Acids Res.*, **9**, 5311–5330.
- Morse, R.H. (1992) *Trends Biochem.Sci.*, **17**, 23–26.
- Felsenfeld, G. (1992) *Nature*, **355**, 219–224.
- Walfie, A.P. (1992) *FASEB J.*, **6**, 3354–3361.
- Long, E.O. and Dawid, I.B. (1980) *Ann.Rev.Biochem.*, **49**, 727–764.
- Long, E.O. and David, I.B. (1979) *Nucl.Acids Res.*, **7**, 205–215.
- Coen, E.S. and Dover, G.A. (1982) *Nucl.Acids Res.*, **10**, 7017–7026.
- Kohorn, B.D. and Rae, P.M.M. (1982) *Nucl.Acids Res.*, **10**, 6879–6886.
- Bach, R., Allet, B. and Crippa, M. (1981) *Nucl.Acids Res.*, **9**, 5311–5330.
- Simeone, A., de Falco, A., Macino, G. and Boncinelli, E. (1982) *Nucl. Acids Res.*, **10**, 8263–8272.
- Baker, S.M. and Platt, T. (1986) *Cell*, **47**, 839–840.
- Grimaldi, G. and Di Nocera, P.P. (1986) *Nucl.Acids Res.*, **14**, 6417–6432.
- Hayward, D.C. and Glover, D.M. (1988) *Nucl.Acids Res.*, **16**, 4253–4268.
- Udvardy, A., Louis, C., Han, S. and Schedl, P.J. (1984) *J.Mol. Biol.*, **175**, 113–130.
- Dzherbashyajan, A.R., Vashakidse, R.P., Karpov, V.L., Kolchinsky, A.M. and Mirzabekov, A.D. (1984) *Molecular Biology (USSR)*, **22**, 231–241.
- Carlson, M. and Brutlag, D. (1977) *Cell*, **11**, 371–381.
- Strauss, F., Varshavsky, A. (1984) *Cell*, **37**, 889–901.
- Karpov, V.L., Preobrazhenskaya, O.V. and Mirzabekov, A.D. (1984) *Cell*, **36**, 423–431.
- Mirzabekov, A.D., Bavykin, S.G., Belyavsky, A.V., Karpov, V.L., Preobrazhenskaya, O.V., Shick, V.V., Ebralidse, K.K. (1989) *Meth.Enzymology*, **170**, 386–408.
- Nacheva, G.A., Guschin, D.Y., Preobrazhenskaya, O.V., Karpov, V.L., Ebralidse, K.K. and Mirzabekov, A.D. (1989) *Cell*, **58**, 27–36.
- Belikov, S.V., Belgovsky, A.I., Preobrazhenskaya, O.V., Karpov, V.L. and Mirzabekov, A.D. (1993) *Nucl.Acids Res.*, **21**, 1031–1034.
- Echalier, G. and Ohanessian, A. (1970) *In Vitro*, **6**, 162–172.
- Gilmour, D.S. and Lis, J.T. (1985) *Mol.Cell.Biol.*, **5**, 2009–2018.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning; a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Church, G.M. and Gilbert, W. (1984) *Proc.Natl.Acad.Sci.USA*, **81**, 1991–1995.
- Marchionni, M. and Gilbert, W. (1980) *Cell*, **46**, 133–141.
- Wu, C., Bingham, P.M., Livak, K.J., Holmgren, R. and Elgin, S.C.R. (1979) *Cell*, **16**, 797–806.
- Maxam, A.M. and Gilbert, W. (1980) *Meth. Enzymology*, **65**, 499–560.
- Fried, M. and Crothers, D.M. (1981) *Nucl.Acids Res.*, **9**, 1605–1625.
- Hsieh, T.S. and Brutlag, D. (1979) *Cell*, **18**, 1359–1370.
- Belikov, S.V., Dzherbashyajan, A.R., Preobrazhenskaya, O.V., Karpov, V.L. and Mirzabekov, A.D. (1990) *FEBS Letters*, **273**, 205–207.
- Dimitrov, S.I., Stefanovsky, V.Yu., Karagyzov, L., Angelov, D. and Pashev, I.G. *Nucl.Acids Res.*, **18**, 6393–6397.
- Boffelli, D., De Santis, P., Palleschi, A., Risuleo, G. and Savino, M. (1992) *FEBS Letters*, **300**, 175–178.
- Shick, V.V., Belyavsky, A.V., Bavykin, S.G. and Mirzabekov, A.D. (1980) *J.Mol.Biol.*, **139**, 491–51.
- Crane-Robinson, C., Staynov, D.Z. and Baldwin, J.P. (1984) *Comm. Mol. Cell.Biophys.*, **2**, 219–265.
- Shick, V.V., Belyavsky, A.V. and Mirzabekov, A.D. (1985) *J.Mol.Biol.*, **185**, 329–339.
- Rodes, D. (1985) *EMBO.J.*, **5**, 3470–3482.
- Grimaldi, G., Fiorentini, P. and Di Nocera, P.P. (1990) *Mol. Cell. Biol.*, **10**, 4667–4677.
- Pruitt, S.C. and Grainger, R.M. (1981) *Cell*, **23**, 711–720.
- Bavykin, S.G., Usachenko, S.I., Zalensky, A.O. and Mirzabekov, A.D. (1990) *J.Mol.Biol.*, **212**, 495–511.
- Mirzabekov, A.D., Pruss, D.V. and Ebralidse, K.K. (1989) *J.Mol.Biol.*, **211**, 479–491.