A Naturally Occurring T14A11 Tract Blocks Nucleosome Formation Over the Human Neurofibromatosis Type 1 (NF1)-Alu Element

(Received for publication, October 24, 1995, and in revised form, December 28, 1995)

Ella W. Englander and Bruce H. Howard
From the Laboratory of Molecular Growth Regulation, NICHD, National Institutes of Health, Bethesda, Maryland 20892-2753

The nature of chromatin organization over Alu repetitive elements is of interest with respect to the maintenance of their transcriptional silencing as well as their potential to influence local chromatin structure. We previously demonstrated that the pattern of nucleosomal organization over Alu elements in native chromatin is specific and similar to the pattern observed with an in vitro reconstituted Alu template. This pattern, distinguished by a nucleosome centered over the 5'-end of the Alu element, is associated with repression of polymerase III-dependent transcription in vitro (Englander, E. W., Wolff, A. P., and Howard, B. H. (1993) J. Biol. Chem. 268, 19565–19573; Englander, E. W., and Howard, B. H. (1995) J. Biol. Chem. 270, 10091–10096). In the current study, additional templates representing both evolutionarily old and young Alu subfamilies were found to direct a similar pattern of nucleosome assembly, consistent with the view that nucleosome positioning in vitro is shared by a majority of Alus. We discovered however, that the specific nucleosome positioning pattern was disrupted over one member of a young Alu subfamily, which recently transposed immediately downstream to a T14A11 sequence in the neurofibromatosis type 1 locus (Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, P. E., Glover, T. W., and Collins, F. S. (1991) Nature 353, 864–866). Upon removal of this sequence motif, the expected pattern of assembly was restored to the neurofibromatosis type 1-Alu template. This finding indicates that, at least in vitro, certain sequences can override the propensity for positioning nucleosomes that is inherent to Alu elements. The finding also raises the possibility that a similar situation may occur in vivo, with potential implications for understanding mechanisms by which certain Alu elements may evade chromatin-mediated transcriptional silencing.

The Alu class of mammalian short interspersed elements (SINEs) is the most abundant family of interspersed DNA repeats in primates. It encompasses more than 500,000 copies of a 300-bp dimeric sequence comprising 5–10% of human genomic DNA (reviewed in Refs. 1 and 2). Since Alu elements carry internal RNA polymerase III promoter motifs and form efficient pol III-dependent transcription templates in vitro (3–5), a subject of continuing interest is how these elements are regulated by the host cell. This issue has been a subject of several lines of investigation, initially leading to the hypothesis that most Alu promoters are mutated and inherently weak and therefore require specialized 5'-flanking sequences for expression (6–8). More recently it was demonstrated, however, that the normally very low level of Alu transcription in vivo can be selectively increased (i.e., relative to other pol III-transcribed genes) in response to viral infection (9–13), inhibition of translation (14), or demethylation of internal Alu CpG dinucleotide residues (15). In addition, template availability experiments demonstrated that only a minor fraction of the transcriptional potential of Alu elements in naked genomic DNA is realized in either native chromatin or chromatin from which histone H1 was removed (16). Taken in combination, this evidence appears to favor the view that Alu elements are subject to one or more levels of transcriptional repression in human cells.

A potentially attractive mode of gene silencing, especially for dispersed repeat elements, is packaging of transcription control regions into nucleosomes. Accordingly, the nucleosomal organization over Alu elements in vitro and in vivo has been the focus of ongoing studies in our laboratory. Analysis of nucleosome assembly in vitro over a naturally occurring Alu element that transposed into the human α-fetoprotein gene (AFP-Alu) (17) revealed that this element directs precise rotational and translational nucleosome positioning (18). Similar translational nucleosome positioning over this Alu was demonstrated in native human chromatin, and examination of human Alu elements as a family revealed that a significant fraction of repeats is associated with rotationally positioned nucleosomes (19). These results strongly suggested that Alu elements carry inherent signals that direct nucleosomes to center over their 5'-ends, which correspond to their potential transcription start site. Moreover, assembly of a nucleosome in this position resulted in total repression of Alu promoter-mediated pol III-dependent transcription (18), leading us to propose that the transcriptional silencing of Alu elements observed in vivo is likely to have a chromatin-mediated component. As noted above, Alu elements are not significantly derepressed by simpler removal of histone H1 from chromatin (16). This contrasts with other pol III-transcribed genes (20–23) and suggests that the unusually high degree of transcriptional repression of Alu repeats may be related to their propensity to direct a non-random pattern of nucleosome assembly.

To the extent that Alu elements have strong inherent signals for positioning nucleosomes, they provide a system for investigating the nature of sequence motifs that may override such signals. In this context, we considered the available experimental and theoretical data showing that homopolymeric tracts of adenine residues are not energetically favorable for wrapping around the histone core and accordingly are excluded from the central regions of nucleosomal DNA (24–26). Here, we examine a naturally occurring sequence of this type (T14A11) juxtaposed immediately 5' to the start site of an Alu element that recently transposed into the neurofibromatosis type 1 (NF1) locus, causing that disorder (27). We analyzed the pattern of in vitro
nucleosome assembly over this Alu element within its natural genomic context and also compared it to the pattern obtained in the context of prokaryotic DNA. It was found that the adjacent region consisting of T_{14}A_{11} interferes with nucleosome assembly over the start site of the NF1-Alu. As expected, revelation of this sequence with non-homopolymeric prokaryotic DNA restored the pattern of nucleosome formation characteristic of Alu templates as well as the in vitro transcriptional repression associated with nucleosome positioning on these elements.

MATERIALS AND METHODS

Template Construction—The NF1-Alu clone (Fig. 1) was generously provided by Dr. M. Wallace. It comprises 794 bp corresponding to the Alu insertion and the surrounding region in the sixth intron of the NF1 gene, generated by polymerase chain reaction and cloned (clone U4) into the BamHI site of the Bluescript 1(ks-) vector (27). The BamHI fragment was subcloned into pUC21 (Boehringer & Mannheim) by digestion with Sty at position −197 relative to the Alu transcription start site and at position 387 within the vector, yielding a ~670-bp fragment-containing plasmid termed NF1-Alu. To create a template labeled uniquely at one end, a fill-in reaction with Klenow was employed with the distal-labeled end of the fragment removed by cutting with BamHI. The resulting template was a 597-bp long Sty/BamHI fragment. To generate the short ~330-bp template lacking the right monomer, the NF1-Alu was cut with HaelI and termed NF1-Alu/l.

The NF1-Alu/d (Fig. 1) template was generated by digesting the NF1-Alu fragment with EcoRI at position −1 relative to the Alu transcription start site and at position 297 within the vector. The ends were filled in with Klenow, and the fragment was cut with MluI at position 310 within the polylinker to generate a MluI/BamHI end fragment, which was then subcloned into pUC21 cut with EcoRV and MluI. The NF1/d-Alu template was generated by cutting this construct with PvuII and EcoRI at positions 54 and 323, respectively, within the vector.

The AluK/pUC19 (Fig. 1) construct was synthesized as a consensus Alu sequence of 285 bp (28). The ~490-bp fragment used as template for nucleosome reconstitution was prepared by cutting the construct with PvuII and EcoRI within the vector. The resulting fragment consisted of ~200 bp of pUC19 DNA and ~285 bp of Alu sequence.

Histone Purification and Nucleosome Reconstitution—Core histones were purified from adult chicken erythrocytes according to the described procedure (29, 30). Assembly of nucleosomes onto the end-labeled DNA fragments was by stepwise sequential dialysis from high salt and urea with purified histones (18, 30, 31). Reconstitution reactions were in a total volume of 180 μl. Histone octamer solutions; the samples were then phenol:chloroform extracted, ethanol precipitated, and finally dialyzed for 4 h at 0.6 M NaCl without urea. Dialysis was continued bations with decreasing NaCl concentrations of 1.2, 1.0, and 0.8 M NaCl, 3 M urea. Salt gradient dialysis was carried out by successive 80-min incubations with 1 M Tris, pH 8, 1 mM EDTA, 10 mM mercaptoethanol. The mixtures were dialyzed 12 h against the same buffer containing 5 mM urea. Salt gradient dialysis was carried out by successive 80-min incubations with 1 M NaCl concentrations of 1.2, 1.0, and 0.8 M NaCl and finally for 4 h at 0.6 M NaCl without urea. Dialysis was continued with two further changes over the next 16 h against 10 mM Tris, pH 8, 1 mM EDTA, 10 mM mercaptoethanol. The extent of reconstitution was monitored by gel retardation analysis of the histone-DNA complexes in 0.7% agarose gels in 0.5 × TBE.

DNase I Analysis—To obtain DNase I digestion, approximately 25 fmol of end-labeled templates were incubated with 25 μg of DNase I (Sigma) in 20 μl of reaction buffer containing 50 mM NaCl, 0.1 mM EDTA, 5 mM MgCl2, 0.1 mg/ml bovine serum albumin, and 150 μg/ml carrier DNA. Following incubation for 1, 2, or 4 min, aliquots of the reaction mixtures were transferred into stop solution; the samples were then phenol:chloroform extracted, ethanol precipitated, and resolved in a 6% sequencing gel (18).

In Vitro Transcription—The NF1-Alu and the NF1/d-Alu fragments, each carrying the NF1-Alu element flanked on the 3'-end by its natural pol III transcriptional start signal, were used as templates for in vitro transcription in a free form or reconstituted with histone octamers. Nuclear transcription extract was prepared according to the described procedure (32), modified by addition of a 2.5 mM ammonium sulfate precipitation step. In vitro transcription reactions were assembled in a final volume of 25 μl with 100 ng of Alu element in 10 mM Hepes, pH 7.9, 5 mM MgCl2, 42 mM KCl, 1 mM dithiothreitol, 10% glycerol. The total amount of DNA was standardized to 300 ng with λ DNA. Mixtures were incubated for 20 min on ice, and transcription was initiated by addition of 15 mM α-32P[GTP (18). Following incubation at 30°C for 60 min, reactions were stopped, extracted with phenol:chloroform, and analyzed by 6% polyacrylamide denaturing gel electrophoresis.

RESULTS

To confirm that nucleosome positioning in vitro is a general property of Alu repeats, a 285-bp synthetic consensus Alu element (AluK) (Fig. 1) inserted into a prokaryotic vector (28) was tested as a template for nucleosome assembly. This consensus element was designed to conform to the Alu major family sequence (33), which represents ~85% of Alu repeats that currently reside in human DNA. In contrast, the AFP-Alu (Fig. 1) element, whose nucleosomal assembly pattern was analyzed previously, represents a more evolutionary recent family designated precise (34), which comprises 10–15% of Alu elements.

The NF1-Alu represents a young family that is currently transposing in the human genome. Although the consensus sequences representing the different families exhibit ~95% similarity, the individual members of the older subfamilies retain on average only ~85% sequence identity (reviewed in Refs. 2 and 35).

Analysis of DNase I footprinting patterns of the reconstituted AluK template revealed a pattern of ~10-bp periodic nicking that is characteristic of nucleosomal DNA. The pattern revealed defined rotational positioning around the start of the Alu consensus sequence (Fig. 2, arrows), consistent with nucleosome positioning observed earlier for the in vitro reconstituted AFP-Alu element within the context of its native genomic sequence (18). This observation further supports our proposal (19) that nucleosome positioning signals are inherent to the Alu sequence and confer the capacity to direct nucleosome positioning to a significant fraction of human Alu repeats.

Previously, others have observed that certain sequences are relatively refractory to nucleosome assembly (24, 36, 37). We observed that one such sequence, a short homopolymeric tract, T_{14}A_{11}, occurred immediately 5' to the recently transposed NF1-Alu and, accordingly, decided to examine its effect on the assembly of a nucleosome predicted to center over the transcription start site of the adjacent NF1-Alu element. The NF1-
Alu element located centrally in a ~600-bp fragment (Fig. 1) was generated and reconstituted with purified histones by salt gradient dialysis. The disappearance of free DNA confirmed the complete association of the template with histones (Fig. 3a, lane 3). The DNaseI footprint, however, failed to generate periodic preferential nicking indicative of nucleosomal DNA (Fig. 3b); instead, the DNaseI pattern corresponding to the Alu sequence was similar for the reconstituted and the free templates. Nonetheless, a marked difference in the footprinting pattern between the reconstituted and the free DNA was detected outside of the Alu sequence, at position -60 relative to the transcription start site (arrow). This suggested that a nucleosome or a nucleosome-like particle that would account for the formation of the observed bandshift was associated with the proximal portion of this template. To rule out a situation in which the observed bandshift was due to a nucleosome associated with the right Alu monomer, the original template was truncated to remove its distal portion. The resulting ~330-bp fragment (NF1/t-Alu), consisting of ~170 bp of a 5’-upstream region, the left Alu monomer, and the intermonomeric linker, was reconstituted with purified histones, and its complete assembly was confirmed by gel retardation assay (Fig. 3a, lane 7). Footprinting differences between DNaseI patterns in the region corresponding to the Alu sequence or the immediately adjacent 5’-flanking sequence T14A11 were likewise absent in this case (Fig. 3c). Differences in the footprinting pattern suggesting the presence of nucleosome-associated DNA that could account for the observed bandshift were apparent only in the proximal portion of the template, i.e. upstream of position -60 (arrows). These results are in striking contrast to those observed with other Alu templates, including AluK (Fig. 2), in which the ~10-bp periodic nicking pattern is most readily seen immediately upstream of the Alu start site. Since this region is occupied by T14A11, the data suggest that this tract is responsible for exclusion of a nucleosome from its otherwise preferred position.

In view of these results, we wished to confirm the possibility that the presence of the T14A11 sequence interferes with nucleosome formation over the start of the Alu element. To accomplish this, the T14A11 tract was removed (see “Materials and Methods”). The resulting construct was used to generate a ~600-bp PvuII-EcoRI fragment termed NF1/d-Alu, carrying ~190 bp of prokaryotic sequence immediately flanking the 5’-end of the NF1-Alu element. The sequence replacing the T14A11 is comprised of ~50% G+C and has no homopolymeric tracts (not shown). Complete reconstitution of this fragment with purified histones was confirmed by the gel retardation assay (Fig. 3a, lane 5). In contrast to its parent templates, DNaseI footprinting analysis did reveal differences in the patterns on comparison of the reconstituted and free templates over the region corresponding to the proximal portion of the Alu left monomer and the flanking upstream sequence (Fig. 4, arrows). Thus, the footprinting pattern obtained with the modified template suggested that formation of nucleosomal DNA over the start site of the Alu element was dependent on the removal of the T14A11 sequence.

Since we previously found that an octamer centered over the start of the AFP-Alu element repressed pol III-mediated transcription (18), we next examined the ability of the NF1 constructs to direct transcription. For this analysis, we used nucleosome-free and fully reconstituted templates (Fig. 3a). Control experiments demonstrated that replacement of the T14A11 sequence had no apparent effect on the ability to direct transcription by the free template (Fig. 5, a (lanes 2 and 3) and b (lane 1)). Interestingly, in these reactions it was found that while the natural NF1-Alu template was transcriptionally active when associated with histones (Fig. 5a, lanes 4–6), the NF1/d-Alu template lacking the 5’-T14A11 tract was reproducibly repressed by the assembly of nucleosomes (Fig. 5b, lanes 2 and 3). As shown in Fig. 4, the pattern of assembly in the latter case indicated that the NF1/d-Alu was associated with a histone octamer located over the transcription start site of this template. These results further confirm that the ability of an Alu element to position a nucleosome over its start site corresponds closely to its transcriptional repression when reconstituted in vitro with core histones.

**DISCUSSION**

Our initial study examined AFP-Alu, a human-specific element that is a member of the Alu precise subfamily representing 10–15% of the Alu elements in the human genome. Here, we extended the analysis using the AluK consensus element, which represents older Alu subfamilies corresponding to 85% of human Alu repeats as well as the NF1-Alu element representing the most recent Alu subfamily (38–40). Combined, our results indicate that the ability of Alus to position nucleosomes precisely is very likely associated with many if not most Alu repeats. In addition, we demonstrated that the most recent Alu
subfamily, represented here by the NF1-Alu, carries the same inherent signals to position nucleosomes. Cumulatively, these observations bolster our previous evidence that Alu elements en masse position nucleosomes centered over their 5'-ends (19) and that Alu transcriptional silencing is in part due to chromatin-mediated repression (16).

Insofar as the formation of a nucleosome over Alu transcription start sites contributes to the silencing of Alu genes in vivo in the same manner that occurs in vitro, the results presented here suggest a view of Alu regulation somewhat different from those previously proposed. In this regard, it is noteworthy that Alu transcripts originate in vivo from all subfamilies (13, 15, 41, 42), rather than, as expected, only from the most recent subfamily members. In fact, speculations that certain Alu templates may be driven to high levels of expression due to strong upstream promoters or enhancers have not been substantiated. Instead, the available data indicate that although many (>100) distinct Alus are transcriptionally active in human cells, they are active at uniformly low levels, and no one particular sequence appears to be overrepresented relative to others (13, 15, 41, 42). An explanation consistent with this finding could be that sequence motifs, which override the nucleosome positioning information within the left Alu monomer, account for the observed in vivo availability of Alu templates for transcription. This possibility provides an additional facet to the idea that certain Alu repeats may become transcriptionally competent as a result of their 5'-flanking regions (6, 8). Our results suggest that some 5'-flanking sequences may exert stimulatory effects by preventing the formation of a nucleosome over the Alu transcription start site. A similar idea has been suggested previously with regard to RNA polymerase II, specifically that the presence of poly(dA-dT) stretches, in the vicinity of promoters, facilitates pol II-mediated transcription in yeast due to exclusion (43) or destabilization of nucleosomes (44).

Some evidence indicates that poly(dA-dT) tracts of varying lengths differ from the regular B-type DNA helix in their conformation (45, 46). These structural differences have been seen as antagonizing the tendency of DNA to fold around the nucleosom surface, consistent with a body of experimental data which indicate that poly(dA-dT) tracts resist folding into nucleosomes (24, 36, 37). Published studies on this topic are not entirely consistent, however, showing on the one hand that

**Fig. 3.** Gel retardation and DNase I footprinting analyses. a, the extent of reconstitution for end-labeled NF1-Alu-derived templates was monitored by gel retardation assay (0.7% agarose gel). NF1 is the 600-bp parent fragment (lane 2), NF1/d is modified to eliminate the T<sub>14</sub>A<sub>11</sub> 5'-flanking sequence (lane 4), and the NF1/t (lane 6) is the parent template lacking the right monomer (see "Materials and Methods"). Reconstitutes are denoted by + and free templates by –. The marker (M) is pBR/MspI digest. DNase I footprinting patterns are as follows: the 5'-end-labeled NF1-Alu (600 bp) (b) and the NF1/t-Alu (330 bp) (c) fragments were subjected to DNase I digestion. The assays were carried out as described in Fig. 2. Digestion products were resolved in a 6% sequencing gel, and a photograph of an autoradiogram is shown. The landmarks of the Alu gene are indicated; in the case of the ~330-bp template, a major portion of the right monomer has been truncated. Solid bars mark the position of the T<sub>14</sub>A<sub>11</sub> sequence, and arrows denote DNase cleavage sites uniquely enhanced in reconstituted templates. The M lane is pBR/MspI digest, and the products of A+G sequencing reaction are shown for sequence alignment.

**Fig. 4.** DNase footprinting pattern for the NF1/d-Alu template. The NF1-Alu fragment was modified to eliminate the T<sub>14</sub>A<sub>11</sub> 5'-flanking sequence. The resulting NF1/d-Alu template was assembled with histone octamers and subjected to DNase I digestion. The assay was conducted as described above, and the digestion products were analyzed in a 6% sequencing gel. A schematic structure of the Alu gene is aligned with the footprint; left and right monomers are depicted by open bars and promoter motifs by A and B. Horizontal arrows denote modifications of the footprinting pattern for the reconstituted (Rec) template.
Fig. 5. Analysis of in vitro transcription from the NF1-Alu and the NF1/d-Alu templates. The templates were assembled with histone octamers and tested for in vitro transcription. Reaction conditions were as follows. a, the NF1-Alu template: in lanes 2 and 3, the transcription reactions were resolved in 6% sequencing gels and visualized by autoradiography. b, transcription reactions were assembled in a final volume of 25 μl with HeLa nuclear extract and 100 ng of template free, or assembled with histone octamers at the ratio of 1:0.9 DNA mass/histone mass, as expected transcript (350 bp) from the naked unlabeled template (RNA). Lanes 2 and 3, no transcript is detected when the end-labeled NF1/d-Alu template (DNA) is assembled with histone octamers. In lane 4 is a control reaction in which the [α-32P]GTP is omitted from the assay. M is the pBR322/MspI digest. poly(dA-dT) sequences can assemble into nucleosomes under certain conditions (47-50) and on the other hand that some types of poly(dA-dT) tract may always resist this type of assembly (48).

While A+T-rich regions are known to be preferred insertion sites for Alu elements (51-53), these regions vary considerably in their nucleotide composition such that only ~3% of Alu repeats in the GenBank data base are flanked by sequences comprising >10 consecutive A or T residues within the immediate 50-bp upstream region. In less than 10% of these sequences (i.e., 0.3% of the 15,000 Alus in the data base) there are runs of >10 consecutive A or T residues immediately adjacent to the Alu element, bringing the total estimated number of Alu repeats in the human genome that would be flanked by this type of sequence to ~1500 or less than 1%. At the present time, with the exception of the natural T(14)A(11) tract sequence examined here, it is not known which types of A/T tract may interfere with the in vivo nucleosome assembly over the Alu start site. Nonetheless, the notion that up to ~1500 Alu elements might be accessible to low level transcription as a result of the absence of a precisely positioned nucleosome over their transcription start sites is provocative and at the same time consistent with the estimated number of Alus believed to be transcribed constitutively in human cells. Acknowledgments—We are grateful to Drs. W. Makalowski and J. Zhang for extraction of Alu-flanking sequences from the GenBank data base. We thank Drs. R. Marais, A. Wolffe, and W. Makalowski for critical reading of the manuscript. We also thank Dr. M. Wallace for the NF1-Alu clone and D. Lee for purified chicken erythrocyte histones.

REFERENCES

1. Deininger, P. (1989) in Mobile DNA (Berg, D. E., and Howe, M. M., eds) Vol. X, pp. 619–636, American Society for Microbiology, Washington, D.C.

2. W. Makalowski, personal communication.