MOLECULAR AND CELLULAR BIOLOGY, Jan. 2008, p. 863–872 Vol. 28, No. 2
0270-7306/08/$08.00+0 doi:10.1128/MCB.00960-07

Naturally Extended CT · AG Repeats Increase H-DNA Structures and Promoter Activity in the Smooth Muscle Myosin Light Chain Kinase Gene[▽]

Yoo-Jeong Han and Primal de Lanerolle*

Department of Physiology and Biophysics, College of Medicine, University of Illinois at Chicago, Chicago, Illinois 60612

Received 30 May 2007/Returned for modification 12 July 2007/Accepted 25 October 2007

Naturally occurring repeat sequences capable of adopting H-DNA structures are abundant in promoters of disease-related genes. In support of this, we found $(CT)_{22} \cdot (AG)_{22}$ repeats in the promoter of smooth muscle myosin light chain kinase (smMLCK), a key regulator of vascular smooth muscle function. We also found an insertion mutation that adds another six pairs of CT · AG repeats and increases smMLCK promoter activity in spontaneously hypertensive rats (SHR). Therefore, we used the smMLCK promoters from normotensive and hypertensive rats as a model system to determine how CT · AG repeats form H-DNA, an intramolecular triplex, and regulate promoter activity. High-resolution mapping with a chemical probe selective for H-DNA showed that the CT · AG repeats adopt H-DNA structures at a neutral pH. Importantly, the SHR promoter forms longer H-DNA structures than the promoter from normotensive rats. Reconstituting nucleosomes on the promoters, in vitro, showed no difference in nucleosome positioning between the two promoters. However, chromatin immunoprecipitation analyses revealed that histone acetylations are greater in the hypertensive promoter. Thus, our findings suggest that the extended CT · AG repeats in the SHR promoter increase H-DNA structures, histone modifications, and promoter activity of the smMLCK, perhaps contributing to vascular disorders in hypertension.

Simple sequence repeats (SSRs), or microsatellites, are genetic loci where one to six bases are tandemly repeated for various numbers of times. They are ubiquitous in prokaryotes and eukaryotes, and they are present even in the smallest bacterial genomes (13). SSRs are inherently unstable, and one of their interesting characteristics is their mutability. SSRs mutate by inserting or deleting a few repeats, and the mutation rates generally increase in concert with the length of repeat tracts (48). Because of their high mutability, SSRs play a significant role in genome evolution by creating and maintaining quantitative genetic variation (16). Moreover, mutations occurring at SSRs are associated with neurodegenerative diseases (fragile X syndrome, Huntington's disease, and myotonic dystrophy) (32) and some human cancers, such as hereditary nonpolyposis colorectal carcinoma (50).

$(CT)_n \cdot (AG)_n$ dinucleotide sequences are among the most abundant dinucleotide repeats in the mammalian genome (43). These sequences are unique in that they form noncanonical DNA secondary structures at promoter regions of many genes. Along with the trinucleotide $(GAA)_n \cdot (TTC)_n$ or $(CTT)_n \cdot (AAG)_n$ repeats, dinucleotide $(CT)_n \cdot (AG)_n$ sequences can adopt a non-B DNA structure, an intramolecular triplex or H-DNA. These structures form at regions containing homopurine-homopyrimidine sequences, where part of the tract can dissociate into single strands using the energy provided by supercoiling. One of the single strands can then swivel its backbone parallel to the duplex, forming a three-stranded helix and leaving its complementary strand unpaired (14, 45).

Naturally occurring repeat sequences capable of adopting H-DNA structures are abundant in mammalian genomes and are found as frequently as 1 in every 50,000 bp in the human genome (37). They are mutagenic in mammalian cells (46) and are preferentially located at promoter regions of many genes, including proto-oncogenes such as human c-*myc* (8, 18), mouse Ki-*ras* (29), and human c-*Src* (33). Because of this preferential distribution, $(CT)_n \cdot (AG)_n$ repeats have been implicated in the formation of H-DNA structures and the regulation of gene expression. Indeed, $(CT)_n \cdot (AG)_n$ dinucleotide sequences function as a positive transcription element to stimulate gene expression (23), and deleting these sequences reduced promoter activity of heat shock genes in *Drosophila melanogaster* (35). Moreover, the *trans*-acting GAGA factor has been shown to bind dinucleotide repeats and to stimulate promoter activities of heat shock genes in *Drosophila melanogaster* by establishing nucleosome-free DNase-hypersensitive sites (19, 20).

Nucleosome structure and positioning are also critical in promoter and enhancer regions because nucleosomes limit accessibility to regulatory factors (21). The packaging of promoter DNA into nucleosomes has been thought to be a major obstacle to transcription. However, many recent reports have shown that chromatin structure is dynamic (1) and a vital component of transcription (42). The dynamic state of chromatin is regulated partly by posttranslational modifications of histones. Residues with histone tails are subject to diverse posttranslational modifications, including phosphorylation, acetylation, and methylation, which together govern the higher-order structure of chromatin and gene expression. Of the modifications, histone acetylation/deacetylation is highly dynamic (11). Acetylation of conserved lysine residues in histone

* Corresponding author. Mailing address: Department of Physiology and Biophysics, University of Illinois at Chicago, 835 S. Wolcott, Chicago, IL 60612. Phone: (312) 996-6430. Fax: (312) 996-1414. E-mail: primal@uic.edu.

▽ Published ahead of print on 8 November 2007.

tails stimulates gene expression by neutralizing positive charges, resulting in unstable histone-histone and histone-DNA interactions that otherwise limit the binding of transcription factors to DNA. While it is well established that acetylation of the core histone plays a fundamental role in transcriptional regulation, there is almost no literature on how histone modifications interact with DNA structure, especially non-B DNA structures like H-DNA.

Therefore, we studied the promoter of the smooth muscle MLCK (smMLCK) gene. smMLCK regulates smooth muscle contraction by phosphorylating myosin light chains (5). Myosin light chain phosphorylation also regulates cell motility (49) and a host of other cellular responses, including cell spreading and cytokinesis (15). smMLCK is an important regulatory enzyme in the vasculature, and we recently showed that it plays a central role in hypertension. The expression of smMLCK is developmentally regulated and correlates with the onset of hypertension in spontaneously hypertensive rats (SHR), and inhibiting smMLCK expression or activity prevents the development of hypertension (12). Importantly, the smMLCK promoter in normotensive Wistar-Kyoto (WKY) rats contains a dinucleotide repeat motif consisting of 22 (CT) · (AG) repeats. SHR, in contrast, have six additional repeats that increase promoter activity and responsiveness to serum response factor (12), an important regulator of vascular smooth muscle differentiation (28). Because they only differ by these six (CT) · (AG) repeats, the promoters from SHR and WKY rats provide a unique opportunity to determine the effects of nucleotide repeats on gene structure and promoter activity and to relate those results to an important disease. We report below that the CT · AG repeats in both promoters form H-DNA structures without affecting their ability to be packed into nucleosomes and that the extended repeats in the SHR promoter expand the H-DNA structure and increase histone modifications. This is, to the best of our knowledge, the first report on the role of H-DNA in regulating gene expression in a vascular disorder.

MATERIALS AND METHODS

Cell culture. A7r5 vascular smooth muscle cells isolated from embryonic rat thoracic aorta were obtained from Kenneth Byron at Loyola University (Chicago, IL) and cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and 1% nonessential amino acids.

Plasmids. The proximal regions of smMLCK promoters were isolated from genomic DNA obtained from SHR and normotensive WKY rats and cloned into a pGL3-Basic firefly luciferase vector, as described previously (12). The smMLCK promoter from WKY rats has $(CT)_{22} \cdot (AG)_{22}$ repeats, while the promoter from SHR rats has six additional repeats. The repeat sequences were counted from the −254 nucleotide at the WKY promoter but from −255 at the SHR promoter because of one nucleotide difference at position −254 between the two promoters. Progressive deletion fragments of the regions from −425 bp to −162 bp of the normotensive promoter were also obtained by PCR and cloned into the pGL3-Basic vectors (12). The ΔCT promoter was constructed by deleting $(CT)_{22} \cdot (AG)_{22}$ in the normotensive promoter using the QuikChange XL site-directed mutagenesis kit (Stratagene, La Jolla, CA). The integrity of these constructs was confirmed by DNA sequencing using an ABI Prism 3100 genetic analyzer in the Research Resource Center at University of Illinois at Chicago. Various SSRs, flanked by SmaI and HindIII restriction enzyme sites, were also cloned into the pGL3-Enhancer vector. The sequences of $(CT)_3$, $(CT)_9$, $(CTT)_6$, and $(C)_{18}$ SSRs were 5′-GAACCCGGGCTCTCTAAGCTTACG-3′, 5′-GAACCCGGGCTCTCTCTCTCTCTCTCTAAGCTTACG-3′, 5′-GAACCCGGGCTTCTTCTTCTTCTTCTTAAGCTTACG-3′, and 5′-GAACCCGGGCCCCCCCCCCCCCCCCCCAAGCTTACG-3′, respectively. The oligonucleotides were digested with the SmaI and HindIII restriction enzymes and ligated into the pGL3-Enhancer vector cut by the same enzymes. The recombinant clones were selected and subjected to DNA sequencing to confirm their integrities.

Reporter activity assays. For dual luciferase reporter gene assays, A7r5 cells grown in 12-well plates were cotransfected with 500 ng of the smMLCK promoter-firefly luciferase vector and 10 ng of the TK-renilla luciferase vector (Promega, Madison, WI) using the Fugene 6 transfection reagent (Roche, Basel, Switzerland). One day after transfection, the cells were harvested and luciferase activity was measured using Luciferase Assay Reagent II and Stop&Glo reagent (Promega, Madison, WI). Variations in transfection efficiency were normalized by dividing the firefly luciferase activity by the *Renilla* luciferase activity.

Low-resolution mapping of S1 nuclease sensitivity. Supercoiled plasmids (1 μg) were incubated with 1 U of S1 nuclease at 30°C for 10 min in a 20-μl reaction volume containing 40 mM sodium acetate, pH 4.5, 300 mM NaCl, and 2 mM $ZnSO_4$. The reaction was stopped by phenol-chloroform extraction, and DNA was precipitated with ethanol. To determine the nucleotides sensitive to S1 nuclease, primer extension analysis was carried out as described previously (44) with slight modifications. Briefly, 5 pmol of the RV primer or −425 primer (5′-AAGCCTAGCCAGGTCTCCCAC-3′) was labeled at the 5′ end using T4 polynucleotide kinase (Fermentas, Inc., Hanover, MD) and [γ-^{32}P]ATP (Amersham Biosciences, Piscataway, NJ) for 30 min at 37°C. The S1-nuclease-digested DNA was denatured at 94°C for 4 min and annealed with the labeled primer for 10 min at 50°C. Polymerization was carried out at 70°C for 20 min using a *Taq* DNA polymerase (Invitrogen Life Technologies, Carlsbad, CA). The reaction mixtures were denatured and loaded on a 1% agarose gel or a 7% polyacrylamide gel containing 7 M urea. The sizes of the primer extension products were determined by comparison with a 10-bp DNA ladder.

High-resolution mapping of chemical probes. Chemical probe analysis under conditions of one hit per molecule was performed as previously described (31). Unpaired regions in supercoiled plasmids (0.25 μg of DNA) in 40 μl of 0.5× TEN buffer (5 mM Tris-HCl, 25 mM NaCl, 0.5 mM EDTA, pH 7.6) were probed with 0.5% chloroacetaldehyde (CAA) for 8 min. CAA reactions were terminated by diethyl ether extraction. After ethanol precipitation, 0.1 μg of DNA was used to reveal chemically susceptible bases by the radiolabeled primer extension technique using a Stoffel fragment of *Taq* polymerase that terminates synthesis at the chemically modified nucleotides. Primer extension products were separated on a 7% sequencing gel alongside a DNA sequence pattern obtained by sequencing unmodified DNA with the same primers and *ReaderTaq* DNA polymerases (Fermentas, Hanover, MD).

Analysis of reconstituted nucleosomes with micrococcal nuclease (MNase) treatment. H1-stripped core histones purified from HeLa cells were obtained from Vaxron Corp. (Rockaway, NJ). Nucleosomes were reconstituted onto plasmids by a high-salt exchange method, according to the manufacture's instructions. Briefly, supercoiled DNA (4 μg) was incubated with 4 μg of HeLa core histones in the presence of 1 M NaCl. The salt concentration was decreased to 100 mM by stepwise addition of 40 mM HEPES-NaOH, pH 7.4, 0.1 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride and 1 mM dithiothreitol. The reconstituted nucleosomes were treated with 1 U of MNase at 37°C for 2.5 min. The nucleosomes and MNases were then removed from the plasmid DNA by phenol-chloroform extraction. After ethanol precipitation, DNA was subjected to primer extension as described above (see "Low-resolution mapping of S1 nuclease sensitivity"), and the products were analyzed on a 7% polyacrylamide gel or a sequencing gel.

ChIP assays. A7r5 cells transfected with a pGL3 vector containing an smMLCK promoter or SSRs were fixed in formaldehyde and subjected to chromatin immunoprecipitation (ChIP) assays as previously described (4) with the following modifications. Antibodies to acetyl-histone H3 on Lys 9 and 14 (no. 06-599) or acetyl-histone H4 on Lys 5, 8, 12, and 16 (no. 06-866) were obtained from Upstate (Lake Placid, NY). Total histone H3 antibodies (ab1791) were obtained from Abcam (Cambridge, MA). The antibodies were prebound for a minimum of 4 h to protein A-Dynal magnetic beads (Invitrogen, Carlsbad, CA), added to the diluted chromatin, and immunoprecipitated overnight. After the magnetic bead-chromatin complexes were washed six times in radioprecipitation assay buffer, DNA was eluted from the beads as previously described (40). The PCR was performed using the primers targeting CT · AG repeats of smMLCK promoter (5′-GAGGGGGTGATGTGTGTTTCT-3′ and 5′-CGGCCAGGCTGTGTTTATAGA-3′), primers targeting nucleotides 2636 to 2656 (5′-CAACCCGGTAAGACACGACT-3′) and 2768 to 2788 (5′-GCAGAGCGCAGATACCAAAT-3′) of the pGL3-Basic vector, or primers targeting nucleotides 4787 to 4807 (5′-AGTGCAGGTGCCAGAACATT-3′) and 76 to 96 (5′-TCTTCCATGGTGGCTTTACC-3′) of the pGL3-Basic vector. To quantitatively analyze PCR products, PCR was carried out for 15 cycles using a ^{32}P-labeled primer. The PCR

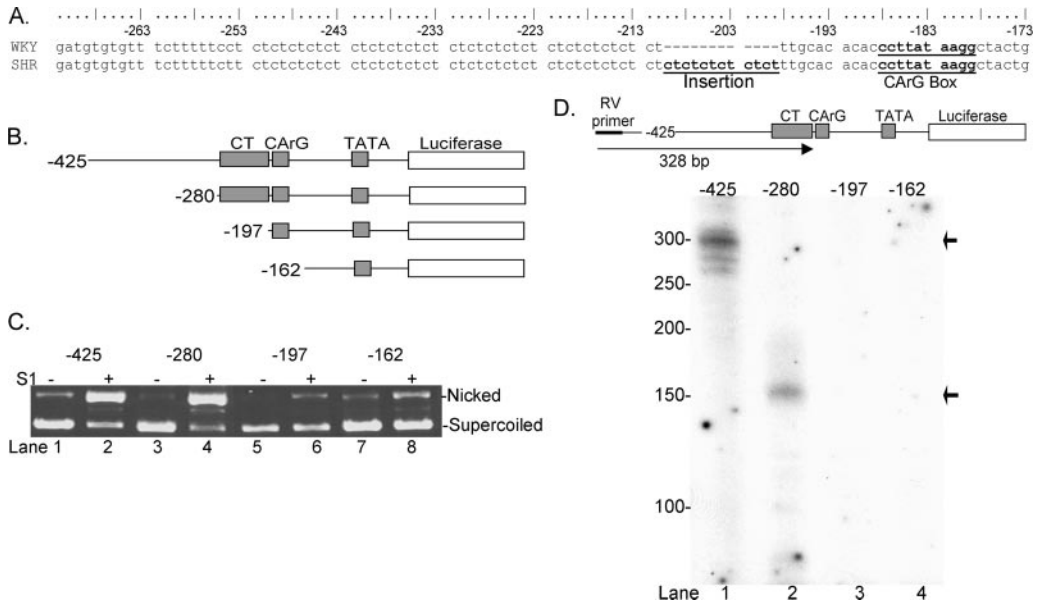

FIG. 1. Low-resolution mapping of S1 nuclease sensitivity. (A) DNA sequences of smMLCK promoters isolated from SHR and normotensive WKY rats. Comparison of these sequences revealed the presence of a 12-bp insertion (underlined) in the sequence from SHR not found in the sequence from the normotensive WKY rats. The numbers indicate the locations of nucleotides relative to the transcription start site (+1) of the smMLCK gene. CArG box indicates a *cis*-acting element for serum response factor. (B) Progressive-deletion mutants were constructed by deleting the regions from bp −425 to bp −162 of the smMLCK promoter isolated from the normotensive WKY rats. These mutants were cloned into the pGL3-Basic luciferase vector and subjected to S1 nuclease treatment. The numbers indicate the location of the first nucleotide in each fragment. CT, CArG, and TATA represent $(CT)_n \cdot (AG)_n$ repeats, CArG box, and TATA box, respectively. These deletion mutants were used in Panels C and D. (C) Agarose gel analysis was performed on the plasmid DNA treated with (lanes 2, 4, 6, and 8) or without (lanes 1, 3, 5, and 7) S1 nuclease. S1 nuclease changed plasmid conformations from supercoiled to nicked forms, as indicated by slower-migrating bands. However, deleting CT repeats abolished the sensitivity to S1 nuclease. (D) Primer extension was performed using the DNA treated with the S1 nuclease. RV primers were labeled at the 5′ end and hybridized with the S1-nuclease-digested DNA. After primer extension, the product was denatured and separated on 7% denaturing polyacrylamide gels that were designed to resolve products that are less than 1 kb. The products were identified by autoradiography, and their sizes were determined by comparison with a ^{32}P-labeled 10-bp ladder. Note the presence of products (arrows) in lanes 1 and 2 and the absence of products in the plasmids not containing $(CT)_n \cdot (AG)_n$ repeats (lanes 3 and 4).

products were separated on a 7% polyacrylamide denaturing urea gel, and the intensities of the bands were quantified using a densitometer.

RESULTS

S1 nuclease mapping of the smMLCK promoter. We previously isolated smMLCK promoters from hypertensive (SHR) and normotensive (WKY) rats and identified their DNA sequences (12). The smMLCK promoter from SHR has a 12-bp insertion consisting of 6 pairs of (CT) · (AG) repeats compared to the sequence of the promoter from the WKY normotensive rats (Fig. 1A). These repeats reside between −255 bp and −199 bp relative to the transcription start site (+1) of the smMLCK gene. They are also adjacent to a CArG box, a binding site for serum response factor, an important regulator of smooth muscle differentiation (28). To investigate the possibility that the $(CT)_n \cdot (AG)_n$ sequences of the smMLCK promoter form H-DNA structures, mutants were constructed by progressively deleting the regions from bp −425 to bp −162 of the smMLCK promoter isolated from normotensive rats and cloning the fragments into the pGL3-Basic luciferase reporter plasmid (Fig. 1B). These mutant plasmids were then treated with S1 nuclease. Since H-DNA structures adopt a three-stranded helix, leaving one strand unpaired, S1 nuclease attacks the unpaired strand and changes the plasmid conformation from a supercoiled form into a nicked form. Nicked DNA can be separated from supercoiled DNA by agarose gel electrophoresis because the uncoiled conformation retards their mobility. Figure 1C shows that more than 90% of these plasmids were in the supercoiled form prior to S1 nuclease digestion (Fig. 1C, lanes 1, 3, 5, and 7). S1 nuclease, however, changed only the migration of the plasmids containing the CT · AG repeats (Fig. 1C, lanes 2 and 4). There were no changes in the migration of the plasmids that did not contain the CT · AG repeats (Fig. 1C, lanes 6 and 8).

We further investigated S1 nuclease sensitivity by performing primer extension and analyzing the extension products on 7% polyacrylamide gels. Because primer extension produces a short DNA fragment that terminates at the nucleotide cut by S1 nucleases, measuring the sizes of the products makes it possible to approximately map the nuclease-sensitive site. Therefore, we designed a forward primer corresponding to RV primer sequences on the coding strand of the pGL3-Basic vector, because all deletion constructs have these sequences in common (Fig. 1D). The primer was labeled at the 5′ end using T4 polynucleotide kinase and [γ-^{32}P]ATP. Primer extension with the constructs containing $(CT)_n \cdot (AG)_n$ repeats produced short DNA fragments (Fig. 1D, lanes 1 and 2). In contrast, there were no products in the plasmids that did not contain $(CT)_n \cdot (AG)_n$ sequences (Fig. 1D, lanes 3 and 4). The sizes of the fragments were approximately 300 bp and 160 bp

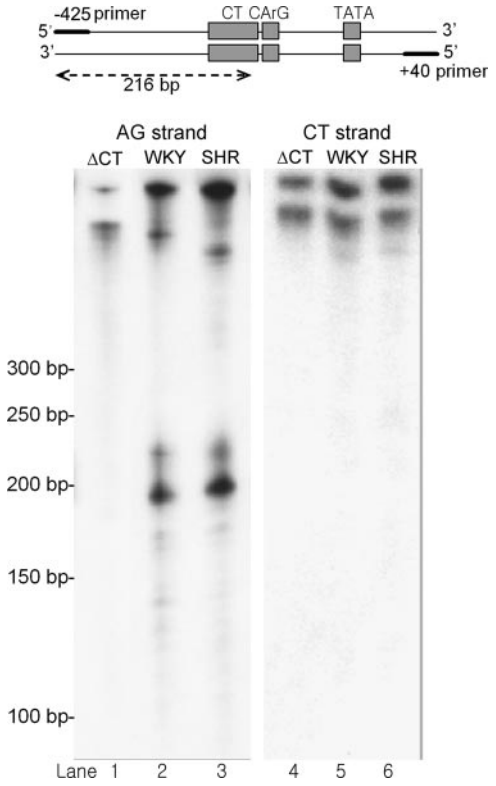

FIG. 2. The smMLCK promoter adopts the H-y structure of H-DNA. The possibility of adopting the H-y or H-r structure was determined by performing primer extension using two separate primers (−425 and +40 primers) targeting each strand of the double helix. SHR and WKY represent the promoters isolated from hypertensive or normotensive rats, respectively. A normotensive promoter without the $(CT)_n \cdot (AG)_n$ sequences was used as a negative control (ΔCT). The data show that the S1 nuclease attacked only $(AG)_n$ sequences, leaving $(CT)_n$ sequences intact.

when plasmids containing nucleotides −425 to +40 and −280 to +40 were used as templates, respectively (Fig. 1D, lanes 1 and 2). Considering that $(CT)_n \cdot (AG)_n$ repeats reside between 282 and 328 bp downstream of the RV primer sequence, these data indicate that the S1 nuclease-sensitive site resides within the repeats.

We next examined which strand was involved in triplex formation and which strand was left unpaired. Homopurine-homopyrimidine repeat sequences adopt either H-y (pyrimidine-purine-pyrimidine triplex) (10) or H-r (pyrimidine-purine-purine triplex) (34) conformations under physiological conditions. We tested this possibility using two primers targeting each strand of the double helix. A forward primer corresponding to nucleotides −425 to −405 on the coding strand was designed to detect S1 digestion on the anticoding strand where the $(AG)_n$ repeats reside. A reverse primer corresponding to nucleotides +40 to +19 on the anticoding strand was also designed to detect S1 digestion on the coding strand where the $(CT)_n$ repeats reside (Fig. 2). These primers were end labeled using [γ-^{32}P]ATP. The smMLCK promoter regions isolated from SHR or normotensive rats were cloned into the pGL3-Basic luciferase reporter gene (12) and used for primer extension analyses. A mutant plasmid that has the same sequences as the normotensive plasmid but does not contain the $(CT)_n \cdot (AG)_n$ sequences was created by site-directed mutagenesis and used as a negative control (ΔCT). Primer extension with the forward primer produced approximately 200-bp fragments (Fig. 2, lanes 2 and 3) in plasmids with the normotensive or hypertensive promoters but not in the ΔCT plasmid. In contrast, there were no products when the extension was performed with the reverse primer (lanes 5 and 6 in Fig. 2). These data confirmed that the repeat in the smMLCK promoter adopts an H-y conformation.

High-resolution mapping using chemical probe sensitivity. CAA was used next to precisely analyze the structure at a neutral pH. CAA has been widely used as a chemical probe to detect the unpaired strand of H-DNA structures (30, 31), because it preferentially interacts with single-stranded adenines and cytosines forming their ethenoderivatives (47). The resulting derivatives were subjected to primer extension using a Stoffel fragment of *Taq* polymerase that is blocked by the derivatized nucleotides (30). The end-labeled forward primer corresponding to nucleotides −425 to −405 on the coding strand was used for the reaction, and the extension products were analyzed on 7% denaturing polyacrylamide sequencing gels (Fig. 3). Primary data on chemical modification are presented in Fig. 3A and summarized graphically in Fig. 3B. A total of 10 adenines were modified by CAA, beginning with the 6th and extending to the 15th AG repeat in the normotensive promoter (Fig. 3A, lane 6). The H-DNA structure was further expanded to the 21st AG repeat in the SHR promoter, making a total of 16 adenines modified by CAA (Fig. 3A, lane 7). There were no CAA-modified derivatives in the ΔCT promoter (Fig. 3A, lane 5). These data demonstrate that extending the $(CT)_n \cdot (AG)_n$ repeats expands the H-DNA structure in SHR.

Nucleosome positioning on H-DNA. We then investigated the effects of H-DNA structures in changing nucleosome organization or positioning. Rotational positioning of DNA about the histone octamer appears to be determined by certain sequence-dependent modulations of DNA structures. In a triple-stranded complex, the major groove of the duplex, which is occupied by the third strand, would not be accessible to histones (7). Therefore, we assembled histone octomers on the plasmids containing the smMLCK promoters and analyzed the histone-DNA complex by digesting them with MNases. Reconstituted nucleosomes were treated with various concentrations of MNases (Fig. 4A). Naked DNA without histones was used as a control. The digested DNA was isolated, labeled with [^{32}P]ATP, denatured, and separated on a 7% polyacrylamide gel electrophoresis (PAGE) gel containing 7 M urea. The plasmids without histones were fragmented into progressively smaller fragments as the MNase concentration was increased (Fig. 4A, lanes 3 to 5). In contrast, the plasmids with assembled histones produced a band of ~150 bp regardless of the MNase concentration (Fig. 4A, lanes 6 to 8). Because histone octomers are reported to span approximately 147 bp (24), the size of this fragment suggests that it comes from DNA protected by the histones. Moreover, there were weak smears at ≤450, 300, and 170 bp as the MNase concentration was increased (Fig. 4A, lanes 6 to 8), suggesting that these were the linker DNA digested by the MNase.

Based on these data, we performed primer extension anal-

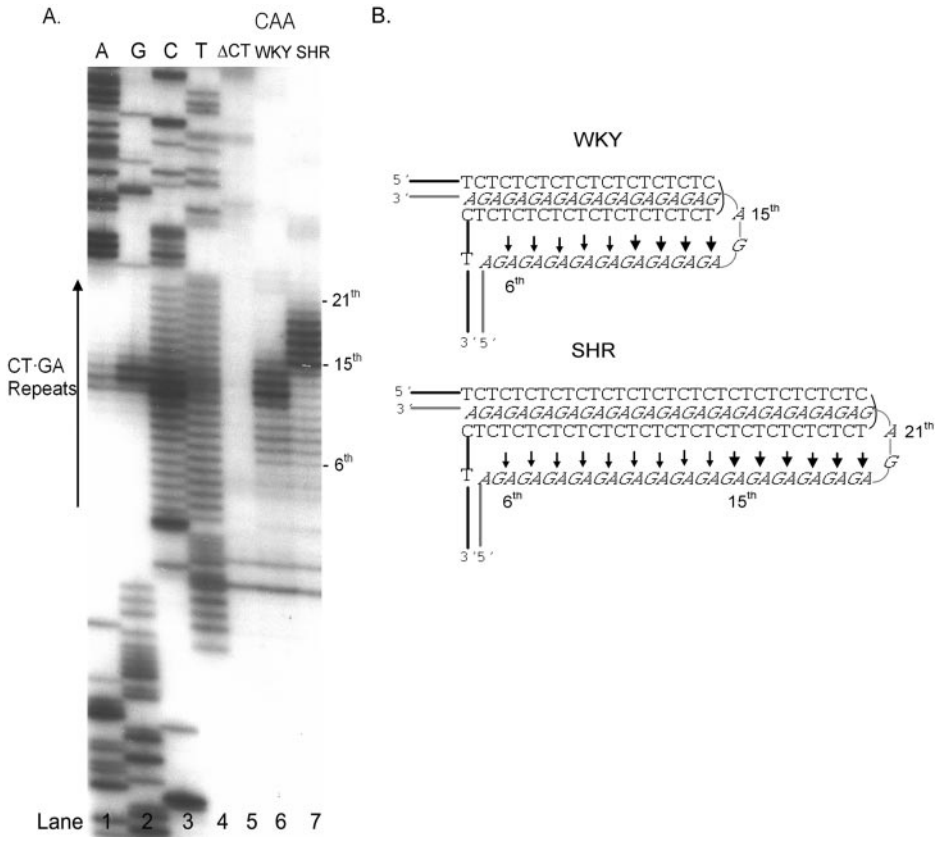

FIG. 3. Expanded H-DNA at the SHR promoter. (A) Plasmids were probed with CAA and used to reveal chemically susceptible bases by primer extension using a Stoffel fragment of *Taq* polymerase. Primer extension products from the modified (lanes 5 to 7) DNA were separated on a 7% sequencing gel along with a DNA sequence marker (lanes 1 to 4). The $(CT)_n \cdot (AG)_n$ repeats are indicated with an arrow. The positions of the chemically modified adenines are indicated with numbers. Note that the SHR promoter (SHR) adopts expanded H-DNA compared to the normotensive promoter (WKY), while the ΔCT promoter did not form H-DNA (lane 5). (B) The sequencing data in panel A are summarized graphically. The arrows indicate the nucleotides modified by CAA. The sizes of the arrowheads indicate whether the nucleotides were modified to a greater (large arrows) or lesser (small arrows) extent.

yses using nucleosomes that were partially digested with 1 U of MNase (Fig. 4B). The digested DNA was separated from histones, denatured, and annealed with the −425 primer labeled with ^{32}P. After extending the primer at 70°C for 20 min using a *Taq* DNA polymerase, the products were analyzed using 7% PAGE (Fig. 4B) or sequencing gels (Fig. 4C). In the absence of MNase treatment, all of the primer products were retained on the top of the well and there were no short fragments. Because 7% PAGE resolves fragments of <1 kb, these data suggest that the polymerase did not terminate prematurely and extended through the repeats (Fig. 4B, lanes 2 and 4). In contrast, primer extension with the digested nucleosomes produced short fragments below 160 bp (Fig. 4B, lanes 3 and 5). Since the primer (−425) is located 170 to 226 bp upstream of the $(CT)_n \cdot (AG)_n$ sequences (see diagram in Fig. 4B) and the MNase-sensitive sites (linker region) are up to 160 bp from the primer, the data show that the $(CT)_n \cdot (AG)_n$ sequences are protected from MNase digestion by the histones. Further analysis using sequencing gels confirmed that histones assembled on the $(CT)_n \cdot (AG)_n$ repeats in both promoters (Fig. 4C, lanes 3 and 4). There was no significant difference in nucleosome positioning between the promoters from SHR and WKY rats.

Increased histone acetylation and promoter activity at the SHR promoter. To determine if histone modifications are changed on H-DNA, we transfected A7r5 cells with the pGL3-Basic vectors containing the smMLCK promoters and performed ChIP assays using antibodies to histone H3, acetylated histone H3, or acetylated histone H4. These antibodies allowed us to semiquantitatively compare the amounts of total and acetylated histone bound to the DNA and to confirm that the transfected plasmids were assembled by histones and suitable for ChIP assays. PCR was performed on immunoprecipitated chromatin-DNA complexes using primers targeting a 190-bp region around the $(CT)_n \cdot (AG)_n$ repeats of the smMLCK promoter or primers targeting nucleotides 2636 to 2788 of the pGL3-Basic vector (Fig. 5A). Standard PCR conditions were modified in several ways to quantitatively measure the amount of PCR products. First, PCR cycles were restricted to 15 to avoid saturating the products. Second, one primer was end labeled with ^{32}P to increase the sensitivity. Third, the PCR products were resolved on 7% denaturing polyacrylamide gels. The specific products were identified by autoradiography and quantified densitometrically. Figure 5B shows that there is no significant difference in the enrichment of total histone H3 between the WKY and SHR promoters, both in

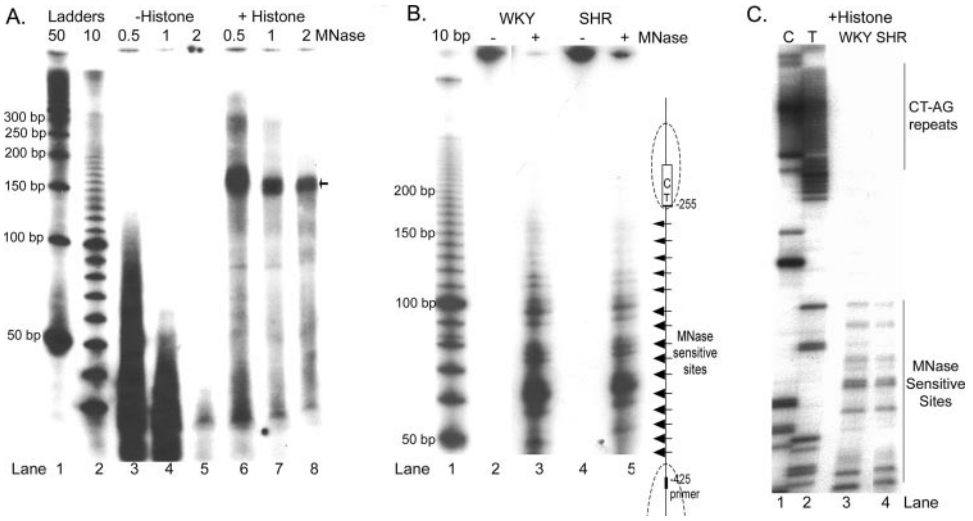

FIG. 4. Nucleosome positioning on the H-DNA. Core histones purified from HeLa cells were reconstituted onto plasmids by a high-salt exchange method and treated with MNase. (A) The nucleosome sensitivity to MNase was tested by treating with 0.5, 1, or 2 U of MNase at 37°C for 2.5 min. Ladders of 50 bp (lane 1) and 10 bp (lane 2) were used to indicate the sizes of the DNA fragments. The DNA without histones (lanes 3 to 5) was digested into shorter fragments with increasing concentrations of MNase, while the DNA assembled by a histone octamer (lanes 6 to 8) was protected from the MNase and produced a band of ~150 bp regardless of the MNase concentration. (B and C) DNA partially digested with 1 U of MNase was separated from histones, subjected to primer extension, and analyzed on 7% polyacrylamide (B) or sequencing (C) gels. The diagram in panel B shows the corresponding DNA sites digested by MNase (arrows) or assembled by histones (ovals). The primer and the CT repeats are located at nucleotides −425 and −255 in the smMLCK gene, respectively. The data show that histone assembly protects the CT · AG sequences from MNase digestion of both promoters.

the CT repeats and in the control vector sequences (lanes 3 and 4). However, the amounts of acetylated histone H4 and H3 were greater in the $(CT)_n \cdot (AG)_n$ sequences of the SHR promoter (lanes 6 and 8) compared to the normotensive promoter (lanes 5 and 7). The ratio of acetylated histone H3 to total histone H3 was 0.483 ± 0.11 in the SHR promoter and 0.126 ± 0.07 in the WKY promoter (Fig. 5C). These data show that histone modifications are greater on the longer repeats in the SHR promoter.

The importance of the repeats in regulating promoter activity was investigated using luciferase reporter gene assays. Firefly luciferase constructs containing SHR or WKY promoters were transfected into A7r5 cells along with TK-renilla luciferase vectors. Luciferase activities were measured and firefly activity was normalized to the *Renilla* activity. Promoter activity was 2.65-fold ± 1.057-fold greater with the SHR promoter than with the WKY promoter (Fig. 5D). Together, these data demonstrate that inserting six more CT · AG repeats expanded the length of the H-DNA structure and increased histone acetylation and promoter activity of the SHR promoter.

We further examined the correlation between histone acetylation and CT · AG repeats using an artificial system. Because it has been reported that inserting CT repeats increased the luciferase activity of pGL3-Enhancer (51), we synthesized $(CT)_3 \cdot (AG)_3$ or $(CT)_9 \cdot (AG)_9$ sequences and cloned them into the pGL3-Enhancer vector (Fig. 6A). ChIP assays showed that the enrichment of acetylated histone H4 was greater in the $(CT)_9 \cdot (AG)_9$ sequences than in the $(CT)_3 \cdot (AG)_3$ sequences (Fig. 6B, lanes 5 and 6). The ratio of acetylated histone H4 to total histone H3 was 0.203 ± 0.04 in the $(CT)_9 \cdot (AG)_9$ sequences and 0.052 ± 0.02 in the $(CT)_3 \cdot (AG)_3$ sequences (Fig. 6C). Luciferase activity was also significantly greater in cells transfected with the plasmids containing $(CT)_9 \cdot (AG)_9$ sequences than those with $(CT)_3 \cdot (AG)_3$ sequences (Fig. 6D). These data add substantial support to the idea that inserting six more CT · AG repeats increases histone acetylation and promoter activity of the artificial vectors.

We then rearranged the $(CT)_9 \cdot (AG)_9$ sequences by inserting $(CTT)_6 \cdot (AAG)_6$ or $(C)_{18} \cdot (G)_{18}$ sequences and examined if other SSRs could change histone acetylation and promoter activity. We used trinucleotide repeats of $(CTT)_6 \cdot (AAG)_6$ as alternative H-DNA-forming sequences and mononucleotide repeats of $(C)_{18} \cdot (G)_{18}$ as SSRs that do not form H-DNA (Fig. 7A). While slightly less acetylated histone H4 was associated with the $(CTT)_6 \cdot (AAG)_6$ sequences than with the $(CT)_9 \cdot (AG)_9$ sequences, the level of acetylated H4 was increased in the H-DNA-forming sequences (Fig. 7B, lanes 7 and 8) compared to the level in the $(C)_{18} \cdot (G)_{18}$ sequences (Fig. 7B, lane 9). The ratio of acetylated histone H4 to total histone H3 was 0.203 ± 0.04 and 0.131 ± 0.03 in the $(CT)_9 \cdot (AG)_9$ and $(CTT)_6 \cdot (AGG)_6$ sequences, respectively, compared to 0.002 ± 0.0005 in the $(CT)_3 \cdot (AG)_3$ sequences (Fig. 7C). Luciferase activities were also greater for $(CT)_9 \cdot (AG)_9$ and $(CTT)_6 \cdot (AAG)_6$ constructs than for the $(C)_{18} \cdot (G)_{18}$ construct (Fig. 7D). These data indicate that although there are some differences between the two H-DNA-forming sequences, the H-DNA-forming sequences have greater histone acetylation and hence greater promoter activity than the non-H-DNA-forming sequences.

DISCUSSION

The discovery that abnormalities in the DNA structure itself can drive genetic instability has changed our understanding of some genetic diseases in humans. Long tracts of triple repeats,

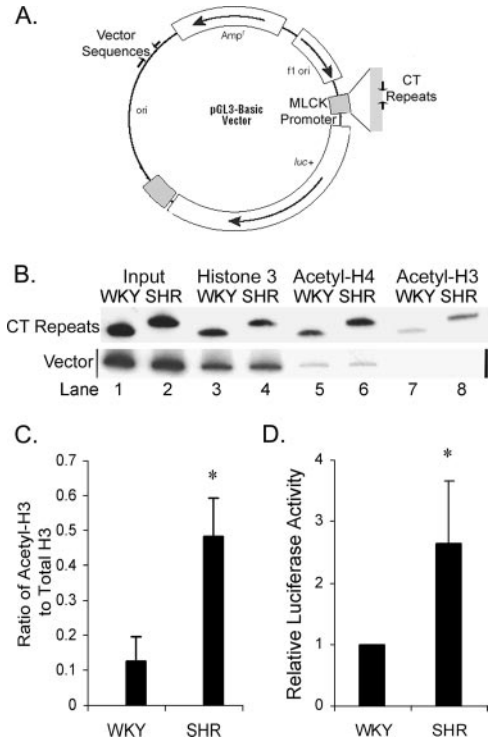

FIG. 5. Increased histone acetylation and promoter activity at the smMLCK promoter from SHR. (A) Schematic of the pGL3-Basic vector containing the smMLCK promoter. Short arrows indicate the primers targeting the CT · AG repeats of the smMLCK promoter and the primers targeting the vector sequences (2636 to 2788) of the pGL3-Basic plasmid. (B) Vectors containing WKY and SHR promoter sequences were transfected into A7r5 cells and subjected to ChIP assays. Cross-linked chromatin was immunoprecipitated with antibodies to total histone H3 (lanes 3 and 4), acetyl-histone H4 (lanes 5 and 6), or acetyl-histone H3 (lanes 7 and 8). Input chromatin (0.2% of total) (lanes 1 and 2) and immunoprecipitated DNA (5% of total) (lanes 3 to 8) were amplified with the primers specific for the smMLCK promoter or vector sequences. Histone acetylation was greater in the SHR promoter (SHR) than in the normotensive promoter (WKY). The difference in band mobility is due to the 12-bp difference between the WKY and SHR promoters. (C) The enrichment of acetylated histone H3 relative to total histone H3 was analyzed after densitometrically quantifying the intensities of the bands shown in panel B. Note that the ratio of acetyl-H3 to total H3 is higher in the SHR promoter than in the WKY promoter (*, $P < 0.05$). These experiments were repeated three times, and the means ± standard deviations are shown. (D) The activities of promoters from SHR or normotensive (WKY) rats were analyzed using luciferase reporter gene assays. The promoter activity was 2.65-fold ± 1.057-fold greater in cells transfected with the SHR promoter than in cells transfected with the normotensive promoter. This experiment was repeated four times, and the means ± standard deviations are shown. *, $P < 0.05$ compared to results for the normotensive promoter.

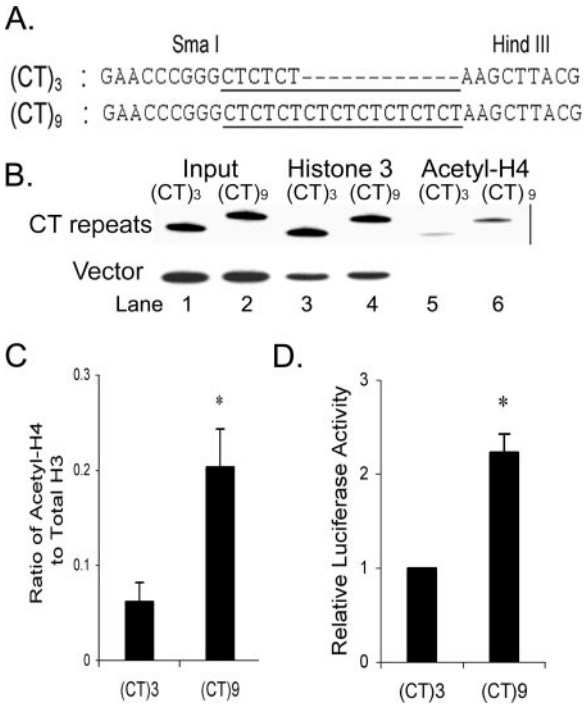

FIG. 6. Increased histone acetylation and promoter activity at the long CT · AG repeats. (A) Oligonucleotides containing $(CT)_3 \cdot (AG)_3$ or $(CT)_9 \cdot (AG)_9$ sequences were synthesized and cloned into pGL3-Enhancer vectors using SmaI/HindIII restriction enzyme sites as shown. (B) ChIP assays were performed on A7r5 cells transfected with the CT_3 and CT_9 repeat constructs using antibodies to total histone H3 (lanes 3 and 4) or acetyl-histone H4 (lanes 5 and 6). Input chromatin (0.15% of total) (lanes 1 and 2) and immunoprecipitated DNA (5% of total) (lanes 3 to 6) were amplified using the primers targeting the CT · AG repeats or the primers targeting nucleotides 2636 to 2788 of the pGL3-Basic vector (see Fig. 5A). The amount of acetylated histone H4 was greater in the $(CT)_9 \cdot (AG)_9$ sequences than in the $(CT)_3 \cdot (AG)_3$ sequences. The difference in band motility is due to the 12-bp difference between the $(CT)_3 \cdot (AG)_3$ and $(CT)_9 \cdot (AG)_9$ sequences. (C) The enrichment of acetylated histone H4 relative to total histone H3 was analyzed. Note that the ratio of acetyl-H4 to total H3 is higher in the $(CT)_9 \cdot (AG)_9$ sequences than in the $(CT)_3 \cdot (AG)_3$ sequences (*, $P < 0.01$). This experiment was repeated three times, and the means ± standard deviations are shown. (D) The luciferase activity of the $(CT)_9 \cdot (AG)_9$ construct was 2.29-fold ± 0.32-fold greater than the activity of the $(CT)_3 \cdot (AG)_3$ construct. This experiment was repeated four times, and the means ± standard deviations are shown. *, $P < 0.01$ compared to results for the $(CT)_3 \cdot (AG)_3$ construct.

which cause genetic instability, are found in the DNA of patients afflicted with more than 15 genetic neurodegenerative disorders. The repeats may be expanded from the normal length up to hundreds or thousands of repeats in diseases. For instance, approximately 98% of Friedreich's ataxia patients have an expanded $(GAA)_n \cdot (TTC)_n$ repeat in the first intron of the Friedreich's ataxia gene (2). Normal alleles have ~33 repeats of GAA · TTC sequences and premutation alleles have 34 to 65 repeats, whereas Friedreich's ataxia-associated alleles have more than 66 repeats (6). These H-DNA-forming sequences are frequently found at promoter regions, and targeting them could be an effective strategy for developing novel therapeutics. Indeed, synthetic triplex-forming oligonucleotides (TFOs) have been studied as a promising carrier for gene therapy. TFOs specifically recognize H-DNA-forming sequences and form an intermolecular triplex by means of Hoogsteen hydrogen bonding of a third strand of DNA to purine-rich regions of the duplex DNA (39). The formation of TFO-induced intermolecular triplex structures can directly inhibit gene transcription and proliferation of carcinoma cells (25). They also induce apoptosis in cancer cells (3, 41). Thus, TFO-based gene therapy could be an effective tool for treating many diseases.

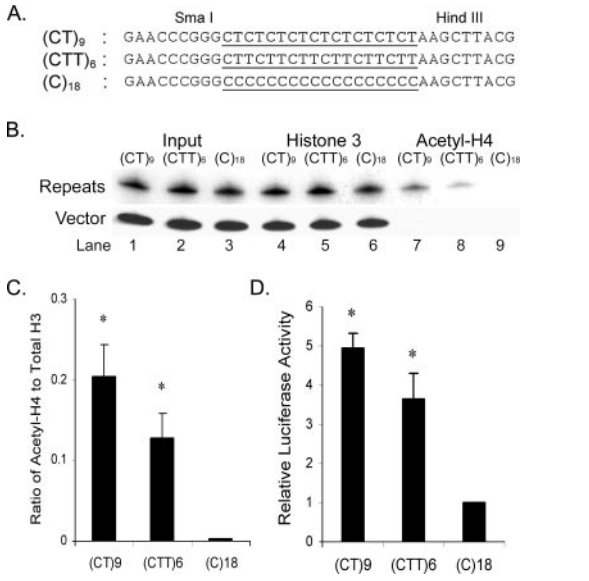

FIG. 7. Histone acetylation and promoter activity at other simple sequence repeats. (A) Oligonucleotides containing $(CT)_9 \cdot (AG)_9$, $(CTT)_6 \cdot (AAG)_6$, or $(C)_{18} \cdot (G)_{18}$ sequences were synthesized and cloned into pGL3-Enhancer vectors. (B) ChIP assays were performed with antibodies to total histone H3 (lanes 4 to 6) or acetyl-histone H4 (lanes 7 to 9). Input chromatin (0.15% of total) (lanes 1 to 3) and immunoprecipitated DNA (5% of total) (lanes 4 to 9) were amplified using the primers targeting the repeats (nucleotides 4784 to 4796 of pGL3-Basic vectors) or the primers targeting nucleotides 2636 to 2788 of pGL3-Basic vectors. Amounts of acetylated histone H4 were greater in the $(CT)_9 \cdot (AG)_9$ and $(CTT)_6 \cdot (AAG)_6$ sequences than in the $(C)_{18} \cdot (G)_{18}$ sequences. (C) The ratios of acetylated histone H4 to total histone H3 are shown. This experiment was repeated three times, and the means ± standard deviations are shown. *, $P < 0.01$ compared to results for the $(C)_{18} \cdot (G)_{18}$ sequences. (D) The luciferase activities were greater for $(CT)_9 \cdot (AG)_9$ and $(CTT)_6 \cdot (AAG)_6$ constructs than for the $(C)_{18} \cdot (G)_{18}$ construct. This experiment was repeated four times, and the means ± standard deviations are shown. *, $P < 0.01$ compared to the $(C)_{18} \cdot (G)_{18}$ sequences.

Our data establish the importance of H-DNA in another important disease, namely, hypertension. Using SHR, the most widely studied model of essential hypertension, we identified an expansion of $(CT)_n \cdot (AG)_n$ dinucleotide repeats in the promoter of the smMLCK gene. SHR were derived by recurrent inbreeding of normotensive WKY rats, selecting for high blood pressure in the absence of dietary or environmental stimuli (26). Interestingly, the insertion mutation occurred during a relatively short period of breeding, perhaps because of the high mutability of the repeated sequences. Moreover, the mutation was conserved in stroke-prone SHR, a substrain of SHR with high susceptibility to stroke (27), while it was not conserved in two normotensive strains, WKY and Sprague-Dawley rats (12). These data strongly suggest that the insertion is one of the genetic factors responsible for the development of high blood pressure in SHR. They also raise the interesting possibility that targeting these sequences with TFOs might be a novel strategy for treating hypertension.

Along with the genetic data, molecular and biochemical studies of the smMLCK promoter also support the notion that the insertion plays an important role in regulating smMLCK promoter activity. We have shown that the insertion increases promoter responsiveness to serum response factor (12), an important transcription factor that regulates smooth muscle differentiation (28). A *cis*-acting element for serum response factor, a CArG element, is located just 10 bp downstream of the $(CT)_n \cdot (AG)_n$ repeats, and various assays showed increased serum response factor binding to the CArG element in the SHR promoter. Consequently, the promoter responsiveness to serum response factor was greater in Cos-7 cells transfected with the SHR promoter than in cells expressing the normotensive promoter (12). Thus, data we have published suggest that the six additional $(CT) \cdot (AG)$ repeats in the SHR promoter affect serum response factor binding by changing the local environment of the promoter (12).

This idea is supported by the observation that the H-DNA structure is expanded at the SHR promoter (Fig. 3). H-DNA was formed from the 6th to the 15th $CT \cdot AG$ repeats at the normotensive promoter, showing that 10 out of 22 repeats were involved in the formation of H-DNA. It was expanded to the 21st $CT \cdot AG$ repeat in the SHR promoter, showing that 16 out of 28 repeats were involved in the structure. Thus, the expansion of the H-DNA structure could stabilize the interaction of the serum response factor with the CArG box and stimulate promoter activity of the SHR promoter.

Furthermore, the formation of H-DNA changes nucleosome organization or dynamics. Unlike the case with $(CT)_n \cdot (AG)_n$ sequences in *hsp70* and *hsp26* genes in *Drosophila melanogaster*, which appear to be nucleosome free (22), our data revealed nucleosomes at sites that form H-DNA in the rat smMLCK promoter (Fig. 4). Similarly, when a number of different DNA sequences from chicken erythrocyte core particles were analyzed, $(CT)_n \cdot (AG)_n$ tracts were occasionally found within nucleosomal DNA, showing no preferential rotational or translational orientation (36). This raises the question of how nucleosomes permit access of the polymerase but block access of the MNase to the $(CT)_n \cdot (AG)_n$ sequences (Fig. 4). While the packaging of promoter DNA into nucleosomes is thought to be an obstacle to both nuclease attack and transcription, nucleosomes are present during the transcription of eukaryotic genes (42). Furthermore, the chromatin structure is dynamic and is a vital component of transcription (1). The importance of histone acetylation in regulating chromatin dynamics and transcription has been demonstrated by genome-wide studies (17, 38), and 97% of active promoters were associated with increased levels of acetylated histone. Our data extend these studies and demonstrate histone acetylation at regions containing $(CT)_n \cdot (AG)_n$ repeat sequences. Our data also show that the longer H-DNA structure of the SHR promoter contains more acetylated histone (Fig. 5B). This is, we believe, the first demonstration that correlates an increasing repeat number with increased histone modifications.

We also confirmed the relationship between the length of repeat sequences and histone acetylation using artificial constructs with two different $(CT)_n \cdot (AG)_n$ repeat sequences. Both the enrichment of acetylated histone and luciferase activity were greater in the long repeats than in the short repeats (Fig. 6). Furthermore, luciferase activity and acetylated histone levels were greater in $(CT)_9 \cdot (AG)_9$ and $(CTT)_6 \cdot (AAG)_6$ constructs than in the $(C)_{18} \cdot (G)_{18}$ construct (Fig. 7). These data suggested that H-DNA-forming sequences have higher promoter activity than the non-H-DNA-forming sequences.

However, there are some unexpected characteristics of the SSRs. Both acetylation and activity were slightly lower in the $(CTT)_6 \cdot (AAG)_6$ sequences than in the $(CT)_9 \cdot (AG)_9$ sequences, suggesting possible differences in formation of H-DNA structures. While the mechanistic basis for this observation remains to be elucidated, our data demonstrate important roles for H-DNA-forming sequences in regulating histone acetylation and promoter activity.

In conclusion, hypertension is a quantitative trait that has a strong genetic component in humans and rats. Quantitative trait loci responsible for high blood pressure have been identified in hypertensive rats, and the MLCK gene is in one of the quantitative trait loci. Based on its chromosomal location and regulation of promoter activity and the demonstration that inhibiting MLCK expression or activity decreases blood pressure (12), we suggested that the regulation of smMLCK gene expression could play a central role in the development of hypertension. The data presented above add credence to this idea. They demonstrate that $(CT)_n \cdot (AG)_n$ repeat sequences found in the naturally occurring rat smMLCK promoter form H-DNA. Moreover, the six extra (CT) · (AG) repeats found in the hypertensive smMLCK promoter increase the H-DNA structure, histone modifications, and promoter activity. While nucleotide repeats have been implicated in many human neurological and neurodegenerative diseases (9), their relevance to vascular diseases is unknown. Our data extend the importance of nucleotide repeat sequences to a vascular disorder and provide new insights into the relationship between DNA structure and gene regulation.

ACKNOWLEDGMENTS

We thank Jennifer Cook and Guo-Fu Hu, Harvard Medical School, for advice on the details of the ChIP protocols and for generously providing repeat constructs, respectively, and Randal C. Jaffe, UIC, for generously providing the pGL3 vectors. We also thank our colleague Wilma Hofmann for critically reviewing the manuscript.

This work was supported in part by a grant from the National Science Foundation (0517468) to P.D.L. Y.-J.H. was supported by a postdoctoral fellowship from the American Heart Association (0525686Z).

We have no competing financial interests.

REFERENCES

1. **Becker, P. B.** 2002. Nucleosome sliding: facts and fiction. EMBO J. **21:**4749–4753.
2. **Campuzano, V., L. Montermini, M. D. Molto, L. Pianese, M. Cossee, F. Cavalcanti, E. Monros, F. Rodius, F. Duclos, A. Monticelli, F. Zara, J. Canizares, H. Koutnikova, S. I. Bidichandani, C. Gellera, A. Brice, P. Trouillas, G. De Michele, A. Filla, R. De Frutos, F. Palau, P. I. Patel, S. Di Donato, J. L. Mandel, S. Cocozza, M. Koenig, and M. Pandolfo.** 1996. Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science **271:**1423–1427.
3. **Carbone, G. M., S. Napoli, A. Valentini, F. Cavalli, D. K. Watson, and C. V. Catapano.** 2004. Triplex DNA-mediated downregulation of Ets2 expression results in growth inhibition and apoptosis in human prostate cancer cells. Nucleic Acids Res. **32:**4358–4367.
4. **Carroll, J. S., X. S. Liu, A. S. Brodsky, W. Li, C. A. Meyer, A. J. Szary, J. Eeckhoute, W. Shao, E. V. Hestermann, T. R. Geistlinger, E. A. Fox, P. A. Silver, and M. Brown.** 2005. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell **122:**33–43.
5. **de Lanerolle, P., and R. J. Paul.** 1991. Myosin phosphorylation/dephosphorylation and regulation of airway smooth muscle contractility. Am. J. Physiol. **261:**L1–L14.
6. **Epplen, C., J. T. Epplen, G. Frank, B. Miterski, E. J. Santos, and L. Schols.** 1997. Differential stability of the (GAA)n tract in the Friedreich ataxia (STM7) gene. Hum. Genet. **99:**834–836.
7. **Espinas, M. L., E. Jimenez-Garcia, A. Martinez-Balbas, and F. Azorin.** 1996. Formation of triple-stranded DNA at $d(GA \cdot TC)_n$ sequences prevents nucleosome assembly and is hindered by nucleosomes. J. Biol. Chem. **271:**31807–31812.
8. **Firulli, A. B., D. C. Maibenco, and A. J. Kinniburgh.** 1994. Triplex forming ability of a c-myc promoter element predicts promoter strength. Arch. Biochem. Biophys. **310:**236–242.
9. **Fortini, M. E.** 2007. Medicine. Anticipating trouble from gene transcription. Science **315:**1800–1801.
10. **Gacy, A. M., G. M. Goellner, C. Spiro, X. Chen, G. Gupta, E. M. Bradbury, R. B. Dyer, M. J. Mikesell, J. Z. Yao, A. J. Johnson, A. Richter, S. B. Melancon, and C. T. McMurray.** 1998. GAA instability in Friedreich's ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. Mol. Cell **1:**583–593.
11. **Grunstein, M.** 1997. Histone acetylation in chromatin structure and transcription. Nature **389:**349–352.
12. **Han, Y. J., W. Y. Hu, O. Chernaya, N. Antic, L. Gu, M. Gupta, M. Piano, and P. de Lanerolle.** 2006. Increased myosin light chain kinase expression in hypertension: regulation by serum response factor via an insertion mutation in the promoter. Mol. Biol. Cell **17:**4039–4050.
13. **Hancock, J. M.** 1995. The contribution of slippage-like processes to genome evolution. J. Mol. Evol. **41:**1038–1047.
14. **Htun, H., and J. E. Dahlberg.** 1989. Topology and formation of triple-stranded H-DNA. Science **243:**1571–1576.
15. **Kamm, K. E., and J. T. Stull.** 2001. Dedicated myosin light chain kinases with diverse cellular functions. J. Biol. Chem. **276:**4527–4530.
16. **Kashi, Y., D. King, and M. Soller.** 1997. Simple sequence repeats as a source of quantitative genetic variation. Trends Genet. **13:**74–78.
17. **Kim, T. H., L. O. Barrera, M. Zheng, C. Qu, M. A. Singer, T. A. Richmond, Y. Wu, R. D. Green, and B. Ren.** 2005. A high-resolution map of active promoters in the human genome. Nature **436:**876–880.
18. **Kinniburgh, A. J.** 1989. A cis-acting transcription element of the c-myc gene can assume an H-DNA conformation. Nucleic Acids Res. **17:**7771–7778.
19. **Leibovitch, B. A., Q. Lu, L. R. Benjamin, Y. Liu, D. S. Gilmour, and S. C. Elgin.** 2002. GAGA factor and the TFIID complex collaborate in generating an open chromatin structure at the *Drosophila melanogaster hsp26* promoter. Mol. Cell. Biol. **22:**6148–6157.
20. **Lu, Q., J. M. Teare, H. Granok, M. J. Swede, J. Xu, and S. C. Elgin.** 2003. The capacity to form H-DNA cannot substitute for GAGA factor binding to a (CT)n*(GA)n regulatory site. Nucleic Acids Res. **31:**2483–2494.
21. **Lu, Q., L. L. Wallrath, and S. C. Elgin.** 1994. Nucleosome positioning and gene regulation. J. Cell Biochem. **55:**83–92.
22. **Lu, Q., L. L. Wallrath, and S. C. Elgin.** 1995. The role of a positioned nucleosome at the Drosophila melanogaster hsp26 promoter. EMBO J. **14:**4738–4746.
23. **Lu, Q., L. L. Wallrath, H. Granok, and S. C. Elgin.** 1993. $(CT)_n \cdot (GA)_n$ repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the *Drosophila hsp26* gene. Mol. Cell. Biol. **13:**2802–2814.
24. **Luger, K.** 2003. Structure and dynamic behavior of nucleosomes. Curr. Opin. Genet. Dev. **13:**127–135.
25. **McGuffie, E. M., D. Pacheco, G. M. Carbone, and C. V. Catapano.** 2000. Antigene and antiproliferative effects of a c-myc-targeting phosphorothioate triple helix-forming oligonucleotide in human leukemia cells. Cancer Res. **60:**3790–3799.
26. **Okamoto, K., and K. Aoki.** 1963. Development of a strain of spontaneously hypertensive rats. Jpn. Circ. J. **27:**282–293.
27. **Okamoto, K., Y. Yamori, and A. Nagaoka.** 1971. Establishment of the stroke-prone spontaneously hypertensive rats (SHR). Circ. Res. **34:**(Suppl. I):I143–I153.
28. **Owens, G. K., M. S. Kumar, and B. R. Wamhoff.** 2004. Molecular regulation of vascular smooth muscle cell differentiation in development and disease. Physiol. Rev. **84:**767–801.
29. **Pestov, D. G., A. Dayn, E. Siyanova, D. L. George, and S. M. Mirkin.** 1991. H-DNA and Z-DNA in the mouse c-Ki-ras promoter. Nucleic Acids Res. **19:**6527–6532.
30. **Potaman, V. N., E. A. Oussatcheva, Y. L. Lyubchenko, L. S. Shlyakhtenko, S. I. Bidichandani, T. Ashizawa, and R. R. Sinden.** 2004. Length-dependent structure formation in Friedreich ataxia (GAA)n*(TTC)n repeats at neutral pH. Nucleic Acids Res. **32:**1224–1231.
31. **Potaman, V. N., and R. R. Sinden.** 1998. Stabilization of intramolecular triple/single-strand structure by cationic peptides. Biochemistry **37:**12952–12961.
32. **Reddy, P. S., and D. E. Housman.** 1997. The complex pathology of trinucleotide repeats. Curr. Opin. Cell Biol. **9:**364–372.
33. **Ritchie, S., and K. Bonham.** 1998. The human c-Src proto-oncogene promoter contains multiple targets for triplex-forming oligonucleotides. Antisense Nucleic Acid Drug Dev. **8:**391–400.
34. **Sakamoto, N., P. D. Chastain, P. Parniewski, K. Ohshima, M. Pandolfo, J. D. Griffith, and R. D. Wells.** 1999. Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia. Mol. Cell **3:**465–475.
35. **Sandaltzopoulos, R., C. Mitchelmore, E. Bonte, G. Wall, and P. B. Becker.**

1995. Dual regulation of the Drosophila hsp26 promoter in vitro. Nucleic Acids Res. **23:**2479–2487.

36. **Satchwell, S. C., H. R. Drew, and A. A. Travers.** 1986. Sequence periodicities in chicken nucleosome core DNA. J. Mol. Biol. **191:**659–675.
37. **Schroth, G. P., and P. S. Ho.** 1995. Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA. Nucleic Acids Res. **23:**1977–1983.
38. **Schubeler, D., D. M. MacAlpine, D. Scalzo, C. Wirbelauer, C. Kooperberg, F. van Leeuwen, D. E. Gottschling, L. P. O'Neill, B. M. Turner, J. Delrow, S. P. Bell, and M. Groudine.** 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev. **18:**1263–1271.
39. **Seidman, M. M., and P. M. Glazer.** 2003. The potential for gene repair via triple helix formation. J. Clin. Investig. **112:**487–494.
40. **Shang, Y., X. Hu, J. DiRenzo, M. A. Lazar, and M. Brown.** 2000. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. Cell **103:**843–852.
41. **Shen, C., A. Buck, B. Polat, A. Schmid-Kotsas, C. Matuschek, H. J. Gross, M. Bachem, and S. N. Reske.** 2003. Triplex-forming oligodeoxynucleotides targeting survivin inhibit proliferation and induce apoptosis of human lung carcinoma cells. Cancer Gene Ther. **10:**403–410.
42. **Studitsky, V. M., G. A. Kassavetis, E. P. Geiduschek, and G. Felsenfeld.** 1997. Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. Science **278:**1960–1963.
43. **Toth, G., Z. Gaspari, and J. Jurka.** 2000. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. **10:**967–981.
44. **Ussery, D. W., and R. R. Sinden.** 1993. Environmental influences on the in vivo level of intramolecular triplex DNA in Escherichia coli. Biochemistry **32:**6206–6213.
45. **Voloshin, O. N., S. M. Mirkin, V. I. Lyamichev, B. P. Belotserkovskii, and M. D. Frank-Kamenetskii.** 1988. Chemical probing of homopurine-homopyrimidine mirror repeats in supercoiled DNA. Nature **333:**475–476.
46. **Wang, G., and K. M. Vasquez.** 2004. Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. Proc. Natl. Acad. Sci. USA **101:**13448–13453.
47. **Wells, R. D., D. A. Collier, J. C. Hanvey, M. Shimizu, and F. Wohlrab.** 1988. The chemistry and biology of unusual DNA structures adopted by oligopurine · oligopyrimidine sequences. FASEB J. **2:**2939–2949.
48. **Wierdl, M., M. Dominska, and T. D. Petes.** 1997. Microsatellite instability in yeast: dependence on the length of the microsatellite. Genetics **146:**769–779.
49. **Wilson, A. K., R. S. Pollenz, R. L. Chisholm, and P. de Lanerolle.** 1992. The role of myosin I and II in cell motility. Cancer Metastasis Rev. **11:**79–91.
50. **Wooster, R., A. M. Cleton-Jansen, N. Collins, J. Mangion, R. S. Cornelis, C. S. Cooper, B. A. Gusterson, B. A. Ponder, A. von Deimling, O. D. Wiestler, et al.** 1994. Instability of short tandem repeats (microsatellites) in human cancers. Nat. Genet. **6:**152–156.
51. **Xu, Z. P., T. Tsuji, J. F. Riordan, and G. F. Hu.** 2003. Identification and characterization of an angiogenin-binding DNA sequence that stimulates luciferase reporter gene expression. Biochemistry **42:**121–128.