Effect of Nuclear Protein HMG1 on *in Vitro* Slippage Synthesis of the Tandem Repeat dTG·dCA[†]

Corinne L. Gibb, Wenjie Cheng, Victor N. Morozov, and Neville R. Kallenbach*

Department of Chemistry, New York University, New York, New York 10003

Received August 14, 1996; Revised Manuscript Received February 10, 1997[⊗]

ABSTRACT: Tandem repeats of simple doublet and triplet sequences occur with high frequency in the DNA of eucaryotes. Among the most frequent is the repeat of dTG, which has unusual structural properties. We show here that HMG1 (modeled by the second HMG box motif from HMG1 of the rat, HMGb) binds to complexes formed from annealing unequal lengths of dTG·dCA and inhibits the *in vitro* elongation of these complexes by the Klenow fragment of DNA polymerase I at 37 °C. At 46 °C, HMGb enhances the elongation. Polylysine inhibits elongation at both temperatures. These results show that the stability of this repeat *in vivo* can be influenced by the presence of basic proteins in general, and more selectively by the abundant nuclear protein HMG1.

The presence in eucaryotic DNA of a variety of dispersed highly repeated simple sequences, examples of which include the doublet repeats dAT and dTG, raises interesting questions concerning the origin, function and evolution of these regions (Singer, 1982). One characteristic of simple sequence repeats is that they exhibit a high level of length polymorphism (Kashi et al., 1990) even among individuals of the same species (Tautz et al., 1986; Levinson & Gutman, 1987b). While several mechanisms can account for this high degree of variability, one possibility is that repeat sequences can "slip" by mispairing during replication or repair, permitting the two strands to shift in register and to elongate readily in the presence of polymerases (Kornberg, 1980; Levinson & Gutman, 1987a). In their studies of DNA synthesis by *E. coli* DNA pol I, Kornberg's group discovered that long chains of poly(dAT) could polymerize *in vitro* from dATP and dTTP in the presence of enzyme without added template (Kornberg et al., 1964). Short oligomers of dAT act as template/primers in analogous extension reactions (Kornberg, 1980) and short primers containing unequal length tri- and tetranucleotide repeats elongate as well (Wells et al., 1965, 1967) to yield high molecular weight DNA duplexes with the corresponding sequence repeats. It has been speculated that both the origin and the variability in length of simple tandem repeats reflect a similar strand slippage mechanism *in vivo* (Levinson & Gutman, 1987b). Schlotterer and Tautz (1992) demonstrated that many primers containing simple tandem sequence repeats including dTG elongate readily in the presence of DNA polymerase and appropriate deoxynucleotide triphosphates. They investigated polymerases that were processive, such as T7 polymerase, as well as nonprocessive enzymes such as *E. coli* pol I. In the case of dTG, they reported extension from 18-mer primers to chains in excess of 400 bases after 120 min of synthesis by *E. coli* pol I at 37 °C. By contrast, the T7 polymerase synthetic rate was much lower, and the distribution of products more heterogeneous.

If these sequences can be extended by means of slippage during repair or replication, what limits the indefinite extension of simple repeat sequences? One answer is that the "slipped strand mispairing" mechanism postulated to allow repeats to extend can also lead to deletions following action by repair enzymes (Levinson & Gutman, 1987b). These workers detected frameshifts with high frequency when they inserted a dTG tract of 40 bp into M13 phage and found that the frequency of frame shifts is strongly reduced by action of methyl directed mismatch repair enzymes (Levinson & Gutman, 1987a). In addition, dTG repeats stimulate genetic recombination (Stringer, 1985; Treco et al., 1985; Bullock et al., 1986). High levels of recombination can lead to extension or deletion of repeats or even excision of unique sequences intervening between direct repeats. Finally, proteins in the nucleus may play a direct role in regulating the length of repeats (Gaillard & Strauss, 1994). The latter reported that 120-mer sequences containing long dTG tracts spontaneously form extended ladders containing branched structures *in vitro*, which bind HMG1 and HMG2 proteins selectively.

The most abundant DNA binding proteins of the nucleus aside from the histones are the HMG1/2 group of proteins (Johns, 1982), also referred to as HMG box proteins (Bianchi et al., 1989). We show here that HMGb, the second of two box motifs present in rat HMG1, binds to structures formed from annealing strands of dCA_{12} with dTG_{21} and that the protein inhibits the synthetic reaction at 37 °C of the Klenow fragment of *E. coli* pol I (Joyce & Grindley, 1983) which catalyzes elongation of this primer *in vitro*. Interestingly, extension by the same enzyme at 46 °C is enhanced by HMGb. A sample of poly-DL-lysine (poly-Lys) was used as a control to assess the effect of nonspecific basic proteins on *in vitro* extension. At both 37 and 46 °C, poly-DL-lysine inhibits extension, at concentrations comparable to that of HMGb. On the basis of these observations, we suggest that basic DNA binding proteins in the nucleus can regulate the

[†] This research was supported by grants from the US NIH (CA 24101).

* Corresponding author: Neville R. Kallenbach, Department of Chemistry, New York University, New York, NY 10003. Telephone: (212) 998-8757. Fax: (212) 260-7905. E-mail: kallnbch@is.nyu.edu.

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1997.

S0006-2960(96)02037-5 CCC: $14.00

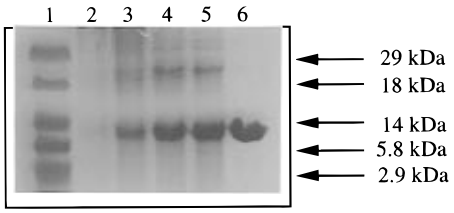

FIGURE 1: Purification of HMGb protein. Samples at various stages of purification were loaded on a 15% SDS−polyacrylamide gel, electrophoresed (200 V, 1 h), and stained with Coomassie Blue (Sigma). Lanes: (1) low molecular weight protein marker set for size calibration (Gibco, BRL); (2) crude lysate from induced BL21(DE3) dissolved directly in loading buffer; (3) supernatant after NaCl precipitation step; (4) supernatant after ammonium sulfate precipitation; (5) protein mixture before application of column chromatography; (6) HMGb after final fractionation on CM-FPLC column (BioRad).

length of simple tandem repeats in eucaryotic DNA, via effects on repair enzymes if not on replicative polymerase activity.

MATERIALS AND METHODS

Proteins. The plasmid pHB1 coding for HMG B domain (amino acids 86−165) of rHMG1 was made available to us by Prof. S. J. Lippard (Department of Chemistry, MIT). The protein was expressed and purified as described by Chow et al. (1995). Purity of the product was verified by observation of a single band on SDS−PAGE (Figure 1). Poly-DL-lysine was purchased from Sigma (lot 41H55231) and used without further purification.

DNA Preparation. Oligonucleotides were synthesized by phosphoramidite chemistry using an Applied Biosystems unit and were purified by preparative ion exchange HPLC. Longer strands requiring further purification were subjected to preparative electrophoresis on denaturing gels. The sequences and designation of the strands used are as follows:

dCA_{12}	(5′)d$[CA]_{12}$-OH(3′)
dCA_{21}	(5′)d$[CA]_{21}$-OH(3′)
dCA_{12cap}	(5′)dCCTGCAG$[CA]_{12}$GGTACCG-OH(3′)
dCA_{21cap}	(5′)dCCTGCAG$[CA]_{21}$GGTACCG-OH(3′)
dTG_{12}	(5′)d$[TG]_{12}$-OH(3′)
dTG_{21}	(5′)d$[TG]_{21}$-OH(3′)
dTG_{12cap}	(5′)dCGGTACC$[TG]_{12}$CTGCAGG-OH(3′)
dTG_{21cap}	(5′)dCGGTACC$[TG]_{21}$CTGCAGG-OH(3′)

Samples of 5 μg of synthetic oligonucleotide were 5′ end-labeled by incubation with 2 μL of [γ-^{32}P]ATP (ICN, 10 μCi/μL), 1 μL of T4 polynucleotide kinase (10 units/μL), and 2 μL of 5x T4 kinase forward reaction buffer (300 mM Tris-HCl, pH 7.8, 50 mM $MgCl_2$, 1.65 μM ATP, 75 mM 2-mercaptoethanol). Labeled strands were purified by electrophoresis on 15% acrylamide denaturing gels.

Primer/templates were prepared by addition of labeled strands to the appropriate concentration of complementary strand in 50 mM Tris-HCl (pH 7.5) and 10 mM $MgCl_2$. The product was annealed by heating the mixture at 90 °C for 2 min and then cooling slowly to room temperature. Where appropriate, the size of the complexes formed was calibrated using a ladder of 10 bp fragments (BRL).

Exo VII Digestion. DNA strands were digested by 0.1 unit of exonuclease VII (Gibco BRL) in 70 mM Tris-HCl (pH 8.0), 8 mM EDTA, 10 mM 2-mercaptoethanol, and 50 μg/mL BSA at room temperature for 10−30 min. The reaction was quenched by adding 5 mM EDTA.

Protein Binding. Electrophoretic mobility shift (EMS) assays were performed after adding proteins to DNA structures in a binding buffer containing 4% glycerol, 10 mM $MgCl_2$, 5 mM KCl, 10 mM HEPES (pH 7.9), 1 mM EDTA, 1 mM spermidine, 0.5 mM DTT, 0.05% Nonidet P-40, and 200 μg/mL BSA in a final volume of 10 μL. After a 20 min incubation on ice, the samples were applied to a precooled (4 °C), 0.5 × TBE, 8% native polyacrylamide gel (30:1 acrylamide:bisacrylamide) that had been extensively prerun. Electrophoresis was performed at 4 °C for 3 h at 10 V/cm with continuous circulation of the buffer. The gel was transferred to a Whatman 3 MM sheet, dried, and autoradiographed with X-ray film (Kodak X-OMAT AR) at −80 °C using an intensifying screen.

Primer Extension Experiments. After annealing labeled strands of interest overnight in annealing buffer as described above, the mixture was chilled on ice for 10 min before reaction buffer was added (10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 7.5 mM DTT, New England Biolabs), polymerization mix (2 mM of each dNTP, Pharmacia) and 1 unit of *E. coli* polymerase I (Klenow) fragment from Biolabs. The total volume was 20 μL. DNA synthesis was carried out for varying times and quenched by adding 2 μL of 0.5 M EDTA and then heating the sample at 75 °C for 10 min. After cooling, 1 unit of proteinase K (Sigma) (in the case of HMGb) or trypsin (for poly-Lys) was reacted for 1 h to hydrolyze the protein, and the reaction was stopped by heating at 90 °C for 5 min. Samples were then subjected to electrophoresis under denaturing conditions (8% acrylamide) to reveal elongation of the primer. Control ladders containing single-stranded 10-mers were used to calibrate the size of the bands on denaturing gels; the results are indicated at the side of the gels in Figures 5 and 6. For investigation of the effect of HMGb or poly-Lys, the protein(s) were added to the reaction mixture prior to adding polymerization buffer, the DNA−protein complexes were stabilized on ice for 20 min, and the polymerase was then added.

Scanning Force Microscopy. Observations were made with a Nanoscope II instrument (Digital Instruments, Santa Barbara, CA) of solutions containing samples of high molecular weight poly(dTG·dCA) (Sigma, mean molecular mass 3.7 MDa) with and without HMGb protein. A stock solution of 350 mg/mL of the polymer was diluted 40-fold with 15 mM Tris buffer and 5 mM $MgCl_2$, heated to 100 °C for 1 min, and then chilled in ice. Droplets containing successive 1:10 dilutions of this material were spread on Ni-treated mica, allowed to stand for 20−30 s, and dried with a stream of N_2. Scanning was in contact mode under dry N_2, using commercial tips that had been plasma treated and coated with dichlorodimethylsilane vapor to render them hydrophobic. Under these conditions the height of the trace corresponds well to the true value for the specimen, while the dimensions along the *x* and *y* axes are amplified as the tip scans in contact mode (Hansma et al., 1993). Protein or protein−DNA complexes were observed using a 4 mM solution of HMGb (2 μL), mixed with 9 μL of annealed dTG·dCA as described above, treated with 0.01% glutaraldehyde, diluted in an equal volume of Tris buffer (10 mM Tris, pH 8, 50 mM NaCl), and applied directly to a Ni−mica surface. The HMGb solution spread with or without glutaraldehyde treatment produced a field of spherical

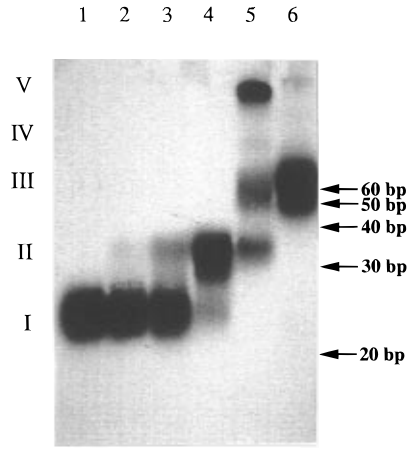

FIGURE 2: PAGE analysis of products of annealing labeled dTG_{21cap} strands with dCA_{12cap} strands on an 8% native gel. The single strand control is shown in lane 1. Lanes 2−6 show different complexes that result as the concentration of CA_{12cap} increases: 0.1, 0.3, 0.5, 1, and 2.5 μM, respectively. The concentration of dTG_{21cap} in each lane was held constant at 0.5 μM. The experiment shown was run in TA-Mg^{2+} buffer; in TBE buffer the mobility of the single-strand shifts (due to hairpin formation) and structures I and II have the same mobility.

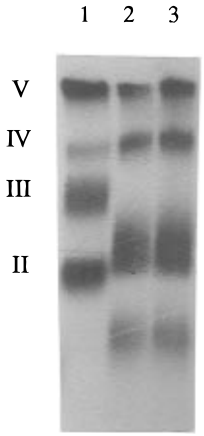

FIGURE 3: Exonuclease VII digestion of dTG_{21cap}·dCA_{12cap} complexes. The products were run on an 8% native acrylamide gel in TBE buffer. Lane 1 shows the four structures **II**−**V** corresponding to lane 5 in Figure 1 without enzyme. Lanes 2 and 3 show the results after 10 and 30 min of exonuclease VII treatment.

particles with height of 1.5 ± 0.3 nm. In the presence of DNA, decorated DNA chains appear, with height of the "beads" averaging 2.2 ± 0.4 nm.

RESULTS

Binding of HMGb to Intermediates Formed by Annealing dTG and dCA Strands of Unequal Length. Observation of the range of structures formed from annealing dTG_{21} with dCA_{12} is facilitated by the greater stability of strands containing unique sequence pairing caps. Thus annealing labeled dTG_{21cap} (band I in Figure 2) with increasing concentrations of dCA_{12cap} produces a total of four additional bands on native PAGE (Figure 2). The position of the dTG_{21cap} band does not correspond to single-stranded 56-mer when run in TBE buffer as opposed to TA-Mg^{2+}, indicating formation of a hairpin structure stabilized by wobble G·T base pairs (Early et al., 1978). The remaining bands and species formed are assigned as follows: II appears at lowest concentration of complementary strand, and is trimmed by Exo VII (see Figure 3), with loss of label, hence represents 1:1 duplex dTG_{21cap}:dCA_{12cap}. Band III appears at higher strand concentrations than II and hence represents a trimer, dTG_{21cap}: 2 dCA_{12cap}, with one set of caps extending from the main duplex. This band shows a reduction in size after Exo VII treatment consistent with removal of the extended strands. Bands IV and V have much lower mobility, as expected for branched intermediates. Band IV is not susceptible to *Exo* VII, indicating it is the three-arm branch structure formed from two strands with caps paired; this band disappears at higher ratios of the shorter strand, in agreement with this assignment. We assign band V to a second branched complex, from its low mobility and insensitivity to Exo VII; since band V appears at low temperature, it probably corresponds to the four-arm branched analog of the three-arm branch of band IV. Low temperature favors four-arm branch formation (Lu et al., 1992). Both bands IV and V disappear at higher ratios of dCA_{12cap}, while band III accumulates as expected. When the label pattern is reversed, and dCA_{21cap} is kinased rather than dTG_{21cap}, the same series of bands seen in Figure 1 results on annealing cold dTG_{12cap}. Similarly, labeling the shorter strand in either case also yields the same series of bands seen in Figure 2.

Gel shift analysis of the interaction of HMGb and poly-Lys with the intermediates formed by dTG_{21cap} and dCA_{12cap} is shown in Figure 4. In this experiment the concentration of strands corresponds to that shown in lane 5 of Figure 2 so that bands II, III, and V are present. Band V shows a clear shift with increasing HMGb concentration, while bands II and III shift only at much higher protein concentration (Figure 4A). Since both contain single-strand extensions, based on the exonuclease VII digestion experiment, this is consistent with the ability of HMGb to interact with single DNA strands (Isackson et al., 1979), albeit more weakly than its interaction with branches. Tight binding of band V confirms our assignment of this band to a branched structure (Bianchi et al., 1989). Similar gel shifts are observed in mixtures containing band IV (data not shown). The corresponding experiment with poly-Lys is shown in Figure 4B, indicating that bands III and V both interact. Poly-Lys binds at micromolar concentration as judged by gel shift assay, but nonspecifically, so that the complexes remain in the wells (Figure 4B).

In Vitro Elongation of dTG·dCA Complexes by E. coli Polymerase I. In contrast to the multiplicity of bands detected in the presence of the caps, only one band is observed when dTG_{21} is annealed with increasing concentrations of dCA_{12}. Nevertheless a variety of structures similar to those detected with the capped sequences are likely to arise during annealing of the uncapped strands TG_{21} and CA_{12}, but their lower stability precludes observation by PAGE analysis. Figure 5 shows the kinetics of extension when labeled dTG_{21} lacking the capping sequence is annealed with dCA_{12} and incubated with Pol I as described in Methods. The mobility of the template strand on a denaturing gel is seen to decrease progressively with time, monitoring slippage synthesis on this primer as reported originally by Schlotterer and Tautz (1992). Negative controls in the same time scale using identical enzyme and triphosphate concentrations were obtained for the polymerization reaction when either dTG_{21} was annealed with dCA_{21} or dTG_{12} was annealed with dCA_{12}. However after 60 min a small extent of elongation could be detected as a series of faint bands with lower mobility than the primer/template strand. This is consistent with the end

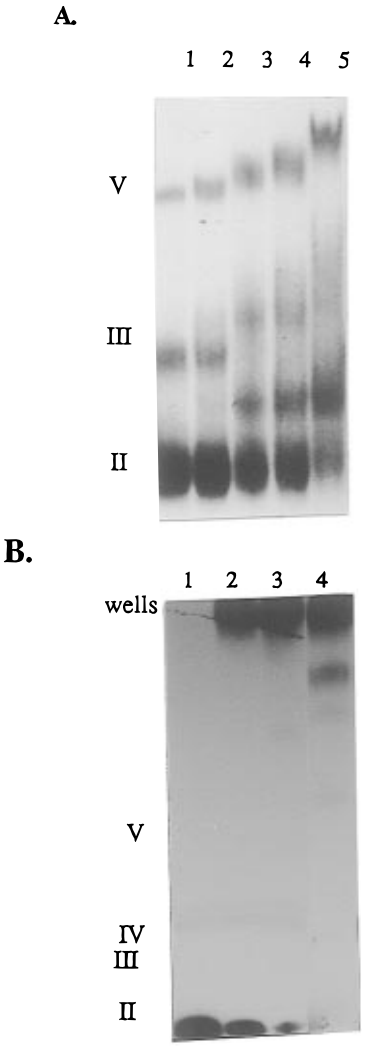

FIGURE 4: Gel shift analysis of HMG-b (A) and poly-DL-lyine (B) binding by dTG_{21cap}·dCA_{12cap} structures on an 8% native acrylamide gel in TBE buffer. Lane 1 in both panels show the structures **II**−**V** (0.1 μM) in the absence of protein. (A) From lane 2 to lane 5 the concentration of HMGb increases: 0.1, 0.4, 0.7, and 2.8 μM, respectively. (B) Poly-Lys concentrations are 1.5, 2, and 2.5 μM, respectively, for lanes 2−4. Note the accumulation of protein−DNA complex in the wells as the poly-Lys concentration increases.

growth mechanism postulated by Schlotterer and Tautz (1992), whereby elongation should proceed from the ends of the duplex at temperatures that promote fraying. The elongation rate is slower at 14 and 46 °C, consistent with the temperature optimum near 37 °C in these conditions for stabilizing intermediates in the dAT extension reaction (Kornberg, 1980).

Effect of HMGb on Strand Elongation. Figure 5 also shows the inhibitory effects of HMGb and poly-Lys on the polymerase-mediated extension of dTG_{21}·dCA_{12} template/primer at 37 °C. Both proteins block the extension reaction, as well as the exonuclease activity of the enzyme, seen by loss of the ladders extending to shorter chain lengths in Figure 5. The effectiveness of HMGb and poly-Lys as inhibitors follows their apparent binding affinity (Figure 4). By contrast however, at 46 °C HMGb enhances the elongation reaction, while poly-Lys still inhibits (Figure 6). Thus the action of HMGb is not simply due to its content of basic amino acids: poly-Lys inhibits under all conditions tested.

Imaging HMGb Complexes with Poly(dTG)·Poly(dCa). The short length of the dTG sequences used in the above

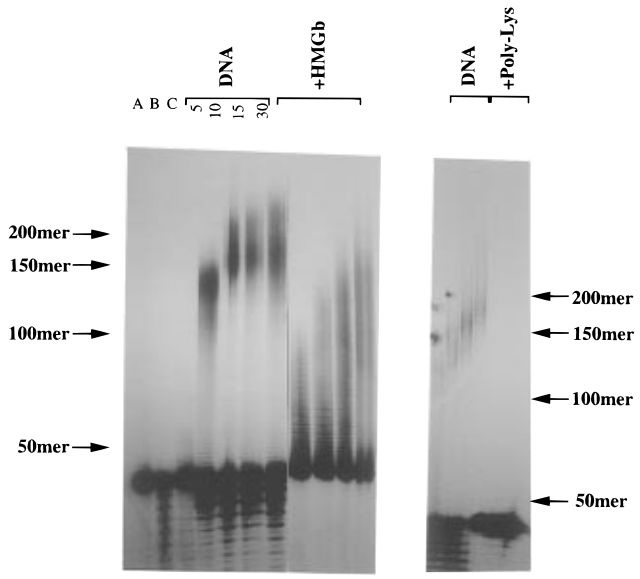

FIGURE 5: The inhibitory effects of the proteins on the extension of labeled dTG_{21}·dCA_{12} template (corresponding to lane 5 in Figure 1, 0.1 μM) at 37 °C. The reaction is monitored on an 8% denaturing gel. For each group, the extension reactions are carried out for 5, 10, 15, and 30 min (left to right). Lane A contains DNA alone; lane B and C shows the exonuclease activity of the Pol I in the absence of triphosphates. The concentration of HMGb was fixed at 6 μM, and that of poly-Lys at 2.5 μM.

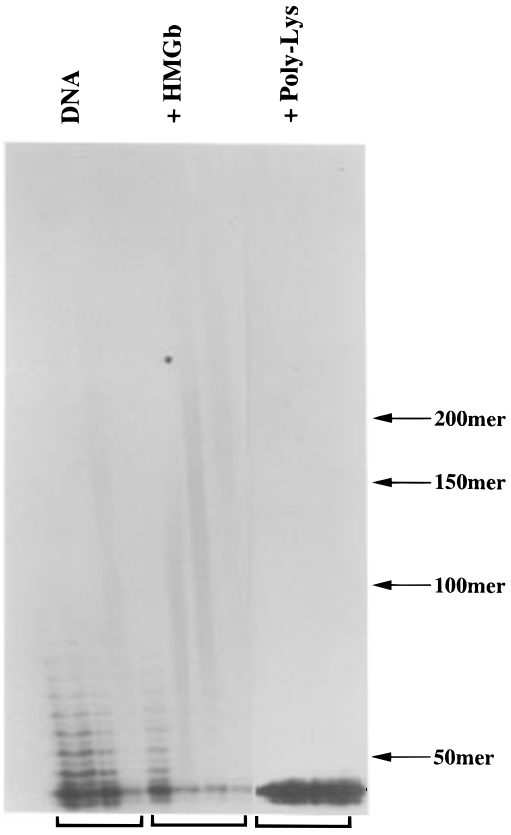

FIGURE 6: Extension of labeled dTG_{21}·dCA_{12} template (corresponding to lane 5 in Figure 1, 0.1 μM) by polymerase I (Klenow fragment) at 46 °C, in the presence of HMGb and poly-Lys. All other reaction conditions are the same as in Figure 6, except for the temperature.

experiments precludes their direct observation by means of scanning force microscopy, which requires longer chain DNAs for imaging (Bezanilla et al., 1994). However, the availability of synthetic poly(dTG)·poly(dCA) allowed us

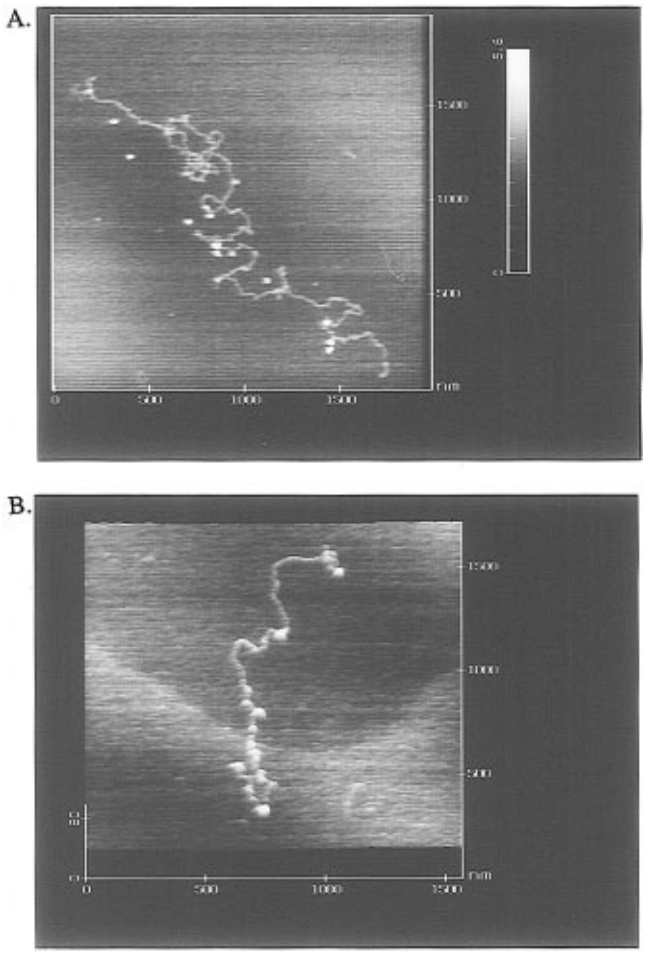

FIGURE 7: Scanning force microscopy of complexes between poly(dTG)·poly(dCA) and HMGb. Two images are shown in panels A and B of the result of an experiment in which the polymer was heated, annealed, and added to a solution of HMGb as described in methods, fixed with glutaraldehyde, spread on Ni-treated mica, and dried. Globules of height 2.2 ± 0.4 nm are seen decorating the DNA strands. Protein alone gives globules of 1.5 ± 0.3 nm, as seen in panel A. The decoration tends to occur at the ends of the chain and at short kinked or branched structures.

to image the binding of HMGb to this material (Figure 7). In the absence of annealing the strands and quenching the mixture, we could detect no interaction between HMGb and the copolymer. After heating and rapidly cooling diluted solutions of the DNA, decoration of the DNA strands was seen, which now contained numerous kinks and bends reflecting the presence of internal loops and branches. Binding of the HMGb to ends could be observed, as seen in panel 7B, but there is indication of binding at sites in the middle of the chains in both panels. Thus the target for binding of HMGb can be associated with regions of branching in the copolymer, as well as with the ends, which would be expected to fray.

DISCUSSION

There is now considerable interest in the properties of simple sequence repeats in DNA, stimulated by the association between triplet repeats such as dCCG and dCTG and inherited neurological disorders (Caskey et al., 1992). In the latter cases, increase in the length of repeats is associated with earlier onset and increased severity of the disorder. The doublet sequence of this study, dTG is the most prevalent repeat in eucaryotes, being represented roughly 10^5 times in mammalian genomes (ca. 0.04% of the total genome), in repeats of up to 30 copies which show a high degree of variability in copy number (Hamada et al., 1982; Gross et al., 1985; Tautz, 1989; Kashi et al., 1990). These repeats can act as recombinational hot spots (Stringer, 1985; Treco & Arnheim, 1986; Wahls et al., 1990) as well as sites for frame-shift mutagenesis (Levinson & Gutman, 1987a).

Two closely related non-histone DNA binding proteins of the nucleus, HMG1 and HMG2, interact with the structure(s) formed by annealing DNA fragments containing 60 bp lengths of dTG: Gaillard and Strauss (1994) reported electron microscopic evidence that HMG1 and HMG2 bind multimeric complexes of DNA containing dTG repeats, which form branched structures in the absence of protein. They attributed the branching to formation of a Z DNA type of intermediate, since the rate of branching was enhanced by addition of ethanol or high salt, while the kinetics they monitored were slow—their *in vitro* process took place over several hours to days. The dTG sequence has been shown to be highly pleotropic, with the potential to form several alternative DNA structures: under the influence of high salt concentration, ethanol, or negative superhelical twist, DNA containing dTG inserts can switch its conformation from the B form to a left-handed form (Vorlickova et al., 1982; Haniford & Pulleyblank, 1983; Nordheim & Rich, 1983; Kladde et al., 1994). Triple-stranded structures have also been identified at acidic pH (Antao et al., 1990). Under physiological conditions, dTG can form hairpins, based on nonstandard Watson−Crick base pairing detected by ^{1}H NMR (Early et al., 1978).

The HMG binding experiments reported here are consistent with the idea that the repeat forms cruciform structures in the absence of conditions known to favor Z or Z-like DNA structures. HMG1 protein interacts only weakly with duplex DNA but strongly with branched species (Bianchi et al., 1989): depending on the ionic conditions, the preference for branches can be about 1000-fold tighter than for duplexes (Cheng et al., 1997). We find that HMGb binds to structures formed by annealing unequal lengths of dTG with dCA, with apparent K_d values in the micromolar range. Control experiments in which the chain length of the dCA is equal to that of dTG show no interaction with HMGb; scanning microscopy also shows no interaction between HMGb and poly(dTG)·poly(dCA) unless the DNA strands are heated and rapidly cooled. Since dTG_{21} annealed with dCA_{21} elongates *in vitro*, albeit much more slowly than when dTG_{21} is annealed with dCA_{12}, it is likely that the intermediates favoring the rapid extension reaction are branches with relatively short arms. Neither HMG1 nor HMGb interacts with branched sequences containing fewer than three base pairs (Cheng et al., 1997), so this should represent a lower limit for such branches, the length of which would be expected to be very sensitive to temperature. We argue then that the intermediates responsive to these proteins are cruciforms (Levinson & Gutman, 1987b) made possible by the ability of dCA to slip with respect to the dTG strand, which itself can form stable hairpins (Early et al., 1978).

HMG1 and its single-box HMGb both interact with single-stranded DNA, but with lower affinity (Cheng et al., 1997). This might suggest that inhibition occurs by direct attachment of HMGb at the ends. Since we detect inhibition at concentrations of HMGb well below levels that interact with single-stranded oligomers, we attribute the inhibition to

cruciform binding rather than attachment to the ends exclusively. The latter would remain a plausible mechanism at higher protein concentrations. End binding cannot explain the enhancement observed at 46 °C (Figure 6). This requires a mechanism whereby stabilizing branches increases the template activity at the end(s). In principle, binding to branched intermediates could inhibit or enhance the extension rate, depending on the lifetime of the complex producing ends and the rate of synthesis at these ends: the apparent extension rate at 46 °C is lower than that at 37 °C. The consistent inhibition by poly-Lys at both 37 and 46 °C might then reflect a different mode of binding by this species. While HMGb selects branched structures over duplexes, several preparations of poly-Lys tested bind tightly also to duplex DNA as seen by gel shifts (see Figure 4B for poly-Lys).

The images in Figure 7 show decoration by HMGb of a polymer of (dTG)·(dCA) pairs, which produces a kinked structure after heating and spreading on Ni-treated mica. The protein alone appears as globules about 1.5 nm in height, while the decorations correspond quite closely to the sum of the protein height and that of the DNA alone, about 0.6 nm. Interactions of the protein with regions of the DNA other than ends can be seen in both panels; the ends appear to bind, however (Figure 7, panel B).

The polymerization of short dAT primers by DNA pol I has been attributed to a reiterative mechanism, in which the primer is bound to the enzyme and grows at an end *via* slippage of the strands (Kornberg, 1980). This is consistent with the existence of a temperature optimum in the extension rate for short oligonucleotide primers: under conditions where the rate of extension for primers of dAT_3 shows a temperature optimum near 0 °C, that for dAT_4 is about 10 °C and for dAT_5 about 20 °C. Above dAT_7, the optimum temperature is 40 °C and insensitive to chain length (Kornberg et al., 1964). The existence of a unique optimum for very short chain lengths provides a strong argument for a slippage mechanism (Kornberg, 1980). The convergence of optimum temperatures for chains longer than 14-mers is consistent with the length of template DNA thought to interact with the enzyme (Beese et al., 1993). In longer chain length template/primers, dissociation of the strands would be replaced by opening at the ends. In the absence of exonuclease activity, the filled-in chain generated from the initial primer species dissociates before slipping again, creating a free 3′ end and allowing a second round of filling in, and so on (Schlotterer & Tautz, 1992). In our experiments, we observe slower extension of primers lacking free ends. The primer extension rates measured in the experiments of Schlotterer and Tautz (1992) with unequal length strands were independent of primer length, consistent with an end-growth mechanism. The extension rate is strongly sequence dependent, however, and higher GC content slows the reaction considerably. Selective binding of putative branched intermediates by HMGb can slow the reaction by restricting slippage, while binding at the free end would of course limit filling in directly. Depending on the stability of the sequence, stabilizing the branch could also promote extension, as seen in the reactions at 46 °C (Figure 6). The fact that poly-Lys inhibits extension under all conditions tested indicates it could interact directly at the ends; under conditions where poly-Lys is added to a reaction in which HMGb enhances the rate, inhibition is observed (data not shown).

In vivo, dTG tracts are found interspersed among nonrepetitive sequences (Hamada et al., 1982), restricting the ability of the ends to slip as required by the simple iterative mechanism discussed above. The observation of branched complexes competent to bind HMGb in duplex DNA consisting of long sequences of dCA and dTG (Figure 7) is consistent with the idea that an asymmetric synthesis originates with hairpin formation, probably on the TG strand. It is interesting in this connection that a single-stranded dTG repeat inserted in one strand only yields heterogeneous length products in the presence of a short dCA template and DNA polymerase (Schlotterer & Tautz, 1992). The reaction we describe here may apply only to repair polymerases, and we have no evidence so far that processive replicases respond to the proteins tested in these experiments. Since changes in bulk solvation or twist are not involved in the shorter template/primer complexes of this study (Haniford & Pulleyblank, 1983; Nordheim & Rich, 1983), we believe our present results do not implicate a transition to left-handed DNA structures (Naylor & Clark, 1990); the alternative state of dTG is probably not classical Z DNA (Kladde et al., 1994). Triple helices have been detected only at acidic pH and seem unlikely candidates in experiments conducted above pH 7 (Antao et al., 1990). The simplest interpretation of the data we have obtained is that interaction with HMGb transiently stabilizes branched intermediates that liberate unpaired bases at the ends of the chain for polymerase extension. This is consistent with the temperature optimum for the extension reaction, as well as the ability of HMGb to enhance the rate of extension at 46 °C. Interaction with poly-Lys on the other hand nonspecifically inhibits polymerase reaction by binding to duplex structure.

Biologically, we believe the results presented point to a role for HMG1 in influencing the length of simple tandem repeats. It has previously been reported that HMG1 accumulates in nuclei during DNA synthesis (Bonne-Andrea et al., 1986) and enhanced replication in ascites nuclei (Alexandrova et al., 1984). The level of HMG1 is estimated to reach 10^6 molecules per nucleus or more (Bonne-Andrea et al., 1986). This protein is thus abundant in nuclei that are synthesizing DNA, and exogeneous HMG1 stimulates DNA synthesis. Given that polymerases can extend the length during a cycle of replication and that repair and recombination processes can further contribute to variability (Levinson & Gutman, 1987b), the ability of abundant HMG1/2 proteins to suppress or enhance elongation could represent an important balancing factor in length polymorphism. Despite the high variability in length of tandem repeats from individual to individual (Tautz, 1989; Kashi et al., 1990), it is striking that length polymorphisms tend to be stable somatically (Schlotterer & Tautz, 1992). Thus there must exist effective mechanisms for suppressing the copy number of repeat sequences as well as for enhancing its variability.

ACKNOWLEDGMENT

We thank Dr. S. J. Lippard for providing the expression vector pHB1 for HMGb. W.C. acknowledges a Dissertation Fellowship from the Graduate School of Arts and Science, NYU.

REFERENCES

Alexandrova, E. A., Marekov, L. N., & Beltchev, B. G. (1984) *FEBS Lett. 178*, 153−156.

Antao, V. P., Ratliff, R. L., & Gray, D. M. (1990) *Nucleic Acids Res. 18*, 4111−4122.

Beese, L. S., Derbyshire, V., & Steitz, T. A. (1993) *Science 260*, 352−355.

Bezanilla, M., Drake, B., Nudler, E., Kashlev, M., Hansma, P. K., & Hansma, H. G. (1994) *Biophys. J. 67*, 2454−2459.

Bianchi, M. E., Beltrame, M., & Paonessa, G. (1989) *Science 243*, 1056−1059.

Bonne-Andrea, C., Harper, F., Puvion, E., Delpech, M., & De-Recondo, A.-M. (1986) *Biol. Cell 58*, 185−194.

Bullock, P., Miller, J., & Botchan, M. (1986) *Mol. Cell. Biol. 6*, 3948−3953.

Caskey, C. T., Pizzuti, A., Fu, Y.-H., Fenwick, R. G., Jr., & Nelson, D. L. (1992) *Science 256*, 784−789.

Cheng, W., Zhong, M., Liaw, Y. C., & Kallenbach, N. R. (1997) (manuscript in preparation).

Chow, C. S., Barnes, C. M., & Lippard, S. J. (1995) *Biochemistry 34*, 2956−2964.

Early, T. A., Olmsted, J. d., Kearns, D. R., & Lezius, A. G. (1978) *Nucleic Acids Res. 5*, 1955−1970.

Gaillard, C., & Strauss, F. (1994) *Science 264*, 433−436.

Gross, D. S., Huang, S.-Y., & Garrard, W. T. (1985) *J. Mol. Biol. 183*, 251−265.

Hamada, H., Petrino, M. G., & Kakunaga, T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6465−6469.

Haniford, D. B., & Pulleyblank, D. E. (1983) *Nature 302*, 632−634.

Hansma, H. G., Bezanilla, M., Zenhausern, F., Adrian, M., & Sinsheimer, R. L. (1993) *Nucleic Acids Res. 21*, 505−512.

Isackson, P. J., Fishback, J. L., Bidney, D. L., & Reeck, G. R. (1979) *J. Biol. Chem. 254*, 5569−5572.

Johns, E. W. (1982) *The HMG chromosomal proteins*, Academic Press, New York.

Joyce, C. M., & Grindley, N. D. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1830−1834.

Kashi, Y., Tikochinsky, Y., Genislav, E., Iraqi, F., Nave, A., Beckmann, J. S., Gruenbaum, Y., & Soller, M. (1990) *Nucleic Acids Res. 18*, 1129−1132.

Kladde, M. P., Kohwi, Y., Kohwi-Shigematsu, T., & Gorski, J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1898−1902.

Kornberg, A. (1980) *DNA replication*, pp 143−150, W. H. Freeman, San Francisco.

Kornberg, A., Bertsch, L. L., Jackson, J. F., & Khorana, H. G. (1964) *Proc. Natl. Acad. Sci. U.S.A. 51*, 315−323.

Levinson, G., & Gutman, G. A. (1987a) *Mol. Biol. Evol. 4*, 203−221.

Levinson, G., & Gutman, G. A. (1987b) *Nucleic Acids Res. 15*, 5323−5338.

Lu, M., Guo, Q., Marky, L. A., Seeman, N. D., & Kallenbach, N. R. (1992) *J. Mol. Biol. 223*, 781−789.

Naylor, L. H., & Clark, E. M. (1990) *Nucleic Acids Res. 18*, 1595−1601.

Nordheim, A., & Rich, A. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1821−1825.

Schlotterer, C., & Tautz, D. (1992) *Nucleic Acids Res. 20*, 211−215.

Singer, M. F. (1982) *Int. Rev. Cytol. 76*, 67−112.

Stringer, J. R. (1985) *Mol. Cell. Biol. 5*, 1247−1259.

Tautz, D. (1989) *Nucleic Acids Res. 17*, 6463−6471.

Tautz, D., Trick, M., & Dover, G. (1986) *Nature 322*, 652−656.

Treco, D., & Arnheim, N. (1986) *Mol. Cell. Biol. 6*, 3934−3947.

Treco, D., Thomas, B., & Arnheim, N. (1985) *Mol. Cell. Biol. 5*, 2029−2038.

Vorlickova, M., Kypr, J., Stokrova, S., & Sponar, J. (1982) *Nucleic Acids Res. 10*, 1071−1080.

Wahls, W. P., Wallace, L. J., & Moore, P. D. (1990) *Mol. Cell. Biol. 10*, 785−793.

Wells, R. D., Ohtsuka, E., & Khorana, H. G. (1965) *J. Mol. Biol. 14*, 221−240.

Wells, R. D., Jacobs, T. M., Narang, S. A., & Khorana, H. G. (1967) *J. Mol. Biol. 27*, 237−263.

BI962037V