Anomalous Rapid Electrophoretic Mobility of DNA Containing Triplet Repeats Associated with Human Disease Genes†

Paul D. Chastain II,‡ Evan E. Eichler,§ Seongman Kang,‡,|| David L. Nelson,§ Stephen D. Levene,⊥ and Richard R. Sinden*,‡

Institute of Biosciences and Technology, Center for Genome Research, Department of Biochemistry and Biophysics, Texas A&M University, 2121 West Holcombe Boulevard, Houston, Texas 77030-3303, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, Program in Molecular and Cell Biology, The University of Texas at Dallas, P.O. Box 830668, Richardson, Texas 75083, and Laboratory of Genetic Disease Research, National Center of Human Genomic Research, National Institute of Health, Building 49, Bethesda, Maryland 20892

Received August 9, 1995; Revised Manuscript Received October 3, 1995⊗

ABSTRACT: Eight human genetic diseases have been associated with the expansion of CTG or CGG triplet repeats. The molecular etiology behind expansion is unknown but may involve participation of an unusual DNA structure in replication, repair, or recombination. We show that DNA fragments containing CTG triplet repeats derived from the human myotonic dystrophy gene migrate up to 20% faster than expected in nondenaturing polyacrylamide gels, suggesting the presence of an unusual DNA helix structure within the CTG triplet repeats. The anomalous migration is dependent upon the number of triplet repeats, the length of the flanking DNA, and the percentage and temperature of the polyacrylamide. The effect could be reduced by the addition of actinomycin D. Applying a reptation model for electrophoresis, the results are consistent with a 20% increase in persistence length of the DNA. PCR products containing CTG or CGG repeats from the spinocerebellar ataxia type I gene (SCA1) or the fragile X FMR1 gene, respectively, also showed higher electrophoretic mobility. These are the first sequences of defined length for which a dramatic increase in mobility can be attributed to sequence-dependent structural elements in DNA.

Defined ordered sequence DNA (dosDNA) can form a variety of alternative structures in living cells, including bends, cruciforms, intramolecular triplexes, Z-DNA, and quadruplexes (Sinden, 1994; Zheng *et al.,* 1991; Ussery & Sinden, 1993; Rahmouni & Wells, 1989; McClellan *et al.,* 1990). The high correlation between dosDNA elements and human disease mutations (Krawczak & Cooper, 1991; Cooper & Krawczak, 1990) suggests that some alternative helical conformations may be inherently mutagenic (Sinden & Wells, 1992; Wells & Sinden, 1993). The unusual expansion of CTG and CGG triplet repeats is associated with eight human genetic diseases, including fragile X syndrome, myotonic dystrophy, and spinocerebellar ataxia type 1, as well as two fragile chromosomal sites which are not pathogenic (Nelson, 1993; Willems, 1994). Triplet repeat instability may be due to an anomalous helical structure giving rise to aberrant DNA replication, repair, or recombination.

The analysis of DNA mobility in polyacrylamide gels has been an extremely powerful tool for characterizing three- and four-way junctions, quadruplex DNA, flexible DNA, and especially bent DNA (Hagerman, 1990; Kahn *et al.,* 1994; Mills *et al.,* 1994). In agarose, both bent and inherently straight DNA molecules migrate similarly, whereas, in polyacrylamide, bent DNA migrates more slowly than expected. The degree of retardation is a function of the angle and position of the bend within the DNA molecule as well as the physical properties of the gel (Kerppola & Kane, 1990; Levene & Zimm, 1989; Zimm & Levene, 1992; Calladine *et al.,* 1991). Recent experiments show that the polyacrylamide-gel mobilities of DNA fragments containing hyperflexible regions, such as mispaired regions (Kahn *et al.,* 1994) or single-stranded gaps (Mills *et al.,* 1994), are reduced relative to their fully double-stranded counterparts. These results indicate that the bending rigidity of DNA is an important determinant of its electrophoretic behavior.

Most theories for the gel electrophoresis of DNA involve the concept of reptation, a model in which the DNA chain is assumed to be tightly confined by the gel so that motion of the chain can occur only along its axis [reviewed by Zimm and Levene (1992)]. The constraints prohibit sideways motion of the chain and therefore restrict motion to occur along a "tube" in the gel, much like the motion of a snake in a burrow. Using principally the assumption of snake-like motion, Lumpkin and Zimm (1982) and Lerman and Frisch (1982) showed that this model leads to a strong DNA-size dependence for the electrophoretic mobility. In the reptation model, mobility is proportional to the ratio $\langle h^2 \rangle / L^2$, where h is the DNA chain's end-to-end length, L is the chain contour length, and the angle brackets denote an average overall conformation. For DNA free in solution, this ratio can be calculated exactly and depends only on the ratio of L to the DNA persistence length, P, a parameter that characterizes the bending rigidity of DNA (Hagerman,

† Work in the authors' laboratories is supported in part by grants from NIEHS [ES05508 to (R.R.S.), NIHGMS (GM47898 to S.D.L.), and NICHD (HD29256 to D.L.N.). S.K. was supported by grants from NSF (9103942), NIHGMS (GM30822), and the Welch Foundation to R. D. Wells. Triplet repeat research by R.R.S., D.L.N., and R. D. Wells is currently supported by PO1 GM52982.

‡ Texas A&M University.
§ Baylor College of Medicine.
|| National Center of Human Genomic Research, NIH.
⊥ The University of Texas at Dallas.

⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1995.

0006-2960/95/0434-16125$09.00/0

1988). For mixed sequence DNA, P has been measured to be 45−50 nm at moderate ionic strength. Making use of the proportionality of mobility and average end-to-end dimensions of DNA fragments in the reptation model, we estimated the apparent persistence length of $(CTG)_n$-containing fragments to be 20% larger than that for electrophoretically normal control fragments.

EXPERIMENTAL PROCEDURES

Preparation of Triplet Repeat Containing DNAs. Plasmids pRW3211, pRW3213, pRW3214, pRW3216, pRW3218, and pRW3222 (derivatives of pUC19) containing 17, 50, 83, 98 ± 1, 175 ± 3, and 255 ± 5 CTG triplet repeats, respectively, with 62 bp[1] of human DNA flanking the DM gene were purified from *Escherichia coli* HB101 as described previously (Sinden *et al.*, 1980). The 17, 50, 83, and 98 bp repeats are pure $(CTG)_n$. The sequences of the triplet repeat region in pRW3218 and pRW3222 are $(CTG)_{27}ACT(CTG)_{40}ACT(CTG)_{106\pm3}$ and $(CTG)_{27}ACT(CTG)_{40}ACT(CTG)_{38}ACT(CTG)_{40}ACT(CTG)_{106\pm5}$, respectively. The length of the triplet repeat was determined by dideoxy sequencing. The length of the $(CTG)_{127}$ repeat (a deletion product in the pRW3222 plasmid preparation) was estimated from its relative mobility in agarose. After restriction, the $(CTG)_n$-containing fragments were extracted twice with phenol and twice with chloroform/isoamyl alcohol (24:1), the solution was adjusted to 0.3 M potassium acetate, and the DNA was precipitated by the addition of 2.5 volumes of 100% ethanol and incubation at −20 °C for at least 2 h. The DNA precipitate was collected by centrifugation, washed twice with 70% ethanol, dried, and resuspended in TE (10 mM Tris-HCl, pH 7.5, 1 mM EDTA). (Identical results were obtained without purification following restriction.)

*Not*I fragments from plasmids pRDW3246 and pRDW3248 containing $(CTG)_{98}$ and $(CTG)_{175}$, respectively (Kang *et al.*, 1995), were cloned into the *Sal*I site of the pBend2 vector (Kim *et al.*, 1989) to generate pPDC88L and pPDC165R, and pPDC169L (both *Not*I and *Sal*I ends were blunt ended by a fill-in reaction prior to ligation). (The *Not*I sites are located outside the *Eco*RI and *Hin*dIII sites shown in Figure 1A.) The number of CTG repeats is denoted in the plasmid designation, and R and L refer to the orientation of the CTG repeat, with 5′ CTG 3′ in the top strand in the R orientation. Plasmids containing spontaneous deletions in the CTG region, derived from pPDC165R, were purified from *E. coli* and transformed into HB101 to produce pPDC25R, pPDC31R, pPDC46R, and pPDC69R. The lengths of the $(CTG)_n$ repeats were estimated from mobilities in agarose gels ($bp_{agarose}$) and are accurate within 5−10%.

The conditions for preparation of the FMR1 gene and SCA1 gene PCR products were described by Chong *et al.* (1994) and Orr *et al.* (1993), respectively. The FMR1 gene products have 196 and 144 bp of flanking sequence, while the SCA1 gene products have 43 and 81 bp of flanking sequence on the 5′ and 3′ sides of the repeats, respectively.

Polyacrylamide and Agarose Gel Electrophoresis. Nondenaturing polyacrylamide gel electrophoresis was performed in 24 × 14 × 0.8 cm gels using an acrylamide/bis-(acrylamide) ratio of 29:1 in TAE buffer (40 mM Tris, pH 8.3, 25 mM acetate, 1 mM EDTA) containing 10% glycerol (which produces sharper bands). (The results were the same without addition of glycerol.) For most gels, the current was kept below 12 mA to avoid heating. For the data shown in Figure 2, a field strength of 10 V/cm was maintained. Two percent agarose gels (BRL) were run in TAE buffer at room temperature at 2 V/cm.

RESULTS

Anomalous Gel Mobility of CTG Triplet Repeats Associated with Myotonic Dystrophy. The gel mobilities of two series of DNA fragments encompassing a region of the human myotonic dystrophy (DM) gene containing $(CTG)_n$, where n ranged from 17 to 255, were analyzed in polyacrylamide and agarose gels. One series of *Sac*I−*Hin*dIII fragments contained 44 and 64 bp of non-triplet-repeat sequences flanking $(CTG)_n$, whereas a second series of *Pvu*II fragments contained 140 and 245 bp on the 5′ and 3′ sides of the repeats, respectively (Figure 1A). Both series of restriction fragments containing $(CTG)_n$ migrated faster (had a higher mobility value) than expected in polyacrylamide gels. The ratios of the apparent size to actual size (bp_{app}/bp_{seq}) decreased with increasing lengths of the $(CTG)_n$ repeat to a 20% lower value than expected for $(CTG)_{255}$ (Figure 1B,C). For a defined length of $(CTG)_n$ repeat, the bp_{app}/bp_{seq} values were slightly lower for the *Pvu*II fragments, which had longer lengths of flanking DNA, compared with the *Sac*I−*Hin*dIII fragments (Figure 1C). The migration was more rapid at higher polyacrylamide gel concentrations (Figure 1D) and at lower temperatures (Figure 1E). The migration of $(CTG)_n$-containing fragments in agarose was slightly faster than expected but independent of repeat length (on average, $bp_{app}/bp_{seq} = 0.95 \pm 0.02$) (Figure 1B).

Analysis of Mobility According to a Reptation Model for Electrophoresis. We have analyzed mobility in terms of a reptation model for electrophoresis (Levene & Zimm, 1989) to determine an apparent value for persistence length of the $(CTG)_n$-containing DNA fragments. The complete expression for the electrophoretic mobility, μ, derived by Lumpkin and Zimm (1982) is

$$\mu = \frac{Q\langle h_z^2\rangle}{\zeta L^2}$$

where Q is the total charge on the DNA chain, ζ is the friction coefficient for translation along the tube, h_z is the component of the chain end-to-end vector in the direction of the electric field (taken to lie along the z axis), and L is the contour length of the DNA. h_z depends on the conformation and flexibility of DNA as well as the orienting effects of the electric field. To a good approximation, both Q and ζ are proportional to chain length, and thus the ratio of these quantities will be independent of chain length. Moreover, in tight gels such as polyacrylamide, the orienting effects of the applied field are negligible and therefore, $\langle h_z^2\rangle = \langle h^2\rangle/3$. We therefore expect that $\mu \propto \langle h^2\rangle/L^2$.

We assume that the bending rigidity of DNA in a gel can be characterized in terms of an effective value for the persistence length. A simple mathematical expression relates the ratio of the DNA chain's mean-squared end-to-end length

[1] Abbreviations: bp, base pairs; SCA1, spinocerebellar ataxia type I; FMR1, fragile X mental retardation gene 1; PCR, polymerase chain reaction.

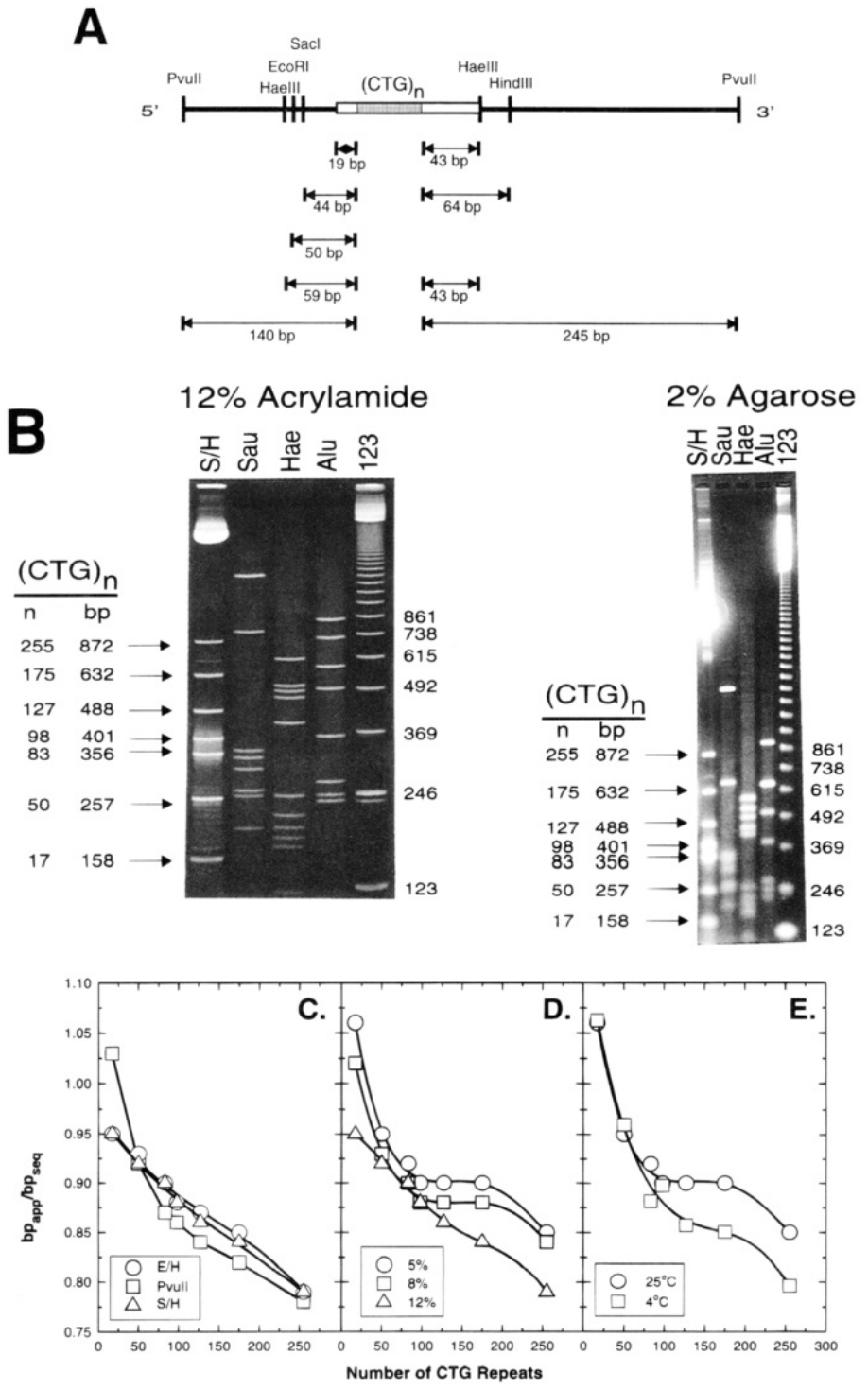


FIGURE 1: Electrophoretic behavior of triplet repeat-containing fragments in agarose and polyacrylamide. (A) The box represents the fragment from the human DM gene and the shaded region shows the location of the CTG triplet repeats. Positions of relevant restriction sites and lengths are indicated. (B) Mobility of *Sac*I−*Hin*dIII fragments in 12% polyacrylamide and 2% agarose at room temperature. Marker lanes are indicated as follows: 123, a 123 bp ladder (BRL); Alu, *Alu*I digest of pBR322; Hae, *Hae*III digest of pBR322; Sau, *Sau*3AI digest of pBR322. The sizes of the *Alu*I fragments are 910, 659, 656, 521, 403, 281, 257, and 226 bp. The sizes of the *Hae*III fragments are 587, 540, 504, 458, 434, 267, 234, 213, 192, 184, 124, and 104 bp. The sizes of the *Sau*3AI fragments are 1374, 665, 359, 341, 272, 258, and 207 bp. Lane S/H refers to the *Sac*I-*Hin*dIII digests of $(CTG)_n$-containing plasmids, where *n* and the lengths of the *Sac*I-*Hin*dIII fragments are indicated. *Sac*I-*Hin*dIII digests of individual plasmid DNAs were mixed together to give a sample containing all of the triplet-repeat-containing fragments. (C) Effect of length of the CTG repeats on bp_{app}/bp_{seq} in 12% polyacrylamide gels. bp_{app}/bp_{seq} (the apparent length in bp ÷ the actual length in bp) is plotted as a function of the number of CTG triplet repeats in each different fragment. bp_{app} is determined from a plot of the log of the length versus the distance migrated for the 123 bp ladder. [Analysis of mobility with respect to the inherently straight restriction fragments from pBR322 (excluding known bent fragments, which do not migrate as the same length on acrylamide and agarose) confirms the rapid mobility of the $(CTG)_n$-containing molecules.] The linear regression of the calibration curves resulted in a correlation coefficient of >0.995. (○) *Eco*RI−*Hin*dIII fragments; (△) *Sac*I−*Hin*dIII fragments; (□) *Pvu*II fragments. (D) bp_{app}/bp_{seq} values are shown for the *SacI*−*HindIII* fragments in 5% (○), 8% (□), and 12% (△) polyacrylamide. (E) bp_{app}/bp_{seq} values are shown for *Sac*I−*Hin*dIII fragments in 5% polyacrylamide at 25 (○) and 4 °C (□).

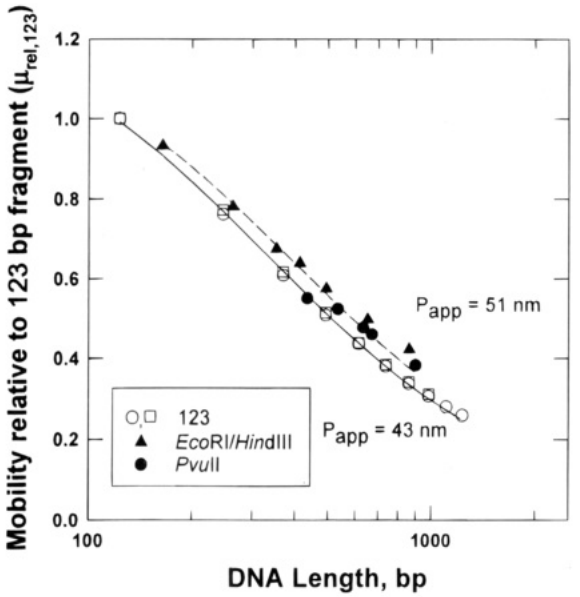


FIGURE 2: Analysis of the electrophoretic mobility of $(CTG)_n$-containing fragments in terms of a reptation model for wormlike DNA chains. Mobilities of DNA fragments in 12% polyacrylamide normalized to that for the 123 bp marker fragment ($\mu_{rel,123}$) were fitted to

$$\mu_{rel} = 2A\left[\frac{P_{app}}{L} - \frac{P_{app}^2}{L^2} + \frac{P_{app}^2}{L^2}e^{-L/P_{app}}\right]$$

where A and P_{app} were the fitted parameters. The nonlinear least-squares curve fitting was done using a standard implementation of the Marquardt−Levenberg algorithm in SigmaPlot (Jandel Scientific). (○, □) 123-bp-ladder fragments from duplicate lanes on the same gel, (▲) *Eco*RI/*Hin*dIII $(CTG)_n$-containing fragments, (●) *Pvu*II $(CTG)_n$-containing fragments. Predicted dependencies of mobility on DNA length from the reptation analysis are given by the smooth curves; (—) "best fit" to the data for the 123-bp-ladder fragments with P_{app} = 43 nm; (- - -) "best fit" to the data for the $(CTG)_n$-containing fragments with P_{app} = 51 nm. Estimated coefficients of variation for P_{app} from the nonlinear least-squares fits were 2.2% and 7.9%, respectively.

to the square of the contour length, $\langle h^2\rangle/L^2$, to the ratio, $X = L/P$, of DNA contour length to persistence length (Hagerman, 1988).

$$\frac{\langle h^2\rangle}{L^2} = \frac{2}{X}\left[1 - \frac{1}{X}(1 - e^{-X})\right]$$

The ratio $\langle h^2\rangle/L^2$ inferred from the mobility is used to estimate the apparent persistence length, P_{app}, using the reptation model. The present method of analysis leads to absolute values of P that are close to, but not necessarily identical with, that measured in solution (S.D.L., unpublished results).

The results of this analysis are shown in Figure 2. In 12% polyacrylamide, P_{app} values for the 123 bp ladder fragments and the $(CTG)_n$-containing fragments are both in the range measured for mixed-sequence DNA in solution. However, the increased mobility of the $(CTG)_n$-containing DNA relative to control fragments can be accounted for by a 20% increase in the apparent persistence length from P_{app} = 43−51 nm.

Effect of Actinomycin D on Mobility of (CTG)-Containing DNA in Acrylamide. If anomalous mobility is a property of a helix structure adopted by $(CTG)_n$ repeats, we reasoned

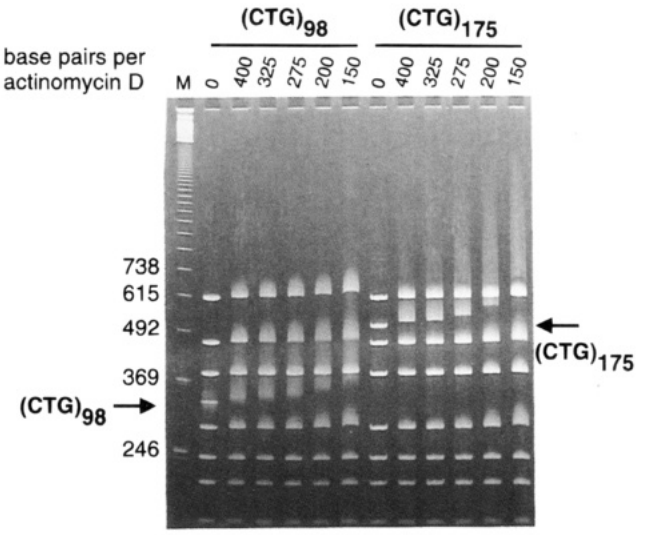
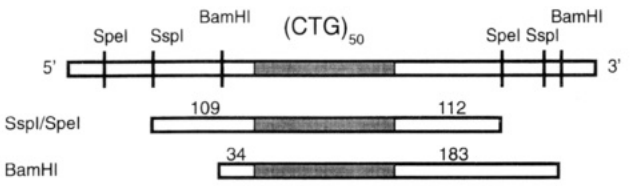


FIGURE 3: Influence of actinomycin D on anomalous mobility. *Hae*III restriction digests of pRW3216 and pRW3218 (containing 98 and 175 CTG repeats, respectively) were incubated at 50 μg/mL for 1 h at 37 °C in the presence of actinomycin D before electrophoresis. Actinomycin D (Sigma) was dissolved in ethanol. The number of base pairs per actinomycin D molecule is listed above each lane. The marker lane, M, contains the 123 bp ladder with sizes shown to the left. The position of the $(CTG)_{98}$- and $(CTG)_{175}$-containing bands are indicated with arrows.

that it may be possible to disrupt or perturb this structure and reduce the mobility value. To determine if the anomalous mobility could be altered by the binding of a DNA intercalating drug, actinomycin D was added to the $(CTG)_n$-containing molecules during electrophoresis. Actinomycin D, at bp/drug ratios of 400−150, decreased the mobility of $(CTG)_n$-containing *Hae*III fragments in polyacrylamide resulting in an increase of bp_{app}/bp_{seq} from 0.8 to ≈1 (Figure 3). Under these conditions, the structural properties of the $(CTG)_n$ repeat leading to the unusual rapid migration were apparently disrupted. However, at the same concentrations, actinomycin D had little effect on the mobility of other [non-$(CTG)_n$-containing] *Hae*III digestion products (Figure 3). This indicates a preferential structural perturbation of the CTG-tract helix structure, which may result from introduction of a bend or the creation of a flexible region within the CTG tract. Moreover, this may reflect a preferential binding of actinomycin D to CTG tracts. At higher concentrations, actinomycin D, which is known to bend DNA (Reinert, 1991), appeared to introduce bends into the $(CTG)_n$-containing molecules as indicated by bp_{app}/bp_{seq} values >2.5 (data not shown).

Analysis of the Mobility of Circularly Permuted DNA Molecules Containing $(CTG)_n$. A family of circularly permuted DNA molecules was generated to analyze the dependence of the anomalous mobility on the position of the $(CTG)_n$ triplet repeat within a linear fragment. For this, DM-derived fragments containing various lengths of $(CTG)_n$ were cloned into the pBend2 vector which allows the facile generation of a family of circularly permuted DNAs. Analysis of a series of three circularly permuted DNAs, in which the 5′ flanking region varied from 109 to 181 bp and the 3′ flanking region varied from 88 to 160 bp, revealed a change in $bp_{app}/bp_{agarose}$ of only 1% for $(CTG)_{165}$- and $(CTG)_{169}$-containing molecules and 3% for a $(CTG)_{88}$-containing molecule (data not shown). Because of the relatively long lengths of the $(CTG)_n$ repeats with respect to



CTG	$bp_{app}/bp_{agarose}$ BamHI	SpeI/SspI
25R	0.93	0.97
31R	0.91	0.94
46R	0.91	0.93
69R	0.89	0.92

FIGURE 4: Effect of location of $(CTG)_n$ within a restriction fragment on mobility. Plasmids pPDC25R, pPDC31R, pPDC46R, and pPDC69R were cut with both *Ssp*I and *Spe*I (sites in pBend2 vector) or with *Bam*HI (which cuts once in the DNA flanking the DM sequence as well as within the pBend2 vector). The *Ssp*I/*Spe*I and *Bam*HI restriction fragments differ in length by 4 bp.

the flanking sequences in these experiments, smaller repeat tracts were analyzed in two molecules that placed the $(CTG)_n$ repeats in the center or near one end of nearly identically sized molecules. $(CTG)_n$ repeat lengths, where $n = 25-69$, showed 2−4% lower $bp_{app}/bp_{agarose}$ values when the triplet repeats were positioned near the end of a restriction fragment compared with when positioned near the center (Figure 4).

Anomalous Mobility of DNA Containing $(CTG)_n$ Repeats from SCA1 and $(CGG)_n$ Repeats from FMR1. $(CTG)_n$- and $(CGG)_n$-containing PCR products of the human SCA1 gene, associated with spinocerebellar ataxia type 1, and the FMR1 gene, associated with fragile X syndrome, respectively, also migrated faster than expected in polyacrylamide gels (Table 1). The bp_{app}/bp_{seq} values for the CTG-containing SCA1 gene products were similar to those shown in Figure 1C for CTG-containing DM-derived fragments, despite having different sequences flanking the triplet repeat. The bp_{app}/bp_{seq} values decreased from 0.89 to 0.85 with increasing repeat length from $(CTG)_{19}$ to $(CTG)_{74}$. This result strongly suggests that the anomalous mobility is not due to the human DNA sequences flanking the $(CTG)_n$ repeats. The anomalous mobility was also observed for all CGG-containing fragments analyzed from the FMR1 gene of humans, gorillas, and chimpanzees, which have various levels of interruptions of a pure CGG repeat. bp_{app}/bp_{seq} values ranged from 0.88 to 0.85 for between 22 and 55 triplet repeats.

DISCUSSION

We demonstrate that $(CTG)_n$ and $(CGG)_n$ triplet repeats from genes associated with DM, FRAXA, and SCA1 confer an anomalous rapid electrophoretic gel mobility on DNA. When a reptation model for electrophoresis is applied [reviewed in Zimm and Levene (1992)], the anomalous rapid electrophoretic mobilities of triplet repeat-containing molecules are consistent with a 20% increase in DNA persistence length. Although we cannot entirely rule out other explanations for the increased mobility of DNA fragments containing $(CTG)_n$ and $(CGG)_n$ triplet repeats, the data are compatible with the observation that increased flexibility of DNA decreases gel mobility (Kahn *et al.*, 1994; Mills *et al.*, 1994). The increase in the apparent persistence length could result from decreased bendability relative to mixed-sequence DNA, diminished sequence-dependent contributions of local roll and tilt components, or a combination of these effects. Similar to the observed behavior of intrinsically bent A-tract DNA, the anomalous electrophoretic mobility associated with CTG and CGG repeats is sensitive to changes in temperature and gel concentration. Taken together with the observation that the binding of low levels of actinomycin D alters the mobility of these triplet-repeat-containing fragments, these data suggest that a special helical structure is responsible for the anomalous mobility. The anomalous rapid mobility

Table 1: bp_{app}/bp_{seq} and $\mu_{rel,123}$ Values of FMR1 and SCA1 Gene Products

PCR product	total repeats	triplet repeat sequence	bp_{app}/bp_{seq}[a]	$\mu_{rel,123}$[b]
		FMR1 $(CGG)_n$ Products		
gorilla	22	$(CGG)_8AGGCGGAGG(CGG)_3AGG(CGG)_3AGG(CGG)_3$	0.86	0.62
SC72	23	$(CGG)_{13}AGG(CGG)_9$	0.86	0.62
SC81	24	$(CGG)_{24}$	0.85	0.62
SC47	29	$(CGG)_9AGG(CGG)_{19}$	0.86	0.60
SC50	29	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	0.87	0.60
chimp	29	$(CGG)_8AGGCGGAGG(CGG)_2AGG(CGG)_2AGG(CGG)_{12}$	0.88	0.60
SC57	31	$(CGG)_{21}AGG(CGG)_9$	0.87	0.60
GD1	34	$(CGG)_{34}$	0.86	0.59
53−654	38	$(CGG)_{38}$	0.86	0.58
SC43	44	$(CGG)_9AGG(CGG)_9AGG(CGG)_{24}$	0.85	0.56
A17	44	$(CGG)_9AGG(CGG)_{34}$	0.88	0.57
XVII-1	47	$(CGG)_9AGG(CGG)_{37}$	0.87	0.56
IV-1	48	$(CGG)_{10}AGG(CGG)_{37}$	0.85	0.56
IX-1	49	$(CGG)_9AGG(CGG)_9AGG(CGG)_{29}$	0.85	0.56
SC70	54	$(CGG)_{54}$	0.87	0.54
1408−08	55	$(CGG)_{55}$	0.86	0.55
		SCA1 $(CTG)_n$ Products		
TL #1	19	$(CAG)_{19}$	0.89	0.92
NC #1	21	$(CAG)_{21}$	0.88	0.91
TH6A	25	$(CAG)_7CATCAGCAT(CAG)_{15}$	0.88	0.89
DM #1	30	$(CAG)_{12}CATCAGCAT(CAG)_{15}$	0.87	0.87
KM3B M	31	$(CAG)_{13}CATCAGCAT(CAG)_{15}$	0.87	0.86
DM #2	36	$(CAG)_{17}CATCAGCAT(CAG)_{16}$	0.87	0.84
EH7 M	55	$(CAG)_{55}$	0.85	0.77
NC #2	56	$(CAG)_{56}$	0.87	0.76
KM3E P	74	$(CAG)_{74}$	0.86	0.70

[a] bp_{app}/bp_{seq} values calculated as described in text. [b] $\mu_{rel,123}$, mobility relative to the 123 bp fragment.

is not simply due to G+C content since *Eco*RI/*Hin*dIII (62.2−65.8% G+C) and *Pvu*II (54.1−61.9% G+C) fragments migrate similarly despite differences in G+C content. Neither is the anomalous mobility simply related to the proportion of DNA length containing (CTG) repeats. Experiments involving cyclization kinetics and rotational relaxation time measurements are in progress to further characterize the structure of these triplet repeats.

The dependence of mobility on gel concentration has been demonstrated for DNA containing hyperflexible regions (Kahn *et al.,* 1994; Mills *et al.,* 1994), although the overall effect of the degree and position of the flexible bend was substantially less than that due to static bends of similar magnitude (Kahn *et al.,* 1994). A similar effect may occur with sequences of increased bending rigidity, which may explain the weak dependence of mobility on the position of the repeats within the DNA molecule. Moreover, a modest effect of increased flexibility on the overall mobility of DNA fragments (Kahn *et al.,* 1994; Mills *et al.,* 1994) raises the possibility that the value of P_{app} obtained here may underestimate the bending rigidity of the CTG region. A detailed understanding of the molecular basis for this electrophoretic anomaly may further the development of DNA electrophoresis theory.

Short single stranded CTG, CAG, CGG, and CCG repeats can form hairpin structures containing two C·G base pairs and T·T, A·A, G·G, or C·C mispairs (Gacy *et al.,* 1995; Yu *et al.,* 1995; Gao *et al.,* 1995). These structures, however, may have no relevance to the duplex structure described here. Heteroduplex duplex DNAs containing CTG and CAG tracts of different length and slipped strand structures between equal length triplet repeat tracts exhibit a 1.5−2.5-fold reduction in the mobility of DNA in polyacrylamide (C. E. Pearson and R.R.S., unpublished results). This would be expected if the triplet repeat tracts in duplex DNA form hairpin structures and create three- and four-way junctions that bend DNA (Cooper & Hagerman, 1987; Duckett & Lilley, 1990).

The anomalous rapid mobilities of a "structural control element" from the DHFR gene (Pierce *et al.,* 1992), poly(dA)·poly(dT) (Koo *et al.,* 1986; Anderson, 1986), and several other A+T- and G+C-rich sequences (Anderson, 1986) as well as $(CTG)_n$ and $(CGG)_n$ repeats suggests that the helical properties responsible for the unusual mobility, while to date rare, may be a biologically important structural feature of DNA. Bending rigidity may be a property of tracts of DNA high in G+C-rich triplets and/or PuPu dinucleotides. The similar mobility of CGG sequences from the FMR1 gene of gorilla and chimpanzee compared to the human sequences (Table 1) demonstrates that long regions of pure repeats are not necessary for the anomalous mobility. The $(CGG)_n$ tracts found in normal and fragile X individuals contain AGG interruptions, and the length of the pure $(CGG)_n$ tract correlates with the genetic instability (Eichler *et al.,* 1994). Although there was no significant difference in the anomalous mobility of sequences containing pure repeats or those with interruptions, the unusual helical structure may contribute to length-dependent events responsible for expansion.

The molecular basis for this electrophoretic anomaly may be an important factor in the instability and expansion of CTG and CGG triplet repeats when transmitted in somatic cells (Wong *et al.,* 1995) or from one generation to the next (Eichler *et al.,* 1994; Nelson, 1993; Willems, 1994). DNA with increased bending rigidity may resist organization into nucleosomes, and, especially upon expansion, the local or global chromosomal architecture may be altered, leading to changes in gene regulation, the timing and/or fidelity of replication, DNA repair, or the probability of genetic recombination.

ACKNOWLEDGMENT

We thank Huda Zoghbi for providing the human SCA1 DNA samples. We thank M. Angela Parsons, Christopher E. Pearson, Vladimir Potaman, and Robert D. Wells for comments regarding the manuscript. We also thank Bozenna Jaworska for technical assistance.

REFERENCES

Anderson, J. N. (1986) *Nucleic Acids Res. 14*, 8513−8533.
Calladine, C. R., Collis, C. M., Drew, H. R., & Mott, M. R. (1991) *J. Mol. Biol. 221*, 981−1005.
Chong, S. S., Eichler, E. E., Nelson, D. L., & Hughes, M. R. (1994) *Am. J. Med. Genet. 51*, 522−526.
Cooper, J. P., & Hagerman, P. J. (1987) *J. Mol. Biol. 198,* 711−719.
Cooper, D. N., & Krawczak, M. (1990) *Hum. Genet. 85*, 55−74.
Duckett, D. R., & Lilley, D. M. J. (1990) *EMBO J. 9,* 1659−1664.
Eichler, E. E., Holden, J. J. A., Popovich, B. W., Reiss, A. L., Snow, K., Thibodeau, S. N., Richards, C. S., Ward, P. A., & Nelson, D. L. (1994) *Nature Genet. 8*, 88−94.
Gacy, A. M., Goellner, G., Juranic, N., Macura, S., & McMurry, C. T. (1995) *Cell 81*, 533−540.
Gao, X., Huang, X., Smith, G. K., Zheng, M., & Liu, H. (1995) *J. Am. Chem. Soc. 117*, 8883−8884.
Hagerman, P. J. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 265−286.
Hagerman, P. J. (1990) *Annu. Rev. Biochem. 59*, 755−781.
Kahn, J. D., Yun, E., & Crothers, D. M. (1994) *Nature 368*, 163−166.
Kang, S., Jaworski, A., Ohshima, K., & Wells, R. D. (1995) *Nature Genet. 10*, 213−218.
Kerppola, T. K., & Kane, C. M. (1990) *Biochemistry 29*, 269−278.
Kim, J., Zwieb, C., Wu, C., & Adhya, S. (1989) *Gene 85*, 15−23.
Koo, H. S., Wu, H. M., & Crothers, D. (1986) *Nature 320*, 501−506.
Krawczak, M., & Cooper, D. N. (1991) *Hum. Genet. 86*, 425−441.
Lerman, L. S., & Frisch, H. L. (1982) *Biopolymers 21*, 995−997.
Levene, S. D., & Zimm, B. H. (1989) *Science 245*, 396−399.
Lumpkin, O. J., & Zimm, B. H. (1982) *Biopolymers 21*, 2315−2316.
McClellan, J. A., Boublikova, P., Palecek, E., & Lilley, D. M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8373−8377.
Mills, J. B., Cooper, J. P., & Hagerman, P. J. (1994) *Biochemistry 33*, 1797−1803.
Nelson, D. L. (1993) in *Genome Analysis Volume 7: Genome Rearrangement and Stability* (Davies, K., & Warren, S., Eds.) pp 1−24, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Orr, H. T., Chung, M., Banfi, S., Kwiatkowski, T. J., Jr., Servadio, A., Beaudet, A. L., McCall, A. E., Duvick, L. A., Ranum, L. P. W., & Zoghbi, H. Y. (1993) *Nature Genet. 4*, 221−226.
Pierce, A. J., Jambou, R. C., Jensen, D. E., & Azizkhan, J. C. (1992) *Nucleic Acids Res. 20*, 6583−6587.
Rahmouni, A. R., & Wells, R. D. (1989) *Science 246*, 358−363.
Reinert, K. E. (1991) *J. Biomol. Struct. Dyn. 9*, 331−352.
Sinden, R. R. (1994) *DNA Structure and Function*, Academic Press, San Diego, CA.
Sinden, R. R., & Wells, R. D. (1992) *Curr. Opin. Biotechnol. 3*, 612−622.
Sinden, R. R., Carlson, J. O., & Pettijohn, D. E. (1980) *Cell 21*, 773−783.

Ussery, D. W., & Sinden, R. R. (1993) *Biochemistry 32*, 6206–6213.
Wells, R. D., & Sinden, R. R. (1993) in *Genome Analysis Volume 7: Genome Rearrangement and Stability* (Davies, K., & Warren, S., Eds.) pp 107–138, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Willems, P. J. (1994) *Nature Genet. 8*, 213–215.
Wong, L.-J. C., Ashizawa, T., Monckton, D. G., Caskey, C. T., & Richards, C. S. (1995) *Am. J. Hum. Genet. 56*, 114–122.
Yu, A., Dill, J., Wirth, S. S., Huang, G., Lee, V. H., Haworth, I. S., & Mitas, M. (1995) *Nucleic Acids Res. 23*, 2706–2714.
Zheng, G., Kochel, T., Hoepfner, R. W., Timmons, S. E., & Sinden, R. R. (1991) *J. Mol. Biol. 221*, 107–129.
Zimm, B. H., & Levene, S. D. (1992) *Q. Rev. Biophys. 25*, 171–204.

BI9518611