EXPERIMENTAL and MOLECULAR MEDICINE, Vol. 34, No. 4, 313-317, September 2002

Human FEN-1 can process the 5'-flap DNA of CTG/CAG triplet repeat derived from human genetic diseases by length and sequence dependent manner

Suman Lee[1,3] and Min S. Park[2]

[1] Genome Research Center for Reproductive Medicine and Infertility, CHA General Hospital, College of Medicine, Pochon University, Seoul, Korea
[2] Life science division, M888, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[3] Correspondence author: Tel, +82-2-3468-3445; Fax, +82-2- 3468-3464; E-mail, suman@cha.ac.kr

Accepted 16 September 2002

Abbreviations: FEN-1; human flap endonuclease 1, TNR; Trinucleotide repea; nt, nucleotides; ss, single strand

Abstract

Trinucleotide repeat (TNR) instability can cause a variety of human genetic diseases including myotonic dystrophy and Huntington's disease. Recent genetic data show that instability of the CAG/CTG repeat DNA is dependent on its length and replication origin. In yeast, the RAD27 (human FEN-1 homologue) null mutant has a high expansion frequency at the TNR loci. We demonstrate here that FEN-1 processes the 5'-flap DNA of CTG/CAG repeats, which is dependent on the length *in vitro*. FEN-1 protein can cleave the 5'-flap DNA containing triplet repeating sequence up to 21 repeats, but the activity decreases with increasing size of flap above 11 repeats. In addition, FEN-1 processing of 5'-flap DNA depends on sequence, which play a role in the replication origin-dependent TNR instability. Interestingly, FEN-1 can cleave the 5'-flap DNA of CTG repeats better than CAG repeats possibly through the flap-structure. Our biochemical data of FEN-1's activity with triplet repeat DNA clearly shows length dependence, and aids our understanding on the mechanism of TNR instability.

Keywords: trinucleotide repeats, trinucleotide expansion, genetics, neuromuscular direases, DNA replication

Introduction

The genetic instability of trinucleotide repeats (TNR) is important for a number of human inherited neuromuscular and neurological diseases, such as Huntington's disease (HD), muscular myotonic dystrophy (MMD), Fragile X syndrome and Spinocerebellar ataxia (Warren and Nelson, 1993). The expansion of TNR in the particular loci of human genome can cause the genetic diseases while normal individuals can maintain a short range of TNR. e.g. 10-35 CTG repeats in the case of Huntington's disease (Oostra and Willems, 1995). The mechanism of TNR expansion is not clear so far, but it has been suggested that the realignment or slippage mediated by the characteristics of the repeat sequence on the Okazaki fragment can lead to the instability of TNR. (Gordenin *et al.*, 1997). Structural studies of TNR DNA point out that single strand (ss) - triplet repeat DNA (CNG, N equals the nucleotides, A, G, C, or T) can partially form hairpin structures, as demonstrated by NMR and biochemical studies (Mitas, 1997; McMurray, 1999). During lagging strand synthesis, which has a greater chance to form ss regions compared to leading strand synthesis, the unusual structural formation (hairpin-type) from ss DNA can give rise to replication errors at the TNR loci. There is evidence to suggest that triplet repeat instability is dependent on the character of repeat DNA itself, such as its length and sequence, in *E.coli* and yeast (Kang *et al.*, 1996; Freudenreich *et al.*, 1997).

Recent yeast genetic data suggest that factors other than DNA structure may affect TNR instability. The disruption of the RAD27 gene of *Saccharomyces cerevisiae* (a homologue of human flap endonuclease 1, FEN-1) did affect the instability of TNR while the disruption of mismatch or recombination related genes did not change the rates of TNR instability (Maurer *et al.*,1996; Freudenreich *et al.*, 1998). Zakian *et al* showed that a deletion mutant of RAD27 caused length-dependent destabilization of CTG tracts and increased in its expansion frequency. FEN-1 is highly conserved between yeast and human, and is a 5' flap specific endonuclease (Li *et al.*, 1995). The 5'-flap structure can occur by strand displacement at 5' ends of Okazaki fragments during lagging strand synthesis, and can be removed by FEN-1 and other replication proteins, resulting in nicked double strand DNA suitable for ligation (Gordenin *et al.*, 1997; Bae and Seo, 2000). On primer-template substrates containing a flap structure, FEN1 employs unknown mechanism to cleave the 5'-tail. FEN1 appears to track along the full length of the flap from the 5'-end to the point of cleavage. (Bornarth *et al.*, 1999). Therefore, the investigation on the FEN-1 protein

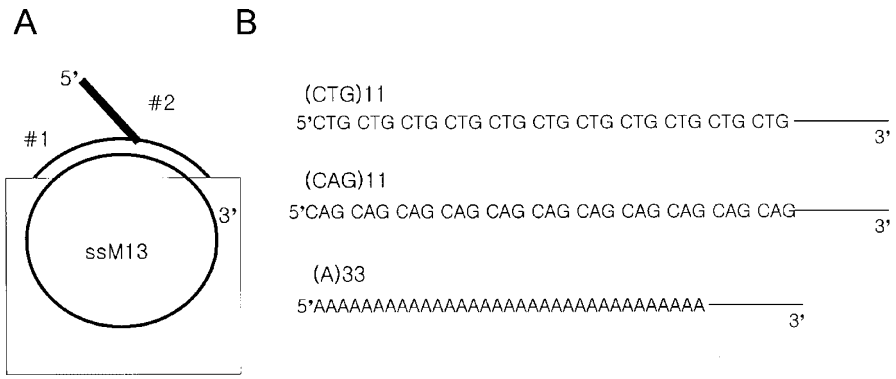

Figure 1. Schematic drawing of M13 based substrates for FEN-1 cleavage assay (A). The details on the preparation of these substrates are described in Materials and Metheds. The eleven CTG, CAG repeats and the poly(adenine) sequences of the flap [(bold line on (A)] are shown in (B).

functions on CTG/CAG repeat DNA is crucial for understanding the mechanism of triplet repeat instability.

Here, our data showed that FEN-1 can process CTG/CAG repeat DNA dependent on its repeat-length and orientation. First, the longer 5'-flap of triplet repeats are inhibitory for FEN-1 cleavage. Our FEN-1 endonuclease assay with different sizes of CTG or CAG repeat flaps (from 11 to 21 repeats) showed that over 11 repeat lengths of CTG or CAG were inhibitory for their processing by FEN-1. This suggests that the individuals who have longer repeats may have a greater chance of avoiding the FEN-1 repair mechanism by making longer 5 -flap structures during lagging strand synthesis, which may result in expansion of TNR. Second, the replication direction has also been shown to affect triplet repeat instability. In the case of the CTG/CAG tract, differential instability showed dependence on which strand is the Okazaki fragment. The unstable orientation is when the CAG tract is the Okazaki fragment. Therefore, it would be interesting to test whether FEN-1 can process 5'-flap DNAs differently depending on sequence (5'CTG3' or 5'CAG3' repeat flaps). Indeed, we report here that FEN-1 can process a 5' CTG-flap better than a 5' CAG-flap which might be due to the hairpin structure.

These biochemical studies of FEN-1 with triplet repeat instability explain previous yeast genetic data, and suggest that FEN-1 has an important role in length- and sequence- dependent triplet repeat instability in human genetic diseases.

Materials and Methods

Proteins and Materials

FEN-1 protein was purified by the method of Nolan *et al.* (Nolan *et al.*, 1996). After purification, the protein was dialyzed into Tris-HCl , pH 7.9, 100 mM NaCl, 10 mM 2-mercaptoethanol and 10% glycerol, and stored at -70°C. The protein concentration was determined by Bradford assay (Bio-Rad, Hircules, CA) using BSA as control, and the purity confirmed by SDS- PAGE (approximately 95% pure). BSA, M13 mp19 ss DNA and T4 polynucleotide kinase were purchased from New England Biolabs (Beverly, MA). Oligonucleotide sizing marker was purchased from Pharmacia Biotech (NJ).

Oligonucleotides

Oligonucleotides were synthesized on an Applied Biosystem (ABI) DNA synthesizer and further purified on a 10% 0.5X TBE Urea-gel by the method indicated in Sambrook *et al.* (Sambrook, 1989; Kim *et al.*, 2000; Lee *et al.*, 2001). The sequences of the oligonucleotides are indicated below. M-13 flap adjacent strand (#1 in Figure 1A): 5' CGG CCA GTG AAT TCG AGC TCG GTA CCC GGG 3', Flap strands (#2 in Figure 1 A) are; CTGn: 5' (CTG)n GAT CCT CTA GAG TCG ACC TGC AGG CAT GCA AGC C3', CAGn: 5' (CAG)n GAT CCT CTA GAG TCG ACC TGC AGG CAT GCA AGC C3', (n=11, 15, 21 repeats). PAL11: 5'CAG CAG CAG CAG CAG TTT CTG CTG CTG CTG CTG GAT CCT CTA GAG TCG ACC TGC AGG CAT GCA AGC C3'. The annealing portion of oligonucleotides to M13 mp19 ss DNA was 30 to 34 nt-long and is indicated by underlining. The quantity was determined on a UV spectrophotometer at 260 nm wavelength, and DNA was labeled by T4 poly-nucleotide kinase with [γ-^{32}P]ATP (NEN) following the method of Sambrook *et al.* (Sambrook, 1989).

M13 endonuclease assay

The M13 flap adjacent strand (#1 in Figure 1A) and ^{32}P-labeled flap oligonucleotide (#2 in Figure 1A) were mixed with M13 mp19 ss DNA (1 μg) at the molar ratio of 5:5:1. The mixtures were annealed in a Perkin Elmer thermocycler by slowly decreasing the temperature from 95°C to 20°C (0.4°C/min). Unannealed oligonucleotides were removed by PCR purification kit (Qiagen)

according to the manufacture's method. The quantity of DNA and radioactivity of final M13-based substrates were measured by GeneQuant (Pharmacia) and Bioscan QC 2000. Each substrate (about 100 ng) was incubated with the amount of FEN-1 protein indicated in the presence of 60 mM Tris, pH7, 5% glycerol, 0.1 mg/ml BSA, 5 mM 2-mercaptoethanol, and 5 mM $MgCl_2$ at room temperature for 20 min. The reaction halted by adding an equal volume of denaturing buffer (95% formamide, 2% SDS, 10 mM EDTA, 1 mg/ml bromophenolblue, and 1 mg/ml xylene cyanol). Each stopped reaction was heated in a heating block (95°C) for 10 min, cooled down on ice and run on a 10% TBE-Urea gel. The gel was run for 2 h at 100 V, and auto-radiographed with Kodak Biomax film and analyzed on a phosphorimager (Molecular Dynamics).

Results

The biochemical study of FEN-1 protein related to triplet repeat instability is important for understanding the mechanism of human genetic diseases. Here, we characterized that FEN-1 protein activity was dependent on flap repeat- size and sequence. The endonuclease assay of FEN-1 protein was done using an M13 endonuclease assay. A schematic representation of the substrate is shown in Figure 1A. All M13 ss based substrates in this experiment had the same sequence except for the 5'-flap. The (CTG)n, (CAG)n and palindrome flap oligonucleotides were synthesized and annealed to M13 mp19 ss DNA (n equals the number of repeats, n = 11, 15, 21). The details of substrate preparation are illustrated in the Materials and Method. The bold line in Figure 1A indicates the 5'-flap region. Eleven repeats of CTG, CAG and similar size of CA and poly A of flap sequences are shown in Figure 1B.

Studies of hairpin formation by triplet repeat sequences have shown that (CNG)n triplet repeats can form hairpin structures by N:N mispairing in the stem-loop (N = the nucleotide, A, G, C, T and n = the number of repeats) (Mitas,1997). T:T mispairing (CTG repeats) can make two hydrogen bonds, which is more stable than A:A mispairing (CAG repeats) with one hydrogen bond. The CA and poly(A) sequences were synthesized as a control with no secondary structure.

FEN-1 can process the 5'-flap DNA of CTG/CAG repeat DNA depending on its size

We have selected the sizes of flap structures between 33 to 63 nucleotides (nt) (11 to 21 triplet repeats) in this experiment based on structural information and experimental limits for preparing substrates. While 11 repeats of CTG forms exclusively a hairpin, the final yield of substrate longer than 21 repeats of flap (97 nt-long, 63 nt for the flap plus 34 nt annealing section; see Figure 1A) was low. Six different M13 based substrates were prepared for the length-dependent cleavage assay shown in Figure 2. The 11, 15, and 21 CTG repeats of flap sub-strate (oligo #2 in Figure 2A) without cleavage are shown in Figure 2A (Lanes 1-3). Each substrate was cleaved with 100 ng of FEN-1 protein at the same molar ratio (by adjusting the concentration of DNA to be around 100 ng) using the M13 endonuclease assay indicated in the Materials and Method. The cleavage products are shown as one distinct band (Lanes 4-6) and the products migrated more slowly with an increasing repeat size. Each cleavage product was confirmed as corresponding the flap sizes (33, 45, 63nt) by oligonucleotide sizing markers. FEN-1 protein cleaves exactly at the junction between double and ss DNA. A similar experiment was also carried out with 5' (CAG)n repeats of flap in Figure 2B (n=11, 15, 21 repeats). The ^{32}P-labeled CAG_{11} (Lanes 1, 4), CAG_{15} (Lanes 2, 5) and CAG_{21} (Lanes 3, 6) are shown before (Lanes 1-3) and after cleavage (Lane 3-6). As is shown in Figure 2, FEN-1 protein can cleave 5'-CTG/CAG flap DNA dependent on the length of the flap. There is similar FEN-1 cleavage activity on repeat length of 5 and 11 of CTG flaps (data not shown), and so we concluded that repeat lengths

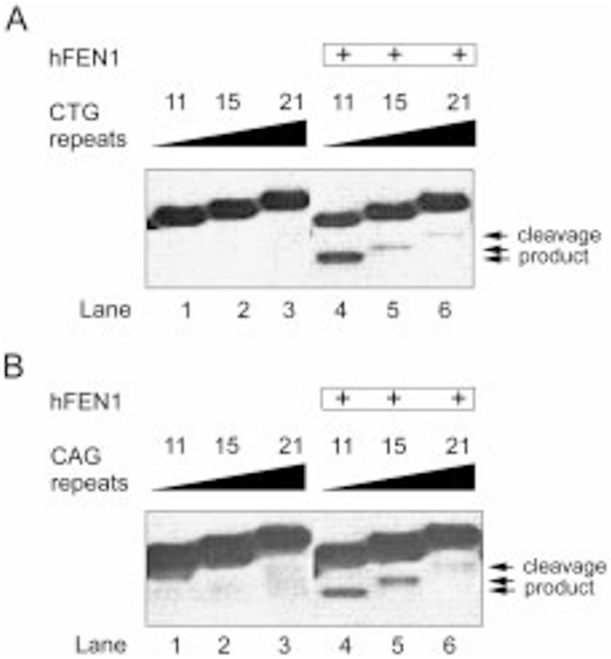

Figure 2. FEN-1 cleavage assay with flap substrates of different length. Six different flap substrates with TNR sequences were prepared. (A)^{32}P-labeled M13-based CTG_{11} (Lane 1), CTG_{15} (Lane 2) and CTG_{21} (Lane 3) repeats-flap DNAs (100 ng) are shown before and after incubation with 100 ng of FEN-1 protein (Lanes 4-6). (B) Same reaction as(A), but with CAG_{11} (Lanes 1, 4), CAG_{15} (Lanes 2, 5) and CAG_{21} (Lanes 3, 6) flap DNAs before (Lanes 1-3) and after (Lanes 4-6) the cleavage. The cleavage products are indicated by arrow.

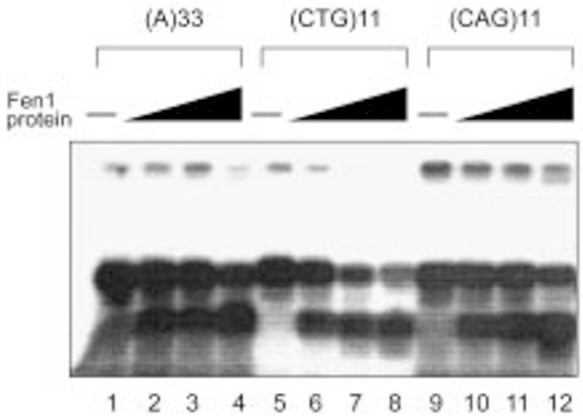

Figure 3. M13-based endonuclease assay performed using Poly$(A)_{33}$, CTG_{11} and CAG_{11} -flap DNAs with increasing amounts of protein. Three different M13 based substrates, poly$(A)_{33}$ (Lanes 1-4), $(CTG)_{11}$ (Lanes 5-8) and $(CAG)_{11}$ (Lanes 9-12) were shown before (Lanes 1, 5, 9) and after adding 100 ng (Lanes 2, 6, 10), 250 ng (Lanes 3, 7, 11) and 500 ng (Lanes 4, 8, 12) of protein.

greater than 11 are inhibitory to FEN-1 cleavage of CTG/CAG repeat flap substrates.

FEN-1 can process the 5'-flap of the CTG/CAG repeat DNA dependent on its orientation

The stability of CTG/CAG tracts has been shown to be dependent on their orientation during replication in *E.coli* and yeast. The stability can be affected by the direction of replication, and therefore is dependent on which strand forms the Okazaki fragment or lagging strand template. The CTG/CAG repeat region was more unstable when the 5' $(CAG)_n$ 3' tract is part of Okazaki fragment (Kang *et al*., 1996; Freudenreich *et al*., 1997). Therefore, we tested whether FEN-1 can process different flap sequence depending on flap structure. Indeed, by a series of experiments, we found that the 5'$(CTG)_n$-flap can be cleaved by FEN-1 more efficiently than the 5' $(CAG)_n$-flap of the same length. This result might be due to different secondary structural formation between CTG and CAG-flaps, because ss CTG repeats can form more stable hairpin than ss CAG repeats can. Therefore, we synthesized three different substrates which may have different ss structure and prepared a sequence-dependent cleavage assay in Figure 3. Three different M-13 based substrates, Poly$(A)_{33}$ (Lanes 1-4), $(CTG)_{11}$ (Lanes 5-8) and $(CAG)_{11}$ (Lanes 9-12) were shown on the figure before (Lanes 1, 5, 9) and after adding 100 ng (Lanes 2, 6, 10), 250 ng (Lanes 3, 7, 11) and 500 ng (Lanes 4, 8, 12) of protein. The cleavage products of the substrates have migrated similarly and been confirmed as 33 nt by sizing markers. The CTG repeat flap can be cleaved by FEN-1 better than the CAG repeat-flap at both protein concentration by about two fold. Interestingly, 5' polyadenosine flap $(A)_n$, n = 33, which has not secondary structure in flap was cleaved by FEN-1 similar to CAG_{11} DNA. These results suggest that FEN-1 can process the 5'- flap DNA dependent on their structures. Therefore, the different cleavage rate between CTG and CAG flap DNAs might come from the stability of the different secondary structures of the flaps. This suggests that FEN-1 may play a role in orientation- dependent CTG/CAG repeat instability.

Discussion

The biochemical studies of human FEN-1 related to CTG/CAG repeat instability was done here. We carried out FEN-1 endonuclease assay using an M-13 based substrates. This method is appropriate for preparing various length- and sequence- flap substrates, and may give high specificity because of substrate itself containing large amount of ss DNA per flap. The results from M13 method were also well correlated with the data from other established methods, such as flow cytometry. We characterized FEN-1 activity dependent on the size and sequence of the 5'-flap DNA on CTG/CAG repeats. The longer 5'-flap DNA on CTG/CAG repeat was inhibitory, and CTG-flap DNA is better substrate than CAG-flap substrate for FEN-1 cleavage assay. These results can explain the details of triplet repeat instability in human genetic diseases, specially for the disease causing mechanism (expansion side) based on unusual triplet repeat DNA structures and yeast genetic data.

The expansion of CTG/CAG repeat loci within the human genome might be made by the slippage, realignment or strand displacement between Okazaki fragments. The unannealed 5'-flap DNA on triplet repeat can be expanded after replication if by-passed the FEN-1 repair process. The CTG/CAG repeat numbers of flap for FEN-1 cleavage assay on this results are the range of normal individuals (~37 repeats in the loci of DM and Huntington's diseases). The frequency of hairpin formation at TNR loci would be highly dependent on its size. The patients with long repeats (up to thousands of CTG/CAG repeats in case of DM) have a great risk to have longer 5'-flap structure during lagging strand synthesis on triplet repeat locus, which have shown to be inhibitory to FEN-1 repair in our results. In the case of DM, CAG expansion starts when the number of repeats reach 50 repeats, while the normal range is 5-37 repeats.

Yeast studies have shown that the direction of replication is important for the stability of CTG/CAG tract. Human FEN-1 and yeast RAD27 are highly conserved, and Rad27 protein has exactly same activity as human FEN-1 protein by this assay. Our biochemical study of FEN-1 also suggests that the replication direction may be important, and there is a two fold increase in the risk of expansion when 5'CAG3' tracts are on the Okazaki fragment compared to the sequence

orientated in the reverse direction. However, this explanation is based on the FEN-1s activity processing with different 5' flap substrates. The mechanism of TNR expansion may be more complicated *in vivo.* Since several proteins can conjugate with FEN-1 protein, which may effect its activity. The study of replication origin near triplet repeat loci in human genome may give important insights into the mechanism of triplet repeat instability.

Acknowledgements

This work was supported by Korean Research Foundation (KRF-2000-015-DP0418)

Reference

Bae SH, Seo YS. Characterization of the enzymatic properties of the yeast dna2 Helicase/endonuclease suggests a new model for Okazaki fragment processing. J Biol Chem 2000; 275;38022-31

Bornarth CJ, Ranalli TA, Henricksen LA, Wahl AF, Bambara RA. Effect of flap modifications on human FEN1 cleavage. Biochemistry 1999;38;13347-54

Freudenreich CH, Kantrow SM, Zakian VA. Expansion and length-dependent fragility of CTG repeats in yeast. Science 1998;279;853-6

Freudenreich CH, Stavenhagen JB, Zakian VA. Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. Mol Cell Biol 1997;17;2090-8

Gordenin DA, Kunkel TA, Resnick MA. Repeat expansion--all in a flap? Nat Genet 1997;16;116-8

Kang S, Ohshima K, Jaworski A, Wells RD. CTG triplet repeats from the myotonic dystrophy gene are expanded in Escherichia coli distal to the replication origin as a single large event. J Mol Biol 1996;258;543-7

Kim JW, Cho EH, Kim YM, Kim JM, Han JY, Park SY. Detection of cryptic Y chromosome mosaicism by coamplification PCR with archived cytogenetic slides of suspected Turner syndrome. Exp Mol Med 2000;32;38-41

Lee SH, Kim CJ, Park HK, Koh JW, Cho MH, Baek MJ, Lee MS. Characterization of aberrant FHIT transcripts in gastric adenocarcinomas. Exp Mol Med 2001;33;124-30

Li X, Li J, Harrington J, Lieber MR, Burgers PM. Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. J Biol Chem 1995;270;22109-12

Maurer DJ, O'Callaghan BL, Livingston DM. Orientation dependence of trinucleotide CAG repeat instability in Saccharomyces cerevisiae. Mol Cell Biol 1996;16;6617-22

McMurray CT. DNA secondary structure: a common and causative factor for expansion in human disease. Proc Natl Acad Sci USA 1999;96;1823-5

Mitas M. Trinucleotide repeats associated with human disease. Nucleic Acids Res 1997;25;2245-54

Nolan JP, Shen B, Park MS, Sklar LA. Kinetic analysis of human flap endonuclease-1 by flow cytometry. Biochemistry 1996;35;11668-76

Oostra BA, Willems PJ. A fragile gene. Bioessays 1995;17; 941-7

Sambrook J, Fritsch E, Maniatis T, *Molecular cloning*: *A Laboratory Manual*, 2nd Ed., 1989, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Warren ST, Nelson DL. Trinucleotide repeat expansions in neurological disease. Curr Opin Neurobiol 1993;3;752-9