Mismatched nucleotides may facilitate expansion of trinucleotide repeats in genetic diseases

Mikiya Nakayabu, Shigeru Miwa, Motoshi Suzuki¹, Shunji Izuta², Gen Sobue and Shonen Yoshida¹,*

Department of Neurology and ¹Laboratory of Cancer Cell Biology, Institute of Disease Mechanism and Control, Nagoya University School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan and ²Department of Biological Science, Faculty of Science, Kumamoto University, Kumamoto 860-0862, Japan

ABSTRACT

We have studied the contribution of mismatch sequences to the trinucleotide repeat expansion that causes hereditary diseases. Using an oligonucleotide duplex, (CAG)₅/(CTG)₅, as a template–primer, DNA synthesis was carried out using either Escherichia coli DNA polymerase I (Klenow fragment) or human immunodeficiency virus type I reverse transcriptase (HIV-RT). Both enzymes expanded the repeat sequence longer than 27 nucleotides (nt), beyond the maximum length expected from the template size. The expansion was observed under conditions in which extension occurs either in both strands or in one strand. In contrast, with another template–primer that contains a non-repetitive flanking sequence 5′-upstream of the repetitive sequence, the reaction products were not extended beyond the template size (45 nt) by these DNA polymerases. We then used mismatched template–primers, in which either 1, 2 or 6 non-complementary nucleotides were introduced to the repeat sequence that is flanked by a non-repetitive sequence. In this case, primers were efficiently expanded over the expected length of 45 nt, in a mismatch-dependent manner. One of the primers with six mismatches extended as long as 72 nt. These results imply that the misincorporation of non-complementary deoxyribonucleoside monophosphates (dNMPs) into the repeat sequence makes double-stranded DNA unstable and triggers the slippage and expansion of trinucleotide repeats by forming loops or hairpin structures during DNA synthesis.

INTRODUCTION

Triplet repeat expansions are associated with various human genetic diseases (1). The expansion of CAG/CTG is found in Huntington’s disease, spinal and bulbar muscular atrophy, spinocerebellar ataxia types 1, 2, 6 and 7, dentatorubral-pallidoluysian atrophy, Machado–Joseph disease in the coding regions and myotonic dystrophy in the non-coding regions (1–6). The in vitron vitro expansion of this repeat, as well as of other types of triplet or dinucleotide repeats, has been demonstrated with bacterial and eukaryotic DNA polymerases (7–9) in either the polymerase chain reaction (8) or a single round of DNA synthesis (7,9). These results were obtained using the template–primer duplexes consisting of short repeats, which might hybridize each other at variable positions. Using the trinucleotide repeats flanked by the unique sequence, however, these in vitro expansions could not be observed (9). In the present study, we have demonstrated that mismatched base-pairs in the CAG/CTG repeat greatly enhanced the expansion of the repeat, even if the repeat was flanked by unique sequence. These results suggest that misincorporation of mismatched deoxyribonucleoside monophosphates (dNMPs) into the repeat contributes to the progression of this unusual entity of genetic changes.

MATERIALS AND METHODS

Materials

Deoxyribonucleoside triphosphates (dNTPs) were purchased from Yamasna Shoyu (Chiba, Japan), radioactive compounds from NEN (MA, USA) or Amersham (Buckinghamshire, UK), large fragment of Escherichia coli DNA polymerase I (Klenow fragment) from Takara Shuzo (Kyoto, Japan) and human immunodeficiency virus type I reverse transcriptase (HIV-RT) from Seikagaku Kogyo (Tokyo, Japan). Enzyme units are defined by companies.

Synthetic DNA

The DNA oligomers were synthesized by BIOSYNTHESIS, INC. (TX, USA), and their sequences were summarized in Table 1.

Annealing of DNA

Oligomers (CAG)₅ and (AAG)₅ were incubated with their complementary counterparts, (CTG)₅ and (CTT)₅, respectively, at a molar ratio of 10:1 at 70°C for 15 min, then cooled slowly. Oligomers containing trinucleotide repeats and a flanking sequence were annealed with their complementary primers at the molar ratio of 1:1, 10:1 or 1:10.

*To whom correspondence should be addressed. Tel: +81 52 744 2453; Fax: +81 52 744 2457; Email: syoshida@tsuru.med.nagoya-u.ac.jp
DNA synthesis reaction

The reaction mixture (25 µl) for Klenow fragment contained 50 mM Tris–HCl (pH 8.3), 3 mM dithiothreitol, 10 mM MgCl₂, 50 mM KCl, 80 µM of dCTP, dATP and dGTP, 40 µM of dTTP, 5 µCi [α-32P]dTTP (111 kBq/pmol) and 3.6 µg/ml (CAG)₅/(CTG)₅ (complete conditions). Enzyme quantities were varied in experiments as indicated in the figure legends. After incubation for 90 min at 37°C, 5 µl of 20 mM EDTA in 1% SDS was added to the reaction mixture, followed by incubation at 37°C for 10 min. Samples were analyzed by 12% polyacrylamide gel electrophoresis containing 8 M urea.

RESULTS

Enzymatic expansion of trinucleotide repeat sequence

We chose two trinucleotide repeat sequences, CAG/CTG and AAG/CTT. The former is a model, in which template–primer could form a hairpin in expansion. This repeat is found in several hereditary neuro-degenerative diseases (1). The latter sequence suggests the involvement of other factors for expansion of the repeat with flanking sequences (1,13). Using this template–primer, we could not observe the expansions of the repeat in reactions using combinations of DNA polymerases and repeat sequences (7–9). Consistently, we also observed the expansion using purified bovine DNA polymerase α and β, though with lower efficiency than Klenow fragment or HIV-RT (data not shown).

Enzymatic expansion of trinucleotide repeats flanked by non-repetitive sequence

In chromosomal DNA, the trinucleotide repeats are flanked by non-repetitive sequences at both ends. The flanking sequences would be against the local stress that disrupts the double-stranded nature of DNA produced by a slippage of short repeats. Slippage was observed using simple repeat duplexes (Fig. 1A and B). To discover the effects of slippage expansion, we have prepared another template–primer, which consists of (CAG)₅/(CTG)₅ flanked by 15 nt of non-repetitive sequence just upstream of the CTG repeat in the gene for the spinal and bulbar muscular atrophy (1,13). Using this template–primer, we could not observe the reaction products longer than 45 nt in length, corresponding to the full length of the template strand (Fig. 3A). Both Klenow fragment and HIV-RT failed in expansion with the template–primers that were annealed at the molar ratios of 1:1, 10:1 and 1:10 (Fig. 3B).
Figure 2. Schematic illustration of template–primers carrying mismatches. Open and closed boxes represent non-repetitive sequences in a primer and a template, respectively. Repeat sequence and mismatches in repeats are illustrated as lines and notches, respectively. Names of each oligomer on the left side correspond with those in Table 1. Since the SB(m1)45 template has a base substitution, template–primers consisting of SB(m1)45–SB(m2)33, SB(m1)45–SB(m1)33 and SB(m1)45–SB(m5)33 would have one, two and six mismatch bubbles, respectively.

The mutated template consists of the flanking sequence and the CAG repeat carrying a substituted nucleotide. A new set of mutated primers consists of the flanking sequence and the CTG repeats carrying 1–5 nt non-complementary to the prototype repeat. When these primers were annealed with the template without forming a hairpin or loop, the number of mismatches was 1, 2 and 6, as illustrated in Figure 2. Using template–primers that have one or two mismatches, HIV-RT was capable of expanding the primer up to 51 nt in length, which is 6 nt longer than the template size (45 nt) (Fig. 3C). Longer expansion was observed using a primer that has six mismatches. In this case, both enzymes extended the primers to ∼72 nt in length (Fig. 3C), which is much longer than the predicted maximal length (63 nt).

DISCUSSION

Expansion of the trinucleotide repeat sequence may not be explained simply, since it may involve an unequal crossing over between repeats, gene conversions or the misalignment of the DNA strand during DNA replication or repair (14,15). Here we have focused on the replication of the trinucleotide repeat sequence. Although DNA synthesis is remarkably accurate, it has been shown that the duplexes of simple repeat sequences are

Figure 1. Expansion of simple triplet repeat duplex by DNA polymerases in vitro. Simple triplet repeat duplex (CAG)3/(CTG)3 was used in lanes 1–4, and (AAG)3/(CTT)3 was used in lanes 5 and 6 as template–primers in (A) and (B). (A) 0.1 U of Klenow fragment was used in lanes 1, 3 and 5, and 0.8 U in lanes 2, 4 and 6. Reactions were performed with three dNTPs (dCTP, dGTP and dTTP) in lanes 1 and 2, and under complete conditions with all four dNTPs in lanes 3–6, as indicated in the figure and described in Materials and Methods. The 32P-labeled products were analyzed by polyacrylamide gel electrophoresis and exposed on X-ray film. Sizes of the reaction products were estimated by the sequencing ladder of M13mp18, indicated on the right side of each lane. Unless slippage occurs, reaction products migrate around the arrow position of 27 nt in length or shorter. (B) 0.32 U of HIV-RT was used in lanes 1, 3 and 5, and 2.6 U in lanes 2, 4 and 6. Reaction conditions were as described in Materials and Methods, and other conditions were the same as described in (A) for each lane number.
Figure 3. Expansion of the CAG/CTG repeat, which is flanked with a non-repetitive sequence. The sequences and combinations of template–primers are shown in Table 1 and schematically illustrated in Figure 2. (A) The reaction with a template–primer with a flanking sequence, which has no mismatch, was performed using the prototype SB45/SB15 as a template–primer in lanes 1, 2, 4 and 5. SB45 was annealed with SB15 at a molar ratio of 1:1. Samples of 0.32 and 2.6 U HIV-RT were used in lanes 1 and 2, respectively. Samples of 0.1 and 0.8 U were used in lanes 4 and 5, respectively. In lanes 3 and 6, an oligonucleotide of 45 nt, having the same sequence as reaction product (SB product), labeled at the 5′-end by [γ-32P]ATP and T-4 polynucleotide kinase, was loaded as a size marker. The position of SB product is indicated by an arrow. Sizes of the reaction products were also estimated by the sequencing ladder of M13mp18, as indicated on the right side of the lanes. (B) The reaction was performed under the same conditions as in (A), except that the ratios between template (SB45) and primer (SB15) were changed to 10:1 (lanes 2 and 5) and 1:10 (lanes 3 and 6), in comparison with the results using the template–primer ratio of 1:1 (lanes 1 and 4). Samples of 2.6 U HIV-RT were used in lanes 1–3, while 0.8 U of Klenow fragment was used in lanes 4–6. (C) The reactions using template–primers with flanking sequences, which have mismatches in repeat sequence, were performed using SB(m.1)45–SB(m.2)33, having one mismatch, in lanes 2–5; SB(m.1)45–SB(m.1)33, having two mismatches, in lanes 6–9; and SB(m.1)45–SB(m.5)45, having six mismatches, in lanes 10–13. Escherichia coli DNA polymerase I Klenow fragment (0.1 U) was used in lanes 3, 7 and 11; HIV-RT (0.32 U) was used in lanes 5, 9 and 13. Lanes 2, 4, 6, 8, 10, 12 and 14 contained no enzyme. In lanes 1 and 14, 32P-labeled SB product (45 nt) was loaded. The position of SB product is indicated by an arrow. Sizes of the reaction products was estimated by the sequencing ladder of M13mp18, as indicated on the right side of the lanes.

In normal individuals, the mean number of repeated units, as targets of hereditary diseases, ranges from 6 to 52 repeats (mean of ~20) (1) and they are flanked by non-repetitive sequences. Under normal conditions, the flanking sequence at the 5′-end of the primer would not tolerate the slippage backward of primer strands during replication (Fig. 3A). Therefore, for the expansion of trinucleotide repeat sequence during replication, some sort of destabilization of the duplex may be required. A defect in the complementality caused by any type of misincorporation might be a candidate for this destabilization. In fact, repeated sequences in the genome are variable in individuals (16). Eichler et al. (17) have reported that the loss of AGG within the CGG trinucleotide repeat is an important mutational event in the generation of unstable alleles predisposed to the Fragile X syndrome. Loss of the CAT within the CAG repeat leads to CAG repeat instability in spinocerebellar ataxia type 1 (18,19). In Huntington’s disease, loss of the CAA and CCA interruptions also influences the probability of expansion (20).
On the other hand, the expansion or deletion of CTG repeats may be strand-specific, depending on the direction of DNA replication in E.coli (21), as was observed with a secondary structure-dependent mutagenesis (22). As for strand specificity, Izuta et al. (23) showed that the replication of the lagging strand causes a higher frequency of misincorporation than the leading strand, by a human cell-free DNA replication system. These mutations might be best explained by a frequent dissociation and reassociation of DNA polymerase to the primer end (7), which has been hypothesized to be a cause of base substitution mutation and of frameshift mutation (24–27).

The mismatch repair system corrects the mismatched base pairs produced by misincorporation during replication (15,28). In this sense, it may prevent the trinucleotide repeat expansion. Furthermore, the fixation of the looping resulting from slippage DNA synthesis of the CTG repeat, and its transfer in generations, may also be closely related to the mismatch repair system (29). Moreover, Umar et al. (30) have reported that the mismatch repair enzyme of human cells can repair DNA loops, and have proposed the importance of the defect in the loop repair for the trinucleotide repeat instability. Our attempt to introduce non-complementary nucleotides into the repeat sequence is based on these observations.

The mismatched primers would mimic the ones that were raised by any error-prone DNA synthesis or recombination. Both Klenow fragment and HIV-RT expanded these primers in a mismatch-dependent manner (Fig. 3C). One mismatch could make a bubble of 3 nt that loops out from double-stranded DNA. In this context, a primer that has a mismatch could be extended as 48 nt in length. Two mismatches could make 51 nt, and six could make 63 nt.

One of the primers that has a single mismatch was extended as long as 45, 48 and 51 nt by HIV-RT. Both enzymes expanded a primer that has six mismatches as long as 72 nt, far beyond the template size (45 nt). This is even longer than the size that is expected by forming the largest loop of primer strand. This unexpected length of expansion may be explained as mismatches facilitating a slippage of primer strand and causing the expansion by a loop formation.

Our results imply that a short stretch of repeat flanked by the unique sequence can be extended as a consequence of misincorporation in DNA synthesis. In other words, short repeat sequences might be maintained unchanged for generations unless the rare event of misincorporation disturbs the double-stranded nature of repeat DNA. In the normal replication, however, it is unlikely that numbers of misincorporation occur in the single round replication of repeat. Therefore, other unidentified factor(s) would destabilize the DNA duplex to induce local denaturation at the site of misincorporation, which triggers the slippage synthesis in the target sequence.

ACKNOWLEDGEMENTS

We are grateful to Drs M.Takemura and K.Furuta for supplying purified mammalian DNA polymerases, to Dr K.Tamiya-Koizumi of our laboratory for discussion and Ms T.Tomita and M.Takahashi for their excellent technical assistance. This work is supported by the Grant-in-Aid for Scientific Research from Ministry of Education, Science and Culture of Japan.

REFERENCES