Journal of Bacteriology, Aug. 1996, p. 5042–5044
0021-9193/96/$04.00+0
Copyright © 1996, American Society for Microbiology
Vol. 178, No. 16

Single-Stranded DNA-Binding Protein Enhances the Stability of CTG Triplet Repeats in *Escherichia coli*

WILLIAM A. ROSCHE,[1] ADAM JAWORSKI,[1]† SEONGMAN KANG,[1]‡ SHANNON F. KRAMER,[1] JACQUELYNN E. LARSON,[1] DAVID P. GEIDROC,[2] ROBERT D. WELLS,[1] AND RICHARD R. SINDEN[1]*

Center for Genome Research, Institute of Biosciences and Technology, Department of Biochemistry and Biophysics, Texas A&M University, Texas Medical Center, Houston, Texas 77030-3303,[1] and Center for Macromolecular Design, Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843-2128[2]

Received 5 February 1996/Accepted 10 June 1996

The stability of CTG triplet repeats was analyzed in *Escherichia coli* to identify processes responsible for their genetic instability. Using a biochemical assay for stability, we show that the absence of single-stranded-DNA-binding protein leads to an increase in the frequency of large deletions within the triplet repeats.

The molecular etiology of at least nine human diseases including myotonic dystrophy, Huntington's disease, and fragile X syndrome has been linked to the expansion of CTG and CGG triplet repeats (23). These diseases exhibit the non-Mendelian inheritance pattern known as anticipation, in which the penetrance of the disease increases with each generation. Anticipation results from a progressive increase in the number of triplet repeats in successive generations. The cause of the expansion is not known, but it may be due to replication slippage (19).

Simple microsatellite sequences with one to four repeat units exhibit genetic instability (3, 12, 22). In bacteria and *Saccharomyces cerevisiae*, mismatch repair pathways are involved in the repair of mismatches and short loops caused by misalignment during DNA replication. In *S. cerevisiae*, mutations in *msh2* or *pms1* result in a decreased stability of dinucleotide repeats (22). The identification of deficiencies in human homologs of the *Escherichia coli mutS* and *mutL* mismatch repair genes in hereditary nonpolyposis colorectal and other cancers (4, 11) suggests that replication slippage at repeats can lead to cancer if uncorrected. We have shown previously that mutations in *mutH*, *mutL*, and *mutS* reduce the frequency of large deletions of long CTG and CGG triplet repeats in *E. coli* (8).

Single-stranded-DNA-binding protein (SSB) is an important component in DNA replication, repair, and recombination (13) since SSB can prevent the formation of DNA secondary structures. Replication can pause at sites of potential DNA secondary structure (7, 18, 24), and pause sites are associated with template misalignment mutagenesis (1, 16). The presence of SSB can prevent replication pausing at DNA secondary structures (7, 17). The potential for slippage and for DNA secondary structure formation within CTG and CGG triplet repeats (5, 21, 26) may contribute to the instability of these sequences observed in individuals with triplet repeat diseases (8, 9, 19, 25). Herein, we used a biochemical assay to determine the extent of deletions of triplet repeats cloned into plasmids in an *E. coli* strain containing a mutant SSB.

A biochemical assay for repeat stability. Plasmids pRW3244, pRW3246, and pRW3248 containing a human myotonic dystrophy-derived fragment with $(CTG)_{17}$, $(CTG)_{100}$, and $(CTG)_{180}$, respectively (8, 9), were grown in *E. coli* RM121, a derivative of strain C600, which contains a temperature-sensitive mutation (*ssb-1*) in the gene encoding SSB. A temperature of 43°C is required to completely inactivate SSB-1 and prevent growth of RM121 (14). The doubling time of RM121 in log phase is about 45 min at 30 to 32°C. At 41.5 to 42°C, the doubling time is increased concomitant with a loss of functional SSB. The introduction into *E. coli* of a plasmid containing a homogeneous-length $(CTG)_n$ triplet repeat for a population stability analysis has been described previously (8). Following transformation and initial growth at 32°C, the culture was grown in parallel at 32 and 42°C to an optical density at 650 nm of ca. 0.9 to 1.0 for the number of generations denoted. The biochemical assay of triplet repeat instability involving analysis of the size of plasmid DNA by agarose gel electrophoresis was described previously (8).

Length-dependent instability of $(CTG)_n$ in *ssb-1 E. coli*. Long $(CTG)_n$ triplet repeats are relatively unstable when cloned near the ColE1 origin in pUC-based plasmids and undergo deletions more frequently than expansions (8, 9). In this plasmid background with a unidirectional mode of replication, triplet repeat stability is orientation dependent with the greatest frequency of deletion observed when the CTG sequence occurs in the lagging strand (called orientation II) (9). Figure 1 shows results of a biochemical assay of $(CTG)_n$ triplet repeat instability for the stable orientation (orientation I) for different $(CTG)_n$ repeat lengths, where $n = 17$, 100, and 180. Following transformation of RM121 (*ssb-1*) with gel-purified plasmid containing a homogeneous length of triplet repeats, samples were removed for analysis from a population of transformants grown for 30 to 90 generations at either 32 or 42°C. Little detectable difference in the amount of nondeleted $(CTG)_{17}$-containing plasmid was observed at either temperature. However, for plasmid containing $(CTG)_{100}$ or $(CTG)_{180}$, the fraction of plasmids that contained the original number of CTG repeats decreased to about 60 and 50%, respectively, after 90

* Corresponding author. Mailing address: Center for Genome Research, Institute of Biosciences and Technology, Department of Biochemistry and Biophysics, Texas A&M University, Texas Medical Center, 2121 W. Holcombe Blvd., Houston, TX 77030-3303.

† Present address: Department of Genetics of Microorganisms, University of Lodz, Banacha 12/16, and Center for Microbiology and Virology, Polish Academy of Sciences, Lodz, Lodowa 106, Poland.

‡ Present address: Laboratory of Genetic Disease Research, National Center of Human Genomic Research, National Institutes of Health, Bethesda, MD 20892.

Downloaded from http://jb.asm.org/ on September 19, 2018 by guest

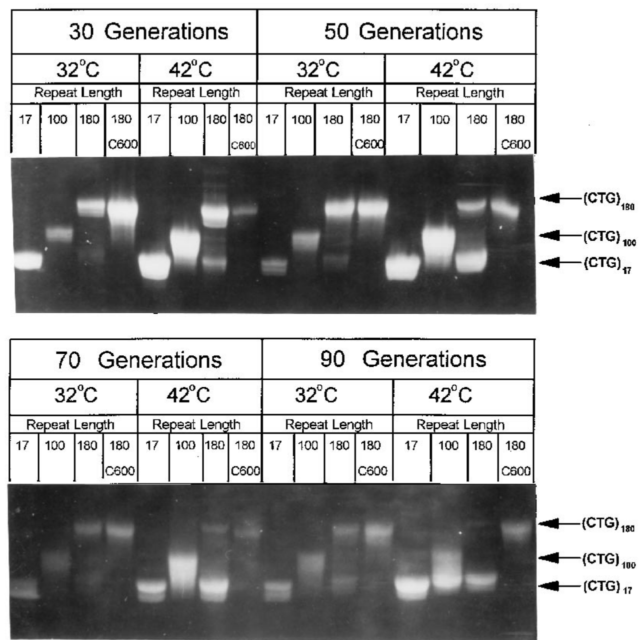

FIG. 1. Increased stability of long (CTG) triplet repeats in vivo with functional SSB. Cells containing plasmids with various-length (CTG) triplet repeats were grown at 32 or 42°C, as indicated, for the specified number of generations. The DNA was then isolated, and the plasmid was analyzed on agarose gels as described previously (8). Except for lanes labeled C600, all DNA was isolated from strain RM121 containing the *ssb-1* mutation. Only the regions of the gel containing the supercoiled plasmid bands are shown. The positions of migration of plasmids containing $(CTG)_{17}$, $(CTG)_{100}$, and $(CTG)_{180}$ are indicated by arrows on the right. Deletions are indicated by the DNA migrating faster than plasmid containing the full-length triplet repeat (typically a smear in longer repeats at greater numbers of generations).

generations at 32°C, similar to that observed for these plasmids in other ssb^{+} strains at 37°C (8). In contrast, there was a dramatic decrease in stability at 42°C, at which the mutant SSB protein is less active. The $(CTG)_{180}$ repeat was rapidly deleted to about $(CTG)_{20}$, and by 70 generations less than 10% of the plasmid molecules appeared to contain nondeleted triplet repeats. When plasmid containing $(CTG)_{180}$ was grown in C600, the parental strain of RM121, the stability of the triplet repeat sequence at both temperatures was greater than that in RM121 at 32°C. Quantitative analyses of agarose gels for the results shown in Fig. 1, as well as for those from other experiments, are shown in Table 1. These results demonstrate that growth at 42°C was not responsible for the instability (Fig. 1, lane 4). Moreover, the greater instability of $(CTG)_{180}$ in RM121 at 32°C is consistent with a partial deficiency in SSB-1 activity at

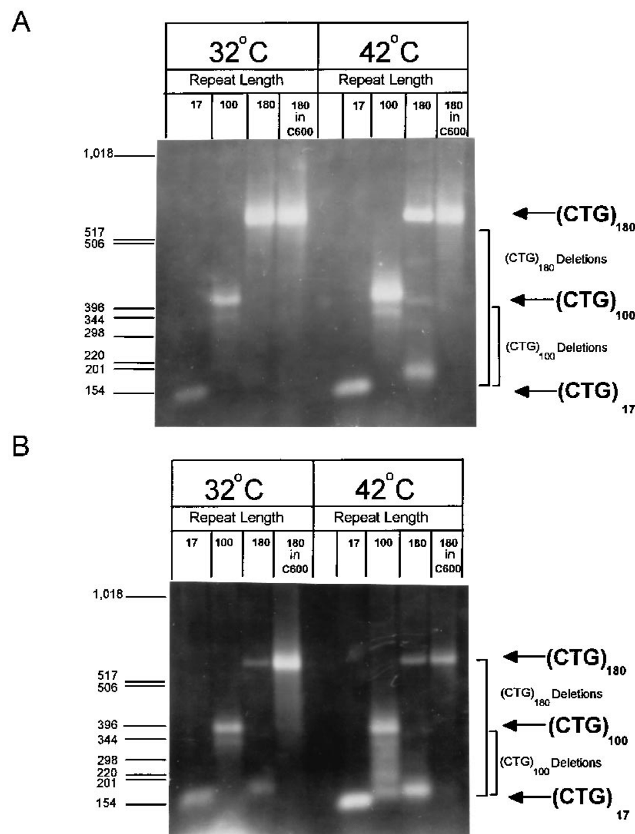

FIG. 2. Deletions in plasmids containing (CTG) triplet repeats are within the repeats. (A) DNA isolated after 30 generations. (B) DNA isolated after 50 generations. In both, cells containing plasmids with various-length (CTG) triplet repeats were grown, and DNA was isolated, digested with *Sac*I and *Hin*dIII, and then analyzed on acrylamide gels (8). Except for lanes labeled C600, all DNA was isolated from strain RM121. The positions of the *Sac*I-*Hin*dIII fragments containing the full-length triplet repeat are indicated by arrows on the right. The positions of migration of deletion products are indicated by brackets on the right. The positions of molecular size markers (in base pairs) are indicated on the left.

this temperature. Purified SSB-1 is only partially active at 30°C (15). Moreover, partial activity at 37°C is suggested by increased sensitivity to UV light and other DNA damage and by a decreased recombination proficiency of *ssb-1* cells (6).

To demonstrate that the deletions were occurring within the triplet repeats in these plasmids, *Sac*I-*Hin*dIII digestion products were analyzed (Fig. 2). As described previously, *Sac*I cuts 64 bp 5′ and *Hin*dIII cuts 44 bp 3′ to the $(CTG)_n$ triplet repeat (8). While no changes were observed for the $(CTG)_{17}$-contain-

TABLE 1. Quantitation of $(CTG)_n$ repeat stability in *ssb-1*

Plasmid	No. of CTG repeats	Strain	% Undeleted plasmid[a]					
			40 generations		60 generations		80 generations	
			32°C	42°C	32°C	42°C	32°C	42°C
pRW3244	17	RM121	>95	>95	>95	>95	>95	>95
pRW3246[b]	100	RM121	100	98	85	46	70	33
pRW3248	180	RM121	87 ± 8	25 ± 4	80 ± 6	13 ± 1	75 ± 11	7 ± 1
pRW3248	180	C600	87 ± 4	82 ± 9	96 ± 3	83 ± 12	90 ± 3	80 ± 2

[a] Values are the means of three experiments ± the standard deviations. The value is the percentage of undeleted plasmid as determined by densitometric analysis of agarose gels.

[b] pRW3246 was quantitated once.

Downloaded from http://jb.asm.org/ on September 19, 2018 by guest

ing plasmid, a major product representing deletion of about 40 to 50 bp (13 to 17 repeats) in $(CTG)_{100}$ and deletion of 450 to 480 bp (150 to 160 repeats) in $(CTG)_{180}$ accumulated at 32°C. After 90 generations at 42°C, the $(CTG)_{180}$ sequence was deleted to about 20 repeats, which has been observed as a deletion endpoint in most strains studied to date (8, 9). Thus in all cases, the decrease in plasmid size can be attributed to deletions within the triplet repeats; no deletions were found in the vectors.

These results demonstrate that functional SSB is important in maintaining the genetic stability of triplet repeats in *E. coli*. In summary, in *E. coli* containing a temperature-sensitive mutation in SSB, long $(CTG)_n$ triplet repeats showed a decreased stability at 42°C compared with cells grown at 32°C or in ssb^+ cells at 42°C. The formation of hairpin structures within the (CTG) or (CAG) repeats (21, 26) in a single strand may lead to the genetic instability of these repeats (8, 9, 19). The formation of DNA secondary structures can lead to increased mutagenesis, presumably by increasing the error frequency during DNA polymerization (10, 19, 20). Hairpin structures are known to pause DNA polymerases (7, 24), and in general, pause sites are associated with hot spots for mutation (1, 2, 16). SSB may stabilize triplet repeats by preventing the formation of DNA secondary structure. The formation of hairpin structures, or other DNA secondary structures, may be an important factor contributing to the instability of triplet repeats associated with certain human hereditary diseases.

This work was supported by grants from the National Institutes of Health (ES 05508 to R.R.S. and GM 52982 to R.D.W.), the National Science Foundation (DMB-9103942 to R.D.W.), and the Robert A. Welch Foundation (to R.D.W.).

REFERENCES

1. **Bebenek, K., J. Abbotts, J. D. Roberts, S. H. Wilson, and T. A. Kunkel.** 1989. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J. Biol. Chem. **264:**16948–16956.
2. **Bierne, H., S. D. Ehrlich, and B. Michel.** 1991. The replication termination signal terB of the Escherichia coli chromosome is a deletion hot spot. EMBO J. **10:**2699–2705.
3. **Edwards, A., H. A. Hammond, L. Jin, C. T. Caskey, and R. Chakraborty.** 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics **12:**241–253.
4. **Fishel, R., M. K. Lescoe, M. R. S. Rao, N. G. Copeland, N. A. Jenkins, J. Garber, M. Kane, and R. Kolodner.** 1993. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell **75:**1027–1038.
5. **Gacy, A. M., G. Goellner, N. Juranic, S. Macura, and C. T. McMurry.** 1995. Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell **81:**533–540.
6. **Glassberg, J., R. R. Meyer, and A. Kornberg.** 1979. Mutant single-stranded binding protein of *Escherichia coli*: genetic and physiological characterization. J. Bacteriol. **140:**14–19.
7. **Hacker, K. J., and B. M. Alberts.** 1994. The rapid dissociation of the T4 DNA polymerase holoenzyme when stopped by a DNA hairpin helix. A model for polymerase release following the termination of each Okazaki fragment. J. Biol. Chem. **269:**24221–24228.
8. **Jaworski, A., W. A. Rosche, R. Gellibolian, S. Kang, M. Shimizu, R. P. Bowater, R. R. Sinden, and R. D. Wells.** 1995. Mismatch repair in E. coli enhances instability of (CTG)n triplet repeats from human hereditary diseases. Proc. Natl. Acad. Sci. USA **92:**11019–11023.
9. **Kang, S., A. Jaworski, K. Ohshima, and R. D. Wells.** 1995. Expansion and deletion of CTG triplet repeats from human disease genes are determined by the direction of replication in E. coli. Nature (London) Genet. **10:**213–218.
10. **Leach, D. R. F.** 1994. Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. BioEssays **16:**893–900.
11. **Leach, F. S., N. C. Nicolaides, N. Papadopoulos, B. Liu, J. Jen, R. Parsons, P. Peltomaki, P. Sistonen, L. A. Aaltonen, M. Nystrom-Lahti, X.-Y. Guan, G. J. Zhang, P. S. Meltzer, J.-W. Yu, F.-T. Kao, D. A. Chen, K. M. Cerosaletti, R. E. K. Fournier, S. Todd, T. Lewis, R. J. Leach, S. L. Naylor, J. Weissenback, J. P. Mecklin, H. Jarvinen, G. M. Petersen, S. R. Hamilton, J. Green, J. Jass, P. Watson, H. T. Lynch, J. M. Trent, A. de la Chapelle, K. W. Kinzler, and B. Vogelstein.** 1993. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell **75:**1215–1225.
12. **Levinson, G., and G. A. Gutman.** 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. **4:**203–221.
13. **Lohman, T. M., and M. E. Ferrari.** 1994. Escherichia coli single-stranded DNA binding protein: multiple DNA binding models and cooperativities. Annu. Rev. Biochem. **63:**527–570.
14. **Meyer, R. R., J. Glassberg, and A. Kornberg.** 1979. An Escherichia coli mutant defective in single-strand binding protein is defective in DNA replication. Proc. Natl. Acad. Sci. USA **76:**1702–1705.
15. **Meyer, R. R., J. Glassberg, J. V. Scott, and A. Kornberg.** 1980. A temperature sensitive single-strand DNA-binding protein from Escherichia coli. J. Biol. Chem. **225:**2897–2901.
16. **Papanicolaou, C., and L. S. Ripley.** 1991. An in vitro approach to identifying specificity determinants of mutagenesis mediated by DNA misalignments. J. Mol. Biol. **221:**805–821.
17. **Samadashwily, G. M., and S. M. Mirkin.** 1994. Trapping DNA polymerases using triplex-forming oligodeoxyribonucleotides. Gene **149:**127–136.
18. **Sinden, R. R.** 1994. DNA structure and function. Academic Press, San Diego, Calif.
19. **Sinden, R. R., and R. D. Wells.** 1992. DNA structure, mutations, and human genetic disease. Curr. Opin. Biotechnol. **3:**612–622.
20. **Sinden, R. R., G. Zheng, R. G. Brankamp, and K. N. Allen.** 1991. On the deletion of inverted repeated DNA in Escherichia coli: effects of length, thermal stability, and cruciform formation in vivo. Genetics **129:**991–1005.
21. **Smith, G. K., J. Jie, G. E. Fox, and X. Gao.** 1995. DNA CTG triplet repeats involved in dynamic mutations of neurologically related gene sequences form stable duplexes. Nucleic Acids Res. **23:**4303–4311.
22. **Strand, M., T. A. Prolla, R. M. Liskay, and T. D. Petes.** 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature (London) **365:**274–276.
23. **Sutherland, G. R., and R. I. Richards.** 1995. Simple tandem DNA repeats and human genetic disease. Proc. Natl. Acad. Sci. USA **92:**3636–3641.
24. **Weaver, D. T., and M. L. DePamphilis.** 1984. Role of palindromic and nonpalindromic sequences in arresting DNA synthesis in vitro and in vivo. J. Mol. Biol. **180:**961–986.
25. **Wells, R. D.** 1996. Molecular basis of genetic instability of triplet repeats. J. Biol. Chem. **271:**2875–2878.
26. **Yu, A., J. Dill, S. S. Wirth, G. Huang, V. H. Lee, I. S. Haworth, and M. Mitas.** 1995. The trinucleotide repeat sequence d(CTG)15 adopts a hairpin conformation. Nucleic Acids Res. **23:**2706–2714.

Downloaded from http://jb.asm.org/ on September 19, 2018 by guest