(CAG)•(CTG) Repeats Associated with Neurodegenerative Diseases are Stable in the Escherichia coli Chromosome

Seung-Hwan Kim, Malgorzata J. Pytlos, William A. Rosche, and Richard R. Sinden*

From the Laboratory of DNA Structure and Mutagenesis, Center for Genome Research, Institute of Biosciences and Technology, Texas A&M University System Health Science Center, 2121 West Holcombe Blvd., Houston, Texas, 77030-3303 and the Department of Biological Science, University of Tulsa, 600 S. College Ave., Tulsa, OK 74104-3126

Running title: (CAG)•(CTG) repeat stability in the E. coli chromosome

Address correspondence to: Richard R. Sinden, Laboratory of DNA Structure and Mutagenesis, Center for Genome Research, Institute of Biosciences and Technology, Texas A&M University System Health Science Center, 2121 West Holcombe Blvd., Houston, TX 77030-3303, Tel. 713 677-7664; Fax. 713 677-7689; E-Mail: RSinden@IBT.tamhsc.edu

(CAG)n•(CTG)n deletion is associated with many neurodegenerative diseases. Repeat instability has been extensively studied in bacterial plasmids, where repeats undergo deletion at high rates. We report an assay for (CAG)n•(CTG)n deletion from the chloramphenicol acetyltransferase gene integrated into the E. coli chromosome. In strain AB1157, deletion rates for 25 to 60 (CAG)•(CTG) repeats integrated in the chromosome ranged from 6.88 x 10^-9 to 1.33 x 10^-10, or approximately 6,300 to 660,000-fold lower than in plasmid pBR325. In contrast to the situation in plasmids, deletions occur at a higher rate when (CTG)43, rather than (CAG)43, comprised the leading template strand, and complete rather than partial deletions were the predominant mutation observed. Repeats were also stable on long term growth following multiple passages through exponential and stationary phase. Mutations in priA and recG increased or decreased deletion rates, but repeats were still greatly stabilized in the chromosome. The remarkable stability of (CAG)n•(CTG)n repeats in the E. coli chromosome may result from the differences in the mechanisms for replication or the probability for recombination afforded by a high plasmid copy number. (CAG)n•(CTG)n tracts are generally comprised of less than 30 repeats in normal individuals. In certain diseases, such as DM1, repeats can expand from less than 30 to 30-80 copies resulting in mild symptoms, and then expand to thousands of copies in severely affected individuals. In the majority of (CAG)n•(CTG)n-associated diseases however, full-expansion mutations causing the disease are less than 80 repeats in length (see (2) for review). With all diseases, repeat expansion is the genetic mutation responsible for the disease and various approaches for causing repeat deletions or for preventing expansions throughout life may provide therapeutic approaches for disease treatment (3-7).

Much of our understanding of the biochemistry and molecular mechanisms of repeat instability has come from studies involving repeats cloned into plasmids in E. coli. Plasmid-based genetic assays for repeat deletion have shown that (CAG)n•(CTG)n repeats are very unstable when inserted into plasmids in E. coli, where they exhibit rates of deletion as high as 10^-4 to 10^-2 deletions per cell per generation (8-11). The repeats can be almost completely lost from the plasmid following several recultivations of the bacterial culture (12-15). Plasmid-based assays have revealed that mutations in genes involved in DNA mismatch repair (mutS), excision repair (uvrA), recombination (recA and recB), and replication restart (priA, recG, ruvAB, recF, and recO) influence repeat instability in E. coli (8,10,11,16). (CAG)n•(CTG)n repeats can block DNA replication in cells (17), and repeat deletions may predominantly result from attempts to restart DNA replication when the replication fork is stalled at trinucleotide repeats (8,10,16). Repeat instability in plasmids may also result from the high plasmid copy number that may promote recombination between repeats (18,19).

The expansion of (CAG)n•(CTG)n repeats is associated with many diseases including myotonic dystrophy type 1 (DM1), Huntington disease (HD), and many spinocerebellar ataxias (SCAs) (1,2).
The high rates of deletion and recombination in plasmids are very different from the situation in humans where intergenerational transmission can lead to expansion from hundreds to thousands of repeats (1,2). In some cases, such as fragile X syndrome, once expanded, the repeat tracts are somatically stable (20). In other diseases such as DM1 and HD, long repeat tracts can increase in length throughout life (21-23). Thus, the study of triplet repeats in bacterial plasmids, while useful for studying the mechanism and genetics of deletion, does not model many characteristics of (CAG)n•(CTG)n instability in human cells where high rates of deletion are not observed.

Here, we have developed a chromosome-based genetic assay for repeat deletion. For this, we placed a chloramphenicol acetyltransferase (CAT) gene, disrupted by insertion of (CAG)n•(CTG)n tracts, into the phage 8 attB attachment site in the E. coli chromosome using the lambda InCh method (24). In both wild type cells and in strains containing mutations affecting replication restart, (CAG)n•(CTG)n tracts were remarkably stable in the E. coli chromosome compared with the very high rate of instability in plasmids. Thus, this model system more closely mimics the situation for repeats in the human chromosome.

EXPERIMENTAL PROCEDURES

Strains and media — Parental E. coli strains, AB1157, MG1655, and SM551, were used as permissive strains for the generation of lambda InCh1 lysates, which were prepared from parental lysogenic strain DH6521 (Table 1). Derivatives of bacterial strain AB1157, AG109 (priA2::kan), RJ1075 (priA4; where the final six amino acids of the PriA protein are replaced by 12 different amino acid residues, resulting in a partially functional PriA (25)), and N3793 (ΔrecG263) were used as target strains for developing the chromosome-based genetic selection assays. Strains containing the original pBR325 plasmids were grown in lambda InCh broth (10 g of bacto-tryptone and 5 g of NaCl per liter of H2O). For reversion and long-term growth experiments, cells were grown at 37°C in K medium (M9 buffer supplemented with 1% casamino acids, 1% glucose, 0.2 mM CaCl2, 1 mM MgSO4, and 0.001% thiamine) containing 100 μg/ml of ampicillin (Amp). M9 buffer contained 1 g of NH4Cl, 5.8 g of Na2HPO4, and 3 g of KH2PO4 per liter. K + Amp plates contained 100 μg/ml of Amp. K + Amp + Cm plates contained 100 mg/ml of Amp and 25 mg/ml of chloramphenicol (Cm).

Construction of the chromosome-based genetic assay — The CAT gene, disrupted by (CAG)n•(CTG)n tracts from pVH plasmids derived from pBR325 (10), was inserted into the E. coli chromosome using lambda (λ) InCh vectors as described by Boyd et al. (24). Different plasmids were used that contained either orientation of the CAT gene, thus reversing the repeats with respect to the direction of replication. The structure at the insertion site was confirmed by PCR and DNA sequencing as recommended (24).

Repeat deletion rate measurements and analysis of revertants — Measurement of the rate of reversion to chloramphenicol resistance (Cmr), which occurs by deletion to ≤24 repeats, was as described (10,16). At least 6 independent reversion assays were performed for each strain. The mutation rate was determined by the method of the median frequency (26). 95% confidence intervals (C.I.) for the individual mutation rates were calculated by the method of the population mean (27). Statistical significance levels (p values) between the mutation rates were determined by the two-sample ranking test of the Mann-Whitney Test (27). The nature of the deletion events in the Cmr revertants was analyzed by PCR as described (10).

Long term growth assay and analysis — A 10 ml culture in K medium containing ampicillin (30 μg/ml) was inoculated from a single colony and grown overnight at 37°C at 250 rpm. The next day, the cultures were recultivated by a 1:1000 dilution into 10 ml of K medium containing ampicillin and grown for another one or two days. This procedure was repeated every day for a total of 7 days and, in a separate experiment, every 2 days for a total of 14 days. Plasmids were purified each time the culture was recultivated. The length of the repeat tract was determined by polyacrylamide gel analysis of PCR products encompassing the repeat region, and the frequency of Cmr revertants in the population was determined.

RESULTS

A genetic assay for studying the genomic instability of DNA repeats inserted into the E. coli chromosome—We have developed and used
pBR325 plasmid-based genetic assays to measure rates of deletion of (CAG)$_n$•(CTG)$_m$ tracts in *E. coli* (8-10,16). Here, we have introduced this system, in single copy, into the *E. coli* chromosome. For this, both orientations of the CAT gene in plasmid pBR325, disrupted by (CAG)$_n$•(CTG)$_m$ tracts, were inserted into the *attB* site of the *E. coli* chromosome (24). The two orientations, with respect to the chromosomal replication origin (oriC), place the (CAG)$_m$ or (CTG)$_m$ tract in the leading template strand for *E. coli* chromosome replication (Figure 1). During the process of introducing the repeats into the chromosome, the (CAG)$_n$•(CTG)$_m$ tract present in the original plasmid was shortened to (CTG)$_{60}$•(CAG)$_{60}$ in the cured lysogens SKA793, SKG793, SKR793, and SKN793, and to (CAG)$_{52}$•(CTG)$_{52}$ in the cured lysogens SKA794, SKG794, and SKN794. However, cured lysogen SKR794 contained the original length (CAG)$_{79}$•(CTG)$_{79}$ (Table 1). The (CAG)$_{25}$•(CTG)$_{25}$ tract was duplicated to (CAG)$_{50}$•(CTG)$_{50}$ in cured lysogens, SKA254, SKG254, SKR254, and SKN254 (Table 1). Presumably, these changes occurred during construction of the primary lysogens. The initial (CAG)$_{43}$•(CTG)$_{43}$ tracts were stable upon introduction into the chromosome in all cured lysogens (Table 1).

As expected in this genetic assay, the in-frame insertion of (CAG)$_m$•(CTG)$_m$ tracts of 25 repeats or longer in the CAT gene resulted in a chloramphenicol-sensitive (Cms) phenotype in either orientation of the CAT gene. In plasmids, deletions that reduce the length too less than 25 (CAG)$n$•(CTG)$m$ repeats restore a chloramphenicol-resistant (Cm$'$) phenotype, allowing measurement of the rates of deletions (10). This was also the case for the CAT gene constructs integrated into the chromosome as revertants containing 23 repeats were obtained for Cm$'$ revertants from cells containing the (CAG)$_{25}$•(CTG)$_{25}$ tract (see below).

**(CAG)$n$•(CTG)$m$ repeats were very stable in the *E. coli* chromosome** — To determine (CAG)$n$•(CTG)$m$ deletion rates, cultures were grown to mid log phase in K medium from single colonies from frozen cultures, as described (10). In parental strain AB1157, mutation rates for (CAG)$_{25}$, (CAG)$_{43}$, and (CAG)$_{52}$ tracts in the leading template strand of the *E. coli* chromosome ranged from 2.05–18.02 ×10$^{-10}$ Cm$'$ revertants per cell per generation (Table 2). Mutation rates for (CTG)$_{43}$ and (CTG)$_{60}$ tracts in the leading template strand of the *E. coli* chromosome were 6.88 ×10$^{-9}$ and 2.29 ×10$^{-9}$ respectively.

Repeat deletion rates were also measured in several derivatives of AB1157 containing mutations in *priA* and *recG*, which reduce repeat deletion rates in plasmids (16). In the *priA2* strain (AG109), mutation rates for the two orientations of (CTG)$_{43}$ and (CAG)$_{43}$ tracts were very similar and they were 14 and 4-fold lower than the rates in the isogenic parental strain AB1157, respectively (Table 2). The differences in rates comparing the parental and *priA2* strain were significant ($p<0.001$). In contrast, mutation rates for (CAG)$_{25}$, (CTG)$_{60}$, and (CAG)$_{32}$ tracts in the *E. coli* chromosome were increased by factors of 2.8, 1.7, and 4.1, respectively ($p<0.01$ to 0.005) compared with rates in parental strain AB1157. In the *priA4* strain (RF1075) mutation rates for most repeats (CAG)$_{25}$, (CAG)$_{43}$, and (CTG)$_{43}$ tracts in the chromosome were decreased 1.6 to 5.5-fold compared with the parental strain ($p<0.01$ to 0.001). The mutation rate for the (CAG)$_{79}$ tract was similar to rates for repeat lengths from 43-60 in this strain. In strain *recG263*, mutation rates for (CAG)$_{25}$, (CTG)$_{43}$, and (CAG)$_{43}$ tracts were decreased 1.6 to 25-fold compared with the parental strain ($p<0.001$), while mutation rates for (CTG)$_{60}$ and (CAG)$_{52}$ were increased 2 and 23.6-fold, respectively, relative to the parental strain ($p<0.001$).

**(CAG)$n$•(CTG)$m$ repeats generally undergo complete deletions in the *E. coli* chromosome during selection in log phase growth** — Mutation spectra of Cm$'$ revertants were characterized by PCR analyses (Table 3). All of the initial 43 to 79 (CAG)$n$•(CTG)$m$ repeats were completely deleted in the Cm$'$ revertants in all strains. When (CAG)$_{25}$ tracts were placed in the leading template strand of the *E. coli* chromosome, complete deletions prevailed in the parental, *priA4*, and *recG263* strains, although several Cm$'$ revertants in wild type and the *recG* strain had deletions of only two repeats, leaving 23 repeats in the chromosome. In the *priA2* strain, only the deletion of two repeats was observed for (CAG)$_{25}$.

**Stability of (CAG)$n$•(CTG)$m$ repeats in the *E. coli* chromosome during repeated passage through log and stationary phase** — In a biochemical assay for repeat instability, long tracts of (CAG)$n$•(CTG)$m$ repeats can be almost completely lost from plasmid following only several recultivations of the bacterial culture in which cells go through many cell divisions and multiple cycles of log and stationary phase (12-
To test the stability of (CAG)(CTG) repeats in the *E. coli* chromosome under these growth conditions, cells were grown for 7 days with recultivation every day, or for 14 days with recultivation every two days. Cells were grown in the presence of ampicillin selection to maintain plasmid, but not chloramphenicol selection to select for repeat deletions. On analysis of the repeat tracts in the total population by PCR analysis, only the original repeat lengths were detected (data not shown). Cmr revertants were detected in the long term growth cultures but at only very low levels, consistent with the rates reported in Table 2.

**DISCUSSION**

We have used a chromosome-based genetic assay to measure deletion rates of repeat tracts from the *E. coli* chromosome. For this, the CAT gene, disrupted by the in-frame insertion of 25, 43, 52, 60, or 79 (CAG)(CTG) repeats, was inserted into the phage *λ* *attB* attachment site of the *E. coli* chromosome using the lambda InCh method (24) (Figure 1). These lengths of repeats encompass the length in normal individuals (<30) as well as the full repeat expansions for the majority of (CAG)(CTG) repeat expansion diseases (2). We have not yet been able to test the stability of very long repeat tracts associated with a few diseases, such as DM1. In our previous plasmid-based genetic assay, mutation rates for deletions to less than 25 repeats, from pBR325-derivative plasmids carrying 25, 43, or 79 (CAG)(CTG) repeats, ranged from $1.1 \times 10^{-4}$ to $2 \times 10^{-6}$ in parental strain AB1157 and its derivatives (16). In the chromosome-based genetic assay, the deletion rates ranged from $6.88 \times 10^{-9}$ to $1.33 \times 10^{-10}$ in parental strain AB1157 and its derivatives (16). In the chromosome-based genetic assay, the deletion rates ranged from $6.88 \times 10^{-9}$ to $1.33 \times 10^{-10}$ in AB1157 (Figure 2, A, B). These rates are approximately 6,300 to 660,000-fold lower than the rates in plasmids. Considering a copy number of about 20 for plasmid (10) compared with a copy number of 2-4 for the chromosome, the differences in mutation rates remain high (630 - 1,260 to 63,000 - 126,000 fold lower). In addition, upon long term growth, through multiple cycles of log and stationary phase the (CAG)(CTG) repeats were stable, unlike the situation in plasmids where they undergo rapid deletion generally within 20-40 generations (12-15). Thus, (CAG)(CTG) repeat tracts are extremely stable in the *E. coli* chromosome relative to the situation in plasmids. Moreover, the mutation spectrum in the chromosome was quite different from that in the plasmid. In the chromosome complete deletions predominated, except for (CAG)$_{25}$ in the priA2 strain. In plasmids, partial deletions were more common (Figure 2, C, D) (16).

These large differences in stability may be explained by one or more of the following differences: 1) ColE1-based plasmids and the *E. coli* chromosome initiate DNA replication very differently. In *E. coli*, the initiation of chromosomal DNA replication occurs at a 254-bp sequence, *oriC*, in a process controlled by the DnaA protein (28). ColE1 plasmids, on the other hand, initiate replication from a very long RNA primer molecule (RNA II) synthesized by RNA polymerase. Transcription begins 555 nucleotides from the origin (where the transition for DNA synthesis is made). DNA pol I then continues DNA synthesis on the long RNA primer for at least 400 nucleotides before DNA pol III begins synthesis. It is not until this point that a primosome assembles to begin lagging strand synthesis (28). Instability, including aberrant replication events, such as intermolecular strand switching during replication, can occur at this Pol I to Pol III junction region (29). 2) Plasmid replication is unidirectional, whereas replication on the chromosome is bidirectional (28). 3) The copy number of the CAT gene in cells is also very different, being about 20 for cells harboring pBR325 derivatives, compared with 2-4 under log phase conditions when the CAT gene is introduced into the bacterial chromosome (28). The high copy number and the PolI-PolIII switch may increase the probability of primer-template misalignment. Moreover, this may also promote the high rate of recombination observed in repeats in plasmids (18,19) while recombination may only rarely be involved in repeat instability in human cells (2). 4) ColE1-based plasmids are not maintained in the priA2 strain (30), suggesting a strict requirement for PriA in plasmid replication or stability, but not for chromosomal replication. 5) Finally, the size difference of the two genomes may influence stability, given that the repeat may constitute 10% of a plasmid, but only a very small proportion of the chromosome. DNA secondary transitions, which might occur within the (CAG)(CTG) tract (31,32), would have very different consequences for supercoiling levels given different topological domain sizes.
Consistent with our results, a large difference between deletion rates of DNA repeats when present on plasmids or on the chromosome was also observed in another genetic assay in E. coli. The rates of deletion between long tandem repeats in the chromosome in E. coli strains carrying mutations in many replication genes ranged from 1.7 x 10^{-5} to 3.1 x 10^{-9}. These rates were approximately 10 to 100,000-fold lower than rates in the plasmid (30). Also, mutation rates of (CTG)_n•(CAG)_n tracts within yeast chromosomes range from 2 x 10^{-3} to 5 x 10^{-5} (33-35), rates that are much higher than observed in the E. coli chromosome.

In plasmids, (CAG)•(CTG) repeats are sometimes more prone to deletions when (CTG)_n tracts are placed in the lagging rather than the leading template strand (10,14). On the other hand, expansions are often favored when the (CTG) tract comprises in the leading template strand (14). However, this strand bias for deletion was not observed for (CTG)•(CAG) repeats in plasmid pBR325 in strain AB1157 grown in K medium (16), and in the present study (CTG)_n tracts were more unstable when they comprised the leading template strand of the E. coli chromosome in AB1157 and derivatives (Table 2). In another chromosome mutation assay system, base substitution mutations occurred at a slightly higher rate in the leading template strand (36), while this pattern was altered in cells when the SOS repair system was induced (37). Differences in orientation-dependent repeat instability in the plasmids and the chromosome may also result from the differences in these two systems, as discussed previously, and as discussed below.

The progression of replication can be blocked at trinucleotide repeats (17,38,39) and by collisions between replication and transcription complexes (40). To compare different repeat orientations in the plasmid- and chromosome-based CAT gene assay systems, the direction of transcription of the CAT gene is reversed with respect to movement of the replication fork. Thus, deletion rates may be influenced by the orientation of the CAT gene. Indeed, in the previous plasmid-based genetic assays, Cmr reversion frequencies were decreased approximately 2-fold when the CAT gene containing symmetrical repeat inserts was reversed, making the direction of transcription the same as the direction of the replication fork (41). In the chromosome-based genetic assay in parental strain AB1157, the deletion rate of 43 (CAG)•(CTG) repeats was 3.8-fold lower when progression of the replication fork was co-directional with transcription of the CAT gene than when replication and transcription collide (Table 2). Our results are consistent with an interpretation that collisions between DNA polymerase and RNA polymerase complexes at the repeats may be responsible for an increased rate of deletion. Indeed, in wild-type and mismatch-repair-deficient mutL strains containing lacZ frameshift alleles, mutation frequencies were increased 1.1 to 1.6-fold and 1 to 5.4-fold, respectively, when the movement of the fork was opposite to the direction of lac transcription in the E. coli chromosome (42).

In the priA2, priA4, and recG263 strains, deletion rates for strains containing 43 or fewer repeats were reduced relative to the parental strain, consistent with observations in plasmids (16). The influence of mutations in these genes on the rates of deletion is consistent with the involvement of replication restart in instability (16). An exception was the (CAG)_{35} tract in the priA2 strain, where the Cmr reversion rate was increased. Moreover, in this strain only partial deletions were observed in strong contrast to all other chromosomal mutation spectra. This is consistent with an increase in replication slippage in this genetic background. Repeat tracts longer than 43 repeats generally showed an increase in deletion rates in priA and recG strains relative to the parental strain. This may result from an increased probability of DNA secondary structure formation leading to replication pausing, or DNA secondary structure formation occurring during replication restart pathways (8). Forty three repeats may represent the length threshold for replication pausing, perhaps mediated by structure formation. This is in the range of repeat thresholds for pausing in (CGG)•(CCG) and (GAA)•(TTC) tracts in plasmid in yeast (39). However, it is shorter than the threshold observed for (CAG)•(CTG) repeats in plasmids in E. coli and yeast (17,39). The length threshold for structure formation may be different in the E. coli chromosome, especially in the presence of mutations in genes involved in replication restart. In summary, these results show that mutations in priA2, priA4, and recG263 can increase or decrease rates of (CAG)•(CTG)_n deletion from the E. coli chromosome depending on the repeat length and particular mutation.

Human disease-associated DNA repeats thousands of base pairs in length can be maintained...
relatively stably in human chromosomes (20-23), whereas they exhibit very high rates of deletion in *E. coli* plasmids (8,10,18,19). The large differences in rates of instability of triplet repeats comparing plasmids with the bacterial chromosome, may be due to differences in genome size, mechanisms of replication, and copy number. This chromosome-based genetic assay should provide a valuable new system for understanding many features of the mechanisms responsible for the instabilities of disease-associated DNA repeats.

REFERENCES

1. Sinden, R. R. (1999) *Am. J. Hum. Genet.* 64, 346-353
2. Cleary, J. D. and Pearson, C. E. (2003) *Cytogenet. Genome Res.* 100, 25-55
3. Hashem, V. I., Pytlos, M. J., Klysik, E. A., Tsuji, K., Khajavi, M., Ashizawa, T., and Sinden, R. R. (2004) *Nucleic Acids Res.* 32, 6334-6346
4. Pineiro, E., Fernandez-Lopez, L., Gamez, J., Marcos, R., Surralles, J., and Velazquez, A. (2003) *Nucleic Acids Res.* 31, 6733-6740
5. Gorbunova, V., Seluanov, A., Dion, V., Sandor, Z., Meservy, J. L., and Wilson, J. H. (2003) *Mol. Cell Biol.* 23, 4485-4493
6. Yang, Z., Lau, R., Marcadier, J. L., Chitayat, D., and Pearson, C. E. (2003) *Am. J. Hum. Genet.* 73, 1092-1105
7. Gomes-Pereira, M. and Monckton, D. G. (2004) *Nucleic Acids Res.* 32, 2865-2872
8. Hashem, V. I., Rosche, W. A., and Sinden, R. R. (2004) *Mutat. Res.* 554, 95-109
9. Hashem, V. I. and Sinden, R. R. (2002) *Mutat. Res.* 508, 107-119
10. Hashem, V. I., Rosche, W. A., and Sinden, R. R. (2002) *Mutat. Res.* 502, 25-37
11. Oussatcheva, E. A., Hashem, V. I., Zou, Y., Sinden, R. R., and Potaman, V. N. (2001) *J. Biol. Chem.* 276, 30878-30884
12. Bowater, R. P., Rosche, W. A., Jaworski, A., Sinden, R. R., and Wells, R. D. (1996) *J. Mol. Biol.* 264, 82-96
13. Jaworski, A., Rosche, W. A., Gellibolian, R., Kang, S., Shimizu, M., Bowater, R. P., Sinden, R. R., and Wells, R. D. (1995) *Proc. Natl. Acad. Sci. USA* 92, 11019-11023
14. Kang, S., Jaworski, A., Oshtima, K., and Wells, R. D. (1995) *Nat. Genet.* 10, 213-218
15. Rosche, W. A., Jaworski, A., Kang, S., Kramer, S. F., Larson, J. E., Geidroc, D. P., Wells, R. D., and Sinden, R. R. (1996) *J. Bacteriol.* 178, 5042-5044
16. Kim, S.-H., Pytlos, M. J., and Sinden, R. R. (2006) *Mutat. Res.* 595, 5-22.
17. Samadashwily, G. M., Rac, G., and Mirkin, S. M. (1997) *Nature Genet.* 17, 298-304
18. Napierala, M., Parniewski, P., Pluciennik, A., and Wells, R. D. (2002) *J. Biol. Chem.* 277, 34087-34100
19. Jakupciak, J. P. and Wells, R. D. (1999) *J. Biol. Chem.* 274, 23468-23479
20. Vohle, D., Hennig, I., Vogel, W., and Steinbach, P. (1993) *Nature Genet.* 4, 140-142
21. Martorell, L., Gamez, J., Cayuela, M. L., Gould, F. K., McAbney, J. P., Ashizawa, T., Monckton, D. G., and Baiget, M. (2004) *Neurology* 62, 269-274
22. Kennedy, L., Evans, E., Chen, C. M., Craven, L., Detloff, P. J., Ennis, M., and Shelbourne, P. F. (2003) *Hum Mol. Genet.* 12, 3359-3367
23. Wong, L.-J. C., Ashizawa, T., Monckton, D. G., Caskey, C. T., and Richards, C. S. (1995) *Am. J. Hum. Genet.* 56, 114-122
24. Boyd, D., Weiss, D. S., Chen, J. C., and Beckwith, J. (2000) *J. Bacteriol.* 182, 842-847
25. Gregg, A. V., McGlynn, P., Jakta, R. P., and Lloyd, R. G. (2002) *Mol. Cell* 9, 241-251
26. Drake, J. W. (1991) *Proc. Natl. Acad. Sci. USA* 88, 7160-7164
27. Zar, J. H. (1999) *Biostatistical Analysis*, Fourth Edition Ed., Prentice-Hall,
28. Kornberg, A. and Baker, T. A. (1992) *DNA Replication*, 2 Ed., W. H. Freeman & Co., New York
29. Backman, K., Betlach, M., Boyer, H. W., and Yanofsky, S. (1978) *Cold Spr. Harbor Symp.*
30. Savesan, C. J. and Lovett, S. T. (1997) Genetics 146, 457-470
31. Pearson, C. E. and Sinden, R. R. (1996) Biochemistry 35, 5041-5053
32. Pearson, C. E., Tam, M., Wang, Y. H., Montgomery, S. E., Dar, A. C., Cleary, J. D., and Nichol, K. (2002) Nucleic Acids Res. 30, 4534-4547
33. Freudenreich, C. H., Kantrow, S. M., and Zakian, V. A. (1998) Science 279, 853-856
34. Callahan, J. L., Andrews, K. J., Zakian, V. A., and Freudenreich, C. H. (2003) Mol. Cell. Biol. 23, 7849-7860
35. Bhattacharyya, S., Rolfsmeier, M. L., Dixon, M. J., Wagoner, K., and Lahue, R. S. (2002) Genetics 162, 579-589
36. Fijalkowska, I. J., Jonczyk, P., Tkaczyk, M. M., Bialoskorska, M., and Schaaper, R. M. (1996) Proc Natl Acad Sci U S A 93, 10020-10025
37. Maliszewska-Tkaczyk, M., Jonczyk, P., Bialoskorska, M., Schaaper, R. M., and Fijalkowska, I. J. (2000) Proc Natl Acad Sci U S A 97, 12678-12683
38. Krasilnikova, M. M. and Mirkin, S. M. (2004) Mol. Cell. Biol. 24, 2286-2295
39. Pelletier, R., Krasilnikova, M. M., Samadashwily, G. M., Lahue, R., and Mirkin, S. M. (2003) Mol. Cell. Biol. 23, 1349-1357
40. Vilette, D., Ehrlich, S. D., and Michel, B. (1995) Mol. Microbiol. 17, 493-504
41. Rosche, W. A., Trinh, T. Q., and Sinden, R. R. (1995) J. Bacteriol. 177, 4385-4391
42. Gawel, D., Jonczyk, P., Bialoskorska, M., Schaaper, R. M., and Fijalkowska, I. J. (2002) Mutat. Res. 501, 129-136

FOOTNOTES

Acknowledgments— This work was supported by National Institutes of Environmental Health Sciences grant ES05508 to RRS.

Abbreviations— DM1, myotonic dystrophy type 1; DM2, myotonic dystrophy type 2; FA, Friedreich ataxia; HD, Huntington disease; CAT, chloramphenicol acetyltransferase

FIGURE LEGENDS

Figure 1. Insertion of the CAT gene containing (CAG)•(CTG) repeats into the E. coli chromosome. The CAT gene containing various lengths of (CAG)•(CTG) repeats was inserted into the E. coli chromosome at the attB site as described (24). By using pBR325-based plasmids containing opposite orientations of the CAT gene (8-10), both orientations of (CTG)n•(CAG)n tracts, with respect to the chromosomal replication origin (oriC), were introduced into the chromosome.

Figure 2. Comparison of (CTG)n•(CAG)n instability in plasmids and the E. coli chromosome. A. Deletion rates for (CTG)n•(CAG)n repeats in the CAT gene in plasmid pBR325 in strains AB1157 and priA4 and recG derivatives (data from (16)). B. Deletion rates for (CTG)n•(CAG)n repeats in the CAT gene but integrated in the chromosome (from Table 2). C. Frequency of complete deletion of the (CTG)n•(CAG)n tract from the CAT gene in plasmid pBR325 in strains AB1157 and priA4 and recG derivatives (data from (16)). D. Frequency of complete deletion of the (CTG)n•(CAG)n tract in the CAT gene integrated in the chromosome (from Table 3).
| Strain          | Relevant genetic marker(s) or features                                             | Reference and source |
|-----------------|----------------------------------------------------------------------------------|----------------------|
| **Parental strains**                                         |                                                                                  |                      |
| AB1157          | thr- araC14 leuB6 Δ(gpt-proA)62 lacY1 tsx-33 gsr' - glnV44(AS) galK2(Oc) λ- Rac-0 hisG4(Oc) rfbD1 mgl-51 rpoS396 (Am) rpsL31(strR) kdgK5 xyl-5 mtl-1 argE3(Oc) thi |                      |
| MG1655          | LAM' rph-1                                                                         |                      |
| SM551           | F- λ- λ' Δlac(MS265) mel NalA' supF58(=sulII')                                    | (24)                 |
| DHB6521         | SM551 λ.lnCh1 (Kan')                                                              |                      |
| **AB1157 derivatives**                                      |                                                                                  |                      |
| AG109           | priA2::kan                                                                         | (25), R. Lloyd       |
| RJ1075          | priA4::kan                                                                         | (25), R. Lloyd       |
| N3793           | ΔrecG263                                                                          | (25), R. Lloyd       |
| **Cured lysogens**                                          |                                                                                  |                      |
| SKA253          | AB1157 Δ(attB-lom)::bla Cm'(CTG)50a                                             | This study           |
| SKA254          | AB1157 Δ(attB-lom)::bla Cm'(CAG)25                                               | This study           |
| SKA433          | AB1157 Δ(attB-lom)::bla Cm'(CTG)43                                               | This study           |
| SKA434          | AB1157 Δ(attB-lom)::bla Cm'(CAG)43                                               | This study           |
| SKA793          | AB1157 Δ(attB-lom)::bla Cm'(CTG)60                                               | This study           |
| SKA794          | AB1157 Δ(attB-lom)::bla Cm'(CAG)52                                               | This study           |
| SKG253          | AG109 Δ(attB-lom)::bla Cm'(CTG)50a                                              | This study           |
| SKG254          | AG109 Δ(attB-lom)::bla Cm'(CAG)25                                                | This study           |
| SKG433          | AG109 Δ(attB-lom)::bla Cm'(CTG)43                                                | This study           |
| SKG434          | AG109 Δ(attB-lom)::bla Cm'(CAG)43                                                | This study           |
| SKG793          | AG109 Δ(attB-lom)::bla Cm'(CTG)60                                                | This study           |
| SKG794          | AG109 Δ(attB-lom)::bla Cm'(CAG)52                                                | This study           |
| SKR253          | RJ1075 Δ(attB-lom)::bla Cm'(CTG)50a                                              | This study           |
| SKR254          | RJ1075 Δ(attB-lom)::bla Cm'(CAG)25                                               | This study           |
| SKR433          | RJ1075 Δ(attB-lom)::bla Cm'(CAG)43                                               | This study           |
| SKR434          | RJ1075 Δ(attB-lom)::bla Cm'(CAG)43                                               | This study           |
| SKR793          | RJ1075 Δ(attB-lom)::bla Cm'(CTG)60                                               | This study           |

Table 1. *Escherichia coli* strains
| Strain       | Relevant genetic marker(s) or features                      | Reference and source |
|-------------|------------------------------------------------------------|----------------------|
| SKR794      | RJ1075 $\Delta(\lambda$attB-lom)::bla Cm'$\text{(CAG)}_{79}$ | This study           |
| SKN253      | N3793 $\Delta(\lambda$attB-lom)::bla Cm'$\text{(CTG)}_{50^a}$ | This study           |
| SKN254      | N3793 $\Delta(\lambda$attB-lom)::bla Cm'$\text{(CAG)}_{25}$  | This study           |
| SKN433      | N3793 $\Delta(\lambda$attB-lom)::bla Cm'$\text{(CTG)}_{43}$  | This study           |
| SKN434      | N3793 $\Delta(\lambda$attB-lom)::bla Cm'$\text{(CAG)}_{43}$  | This study           |
| SKN793      | N3793 $\Delta(\lambda$attB-lom)::bla Cm'$\text{(CTG)}_{60}$  | This study           |
| SKN794      | N3793 $\Delta(\lambda$attB-lom)::bla Cm'$\text{(CAG)}_{52}$  | This study           |

$^a$ After transfer of a CAT gene containing an insertion of $(\text{CTG})_{25}•(\text{CAG})_{25}$ into the *E. coli* chromosome, the repeat tracts were duplicated to 50 in all cured lysogens. Presumably, these changes occurred at the level of the primary lysogens for this original plasmid construct.
Table 2. Deletion rates of (CAG)$_n$•(CTG)$_n$ tracts on the *E. coli* chromosome

| Repeats | AB1157 (parental) | AG109 (priA2) | RF1075 (priA4) | N3793 (recG263) |
|---------|-------------------|---------------|----------------|-----------------|
|         | Cm$^r$ reversion rate $\times 10^{-10}$ (95% C.I.) | Rel. rate | Cm$^r$ reversion rate $\times 10^{-10}$ (95% C.I.) | Rel. rate | Cm$^r$ reversion rate $\times 10^{-10}$ (95% C.I.) | Rel. rate |
| (CAG)$_{25}$ | 1.32 (1.1-1.56) | 1 | 3.73 (2.36-5.11) | 2.81 | 0.368 (0.23-0.51) | 0.27 |
| (CTG)$_{43}$ | 68.83 (64.93-72.73) | 1 | 5.01 (4.5-5.12) | 0.07 | 12.79 (7.71-17.87) | 0.18 |
| (CAG)$_{43}$ | 18.02 (15.35-20.69) | 1 | 4.26 (3.73-4.8) | 0.23 | 11.3 (5.98-16.63) | 0.62 |
| (CTG)$_{60}$ | 22.9 (20.94-24.86) | 1 | 39.14 (13.17-65.11) | 1.7 | 9.89 (7.29-12.49) | 0.43 |
| (CAG)$_{52}$ | 2.054 (1.77-2.34) | 1 | 8.39 (4.4-12.37) | 4.1 | 48.49 (41.98-55) | 23.6 |
| (CAG)$_{79}$ | 14.17 (10.19-18.16) | NA |

$^a$ Mutation rates (Cm$^r$ revertants per cell per generation) were calculated by the method for the median (26)95% confidence intervals (C.I.) for the individual rate calculations (27) are shown in parentheses.

$^b$ Relative rates ([mutation rate in mutant strain]/[mutation rate in parent strain]) were determined by comparison to the parental strain.
Table 3. Deletion spectra for \((\text{CTG})_n•(\text{CAG})_n\) tracts in the CAT gene in the \textit{E. coli} chromosome in AB1157 derivatives

| Repeats  | \(\Delta_c^a\) | \(\Delta_p^b\) | \(\Delta_c^a\) | \(\Delta_p^b\) | \(\Delta_c^a\) | \(\Delta_p^b\) | \(\Delta_c^a\) | \(\Delta_p^b\) |
|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| \((\text{CAG})_{25}\) | 16 | 4 | 23 | 0 | 20 | 23 | 20 | 0 | NA | 19 | 1 | 23 |
| \((\text{CTG})_{43}\) | 20 | 0 | NA | 19 | 0 | NA | 20 | 0 | NA | 20 | 0 | NA |
| \((\text{CAG})_{43}\) | 20 | 0 | NA | 19 | 0 | NA | 19 | 0 | NA | 13 | 0 | NA |
| \((\text{CTG})_{60}\) | 20 | 0 | NA | 20 | 0 | NA | 20 | 0 | NA | 19 | 0 | NA |
| \((\text{CAG})_{52}\) | 19 | 0 | NA | 20 | 0 | NA | 20 | 0 | NA |
| \((\text{CAG})_{79}\) | 19 | 0 | NA |

\(^a\) \(\Delta_c\), complete deletion; \(^b\) \(\Delta_p\), partial deletion; \(^c\) ND, not done; \(^d\) NA, not applicable
Figure 1
Figure 2

Kim, et al: (CAG)•(CTG) repeat stability in the E. coli chromosome
(CAG)\textsuperscript{5}(CTG) repeats associated with neurodegenerative diseases are stable in the Escherichia coli chromosome
Seung-Hwan Kim, Malgorzata J. Pytlos, William A. Rosche and Richard R. Sinden

J. Biol. Chem. published online July 27, 2006

Access the most updated version of this article at doi: 10.1074/jbc.M601129200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts