J. Mol. Biol. (1996) **264**, 82–96

JMB

Relationship Between *Escherichia coli* Growth and Deletions of CTG·CAG Triplet Repeats in Plasmids

Richard P. Bowater, William A. Rosche, Adam Jaworski Richard R. Sinden and Robert D. Wells*

Albert B. Alkek Institute of Biosciences and Technology Texas A and M University Center for Genome Research, Department of Biochemistry and Biophysics Texas Medical Center 2121 W. Holcombe Blvd. Houston TX 77030, USA

**Corresponding author*

Instabilities that are intrinsic to natural repetitive DNA sequences produce high frequencies of length changes *in vivo*. Triplet repeats cloned in plasmids in *Escherichia coli* undergo expansions and deletions, and this instability is affected by multiple factors. We show that CTG·CAG repeats in plasmids can influence the growth of *E. coli*, which affects the observed stabilities. At extended growth periods, the observed frequencies of deletions were dramatically increased if the cells passed through stationary phase before subculturing. Deletions were particularly pronounced for a plasmid containing the longest repeat, 525 bp in total, with the CTG sequence as the lagging strand template for replication. Measurements of cell growth showed that the lag phase associated with *E. coli* growth was increased for cultures containing plasmids with long CTG·CAG repeats, particularly when the CTG-containing strand was the lagging template. High frequencies of deletions were observed because of a growth advantage of cells containing plasmids with deleted triplet repeats. Incubation conditions that reduced the bacterial growth-rate produced a decreased extent of deletions, presumably because they alleviated the growth advantage of cells harboring plasmids with deleted triplet repeats. The experimental observations were simulated by a model in which shorter triplet repeats provided a growth advantage due to a shorter lag phase. We demonstrate that the accumulation of deletions within repeating sequences during growth of *E. coli* can be prevented, and discuss these findings in relation to the studies of repetitive DNA sequences. These are the first observations to show a direct influence between a plasmid-based DNA sequence or structure and factors controlling bacterial growth.

© 1996 Academic Press Limited

Keywords: triplet repeats; plasmid instability; DNA structure; deletions; *E. coli* growth

Introduction

Natural DNAs contain a wide array of repeating sequences (Charlesworth *et al*., 1994; Tautz & Schlötterer, 1994; Sutherland & Richards, 1995). Repeat sequences with 1 to 5 bp in their unit structure are termed simple repeating sequences. More complex repeat units, known as minisatellites, are also ubiquitous and have been particularly well characterized in higher eukaryotes (Armour & Jeffreys, 1992).

Simple repetitive DNA sequences have an intrinsic instability in all genomes, manifested as frequent length changes due to expansions or deletions of repeat units (Richards & Sutherland, 1992; Lustig & Petes, 1993; Charlesworth *et al*., 1994; Tautz & Schlötterer, 1994). Instabilities within repetitive DNA sequences have been linked to a variety of human diseases (Krontiris, 1995; Sutherland & Richards, 1995). For example, some cancer cell lines have an increased frequency of length changes in specific dinucleotide repeats (Aaltonen *et al*., 1993; Thibodeau *et al*., 1993) due to

Present addresses: W. A. Rosche, Department of Environmental Health, School of Public Health, Boston University School of Medicine, 80 E. Concord St., Boston, MA 02118, USA; A. Jaworski, Department of Genetics and Microorganisms, Institute of Microbiology and Immunology, University of Lodz, Banacha Str. 12/16, 90–237, and Center for Microbiology and Virology, Polish Academy of Sciences, Lodowa Str. 106, 93–232, Lodz, Poland.

Abbreviations used: TRS, triplet repeat sequence(s); rpm, revolutions per minute.

0022–2836/96/460082–15 $25.00/0

© 1996 Academic Press Limited

defects in DNA repair proteins (Fishel *et al*., 1993; Leach *et al*., 1993; Parsons *et al*., 1993). Of particular interest are triplet repeat sequences (TRS) such as CTG·CAG, CGG·CCG and AAG·CTT, which are associated with at least nine human hereditary neuromuscular or neurodegenerative disorders (for recent reviews, see Caskey *et al*., 1992; Willems, 1994; Ashley & Warren, 1995; Warren, 1996). Most of these disorders show the clinical behavior of anticipation (the increased severity and/or decreased age of onset of a hereditary disease with progression through a pedigree). The mutation events correlated with these diseases involve an increase in the copy number of the TRS. These expansions can be modest, as in Huntington's and Kennedy's diseases, or massive, as in fragile X syndrome and myotonic dystrophy (Ashley & Warren, 1995). Since the molecular mechanisms responsible for these expansions are unknown, the development of a genetically tractable system for investigating these instabilities would be a significant advance.

Repetitive DNA sequences are susceptible to slipped-strand mispairing (for a review, see Levinson & Gutman, 1987) which provides a plausible mechanism for generating unusual DNA conformations (Sinden & Wells, 1992; Kang *et al*., 1995a; Pearson & Sinden, 1996; Wells, 1996). The formation of folded structures within short single-stranded oligonucleotides of TRS has been shown to occur *in vitro* (Fry & Loeb, 1994; Gacy *et al*., 1995; Gao *et al*., 1995; Mitas *et al*., 1995; Mitchell *et al*., 1995; Smith *et al*., 1995; Yu *et al*., 1995a,b). These structures are stabilized by the formation of intra-strand base-pairs, although they also contain mismatched bases. Also, stable non-*B* DNA structures were observed in sequences containing long CTG·CAG and CGG·CCG repeats (Jaworski *et al*., 1995; Pearson & Sinden, 1996; and unpublished results). Therefore, the indications are that the instability of repetitive DNA sequences is related to an alternative DNA structure, but the exact nature of this structure(s) remains to be identified.

Direct evidence for the *in vivo* occurrence of slipped-strand mispairing in generating the instabilities within TRS has not been obtained, hence other mechanisms may be involved. In *Escherichia coli*, the usual mechanisms for homologous recombination (Kowalczykowski *et al*., 1994; Camerini-Otero & Hsieh, 1995) seem to be ruled out, since expansions have been observed in *recA*$^-$ strains (Kang *et al*., 1995a, 1996; Ohshima *et al*., 1996a,c) and similar extents of deletions were observed in isogenic strains with a functional or deficient *recA* gene (Jaworski *et al*., 1995). However, it is possible that other pathways of recombination have some influence. For example, *in vitro* experiments observed pausing of DNA polymerases at TRS containing CTG·CAG or CGG·CCG (Kang *et al*., 1995b; Ohshima *et al*., 1996a,b) and increased recombination has been proposed to occur at stalled replication forks (Kuzminov, 1995). A number of studies have suggested an involvement of replication and recombination in the deletion of direct repeat sequences (Gordenin *et al*., 1992; Scearce & Masker, 1993; Trinh & Sinden, 1993; Kong & Masker, 1994; Tran *et al*., 1995; Bi & Liu, 1996). Another alternative is gene conversion, which could account for instabilities at TRS *via* mechanisms similar to those used to explain polymorphisms at minisatellites (Jeffreys *et al*., 1994).

To help identify the factors affecting the instability of TRS, studies have been initiated in *E. coli* (Kang *et al*., 1995a). Maintenance of long TRS in bacterial systems is difficult, since they are deleted readily to sequences of heterogeneous length (Kang *et al*., 1995a; Ohshima *et al*., 1996a,b; Shimizu *et al*., 1996). The orientation of the TRS with respect to the origin of replication was observed to affect expansions and deletions in *E. coli* (Kang *et al*., 1995a) and was explained by the involvement of intra-strand structures, since each strand of the TRS produces structures of different stability (Gacy *et al*., 1995; Gao *et al*., 1995; Yu *et al*., 1995a,b).

During our recent investigations on the cloning of all ten TRS in *E. coli* (Kang *et al*., 1995a; Ohshima *et al*., 1996a,b; Shimizu *et al*., 1996), we observed an influence of growth phase on the frequency of deletions in the TRS. This study provides further characterization of the stability of CTG·CAG repeats in *E. coli*, and demonstrates that the frequency of deletions is influenced by interactions between cellular growth characteristics and the plasmid-based repeats.

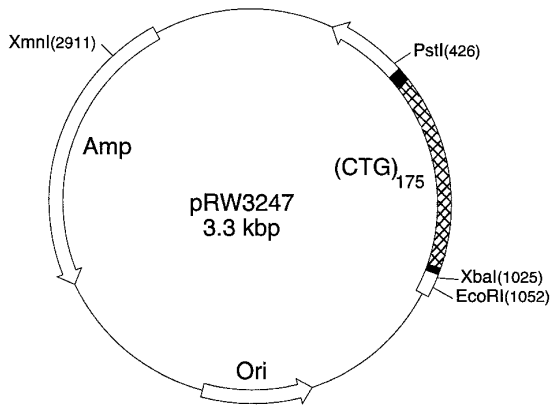
Results

Analysis of CTG repeats in plasmids in *Escherichia coli*

The maintenance of long TRS in bacterial systems is difficult and their stability is influenced by many factors (reviewed by Wells, 1996). This study utilizes plasmids containing different lengths of CTG·CAG repeats (referred to as CTG tracts) to further characterize their stability in *E. coli*. The plasmids harbored triplet repeat lengths of 17, 98 and 175 in the polylinker of pUC19 *Not*I (Herrero *et al*., 1990; Kang *et al*., 1995a), although the longest TRS was not completely homogeneous (Figure 1). Since we showed previously that the stability of long CTG repeats is dependent on their orientation with respect to the direction of replication (Kang *et al*., 1995a), each sequence was studied in both orientations.

The instability of triplet repeats in plasmids was analyzed under a variety of growth conditions. Plasmids were purified from the experimental cultures and digested with appropriate restriction enzymes, generally *Eco*RI and *Pst*I. Alterations in the length of the fragment containing the CTG tract were determined by polyacrylamide gel electro-

A



B

```
390                    NotI
  gccagtgccaagctGGGCGGCCGCAAGCTtgcatgcctgcaggtcggCCTCAGCCTG
  cggtcacggttcgaCCCGCCGGCGTTCGAacgtacggacgtccagccGGAGTCGGAC

  GCCGAAAGAAAGAAATGGTCTGTGATCCCCC(CAG)nCATTCCCGGCTACAAGGACg
  CGGCTTTCTTTCTTTACCAGACACTAGGGGG(GTC)nGTAAGGGCCGATGTTCCTGc

                                          NotI
  actctagaggatccccgggtaccgagctcgaattCGCGCGGCCGCAATTcgtaatca
  tgagatctcctaggggcccatggctcgagcttaaGCGCGCCGGCGTTAAgcattagt
```

Figure 1. A diagram of plasmids containing the CTG triplet repeats. A, pRW3247 that contains the $(CTG)_{175}$ on the lagging strand template for replication, termed orientation II (Kang *et al*., 1995a); this sequence is not completely homogeneous but is actually $(GCT)_{27}ACT(GCT)_{40}ACT(GCT)_{106}$. pRW3248 contains exactly the same sequence, but the sequence between the *Not*I sites is reversed so that $(CTG)_{175}$ is on the leading strand template, termed orientation I. All DNA except that shown by the filled regions is derived from pUC18 *Not*I (Herrero *et al*., 1990). The approximate positions and direction of the origin of replication (Ori) and the gene encoding resistance to ampicillin (Amp) are shown. The TRS is shown by the cross-hatched box, and the filled area is human genomic DNA that arises from the cloning procedure. These plasmids do not have a functional *lacZ'* gene. The other plasmids used in this study contain different lengths of $(CTG)_n$ in the *Hin*cII site of pUC18 *Not*I: pRW3243 and pRW3244 contain $(CTG)_{17}$ in orientations II and I, respectively; pRW3245 and pRW3246 contain $(CTG)_{98}$ in orientations II and I, respectively. B, The TRS and the flanking DNA of the plasmids used in this study, showing the relationship between plasmids in orientations I and II; the sequence shown is that of plasmids in orientation II, with base-pair number 1 designated as for pUC18 (Yanisch-Peron *et al*., 1985). The bases shown in lower case, and the remainder of the sequence that is not shown, correspond to the sequence of pUC18. Sequences shown in capitals around the *Not*I site were constructed in pUC18 *Not*I (Herrero *et al*., 1990). Plasmids in orientation I contain exactly the same sequences, except that the region between the *Not*I sites is in the opposite orientation. The subscript *n* corresponds to the number of triplet repeats.

phoresis. Numerous plasmids with different lengths of the fragment containing TRS were characterized by DNA sequencing and, in all cases, the length changes were due to an insertion or deletion of unit lengths (3 bp) of TRS (Kang *et al*., 1995a, 1996; Ohshima *et al*., 1996a; Shimizu *et al*., 1996; and unpublished results); no mutation was found in any other region of the plasmids. Tracts of TRS have been observed to have a faster mobility than random sequence DNA on polyacrylamide gels (Chastain *et al*., 1995; and unpublished results). However, since the gel mobility of a fragment containing TRS is proportional to its size, the mobilities of fragments on polyacrylamide gels can be used to estimate the lengths of the CTG tracts.

Plasmids containing long CTG repeats have heterogeneous lengths due to errors incurred during bacterial growth

Growth of *E. coli* after transformation with vectors containing $(CTG)_{175}$ showed that plasmids isolated from separate cultures contained TRS of heterogeneous length (Kang *et al*., 1995a; data not shown). Such length alterations could be caused by (1) some process during transformation (or some other event taking place before the first cell division) or (2) a high frequency of errors incurred during DNA metabolism throughout the growth period. To differentiate between these two possibilities, we transformed *E. coli* HB101 with plasmids containing CTG repeats, grew the mixtures overnight and then diluted them so that fresh cultures would grow from an average of one cell. Assuming that only one plasmid was introduced into each cell upon transformation (Hanahan *et al*., 1991), faithful DNA replication would produce a homogeneous population of plasmids in each cell even after further growth of the culture.

Plasmids were isolated from overnight growth of liquid cultures that had begun from one cell, digested with *Eco*RI and *Pst*I and analyzed by polyacrylamide gel electrophoresis. For pRW3247 ($(CTG)_{175}$ in orientation II), cultures that had begun from a single cell contained major bands of defined, but different, CTG tract lengths (Figure 2). Also, there were minor bands in the DNA samples, and a greater number of bands was observed for plasmids containing $(CTG)_{175}$ in orientation II compared with those with $(CTG)_{175}$ in orientation I (data not shown). In similar experiments, more bands were produced from pRW3247 when the dilutions to a single cell were made from cultures that had been allowed to grow for more generations (labeled early stat in Figure 2) compared with those in exponential growth phase (labeled late log in Figure 2). When these bacterial mixtures were further subcultured, the deletion products in each culture became a larger proportion of the purified plasmid, and new deletion products also appeared (lanes numbered 2 in Figure 2). The starter cultures used in these experiments were grown directly from the transformation mixtures; similar results were obtained if the transformation mixtures were

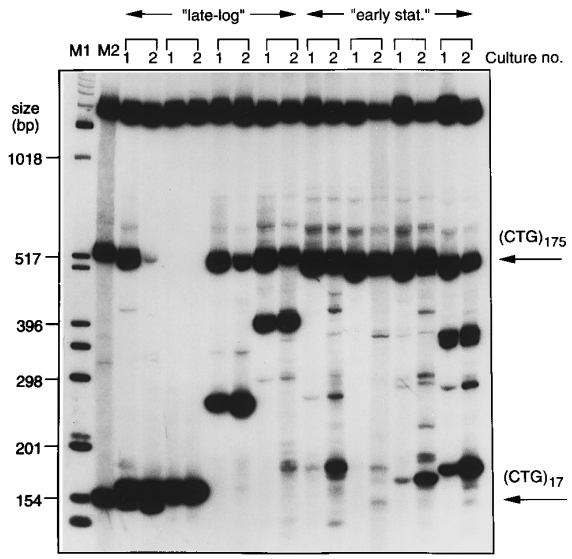


Figure 2. Analysis of the stability of plasmids containing $(CTG)_{175}$ when grown from a single cell of *E. coli* HB101. Cells were transformed with purified supercoiled monomer of pRW3247 ($(CTG)_{175}$ in orientation II), grown overnight in LB containing ampicillin (100 µg / ml) at 37°C, and at an A_{600} of 1.1 (late log) or 1.6 (early stationary), the culture was diluted 10^9-fold into fresh medium. (Since this dilution factor is greater than the number of viable cells in stationary phase cultures (Table 1), any observed growth from these inoculations must have begun from only one cell.) These cultures (number 1s) were grown into stationary phase (24 hours) and then subcultured into fresh medium with a dilution factor of 10^7, to give cultures numbered 2. The second cultures were grown for 24 hours and then harvested. DNA from all cultures was purified, digested with *Eco*RI and *Pst*I, labeled with ^{32}P using *E. coli* Klenow DNA polymerase I, and electrophoresed through a 5% polyacrylamide gel in TBE buffer. The samples are shown, with numbers 1 and 2 indicating the first and second growth cultures, respectively. M1 is marker DNA (1 kb ladder, GlBCO BRL); the sizes of these bands are shown at the side of the Figure. M2 is a marker containing a mixture of pRW3244 and pRW3248 (plasmids with $(CTG)_{17}$ and $(CTG)_{175}$, respectively) that had been digested with *Eco*RI and *Pst*I and labeled with radioactivity. Note that fragments of DNA containing TRS migrate faster than expected from their known size, in agreement with other studies (Chastain *et al*., 1995; and unpublished results).

first grown overnight on a plate and a single colony was used for the starter culture (data not shown).

Hence, the deletions in plasmid CTG repeat sequences are due to the unfaithful propagation of the CTG tract. The wide variety of TRS lengths from single cells suggests that transformation (or some event taking place soon afterwards) may act to promote deletion events. However, the continuous appearance of a variety of different lengths of TRS at long growth times showed that deletion events were taking place throughout the bacterial growth period.

Deletions of long CTG repeats are influenced by the growth phase

A high frequency of deletions within TRS in plasmids during long-term growth of *E. coli* cultures was observed previously (Jaworski *et al*., 1995; Kang *et al*., 1995a; Rosche *et al*., 1996; Shimizu *et al*., 1996). Since the experiments discussed above showed that the growth phase at which subculturing occurred had an influence on the deletion events, we analyzed the extent of deletions after long-term growth of bacteria maintained in exponential phase or allowed to go through stationary phase approximately every 20 to 25 generations. *E. coli* HB101 was transformed with plasmids harboring CTG repeats, then each transformation mixture was halved and grown in two identical liquid cultures. One of the cultures was maintained in exponential growth by dilution into fresh medium at $A_{600} = 0.7$, and the other was allowed to grow into stationary phase before dilution into fresh medium.

At each dilution, the cells from a portion of the culture were harvested, and the plasmids were purified and analyzed by restriction mapping. For pRW3247 ($(CTG)_{175}$ in orientation II), the deletions accumulated more readily in cultures that passed through stationary phase (Figure 3A). Similar results were obtained for bacteria harboring pRW3248 ($(CTG)_{175}$ in orientation I: see Figure 6 for representative gels of cultures allowed to pass through stationary phase). Quantification of the data from Figure 3A showed that about 1% of plasmid contained the full-length CTG repeat by the third stationary phase (60 generations). By comparison, there was 24% of full-length plasmid in the DNA isolated after 100 generations of exponential growth of *E. coli*. Similar experiments performed with pRW3247 and pRW3248 in *E. coli* HB101 were quantified (Figure 3B). For both plasmids, there was an increased accumulation of deletions for cultures that passed through stationary phase. Bacteria harboring pRW3248 amassed deletions more slowly than those harboring pRW3247 under all incubation conditions. Note that the pRW3247 samples were deleted more quickly in the experiments shown in Figure 3A compared with Figure 3B, but in both cases there was faster accumulation of deletions for samples passing through stationary phase. The trend of the data was the same, but the two experiments had different amounts of deletions after specific numbers of generations. These deviations are often observed in different experiments and illustrate the random nature of deletion events.

In summary, these results confirm our earlier observations that an increased rate of deletions occurred within CTG tracts in orientation II (Kang *et al*., 1995a). Also, these experiments show that the deletions within the CTG repeats were markedly influenced by the bacterial growth phases; those bacteria maintained in exponential growth phase had a lower extent of deleted products compared

with those that were allowed to go through stationary phase.

Deletions within CTG tracts are not influenced by *recA*

The experiments shown above were performed in *E. coli* HB101, a strain with a mutation in its *recA* gene (*recA13*: Sambrook *et al*., 1989). The stability of pRW3247 and pRW3248 was examined during long-term growth in *E. coli* RR1, a strain that is isogenic to HB101 except that it contains a functional *recA* gene (Sambrook *et al*., 1989). Similar extents of deletions were observed in both *E. coli* HB101 and RR1 (data not shown). Thus, in agreement with previous observations (Jaworski *et al*., 1995), *recA* is not required to generate deletions in TRS.

Cells containing long CTG repeats in orientation II exhibit a reduction in viability upon entering stationary phase

The viability of *E. coli* HB101 harboring the plasmids with CTG repeats was determined at various times of growth by counting the number of colonies on LB plates containing ampicillin. At mid-exponential growth (A_{600} = 0.6), *E. coli* had viability counts of 1.4×10^8 to 2.0×10^8 cells / ml for all plasmids studied (data not shown). Cultures grown into stationary phase (26 hours, A_{600} of 3.0 to 4.0) had viability counts of approximately 4×10^8 cells / ml for bacteria harboring each of the plasmids except pRW3247 (Table 1); these measurements are in agreement with earlier studies on the influence of plasmids on the viability of *E. coli* grown into stationary phase (Kolter *et al*., 1993). Control experiments with bacteria harboring pUC19 *Not*I had a viability similar to those of the shorter CTG repeats at all growth times (data not shown). For pRW3247 ($(CTG)_{175}$ in orientation II), the culture at 26 hours had a viability count of only 0.22×10^8 cells / ml, an order of magnitude lower than for the other plasmids containing CTG repeats. Cells harboring pRW3247 showed a further decline in viability at longer growth times (data not shown). Plating of an aliquot of all cultures on LB plates without antibiotic showed that the total number of cells was equivalent to the viability count in the presence of ampicillin (data not shown). Therefore, the decreased viability of stationary phase cells

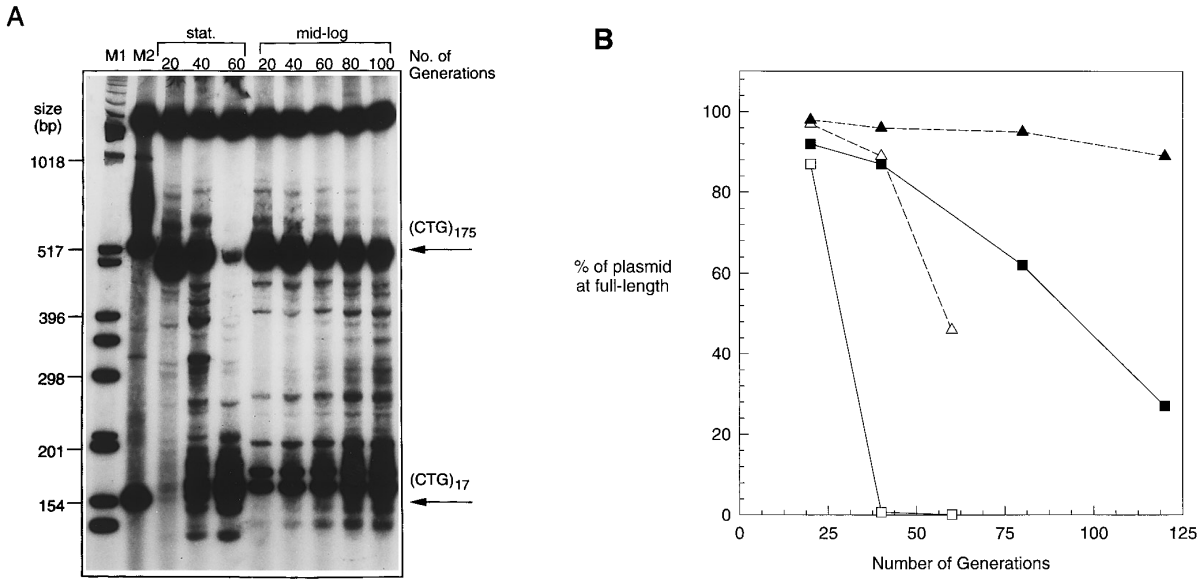


Figure 3. The extent of deletions in plasmids containing $(CTG)_{175}$ during long-term growth of *E. coli* HB101 is influenced by the growth phase at subculturing. A, *E. coli* HB101 was transformed with purified supercoiled monomer of pRW3247 ($(CTG)_{175}$ in orientation II) and incubated in LB containing ampicillin (100 μg / ml) at 37°C. The bacterial cells were grown to stationary phase (A_{600} of 1.7 to 2.5) or mid-logarithmic phase (A_{600} of 0.7 to 1.0) and then subcultured into fresh medium with a dilution factor of 10^6. At each subculturing, cells were harvested and the pellets stored at − 80°C. DNA from all cultures was purified, digested with *Eco*RI and *Pst*I, labeled with ^{32}P using *E. coli* Klenow DNA polymerase I, and electrophoresed through a 5% polyacrylamide gel in TBE buffer. The samples are shown, with the numbers above each lane indicating the approximate number of generations at the time of harvesting. Markers M1 and M2 were as described for Figure 2. B, Quantification of deletions in pRW3247 and pRW3248 during long-term growth. Data from long-term growth experiments that were subcultured at stationary or mid-exponential phase were quantified by densitometry. The data are not from experiment A, but the only difference is that the bacterial cells were diluted by a factor of 10^7 upon subculturing. The extent of full-length plasmid ($(CTG)_{175}$ was determined as the proportion to all DNA of this size or smaller. Reproducibility of these measurements was estimated to be ± 1%. Squares, pRW3247; triangles, pRW3248; open symbols, samples that were subcultured in stationary phase; filled symbols, samples that were subcultured at mid-exponential phase.

Table 1. *E. coli* HB101 containing pRW3247 have a reduced viability in stationary phase

Plasmid	Length of CTG repeat	Orientation	Viability count[a] ($\times 10^8$/ml)	Standard deviation[b]
pRW3243	17	II	3.88	0.38 (3)
pRW3244	17	I	3.83	0.83 (4)
pRW3245	98	II	4.45	0.66 (3)
pRW3246	98	I	4.47	0.62 (3)
pRW3247	175	II	0.22	0.12 (4)
pRW3248	175	I	3.25	1.01 (4)

Viability measurements were made on stationary phase cultures (A_{600} = 3.0 to 4.0). At mid-logarithmic growth (A_{600} = 0.6), the viability counts for all of the cultures was 1.4×10^8 to 2.0×10^8/ml.

[a] Total number of viable cells was the same on LB plates with and without ampicillin.

[b] The number of times the experiment was performed are in parentheses.

harboring pRW3247 cannot be explained by an altered sensitivity to antibiotic.

Thus, the length of the CTG repeat has no influence on cell viability except for those cells in stationary phase with $(CTG)_{175}$ in orientation II.

CTG repeats in plasmids affect the entry of *E. coli* into exponential growth phase

The growth rates of *E. coli* harboring plasmids that contained various tracts were analyzed. *E. coli* HB101 was transformed with plasmids harboring CTG repeats and the growth of the liquid cultures was monitored spectrophotometrically at A_{600}; pUC19 *Not*I with no CTG repeat (Herrero *et al*., 1990; Kang *et al*., 1995a) was transformed to act as a control.

Growth curves for cultures containing pRW3244, pRW3247, pRW3248 and pUC19 *Not*I are shown in Figure 4. The semi-log plot of the data (Figure 4, inset) shows that the rate of cell growth during exponential phase was similar for all plasmids studied. However, the entry into exponential growth was delayed for cells harboring plasmids with CTG repeats, and this is shown more clearly on the large graph, which contains linear scales on both axes. Since the transformations were performed with a large excess of cells to DNA molecules (estimated to be at least a factor of 10^3), at least ten generations of growth were required before an increase in A_{600} was detected. This experimental protocol means that the lag period of the culture was over by the time an increase in absorbance was detected. However, since the exponential growth rate of all cultures was similar, extrapolation of the curves down to the first doubling event shows that the cultures differed only in their lag period. The length of the lag phase was proportional to the size of the triplet repeat; cells harboring plasmids with $(CTG)_{175}$ had a longer lag phase than cells harboring plasmids with $(CTG)_{17}$ (Figure 4). Cells harboring plasmids with $(CTG)_{98}$ had an intermediate length of lag phase (data not shown). For the greatest length of CTG repeats, there was an effect of the sequence orientation on the lag phase, since pRW3247 ($(CTG)_{175}$ in orientation II) had a longer lag phase than pRW3248 ($(CTG)_{175}$ in orientation I). Hence, there is an effect of the length and orientation of TRS in plasmids on the growth kinetics of *E. coli*, as manifest by the length of the lag phase.

Cultures beginning from smaller numbers of live bacterial cells have longer lag phases (Neidhardt

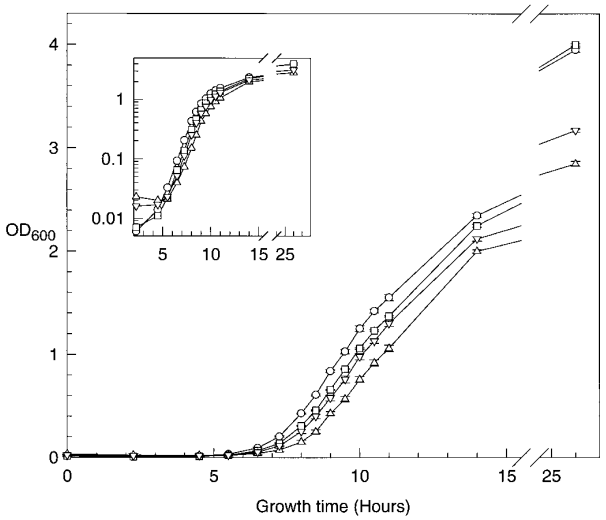


Figure 4. Growth curves of *E. coli* HB101 harboring plasmids containing CTG repeats. *E. coli* HB101 was transformed with purified supercoiled monomer of the designated plasmids, and grown in LB containing ampicillin (100 μg/ml) at 37°C. The growth of each culture was monitored by A_{600} readings at the times indicated. Each culture was grown in triplicate and contained approximately 3000 to 10,000 colony-forming units. The graph is plotted with linear axes to highlight the differences in the curves as the cultures begin to grow in exponential phase. The inset shows the traditional method of plotting growth curves using a semi-log axis for A_{600} measurements. Note that the slopes of the curves (the exponential phase growth rates) are similar in all cases. Error bars are shown in the main part of the Figure but are omitted from the inset to improve clarity. The symbols designate the following plasmids: circle, pUC19 *Not*I; square, pRW3244 ($(CTG)_{17}$ in orientation I); triangle, pRW3247 $(CTG)_{175}$ in orientation II; inverted triangle, pRW3248 ($(CTG)_{175}$, orientation I).

et al., 1990). We observed similar effects for *E. coli* harboring plasmids with CTG repeats (data not shown). The varying lag phases for cells harboring the different plasmids could be caused by altered viability of cells in the cultures, but this is not supported by measurements of cell viability (Table 1). Also, in experiments comparing the growth rates of *E. coli* harboring different plasmids, approximately the same number of viable cells was added to the cultures to prevent any influence of differential cell number. Moreover, cells containing plasmids with deleted CTG repeats dramatically increased their proportion within a culture even though they were present in a very low proportion at the time of the initial inoculation. Thus, the observed effects on growth are due to the influence of CTG repeats on lag phase rather than due to experimental artifacts brought about by altered viability of cells containing different plasmids.

Simulations of the growth of *E. coli* containing a mixture of cell types with different growth characteristics

The observed effects of CTG repeats on the growth kinetics and viability of *E. coli* suggest that deleted products may accumulate because cells within the culture harboring deleted plasmids have a growth advantage over those containing nondeleted plasmids. Over long growth times, there would be a significant increase in the proportion of deleted plasmids, which would make the deletion events appear to occur at a higher frequency than their actual rate.

To test this proposal, we performed calculations to identify potential changes in a bacterial culture containing a mixture of cells with different growth characteristics. Presumably, the growth characteristics of bacterial cells in these experiments are affected by their constituent plasmids. In particular, the length of the CTG tract and its orientation in a plasmid influences the growth of the host. In our experiments, plasmids with many different CTG lengths were produced within each culture, and thus would give rise to many different growth characteristics. However, for simplicity, the calculations were performed only for cultures containing cells with two types of growth characteristics.

The simulations were conducted with a mixed culture which contained two types of cells at a ratio 99:1. Starting from 100 cells, the bacteria were allowed to undergo doubling with a specified lag period (equivalent to no growth), after which they doubled at a constant steady-state rate. Growth within the simulated cultures was stopped when the total number of cells reached 5×10^8, which was equivalent to the number of cells observed in our stationary phase cultures (Table 1). The proportions of each cell type at this stage were used in another cycle of calculated growth; this simulated long-term growth experiments in which the bacterial cultures were allowed to pass through stationary phase between each subculturing. Calculations were performed in which the cell type that began at 1% of the culture was allowed to have either a shorter lag phase (Figure 5A) or a faster rate of doubling (Figure 5B). Our calculations showed that differences in lag phases for the two cell types produced large changes in the proportions of the cells as the culture passed through a number of stationary phases. For example, if the differences in the lag periods was a factor of 2, the cells with the longer lag phase decreased their proportion of the culture from 99% to about 60% at the seventh culture (Figure 5A). The proportion decreased to this level at an earlier time if the ratio of lag phases of the two cell types was higher. If the cell types were given differences in their doubling time, dramatic changes in their ratio were observed; for doubling times of the cell types with ratios of 1:1.1, at the fifth culture the cell type with the slower doubling time decreased its proportion of the culture from 99% to about 20% (Figure 5B).

Our long-term growth experiments (Figure 3) that were maintained in exponential growth were simulated. Calculations were performed in which the cultures were allowed to go through 100 doublings (generations) after the lag phase (data not shown). For cultures containing cell types with different lag phases, there was no change to the composition of the culture after the lag phase. Consequently, at long generation times, these cultures did not have changes in their composition as dramatic as those shown in Figure 5A. This is in agreement with our experimental observation (Figure 3). For mixed cultures of cell types with different doubling times, simulations maintained in exponential growth continued to show changes in their composition at long generation times; these changes occurred at the same rate with respect to the number of generations as calculated for mixed cultures that passed through stationary phase upon reaching 5×10^8 cells (Figure 5B). This prediction, that the composition of mixed cultures does not depend on whether they are maintained in exponential growth or pass through stationary phase, is not in agreement with our experimental observations (Figure 3).

Our experimental observations cannot be explained completely by such simple growth simulations, as expected, since many factors influence the growth kinetics of a bacterial culture. However, these calculations suggest that the lag phase of the cells is a major determining factor in explaining our observations of deletions within plasmid-based CTG repeats during long-term growth.

CTG repeats in plasmids influence the growth of *E. coli*

To provide an experimental test of the predictions from the simulations, the rate of propagation

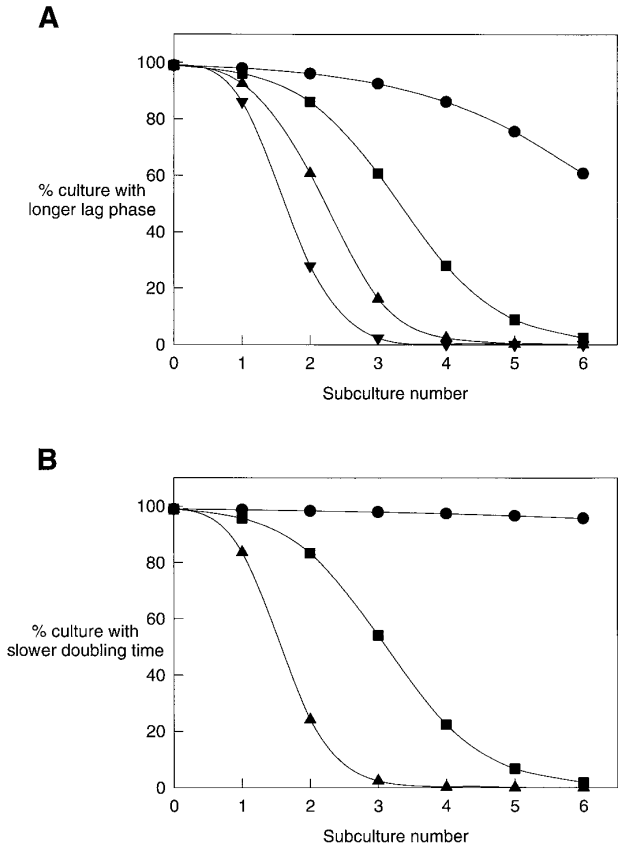


Figure 5. Simulations of the growth of *E. coli* containing a mixture of plasmids. Simulations of the growth characteristics of cultures containing two types of *E. coli* with different growth kinetics were performed. The starting-point was a mixed culture of 100 cells, which contained two cell types in the ratio of 99:1. The bacteria were allowed to undergo doubling with a specified time of lag phase (during which no growth occurred), and then they doubled at a constant steady-state rate until the cell density within the culture reached 5×10^8. The proportions of each cell type at this stage were used as the starting point for the next cycle of growth within the simulation, which continued with the same growth kinetics for each of the two sets of cells in the mixed culture. These cycles of growth were repeated a number of times in order to simulate our long-term growth experiments in which the bacterial cultures were allowed to pass through stationary phase between each subculturing. For each simulation, the bacteria were allowed to differ in the extent of their lag phase or in their growth-rate (doubling time). The cells with the growth disadvantage are shown as a proportion of the total culture; the subculture number refers to the cycle of simulated growth. A, Simulations were performed in which the cells at 1% of the initial culture were given a shorter lag phase; for computation purposes this was defined as 60 minutes. The two cell types had lag phases according to the following ratios: circle, 1:2; square, 1:3; triangle, 1:4; inverted triangle, 1:5. For example, a ratio of 1:2 designates lag phases of 60 minutes and 120 minutes for the two sets of cells. The doubling time of both cell types was set at a constant 60 minutes. B, Simulations were performed in which the cells at 1% of the initial culture were given a faster rate of doubling; for computation purposes this was defined as 60 minutes. The two cell types had doubling times according to the following ratios: circle, 1:1.017; square, 1:1.1; triangle, 1:1.2. For example, a ratio of 1:1.1 designates doubling times of 60 minutes and 66 minutes for the two sets of cells. The lag phase of both cell types was set at a constant 60 minutes.

of cells containing plasmids with different lengths of CTG repeats was determined. *E. coli* HB101 was transformed with one of two types of plasmids, each containing a different length or orientation of CTG repeat. Varying amounts of the two transformation mixtures were grown in the same liquid culture for many generations, and samples of the culture were removed at various times for analysis of the plasmids.

In experiments where the cultures passed through stationary phase every 25 generations (Figure 6A), bacteria harboring pRW3244 $((CTG)_{17}$ in orientation I) were propagated at a faster rate than those containing pRW3248 $((CTG)_{175}$ in orientation I); pRW3244 increased its proportion of the culture from 33% to 90% between the first and third growth cycles (compare generations 25 with 75 for lanes labeled 1:1). Experiments with pRW3248 in excess at the beginning of the inoculations produced similar observations, with pRW3244 increasing its proportion of the culture from 10% to 65% between the first and third growth cycles (compare generations 25 with 75 for lanes labeled 1:5). If similar cultures were maintained in exponential growth through the same number of generations, no change in the proportions of each set of bacteria within the cultures was detected (data not shown). Thus, *E. coli* harboring plasmids that are identical except for the length of their CTG repeat are propagated at different rates, but only if the bacteria pass through stationary phase. These experiments show that the growth advantage obtained with bacteria harboring plasmids containing shorter CTG repeats is because of their reduced lag phase, as simulated above (Figure 5).

In a similar experiment (Figure 6B), the growth rates of bacteria harboring pRW3248 $((CTG)_{175}$ in orientation I) were compared with those containing pRW3247 $((CTG)_{175}$ in orientation II). To differentiate between the two plasmids in these cultures, analyses were performed with *Xmn*I, which digests the vector sequence, and *Xba*I, which cleaves the polylinker. Since the polylinker (and the TRS) are in the reverse orientation in pRW3247 and pRW3248, an *Xmn*I + *Xba*I digest of these plasmids produces different patterns. This experiment showed that almost no $(CTG)_{175}$ in orientation II remained (<3%) by the third stationary phase of the growth culture. Although a small proportion of plasmids purified from the culture were derived from pRW3247, these contained a shorter CTG repeat (Figure 6B). The only difference between the two plasmids in the bacteria is the orientation of their TRS. Thus, there is an intrinsic property of $(CTG)_{175}$ in orientation II that makes it more difficult for the bacteria to grow.

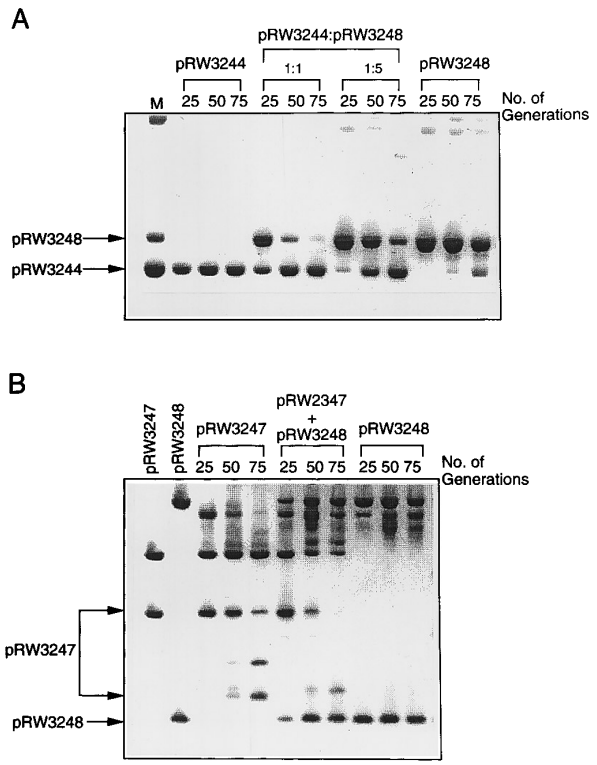


Figure 6. Comparison of the rate of growth of *E. coli* containing plasmids with different CTG repeats. *E. coli* HB101 was transformed with purified supercoiled monomer of the designated plasmids and grown separately, or in combination with other cells as indicated, in LB containing ampicillin (100 μg/ml) at 37°C. When cultures reached stationary phase (A_{600} of 2.1 to 3.0) they were subcultured into fresh medium, with a dilution factor of 10^7. Cells from each growth were harvested and the pellets were stored at −80°C. DNA from all cultures was purified and analyzed as described in Materials and Methods. The numbers above each experimental sample indicate the number of generations at the time of harvesting, with 25 generations being equivalent to each cycle of growth into stationary phase. A, Growth of pRW3244 ($(CTG)_{17}$ in orientation I), pRW3248 ($(CTG)_{175}$ in orientation I) and a mixture of the two plasmids. Cultures containing both plasmids were prepared with two starting amounts of pRW3244 (approximately 1:1 and 1:5 for pRW3244:pRW3248). Supercoiled DNA was electrophoresed through a 1.5% agarose gel in TBE buffer. M is a marker containing supercoiled pRW3244 and pRW3248. B, Growth of pRW3247 ($(CTG)_{175}$ in orientation II), pRW3248 ($(CTG)_{175}$ in orientation I)) and a mixture of the two plasmids. DNA from all cultures was purified, digested with *Xmn*I and *Xba*I, and electrophoresed through a 1.5% agarose gel in TBE buffer. Samples of supercoiled pRW3247 and pRW3248 were processed in the same manner to serve as markers. Note that, because of the altered position of the *Xba*I site and the inverted orientation of the CTG repeat in the two plasmids, deletions within the TRS are observed from different fragments for pRW3247 and pRW3248. The arrows at the side designate the bands that contain the TRS for each plasmid; the bracket indicates that fragments arising from pRW3247 have different sizes due to deletions within the TRS.

Alterations of the growth rates of *E. coli* produce different rates of deletions of CTG repeats

The growth states of *E. coli* profoundly alter the observed extent of deletions in CTG tracts. Hence, we investigated the effect of various growth conditions on the stability of the repeats. Bacteria were transformed with either pRW3247 or pRW3248 and long-term growth experiments were performed as described above (see Figures 3 and 6). The cultures were grown in liquid culture under three types of environmental conditions: (1) incubation temperature of 30°C, shaking rate of 250 rpm; (2) incubation temperature of 37°C, shaking rate of 250 rpm; (3) incubation temperature of 37°C, shaking rate of 175 rpm. At each subculturing, a portion of the bacteria was harvested and the plasmids purified and analyzed by restriction mapping. For pRW3247 and pRW3248, the greatest extent of deletions was observed upon incubation at 37°C and 250 rpm, conditions that allowed the fastest growth-rates (Figure 7). When the bacteria were incubated under conditions that allowed a slower growth-rate (lower temperature, slower shaking speed), plasmids with deleted CTG repeats were formed, but these did not take over the culture to the same extent as those inocula at 37°C and 250 rpm.

Bacterial growth is influenced by the medium used for the cultures, and is slower in minimal medium compared with LB. In experiments similar to those described here, bacteria harboring plasmids containing long CTG tracts showed a decreased extent of deletions when grown in minimal medium compared with those grown in LB (unpublished results).

Thus, bacteria that harbored deleted plasmids did not gain as much of a growth advantage under slower growth-rate conditions. Despite complexities introduced by cell physiology changes at different growth conditions, these observations are consistent with our proposal that the accumulation of deleted products over many generations of growth is due to the growth advantage of bacteria harboring deleted plasmids.

Discussion

E. coli is an effective model system for the study of the molecular mechanisms of expansion and deletion of trinucleotide repeats (Wells, 1996). Deletions were frequently observed to occur in an unpredictable fashion in long CTG tracts in plasmids, producing different lengths of TRS from an initially homogeneous sample. As found previously (Kang *et al*., 1995a), we observed an increased frequency of deletions from plasmids containing longer TRS, especially when the CTG strand was the lagging strand template (orientation II). The stability of CGG repeats in plasmids was influenced also by their length and orientation (Shimizu *et al*., 1996). Deletion events occurred

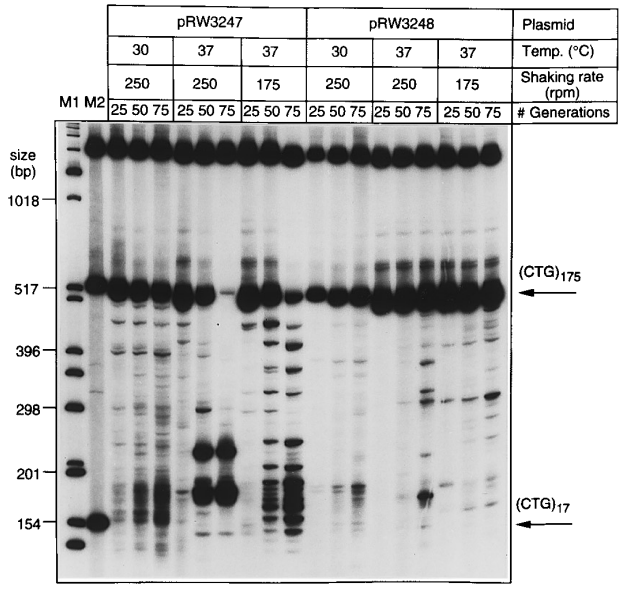


Figure 7. The effect of growth conditions of *E. coli* on the stability of $(CTG)_{175}$ in plasmids. *E. coli* HB101 was transformed with purified supercoiled monomer of pRW3247 ($(CTG)_{175}$ in orientation II) or pRW3248 ($(CTG)_{175}$ in orientation I) and grown in LB containing ampicillin (100 μg/ml). For each plasmid, three cultures were grown under different conditions: 30°C, shaking at 250 rpm; 37°C, shaking at 250 rpm; 37°C, shaking at 175 rpm. At stationary phase (A_{600} of 1.7 to 2.5), the mixtures were subcultured with a dilution of 10^7 and grown for 24 hours under the same conditions. At each subculturing, cells were harvested and the pellets stored at −80°C. DNA from all cultures was purified, digested with *Eco*RI and *Pst*I, labeled with ^{32}P using *E. coli* Klenow DNA polymerase I, and electrophoresed through a 5% polyacrylamide gel in TBE buffer. The samples are shown with the numbers directly above the samples indicating the approximate number of generations at the time of harvesting, with 25 generations being equivalent to each cycle of growth into stationary phase. The growth temperature and shaking rate of each sample were as indicated. Markers M1 and M2 were as described for Figure 2.

continuously during culture incubation and were particularly prominent at extended growth times. Also, the conditions of growth were critical and the extent of deletions was influenced by the bacterial growth phases, being greatly enhanced when the cells passed through stationary phase. To our knowledge, these are the first observations showing the direct influence of a plasmid-based DNA sequence or conformation on the growth characteristics of the host cell.

Several other factors influence the stability of long CTG repeats in plasmids in *E. coli*. Dramatic effects have been observed in different strains, with an enhancement of triplet repeat stability in strains with a deficient methyl-directed mismatch repair system (Jaworski *et al*., 1995) or a functional single-stranded binding protein (Rosche *et al*., 1996). Different stabilities for TRS have been observed when they are cloned at different plasmid locations (Kang *et al*., 1995a; Shimizu *et al*., 1996; unpublished results). All of these observations may be explained by altered growth characteristics, or by the action of specific gene products on the TRS. Taken together, these data indicate that the stability of triplet repeats in *E. coli* is a complex process involving many factors. We expect that the ability of the TRS to form unusual DNA structures (Pearson & Sinden, 1996; Wells, 1996; unpublished results), along with their specific interactions with cellular proteins, are critical. Regardless of the mechanism responsible for deletions, the length changes within TRS produce *E. coli* that are propagated with different growth characteristics (i.e. different lag phases but similar rates of exponential growth).

Interactions between *E. coli* growth characteristics and the stability of CTG repeats

A culture of *E. coli* is under constant pressure to evolve due to the production of mutant cell types, even in a fixed environment (Marr, 1991). In our experiments, the accumulation of deleted products over many generations is due to the growth advantage obtained by bacteria harboring deleted plasmids. Analysis of the growth of bacterial cells harboring plasmids with different CTG repeats showed differences in their lag phase; cells had a longer lag phase if they contained plasmids with longer repeats, or with repeats in orientation II. The composition of the culture changed most dramatically when it passed through a number of stationary phases, and our simulations of bacterial growth rates support the notion that the main determinant for a growth advantage was due to the effects at lag phase.

Our data show that, upon inoculation of fresh cultures, those cells containing the deleted products reached exponential growth earlier than those with non-deleted CTG tracts. During growth that cycled through stationary phase a number of times, cells containing plasmids with deletions increased their proportion of the culture. The observed frequency of deletions was proportional to the actual rate, but it was artificially increased due to the growth advantage of deleted products. High rates of deletions were not seen for plasmids with short repeat lengths ($(CTG)_{17}$) or in orientation I) because deleted products did not give an advantage to their host cell. The extremely high rate of deletions of pRW3247 most likely occurred because of the reduced viability of stationary phase cells that contained plasmid with the full-length CTG repeat; the smaller number of cells would give rise to a longer lag phase (Neidhardt *et al*., 1990) This reduced viability may actually have promoted deletion events as the cells endeavored to survive.

Although the factors causing deletions are not well defined, it is clear that the growth characteristics of *E. coli* influence the observed extent of

deletions in CTG tracts. Those conditions allowing faster growth-rates generated a higher accumulation of deleted products (Figure 7 and unpublished results). These observations may be explained by a greater growth advantage for deleted products at increased growth-rates, or there may be some effects due to the changes in cellular physiology that occur under different growth conditions.

Bacterial cells grown in batch cultures have complex growth patterns, particularly in rich media (Åkerlund *et al*., 1995). Changes in cellular growth occur in a gradual fashion and are dictated by the environmental conditions. Therefore, the definitions of bacterial growth phases are somewhat arbitrary and are not necessarily the same for different experimental procedures (Kolter *et al*., 1993). In general, the onset of stationary phase is indicated when there is no overall growth of the bacteria within the batch culture.

Stationary phase cultures continue to undergo changes that alter the metabolism within the cell. Growth advantages were previously observed for ten-day old bacterial cultures when mixed with one-day old cultures (Zambrano & Kolter, 1993; Zambrano *et al*., 1993), which were attributed to specific mutations. Certain mutations in *rpoS*, the gene coding for the stationary phase-specific sigma factor of RNA polymerase, also confer advantages to stationary phase cells (Hengge-Aronis, 1993). Cell types with a growth advantage will be favored within any culture because of their optimal use of nutrients; the pressures for such selections are expected to arise from changes in bacterial cell physiology (Marr, 1991). In this study, the selection pressures are due to the presence of the TRS, and the growth advantages must arise from an intrinsic property of the TRS (sequence or structure). These novel observations suggest that specific changes in the TRS are induced in stationary phase.

Stationary phase bacteria have a different physiology compared with those in exponential growth (Siegele & Kolter, 1992; Kolter *et al*., 1993). A wide variety of novel proteins are induced upon entry into stationary phase; possibly, one of these promotes deletion events. Alternatively, if the deletion events are associated with replication, the formation of intermediates may occur as the cells leave stationary phase and begin replication again. Many proteins play important roles in replication, and the cellular concentrations of some of these change under different growth conditions. Elucidation of the details of the mechanism(s) by which stationary phase influences deletion events requires further analysis.

Factors influencing deletion frequency

Like all simple repeating sequences cloned in bacteria so far, TRS delete readily. The extent of deletions increases with the length of the repeat, making it difficult to clone long TRS that are stable. Expansion of some TRS has been observed in *E. coli*, but technical intervention is required to increase the proportion of expanded product (Kang *et al*., 1995a, 1996; Ohshima *et al*., 1996c). Presumably, the extensive manipulation is necessary because the growth advantage of *E. coli* harboring the original (shorter) plasmid will not allow the expanded products to build up to a high proportion of the total plasmids.

Factors influencing the stability of TRS include: (1) host strain. A number of genetic factors are known to influence TRS stability (Jaworski *et al*., 1995; Rosche *et al*., 1996), and, in general, we observe fewest deletions in *E. coli* strains HB101 and SURE. (2) The location of the TRS in the vector influences the extent of deletions (Kang *et al*., 1995a; Shimizu *et al*., 1996; and unpublished results). These effects are probably due to the TRS disturbing processes taking place on the DNA, such as replication, recombination, repair or transcription. (3) Orientation of the TRS. The stability of long CTG and CGG repeats is different in their two orientations (Kang *et al*., 1995a; Shimizu *et al*., 1996). This is possibly due to the two strands forming folded structures of different stabilities (Gacy *et al*., 1995; Smith *et al*., 1995; Yu *et al*., 1995a; Pearson & Sinden, 1996). (4) Vectors. Cloning of the TRS in different vectors alters their stability (Kang *et al*., 1995a; Shimizu *et al*., 1996). This is due, at least in part, to vector factors such as copy number, but may be due to the locations of the TRS in various, vectors. (5) Homogeneity of TRS. Interruptions within trinucleotide repeats have important consequences for triplet repeat diseases due to their influence on the stability of the repeating sequence (Chung *et al*., 1993; Eichler *et al*., 1994). The longest triplet repeat sequence used in these experiments contains two non-CTG units (Kang *et al*., 1995a). These polymorphisms reduce the greatest length of homogeneous triplet repeat, and undoubtedly act to increase the stability of the sequences. (6) Growth conditions influence the stability of TRS in plasmids contained in *E. coli*. To reduce the extent of deletions, it is beneficial to grow the bacterial cultures at slower rates. If it is necessary to grow the cultures for many generations, the extent of deletions is reduced by maintaining the cultures in exponential growth phase.

Relationship of experiments in *E. coli* to human triplet repeat diseases

Investigations into the stability of repeating sequences in *E. coli* are important from a practical standpoint, since the cloning of TRS in bacteria enables their subsequent manipulation in more complex systems. Such studies have the potential to provide models of the molecular mechanisms of processes found in humans. Human sequences contain much heterogeneity at TRS (Ashley & Warren, 1995; Sutherland & Richards, 1995), but expanded sequences are tolerated better than in

E. coli. At the moment, the reasons for such differences between repeat instability in humans and bacterial cells remain to be determined.

The propagation of triplet repeats in plasmids in *E. coli* is difficult because of frequent deletions within the repeat sequences. This study shows that this high instability relates to the growth conditions. At times of rich nutrition, such as growth in LB, the bacterial cell divides as fast as possible; thus, growth advantages associated with shorter triplet repeats will always be propagated. In multicellular organisms such as humans, the survival of the whole organism is paramount. Thus, the relationship between repeat instability and cellular growth that was found for *E. coli* in this study is unlikely to hold for the propagation of triplet repeats in humans.

Distribution of repeating sequences within natural genomes

Repeating DNA sequences (direct repeats, inverted repeats, mirror repeats) are widely dispersed in natural genomes (Charlesworth *et al*., 1994; Tautz & Schlötterer, 1994; Sutherland & Richards, 1995). These sequences can form a number of different non-*B* DNA structures (Sinden & Wells, 1992), which would be expected to have specific consequences (either good or bad) within the cell. There is an over-abundance of some sequences in eukaryotes, but prokaryotes do not contain repetitive DNA sequences of substantial length (Behe, 1995; Schroth & Ho, 1995; Karlin & Burge, 1996). This suggests that simple repeating sequences provide no beneficial functions in prokaryotes and that they have been discarded (deleted) throughout evolution. This study shows that the removal of repeating sequences may be a consequence of their effects on the cell. If it is assumed that simple repeating sequences confer a growth disadvantage for their host cell, then cells containing deletions would be propagated more efficiently. Hence, it is not surprising that they have been almost totally removed from the genome.

The growth advantages of *E. coli* containing deletions of TRS accounts for the observed high frequencies of deletions in bacterial systems. Studies in *E. coli* have provided numerous insights into the factors affecting stability of repeated DNA sequences, although it is clear that complete comprehension of the biology of the triplet repeat diseases requires studies in complex systems.

Materials and Methods

Plasmids and bacterial strains

The plasmids used in these experiments contained repeating (CTG)·(CAG) inserts and have been described elsewhere (Kang *et al*., 1995a). The (CTG)·(CAG) inserts may be designated (TGC)·(GCA) or (GCT)·(AGC), and are referred to as CTG for convenience. The orientation of this TRS with respect to the direction of replication within these plasmids has been defined (Kang *et al*., 1995a) as follows: plasmids containing the CTG sequence as the leading strand template are designated orientation I; plasmids containing the CTG sequence as the lagging strand template are designated orientation II. All plasmids used in this study are derivatives of pUC18 *Not*I (Herrero *et al*., 1990; Kang *et al*., 1995), with different lengths of $(CTG)_n$ cloned in the *Hin*cII site: pRW3243 and pRW3244 contain $(CTG)_{17}$ in orientations II and I, respectively; pRW3245 and pRW3246 contain $(CTG)_{98}$ in orientations II and I, respectively; and pRW3247 and pRW3248 contain $(CTG)_{175}$ in orientations II and I, respectively. It should be noted that the longest sequence is not completely homogeneous but is actually $(GCT)_{27}ACT(GCT)_{40}ACT(GCT)_{106}$. Figure 1 shows a map of pRW3247 with the restriction sites relevant to this study; pRW3248 is the same except that the sequence between the *Not*I sites is cloned in the reverse orientation.

Large amounts of plasmids were prepared by alkaline lysis of a one liter culture (LB containing 100 μg/ml ampicillin) grown to stationary phase and were purified by CsCl/ethidium bromide ultracentrifugation overnight (Sambrook *et al*., 1989). Plasmid identity was confirmed using restriction enzyme mapping and dideoxy sequencing of both strands with Sequenase (version 2.0). Purification of plasmids from 10 ml cultures was performed using the standard alkaline lysis miniprep procedure (Sambrook *et al*., 1989).

All cloning procedures and experiments were conducted in *E. coli* HB101. All DNA transformations of *E. coli* were performed by electroporation (Hanahan *et al*., 1991). To obtain plasmids for transformation without deletions, a sample of CsCl-purified DNA (about 2 μg) was electrophoresed on a 1.5% (w/v) agarose gel and, after staining of the gel with ethidium bromide (1 μg/ml), a section of the gel containing the full-length supercoiled monomer plasmid was excised. The DNA was recovered from the crushed gel slice by extraction with phenol, precipitated with sodium acetate and ethanol, and resuspended in sterile water. For each experiment, electroporations were performed with an excess of number of cells to DNA molecules so that, in general, transformation of a cell should have occurred by a single plasmid molecule (Hanahan *et al*., 1991). The bacteria were allowed to recover at 37°C in SOC medium (2% (w/v) Bacto-tryptone, 0.5% yeast extract (w/v), 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose, pH 7.0) for 30 to 45 minutes and then grown in liquid culture with antibiotic selection. To find the number of potential colony-forming units within each transformation, 20 μl of the transformation mixture was grown overnight at 37°C on LB-agar plates containing antibiotic.

Conditions of bacterial growth

The stability of plasmids containing various lengths of repeating CTG sequences was analyzed under a number of growth conditions. In the basic procedure, *E. coli* HB101 was transformed with the appropriate plasmid and an aliquot of this mixture was inoculated into glass tubes (150 mm length, 17 mm diameter) containing 10 ml of LB with ampicillin at 100 μg/ml. To provide an estimate of the number of viable cells in the transformation mixture, another aliquot was plated on LB with ampicillin at 50 μg/ml and grown overnight at 37°C. Incubations of the liquid cultures were continued for the required time at 37°C and a shaking rate of 250 rpm, unless indicated otherwise in the text. Growth of the

cultures was monitored by absorbance readings at 600 nm. In order to correct for deviations from linearity in A_{600} readings above 0.5, the measured absorbance readings were converted to the true values by use of a calibration curve specific for the spectrophotometer and strain of *E. coli*.

The bacteria were subcultured into fresh liquid medium as appropriate. The cells from each culture were harvested and kept at –80°C until the end of the experiment, and then all DNA mini-preps were performed at the same time. The number of generations of growth (n) for each sample to produce N_t cells from N_0 cells was estimated from the equation:

$$N_t / N_0 = 2^n$$

assuming an A_{600} of 0.6 corresponds to 2×10^8 cells / ml (Table 1). For example, if one cell is grown for 20 division times (generations) there will be approximately 10^6 progeny cells.

The viability of cultures containing the plasmids was determined at mid-exponential and stationary phases of growth. Aliquots of the cultures were diluted (4×10^4) and 50 µl of this solution was spread on LB-ampicillin plates. The number of colonies were counted after overnight incubation of the plates at 37°C, and the number of viable cells per ml of the original solution was calculated.

Analysis of DNA

Supercoiled plasmids were electrophoresed on 1.5% agarose gels in TBE buffer (90 mM Tris-borate, 1 mM EDTA, pH 8.3). Restriction digests were performed following the manufacturer's instructions, and the samples were analyzed by electrophoresis through 1.5% agarose or 5% polyacrylamide gels in TBE buffer. DNA samples electrophoresed through agarose gels were observed by staining of the gel with ethidium bromide (1 µg / ml) and visualized on a trans-illuminator. Gels were photographed through a red filter using Polaroid 55 film. DNA samples that were electrophoresed through polyacrylamide gels were first radio-labeled using the standard procedure of end-filling with *E. coli* DNA polymerase Klenow fragment incubated with [α-^{32}P]dATP (Sambrook *et al*., 1989); radioactive fragments were observed with autoradiography or by a phosphorimager (400S, Molecular Dynamics). Quantification of the gels was performed directly upon the phosphorimage.

Simulations of bacterial growth

Calculations were performed that simulated the bacterial growth in a culture containing a mixture of cells, each with different growth characteristics with respect to their steady state growth-rates or lag phases. For simplicity, in these calculations only two types of cells were included and we considered that no growth took place in the lag phase. After the specified time of lag phase, the cells were allowed to double at a constant steady-state rate. Thus, the total number of cells in the culture, X_T, was calculated according to the equation:

$$X_T = X_a(2^{tk_a}) + X_b(2^{tk_b})$$

where X_a and X_b are the initial number of cells for the two cell types, t is the time of growth (in minutes), and k_a and k_b are the growth-rate constants for the two cell types (in minutes^{-1}, since it is the reciprocal of the doubling time).

The simulations began with a mixed culture containing two types of cells at a ratio of 99:1. For all calculations, the cell type that began at 1% of the culture was given a growth advantage, either a shorter lag phase or a shorter doubling time. Growth within the culture was abruptly stopped when the number of cells reached a specified cell density. The proportions of each cell type at this stage were used in another cycle of calculated growth, in order to simulate our long-term growth experiments in which the bacterial cultures were allowed to pass through stationary phase between each subculturing. Calculations were performed for variations in the relative lag phases and steady-state growth rates for the two cell types. Calculations were performed to simulate long-term growth in exponential phase, in which the cultures were allowed to go through 100 doublings (generations) after the lag phase.

Acknowledgements

This work was supported by grants from NIH (GM52982) and the Robert A. Welch Foundation. We thank S. Kang, K. Ohshima, R. Iyer and P. Parniewski for helpful discussions, and D. Siegele for constructive suggestions regarding the manuscript.

References

Aaltonen, L. A., Peltomäki, P., Leach, F. S., Sistonen, P., Pylkkänen, L., Mecklin, J. P., Järvinen, H., Powell, S. M., Jen, J., Hamilton, S. R., Petersen, G. M., Kinzler, K. W., Vogelstein, B. & de la Chapelle, A. (1993). Clues to the pathogenesis of familial colorectal cancer. *Science*, **260**, 812–816.

Åkerlund, T., Nordström, K. & Bernander, R. (1995). Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. *J. Bacteriol.* **177**, 6791–6797.

Armour, J. A. & Jeffreys, A. J. (1992). Biology and applications of human minisatellite loci. *Curr. Opin. Gen. Dev.* **2**, 850–856.

Ashley, C. T. & Warren, S. T. (1995). Trinucleotide repeat expansion and human disease. *Annu. Rev. Gen.* **29**, 703–728.

Behe, M. (1995). An overabundance of long oligopurine tracts occurs in the genome of simple and complex eucaryotes. *Nucl. Acids Res.* **23**, 689–695.

Bi, X. & Liu, L. F. (1996). A replicational model for DNA recombination between direct repeats. *J. Mol. Biol.* **256**, 849–858.

Camerini-Otero, R. D. & Hsieh, P. (1995). Homologous recombination proteins in prokaryotes and eukaryotes. *Annu. Rev. Gen.* **29**, 509–552.

Caskey, C. T., Pizzuti, A., Fu, Y. H., Fenwick, R. G., Jr & Nelson, D. L. (1992). Triplet repeat mutations in human disease. *Science*, **256**, 784–789.

Charlesworth, B., Sniegowski, P. & Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eucaryotes. *Nature*, **371**, 215–220.

Chastain, P. D., Eichler, E. E., Kang, S., Nelson, D. L., Levene, S. D. & Sinden, R. R. (1995). Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. *Biochemistry*, **34**, 16125–16131.

Chung, M.-y., Ranum, L. P. W., Duvick, L. A., Servadio, A., Zoghbi, H. Y. & Orr, H. T. (1993). Evidence for

a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nature Genet.* **5**, 254–258.

Eichler, E. E., Holden, J. J., Popovich, B. W., Reiss, A. L., Snow, K., Thibodeau, S. N., Richards, C. S., Ward, P. A. & Nelson, D. L. (1994). Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nature Genet.* **8**, 88–94.

Fishel, R., Lescoe, M. K., Rao, M. R. S., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M. & Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell,* **75**, 1027–1038.

Fry, M. & Loeb, L. A. (1994). The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc. Natl Acad. Sci. USA,* **91**, 4950–4954.

Gacy, A. M., Goellner, G., Juranic, N., Macura, S. & McMurray, C. T. (1995). Trinucleotide repeats that expand in human disease form hairpin structures *in vitro*. *Cell,* **81**, 533–540.

Gao, X. L., Huang, X. N., Smith, G. K., Zheng, M. X. & Liu, H. Y. (1995). New antiparallel duplex motif of DNA CCG repeats that is stabilized by extrahelical bases symmetrically located in the minor groove. *J. Am. Chem. Soc.* **117**, 8883–8884.

Gordenin, D. A., Malkova, A. L., Peterzen, A., Kulikov, V. N., Pavlov, Y. I., Perkins, E. & Resnick, M. A. (1992). Transposon Tn5 excision in yeast: influence of DNA polymerases alpha, delta, and epsilon and repair genes. *Proc. Natl Acad. Sci. USA*, **89**, 3785–3789.

Hanahan, D., Jessee, J. & Bloom, F. R. (1991). Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol.* **204**, 63–113.

Hengge-Aronis, R. (1993). Survival of hunger and stress: the role of *rpoS* in early stationary phase gene regulation in *E. coli*. *Cell,* **72**, 165–168.

Herrero, M., de Lorenzo, V. & Timmis, K. N. (1990). Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in Gram-negative bacteria. *J. Bacteriol.* **172**, 6557–6567.

Jaworski, A., Rosche, W. A., Gellibolian, R., Kang, S. M., Shimizu, M., Bowater, R. P., Sinden, R. R. & Wells, R. D. (1995). Mismatch repair in *Escherichia coli* enhances instability of $(CTG)_n$ triplet repeats from human hereditary diseases. *Proc. Natl Acad. Sci. USA,* **92**, 11019–11023.

Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L. & Armour, J. A. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nature Genet.* **6**, 136–145.

Kang, S., Jaworski, A., Ohshima, K. & Wells, R. D. (1995a). Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nature Genet.* **10**, 213–218.

Kang, S. M., Ohshima, K., Shimizu, M., Amirhaeri, S. & Wells, R. D. (1995). Pausing of DNA synthesis *in vitro* at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *J. Biol. Chem.* **270**, 27014–27021.

Kang, S., Ohshima, K., Jaworski, A. & Wells, R. D. (1996). 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.* **258**, 543–547.

Karlin, S. & Burge, C. (1996). Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. *Proc. Natl Acad. Sci. USA,* **93**, 1560–1565.

Kolter, R., Siegele, D. A. & Tormo, A. (1993). The stationary phase of the bacterial life cycle. *Annu. Rev. Microbiol.* **47**, 855–874.

Kong, D. & Masker, W. (1994). Deletion between direct repeats in T7 DNA stimulated by double-strand breaks. *J. Bacteriol.* **176**, 5904–5911.

Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* **58**, 401–465.

Krontiris, T. G. (1995). Minisatellites and human disease. *Science,* **269**, 1682–1683.

Kuzminov, A. (1995). Instability of inhibited replication forks in *E. coli*. *Bioessays,* **17**, 733–741.

Leach, F. S., Nicolaides, N. C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomäki, P., Sistonen, P., Aaltonen, L. A., Nyström-Lahti, M., Guan, X.-Y., Zhang, J., Meltzer, P. S., Yu, J.-W., Kao, F.-T., Chen, D. J., Cerosaletti, K. M., Fournier, R. E. K., Todd, S., Lewis, T., Leach, R. J., Naylor, S. L., Weissenbach, J., Mecklin, J. P., Järvinen, H., Petersen, G. M., Hamilton, S. R., Green, J., Jass, J., Watson, P., Lynch, H. T., Trent, J. M., de la Chapelle, A., Kinzler, K. W. & Vogelstein, B. (1993). Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. *Cell,* **75**, 1215–1225.

Levinson, G. & Gutman, G. A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**, 203–221.

Lustig, A. J. & Petes, T. D. (1993). Genetic control of simple sequence stability in yeast. In *Genome Rearrangement and Stability* (Davies, K. E. & Warren, S. T., eds), vol. 7, pp. 79–106, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Marr, A. G. (1991). Growth rate of *Escherichia coli*. *Microbiol. Rev.* **55**, 316–333.

Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J. & Haworth, I. S. (1995). Hairpin properties of single-stranded DNA containing a GC-rich triplet repeat: $(CTG)_{15}$. *Nucl. Acids Res.* **23**, 1050–1059.

Mitchell, J. E., Newbury, S. F. & McClellan, J. A. (1995). Compact structures of $d(CNG)_n$ oligonucleotides in solution and their possible relevance to fragile X and related human diseases. *Nucl. Acids Res.* **23**, 1876–1881.

Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. (1990). *Physiology of the Bacterial Cell: A Molecular Approach*, Sinauer Associates, Inc., Sunderland, MA.

Ohshima, K., Kang, S., Larson, J. E. & Wells, R. D. (1996a). Cloning, characterization, and properties of seven triplet repeat DNA sequences. *J. Biol. Chem.* **271**, 16773–16783.

Ohshima, K., Kang, S., Larson, J. E. & Wells, R. D. (1996b). TTA·TAA triplet repeats in plasmids form a non-H bonded structure. *J. Biol. Chem.* **271**, 16784–16791.

Ohshima, K., Kang, S. & Wells, R. D. (1996c). CTG triplet repeats from human hereditary diseases are dominant genetic expansion products in *Escherichia coli*. *J. Biol. Chem.* **271**, 1853–1856.

Parsons, R., Li, G.-M., Longley, M. J., Fang, W.-h., Papadopoulos, N., Jen, J., de la Chapelle, A., Kinzler, K. W., Vogelstein, B. & Modrich, P. (1993). Hypermutability and mismatch repair deficiency in RER^+ tumor cells. *Cell,* **75**, 1227–1236.

Pearson, C. E. & Sinden, R. R. (1996). Alternative structures in duplex DNA formed within the

trinucleotide repeats of the myotonic dystrophy and fragile X loci. *Biochemistry*, **35**, 5041–5053.
Richards, R. I. & Sutherland, G. R. (1992). Dynamic mutations: a new class of mutations causing human disease. *Cell*, **70**, 709–712.
Rosche, W. A., Jaworski, A., Kang, S., Kramer, S. F., Larson, J. E., Giedroc, D. P., Wells, R. D. & Sinden, R. R. (1996). Single strand DNA binding protein enhances the stability of CTG triplet repeats in *Escherichia coli*. *J. Bacteriol*. **178**, 5042–5044.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Scearce, L. M. & Masker, W. (1993). Deletion between direct repeats in bacteriophage T7 gene 1.2. *Mutat. Res.* **288**, 301–310.
Schroth, G. P. & Ho, P. S. (1995). Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA. *Nucl. Acids Res.* **23**, 1977–1983.
Shimizu, M., Gellibolian, R., Oostra, B. A. & Wells, R. D. (1996). Cloning, characterization and properties of plasmids containing CGG triplet repeats from the FMR-1 gene. *J. Mol. Biol.* **258**, 614–626.
Siegele, D. A. & Kolter, R. (1992). Life after log. *J. Bacteriol.* **174**, 345–348.
Sinden, R. R. & Wells, R. D. (1992). DNA structure, mutations, and human genetic disease. *Curr. Opin. Biotechnol.* **3**, 612–622.
Smith, G. K., Jie, J., Fox, G. E. & Gao, X. (1995). DNA CTG triplet repeats involved in dynamic mutations of neurologically related gene sequences form stable duplexes. *Nucl. Acids Res.* **23**, 4303–4311.
Sutherland, G. R. & Richards, R. I. (1995). Simple tandem DNA repeats and human genetic disease. *Proc. Natl Acad. Sci. USA*, **92**, 3636–3641.
Tautz, D. & Schlötterer. (1994). Simple sequences. *Curr. Opin. Genet. Dev.* **4**, 832–837.
Thibodeau, S. N., Bren, G. & Schaid, D. (1993). Microsatellite instability in cancer of the proximal colon. *Science*, **260**, 816–819.
Tran, H. T., Degtyareva, N. P., Koloteva, N. N., Sugino, A., Masumoto, H., Gordenin, D. A. & Resnick, M. A. (1995). Replication slippage between distant short repeats in *Saccharomyces cerevisiae* depends on the direction of replication and the RAD50 and RAD52 genes. *Mol. Cell. Biol.* **15**, 5607–5617.
Trinh, T. Q. & Sinden, R. R. (1993). The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in *Escherichia coli*. *Genetics*, **134**, 409–422.
Warren, S. T. (1996). The expanding world of trinucleotide repeats. *Science*, **271**, 1374–1375.
Wells, R. D. (1996). Molecular basis of genetic instability of triplet repeats. *J. Biol. Chem.* **271**, 2875–2878.
Willems, P. J. (1994). Dynamic mutations hit double figures. *Nature Genet.* **8**, 213–215.
Yanisch-Peron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.
Yu, A., Dill, J. & Mitas, M. (1995a). The purine-rich trinucleotide repeat sequences $d(CAG)_{15}$ and $d(GAC)_{15}$ form hairpins. *Nucl. Acids Res.* **23**, 4055–4057.
Yu, A. D., Dill, J., Wirth, S. S., Huang, G., Lee, V. H., Haworth, I. S. & Mitas, M. (1995b). The trinucleotide repeat sequence $d(GTC)_{15}$ adopts a hairpin conformation. *Nucl. Acids Res.* **23**, 2706–2714.
Zambrano, M. M. & Kolter, R. (1993). *Escherichia coli* mutants lacking NADH dehydrogenase I have a competitive disadvantage in stationary phase. *J. Bacteriol.* **175**, 5642–5647.
Zambrano, M. M., Siegele, D. A., Almirón, M., Tormo, A. & Kolter, R. (1993). Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. *Science*, **259**, 1757–1760.

Edited by K. Yamamoto

(Received 7 May 1996; received in revised form 12 August 1996; accepted 28 August 1996)