Nucleic Acids Research, 1997, Vol. 25, No. 14 **2861–2868**

Transcription increases the deletion frequency of long CTG·CAG triplet repeats from plasmids in *Escherichia coli*

Richard P. Bowater[+], Adam Jaworski[§], Jacquelynn E. Larson, Pawel Parniewski and Robert D. Wells*

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

Received March 20, 1997; Revised and Accepted May 22, 1997

ABSTRACT

Induction of transcription into long CTG·CAG repeats contained on plasmids in *Escherichia coli* is shown to increase the frequency of deletions within the repeat sequences. This elevated genetic instability was detected because active transcription into the triplet repeat influenced the growth transitions of the host cell, allowing advantageous growth for cells harboring plasmids with deleted repeat sequences. The variety of deletion products observed in separate cultures suggests that transcription altered the metabolism of the DNA in a manner that produced random length changes in the repeat sequence. For cultures containing plasmids without active transcription into the triplet repeat, or those maintained in exponential growth, deletions occurred within the repeat at a lower frequency (5–20-fold lower). In these incubations the extent of deletions was proportional to the number of cell divisions and many repeat lengths were observed within each culture, suggesting that the decrease in average repeat length at long incubation times was due to multiple small deletions. These observations show that deletions within long CTG·CAG repeats contained on plasmids in *E.coli* occur via more than one pathway and their level of genetic instability is altered by the enzymatic processes occurring upon the DNA.

INTRODUCTION

Simple repetitive DNA sequences, which are widespread throughout natural genomes, are polymorphic in length due to intrinsic genetic instabilities (1–3). The recent surge of interest in the factors influencing such instabilities is due to the association of simple repetitive DNA sequences with a variety of human diseases (4,5). For example, hereditary non-polyposis colon cancer cell lines have a high frequency of length changes in specific dinucleotide repeats (6,7) and unusual mutation events, involving the expansion of triplet repeat sequences (TRS), are associated with a number of neurogenetic disorders (for recent reviews see 8–12). The propensity for TRS to expand provides an explanation for the non-Mendelian inheritance patterns (termed anticipation) of their associated diseases; with progression through a pedigree, the TRS become longer and the disease has an earlier age of onset and an increased severity.

Some features of the events leading to certain diseases associated with expansion of TRS are known (11,12). The expansion events fall into two categories, termed Types 1 and 2 (12). In Type 1 expansions, the TRS is always CTG·CAG and is in a coding segment of the gene. In this case expansion events occur in small steps and the net effect is to produce proteins with longer tracts of polyglutamines. It is likely that these mutations lead to disease proteins with a gain of function that is particularly deleterious in neurons. Type 2 expansion events are associated with multi-system disorders and have a number of differences; various types of sequences are found and the TRS is not located within a coding region of the gene. All of the Type 2 expansion events are extremely large (hundreds of copies of the repeat for disease alleles).

Although progress has been made in understanding the pathologies of diseases associated with genetic instability of simple repetitive DNA sequences, the mechanism of expansion of TRS remains unclear. A number of studies have shown that sequences which undergo expansion can adopt unusual structures (reviewed in 13) which may cause aberrant metabolism of the DNA. To attempt to elucidate these mechanisms, we have initiated *in vivo* studies in *Escherichia coli*. Maintenance of long TRS in bacterial systems is difficult since they are deleted readily to sequences of heterogeneous length when contained on plasmids (14,15). Despite these problems, expanded products of some TRS have been obtained in *E.coli* (14,16), although the relationship of these observations to eukaryotic systems remains to be determined.

*To whom correspondence should be addressed. Tel: +1 713 677 7651; Fax: +1 713 677 7689; Email: rwells@ibt.tamu.edu

Present addresses: [+]Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK and [§]Department of Genetics of Microorganisms, Institute of Microbiology and Immunology, University of Lodz, 12/16 Banacha St., 90-237 Lodz and Center for Microbiology and Virology, Polish Academy of Sciences, 106 Lodowa St., 93-232 Lodz, Poland

Herein we show that the deletion of long CTG·CAG tracts from plasmids in *E.coli* is increased upon active transcription into the repeat sequence. An elevated frequency of deletions is observed because cells harboring plasmids with deleted repeat sequences have a growth advantage as the cultures pass through stationary phase (17). In the absence of transcription, a lower rate of deletions is observed and their extent was dependent on the number of cell division events. Hence, at least two pathways leading to deletions of TRS exist in *E.coli*.

MATERIALS AND METHODS

Plasmids and bacterial strains

Plasmids used in these experiments contained repeating CTG·CAG inserts, which may also be designated TGC·GCA or GCT·AGC, have been described previously (14,17). The orientation of this TRS with respect to the direction of replication within these plasmids has been defined (14) 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. Figure 1 shows maps of pRW3247 and pRW3268 with the restriction sites relevant to this study. Both plasmids contain the sequence $(GCT)_{27}ACT(GCT)_{40}ACT(GCT)_{106}$ [referred to as $(CTG)_{175}\cdot(CAG)_{175}$ for convenience] in orientation II; note that pRW3247 has an intact promoter for *lacZ′*, but pRW3268 does not. Equivalent plasmids with the same TRS in orientation I are pRW3248 (intact *lacZ′* promoter) and pRW3269 (deleted *lacZ′* promoter). In some experiments plasmid pI^Q-kan (18) was used to provide the $lacI^Q$ repressor inside the cell.

Large amounts of plasmids were prepared by alkaline lysis of a 1 l culture (LB + antibiotic) grown to stationary phase and were purified by CsCl/ethidium bromide ultracentrifugation overnight (19). Purification of plasmids from 10 ml cultures was performed using the standard alkaline lysis miniprep procedure (19).

All cloning procedures and experiments were conducted in *E.coli* HB101 as described previously (17). After transformations, bacteria were allowed to recover at 37°C in SOC medium (2% Bacto-tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose, pH 7.0) for 30–45 min and then grown up in liquid culture with antibiotic selection; ampicillin and kanamycin were included at final concentrations of 100 and 50 µg/ml respectively. When required, IPTG was included in the growth medium at a final concentration of 1 mM. To check the efficiency of each electroporation, 20 µl of the transformation mixture were grown overnight at 37°C on LB–agar plates containing antibiotic.

Conditions of bacterial growth

The genetic stability of plasmids containing TRS was analyzed under a variety of conditions. In the basic procedure, *E.coli* HB101 was transformed with the appropriate plasmid and an aliquot was inoculated into glass tubes (150 mm length × 17 mm diameter) containing 10 ml LB with antibiotic. For incubations containing IPTG, the transformed cells were first grown for 2 h without IPTG to allow cell recovery in the absence of transcription from the *lacZ′* promoter and this culture was then diluted into fresh medium with or without IPTG.

Incubation of the liquid cultures was continued for the required time at 37°C and a shaking rate of 250 r.p.m. Growth of the cultures was monitored by optical density readings at 600 nm. The bacteria were subcultured into fresh liquid medium as appropriate: inoculations from stationary phase cultures were performed after 24 h (OD_{600} >2) with a dilution factor of 10^7; inoculations from logarithmic phase cultures were performed after 12 h (OD_{600} 0.3–0.8) with a dilution factor of 4×10^6. The cells from each culture were harvested and stored at –80°C until the end of the experiment and then all DNA minipreps were performed at the same time.

Analysis of DNA

Restriction digests were performed following the manufacturer's instructions and the samples were radiolabeled at the *Eco*RI site using *E.coli* DNA polymerase I Klenow fragment incubated with [α-^{32}P]dATP (19). The samples were analyzed by electrophoresis through 5% polyacrylamide gels in TBE buffer and the radioactive fragments were observed by autoradiography or with a phosphorimager (400S, Molecular Dynamics). Quantitation of the gels was performed directly upon the phosphorimage.

RESULTS

Transcription into $(CTG)_{175}\cdot(CAG)_{175}$ increases its frequency of deletions

Long tracts of CTG·CAG repeats are prone to deletion over extended periods of growth in *E.coli*. Previous investigations (14,17) showed that $(CTG)_{175}\cdot(CAG)_{175}$ was more stable when the CTG strand was the leading strand template for replication (orientation I) compared with when it was the lagging strand template (orientation II). For all plasmids containing long CTG·CAG tracts, deletions were observed within the TRS. However, the frequency of deletions varied widely between different vectors. Diverse aspects of DNA metabolism (e.g. recombination, repair, replication, transcription) were proposed to explain such variations, but the relative importance of each was not known. Thus the present experiments were undertaken to identify the factors influencing the frequencies of deletions of long CTG·CAG repeats from plasmids in *E.coli*.

In an earlier study (14) the following pairs of plasmids were produced: pRW3247 (orientation II) and pRW3248 (orientation I); pRW3268 (orientation II) and pRW3269 (orientation I). These plasmids all contain $(GCT)_{27}ACT(GCT)_{40}ACT(GCT)_{106}$ [referred to as $(CTG)_{175}\cdot(CAG)_{175}$], but growth of *E.coli* harboring these plasmids produced different frequencies of deletions within each of the TRS (14). An obvious difference in the two sets of plasmids is that pRW3247 and pRW3248 have an intact promoter for *lacZ′*, whereas pRW3268 and pRW3269 do not; thus the TRS is located inside a transcription unit for pRW3247 and pRW3248, but not for pRW3268 and pRW3269 (Fig. 1).

A number of different experiments were performed to assess the influence of transcription on the genetic stability of the TRS in these plasmids. Repression of transcription from the *lacZ′* promoter by addition of glucose to the medium (up to 20 mM) decreased the frequency of deletions for pRW3247 and pRW3248 (plasmids containing the *lacZ′* promoter) but not pRW3268 and pRW3269 (plasmids without the *lacZ′* promoter) (data not shown). These observations suggested that transcription into the TRS increased the frequency of deletions. However, the various concentrations of glucose produced dramatically different growth rates of the bacteria, which could alter genetic propagation of the TRS.

A more discriminating methodology was designed to analyze the influence of transcription on deletions from TRS which

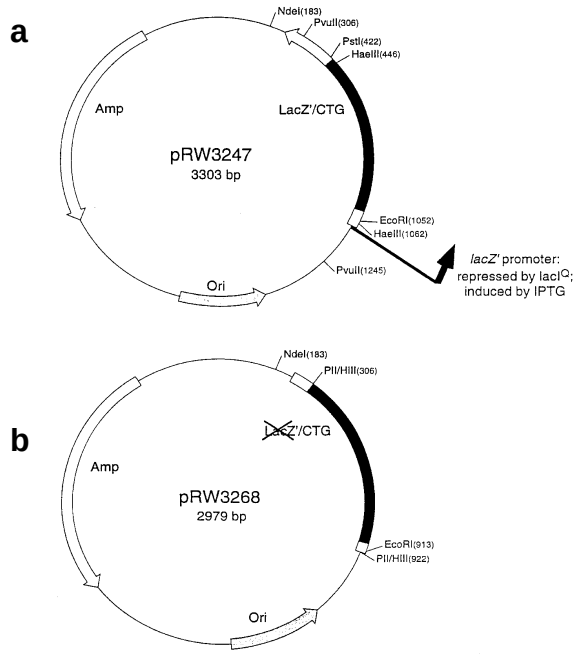

Figure 1. Schematic diagram of plasmids containing CTG·CAG triplet repeats. All plasmids used in this study contain the sequence $(GCT)_{27}ACT(GCT)_{40}ACT(GCT)_{106}$, referred to as $(CTG)_{175}\cdot(CAG)_{175}$ for convenience. The orientation of this TRS with respect to the direction of replication within these plasmids has been defined (14) as follows: those with the CTG sequence as the leading strand template are designated orientation I; those with the CTG sequence as the lagging strand template are designated orientation II. Two pairs of plasmids were used (14): (**a**) pRW3247 (orientation II) and pRW3248 (orientation I), which contain a functional *lacZ′* promoter (shown by the arrow); (**b**) pRW3268 (orientation II) and pRW3269 (orientation I), which do not have the promoter for *lacZ′*. None of these plasmids has a functional *lacZ′* gene. Within each pair, the plasmids are equivalent except that the orientation of the TRS with respect to the direction of replication is reversed. The approximate positions and direction of the origin of replication (Ori), the gene encoding resistance to ampicillin (Amp) and restriction sites relevant to this study are shown. In pRW3268 the positions designated PII/HIII identify the sequences produced during cloning of blunt-ended fragments from *Pvu*II and *Hae*III digests: the *Hae*III fragment containing the TRS was ligated to vector DNA digested with *Pvu*II. The filled box shows the position of the human genomic DNA, including the TRS; the sequence of this section of the DNA is given in Bowater *et al.* (17).

involved co-transformation into the bacterial cells of a compatible plasmid, pI^Q-kan (18). The $lacI^Q$ repressor is synthesized in all cells containing pI^Q-kan and, therefore, transcription from the *lacZ′* promoter is normally repressed, but it can be induced by addition of IPTG (18). Analysis of RNA transcripts from cultures of *E.coli* HB101 containing pI^Q-kan and pRW3247 in exponential growth detected an increase in transcription in the presence of IPTG (Lawson and Wells, unpublished data). We estimate that transcription from the *lacZ′* promoter was 3–10-fold higher in the presence of IPTG compared with cells which did not have pI^Q-kan co-transformed (see Fig. 4).

Escherichia coli HB101 transformed with pI^Q-kan and plasmids containing the TRS was grown in two identical liquid cultures, with and without IPTG, and these cultures were grown to stationary phase and then diluted into fresh medium. At each subculturing a portion of the culture was harvested. The plasmids were purified and the lengths of the fragments containing the TRS were determined by polyacrylamide gel electrophoresis of restriction digests (Fig. 2A and B). Although tracts of TRS have a faster mobility than random sequence DNA on polyacrylamide gels, the mobility of a fragment containing TRS is proportional to the size of the repeat sequence (20,21). We have shown previously that all changes to the length of this fragment were due to insertion or deletion of unit lengths (3 bp) of TRS (14).

Quantitation of the amount of deletion products showed that, upon induction of transcription of *lacZ′* with IPTG, there was a dramatic increase in deletions within the TRS for the plasmid harboring an intact *lacZ′* promoter and with the TRS in orientation II (pRW3247) (Fig. 2C). By the third stationary phase, there were relatively few deletions in cultures without IPTG (~95% of plasmids were full length), whereas <10% of plasmids were full length in cultures grown in the presence of IPTG. The difference in genetic stability of the TRS upon induction of transcription was not due to the presence of IPTG in the medium, since similar experiments with other plasmids showed no difference in the stability of the TRS upon addition of IPTG (Fig. 2C). For the plasmid harboring an intact *lacZ′* promoter and with the TRS in orientation I (pRW3248), the TRS was almost completely stable with or without IPTG (~96% full length in each case after the third stationary phase).

Cultures containing the plasmid without the promoter for *lacZ′* and with the TRS in orientation II (pRW3268) did not have a different extent of deletions upon addition of IPTG (Fig. 2B and C). This absence of an effect of IPTG on genetic stability was expected, since transcription through the TRS would not be induced by the presence of IPTG. However, it should be noted that in all experiments there was a measurable extent of deletions for pRW3268 (no *lacZ′* promoter). With or without IPTG the cultures from the third stationary phase contained ~65% of plasmids with full length TRS. The similar plasmid with no *lacZ′* promoter and $(CTG)_{175}\cdot(CAG)_{175}$ in orientation I (pRW3269) had a completely stable TRS under all conditions (data not shown). Thus experiments with both pairs of plasmids (pRW3247, pRW3248 and pRW3268, pRW3269) support earlier studies which showed that long CTG·CAG repeats are more stable in orientation I (14,17).

Comparison of the two plasmids with $(CTG)_{175}\cdot(CAG)_{175}$ in orientation II showed that they had a different response to the various incubation conditions. In cultures without IPTG, the plasmid harboring an intact *lacZ′* promoter (pRW3247) had a lower frequency of deletions than the plasmid with a deleted *lacZ′* promoter (pRW3268). However, upon addition of IPTG, pRW3247 had a frequency of deletions that was higher than pRW3268. Hence, the plasmid harboring an intact *lacZ′* promoter (pRW3247) was relatively stable without transcription, but very unstable when there was active transcription into the TRS. The good genetic stability of pRW3247 in the absence of IPTG was probably because binding of the $lacI^Q$ repressor prevented all transcription complexes, even from upstream promoters, from entering the TRS. Similarly, the intermediate level of genetic instability of pRW3268 (no *lacZ′* promoter), which was independent of transcription from the *lacZ′* promoter, was likely due to residual transcription from other promoters or to other differences in unidentified aspects of DNA metabolism of pRW3268 and pRW3247.

These experiments show that induction of transcription through $(CTG)_{175}\cdot(CAG)_{175}$ in orientation II increases the frequency of deletions within the TRS. This decreased genetic stability of the

repeat sequence could be dependent on transcription events alone or other associated processes, such as translation.

Transcription influences the genetic stability of triplet repeats in cultures that pass through stationary phase

Earlier experiments on *E.coli* HB101 cultures not containing the lacI^{Q} repressor showed that passage of the cultures through stationary phase increased the frequency of deletions in the TRS of the plasmids harboring an intact *lacZ′* promoter (pRW3247 and pRW3248) (17). Since the above experiments showed a clear influence of transcription on the stability of pRW3247, we wondered if transcription altered the effects of growth phase on deletion frequency. Long-term growth experiments were performed as outlined above, but the bacterial cultures were either maintained in exponential phase or allowed to pass through stationary phase (after every 24 generations). This experiment was performed on the plasmid harboring an intact *lacZ′* promoter with the TRS in orientation II (pRW3247), since it was the only plasmid for which addition of IPTG influenced the stability of its TRS (Fig. 2). DNA was purified at each subculturing and the samples were analyzed by restriction mapping (Fig. 3A). Quantitation of the restriction digests (Fig. 3B) confirmed that IPTG increased the extent of deletions in cultures that passed through stationary phase; by the fourth stationary phase (~100 generations) there was >90% full-length product remaining in cultures without IPTG, but almost no full-length product in the presence of IPTG. However, for the cultures maintained in exponential growth, the stability of the TRS was influenced less by transcription; under exponential growth conditions the amount of deletions was ~2-fold higher in the presence of IPTG. It should be noted that the frequency of deletions within the TRS for cultures maintained in exponential growth (with or without IPTG) was lower than stationary phase cultures which contained IPTG but higher than stationary phase cultures without IPTG.

Therefore, these results show that in exponential growth or in the absence of transcription into the TRS there is a basal level of deletions for pRW3247 (intact *lacZ′* promoter, TRS in orientation II). Cultures that pass through stationary phase with active transcription into the TRS have an increased frequency of deletions, suggesting that an additional pathway promotes deletions in this situation.

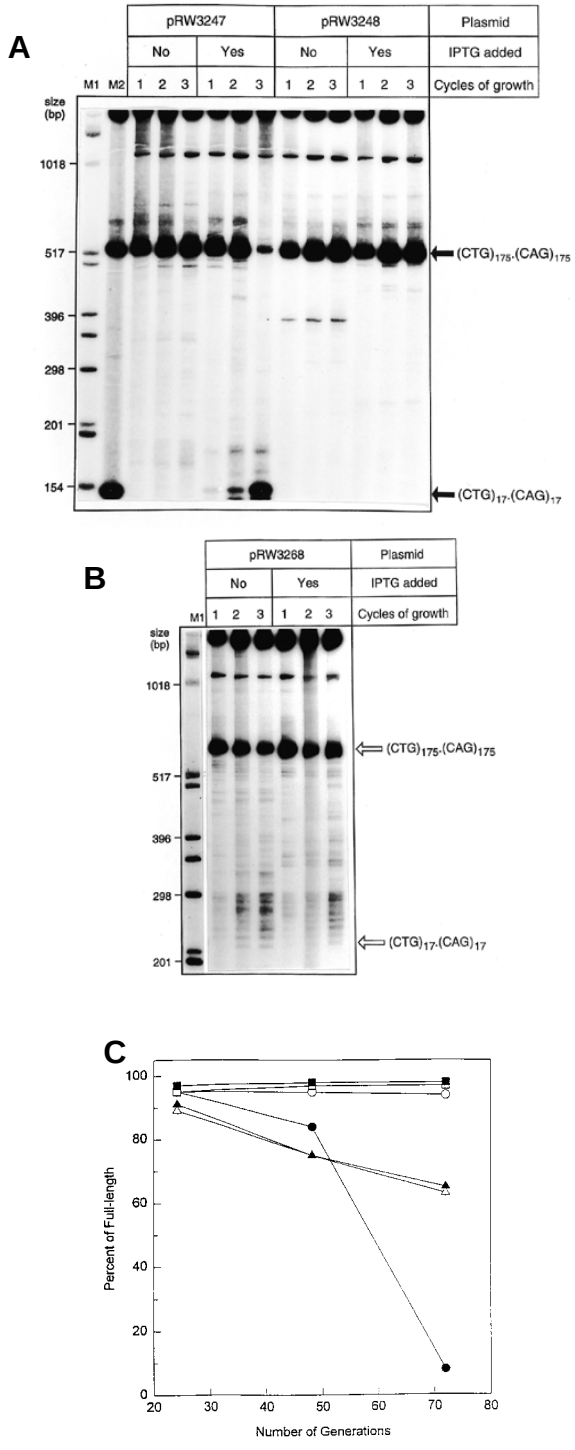

Figure 2. Transcription increases the frequency of deletions within $(CTG)_{175}\cdot(CAG)_{175}$ contained on plasmids in *E.coli*. *Escherichia coli* HB101 harboring pI^{Q}-kan was transformed with purified supercoiled monomer of the designated plasmids, grown to stationary phase in medium with or without IPTG and the growth cycles repeated as required. DNA from all cultures was purified and analyzed as described. The samples are shown with the numbers above each lane indicating the number of cycles of growth into stationary phase at the time of harvesting; each growth cycle is equivalent to ~24 generations of growth. Arrows designate the position of migration of bands containing $(CTG)_{17}\cdot(CAG)_{17}$ and $(CTG)_{175}\cdot(CAG)_{175}$; an open arrow refers to pRW3268 and a filled arrow refers to pRW3247 and pRW3248. (**A**) Restriction analysis of deletions from plasmids with an intact *lacZ′* promoter: pRW3247 (orientation II) and pRW3248 (orientation I). DNA samples were digested with *Eco*RI and *Pst*I, labeled with ^{32}P and electrophoresed through a 5% polyacrylamide gel in TBE. M1 is marker DNA (1 kb ladder; Gibco BRL); the sizes of these bands are shown at the side. M2 is a marker containing a mixture of pRW3244 $[(CTG)_{17}\cdot(CAG)_{17}]$ and pRW3248 $[(CTG)_{175}\cdot(CAG)_{175}]$ digested with *Eco*RI and *Pst*I. The band that is present in all experimental samples at ~1 kb is a digestion product from pI^{Q}-kan. Note that fragments of DNA containing TRS migrate faster than expected from their known size, in agreement with other studies (20,21). (**B**) Restriction analysis of deletions from a plasmid without the *lacZ′* promoter, pRW3268 (orientation II). DNA samples were digested with *Eco*RI and *Nde*I, labeled with ^{32}P and electrophoresed through a 5% polyacrylamide gel in TBE. M1 is a marker as described in (A). (**C**) Quantitation of deletions in pRW3247, pRW3248 and pRW3268. Quantitation was performed on the digested samples shown in (A) and (B). The extent of full-length TRS $[(CTG)_{175}\cdot(CAG)_{175}]$ was determined as the proportion of all DNA of this size or smaller. Reproducibility of these measurements was estimated to be ±2%. Circles, pRW3247; squares, pRW3248; triangles, pRW3268; open symbols, samples grown in the absence of IPTG (i.e. no transcription through the TRS); filled symbols, samples grown in the presence of IPTG (i.e. active transcription through the TRS).

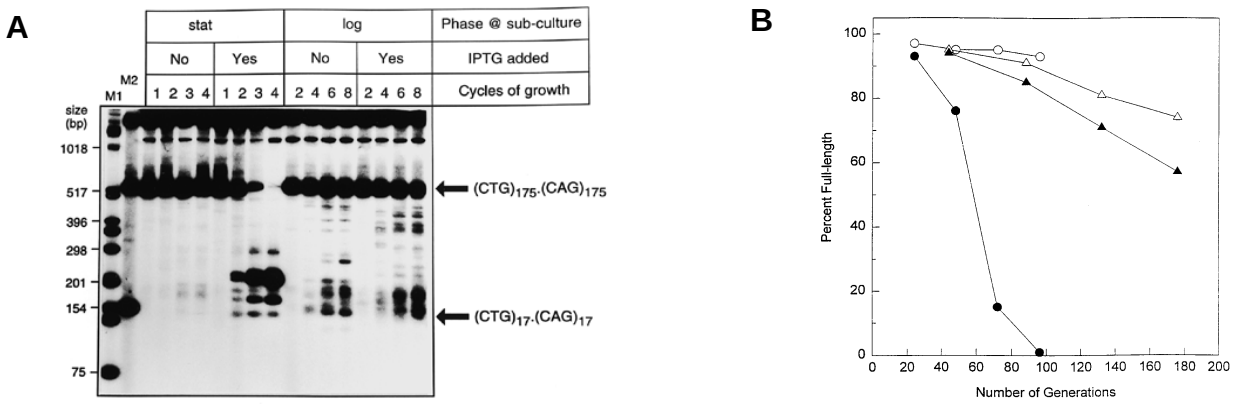

Figure 3. Transcription increases the frequency of deletions in $(CTG)_{175}\cdot(CAG)_{175}$ in cultures that pass through stationary phase. *Escherichia coli* HB101 harboring pI^{Q}-kan was transformed with purified supercoiled monomer of pRW3247 (orientation II, intact *lacZ′* promoter), grown with or without IPTG to stationary phase or mid-logarithmic phase and the growth cycles were repeated as required. DNA from all cultures was purified and analyzed as described. The samples are shown with the numbers above each lane indicating the number of cycles of growth at the time of harvesting; each growth cycle is equivalent to ~24 generations of growth for samples subcultured in stationary phase and to ~22 generations of growth for samples subcultured in logarithmic phase. Arrows designate the position of migration of bands containing $(CTG)_{17}\cdot(CAG)_{17}$ and $(CTG)_{175}\cdot(CAG)_{175}$. (**A**) Analysis by restriction digestion: DNA samples were analyzed as in Figure 2A. (**B**) The data shown in (A) were quantitated as outlined in Figure 2C. Circles, samples subcultured in stationary phase; triangles, samples subcultured at mid-exponential phase; open symbols, samples grown in the absence of IPTG (i.e. no transcription through the TRS); filled symbols, samples grown in the presence of IPTG (i.e. active transcription through the TRS).

Growth phase influences deletion frequency only when the TRS is located within an active transcription unit

Thus far, we have shown that active transcription into the TRS of plasmids in *E.coli* HB101 produced an increased frequency of deletions in repeat sequences from cultures that passed through stationary phase compared with those maintained in exponential growth. Since pRW3268 and pRW3269 do not have a promoter for *lacZ′*, the frequency of deletions from the TRS in these plasmids should not be influenced by growth phase. Therefore, experiments were performed with pRW3268 and pRW3269 to test this hypothesis. Since transcription from the *lacZ′* promoter was not possible for these plasmids, in these experiments pI^{Q}-kan was not co-transformed into the bacteria and IPTG was not included in the growth medium.

After electroporation of plasmids containing the TRS into *E.coli* HB101, the transformation mixture was halved and grown for many generations with the cultures either maintained in exponential phase or allowed to pass through stationary phase (after every 24 generations). At each subculturing a portion of the culture was harvested. The plasmids were purified and the lengths of fragments containing the TRS from pRW3268 (TRS in orientation II, no *lacZ′* promoter) were determined (Fig. 4); samples from cultures harboring the plasmid with an intact *lacZ′* promoter and the TRS in orientation II (pRW3247) that were maintained in exponential growth were also analyzed for comparison with our previous observations (17). Quantitation of the data showed that there was little influence of growth phase on the frequency of deletions in the plasmid without the *lacZ′* promoter and the TRS in orientation II (pRW3268), since cultures maintained in exponential growth or allowed to pass through stationary phase had ~80% full-length product after 72 generations. This is in contrast to cultures harboring the plasmid with the *lacZ′* promoter and the TRS in orientation II (pRW3247), which had a dramatically higher extent of deletions when allowed to pass through stationary phase (17). For both types of growth conditions, no deletions were detected in pRW3269 (TRS in orientation I, no *lacZ'* promoter) (data not shown).

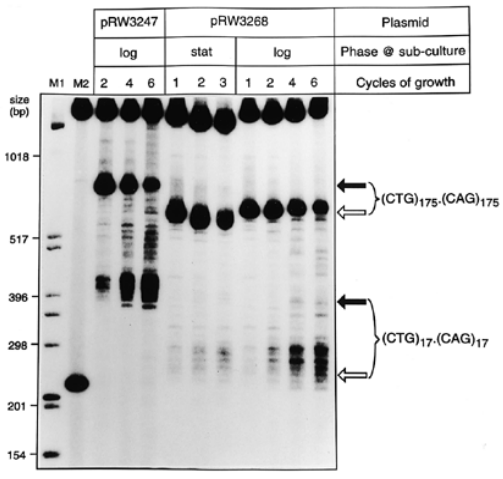

Figure 4. Bacterial growth phase does not influence the extent of deletions in $(CTG)_{175}\cdot(CAG)_{175}$ that are located outside an active transcription unit. *Escherichia coli* HB101 was transformed with purified supercoiled monomer of pRW3247 (orientation II, intact *lacZ′* promoter) or pRW3268 (orientation II, no *lacZ′* promoter), grown to stationary phase or mid-logarithmic phase and the growth cycles repeated as required. (Note that these cells did not contain the $lacI^{Q}$ repressor and so transcription could occur from *lacZ′*). DNA from all cultures was purified and analyzed by restriction digestion as in Figure 2B. M2 is a marker containing a mixture of pRW3244 and pRW3248 [plasmids with $(CTG)_{17}\cdot(CAG)_{17}$ and $(CTG)_{175}\cdot(CAG)_{175}$ respectively]. Arrows designate the position of migration of bands containing $(CTG)_{17}\cdot(CAG)_{17}$ and $(CTG)_{175}\cdot(CAG)_{175}$; when two arrows are shown for each length, the open arrow refers to pRW3268 and the filled arrow refers to pRW3247.

In the experiments with cultures which contained pI^{Q}-kan as well as the plasmids with the TRS (Fig. 2), cultures containing pRW3268 and maintained in exponential growth or allowed to pass through stationary phase had ~65% full-length product after 72 generations. Thus although co-transformation of pI^{Q}-kan decreased the growth rate of the host cell (data not shown), this had little effect upon genetic stability of the TRS in pRW3268 (no *lacZ′* promoter). In contrast, comparison of the experiments with

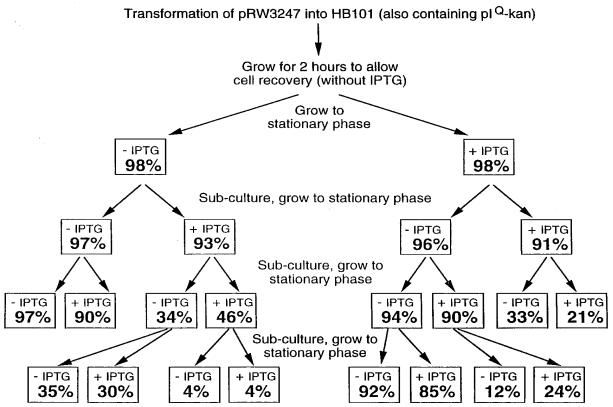

Figure 5. Transcription produces a growth advantage for bacteria harboring plasmids with shorter triplet repeats upon passage of the culture through stationary phase. *Escherichia coli* HB101 harboring pI^{Q}-kan was transformed with purified supercoiled monomer of pRW3247 (orientation II, intact *lacZ′* promoter) and grown with or without IPTG to stationary phase. Each sample was subcultured into medium with or without IPTG and the growth cycles repeated as required; thus each culture gives rise to two daughter cultures. Every cycle of growth into stationary phase is equivalent to ~24 generations of growth. DNA from all cultures was purified, digested with *Eco*RI and *Pst*I and analyzed as in Figure 2A. Quantitation was performed on the restriction digests shown as outlined in Figure 2C and the percentage of full-length product in each sample is shown. Reproducibility of these measurements was estimated to be ±2%.

plasmids containing an intact *lacZ′* promoter (pRW3247 and pRW3248) (compare Fig. 2 with fig. 3 of 17) shows that co-transformation of pI^{Q}-kan did alter the genetic stability of the TRS. In *E.coli* HB101 not containing pI^{Q}-kan there was transcription into the TRS of pRW3247 and pRW3248 (data not shown), probably due to the high copy number of the plasmid titrating out the normal level of repressor in cells; the increased amount of deletion in cultures that passed through stationary phase was due to a growth advantage for cells harboring plasmids with deletions (17). The reduced effects on the genetic stability of pRW3247 and pRW3248 by induction of transcription from the *lacZ′* promoter with IPTG (Fig. 2) were because the decreased growth rate of bacteria harboring pI^{Q}-kan (data not shown) reduced the growth advantage of cells harboring plasmids with deleted TRS. Thus, although transcription clearly affects the frequency of deletions within the TRS, other factors also influence the observed genetic stability of the repeat sequence.

Growth curve measurements of cultures harboring each of the plasmids showed that there was less interaction between the growth characteristics of *E.coli* HB101 and plasmids with long TRS if the plasmids did not contain the *lacZ′* promoter (pRW3268 and pRW3269) (data not shown). Hence, the absence of an effect of growth phase on the deletion frequency within the TRS of plasmids without an intact *lacZ′* promoter (Fig. 4) was as predicted. Plasmids without transcription from *lacZ′* are almost completely stable in orientation I (pRW3269) and in orientation II (pRW3268) have a deletion frequency that is dependent only on the number of cell division events.

Active transcription into TRS produces a growth advantage for cells with deleted products after passage of the culture through stationary phase

We wished to determine whether the advantageous propagation of cells harboring deleted repeats was due to effects of transcription at a specific growth phase of the culture. Thus, an experiment was performed in which the plasmid harboring an intact *lacZ′* promoter with the TRS in orientation II (pRW3247) was transformed into *E.coli* HB101 containing pI^{Q}-kan and at each subculturing the cells were diluted into medium with or without IPTG.

The strategy and results of the experiment are shown in Figure 5. The proportion of plasmid containing full-length TRS was quantitated from restriction digests analyzed on polyacrylamide gels (data not shown). The results showed that the genetic stability of the TRS was determined by the state of the parent culture regarding transcription into the TRS. If the parent culture had transcription into the TRS then, after passage through stationary phase, its daughter cultures had increased amounts of deletions with and without IPTG. Conversely, if there was no active transcription into the TRS in the parent culture, the daughter cultures had relatively little change in their frequency of deleted products, even when active transcription into their TRS was induced by the presence of IPTG. In a control experiment the cultures were washed and incubated with medium with or without IPTG immediately prior to subculturing; as expected, the presence or absence of IPTG had no effect at this stage (data not shown).

Thus active transcription into $(CTG)_{175}\cdot(CAG)_{175}$ contained on plasmids influences the transitions between stationary and exponential growth. This produces advantageous growth for cells harboring plasmids with deleted TRS and increases the observed amount of deletions in the TRS in cultures that pass through stationary phase.

DISCUSSION

The genetic stability of simple repeating sequences contained on plasmids in *E.coli* is influenced by many factors. Long CTG·CAG repeats were shown to have a higher frequency of deletions upon active transcription into the TRS. This elevated frequency of deletions (5–20-fold increase for each growth cycle, depending on the experiment) was not proportional to the number of generations of the culture, but was produced upon passage of cultures through stationary phase. This observation explains earlier experiments which detected a growth advantage for cells harboring plasmids which contained deleted TRS (17). These results are unlikely to be due to an elevated copy number of plasmids harboring deletions because the size of the TRS has little effect on the copy number of these plasmids (data not shown). Although growth advantages were obtained only for plasmids with active transcription into the TRS, we are uncertain whether the destabilizing influence was transcription or some other event dependent upon transcription, such as translation. Plasmids without active transcription into the TRS also produced deletions within the repeat, but at the lower frequency of cultures maintained in exponential growth and the extent was proportional to the number of cell division events. Hence, deletions from long CTG·CAG repeats on plasmids in *E.coli* occur via more than one pathway.

The molecular processes responsible for the effect of transcription on deletion frequency are uncertain from this *in vivo* study. However, examination of the patterns of deletions from different plasmids with the TRS in orientation II (CTG sequence as lagging strand template) reveals details of the mechanism by which transcription influences the stability of long TRS. For example, in conditions providing a growth advantage for cells with deleted products, passage through stationary phase of cultures which

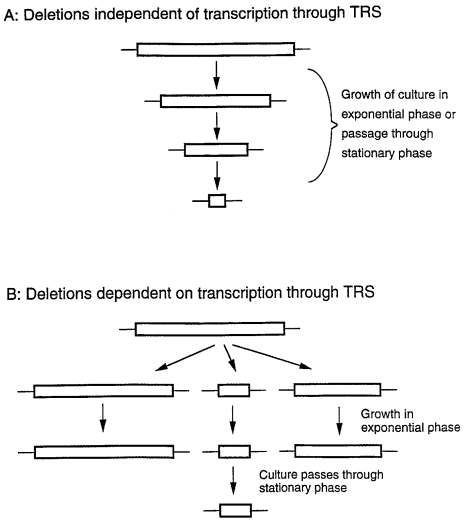

Figure 6. Deletions within TRS occur via more than one pathway. For all bacterial cultures containing plasmids harboring TRS, deletions occur within the TRS with a frequency which is influenced by the orientation of the TRS with respect to the direction of replication. (**A**) In cultures without transcription into the TRS, the extent of deletions is proportional to the age of the culture, but is not influenced by the growth history of the culture (whether or not it was maintained in exponential growth). Small changes to repeat length occur continuously, producing a decrease in average repeat length for longer incubations. (**B**) Active transcription through the TRS alters the pattern and frequency of deletions from the repeating sequence. Deletions of random size occur when transcription (or related processes) encounters long TRS and, after the culture passes through stationary phase, a growth advantage is produced for cells harboring plasmids with deletions. Thus cultures that pass through stationary phase have a much greater frequency of deletions.

harbor plasmids with transcription into the TRS (pRW3247 co-transformed with pI^Q-kan and in the presence of IPTG), the pattern of deletions varied in different experiments and the frequency of deletions varied by up to 2-fold (see for example Figs 2 and 3). However, each culture contained plasmids with relatively few different lengths of TRS. Comparison of separate inoculations showed that a wide variety of lengths of TRS were produced (compare Figs 2 and 3); frequently the deletions produced only plasmids containing small repeats.

Analysis of growth conditions which did not give an advantage for cells with deleted products, plasmids without transcription into the TRS (e.g. pRW3268 and pRW3247 co-transformed with pI^Q-kan and in the absence of IPTG) or cultures maintained in exponential growth, shows that the patterns of deletion products were similar in different experiments, in agreement with earlier observations (14). Also, the observed frequency of deletions was similar in separate experiments. In these samples it was usual for many lengths of TRS to be observed in each culture. Thus deletions occurred via a step-wise mechanism which produced a gradual decrease in the average length of repeats as the incubation time of the culture was increased (Fig. 6A). These types of deletion events were influenced by the orientation of the repeat sequence with respect to the direction of replication, suggesting that they were generated by slippage of the DNA at the replication fork (14). It has been proposed previously that the intrahelical process of slipped strand mispairing is likely to be the major factor in initial length changes of short repeated motifs (22); multiple events would be required to produce the short TRS observed at long generation times of the cultures in our experiments.

An additional mechanism for generating deletions was dominant under growth conditions which produced an advantage for plasmids harboring a smaller TRS. Upon passage of the cultures through stationary phase, the growth advantage associated with plasmids harboring shorter lengths of TRS meant that the bacterial culture became biased towards the shortest length produced during that incubation (Fig. 6B). Induction of transcription into the repeat sequence produced random length changes, giving a variety of deletion sizes in different experiments. A plausible mechanism for generating random length changes which could encompass almost the full length of the TRS is the formation of hairpin structures within single-stranded regions of the DNA. Stable hairpin structures have been observed *in vitro* within short single-stranded oligonucleotides of CTG·CAG repeats and in plasmids containing long CTG·CAG repeats (reviewed in 13).

Our data cannot determine whether the deletions due to transcription into the TRS were produced during exponential growth of the culture (as illustrated in Fig. 6B) or only if the cultures entered stationary phase. Rare deletion events that occurred during exponential growth may not have been detected by our analysis, but the reduced lag phase of cells harboring such plasmids (17) could have produced their accumulation upon passage through stationary phase. The cellular physiology of stationary phase cells is dramatically different from cells in exponential growth (23,24) and the fitness of a culture continues to change during long incubations in stationary phase (25). It is possible that interactions between the replication and transcription machinery produced the growth disadvantages that we observed for cultures harboring plasmids with long TRS as they passed through stationary phase.

Studies in a variety of experimental systems have suggested that problems occur when a replication fork meets a transcription complex on plasmid DNA (26–29). Such encounters occur during normal cell growth, so the cell must have evolved efficient ways to circumvent these interactions (30). However, under certain conditions errors may be more likely. The observation that DNA synthesis does not proceed efficiently through long CTG·CAG repeats *in vitro* (31,32) suggests that polymerase reactions through TRS may be prone to aberrations *in vivo*. The presence of transcription complexes within the TRS may enhance the error rate of replication and, therefore, produce the high frequency of deletions reported here.

Our experiments show only that induction of transcription from the *lacZ′* promoter had an effect on the genetic stability of $(CTG)_{175}\cdot(CAG)_{175}$. Such effects have been demonstrated for other promoters and simple repeating sequences. In *E.coli* the genetic stability of plasmids containing monomeric repeat sequences $[(A)_{34}\cdot(T)_{34}$ and $(G)_{34}\cdot(C)_{34}]$ cloned downstream of the *lacZ′* promoter was decreased upon induction of transcription (28). Also, recent studies in yeast showed that transcription produced an increase in the polymorphism of dinucleotide repeats (33) and that high levels of transcription increased the rate of spontaneous mutations (34).

The discovery described herein that transcription can influence the genetic stability of simple repeating sequences prompted us to re-evaluate earlier investigations which used repetitive DNA to study unusual DNA structures. Other TRS associated with human diseases, CGG·CCG and AAG·CTT, were also observed to be less stable in plasmids in *E.coli* when cloned within a

transcription unit (15; Ohshima, Montermini, Pandolfo and Wells, unpublished data). Similarly, during analysis of the formation of Z-DNA, dinucleotide repeats were observed to undergo more deletions when located in transcribed regions of plasmids (35–38). Thus transcription has destabilizing effects on the genetic propagation of a number of different simple repetitive DNA sequences.

The relevance of our observations on the stability of TRS to their associated human diseases is unclear. The identification that more than one pathway produces genetic instabilities in long TRS is consistent with the suggestion that different mechanisms produce the distinct characteristics of disease Types 1 and 2 (12). Since the regulation of transcription is specific for each promoter, this may provide one biochemical explanation by which different pathways can produce expansion. In all situations in which the expansion of a TRS is associated with a human disease, the TRS is located within a gene (11,12). Thus there is potential for transcription-associated events to influence the stability of TRS in all disorders. Our observations in *E.coli* show that the type of genetic instability associated with a particular TRS is determined by the processes occurring upon it. The fundamental nature of these processes (transcription and replication) suggests that such instabilities are likely to be conserved through evolution, although the details may vary in different systems.

ACKNOWLEDGEMENTS

These studies were supported by grants from the National Institutes of Health (GM52982) and the Robert A.Welch Foundation. We thank D.M.J.Lilley and D.Chen for provision of the plasmid pI^{Q}-kan and S.Kang, K.Ohshima, E.Lawson, R.Iyer and L.Bowater for helpful discussions.

REFERENCES

1 Charlesworth,B., Sniegowski,P. and Stephan,W. (1994) *Nature*, **371**, 215–220.
2 Richards,R.I. and Sutherland,G.R. (1992) *Cell*, **70**, 709–712.
3 Tautz,D. and Schlötterer,C. (1994) *Curr. Opin. Genet. Dev.*, **4**, 832–837.
4 Sutherland,G.R. and Richards,R.I. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 3636–3641.
5 Krontiris,T.G. (1995) *Science*, **269**, 1682–1683.
6 Thibodeau,S.N., Bren,G. and Schaid,D. (1993) *Science*, **260**, 816–819.
7 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., *et al.* (1993) *Science*, **260**, 812–816.
8 Caskey,C.T., Pizzuti,A., Fu,Y.H., Fenwick,R.G.,Jr and Nelson,D.L. (1992) *Science*, **256**, 784–789.
9 Willems,P.J. (1994) *Nature Genet.*, **8**, 213–215.
10 Warren,S.T. (1996) *Science*, **271**, 1374–1375.
11 Ashley,C.T. and Warren,S.T. (1995) *Annu. Rev. Genet.*, **29**, 703–728.
12 Paulson,H.L. and Fischbeck,K.H. (1996) *Annu. Rev. Neurosci.*, **19**, 79–107.
13 Wells,R.D. (1996) *J. Biol. Chem.*, **271**, 2875–2878.
14 Kang,S., Jaworski,A., Ohshima,K. and Wells,R.D. (1995) *Nature Genet.*, **10**, 213–218.
15 Shimizu,M., Gellibolian,R., Oostra,B.A. and Wells,R.D. (1996) *J. Mol. Biol.*, **258**, 614–626.
16 Ohshima,K., Kang,S. and Wells,R.D. (1996) *J. Biol. Chem.*, **271**, 1853–1856.
17 Bowater,R.P., Rosche,W.A., Jaworski,A., Sinden,R.R. and Wells,R.D. (1996) *J. Mol. Biol.*, **264**, 82–96.
18 Bowater,R.P., Chen,D. and Lilley,D.M.J. (1994) *Biochemistry*, **33**, 9266–9275.
19 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20 Chastain,P.D., Eichler,E.E., Kang,S., Nelson,D.L., Levene,S.D. and Sinden,R.R. (1995) *Biochemistry*, **34**, 16125–16131.
21 Bacolla,A., Gellibolian,R., Shimizu,M., Amirhaeri,S., Kang,S., Ohshima,K., Larson,J.E., Harvey,S.C., Stollar,B.D. and Wells,R.D. (1997) *J. Biol. Chem.*, **272**, 16783–16792.
22 Levinson,G. and Gutman,G.A. (1987) *Mol. Biol. Evol.*, **4**, 203–221.
23 Hengge-Aronis,R. (1993) *Cell*, **72**, 165–168.
24 Kolter,R., Siegele,D.A. and Tormo,A. (1993) *Annu. Rev. Microbiol.*, **47**, 855–874.
25 Zambrano,M.M. and Kolter,R. (1996) *Cell*, **86**, 181–184.
26 French,S. (1992) *Science*, **258**, 1362–1365.
27 Vilette,D., Ehrlich,S.D. and Michel,B. (1995) *Mol. Microbiol.*, **17**, 493–504.
28 Kiyama,R. and Oishi,M. (1994) *Gene*, **150**, 57–61.
29 Deshpande,A.M. and Newlon,C.S. (1996) *Science*, **272**, 1030–1033.
30 Stillman,B. (1994) *Cell*, **78**, 725–728.
31 Kang,S.M., Ohshima,K., Shimizu,M., Amirhaeri,S. and Wells,R.D. (1995) *J. Biol. Chem.*, **270**, 27014–27021.
32 Ohshima,K. and Wells,R.D. (1997) *J. Biol. Chem.*, **272**, 16798–16806.
33 Wierdl,M., Greene,C.N., Datta,A., Jinks-Robertson,S. and Petes,T.D. (1996) *Genetics*, **143**, 713–721.
34 Datta,A. and Jinks-Robertson,S. (1995) *Science*, **268**, 1616–1619.
35 Klysik,J., Stirdivant,S.M., Larson,J.E., Hart,P.A. and Wells,R.D. (1981) *Nature*, **290**, 672–676.
36 Klysik,J., Stirdivant,S.M. and Wells,R.D. (1982) *J. Biol. Chem.*, **257**, 10152–10158.
37 Jaworski,A., Hsieh,W.-T., Blaho,J.A., Larson,J.E. and Wells,R.D. (1987) *Science*, **238**, 773–777.
38 JaworskiA., Blaho,J.A., Larson,J.E., Shimizu,M. and Wells,R.D. (1989) *J. Mol. Biol.*, **207**, 513–526.