Gene Conversion (Recombination) Mediates Expansions of CTG-CAG Repeats*

Received for publication, August 7, 2000, and in revised form, September 18, 2000
Published, JBC Papers in Press, September 25, 2000, DOI 10.1074/jbc.M007153200

John P. Jakupciak‡ and Robert D. Wells§

From the Institute of Biosciences and Technology, Center for Genome Research, Texas A & M University, Texas Medical Center, Houston, Texas 77030

Genetic recombination is a robust mechanism for expanding CTG-CAG triplet repeats involved in the etiology of hereditary neurological diseases (Jakupciak, J. P., and Wells, R. D. (1999) J. Biol. Chem. 274, 23468–23479). This two-plasmid recombination system in Escherichia coli with derivatives of pUC19 and pACYC184 was used to investigate the effect of triplet repeat orientation on recombination and extent of expansions; tracts of 36, 50, 80, and 100, and 175 repeats in length, respectively, in all possible permutations of length and in both orientations (relative to the unidirectional replication origins) revealed little or no effect of orientation of expansions. The extent of expansions was generally severalfold the length of the progenitor tract and frequently exceeded the combined length of the two tracts in the cotransformed plasmids. Expansions were much more frequent than deletions. Repeat tracts bearing two G-to-A interruptions (polymorphisms) within either 171- or 219-base pair tracts substantially reduced the expansions compared with uninterrupted repeat tracts of similar lengths. Gene conversion, rather than crossing over, was the recombination mechanism. Prior studies showed that DNA replication, repair, and tandem duplication also mediated genetic instabilities of the triplet repeat sequence. However, gene conversion (recombinational repair) is by far the most powerful expansion mechanism. Thus, we propose that gene conversion is the likely expansion mechanism for myotonic dystrophy, spinocerebellar ataxia type 8, and fragile X syndrome.

Several hereditary neurological diseases including myotonic dystrophy, fragile X syndrome, spinocerebellar ataxia type 8, and Friedreich's ataxia result from expanded TRS1 CTG repeats. It is known that the majority of errors that accumulate in the repeats are susceptible to strand breaks. Yeast suggested that (CTG ) re- combination system in E. coli. The expansion events were dependent on the presence of long CTG-CAG sequences in the two-plasmid recombination system and required recombinational proficient cells to give frequent, severalfold expansions. Recombination was proven genetically and biochemically.

* This work was supported by National Institutes of Health Grants GM52982 and NS37554 and by funds from the Robert A. Welch Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.
‡ Member of the Genetics Graduate Program at Texas A & M University.
§ To whom correspondence should be addressed: Inst. of Biosciences and Technology, Center for Genome Research, Texas A & M University, Texas Medical Center, 2121 Holcombe Blvd., Houston, TX 77030. Tel.: 713-677-7600; Fax: 713-677-7689; E-mail: rwells@ibt.tamu.edu.
1 The abbreviations used are: TRS, triplet repeat sequence(s); bp, base pair(s).

This paper is available on line at http://www.jbc.org
Thus, it was concluded that if these reactions occur in humans, recombination may contribute, along with replication and repair, to the expansions responsible for anticipation associated with the hereditary neurological syndromes. Herein, we describe studies on the molecular mechanism of the recombination process (gene conversion or unequal crossing over) and the lack of effect of TRS orientation on the multiple fold expansions.

**EXPERIMENTAL PROCEDURES**

**Plasmids**—The plasmids listed in Fig. 1 are derivatives of pUC19 and pACYC184, which have unique unidirectional origins of replication. The plasmids contain human DNA with (CTG)\textsubscript{z}CAG\textsubscript{36}, (CTG)\textsubscript{z}CAG\textsubscript{50}, (CTG-CAG)\textsubscript{36}, (CTG-CAG)\textsubscript{50}, (CAG-CTG)\textsubscript{100}, (CAG-CTG)\textsubscript{130}, and (CAG-CTG)\textsubscript{160}, and transient repeats in orientations I or II. For some plasmids, the sequence may actually be (GCT-AAG-GCU) but is referred to as (CAG-TAG-CAG).
The TRS were derived from the genomes of myotonic dystrophy patients (2, 3, 5, 6, 11, 12). The construction and characterization of the plasmids were previously described (2, 3, 5, 6, 11, 12).

pRW4436 and pRW4480, which harbor the inserts in orientation II, were constructed by digesting 5 μg of pRW3036, which harbors (CTG)\textsubscript{z}CAG\textsubscript{39}, and pRW3080, which harbors (CTG)\textsubscript{z}CAG\textsubscript{80}, respectively, with NheI and XbaI (New England Biolabs Inc.). The digested DNAs were electrophoresed on a 6% polyacrylamide gel. The fragments containing the linearized vector were eluted from the gel and purified by phenol extraction. The inserts and vectors were ligated by treatment for 16 h at 16 °C with one unit of T4 DNA ligase (U.S. Biochemical Corp.). The ligation mixture was precipitated and resuspended in 5 μl of TE (45) and used to transform *E. coli* JC10289 via electroporation. Transformants were selected on LB agar plates and grown in 10-ml culture tubes at 37 °C in the presence of ampicillin (75 μg/ml). Plasmid DNAs were isolated and characterized by restriction mapping. The insert size and orientation were confirmed to be the expected size in orientation II.

pRW4437 and pRW4439 were deletion products of pRW3239 and pRW3238, respectively. They were obtained by repeated recultivation of *E. coli* AB1157 transformed with either pRW2239 or pRW3238, respectively (2, 3). Initially, *E. coli* AB1157 was transformed with either pRW2239 or pRW3238, and then transformants were selected on LB plates and subsequently grown in LB medium containing tetracycline (20 μg/ml) until the end of log phase (A\textsubscript{600} = 0.9). The cells were harvested, and the DNA was isolated and purified. The DNA was electrophoresed on 1.0%-agarose gels, and deletion products were identified. If no deletion products were detected, then 1 μl of the culture was used to inoculate fresh medium, and the cells were grown again until late log phase. The cultures were kept growing logarithmically by diluting them 10-fold with fresh, sterile medium. The recultivation was repeated until deletion products were observed. The deletion products were detected as faster migrating supercoiled DNA. Plasmids that contain deletions of TRS are smaller than their respective parent plasmids and thus migrate faster. The TRS in the pACYC184 derivatives were unusually stable and required five or more recultivations to generate a deletion product. The band that contained the deletion was excised from the gel, purified, and used to transform *E. coli* HB101 (2, 3, 5, 6, 11, 12). Transformants were subsequently characterized by restriction mapping and DNA sequencing to confirm that the deletion occurred within the TRS.

All of the pUC19 derivatives contain no interruptions (mutations) within the CTG-CAG sequences, except pRW3755 (which contains (GCT)\textsubscript{11}ACT(GCT)\textsubscript{11}ACT(GCT)\textsubscript{11}) and pRW3755 (which contains (GCT)\textsubscript{27}ACT(GCT)\textsubscript{14}ACT(GCT)\textsubscript{14}). All of the pUC19 derivatives contain non-human flanking sequences 19 bp proximal to and 13 bp distal to the origin of replication (5). The pACYC184 derivatives also contain human flanking sequences 19 bp proximal to and 43 bp distal to the origin of replication (2).

**Bacterial Strains**—The genotypes of *E. coli* AB1157 and the recombinant deficient strains JC10289, JC5519, and JC10287 used in the two-plasmid recombination system were described elsewhere (44). All strains were obtained from the *E. coli* Genetic Stock Center, Yale University, New Haven, CT.

**Standard Genetic Techniques**—Unless otherwise noted, DNA isolation and agarose gel and polyacrylamide gel electrophoreses were carried out according to standard laboratory protocols (45). Transformation were performed by electroporation (46, 47). All plasmids were initially transformed into *E. coli* JC10289 or HB101 (which are recA strains), cultured until mid-log phase, and purified by alkali lysis. The plasmids (see Fig. 1) were then used for cotransformation of the above strains in the two-plasmid recombination system. For the cotransformation experiments, each strain was cotransformed with a variety of plasmid combinations as shown in Table I. The experiments involving the cotransformation of both pUC19 and pACYC184 derivatives containing TRS in orientation II have been published (44). Additional results are listed in Table III. 40 μl of washed cells of each strain (5 × 10\textsuperscript{9} cells/ml) were transformed with 1 μl of the supercoiled DNAs (0.5 μg/ml) listed in Fig. 1. For the experiments involving cotransformation, the supercoiled plasmid volume was equally divided between the two test plasmids. A voltage of 1950 V was delivered for 4.1–5.8 ms. The current size was 0.2 mm.

Cotransformants were selected on LB agar plates containing ampicillin and tetracycline. After cotransformation, the cells were allowed to recover in 800 μl of SOC medium and kept at 37 °C for 1 h or longer. The cells were plated on LB agar that contained ampicillin (75 μg/ml) and tetracycline (20 μg/ml) and individual colonies were grown for 4–16 h at 37 °C under aerobic conditions.

Plasmid purification, gel electrophoresis, and DNA sequence analyses were conducted as described (3, 5, 6, 11, 12). Restriction mapping reactions and ligase reactions were conducted as described (45). The uncertainty of the accurate measurement of repeat size varies with repeat length. The error for lengths greater than 130 repeats, between 65 and 130 repeats, and less than 65 repeats was 6, 4, and 2%, respectively. The plasmid inserts and linking sequences were characterized by restriction analyses on both strands with Sequenase (version 2.0). The pACYC184 primers, purchased from Genosys Inc. were the following: primer 4244 (ACGGTCTTTAAAAAGGCCG), which 3′-terminates at map position 95; primer 4245 (CGTCAGTAGCTGAAACAGGAGG), which 5′-terminates at map position 522.

Percentage of Gene Conversion versus Unequal Crossing Over—The products of cotransformation of *E. coli* AB1157 with pRW3036 and pRW3239 were analyzed by DNA sequencing. The pUC19 primer 1211 (New England Biolabs Inc.), which binds 21 bp upstream of the polylinker, was used in the Genosys sequence 2.0 kit; the triplet repeats were cloned into the *Hinc*II site of the polylinker of pUC19. The recombination pathway was determined by analyses of the distals (3′ of the TRS) flanking sequences. The recombination products were scored as gene conversion events if the TRS flanking sequences were both derived from pUC19. Alternatively, the recombination products were scored as unequal crossing over events if the distal flanking sequence was derived from pACYC184.

Restriction mapping was conducted to determine whether crossing over or gene conversion cause expansion(s). Certain pairs of restriction enzymes were used. Products of cotransformation that had expansion(s). Cross over events would generate products that contain two TRS tracts (see Fig. 3B in Ref. 44). One of the tracts would be detected by BstZ171 and NdeI. A second of enzymes, BspE1 and PstI, would destroy the other TRS tract. The products from these digestions were separated on 7% polyacrylamide gels, and the sizes of the fragment(s) were compared with the EcoRI/PstI digest. A third of enzymes, NdeI and Hinfl, was used to destroy fragments derived by homologous recombination. When these products contained repeat tracts in orientation II, Sap I and HindIII were used. The fragments were separated on 7% polyacrylamide gels, and the sizes of the fragment(s) were compared with the EcoRI/PstI control fragments.

**Sequential Transformation**—The pUC19 derivatives pRW3036 and pRW4011 were used to separately transform *E. coli* AB1157. Transformants were selected on LB agar plates that contained ampicillin (75 μg/ml). Individual colonies were selected and grown until mid-log phase, after which they were prepared as competent cells for electroporation (45). Hence these cells harbored a pUC19 derivative containing a TRS, namely, (CTG-CAG)\textsubscript{z} or (CAG-CTG)\textsubscript{z}. An aliquot of each new strain was analyzed, and the size of the TRS was confirmed by restriction analyses.

Subsequently, the pACYC184 derivatives pRW3238 and pRW3239 were used to transform the new strains (*E. coli* AB1157 harboring pRW3036 or *E. coli* AB1157 harboring pRW4011). Cotransformants were selected on LB agar plates that contained ampicillin (75 μg/ml) and tetracycline (20 μg/ml) and cultured as described above. Conversely, the pACYC184 derivative, pRW3238 (which contains (CAG-CTG)\textsubscript{z}) was used to transform *E. coli* AB1157. Individual clones that formed on plates that contained tetracycline (20 μg/ml) were cultured and prepared as competent cells (45). This new strain (*E. coli* AB1157 harboring pRW3238) was used for the stepwise transformation of the pUC19 derivatives (see Fig. 1). Cotransformants were selected on LB agar plates that contained ampicillin (75 μg/ml) and tetracycline (20 μg/ml).
Gene Conversion and Triplet Repeat Expansion

RESULTS

Effect of TRS Orientation on Interplasmid Recombination—The potential influence of TRS orientation on recombination-mediated expansions was studied biochemically and genetically in a two-plasmid recombination system in E. coli (44). Derivatives of nonhomologous plasmids were used to cotransform isogenic, recombination proficient and deficient strains. Each strain was cotransformed with plasmids containing different lengths of (CTG-CAG) tracts cloned in either orientation I or orientation II (defined in Fig. 1). In addition, the number of interruptions (defined under “Experimental Procedures”) within the tracts also varied. The plasmids were derivatives of pUC19, which harbors the ampicillin resistance gene, and pACYC184 derivatives, which harbors the tetracycline resistance gene (Fig. 1). Thus, the only segments of the nonhomologous vectors are identical except that the orientation of the repeat sequence is switched from orientation I to orientation II. The double-stranded plasmids are shown schematically on the right. The pUC19 derivatives are represented by heavy black circles, and the orientations of the TRS are indicated as either orientation I or orientation II. The double-stranded pACYC184 derivatives are represented by thin circles, and the TRS orientations are indicated as either orientation I or orientation II. Relevant restriction sites are shown. All plasmids have SacI (S) and HindIII (H) sites in common, but the pUC19 derivatives also contain single PstI (P), NdeI (N), and SphI (S′) sites, whereas all pACYC184 plasmids contain single BtsI17 (B) and BspEI (B′) sites. The origins of replication (44), which are unique for the two nonhomologous plasmids, are shown as open and striped boxes, respectively. The ampicillin resistance gene (Am′) is designated as a filled box on pUC19. The tetracycline resistance gene (Tc′) is designated as a checkered box on pACYC184. pUC19 is a high copy number plasmid with ~500 copies/cell (45), whereas pACYC184 has a copy number of ~10/cell (44). The plasmids are not drawn to scale.

Fig. 1. Plasmids containing (CTG-CAG) and (CAG-CTG) repeats used in the two-plasmid recombination system. These plasmids are derivatives of the unidirectional replicating pUC19 and pACYC184. Plasmids pRW3036 through pRW4437 are derivatives of pACYC184 (2, 3, 5, 6, 11, 12). The types of sequences and their lengths are listed. For example, pRW4011 is a pUC19 derivative that contains a pure insert of (CAG-CTG)n in orientation II. All pUC19 derivatives contain pure uninterrupted repeating tracts of (CTG-CAG) or (CAG-CTG) sequences except pRW3753 and pRW3755 (3). The sequence in pRW3239 is not homogenous but contains two G-to-A point mutations at positions 28 and 69 as described previously (3, 11). The orientation of the TRS with respect to the direction of replication is defined as follows: plasmids that contain the (CTG) sequence as the lagging strand template are designated CTG-CAG and are in orientation I, plasmids that contain the (CTG) sequence as the lagging strand template are designated CAG-CTG and are in orientation II. For each pair of plasmids with the same length of TRS (i.e. pRW4015 and pRW4011), the relative locations of the origin of replication (Ori), the antibiotic resistance gene, and the TRS are identical except that the orientation of the repeat sequence is switched from orientation I to orientation II. The double-stranded plasmids are shown schematically on the right. The pUC19 derivatives are represented by heavy black circles, and the orientations of the TRS are indicated as either orientation I or orientation II. The double-stranded pACYC184 derivatives are represented by thin circles, and the TRS orientations are indicated as either orientation I or orientation II. Relevant restriction sites are shown. All plasmids have SacI (S) and HindIII (H) sites in common, but the pUC19 derivatives also contain single PstI (P), NdeI (N), and SphI (S′) sites, whereas all pACYC184 plasmids contain single BtsI17 (B) and BspEI (B′) sites. The origins of replication (44), which are unique for the two nonhomologous plasmids, are shown as open and striped boxes, respectively. The ampicillin resistance gene (Am′) is designated as a filled box on pUC19. The tetracycline resistance gene (Tc′) is designated as a checkered box on pACYC184. pUC19 is a high copy number plasmid with ~500 copies/cell (45), whereas pACYC184 has a copy number of ~10/cell (44). The plasmids are not drawn to scale.

The recombination behavior of these plasmids (Fig. 1) was analyzed in four E. coli strains that differed in their recombination capacity. Cotransformation of the plasmids is facile (44, 48). Each E. coli strain (AB1157, JC5519, JC10287, and JC10289) was transformed and cotransformed with the individual plasmids (Fig. 1). All cotransformants were grown to an A600 of between 0.5 and 0.9. As a control for the cotransformation of the plasmid pairs, each strain was cotransformed with the vector pACYC184, along with various pUC19 derivatives that contained CTG-CAG of different lengths, orientations, and extents of interruptions. Finally, the four strains were cotransformed with all possible pairwise permutations of the plasmids shown in Fig. 1. Thus, the only segments of the nonhomologous plasmids that had identical sequences were the CTG-CAG inserts, albeit in different orientations. Cotransformants were cultured in parallel and harvested under identical conditions. The DNA was purified and analyzed on agarose as well as native and denaturing polyacrylamide gels. The number of CTG-CAG expansion products and the magnitude of the expansions were compared between the different strains.

Instability of Triplet Repeats—The eight pUC19 derivatives (Fig. 1) were separately cotransformed with the four pACYC184 derivatives. These 32 experiments were conducted in four different isogenic strains. First, the pACYC184 derivatives that contained different lengths of CTG-CAG in both orientations were cotransformed with various pUC19 derivatives that contained different lengths of CTG-CAG in both orientations. Second, each strain was cotransformed with the vector pACYC184 along with various pUC19 derivatives that contained different lengths of CTG-CAG in both orientations. Third, the vector pUC19 was cotransformed along with various pACYC184 derivatives that contained different lengths of CTG-CAG in both orientations. Hence, 128 experiments (32 × 4, for each strain) were conducted.

Restriction analyses of plasmid DNA isolated from individual clones grown for ~6 h (15 generations) in the parent E. coli AB1157 showed that multiple fold expansions occurred within the (CTG-CAG) or (CAG-CTG) tracts. No significant differences in the number of expansions nor the magnitude of the expan-
Gene Conversion and Triplet Repeat Expansion

FIG. 2. Polyacrylamide gel electrophoresis analyses of transformation and cotransformation products from E. coli. Purified supercoiled plasmids containing (CTG-CAG) or (CAG-CTG) (Fig. 1) were used to transform or cotransform E. coli AB1157. The DNA was isolated, and the (CTG-CAG) or (CAG-CTG) containing restriction fragments were subsequently electrophoresed. The electrophoretic mobility of the TRS containing restriction fragments is difficult to interpret because the change in size of the products compared with the (CTG-CAG) or (CAG-CTG) containing restriction fragments is small. The bands were detected by the electrophoretic mobility of the TRS containing restriction fragments comparing the electrophoretic mobility of the TRS containing restriction fragments with the electrophoretic mobility of the TRS containing restriction fragments expected for a Sac1/I fragment containing 36, 50, 80, 104, or 206 repeats of (CTG-CAG) or (CAG-CTG) on a 7.0% polyacrylamide gel. Lane A contains the 1-kilobase pair DNA ladder (Life Technologies, Inc.). The sizes of these bands are 517, 502, 398, 344, 298, 220, 201, and 154 bp. All plasmid products (lanes B–Q) were digested with Sac1/I, Pst1/Lane B, the Sac1/I restriction fragment of pRW3036 (which contains (CTG-CAG) in orientation I) Lane C, pRW4011 (which contains (CAG-CTG) in orientation II). Lane D, pRW3080 (which contains (CTG-CAG) in orientation I) Lane E, cotransformation of pRW4011 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CTG-CAG) in orientation I) Lane F, transformation of pRW4011 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CAG-CTG) in orientation II). Lane G, transformation of pRW4011 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CAG-CTG) in orientation II). Lane H, transformation of pRW4011 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CAG-CTG) in orientation II). Lane I, transformation of pRW4011 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CAG-CTG) in orientation II). Lane J, transformation of pRW3036 (which contains (CTG-CAG) in orientation I and pRW3238 (which contains (CAG-CTG) in orientation II). Lane K, transformation of pRW3036 (which contains (CTG-CAG) in orientation I and pRW3238 (which contains (CAG-CTG) in orientation II). Lane L, transformation of pRW4436 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CAG-CTG) in orientation II). Lane M, transformation of pRW4436 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CAG-CTG) in orientation II). Lane N, transformation of pRW4480 (which contains (CTG-CAG) in orientation I and pRW3238 (which contains (CAG-CTG) in orientation II). Lane O, transformation of pRW4480 (which contains (CTG-CAG) in orientation I and pRW3238 (which contains (CAG-CTG) in orientation II and pRW3238 (which contains (CAG-CTG) in orientation II). Lane P, transformation of pRW3080 (which contains (CTG-CAG) in orientation I and pRW4439 (which contains (CAG-CTG) in orientation II). Lane Q, transformation of pRW3080 (which contains (CTG-CAG) in orientation II and pRW4439 (which contains (CAG-CTG) in orientation II). The plasmids used for transformation and cotransformation are listed above each lane. The expected positions of five Sac1/Pst1 fragments of various lengths of CTG-CAG in orientation I or II (36, 50, 80, 104, and 206 repeats) are indicated on the left of the 1-kilobase pair ladder shown in lane A. The first three lanes (lanes B–D) contain the products of cotransformations of various pUC19 and pACYC184 derivatives containing the repeat inserts of different lengths and orientations.

Considered the singly transformed cells, lane B shows a 169-bp band, which is the expected size of a fragment containing 36 repeats of (CTG-CAG) plus the myotonic dystrophy human flanking sequences. Similarly, lanes C and D show 240- and 301-bp bands, which are the expected sizes of fragments containing 50 and 80 repeats, respectively, of (CAG-CTG) plus human flanking sequences.

Considering the cotransformed cells, lane E shows the Sac1/Pst1 fragments from E. coli AB1157 cotransformed with pRW4011 (a pUC19 derivative with (CAG-CTG)36 in orientation I) and pRW3238 (a pACYC184 derivative with a (CAG-CTG)100 tract in orientation II). Surprisingly, instead of recovering the original length insert (lane C), new products that were severalfold longer than the progenitor starting material length were found. Comparison between the product in lane C with the products in lane E clearly demonstrate that the original starting material length (CAG-CTG)50 was converted to several expansion products, some of which were expanded by more than 2-fold. The expansion products, (CAG-CTG)79 and (CAG-CTG)99, represent the majority of the species. The difference in size between these expansion products suggests an incremental length increase because the progenitor insert length of (CAG-CTG)50 was expanded to (CAG-CTG)59 (an increase of 29 repeats) and/or expanded to (CAG-CTG)99 (an increase of 30 repeats). However, additional sample analyses revealed this to be coincidental because there was no correlation between the size of the insert in the progenitor plasmid and their expansion products.

Lane F shows the analysis of another isolate of the experiment shown in lane E; however, the magnitude of expansion and the number of expansion products are not the same as those contained in lane E. Therefore, a stochastic process that is recombination-dependent must be responsible for generating these expansion products. Again, no deletion products were observed. Another example of expansions that occur regardless of the insert orientation is shown by the cotransformation of pRW4011 and pRW2329 (which contains (CTG-CAG)175 in orientation I) (lane G). Unlike the previously described products (lanes E and F), the sample in lane G contains deletion products as well as expansions. The extent and the magnitude of the deletion products were similar to those recovered from the recombinant deficient strains (data not shown). The size of the deletion from the progenitor starting material compared with the sizes of the expansions is not an equal number (deletion of 10 repeats and expansions of 34 and 59 repeats, respectively). Thus, reciprocal exchange between the pUC19 and pACYC184 derivatives is unlikely. Lane G shows the presence of expansions as found in lanes E and F where the orientation of
### TABLE I

**Summary of expansions of (CTG-CAG) or (CAG-CTG)**

Plasmids (Fig. 1) were cotransformed into recombination-proficient *E. coli*, and DNA was isolated from individual colonies grown as described under “Experimental Procedures.” The lengths of the repeat expansions in the pUC19 derivatives were determined by DNA sequencing and/or by restriction mapping. Each line in the table represents a length determination on the recombinant DNA product from an individual clone that was grown on plates containing ampicillin and tetracycline.

Section A contains the products of three pUC19 derivatives harboring (CTG-CAG) in orientation I cotransformed with a pACYC184 derivative (pRW3238) harboring (CAG-CTG)$_{100}$ in orientation II. Section B contains the products of a pUC19 derivative (pRW3080) harboring (CTG-CAG)$_{80}$ in orientation I cotransformed with a pACYC184 derivative (pRW4439) harboring a (CAG-CTG)$_{39}$ sequence in orientation II. Section C contains the products of five pUC19 derivatives with different lengths of (CAG-CTG) in orientation II cotransformed with a pACYC184 (pRW3239) derivatives harboring (CTG-CAG)$_{175}$ in orientation I. Section D contains the products of three pUC19 derivatives with different lengths of (CAG-CTG) in orientation II cotransformed with a pACYC184 (pRW3238) derivatives harboring (CAG-CTG)$_{100}$ in orientation II.

| Plasmids | Length of repeat tracts | Size of repeats in products of pUC19 derivative | Increase in length of TRS (Number of repeats) |
|----------|------------------------|-----------------------------------------------|---------------------------------------------|
| **A. Cotransformation of TRS in orientation I (pUC19 derivatives) + orientation II (pRW3238; (CAG - CTG)$_{100}$)** | | | |
| pRW3036 (CTG-CAG)$_{36}$ | 210 | 174 | 170 |
| | 209 | 173 | 160 |
| | 186 | 150 | 140 |
| | 185 | 149 | 139 |
| | 184 | 148 | 138 |
| | 170 | 134 | 124 |
| | 165 | 129 | 119 |
| | 146 | 110 | 90 |
| | 85 | 49 | 0 |
| | 80 | 44 | 0 |
| | 72 | 36 | 0 |
| | 58 | 22 | 0 |
| | 36 | 0 | 0 |
| **pRW4015 (CTG-CAG)$_{50}$** | 80 | 30 | 0 |
| | 78 | 28 | 0 |
| | 66 | 16 | 0 |
| | 50 | 0 | 0 |
| **pRW3080 (CTG · CAG)$_{50}$** | 195 | 115 | 175 |
| | 170 | 90 | 65 |
| | 106 | 26 | 26 |
| | 97 | 17 | 17 |
| | 85 | 5 | 5 |
| | 80 | 0 | 0 |
| **B. Cotransformation of TRS in orientation I (pRW3080; (CTG-CAG)$_{39}$) + orientation II (pRW4439; (CAG-CTG)$_{39}$)** | | | |
| pRW3080 (CTG-CAG)$_{36}$ | 83 | 3 | 3 |
| | 80 | 0 | 0 |
| | 79 | 1 | 1 |
| | 78 | 2 | 2 |
| | 75 | 5 | 5 |
| | 72 | 8 | 8 |
| | 71 | 9 | 9 |
| | 70 | 10 | 10 |
| | 52 | 18 | 18 |
| | 57 | 23 | 23 |
| | 31 | 49 | 49 |
| **C. Cotransformation of TRS in orientation II (in pUC19 derivatives) + orientation I (pRW3239; (CTG-CAG)$_{175}$)** | | | |
| pRW4436 (CAG-CTG)$_{36}$ | 179 | 163 | 150 |
| | 175 | 139 | 130 |
| | 120 | 84 | 80 |
| | 119 | 83 | 75 |
| | 116 | 80 | 70 |
| | 106 | 66 | 60 |
| | 91 | 55 | 50 |
| | 86 | 50 | 46 |
| | 64 | 34 | 30 |
| | 67 | 31 | 27 |
| | 41 | 5 | 5 |
| | 40 | 4 | 4 |
| **pRW4011 (CAG-CTG)$_{50}$** | 300 | 250 | 250 |
| | 175 | 125 | 125 |
| **D. Cotransformation of TRS in orientation II (in pUC19 derivatives) + orientation II (pRW3238; (CAG-CTG)$_{100}$)** | | | |
| pRW4436 (CAG-CTG)$_{36}$ | 212 | 176 | 172 |
| | 211 | 175 | 171 |
| | 208 | 172 | 170 |
| | 205 | 169 | 168 |
| | 175 | 139 | 138 |
| | 120 | 84 | 83 |
| | 117 | 81 | 80 |
| | 102 | 66 | 65 |
| | 91 | 55 | 54 |
| | 70 | 34 | 33 |
| | 66 | 30 | 29 |
| | 39 | 3 | 3 |
| | 36 | 0 | 0 |
| **pRW4011 (CAG-CTG)$_{50}$** | 175 | 125 | 125 |
the insert is reversed in the pACYC184 derivative.

Nevertheless, in some cases, similar products were formed for certain cotransformations (lane H). Lane H shows the restriction fragment products of pRW4011 and pRW3238 (which contains (CAG-CTG)_{100}) cotransformed into E. coli AB1157. Expansions are the only type of products recovered. Second, as in lane G, the progenitor insert length is completely absent. Although no deletion products were recovered, the expansion product seems to be identical to lane G. Thus, it is possible that different cotransformed DNAs yield the same products, indicating the random nature of the recombination events. Lane I contains another example of the restriction fragments from the cotransformation of pRW4011 and pRW3238 in E. coli AB1157. As for the duplicate experiment in lane F, the progenitor starting length of pRW4011 is recovered, but the expansion product is similar to one of the products shown in lanes E and G. However, unlike lane F, a small amount of a deletion product was detected in lane I. In summary, lanes E–I demonstrate that the recombination mechanism results in the formation of highly variable length alterations in the product (CTG-CAG) tracts regardless of repeat orientation.

Considering the cotransformation of other pUC19 derivatives, lane J shows that a 6-fold expansion product was recovered from the cotransformation of pRW3036 (which contains (CTG-CAG)_{57}) with pRW3238. This expansion product was longer than the sum of the lengths of the (CTG-CAG) tracts from both plasmids; other examples of this behavior were frequent (Table I, sections A, C, and D). Another clone from the cotransformation of pRW3036 with pRW3238 in E. coli AB1157 is shown in lane K. This sample did not contain expansions or deletions and is consistent with previous reports that shorter inserts tend to be more stable than longer repeat inserts in E. coli, yeast, or in humans (1, 3, 16, 44, 49–51).

Considering the products from the cotransformation of pRW4436 (which contains (CAG-CTG)_{56} in orientation II) with pRW3238, multiple fold expansions and a two repeat deletion product were recovered. Lane L shows the products of (CAG-CTG)_{56} that is expanded to (CAG-CTG)_{107} and (CAG-CTG)_{207}. The results from a different clone of pRW4436 cotransformed with pRW3238 (lane M) shows expansion products of (CAG-CTG)_{57} and (CAG-CTG)_{205}. In summary, the results of lanes J–M illustrate the random formation of multiple fold expansions and show that an expansion bias is associated with the recombination events.

Considering the cotransformation of pUC19 derivatives harboring longer (CTG-CAG) tracts than (CTG-CAG)_{56}, lane N shows the restriction fragment products from the cotransformation of pRW4480 (which contains (CAG-CTG)_{80}) in orientation II and pRW3238 (which contains (CAG-CTG)_{100} in orientation II). The (CAG-CTG)_{80} tract in the pUC19 derivative was expanded by 133 repeats. Another clone from the cotransformation of pRW4480 with pRW3238 (lane O) was expanded by 132 repeats. These similar expansion products represent some of the largest changes in repeat size observed in the two-plasmid recombination system. There seems to be an apparent maximum length to the expansion of the (CTG-CAG) repeat tract. Indeed, the largest expansion product detected was a (CAG-CTG)_{100} tract (Table I, section C). Although the reason for this length limit is uncertain, it was shown that long tracts of TRS have an increased frequency of deletions (1, 7, 8, 10–12, 16, 49, 51).

Surprisingly, when pACYC184 derivatives harboring shorter lengths of (CAG-CTG) than those harbored in the pUC19 derivatives were used in the two-plasmid system, the expansion bias of the pUC19 derivative was dramatically reduced. Lanes P and Q show the products from the cotransformation of pRW3080 (which contains (CTG-CAG)_{80}) and pRW4439 (which contains (CAG-CTG)_{90}). A variety of deletion products and an expansion of only three repeats were detected in this batch of experiments. In summary, we conclude that the E. coli recombination system significantly promotes the expansion of (CTG-CAG) triplet repeats in vivo for sequences cloned in either orientation I or II.

Quantitation of Expansions—More than 110 examples of expansions are listed in Table I, which is divided into four subsections. Section A contains the products from the cotransformation of pRW3238 (which contains (CAG-CTG)_{100} in orientation II) with various pUC19 derivatives containing (CTG-CAG) in orientation I. Section B contains the products from the cotransformation of pRW4439 (a pACYC184 derivative that contains (CAG-CTG)_{90} in orientation II) with pRW3080. Section C contains the products from the cotransformation of pRW3239 with various pUC19 derivatives containing (CAG-CTG) in orientation II. Each subsection is divided into individual entries grouped according to the plasmid pairs used for cotransformation of E. coli AB1157. Each entry represents the analysis of a single colony characterized by restriction mapping and DNA sequencing where appropriate.

Table I summarizes the magnitude of the expansion events. Several conclusions may be drawn: 1) expansions are found regardless of the orientation of the inserts in the vectors (Table I, sections A, C, and D); 2) shorter repeat lengths (36 (data not shown) and 39 units) in the pACYC184 derivatives do not support recombination-mediated expansions (Section B); 3) multiple fold expansions were routinely found (i.e. severalfold longer than the length of the pUC19 progenitor repeat as well as longer than the sum of the two progenitor repeat lengths); 4) the lengths of the expanded repeats were essentially random; 5) the expansion products were of discrete lengths, not a family of lengths as detected by smears of bands; 6) repeat tracts containing interruptions in the pUC19 derivatives (pRW3753 and pRW3755) are much less effective in generating expansion products than similar lengths of pure repeats (pRW4011 and pRW3080), respectively; and 7) the lengths of the interrupted repeat tracts in the low copy number pACYC184 derivatives were unchanged for the cases investigated (~20 isolates).

Effect of Intermittences—Point mutations (polymorphisms) within TRS have an important consequence for hereditary diseases because they increase the stability of the triplet repeat sequences (52–54). Hence, we made direct comparisons of the extent and magnitude of expansions between interrupted and pure (CAG-CTG) repeat tracts used in the recombinase system. E. coli AB1157 was cotransformed with pRW3753 (which contains (CAG-CTG)_{57} with two G-to-A point mutations within the repeat at positions 28 and 43) and pRW3239 and cultured as described above (under “Experimental Procedures”). Parallel control experiments using pRW4011 (which contains a pure repeat tract of similar length as pRW3753) and pRW3239 were also conducted. The products were purified and analyzed by restriction mapping. The products from the progenitor
TABLE II

| Expansions | Percentage of products | Deletions | Original length |
|------------|------------------------|-----------|----------------|
| 53         | 0                      | 47        |                |
| 40         | 23                     | 37        |                |
| 63         | 0                      | 37        |                |
| 26         | 36                     | 38        |                |
| 59         | 0                      | 41        |                |
| 24         | 36                     | 41        |                |
| 36         | 23                     | 42        |                |
| 49         | 6                      | 45        |                |
| 54         | 0                      | 46        |                |
| 45         | 0                      | 55        |                |

pRW3753 repeats were predominately deletions; less than 10% of the products were expansions of less than 15 repeats (approximately one-fourth of the original length). No multiple fold expansion products were recovered when pUC19 derivatives contained point mutations within the repeat tract.

The effect of interruptions was studied further with pRW3755 (which contains (CAG-CTG)36, which harbors a longer repeat sequence than pRW3753 and also contains two G-to-A interruptions (point mutations) at positions 28 and 59); this plasmid was cotransformed with pRW3239 into E. coli AB1157. Individual clones were selected at random, grown to mid-log phase, and harvested, and the DNA was purified. To quantitate the products, and EcoRI/PstI double digestion was performed in EcoRI buffer at 37 °C for 4 h to release the fragment containing the TRS. The restriction fragment was subsequently labeled with the Klenow fragment of the E. coli DNA polymerase I and [α-32P]dATP. The extent of the instabilities of (CAG-CTG) was measured by exposing the dried 7% polyacrylamide gel to a Molecular Dynamics PhosphorImager screen followed by scanning. The amount of radioactivity estimated by the signal intensity in the bands corresponding to the restriction fragments was measured as a proportion of the radioactivity in the lane below the vector band. The disappearance of the restriction fragment corresponding to the original starting length ((CAG-CTG) repeats) was compared to the appearance of restriction fragments that were longer than the original starting length (expansions) and restriction fragments that were shorter than the original starting length (deletions).

Expansions are the Dominant Products in the Two-plasmid Recombination System—When the products of recombination-mediated instability were categorized according to the number of expansion events versus the number of deletion events, there was an expansion bias. To specifically address this issue, a cotransformation experiment using pRW4011 (which contains (CAG-CTG)36) and pRW3239 (which contains (CAG-CTG)175) was conducted in E. coli AB1157 and compared with control experiments. After 20 generations, the progenitor plasmid (pRW4011) was analyzed by restriction mapping and DNA sequencing (Table II). All of the clones contained expansions of the (CAG-CTG) tract. Quantitation of the expansion and deletion products from the cotransformation of E. coli AB1157 with pRW4011 and pRW3239 was performed by scanning polyacrylamide gels containing the end-labeled fragments. The (CAG-CTG) tracts were excised from the products with EcoRI and PstI, end-labeled with [α-32P]dATP, and electrophoresed on a 7% polyacrylamide gel. The amount of instability was measured by exposing the dried gel to a Molecular Dynamics PhosphorImager screen. The percentage of molecules that contained expansions and/or deletions was measured as the proportion of the total radioactivity in the lane below 1.6 kilobase pairs (the linearized vector backbone).

Ten clones were chosen at random, and five of the clones contained only expansions; no deletions were detected by separation of end-labeled fragments on polyacrylamide gels (Table II). The other five clones contained both expansions and deletions as well as the original length fragment. The detection limit of radiolabeled materials is ~1 × 10^-12 mol. Thus, when no deletion products were detected and the only products present were expansions, the frequency of formation of expansions compared with that of deletions is as great as 3 × 10^3. One out of the 10 samples contained a small amount of deletion products (about 6%), whereas the majority of instability was manifested as expansions. The remaining four clones contained an equal number of expansion and deletion products. In summary, the average expansion/deletion ratio of these 10 samples was 1.5 × 10^1. Although these biochemical determinations give a reasonable estimate of the frequency of expansions to deletions, we acknowledge that a genetic assay would be desirable. A genetic assay is under current investigation.2

Gene Conversion Is the Recombination Mechanism—The mechanism of recombination responsible for generating the observed multiple fold expansions was determined. DNA sequencing analyses were performed on the products formed by the cotransformation of pRW3036, which contains (CAG-CTG)36, and pRW3239, which contains (CAG-CTG)175, in E. coli AB1157. The cotransformations were repeated numerous times on different days, and individual colonies were analyzed. The flanking sequences adjacent to the repeat inserts are unique to the pUC19 and pACYC184 derivatives. Plasmid pRW3036 contains nonhuman flanking sequences adjacent to the repeat, whereas pRW3239 contains human flanking sequences adjacent to the repeat. Therefore, the identification of

---

2 A. Pluciennik and R. D. Wells, unpublished data.
The expansion products from the cotransformation of pRW3036 (which contains (CTG-CAG)36) and pRW3239 (which contains (CTG-CAG)10) were sequenced using the pUC19 primer 1211 that binds to the human flanking sequence near to the 5' end of the (CTG-CAG) insert. Both plasmids contain (CTG-CAG) inserts in orientation I. The isolated DNA was electrophoresed through a 7% denaturing polyacrylamide gel in Tris-Borate-EDTA buffer at 30 V/cm at 44 °C. The gels were washed with 15% acetic acid, 5% methanol, dried, and exposed to film or to a PhosphorImager screen. The numbers show the length of each expansion product and indicate how many repeats were added to the progenitor length of (CTG-CAG)10. The flanking sequence 3' to the (CTG-CAG) insert is listed for each isolate. When the 3'-flanking sequences were identical to that of pUC19, the fate of flanking sequences was scored as "retention." When the 3'-flanking sequences were identical to that of pACYC184, they were scored as "exchange."

| Isolate (clone) | Sequence at 3' end of (CTG) repeat | Number of (CTG) repeats | Increase in length of (CTG) | Fate of flanking sequences (retention or exchange) |
|-----------------|------------------------------------|-------------------------|--------------------------|---------------------------------------------|
| 1               | CTGGGAAAGAGCT                      | 47                      | 11                       | Retention                                  |
| 2               | CTGGGAAAGAGCT                      | 41                      | 5                        | Retention                                  |
| 3               | CTGGGAAAGAGCT                      | 43                      | 7                        | Retention                                  |
| 4               | CTGGGAAAGAGCT                      | 58                      | 22                       | Retention                                  |
| 5               | CTGGGAAAGAGCT                      | 45                      | 4                        | Retention                                  |
| 6               | CTGGGAAAGAGCT                      | 40                      | 4                        | Retention                                  |
| 7               | CTGGGAAAGAGCT                      | 56                      | 20                       | Retention                                  |
| 8               | CTGGGAAAGAGCT                      | 53                      | 17                       | Retention                                  |
| 9               | CTGGGAAAGAGCT                      | 43                      | 7                        | Retention                                  |
| 10              | CTGGGAAAGAGCT                      | 49                      | 13                       | Retention                                  |
| 11              | CTGGGAAAGAGCT                      | 50                      | 14                       | Retention                                  |
| 12              | CTGGGAAAGAGCT                      | 53                      | 17                       | Retention                                  |
| 13              | CTGGGAAAGAGCT                      | 42                      | 6                        | Retention                                  |
| 14              | CTGGGAAAGAGCT                      | 46                      | 10                       | Retention                                  |
| 15              | CTGGGAAAGAGCT                      | 44                      | 8                        | Retention                                  |

the flanking sequences on both sides of the repeat inserts contained in the expansion products enables differentiation between recombination that occurs via crossing over or gene conversion. Gene conversion is defined as nonreciprocal homologous recombination with retention of the flanking sequences.

Dideoxy sequencing of the expansion products was conducted using T7 DNA polymerase (Sequenase) and the M13/pUC forward primer 1211, which is unique to the 5' end of the polylinker region of pUC19. Table III shows the analyses of 15 colonies that had varying extents of expansions. The flanking pUC19 sequences were retained for all 15 products; no cases of exchange were found. Although other samples with multiple fold expansions were also sequenced, these samples were of an appropriate size to enable visualization of the entire (CTG-CAG) tracts and the 5'- and 3'-flanking sequences on a single gel. Larger expansion products (as determined by restriction mapping) were also sequenced, and the 3' end of the repeat did contain the pUC19 vector flanking sequence, but the exact length of the expansion could not be determined because the 5' end of the repeat had migrated off the bottom of the gel.

When the flanking sequence 3' to the repeat tract contained CTGGGAAAGAGCT as well as the rest of the nonhuman flanking sequence followed by the pUC19 vector sequence, the type of recombination was classified as gene conversion. In summary, DNA sequence analyses confirmed that all 15 products were expanded by gene conversion. However, crossing over does occur in this system because long cointegrated DNAs were reported (44).
Gene Conversion and Triplet Repeat Expansion

The DNA was isolated and analyzed by restriction mapping, which confirmed the presence of pRW3036 or pRW4011, as expected. Thus, the progenitor plasmids were transmitted stably. These new strains (E. coli AB1157 harboring pRW3036 and AB1157 harboring pRW4011) were rendered competent. The competent cells were then used for transformation by pACYC184 derivatives, pRW3238 or pRW3239 (Fig. 1) and selected on agar plates containing ampicillin and tetracycline. The DNAs from these clones were digested with SacI/PstI, and the fragments were separated on 7% polyacrylamide gels. The products from these stepwise transformations were expansions and deletions with a bias in favor of expansions. The number and magnitude of the expansions were very similar, if not identical, to those products obtained from numerous experiments described in Fig. 1 and Table I.

For a thorough study of stepwise transformation, the order of plasmid introduction was reversed. E. coli AB1157 was transformed initially with the pACYC184 derivatives pRW3238 and pRW3239 and cultured on tetracycline. The cells were rendered competent and used for the transformation by the pUC19 derivatives (Fig. 1) and selected on plates containing ampicillin and tetracycline. The products were digested with SacI/PstI, and the fragments were separated on 7% polyacrylamide gels. The products of these transformations were expansions and deletions with a bias in favor of expansions. Thus, the order of the plasmid transformation (either pUC19 derivatives transformed first followed by pACYC184 derivatives or pACYC184 derivatives transformed first followed by pUC19 derivatives or both introduced simultaneously) did not influence the number or magnitude of the expansions.

**DISCUSSION**

This report describes the direct demonstration that gene conversion mediates the expansions of CTG-CAG\(^3\) repeats. Gene conversion is the principal recombination mechanism rather than unequal crossing over. Also, the results from a large number of gene conversion-mediated expansion studies with plasmids containing TRS inserts in all possible permutations of the orientations revealed little or no influence of orientation. Hence, these data provide further verification that recombination (not complementary strand slippage at the replication fork) is the responsible mechanism. Gene conversion (recombinational repair) is the nonreciprocal transfer of genetic information from one DNA duplex to another with no exchange of flanking sequences (55–57). Hence, this would explain the linkage disequilibrium of flanking markers in haplotype analyses (1, 52–54), if this recombination mechanism is also responsible for expansions in human cells. Whereas our paper (44) was the first to demonstrate the involvement of recombination in TRS expansions, other workers (58) demonstrated that gene conversion plays the major role in controlling the instability of large tandem repeats of ribosomal DNA sequences in yeast. Although the ribosomal tandem repeat sequences are much longer than triplets, the expansion mechanism is probably the same.

The gene conversion (recombinational repair) mechanism responsible for the repeat tract expansions is modeled in Fig. 3. The DNA replication fork stalls when it encounters a CTG-CAG sequence (8, 13, 27), which can result in double-stranded DNA breaks that widen into double-stranded gaps (41, 56, 59). When breaks occur in the repeat tract (15, 19) of the progenitor, the resultant gap may range in size from 40 to 400 repeats (55, 60). A search for homology then takes place, and the ends of the progenitor strands invade the TRS in the pACYC184 derivative (Fig. 3, upper right panel). After annealing, the gap is filled by recombination repair and ligated to finalize the expansion of the TRS tract in the pUC19 derivative. In addition to these standard recombination repair steps, it is possible with the triplet repeat tracts that DNA slippage may enable the formation of hairpin structures (center, right) (2–13, 21, 24–26, 61–64). Thus, these two processes may act in concert to enhance the formation of expansions. Others have suggested that gene conversion and break-induced replication are cooperative processes (65). Because double-stranded breaks occur frequently under normal growth conditions and recBCD and recA proteins are essential for recombinational repair (55, 66, 67), this also suggests a coupling of replication and recombination.

Our data (Fig. 2 and Table I) demonstrate that gene conversion mediates the multiple fold expansions of CTG-CAG sequences, irrespective of the orientation of the TRS tracts. These results are in stark contrast to some prior studies in E. coli (2–13) and yeast (15, 16, 68), which demonstrated a marked influence of the orientation of the repeat tracts on deletions and expansions. These results were explained by slippage of the complementary repeat tracts at the replication fork, which was enhanced by the preferential formation of DNA looped conformations (13, 63, 64, 69) that enabled bypass synthesis at the replication fork. When CTG repeats were in the lagging strand template, deletions were more likely than expansions. This behavior could also occur on the leading strand template (49). However, in the studies described herein in recombination-proficient cells with the two-plasmid system, the inversion of the repeat tract does not cause an alteration in the instability,
as expected from recombination mechanisms (55, 57, 65, 70). Recombinational repair (Fig. 3) is dependent on sequence homology but not sequence orientation. Hence, these results further verify that recombination is the mechanism involved in the multiple fold expansion process. Although gene conversion is the simplest mechanism to explain our data, we cannot eliminate the possibility of a double unequal cross-over event. Also, we demonstrated the requirement for recA as well as recBC functions. As shown in Fig. 3, recombination and DNA synthetic processes may act in concert to generate the instabilities.

The lengths of the repeat tracts as well as the absence or the presence of interruptions influences the expansion process. In our two-plasmid recombination system, the pACYC184 derivatives containing only 36 or 39 repeats were unable to effect expansions, whereas 100 or 175 repeats were extremely active, regardless of the length of the TRS in the pUC19 derivatives. Hence, this effect of length is consistent with prior reports (57) on recombination mechanisms and may explain why other workers observed only deletions in their yeast recombination system (30).

A hallmark of triplet repeat diseases is that intermediate size repeats lose a snip (single nucleotide polymorphism) and subsequently expand by severalfold (52–54). Without the loss of the interruption, the repeat is stably transmitted through a pedigree. We compared the effect of interruptions on gene conversion-mediated expansions with sequences containing no interruptions. However, when the interrupted pUC19 derivatives (pRW3753 and pRW3755) were analyzed, few or no expansions occurred. These plasmids contain two G-to-A point mutations; thus, if the complementary strands slip relative to each other, four mismatches will be generated. Prior work demonstrated that these mismatches effectively destabilize the formation of slipped structures (69). Thus, the polymerase complex is not likely to pause, and the replication fork will not collapse which will limit the number of double-stranded breaks that are inflicted. Whereas four mismatches in 219 bp is a small percentage, this extent of heterology (1.8%) was quite effective in eliminating the formation of multiple fold expansion products in our system. Similarly, other workers have recently demonstrated that as little as 1.2% heterology in a mammalian recombinant system reduces the effects of recombination (71).

Little or no correlation was found between the length of the CTG-CAG tracts in pACYC184 (100 and 175 repeats in length) and pUC19 derivatives with the size of the expansion products. Specifically, any individual clone from a single transformation may contain products of many different sizes. Some clones had only one expansion product, and others had several products. A few clones had both expansions and deletions of the TRS. However, the lengths of these instabilities were essentially random. Interestingly, the repeat tracts were expanded to generate discrete lengths rather than smears representing a large number of related-length molecules. The reason for this behavior is uncertain and is under further investigation. Furthermore, the frequency of expansions to deletions in this gene conversion system is extremely high (approximately 1.5 × 10⁻⁸-fold). This behavior is in marked contrast to the prior investigations in recombination-deficient E. coli and in yeast where expansions were substantially less frequent than deletions by a ratio of approximately 1:100 (2–13, 15, 16, 68). Thus, we believe that powerful recombination processes may be responsible for the large expansions that are observed in human genetic studies of myotonic dystrophy, spino-cerebellar ataxia type 8, Friedreich’s ataxia, and fragile X syndrome (reviewed in Ref. 1).

In summary, gene conversion is a robust mechanism for effecting multiple fold expansions of CTG-CAG tracts in this genetically and biochemically tractable system. Of course, it is desirable to extend these investigations into human systems to evaluate its role in disease pathogenesis. The development of a suitable human gene conversion system will be required to rigorously evaluate this question. Whereas this will present challenges, it may be noted that a recent investigation on individuals with an Ataxia-Telangiectasia-like disorder showed a link between genetic defects in double-stranded break repair genes and this disease of chromosomal instability (72). Hence, if a link is established between gene conversion mechanisms and the expansions (anticipation) observed in certain human hereditary neurological diseases, new targets for therapeutic interventions may be established.

Acknowledgments—We thank Adam Jaworski, Pawel Parniewski, and Michael Hebert for advice.

REFERENCES

1. Wells, R. D., and Warren, S. T. (eds) (1998) Genetic Instabilities and Hereditary Neurological Diseases, Academic Press, Inc., San Diego
2. Kang, S., Jaworski, A., Oohshima, K., and Wells, R. D. (1995) Nat. Genet. 10, 213–218
3. Bowater, R. P., Rosche, W. A., Jaworski, A., Sinden, R. R., and Wells, R. D. (1996) J. Mol. Biol. 264, 82–96
4. Shimizu, M., Gellibolian, R., Oostra, B. A., and Wells, R. D. (1996) J. Mol. Biol. 258, 614–626
5. Shimizu, M., Gellibolian, R., Oostra, B. A., and Wells, R. D. (1996) J. Biol. Chem. 271, 16773–16783
6. van den Ouweland, A. M. W., Deelen, W. H., Kunst, C. B., Uzielli, M. L. G., van Groenendael, J. B., and Zoghbi, H. Y. (1985) J. Mol. Biol. 191–193
7. van den Ouweland, A. M. W., Deelen, W. H., Kunst, C. B., Uzielli, M. L. G., van Groenendael, J. B., and Zoghbi, H. Y. (1985) J. Mol. Biol. 191–193
8. Bowater, R. P., Jaworski, A., Sinden, R. R., and Griffith, J. (1994) Science 265, 669–671
9. Kang, S., Oohshima, K., Jaworski, A., and Wells, R. D. (1996) J. Biol. Chem. 271, 16773–16783
10. Rowen, C. H., Kantrow, S. M., and Zakian, V. A. (1998) Science 278, 853–856
11. Schweitzer, J. K., and Livingston, D. M. (1997) Hum. Mol. Genet. 6, 349–355
12. Kohros, K., R., Stelenovic, L., Tran, H. T., Resnick, M. A., Gordenin, D. A., and Petes, T. D. (1999) Mol. Cell. Biol. 19, 2779–2789
13. Gordenin, D. A., Kunkel, T. A., and Resnick, M. A. (1997) Nat. Genet. 15, 116–119
14. Balakumaran, B. S., Freudenreich, C. H., and Zakian, V. A. (2000) Hum. Mol. Genet. 9, 93–100
15. Moore, H., Greenwell, P. W., Liu, C-P., Arnheim, N., and Petes, T. D. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 1504–1509
16. Maurer, D. J., O’Callaghan, B. L., and Livingston, D. M. (1996) Mol. Cell. Biol. 16, 6617–6622
17. Bringham, P. M., Scott, M. O., Wang, S., McPaul, M. J., Wilson, E. M., Garber, J. Y., Merry, D. E., and Fischbeck, K. H. (1995) Nat. Genet. 9, 191–193
18. Meine, D. G., Cobaugh, M. I., Ashizawa, J., and Caskey, C. T. (1995) Nat. Genet. 15, 193–196
19. Klement, I. A., Skinner, P. J., Kaytor, M. D., Yi, H., Hersch, S. M., Clark, H. B., Zoghbi, H. Y., and Orr, H. T. (1998) Cell 95, 41–53
20. Mangiagare, L., Sathisham, K., Majal, A., Mott, R., Seller, M., and Bates, D. P. (1997) Nat. Genet. 17, 197–200
21. Cummings, C. J., Orr, H. T., and Zoghbi, H. Y. (1999) Philos. Trans. R. Soc. Lond. Biol. Sci. 354, 1079–1081
22. Samadashwily, G. M., Raco, G., and Mirkin, S. M. (1997) Nat. Genet. 17, 298–304
23. Schumacher, J., Fuchs, R. P., and Bichara, M. (1998) J. Mol. Biol. 279, 1101–1110
24. Drouin-Lyons, T. and Topal, M. D. (1999) J. Biol. Chem. 274, 25975–25978
25. Richard, G. F., Dujon, B., and Haber J. E. (1995) Mol. Gen. Genet. 246, 871–882
26. Tbilidz, D., Mackenzie, A. E., Mettler, G., Barcelo, J., and Korneluk, R. G. (1992) Nat. Genet. 2, 192–195
27. O’Hoy, L. L., Tbilidz, C., Mahadevan, M. S., Neville, C. E., Barcelo, J., Hunter, A. G. W., and Korneluk, R. G. (1993) Science 259, 809–810
28. Deen, W. A., Deelen, W. H., Kunz, C. B., Uziel, M. L. G., Nelson, D. L., Warren, S. T., Oostra, G. A., and Halley, J. J. (1994) Hum. Mol. Genet. 3, 1823–1827
29. Losekoot, M., Hoogendoorn, E., Olmer, R., Jansen, C. C. A. M., Oostervik, J. C., Van den Ouweland, A. M. W., Halley, J. D., Warren, S. T.,
Gene Conversion and Triplet Repeat Expansion

40013

Willemsen, R., Oostra, B. A., and Bakker, E. (1997) J. Med. Genet. 34, 924–926
35. Brown, W. T., Houck, G. E., Jr., Ding, X., Zhong, N., Nolin, S., Glicksman, A., Dobkin, C., and Jenkins, E. C. (1996) Am. J. Med. Genet. 64, 287–292
36. Zhong, N., Kayanoja, E., Smiths, B., Pietrofesa, J., Curley, D., Wang, D., Ju, W., Nolin, S., Dobkin, C., Ryynannen, M., and Brown, W. T. (1996) Am. J. Med. Genet. 64, 226–233
37. Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., Howeler, C., Brunner, H., and Wieringa, B. (1994) Am. Soc. Hum. Genet. 54, 575–585
38. Richards, R. I., and Sutherland, G. R. (1992) Nat. Genet. 1, 7–9
39. Richards, R. I., Holman, K., Kozman, H., Kremer, E., Lynch, M., Pritchard, M., Yu, S., Mulley, J., and Sutherland, G. R. (1992) J. Med. Genet. 28, 818–823
40. Szemraj, J., Plucienniczak, G., Jaworski, J., and Plucienniczak, A. (1995) Gene (Amst.) 152, 261–264
41. Tishkoff, D. X., Filosi, N., Gaida, G. M., and Kolodner, R. D. (1997) Cell 88, 253–263
42. Richardson, C., Moynahan, M. E., and Jasin, M. (1998) Genes Dev. 12, 3831–3842
43. Lin, Y., Lukassovich, T., and Waldman, A. S. (1999) Mol. Cell. Biol. 19, 8353–8369
44. Jakupciak, J. P., and Wells, R. D. (1999) J. Biol. Chem. 274, 23468–23479
45. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
46. Hanahan, D., Jessee, J., and Bloom, F. R. (1991) Methods Enzymol. 204, 63–113
47. Dower, W. J., Miller, J., and Ragsdale, C. W. (1988) Nucleic Acids Res. 16, 6127–6145
48. Jaworski, J., Hsieh, W.-T., Blaho, J., A., Larson, J., E., and Wells, R. D. (1987) Science 238, 773–777
49. Iyer, R. R., and Wells, R. D. (1999) J. Biol. Chem. 274, 3865–3877
50. Kramer, P. R., Pearson, C. E., and Sinden, R. R. (1996) Hum. Genet. 98, 151–157
51. Ashizawa, T., Monckton, D. G., Vaishnav, S., Patel, B. J., Voskova, A., and Caskey, C. T. (1996) Genomics 36, 47–53
52. Kunst, C. B. and Warren, S. T. (1994) Cell 77, 853–861
53. Eichler, E. E., Holden, J. J., Popevich, B. W., Reiss, A. L., Snow, K., Thibodeau, S. N., Richards, C. S., Ward, P. A., and Nelson, D. L. (1994) Nat. Genet. 8, 88–94
54. Kunst, C. B., Zerynlick, C., Karickhoff, L., Eichler, E., Bullard, J., Chalifoux, M, Holden, J. J., Torreni, A., Nelsen, D. L., and Warren, S. T. (1996) Am. J. Hum. Genet. 58, 513–522
55. Szostak, J. W., Orr-Weaver, R. L., and Rothstein, R. J. (1983) Cell 33, 773–777
56. Cox, M. M. (1998) Genes Cells 3, 65–78
57. West, S. C. (1991) Annu. Rev. Biochem. 61, 603–640
58. Gangloff, S., Zou, H., and Rothstein, R. (1996) EMBO J. 15, 1715–1725
59. Harmon, F. G., and Kowalczykowski, S. C. (1998) Genes Dev. 12, 1134–1144
60. Cordiero-Stone, M., Makarov, A. M., Zaritskaya, L. S., and Griffith, J. D. (1999) J. Mol. Biol. 289, 1207–1218
61. Usdin, K., and Woodford, K. J. (1995) Nucleic Acids Res. 23, 4202–4209
62. Koo, J. D., Moseley, M. L., Schut, L. J., Benowitz, K. A., Bird, T. D., Day, J. W., and Ranum, L. P. W. (1999) Nat. Genet. 21, 379–383
63. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) Biochemistry 34, 12803–12811
64. Gacy, A. M., Geollner, G., Juranie, N., Macura, S., and McMurray, C. T. (1995) Cell 81, 533–540
65. Haber, J. E. (1999) Trends Biochem. Sci. 24, 271–275
66. Michel, B., Ehrlich, S. D., and Uzest, M. (1997) EMBO J. 16, 430–438
67. Uzest, M., Ehrlich, S. D., and Michel, B. (1995) Mol. Microbiol. 17, 1177–1187
68. Miret, J. J., Pesson-Brandao, L. P., and Lahue, R. S. (1997) Mol. Cell. Biol. 17, 3392–3397
69. Pearson, C. E., and Sinden, R. R. (1996) Biochemistry 35, 5041–5053
70. Cohen, H., Sears, D. D., Zenvirth, D., Hieter, P., and Simchen, G. (1999) Mol. Cell. Biol. 19, 4143–4158
71. Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J. A., and Jasins, M. (1998) Mol. Cell. Biol. 18, 93–101
72. Stewart, S. S., Masser, R. S., Stankovir, T., Bressan, D. A., Kaplan, M. I., Jaepsers, N. G. N., Rasens, A., Byrd, P. J., Petrini, J. H., and Taylor, A. M. R. (1999) Cell 99, 577–587
Gene Conversion (Recombination) Mediates Expansions of CTG·CAG Repeats
John P. Jakupciak and Robert D. Wells

J. Biol. Chem. 2000, 275:40003-40013.
doi: 10.1074/jbc.M007153200 originally published online September 25, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M007153200

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

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

This article cites 70 references, 24 of which can be accessed free at
http://www.jbc.org/content/275/51/40003.full.html#ref-list-1