Nonlinearity in genetic decoding: Homologous DNA replicase genes use alternatives of transcriptional slippage or translational frameshifting

Bente Larsen*†, Norma M. Wills†‡, Chad Nelson, John F. Atkins, and Raymond F. Gesteland

Department of Human Genetics, University of Utah, 15 N 2030 East Building 533, Room 7410, Salt Lake City, UT 84112-5330

Communicated by Michael J. Chamberlin, University of California, Berkeley, CA, December 3, 1999 (received for review October 1, 1999)

The τ and γ subunits of DNA polymerase III are both encoded by a single gene in *Escherichia coli* and *Thermus thermophilus*. γ is two-thirds the size of τ and shares virtually all its amino acid sequence with τ. *E. coli* and *T. thermophilus* have evolved very different mechanisms for setting the approximate 1:1 ratio between τ and γ. Both mechanisms put ribosomes into alternate reading frames so that stop codons in the new frame serve to make the smaller γ protein. In *E. coli*, ≈50% of initiating ribosomes translate the *dnaX* mRNA conventionally to give τ, but the other 50% shift into the −1 reading frame at a specific site (A AAA AAG) in the mRNA to produce γ. In *T. thermophilus* ribosomal frameshifting is not required: the *dnaX* mRNA is a heterogeneous population of molecules with different numbers of A residues arising from transcriptional slippage on a run of nine T residues in the DNA template. Translation of the subpopulation containing nine As (or +/− multiples of three As) yields τ. The rest of the population of mRNAs (containing nine +/− nonmultiples of three As) puts ribosomes into the alternate reading frames to produce the γ protein(s). It is surprising that two rather similar *dnaX* sequences in *E. coli* and *T. thermophilus* lead to very different mechanisms of expression.

Life is based on linear information that is copied, transcribed and translated. Typically we see that when this linearity is disturbed—insertion/deletion mutations in DNA or frameshift mutations in mRNAs—errors result that adversely affect gene function. However, nature orchestrates specific nonlinearities in the readout processes for the purpose of regulating expression of some genes. Here there is no longer a strict correspondence between linear "codons" in the DNA sequence and the amino acid sequence of proteins because alternative reading frames are exploited.

Intriguing examples of nonlinearity are found by comparison of the synthesis of two subunits of the major DNA replicase, DNA polymerase III, of *E. coli* and *T. thermophilus* (Fig. 1). In *E. coli* two subunits, τ and γ, present in a 1:1 ratio in the holoenzyme, are encoded by a single mRNA expressed from the *dnaX* gene (1–3). τ, the long form, is translated by standard decoding. γ, the short form, shares amino acid sequence for the first two-thirds of the τ sequence but then has one additional amino acid encoded by the −1 frame of the mRNA. Fifty percent of the ribosomes respond to signals in the mRNA and shift to the −1 frame at A AAA AAG (4–6) and after one codon, reach a UGA terminator. As a result, γ lacks a crucial domain of τ, so that the two proteins have different functions in DNA polymerase III. τ forms a dimer; the additional 213 aa in τ not present in γ confer the ability to bind core polymerases and the DnaB helicase (7–10). γ dimers are the "motor" for the clamp loader (11) and interact with the core polymerase through the τ dimer (12). γ may also interact with topoisomerase IV (13).

In *T. thermophilus*, there are proteins analogous to τ and γ. They are also both made by translation of mRNA from the same gene. It has been assumed that the γ protein arises by a ribosomal shift of reading frame at a suitably positioned A-rich sequence in the mRNA (AA AAA AAA A) (14, 15). However, this sequence is known to be less prone to ribosomal frameshifting than A AAA AAG in *E. coli*. In addition, stimulatory signals found in the mRNA of *E. coli* responsible for the 50% frameshift efficiency are not found in corresponding positions in the *T. thermophilus* mRNA (Fig. 1) (5). We show here that in *T. thermophilus*, ribosomal frameshifting is not necessary for production of the γ subunit. Instead the *dnaX* mRNA population is a mixture of sequences resulting from transcriptional slippage on the run of T residues and γ is produced by standard translation of a subpopulation of RNA molecules that accesses the alternate reading frames.

Materials and Methods

Isolation and Sequencing of RNA from *E. coli*. RNA was isolated by using an RNeasy kit from Qiagen. Forty micrograms of RNA was used for sequencing as previously described (16).

Preparation of *T. thermophilus* DNA and RNA. *T. thermophilus s*train HB27 Pro$^-$ (17, 18) was grown at 60°C in 0.5% tryptone, 0.4% yeast extract media (19) supplemented with Castenholz salts (20). Nucleic acids were extracted from whole cells by treatment with hot phenol (16) and used as the source of genomic DNA for PCR. A portion of the nucleic acid preparation was incubated with RNase-free DNase (Promega), extracted with phenol-chloroform, and precipitated with ethanol. This RNA was the template for reverse transcriptase before PCR.

Reverse Transcription of RNA. *dnaX*-specific cDNA was synthesized from total cellular RNA using primer TT72 (see below) according to standard protocols (21).

PCR Reactions. Sequences of the PCR primers were: TT68 TATCGTAATAAGCTTCGGAACAGAAGGTGAGGCTC (*Hin*dIII site underlined); TT69 TATAATTCTAGAGCTTCAGGAGGTGGGGCCGG (*Xba*I site underlined); TT71 TCGT*CCTCGTCCTGGAGG; *reported as C in ref. 14 but as T in ref. 15. All of the PCR clones sequenced had a T at this position (see Fig. 3); TT72 CGGGCGGGGCCTTGGGCT. PCR reactions were performed using Biolase Platinum polymerase (Bioline, London) as recommended by the manufacturer. The [Mg^{+2}] was 4 mM and spermidine was added to 0.5 mM. The annealing temperature was 58°C for primer sets TT68/TT69 and TT71/TT72, and extension was for 35 cycles. The genomic DNA template concentration was 1 ng/μl and the cDNA template concentration corresponded to 0.3 ng/μl of RNA. The RNA preparations gave no PCR signal at 0.3 ng/μl before reverse transcription indicating that there was no DNA

*Present address: Bakteriologiska enheten, Smittskyddsinstitutet, S-171 82 Solna, Sweden.

†These authors contributed equally to this work.

‡To whom reprints requests should be addressed. E-mail: nwills@genetics.utah.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

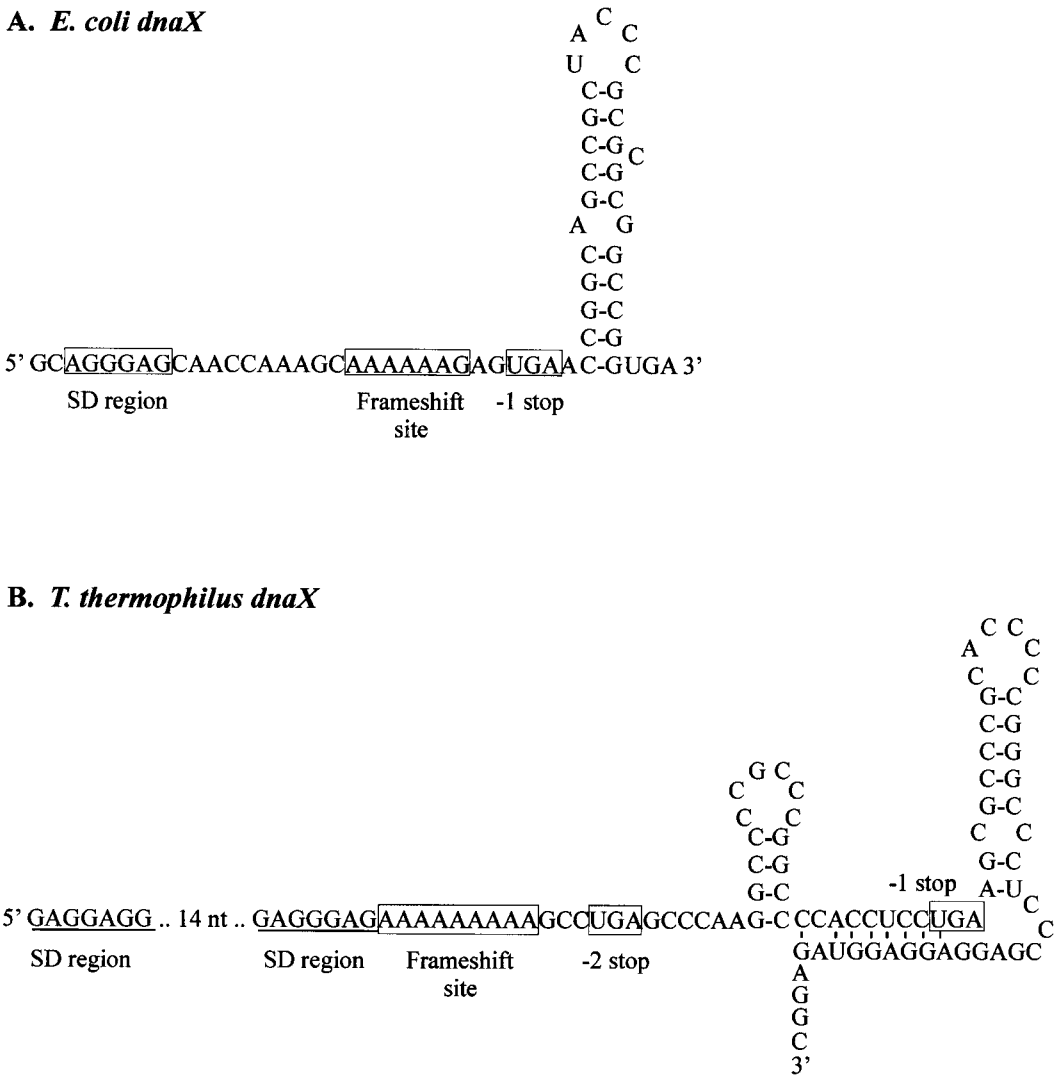


Fig. 1. Slippage sites in the *E. coli* (*A*) and *T. thermophilus dnaX* (*B*) genes. (*A*) The frameshift site is boxed along with the stimulatory 5′ Shine–Dalgarno-like sequence and the −1 frame termination codon. The stem-loop structure located 3′ to the shift site stimulates −1 frameshifting. (*B*) The slippage site and −2 and −1 termination codons are boxed. The two Shine–Dalgarno-like sequences are underlined. Potential secondary structures are shown 3′ of the slippage site.

contamination. For mass spectrometry analysis, the [Mg^{+2}] was lowered to 1 mM to minimize nontemplate A additions at 3′-ends. The annealing temperature was lowered to 34°C for the cDNA template derived from the synthetic RNA due to less overlap with primer TT71.

Cloning and Sequencing of PCR Products. PCR products were extracted with phenol-chloroform and precipitated with ethanol, digested with restriction endonucleases *Hin*dIII and *Xba*I, and ligated with *Hin*dIII–*Xba*I-digested pUC18 DNA. Thirty-six clones were selected randomly from each pool and the plasmid DNA was sequenced.

Electrospray Mass Spectrometry of PCR Products. The nucleotide composition of the poly(A) region was accurately determined by measuring the molecular weight of PCR products using electrospray mass spectrometry (22–24). The electrospray ionization process generates a series of multiply-charged molecular ions from which very accurate (better than 0.01%) mass assignments are derived for each DNA molecule in a PCR experiment. As a result, the exact nucleotide composition of both the sense and antisense strands are determined.

Following the PCR amplification, samples were prepared for electrospray mass spectrometry by a simple ethanol precipitation procedure. The ammonium-salt form of the DNA products were generated by ion-exchange in 2.3 M ammonium acetate, which was followed by overnight precipitation at −20°C through addition of three volumes of absolute ethanol. The precipitates were centrifuged for 30 min, the supernatant was poured off, and traces of liquid were carefully removed. As a wash step, 200 μl of 80% absolute ethanol was added, followed by gentle mixing, centrifugation again, and removal of the supernatant.

Before mass spectrometry analysis, PCR products were reconstituted in 15 μl of a solution consisting of 48% HPLC-grade acetonitrile, 48% HPLC-grade water, and 4% triethylamine (Fluka: research grade, >99.5%). Samples were introduced into an electrospray ionization source of a Quattro-II mass spectrometer (Micromass; Beverly, MA) by direct infusion at a rate of 4 μl/min, using a syringe pump for solvent delivery. Mass spectra were obtained from ≈20 accumulated spectra by using multi-

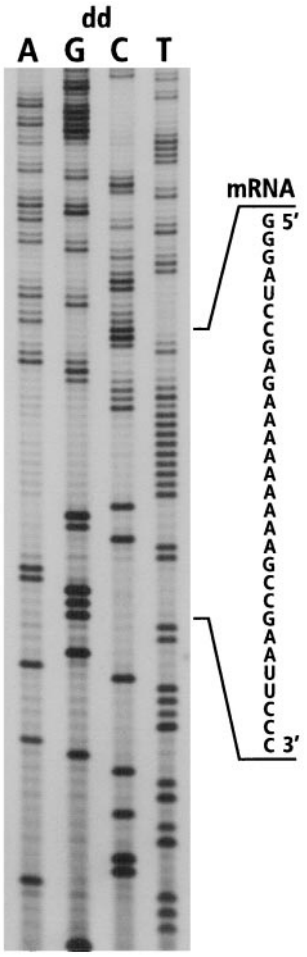


Fig. 2. RNA sequence of the *T. thermophilus dnaX* slippage region expressed in *E. coli*. Each lane contained a chain-terminating dideoxynucleotide as indicated. The predominant RNA sequence, corresponding to nine As in the DNA clone, is shown on the right.

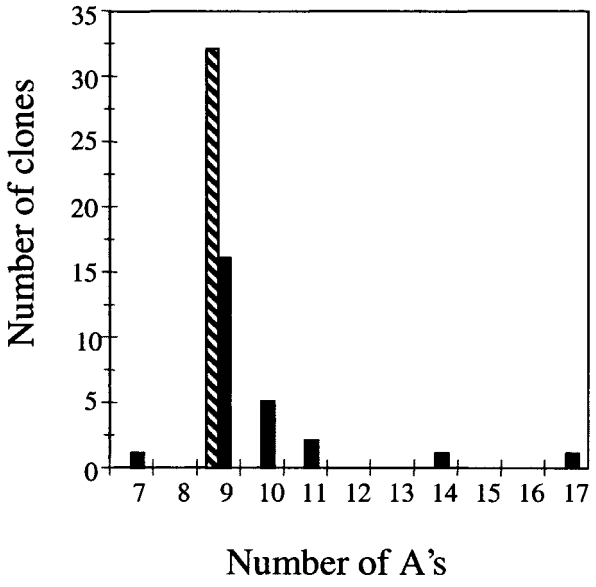


Fig. 3. Graph of the distribution of clones derived from PCR of *T. thermophilus dnaX* mRNA (cDNA) or genomic DNA. The mRNA (cDNA) clones are shown by black rectangles and the genomic DNA clones by the hatched rectangle. The number of A residues found in each clone is shown on the *x* axis, and the number of clones is indicated on the *y* axis. All of the clones sequenced contained a C instead of T at position 2,121 (GenBank accession no. AF025391) as reported in ref. 14, but no other sequence differences outside of the run of A/T were observed.

channel analysis data storage in the negative-ion mode while scanning 640–940 Da in 3 sec, with a cone voltage of 32 eV and a spray voltage of 2.7 kV. Molecular mass spectra showing measured DNA molecular weights were generated by deconvolution of multiply-charged molecular-ion series using MAXENT software (Micromass).

Results and Discussion

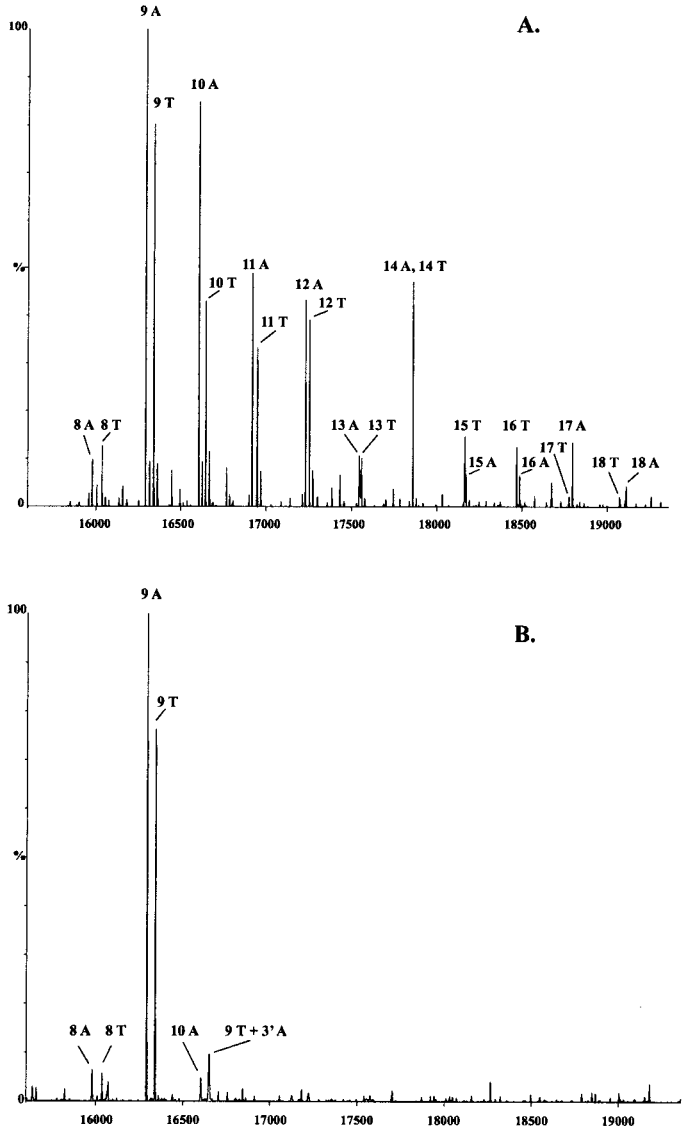
A region of the *T. thermophilus dnaX* gene, including the run of nine A residues, allows switching to both +1 and −1 reading frames when assayed in *E. coli* (14). When the run of As was disrupted by substituting two A residues with Gs, the shifts were eliminated (14). By deletion analysis, the critical region of the *T. thermophilus* mRNA was narrowed to the run of nine A residues, excluding the potential upstream and downstream sequence elements known to be necessary for efficient ribosomal frameshifting in the *E. coli* counterpart (data not shown). The shift to both frames and the sole importance of the A run raised the possibility that slippage of the RNA polymerase during transcription of the *T. thermophilus* gene gives a population of mRNAs with different numbers of A residues. A population of mRNAs with varying numbers of A residues could make all three reading frames accessible without ribosomal frameshifting. RNA polymerase slippage is known to occur on runs of A or T residues (16, 25, 26) by realignment of the nascent RNA chain and DNA template.

To determine if translation into alternate reading frames was due to translational frameshifting or transcriptional slippage, the *T. thermophilus dnaX* region of the mRNA population expressed from a plasmid in *E. coli* was analyzed by sequencing. As seen in Fig. 2, the RNA gives a single sequence up to the run of A residues, but the sequence then becomes complex indicating a mixture of RNAs with variable numbers of A residues. In contrast, the sequence of the plasmid DNA is uniform with the expected nine As (data not shown), implying that the heterogeneity is not in the DNA template but arose during transcription.

To investigate whether transcriptional slippage occurs in *T. thermophilus*, two methods were used to monitor the number of A residues in the *dnaX* genomic DNA and mRNA sequences. One method sampled the mRNA population by sequencing clones derived from the mRNA pool. PCR products spanning the A run from either cDNA or control genomic DNA were cloned and sequenced. All of the 32 clones from the control genomic DNA template had the expected nine A residues (Fig. 3). In contrast, 10 of the 25 clones from the cDNA (mRNA) template had other than nine As, ranging from 7 to 17 (Fig. 3).

In the second more direct method, molecular masses of the population of PCR products were measured by electrospray mass spectrometry. Products with masses corresponding to molecules with 8–18 As in the run are present with the cDNA (mRNA) template (a decreasing abundance distribution) consistent with the cloning results (Fig. 4 *A* and *C*). Each size class gives two mass spectral peaks corresponding to each of the DNA single strands.

The pattern is strikingly different for PCR products from genomic DNA in which the predominant signals are the two single strands from the expected product containing nine As (Fig. 4 *B* and *C*). However, there were minor products corresponding to 8 and 10 A/T. PCR amplification was carried out at 1 mM Mg^{2+} to minimize the well-known nontemplated end addition of A residues. (This also enabled the more accurate



C.

Sequence	$Mass_{measured}$	$Mass_{theoretical}$	% error
Sense strand of *T. thermophilus* RNA/cDNA PCR products			
8A	15978.4	15978.4	0.000
9 A	16291.6	16291.6	0.000
10 A	16604.2	16604.8	0.004
11 A	16917.6	16918.0	0.002
12 A	17230.3	17231.2	0.005
13 A	17545.7	17544.5	0.007
14 A	[c]17859.0	17857.7	0.007
15 A	18171.1	18170.9	0.001
16 A	18485.0	18484.1	0.005
17 A	18795.7	18797.3	0.009
18 A	19111.6	19110.5	0.006
Anti-sense strand of *T. thermophilus* RNA/cDNA PCR products			
8 T	16036.0	16035.4	0.004
9 T	16339.1	16339.6	0.003
10 T	16644.7	16643.8	0.005
11 T	16947.5	16948.0	0.003
12 T	17253.2	17252.2	0.006
13 T	17558.7	17556.4	0.013
14 T	[c]17859.0	17860.6	0.009
15 T	18163.9	18164.8	0.005
16 T	18468.1	18469.0	0.005
17 T	18775.0	18773.2	0.010
18 T	19076.4	19077.3	0.005
Sense strand of *T. thermophilus* DNA PCR products			
8 A	15979.8	15978.4	0.009
9 A	16292.0	16291.6	0.003
10 A	16604.6	16604.8	0.001
Anti-sense strand of *T. thermophilus* DNA PCR products			
8 T	16036.4	16035.4	0.006
9 T	16340.4	16339.6	0.005
10 T	[a] -	16643.8	-
9 T+3' A	[b]16651.5	16652.8	0.008

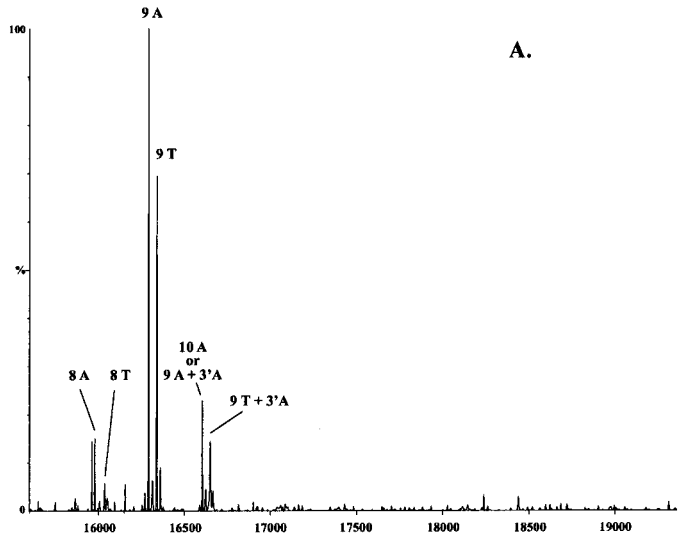
[a] The mass peak due to 10 T's is relatively small or not present.
[b] This peak corresponds to 9 T's with the addition of a non-template A at the 3' end of the molecule during PCR.
[c] The sense and anti-sense strands are unresolved mass spectrometrically, ie., the theoretical mass difference of 2.9 daltons could not be resolved at a molecular weight of 17.8 kDa. Therefore, the sense and anti-sense peaks appear as a single peak located between the two theoretical masses, of roughly twice the relative intensity.

Fig. 4. Mass spectra of PCR products generated from *T. thermophilus dnaX* mRNA (cDNA) or genomic DNA. The molecular mass is shown on the *x* axis, and the relative intensity of the signal is indicated on the *y* axis. (*A*) Distribution of PCR products from the cDNA template. Two single-strand products are evident for each PCR fragment. The number of A or T residues in the homopolymeric run is shown above each peak. The measured and predicted masses of each peak are shown in *C*. When the PCR products contain 14 A/T, the masses of the single strands are not resolved and appear as a single peak. (*B*) Distribution of PCR products from genomic DNA. The number of A or T residues is shown above each peak. The measured and predicted masses of each peak are shown in *C*. (*C*) Tabulation of the measured and predicted masses of the PCR product single strands. The difference between predicted and measured masses is expressed as the percentage error for each peak.

assignment of DNA molecules differing by only 9 Da, such as the 9 T +3′ A vs. the 10 T products.) The small amount of possible heterogeneity detected in the genomic DNA PCR products cannot account for the expression of equal amounts of τ and γ. The lack of a large distribution of products from the genomic DNA template also indicates that significant DNA polymerase slippage does not occur during the PCR amplification.

A remaining possibility is that variation in the number of A/T residues arose during reverse transcription of the mRNA. This possibility was tested directly by using a chemically synthesized RNA template with nine As that corresponded to the sequence of the *T. thermophilus dnaX* slippage region. The mass spectrum of the synthetic RNA (theoretical mass of 14,225.76) has a major peak at 14,225.8, with trace peaks corresponding to molecules lacking an A, C, G, or U residue (data not shown). When this synthetic RNA is reverse transcribed and PCR amplified, extensive heterogeneity in the run of A/Ts is not seen in the mass spectrum of the products (Fig. 5). The minor products with an additional A are due primarily to 3′-end addition because they are seen for both strands containing either the A or the T runs. Other minor products (such as nine As lacking single nucleotides) are consistent with the mass spectrum of the synthetic RNA. We conclude that the heterogeneity observed in the cDNA-derived PCR products reflects a mixed mRNA population *in vivo* and is not a consequence of slippage during reverse transcription.

Therefore, *T. thermophilus* γ can arise by standard translation of a subset of *dnaX* mRNA molecules, in contrast to *E. coli* γ, which is synthesized from a homogeneous population of mRNA by ribosomal frameshifting. The two organisms use different



B.

Sequence	$Mass_{measured}$	$Mass_{theoretical}$	% error
Sense strand of synthetic RNA/cDNA PCR products			
8 A	15979.0	15978.4	0.004
9 A	16292.0	16291.6	0.003
10 A	16605.0	16604.8	0.001
Anti-sense strand of synthetic RNA/cDNA PCR products			
8 T	16036.2	16035.4	0.005
9 T	16340.0	16339.6	0.002
10 T	[a] -	16643.8	-
9 T +3' A	[b]16651.2	16652.8	0.010

[a] The mass peak due to 10 T's is relatively small or not present.

[b] This peak corresponds to 9 T's with the addition of a non-template A at the 3' end of the molecule during PCR.

Fig. 5. (*A*) Mass spectrum of PCR products generated from the cDNA template derived from synthetic RNA. The molecular mass is shown on the *x* axis, and the relative intensity of the signal is indicated on the *y* axis. Two single-strand products are present for each PCR fragment. The number of A or T residues in the homopolymeric run is shown above each peak. (*B*) Tabulation of the measured and predicted masses of the PCR product single strands. The difference between predicted and measured masses is expressed as the percentage error for each peak.

means to the same end. In *E. coli*, 50% of the ribosomes translate by standard rules to make τ and 50% frameshift to make the shorter γ and the frameshift ratio is set by cis-acting elements in the mRNA. In *T. thermophilus*, the ratio of γ to τ is set by the distribution of different mRNAs. Those mRNAs with 9, 12, or 15 As give the full length τ. Because there are stop codons in both the +1 and −1 frames shortly after the run of As, mRNAs with 8, 11, and 14, in one frame and 10, 13, and 16 in the other will all yield γ protein (differing by the number of Lys residues and the carboxyl-terminal amino acids encoded by the −1 or +1 reading frames). This would predict that the ratio of the sum of 9, 12, and 15 A-containing mRNAs to the sum of 8, 10, 11, 13, 14, 16, and 17 A-containing mRNAs should be 1:1. To a first approximation this is true (Fig. 4*A*) but the measurements are not strictly quantitative. It will be of interest to determine if multiple forms of γ are present in the DNA polymerase III holoenzyme of *T. thermophilus*. Conventional translation of the distribution of different mRNA molecules can account for the observed ratio of τ and γ. It is not necessary to invoke frameshifting. (We have no reason to expect that ribosomal frameshifting also contributes to production of γ, but our experiments cannot rule this out.)

There is another potential example of the alternatives of translational frameshifting and transcriptional slippage being used in similar genes by different organisms. Many members of the IS*3*-family of insertion elements exploit −1 translational frameshifting to produce two products (27). IS*120* from *Clostridium thermocellum* (28) contains a run of nine As at the expected frameshift site and is a candidate for transcriptional slippage.

During transcription, an ≈8 bp hybrid between the nascent RNA chain and its DNA template helps to keep the RNA transcript in register (29). With a run of nine Ts in the template, the growing chain with eight or nine As can dissociate from the template and re-pair in a new register, still maintaining the required 8-bp hybrid. The minimum length of the T run to promote slippage is nine. As predicted from the length of the hybrid, eight Ts does not work in *E. coli* (ref.16 and data not shown). The crystal structure of *Thermus aquaticus* RNA polymerase suggests that the upstream rudder of the protein core complex is near the edge of the transcription bubble (30) indicating a possible role in slippage.

T. thermophilus RNA polymerase appears to function similarly to *E. coli* RNA polymerase in adding or deleting A/T residues by simple realignment of the template with the nascent transcript. This also appears to be the mechanism for transcriptional slippage in an apolipoprotein B mutant that contains a run of eight As (31, 32) and in the expression of multiple lipooligosaccharides in *Neisseria gonorrhoeae* due to a polyguanine tract (33). In cases of transcriptional slippage in paramyxoviruses, the number of added residues is controlled by a signal positioned immediately 5′ to the insertion site (34, 35). The base pairing possibilities between the signal sequence and the nascent transcript determine the preferred realignment position and hence the number of additional residues inserted. In Ebola virus, transcriptional slippage appears to require complex signals in the template RNA. The insertion site resembles a polyadenylation site and the 3′ sequence can form a stem-loop structure that may serve to pause the polymerase (36). It is also possible that transcriptional slippage is involved in the predisposition for colorectal cancer due to a particular allele of the *APC* gene containing eight contiguous A residues (37), although differences in the genomic sequences within and near the run of As have been detected in tumors (38) implicating DNA polymerase errors. Whether there are specific signals within the *T. thermophilus dnaX* gene that promote transcriptional slippage will require further study.

Several cases of transcription slippage are known to be used for gene expression purposes. Because the protein motif Lys-Lys-Lys (which can be encoded by a run of nine A residues) is underrepresented in the human protein database (32) but is nonetheless quite widespread, the role of transcriptional slippage is likely much greater than currently appreciated.

It is striking that *E. coli* and *T. thermophilus dnaX* expression occurs by very different mechanisms to produce the equivalent functional end result: a set amount of two proteins from the same gene. With the genome project well underway, thoughts are turning to the potential for diversity in the expression of protein products. As the present work hints, many surprises are likely in store.

We thank Dr. Shuang-yong Xu and New England Biolabs for providing *Thermus thermophilus* strain, HB27 Pro$^-$. We thank Dr. Mike O'Donnell for making available unpublished information and also he and Dr. Charles McHenry for encouraging us to proceed with this project. We appreciate Dr. Olivier Fayet providing us with the IS*120* sequence.

1. Flower, A. M. & McHenry, C. S. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 3713–3717.
2. Tsuchihashi, Z. & Kornberg, A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2516–2520.
3. Blinkova, A. L. & Walker, J. R. (1990) *Nucleic Acids Res.* **18,** 1725–1729.
4. Tsuchihashi, Z. & Brown, P. O. (1992) *Genes Dev.* **6,** 511–519.
5. Larsen, B., Wills, N. M., Gesteland, R. F. & Atkins, J. F. (1994) *J. Bacteriol.* **176,** 6842–6851.
6. Larsen, B., Gesteland, R. F. & Atkins, J. F. (1997) *J. Mol. Biol.* **271,** 1–14.
7. McHenry, C. S. (1982) *J. Biol. Chem.* **257,** 2657–2663.
8. Studwell-Vaughan, P. S. & O'Donnell, M. (1991) *J. Biol. Chem.* **266,** 19833–19841.
9. Kim, S., Dallmann, H. G., McHenry, C. S. & Marians, K. J. (1996) *Cell* **84,** 643–650.
10. Yuzhakov, A., Turner, J. & O'Donnell, M. (1996) *Cell* **86,** 877–886.
11. Turner, J., Hingorani, M., Kelman, Z. & O'Donnell, M. (1999) *EMBO J.* **18,** 771–783.
12. Kelman, Z. & O'Donnell, M. (1995) *Annu. Rev. Biochem.* **64,** 171–200.
13. Levine, C. & Marians, K. J. (1998) *J. Bacteriol.* **180,** 1232–1240.
14. Yurieva, O., Skangalis, M., Kuriyan, J. & O'Donnell, M. (1997) *J. Biol. Chem.* **272,** 27131–27139.
15. McHenry, C. S., Seville, M. & Cull, M. G. (1997) *J. Mol. Biol.* **272,** 178–189.
16. Wagner, L. A., Weiss, R. B., Driscoll, R., Dunn, D. S. & Gesteland, R. F. (1990) *Nucleic Acids Res.* **18,** 3529–3535.
17. Koyama, Y., Hoshino, T., Tomizuka, N. & Furukawa, K. (1986) *J. Bacteriol.* **166,** 338–340.
18. Wayne, J. & Xu, S. (1997) *Gene* **195,** 321–328.
19. Oshima, T. & Imahori, K. (1974) *Int. J. Syst. Bacteriol.* **24,** 102–112.
20. Ramaley, R. F. & Hixson, J. (1970) *J. Bacteriol.* **103,** 527–528.
21. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
22. Wunschel, D. S., Fox, K. F., Fox, A., Bruce, J. E., Muddiman, D. C. & Smith, R. D. (1996) *Rapid Commun. Mass Spectrom.* **10,** 29–35.
23. Laken, S. J., Jackson, P. E., Kinzler, K. W., Vogelstein, B., Strickland, P. T., Groopman, J. D. & Friesen, M. D. (1998) *Nat. Biotechnol.* **16,** 1352–1356.
24. Nelson, C., Nelson, L. & Ward, K. (1998) *Am. J. Obstet. Gynecol.* **178,** S26 (abstr.).
25. Chamberlin, M. & Berg, P. (1962) *Proc. Natl. Acad. Sci. USA* **48,** 81–93.
26. Uptain, S., Kane, C. & Chamberlin, M. (1997) *Annu. Rev. Biochem.* **66,** 117–172.
27. Chandler, M. & Fayet, O. (1993) *Mol. Microbiol.* **7,** 497–503.
28. Snedecor, B., Chen, E. & Gomez, R. (1983) in *Proceedings of the Fourth International Symposium on Genetics of Industrial Microorganisms* (Societe Francaise des Mecaniciens Publications, Paris), pp. 356–360.
29. Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. (1997) *Cell* **89,** 33–41.
30. Zhang, G., Campbell, E., Minakhin, L., Richter, C., Severinov, K. & Darst, S. (1999) *Cell* **98,** 811–824.
31. Linton, M., Pierotti, V. & Young, S. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 11431–11435.
32. Linton, M. F., Raabe, M., Pierotti, V. & Young, S. G. (1997) *J. Biol. Chem.* **272,** 14127–14132.
33. Burch, C. L., Danaher, R. J. & Stein, D. C. (1997) *J. Bacteriol.* **179,** 982–986.
34. Kolakofsky, D. & Hausmann, S. (1998) in *Modification and Editing of RNA*, eds. Grosjean, H. & Benne, R. (Am. Soc. Microbiol., Washington, DC), pp. 413–420.
35. Hausmann, S., Garcin, D., Morel, A. S. & Kolakofsky, D. (1999) *J. Virol.* **73,** 343–351.
36. Volchkov, V. E., Becker, S., Volchkova, V. A., Ternovoj, V. A., Kotov, A. N., Netesov, S. V. & Klenk, H. (1995) *Virology* **214,** 421–430.
37. Raabe, M., Linton, M. & Young, S. (1998) *Am. J. Med. Genet.* **76,** 101–102.
38. Laken, S. J., Petersen, G. M., Gruber, S. B., Oddoux, C., Ostrer, H., Giardiello, F. M., Hamilton, S. R., Hampel, H., Markowitz, A., Klimstra, D., *et al*. (1997) *Nat. Genet.* **17,** 79–83.