Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III γ subunit from within the τ subunit reading frame

Aleksandra L.Blinkowa and James R.Walker
Department of Microbiology, The University of Texas, Austin, TX 78712, USA

Received January 5, 1990; Revised and Accepted February 22, 1990

ABSTRACT

The *Escherichia coli dnaX* gene encodes both the τ and γ subunits of DNA polymerase III holoenzyme in one reading frame. The 71.1 kDa τ and the shorter γ share N-terminal sequences. Mutagenesis of a potential ribosomal frameshift signal located at codons 428 – 430 without changing the amino acid sequence of the τ product, eliminated detectable synthesis of the γ subunit, suggesting that the reading frame is shifted at that sequence and γ is terminated by a nonsense codon located in the −1 frame 3 nucleotides downstream of the signal. This seems to be the first known case of a frameshift which is used, along with the termination codon in the −1 frame, to terminate a peptide within a reading frame. [Mutagenesis of a dibasic peptide (lys-lys) at codons 498 – 499, the site at which a τ'-'LacZ fusion protein was cleaved *in vitro* (1) had no effect on γ formation *in vivo*, suggesting that cleavage observed *in vitro* is not the mechanism of γ formation *in vivo*.]

INTRODUCTION

The *E. coli* DNA polymerase III holoenzyme consists of a three-subunit core plus four to seven auxiliary subunits (2,3). Two of the auxiliary factors, τ and γ are encoded by one reading frame; τ is thought to be the 71.1 kDA full length product of the *dnaX* reading frame but γ is a factor which migrates at about 56 kDA (4,5,6,7) in electrophoresis. The N-terminal sequences of both subunits are identical to the N-terminal amino acid sequence deduced from the reading frame (1,8). Antibody to a synthetic peptide corresponding to amino acids 420−440 reacted with both τ and γ, confirming that the C-terminal region of γ shares sequences over at least part of the 420−440 residue region with τ. A τ'-'LacZ fusion protein was cleaved *in vitro* between adjacent lysine residues at positions 498 and 499 (1) but the mechanism by which γ synthesis is terminated *in vivo* has been unclear.

The τ reading frame contains a run of six A residues followed by a G, extending over codons 428−430, a tandem slippery sequence (9,10) which is known to cause about 50% −1 ribosomal frameshifting in *E. coli* (11). Moreover, a potential stem and loop structure over codons 432−442 could enhance shifting (10,12,13), perhaps as a pseudoknot (13,14,15). These considerations suggest that γ could be terminated *in vivo* by a frameshift followed by a nonsense codon. This model was tested by altering the slippery codons without changing the amino acid sequence of τ with the result that γ synthesis was no longer detectable. This strongly supports a model previously proposed by McHenry (16) that γ is terminated within the τ reading frame by a programmed ribosomal frameshift followed by a nonsense codon.

METHODS

Strain, phage and plasmids

E. coli strain JM103Y (17) was used for all experiments except growth of phage AB5 for which the *dut ung* strain (J236 (18)) was used. The M13mp19 derivative AB5 carried a 3.1 kb *Eco*RI-*Pst*I fragment containing the *apt* and *dnaX* genes (6,7,19) and was used for site-directed mutagenesis. Single-strand DNA prepared from AB5 contains the *dnaX* strand complementary to messenger. Wild-type and mutant *dnaX* fragments (Fig. 1) were cloned as 2.77 kb *Hin*dIII-*Pst*I *dnaX*- containing fragments downstream of the pUC19 *lac* promoter for expression.

Recombinant DNA and DNA sequencing procedures

Standard restriction, ligation, tranformation, transfection, DNA isolation, and electrophoresis procedures were used (20). Chain termination (21) sequencing was performed using wild-type and mutant AB5 single-strand DNA as template and a primer complementary to nucleotides 2104 to 2120, numbering from the first nucleotide of the *Eco*RI site of the 3.1 kb *Eco*RI-*Pst*I *apt-dnaX* fragment mentioned above.

Site-directed mutagenesis

The Kunkel (22) procedure was used with single-strand DNA from phage AB5 grown in the *ung dut* mutant. Mutagenic oligonucleotides consisted of the sequences 5'GCGCTGAAAACAGCGCTG (positions 2355−2372; pAB6) and 5' GCGCTGACAACAGCGCTG (2355−2372; pAB7) from Operon Technologies, San Pablo, CA and 5' AACCAAAGCGAAGAAGAGTGA (positions 2144−2164; pAB9) and 5' GAAAGACTAGCTTCGGTC (positions 2211−2228; pAB10) from Midland Certified Reagents, Midland, TX.

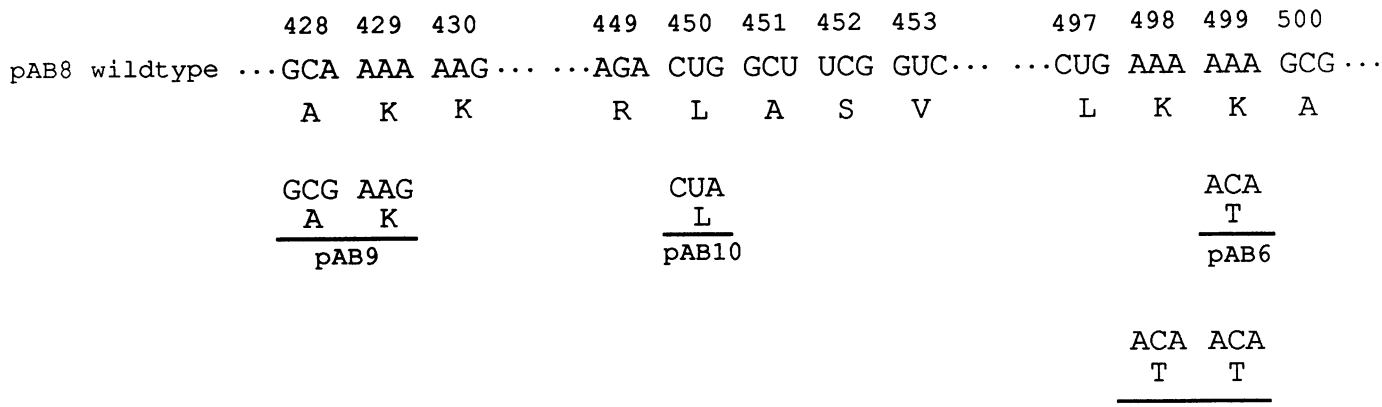


Fig. 1. Relevant codons and translation products of the wild-type pAB8 and mutant plasmids. The codons (triplets) are numbered; single letter amino acid symbols are used. Only changed codons (underlined) are represented for the mutant plasmids. The plasmids all contained a 2.77 kb *Hin*dIII-*Pst*I *dnaX* containing subfragment from the wild-type or mutated phage AB5.

Radioimmune precipitation of τ and γ

Plasmid-containing strains were induced to synthesize *dnaX* products by isopropyl-thio-β-D-galactoside (IPTG); τ and γ were labelled with [^{35}S] methionine and precipitated with antibody to a synthetic peptide corresponding the τ amino acids 420–440, electrophoresed, and autoradiographed as described by Lee et al. (1). These conditions detect τ and γ from the multi-copy plasmids but not from the chromosomal *dnaX* gene.

RESULTS

Three regions of *dnaX* were altered by site-directed mutagenesis

Two ribosomal frameshift and a proteolysis model for generating γ were tested. The two potential frameshift signals were a known 'slippery' sequence (11) and an unusually stable stem-loop structure downstream of a rare codon (16). The proteolysis model was based on the observation (1) that a τ'-fusion protein was cleaved *in vitro* between two lysine residues.

The wild-type and mutant plasmids were transformed into strain JM103Y which was induced with IPTG, labelled with [^{35}S] methionine, and τ and γ quantitated by radioimmune precipitation and electrophoresis (1). The first candidate was the 'slippery' A-AAA-AAG sequence extending over codons 428–430 and preceeding a potential stem-loop. If this sequence signals a -1 shift, as the same sequence from a retrovirus is known to do when cloned into *E. coli* (11), γ would terminate after 431 amino acids at a UGA (Fig. 2). Altering the A-AAA-AAG to G-AAG-AAG in pAB9 did not change the amino acids (alanine and lysine) encoded by codons 428 and 429 respectively, but did eliminate detectable γ synthesis (Fig.3). Thus, we conclude that γ is likely to be terminated by a frameshift over codons 428–430, a sequence already known to cause -1 shifts, (10,11) followed by the UGA in the -1 frame.

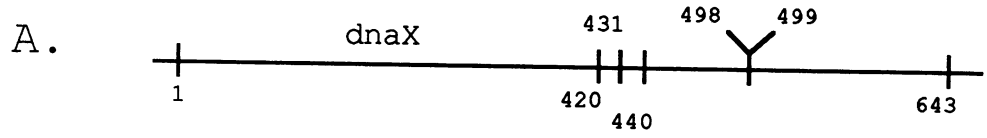
A second candidate for a frameshift signal was the unusually stable stem-loop containing CUUCGG (23). This stem-loop, mentioned as a possible frameshift signal by McHenry (16), is adjacent to a rare arginine codon, number 449, which might slow translation. However, reducing the (predicted) stability of the stem by changing the leucine codon 450 from CUG to CUA, also encoding leucine, in pAB10 (Fig.2) had no effect on γ synthesis (Fig. 3). Thus, it is unlikely that the potential stem-loop over codons 450–454 contributes to γ termination.

Third, the τ sequence of two adjacent lysines over positions 498 and 499 is known to be cleaved *in vitro* and it was proposed (1) that the *in vitro* cleavage might reflect *in vivo* processing. Accordingly, these lysines were altered—lysine 499 was mutated to threonine in pAB6 and both lysines at 498 and 499 were changed to threonines in pAB7 (Fig. 2). Both of these mutant *dnaX* genes continued to direct γ synthesis (Fig. 3) and it is unlikely that proteolysis between these lysines contributes to γ biosynthesis *in vivo*.

DISCUSSION

The *dnaX* reading frame of 643 amino acids was known to encode τ and γ proteins of about 71 and 56 kDA (1,4,5,6,7). These proteins share N-terminal sequences (1,8) which match the sequence deduced from the gene sequence (6,7). In addition, γ reacts with antibody against a synthetic peptide corresponding to amino acids 420–440, which suggests that γ contains at least a portion of this sequence (1). However, the mechanism by which the C-terminal end of γ is generated has been unclear. We report here that the *dnaX* gene contains, near the predicted C-terminal codon of γ based on molecular weight, a sequence which generates -1 ribosomal frameshifts in animal viruses (10) and in *E. coli* (11). This consists of a slippery pair of tandem codons followed by a potential stem-loop-a structure which enhances frameshifting in certain virus systems (10,12,13). Moreover, an RNA pseudoknot (14,15), required for frameshifting of the avian coronavirus, infectious bronchitis virus, in rabbit reticulocyte lysates and in *Xenopus* oocytes (13) could be formed by this stem-loop and nucleotides nearby (Fig. 2). Finally, alteration of the 'slippery' tandem codons by mutagenesis abolished the detectable synthesis of γ *in vivo*, suggesting that in wild-type *E. coli*, translation of τ occurs with a high level programmed ribosomal frameshift over codons 428–430. The predicted -1 frameshift is followed by a termination codon which overlaps 0 frame codons 431–432, resulting in termination of γ as a 431 amino acid, 47.5 kDa protein identical in its N-terminal 430 amino acids to those of τ.

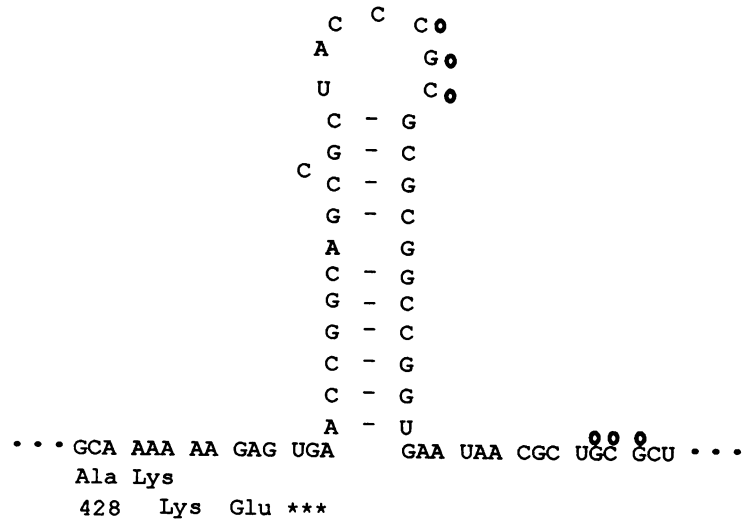
This model is consistent with earlier findings (1) that the steady state τ:γ ratio (in overproducing cells) is about 1 because the



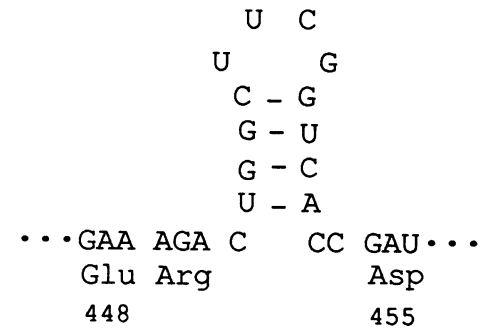
B.

```
···GCA AAA AAG AGU GAA CCG GCA GCC GCT ACC···
   Ala Lys Lys Ser Glu Pro Ala Ala Ala Thr
   428                                  437
```

C.



D.



E.

```
···GCG CUG AAA AAA GCG CUG···
   Ala Leu Lys Lys Ala Leu
         498 499
```

Fig. 2. Potential termination mechanisms for γ. A. The 643 amino acid *dnaX* reading frame. 420–440 mark the amino acid sequence used to generate antibody to τ and γ. 431 marks the probable termination point for γ translation. 498–499 mark the sequence of τ cleavage in vitro (1). B. Wild-type τ sequence (0 reading frame) over codons 428–437. C. Potential signals for −1 shift over codons 428–430 including a stem-loop with one of several pseudoknot possibilities indicated by the small circles. D. Stable stem-loop structure downstream of rare AGA codon. E. Sequence cleaved *in vitro* (1).

slippery sequence A-AAA-AAG cloned into *E. coli* has been shown to generate about 50% frameshifting (11). This model also is consistent with the report (1) that in pulse-labelled *E. coli* growing slowly at 20°C, label appeared in the γ product before it appeared in τ.

We previously reported (1) that τ, as a τ'-'LacZ fusion protein

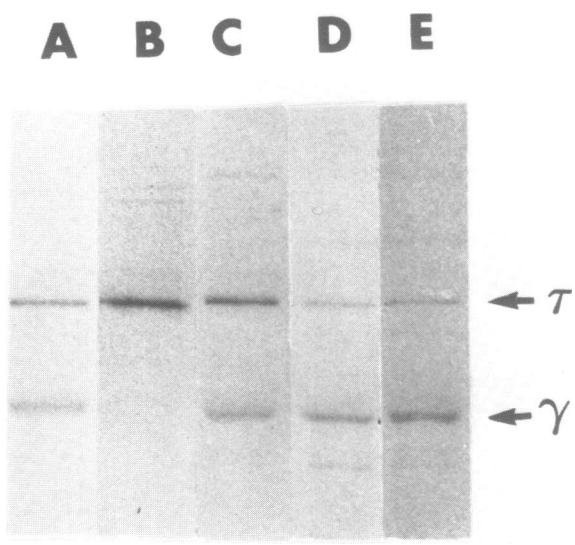


Fig. 3. Formation of τ and γ from wild-type and mutant *dnaX* genes. Labelled τ and γ were precipitated by immune serum, electrophoresed, and autoradiographed. A. wild-type *dnaX* in pAB8. B. pAB9, mutant codons 428–430 (A-AAA-AAG to G-AAG-AAG). C. pAB10 mutant codon 450 (CUG to CUA). D. pAB6 mutant codon 499 (AAA to ACA). E. pAB7 mutant codons 498 and 499 (AAA-AAA to ACA-ACA).

We previously reported (1) that τ, as a τ'-'LacZ fusion protein was cleaved *in vitro* between the two lysines at positions 498–499. The findings that alteration of the lysines to threonines did not influence γ formation and that altering a potential ribosomal frameshift signal did inhibit γ synthesis suggest that the cleavage observed *in vitro* is not the mechanism for γ formation *in vivo*. However, it is not excluded that normal τ turnover, if it occurs, is initiated by this cleavage.

The specific functions of τ and γ are unclear. DNA polymerase III holoenzyme preparations contain as many as ten subunits, including both τ and γ (24). Polymerases might function as dimers (25) which are asymmetric (26,27) and the τ and γ subunits could reside in different holoenzyme halves (16). However, processive holoenzyme complexes lacking τ have been prepared (28) although addition of τ stimulated synthesis by 4–8 fold (24). Processive complexes consisting only of the five subunits α, ϵ γ, δ, and β have been prepared also (29). Moreover, DNA polymerase core (α, ϵ, and β) plus τ can rapidly recycle to primed templates complexed with γ, δ, and β and catalyze highly processive polymerization (30,31).

Examples of programmed ribosomal frameshifting in the enterobacteria include a +1 frameshift near the C terminal end of the phage MS2 coat reading frame to couple translation of the coat protein to that of the downstream lysis protein (32,33). This phage encodes only four proteins with the 3′ end of the coat protein gene overlapping the 5′ end of the lysis gene. The lysis protein reading frame lacks an independent entry site for ribosomes and its translation depends on coat protein translation. A low level +1 shift in the coat protein reading frame directs ribosomes to a termination codon a few nucleotides upstream of the lysis protein initiation codon. Termination then allows re-initiation of lysis protein translation. This translational coupling is dependent on the +1 frameshift and is thought to be a mechanism for limiting the amount of lysis protein (compared to coat protein) and for ensuring that lysis protein is not made unless ample amounts of coat protein are available (33). Phage T7 produces two forms of gene 10 capsid protein, one the result of frameshifting. Phage particles contain both capsid protein and minor amounts of a longer protein produced by a −1 frameshift upstream of the normal termination codon to add 53 amino acids to the major capsid protein (34). Translation of *E. coli* peptide chain release factor two (RF2) requires a frameshift. A UGA present 26 codons from the initiation codon is by-passed by a +1 shift at the codon preceeding the stop (35). The amount of RF2 synthesis is autogenously controlled by RF2 (which releases nascent peptides stalled at UGA and UAA codons) terminating its own translation at the UGA (36). This high level frameshift occurs on about 50% of the ribosomes (36). Transposase of insertion sequence 1 is a trans-frame protein translated by a −1 shift ahead of the nonsense codon of an upstream reading frame which appears to encode a transposase inhibitor. Ribosomes which terminate at the upstream frame nonsense codon produce the inhibitor; the minor fraction which shift translate the longer fusion product which acts as the transposase (37).

In animal (e.g., 10,12,13,38) and yeast (39) viruses and retrotransposons (e.g., 40,41,42) programmed frameshifts bypass stop codons and direct ribosomes to overlapping reading frames to generate trans-frame polyproteins. The retroviral, and possibly the retrotransposon, polyproteins, such as *gagpol*, serve as precursors from which individual proteins are subsequently cleaved.

Thus, programmed ribosomal frameshifting to bypass a stop codon is a common theme in eucaryotic viruses, retroviral-like transposons, and procaryotes. The RNA phage MS2 is a special case in which the shift directs ribosomes to initiate at an overlapping reading frame. The *dnaX* gene of *E. coli* employs a novel mechanism by shifting about half the ribosomes to encounter a nonsense codon in the −1 frame thereby generating the shorter γ protein from the N-terminal segment of the τ reading frame.

A note added in proof to reference 43 recently reported the conclusion that γ is terminated as a result of a −1 frameshift over codons 429–430.

ACKNOWLEDGEMENTS

We thank R.B. Weiss for discussions about slippery codons, C.W.A. Pleij for suggesting the potential pseudoknot structure, and G. Valler for advice and strains JM103Y and CJ236. This work was supported by American Cancer Society grant MV-429M and, in part, by Public Health Service grant GM34471 from the National Institutes of Health and Welch Foundation grant F949.

REFERENCES

1. Lee, S.-H., Kanda, P., Kennedy, R.C. and Walker, J.R. (1987) *Nucleic Acids Res. 15*, 7663–7675.
2. McHenry, C.S. and Kornberg, A. (1977) *J. Biol. Chem. 252*, 6478–6484.
3. Maki, S. and Kornberg, A. (1988) *J. Biol. Chem. 263*, 6561–6569.
4. Mullin, D.A., Woldringh, C.L., Henson, J.M. and Walker, J.R. (1983) *Mol. Gen. Genet. 192*, 73–79.
5. Kodaira, M., Biswas, S.B. and Kornberg, A. (1983) *Mol. Gen. Genet. 192*, 80–86.
6. Yin, K.-C., Blinkowa, A. and Walker, J.R. (1986) *Nucleic Acids Res. 14*, 6541–6549.
7. Flower, A.M. and McHenry, C.S. (1986) *Nucleic Acids Res. 14*, 8091–8101.
8. Maki, S. and Kornberg, A. (1988) *J. Biol. Chem. 263*, 6547–6554.
9. Weiss, R.B., Dunn, D.M., Atkins, J.F. and Gesteland, R.F. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 687–693.

10. Jacks, T., Madhani, H.D., Masiarz, F.R. and Varmus, H.E. (1988) *Cell 55*, 447–458.
11. Weiss, R.B., Dunn, D.M., Shuh, M., Atkins, J.F. and Gesteland, R.F. (1989) *The New Biologist 1*, 159–170.
12. Rice, N.R., Stephens, R.M., Burny, A. and Gilden, R.V. (1985) *Virology 142*, 357–377.
13. Brierley, I., Digard, P. and Inglis, S.E. (1989) *Cell 57*, 537–547.
14. Studnicka, G.M., Rahn, G.M., Cummings, I.W. and Salser, W.A. (1978) *Nucleic Acids Res.* 5, 3365–3387.
15. Pleij, C.W.A., Rietveld, K. and Bosch, L. (1985) *Nucleic Acids Res. 13*, 1717–1731.
16. McHenry, C.S. (1988) *Anm. Rev. Biochem. 57*, 519–550
17. Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. Kunkel, T.A., Roberts, J.D. and Zakour, R.A. (1987) *Methods* Enzymol. *154*, 367–382.
19. Hershey, H.V. and Taylor, M.W. (1986) *Gene 43*, 287–293.
20. Maniatis, T., Fritsch, E.F. aand Sambrook, J. (1982) *Molecular Cloning, a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
22. Kunkel, T.A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488–492
23. Tuerk, C., Gauss, P., Thermes, C., Groebe, D.R., Gaybe, M., Guild, N., Stormo, G., d'Aubenton-Carafa, Y., Uhlenbeck, O.C., Tinoco, I., Brody, E.N. and Gold, L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1364–1368.
24. Maki, S. and Kornberg, A. (1988) *J. Biol. Chem. 263*, 6561–6569.
25. Sinha, N. K., Morris, C. F. and Alberts, B. M. (1980) *J. Biol. Chem. 255*, 4290–4303.
26. Johanson, K. O. and McHenry, C. S. (1984) *J. Biol. Chem. 259*, 4589–4595.
27. Maki, H., Maki, S. and Kornberg, A. (1988) *J. Biol. Chem. 263*, 6570–6578.
28. Maki, S. and Kornberg, A. (1988) *J. Biol. Chem. 263*, 6555–6560.
29. O'Donnell, M. and Studwell, P. S. (1990) *J. Biol. Chem. 265*, 1179–1187.
30. Wickner, S. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3511–3515.
31. O'Donnell, M. (1987) *J. Biol. Chem. 262*, 16558–16565.
32. Atkins, J.F., Gesteland, R.F., Reid, B.R. and Anderson, C.W. (1979) *Cell 18*, 1119–1131.
33. Kastelein, R.A., Remaut, E., Fiers, W. and van Duin, J. (1982) *Nature 295*, 35–40.
34. Dunn, J.J. and Studier, F.W. (1983) *J. Mol. Biol. 166*, 477–535.
35. Craigen, W.J., Cook, R.G., Tata, W.P. and Casken, C.T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3616–3620.
36. Craigen, W.J. And Caskey, C.T. (1986) *Nature 322*, 273–275.
37. Sekine, Y. and Ohtsubo, E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4609–4613.
38. Moore, R., Dixon, M., Smith, R., Peters, G. and Dickson, C. (1987) *J. Virol. 61*, 480–490.
39. Icho, T. and Wickner, R.B. (1989) *J. Biol. Chem. 264*, 6716–6723.
40. Saigo, K., Kugimiya, W., Matsuo, Y., Inouye, S., Yoshioka, K. and Yuki, S. (1984) *Nature 312*, 659–661.
41. Clare, J. J. and Farabaugh, P. J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2829–2833.
42. Mellor, J., Fulton, S. M., Dobson, M. J., Wilson, W., Kingsman, S. M. and Kingsman, A. J. (1985). *Nature 313*, 243–246.
43. Tsuchihashi, Z. and Kornberg, A. (1989) *J. Biol. Chem.* 264, 17790–17795.