Translational frameshifting generates the γ subunit of DNA polymerase III holoenzyme

ZENTA TSUCHIHASHI* AND ARTHUR KORNBERG

Department of Biochemistry, Stanford University, Stanford, CA 94305-5307

Contributed by Arthur Kornberg, January 5, 1990

ABSTRACT The *dnaX* gene (previously called *dnaZX*) of *Escherichia coli* has only one open reading frame for a 71-kDa polypeptide from which two distinct DNA polymerase III holoenzyme subunits, τ (71 kDa) and γ (47 kDa), are produced. To determine how the γ subunit is generated, we examined the influence of mutations in the *dnaX* gene on the pattern of τ and γ production in overproducing cells. Important structural elements in *dnaX* mRNA include a stretch of six adenines (nucleotides 1425–1430), a stable hairpin structure (nucleotides 1437–1466), and a UGA stop codon in a -1 frame (nucleotides 1434–1436) between the stretch of adenines and the hairpin structure. Disruption of this stop codon generates a slightly larger γ subunit, indicative of the use of a -1 stop codon farther downstream (nucleotides 1470–1472). These results suggest that a -1 frameshift during translation allows the use of this UGA codon to terminate translation of the γ polypeptide. The amino acid composition, sequence, and mass spectra of a C-terminal peptide from mild digestion of the purified γ protein with endoproteinase Lys-C confirms that this frameshift occurs at either of the two lysine codons in the region of the adenine stretch. Remarkable features of this frameshifting are its high frequency (i.e., about 80% in an overproducing cell) and the striking structural similarity to the frameshifting signal responsible for expression of the *pol* and *pro* genes in many retroviruses.

DNA polymerase III holoenzyme of *Escherichia coli* contains 10 distinctive subunits (1, 2). The core subunits (α , ϵ , and θ) supply the synthesis and proofreading activities; the accessory subunits are responsible for highly processive synthesis of the leading strand and modulated processivity of the lagging strand (3, 4).

Among the accessory subunits, τ (71 kDa) and γ (47 kDa) are essential for DNA replication *in vivo* (5, 6) and *in vitro* (3, 4). Both τ and γ are products of a single gene locus, *dnaX* (or *dnaZX*) (5, 6), which contains only one open reading frame for a 71-kDa polypeptide (643 amino acid residues), the size of τ (Fig. 1) (7, 8). The N-terminal amino acid sequences of purified τ and γ are identical; peptide mapping data indicate that γ is a part of the τ polypeptide (9). The ability of τ and γ to bind ATP and the generation of a γ -like polypeptide by limited tryptic cleavage of τ also support the idea that γ is included within τ (10). Further, the conversion of a fusion protein of τ to β -galactosidase to a γ -like molecule by an *E. coli* crude extract suggested that γ is a product of proteolytic cleavage (11). Based on the amino acid sequence of the product, the cleavage was judged to be between Lys⁴⁹⁸ and Lys⁴⁹⁹ (11). Whether such a cleavage is the means of γ generation in the cell remained to be proven.

In the present study, we show that the generation of the γ subunit does not require proteolytic cleavage. Rather, it is the result of a remarkably frequent -1 ribosomal frameshift which places a nearby downstream UGA stop codon in-frame

to form the truncated polypeptide. Frameshifting at another position had been speculated as a mechanism for producing the γ subunit (1).

MATERIALS AND METHODS

Bacterial Strains and Plasmids. The *E. coli* strain K38 [HFrc(λ)] was used for expression of the *dnaX* gene (12). Plasmid pGP1-2 contains a heat-inducible T7 RNA polymerase gene under the control of the λ P_L promoter (12); pZT3 has the *dnaX* gene under the control of the T7 ϕ 10 promoter and the T7 DNA polymerase Shine–Dalgarno sequence (10); pMK212 bears a cloned *dnaX* gene (5).

DNAs, Proteins, and Other Reagents. Sources were as follows: M13mp19 replicative-form (RF) DNA from Pharmacia; oligonucleotides synthesized by K. Relloma and P. Patek (Stanford University); γ and τ subunits of DNA polymerase III holoenzyme purified as described (10); rabbit antiserum against the γ subunit raised against the purified γ subunit; nitrocellulose filters from Schleicher & Schuell; nonfat dried milk from Lucerne; goat anti-rabbit antibody conjugated with horseradish peroxidase and 4-chloro-1-naphthol from Bio-Rad; and endoproteinase Lys-C from Boehringer Mannheim.

Linker-Scanning Mutants and Expression of Mutated Genes. The 2.1-kilobase *Pst* I–*Sma* I fragment of pMK212 DNA containing the *dnaX* gene was cloned into M13mp19 phage to make a recombinant phage M13mpZX1. Single-stranded M13mpZX1 DNA was used as a template for all the *in vitro* mutagenesis reactions. A 30-mer oligonucleotide containing the *Bam*HI site was hybridized to the template and replication was carried out with DNA polymerase III holoenzyme as described (13). Mutated phage were screened for sensitivity to *Bam*HI. The mutant *dnaX* gene was excised by *Nar* I and *Pst* I cleavage and then introduced into *Nar* I- and *Pst* I-digested pZT3, so that the mutated gene could be expressed by T7 RNA polymerase. Mutant derivatives of pZT3, designated S03–S17 (see Table 1) were introduced into the *E. coli* K38 strain bearing a pGP1-2 plasmid. These cells, which overproduce the *dnaX* gene products 500-fold (10), were grown at 30°C in 5 ml of LB medium supplemented with ampicillin and kanamycin as described (10). At OD₅₉₅ of 0.5, the temperature was raised to 42°C for 1 hr, after which the cells were harvested by centrifugation.

Deletion Mutants. Deletion mutants were derived from linker-scanning mutants described above; digestion by *Bam*HI cleaved the linker and another site in the vector. Digested DNAs were then self-ligated by T4 DNA ligase to generate the deletion between those two *Bam*HI sites.

Western Blotting. Heat-induced overproducing cells from 1 ml of culture were harvested by centrifugation and resuspended in 50 μ l of buffer (0.4 M Tris Cl, pH 6.8/1% SDS/1% 2-mercaptoethanol/0.02% bromophenol blue/30% glycerol), heated 3 min at 90°C, and then sonicated. This lysate was

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.

Abbreviation: FAB, fast-atom bombardment.

*Present address: Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305-5307.

FIG. 1. *dnaX* gene and its products. The *dnaX* gene has one large open reading frame, which can produce two distinct polypeptides: τ (71 kDa) and γ (47 kDa).

diluted 20-fold with the same buffer, and 2 μ l of the diluted sample was analyzed electrophoretically in a 10% polyacrylamide gel (14). Proteins were transferred to a nitrocellulose filter in the Bio-Rad electroblotting apparatus at 200 mA for 3 hr in 25 mM Tris/192 mM glycine/0.25% SDS/20% methanol. The filter was blocked with 5% nonfat dried milk and then incubated with 15 ml of 5% nonfat dried milk plus 30 μ l of anti- γ antiserum for 2 hr at room temperature. The filter was then washed with TBS (10 mM Tris Cl, pH 7.5/0.5 M NaCl) twice for 15 min each, incubated with 15 ml of 5% nonfat dried milk plus 30 μ l of horseradish peroxidase-conjugated goat anti-rabbit antibody for 30 min at room temperature, washed again with TBS as before, and developed briefly in 120 ml of TBS containing 0.015% H_2O_2 , 20% methanol, and 60 ng of 4-chloro-1-naphthol. The developed filter was scanned with a densitometer (Hoefer); the peak on the chart was cut out and its weight was compared with the weights of peaks representing known quantities of purified γ or τ subunit.

Lys-C Cleavage of γ . The reaction mixture (100 μ l) contained 100 μ g of purified γ and 50 ng of endoproteinase Lys-C in 20 mM sodium phosphate, pH 8.5/1 mM EDTA. Digestion was for 10 min at 30°C and was stopped by heating at 90°C for 3 min. When the reaction was performed on a larger scale, more reaction tubes were used rather than changing the reaction volume. After digestion, 900 μ l of cold ethanol was added and the mixture was kept at -20°C for 10 hr. The mixture was centrifuged for 5 min in an Eppendorf microcentrifuge; dried supernatant was analyzed with an N-terminal microsequencer (Applied Biosystems). The same sample was also hydrolyzed in 6 M HCl containing 0.2% 2-mercaptoethanol at 110°C for 24 hr and analyzed on a Beckman 6300 amino acid analyzer.

Mass Spectral Analysis. Samples collected from reverse-phase HPLC in polypropylene microcentrifuge tubes were concentrated to dryness in a vacuum centrifuge. The sample was redissolved in a few microliters of dimethyl sulfoxide. Approximately 2 μ l of the sample solution were added to 1 μ l of the sample matrix on a 1.5 $\times$ 6-mm stainless-steel sample stage. The sample matrix was a mixture of dithiothreitol/dithioerythritol (5:1) (15) and camphorsulfonic acid (6 mM) (16). Positive-ion fast-atom bombardment (FAB) mass spectra were obtained using a JEOL HX-100HF high-resolution, double-focusing, magnetic-sector mass spectrometer operating at 5-kV accelerating potential and a nominal resolution of 3000. Sample ionization was accomplished using a 6-keV (1 eV = 1.602×10^{-19} J) Xe atom beam. A JEOL DA5000 data system was used to control instrument parameters and collect the spectral data.

RESULTS

Termination of the γ Subunit Is Not at Putative Proteolytic Processing Sites. Replacement of the Lys⁴⁹⁸ and Lys⁴⁹⁹ codons, which specify a possible site for proteolytic processing (11), with a *Bam*HI recognition sequence (GGATCC) did not affect the generation of the γ subunit, nor did replacement of an upstream pair of lysine codons (Lys⁴⁷² and

Lys⁴⁷³) (data not shown). These results weaken the argument that proteolytic cleavage of τ is the source of the γ subunit.

A Stretch of Adenines and a Stem-Loop Structure Are Involved in the Generation of the γ Subunit. Several substitutions (mutants S03–S08, S17) disclosed little effect on γ subunit generation relative to that of τ (Table 1; Fig. 2). On the other hand, a profound decrease in γ production was observed for substitutions (mutants S09–S14, S16) in regions which include a stretch of six adenines and one which has the potential for forming a stem-loop structure (Fig. 3).

In addition to τ and γ , a 58-kDa polypeptide reactive with anti- γ antiserum was also apparent in the *dnaX* overproducers (Fig. 2). This polypeptide was also generated in translation *in vitro* (data not shown), but was not present in the purified DNA polymerase III holoenzyme complex (2). How this peptide is generated and what its function might be are unclear.

A Frameshift of -1 in the Stretch of Adenines. Several lines of evidence support a -1 frameshift mechanism as the basis for termination of the γ subunit. A similar stretch of adenine residues preceding a stem-loop structure is required for the frameshifting observed in the expression of the *pol* and *pro* genes in many retroviruses (18). Presumably, pausing at the secondary structure causes the ribosome to slip back over a stretch of iterated nucleotides (A_n or U_n) to shift the frame by one nucleotide. A second line of evidence is the result obtained upon eliminating the potential UGA stop codon (nucleotides 1434–1436) downstream from the adenine stretch (mutant S15). The size of the γ subunit was increased without affecting the level of its production. Implied in this result is the use of the next stop codon downstream (nucleotides 1470–1472), accounting nicely for the enlargement of the γ subunit by 12 amino acid residues.

Table 1. Expression of τ and γ from the linker-scanning mutants

Mutant	Position of mutation*	% γ [†]
pZT3	Wild type	80
S03	1489–1494	84
S04	1481–1486	76
S05	1472–1477	89
S06	1469–1474	84
S07	1466–1471	65
S08	1465–1470	67
S09	1463–1468	49
S10	1460–1465	26
S11	1457–1462	33
S12	1454–1460	41
S13	1445–1450	50
S14	1439–1444	23
S15	1433–1438	71
S16	1427–1432	0
S17	1418–1423	71

*Numbering of nucleotides follows Yin *et al.* (7). These regions were replaced with the GGATCC *Bam*HI recognition sequence except for S12, in which AGGATCC was used instead.

[†]Amounts of γ and τ were determined from densitometric scanning of Western blots shown in Fig. 2; % γ = ($\gamma \times 100$) / ($\gamma + \tau$).

FIG. 2. Expression of τ and γ from linker-scanning mutants. Overproducing cells containing mutated *dnaX* genes were harvested after heat induction and lysates were electrophoresed in an SDS/10% polyacrylamide gel; *dnaX*-derived products (τ and γ) were measured by Western blotting using anti- γ antiserum. Three separate experiments are designated I, II, and III.

Additional and crucial proof of the -1 frameshift mechanism and the site of its action was the amino acid analysis of the C-terminal region of the γ subunit. Although the protein was resistant to the action of carboxypeptidase Y under a variety of conditions, a partial digest by endoproteinase Lys-C, which specifically cleaves on the carboxyl side of a lysine residue, proved successful. Mild digestion generated a product only slightly smaller than the γ subunit; a 50-fold larger amount of the enzyme was needed for further digestion (data not shown). Thus, it appeared that a hypersensitive cleavage site for Lys-C was near an end, either the N or the C terminus. Proximity to the C terminus seemed more likely inasmuch as cleavage at the first lysine near the N terminus, located 49 residues away, would have shortened the γ product far more. The small oligopeptide generated by mild Lys-C digestion was removed after removal of the large product by ethanol precipitation.

Amino acid analysis of the dried supernatant derived from 200 pmol of γ gave the following (in pmol): Ala (208), Lys (408), Glx (62). N-terminal sequence analysis of the same sample derived from 500 pmol of γ gave the following: Ala (340), Lys (270), Lys (220), Glu (70). These results are consistent with the structure H-Ala-Lys-Lys-Glu-OH. This oligopeptide was eluted in the void volume on reversed-phase

HPLC due to its hydrophilicity, and could not be directly analyzed by FAB MS. Therefore, the sample was treated with dansyl chloride and subjected to rechromatography before FAB MS analysis. The results gave a protonated molecular ion of m/z 708.8, consistent with a monodansylated derivative. Thus, the FAB MS data confirm the amino acid composition and sequence analyses.

This sequence suggests a cleavage by Lys-C between Lys⁴²⁷ and Ala⁴²⁸ (Fig. 3). That the C-terminal residue is glutamic rather than serine and that the N-terminal residue is alanine rather than serine indicates that a -1 frameshift takes place at the codon for either Lys⁴²⁹ or Lys⁴³⁰ (Fig. 3).

Regions Downstream from the Stem-Loop Structure Are Dispensable. The requirement for the stem-loop region (nucleotides 1437–1466) immediately downstream from the adenine stretch does not preclude the need for an additional secondary structure farther downstream. Such a requirement has been observed in the case of infectious bronchitis virus, an avian coronavirus that requires a "pseudoknot" structure for frameshifting (19). This possibility was addressed by examining deletion mutants lacking the region downstream of the *Bam*HI linker inserted in the series of substitution mutants described above. Removal of the regions downstream of nucleotide 1465 (mutants D07 and D08) did not interfere with

FIG. 3. mRNA structure of the essential region for γ production. The thick bar indicates the sequence similar to that of retrovirus frameshift sites (18). UGA (in the black box) is the stop codon in the -1 frame. The stem-loop structure was predicted by the FOLD computer program; ΔG was calculated to be -28.4 kcal (17).

Table 2. Frameshifting efficiency in deletion mutants

Mutant	Deletion	Efficiency*, %
pZT3	Wild type	75
D07	1466→	68
D08	1465→	63
D11	1457→	8
D13	1445→	7
D14	1439→	9

Deletions in D07, D08, D11, D13, and D14 were generated using the *Bam*HI site of S07, S08, S11, S13, and S14.

*Products were determined as in Table 1. Frameshift efficiency (%) = (frameshift product × 100)/(frameshift product + in-frame product).

production of the γ subunit, whereas deletions (mutants D11, D13, and D14) which removed the stem-loop decreased γ formation dramatically (Table 2). It seems likely then that no additional secondary structure is needed for the frameshifting that leads to the generation of γ .

DISCUSSION

The 47-kDa γ subunit is encoded by the N-terminal two-thirds of the sequence responsible for the 71-kDa τ subunit (Fig. 1). We learn from the present study that the γ subunit is generated from this sequence by a -1 translational frameshift rather than, as considered earlier (11), by posttranslational proteolysis of the τ subunit. The frameshifting is remarkable for the high frequency that enables γ to be produced in amounts exceeding that of τ , the product of translating the entire *dnaX* gene.

Other examples of programmed frameshifting in *E. coli* include translation of certain genes of phages T7 (20) and ϕ X174 (21), the transposase gene of insertion element IS1 (22), and the gene encoding the translation termination factor RF2 (23–25). In eukaryotes, instances of frameshifting are known in the reverse transcriptase (*pol*) gene of retroviruses and the protease (*pro*) gene of mouse mammary tumor virus (18), in coronavirus (19), and in the yeast transposon Ty1 (26).

The structural features in the mRNA responsible for the -1 frameshift in translation of the *dnaX* gene are strikingly similar to those responsible for gene expression in retroviruses. A stem-loop structure is immediately downstream from a stretch of six adenine residues. Such stretches of iterated nucleotides (i.e., A_n or U_n) preceding a stem-loop are common to many retroviruses (18). These structures allow the tRNA in both the aminoacyl and peptidyl sites of the ribosome to pair with at least two of the three bases after the slippage that generates a -1 frameshift (18).

An interesting distinction between the frameshifting of *dnaX* and that of retroviruses is that the generation of the γ subunit depends on termination, whereas the generation of reverse transcriptase and protease depends on avoiding termination. Another distinction is the presence of a guanine residue following the adenine stretch in *dnaX* as contrasted with a cytosine residue in the several examples of the retroviral genes (18). The importance of this residue in *E. coli* frameshifting is indicated by the more vigorous expression of the mouse mammary tumor virus *pro* gene in *E. coli* when the cytosine residue was replaced by a guanine (27).

Factors that regulate the frameshifting responsible for the generation of the γ subunit remain to be determined. In nonoverproducing cells, τ and γ are present in roughly equal amounts (1). In contrast to this, in cells that overproduce *dnaX* gene products 500-fold, the abundance of the τ subunit is only one-fourth that of γ (Tables 1 and 2, pZT3), suggestive of more frequent frameshifting when expression of the *dnaX* gene is highly exaggerated. In a cell-free transcription–

translation system prepared from *E. coli*, the γ and τ subunits appeared in roughly equal amounts; in this system, the appearance of γ preceded τ (unpublished results). On the other hand, in a eukaryotic system such as wheat germ extract or rabbit reticulocyte lysate, γ subunit production (and the inferred frameshifting) was too feeble to be detected (data not shown). Whether the presence of a guanine residue in place of the cytosine residue following the iterated stretch is responsible for the failure of eukaryotic systems to recognize the *dnaX* gene for frameshifting needs to be examined.

Availability of mutant *dnaX* gene constructions that impede frameshifting and hence prevent the generation of the γ subunit should help clarify whether the γ subunit is essential for the function of DNA polymerase III holoenzyme. Whereas the τ subunit can complement mutations in the *dnaZ* (γ subunit) region of the *dnaX* gene *in vitro*, the γ subunit cannot replace a deficiency in the τ subunit (5, 6, 9). *In vitro*, both subunits bind ATP, but only τ , by virtue of its C-terminal, DNA-binding domain, functions as a DNA-dependent ATPase (10).

In reconstituting the holoenzyme from its core and accessory subunits, τ appears to be able to replace γ (28). Moreover, the S16 mutant *dnaX* gene (Fig. 2), which expressed only the τ subunit, can still rescue *dnaZ* mutants *in vivo* (unpublished results), suggesting that the γ subunit is dispensable when τ is present. Nevertheless, more refined assays that probe the numerous interactions and functions of the complex holoenzyme may show the advantages of having both τ and γ subunits.

We are especially grateful to Mike Ronk and John E. Shively for performing mass spectral and amino acid analyses, which were supported by Cancer Center Core Grant CA33572 to J. E. Shively. We thank Alan J. Smith for valuable suggestions and the protein sequence analyses, Douglas Brutlag for computer analyses, and Robert B. Weiss for communicating results prior to publication. This work was supported by grants from the National Institutes of Health and the National Science Foundation.

- McHenry, C. S. (1988) *Annu. Rev. Biochem.* **57**, 519–550.
- Maki, H., Maki, S. & Kornberg, A. (1988) *J. Biol. Chem.* **263**, 6570–6578.
- Maki, S. & Kornberg, A. (1988) *J. Biol. Chem.* **263**, 6561–6569.
- O'Donnell, M. (1987) *J. Biol. Chem.* **262**, 16558–16565.
- Kodaira, M., Biswas, S. & Kornberg, A. (1983) *Mol. Gen. Genet.* **192**, 80–86.
- Mullin, D. A., Woldringh, C. L., Henson, J. M. & Walker, J. R. (1983) *Mol. Gen. Genet.* **192**, 73–79.
- Yin, K.-C., Blinkowa, A. & Walker, J. R. (1986) *Nucleic Acids Res.* **14**, 6541–6549.
- Flower, A. & McHenry, C. S. (1986) *Nucleic Acids Res.* **14**, 6541–6549.
- Maki, S. & Kornberg, A. (1988) *J. Biol. Chem.* **263**, 6547–6554.
- Tsuchihashi, Z. & Kornberg, A. (1989) *J. Biol. Chem.* **264**, 17790–17795.
- Lee, S.-H., Kanda, P., Kennedy, R. C. & Walker, J. R. (1987) *Nucleic Acids Res.* **15**, 7663–7675.
- Tabor, S. & Richardson, C. C. (1989) *J. Biol. Chem.* **264**, 6447–6458.
- Tsurushita, N., Maki, H. & Korn, L. J. (1988) *Gene* **62**, 135–139.
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
- Witten, J. L., Schaffer, M. H., O'Shea, M., Cook, J. C., Hemling, M. E. & Rinehart, K. L. (1984) *Biochem. Biophys. Res. Commun.* **124**, 350–358.
- DePauw, E., Pelzer, G., Dang, D. V. & Marien, J. (1984) *Biochem. Biophys. Res. Commun.* **123**, 27–32.
- Zuker, M. & Stiegler, P. (1981) *Nucleic Acids Res.* **9**, 133–148.
- Jacks, T., Madhami, H. D., Masiarz, F. R. & Varmus, H. E. (1988) *Cell* **55**, 447–458.
- Brierley, I., Digard, P. & Inglis, S. C. (1989) *Cell* **57**, 537–547.

20. Dunn, J. J. & Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477–535.
21. Buckley, K. J. & Hayashi, M. (1987) *J. Mol. Biol.* **198**, 599–607.
22. Sekine, Y. & Ohtsubo, E. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4609–4613.
23. Craigen, W. J. & Caskey, C. T. (1986) *Nature (London)* **322**, 273–275.
24. Curran, J. F. & Yarus, M. (1988) *J. Mol. Biol.* **203**, 75–83.
25. Weiss, R. B., Dunn, D. M., Atkins, J. F. & Gesteland, R. F. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52**, 687–694.
26. Clare, J. J., Belcourt, M. & Farabaugh, P. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6816–6820.
27. Weiss, R. B., Dunn, D. M., Shuh, M., Atkins, J. F. & Gesteland, R. F. (1989) *New Biologist* **1**, 159–170.
28. Studwell, P. S., Stukenberg, P. T., Onrust, R., Skangalis, M. & O'Donnell, M. E. (1989) *UCLA Symposia on Molecular and Cellular Biology*, eds. Richardson, C. C. & Lehman, I. R. (Liss, New York), Vol. 127, pp. 153–164.