Mol Gen Genet (1991) 227:1–8
0026892591001350

MGG
© Springer-Verlag 1991

# Specificity of recognition sequence for *Escherichia coli* primase

**Kin-ya Yoda and Tuneko Okazaki**

Department of Molecular Biology, School of Science, Nagoya University, Nagoya 464, Japan

Received October 18, 1990

**Summary.** We have surveyed the frequency of each of 64 trinucleotide permutations at every nucleotide frame located from 1 to 15 nucleotides upstream of primer RNA-DNA transition sites mapped within a 1.5 kb region of the bacteriophage lambda genome and a 1.4 kb region of the *Escherichia coli* genome. We have demonstrated that in both systems initiation of DNA synthesis strongly correlates with a CAG sequence located 11 nucleotides upstream of the DNA start sites. Based on the examination of various reports of the priming reaction catalyzed by *E. coli* primase in vivo and in vitro, we propose that (i) *E. coli* primase itself recognizes a 3′GTC 5′ sequence on the template strand, (ii) DnaB helicase releases the specificity of *E. coli* primase and, (iii) the consensus recognition sequence for *E. coli* primase associated with DnaB helicase is 3′PuPyPy 5′.

**Key words:** *E. coli* primase – 3′GTC 5′ – 3′PuPyPy 5′ – Specificity of recognition sequence

## Introduction

DNA primase synthesizes RNA primers at multiple sites on a template DNA strand for discontinuous DNA replication (Ogawa and Okazaki 1980; DePamphilis and Wassarman 1980). An important question to be resolved concerns the structure on the template strand which DNA primase recognizes in order to catalyze frequent initiation of primer RNA synthesis while a replication fork moves processively. Essential trinucleotide sequences that are recognized by primases from bacteriophage T7 or T4 have been elucidated by mapping of the initiation sites of Okazaki fragments in vivo and in vitro (Fujiyama et al. 1981; Sugimoto et al. 1988; Tabor and Richardson 1981; Cha and Alberts 1986). Bacteriophage T7 primase (gene 4 protein, which also acts as a helicase) synthesizes predominantly tetranucleotide primers, starting with pppAC, at the trinucleotide sequence 3′CTG 5′ of the template strand (Fujiyama et al. 1981; Sugimoto et al. 1988; Tabor and Richardson 1981; Seki and Okazaki 1979; Ogawa and Okazaki 1979; Scherzinger et al. 1977; Romano and Richardson 1979). Bacteriophage T4 primase (gene 61 protein) synthesizes pentanucleotide primers, starting with pppAC or pppGC, at the trinucleotide sequences 3′TTG 5′ or 3′TCG 5′, respectively, of the template strand in a reaction requiring simultaneous action of gene 41 helicase (Cha and Alberts 1986; Kurosawa and Okazaki 1979; Hinton and Nossal 1987; Nossal and Hinton 1987). Thus, essential nucleotide sequences recognized by both bacteriophage systems are trinucleotide stretches composed of one untranscribed nucleotide followed by the sequence complementary to the first dinucleotide sequence of primer RNA molecules.

Intensive studies have been performed on *Escherichia coli* primase (*dnaG* gene product) for the last two decades, and the fundamental mechanism of the priming reaction has been established using in vitro systems (Kornberg 1974, 1982). However consistent interpretation of the sequence specificity of *E. coli* primase has not yet been obtained. In conversion of viral strand (v-strand) DNA of single-strand phage G4, st-1, $\phi$K, or $\alpha$3 to replicative form molecules (RF), DnaG protein synthesizes an RNA primer starting with pppAG at a unique site within the 135-base origin sequence in the presence of SSB (single-stranded DNA binding protein) (Bouche et al. 1987; Fiddes et al. 1978; Sims and Dressler 1978; Sims et al. 1979). With all these phage DNAs, primer synthesis starts at the T residue in a 3′GTC5′ sequence located at the foot of a hairpin structure (v-strand). On the other hand in most other systems examined, such as the general priming reaction (Arai and Kornberg 1979, 1981b), the primosome system for phage $\phi$X174 (Arai and Kornberg 1981c; Arai et al. 1981) and priming systems for phage $\lambda$ (Dodson et al. 1985, 1986) and *E. coli oriC* (Kaguni and Kornberg 1984; Ogawa et al. 1985; van der Ende et al. 1985; Bramhill and Kornberg 1988), DnaG primase synthe-

*Offprint requests to:* K. Yoda

sizes RNA primers at multiple sites on a template strand; in all these reactions primer synthesis by DnaG primase is aided by DnaB helicase. Although the structure of the primer RNAs synthesized by these systems has been reported (Arai and Kornberg 1981b; Ogawa et al. 1983), it is clear that no simple recognition sequence like that of T7 or T4 phage primase exists for DnaG primase action. However, the primer RNA molecules of Okazaki DNA fragments synthesized in *E. coli* in vivo systems have a chain length of $11 \pm 1$ nucleotides and the 5′ terminal dinucleotide sequence is composed of predominantly (80%) pppPuPu (Pu, purine), 50% of which is pppAG (Kitani et al. 1985). Thus, the 5′ terminal dinucleotide structure made in vivo is not random, even if not unique.

We have reported previously the distribution of the initiation sites of DNA synthesis (RNA-DNA transition sites) in vivo on both strands in a 1.4 kb region of the *E. coli* genome (Kohara et al. 1985) and in a 1.5 kb region of the $\lambda$ phage genome (Yoda et al. 1988). We have investigated the properties of the nucleotide sequences in the region upstream of the transition sites in terms of the structural information required for primer RNA synthesis for discontinuous DNA replication in *E. coli* (Kitani et al. 1985). We found from the mapping data of the $\lambda$ phage genome that a CPuPu (especially CAG) sequence frequently appears 12 to 10 nucleotides upstream of the initiation sites of DNA synthesis. We suggested the possibility that *E. coli* primase recognizes the trinucleotide sequences 3′GPyPy 5′ (especially 3′GTC 5′) (Py, pyrimidine) located 12 to 10 nucleotides upstream of the initiation sites (Yoda et al. 1988). The results also suggest that the postulated sequence specificity of *E. coli* primase is much looser compared with that of T7 or T4 primase indicating that we might have overlooked the significance of the sequence specificity in the upstream region of the initiation sites of the *E. coli* genome (Kohara et al. 1985). A statistical analysis was required in order to evaluate the available data and to elucidate the specificity of the recognition sequence of *E. coli* primase.

In this report, we have extensively surveyed the frequency of appearance of each of 64 trinucleotide sequences at every nucleotide position between positions $-1$ and $-15$ from RNA-DNA transition sites located in the replication origin regions of the genomes of *E. coli* (Kohara et al. 1985) and bacteriophage lambda (Yoda et al. 1988) (triplet analysis). The results reveal that the distribution of the trinucleotide CAG shows

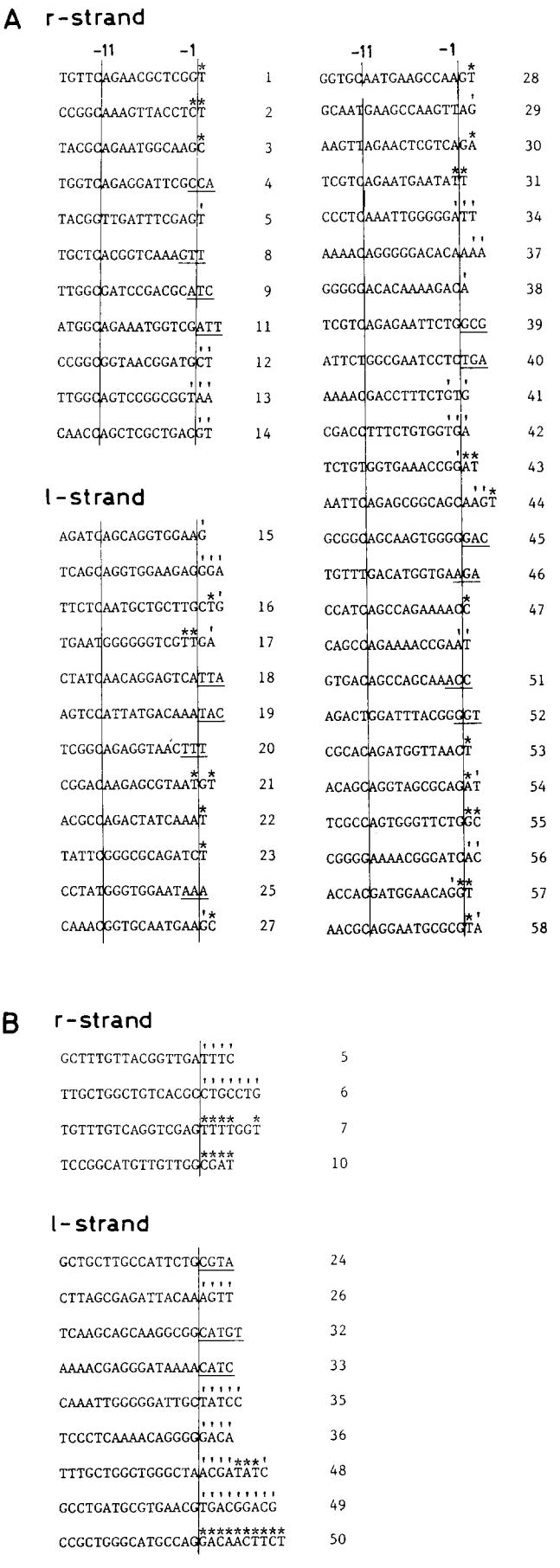


**Fig. 1A and B.** Alignment of the initiation sites of DNA synthesis in vivo mapped within a 1.5 kb region containing *ori*$\lambda$. The principles of the arrangement shown in **A** are described in the text. The transition sites from primer RNA to DNA synthesis and the serial number affixed to each site are indicated according to Yoda et al. (1988). Nucleotides marked by an *asterisk* indicate sites located by both size of bands and nucleotide compositions of 5′ ends; nucleotides marked by an *apostrophe* indicate sites located by size of bands only, but with high accuracy; *underlined* nucleotides indicate sites located by size of bands which may contain molecules of different chain lengths

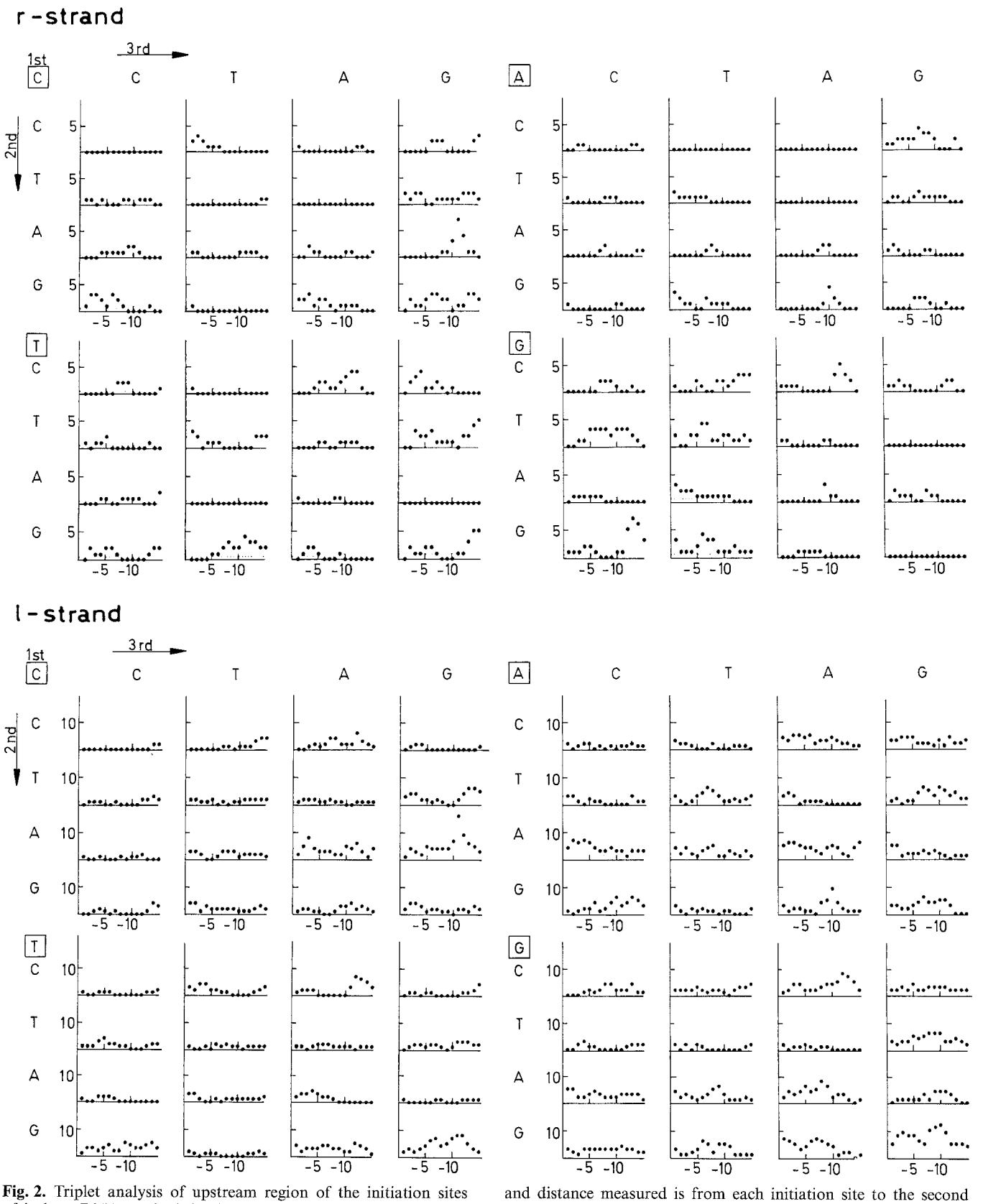


**Fig. 2.** Triplet analysis of upstream region of the initiation sites of $\lambda$ phage DNA synthesis in vivo. Trinucleotide sequences located within positions $-1$ to $-15$ of every initiation site shown in Fig. 1A and B were counted as described in the Materials and methods. Each trinucleotide sequence is numbered from the 5′ side and distance measured is from each initiation site to the second letter of a triplet. A *dotted line* in each column indicates the frequency of appearance of the corresponding triplet if primer initiation occurred at random

a sharp peak at position −11 in both systems, strongly supporting the possibility that during the start of DNA replication of *E. coli* and bacteriophage λ in vivo, *E. coli* primase most frequently recognizes a 3′GTC 5′ sequence on a teplate strand and synthesizes an 11±1 nucleotide primer RNA starting with an AG sequence.

## Materials and methods

All the data on the initiation sites of DNA synthesis in vivo were taken from Yoda et al. (1988) and Kohara et al. (1985).

*Triplet analysis.* The frequency of each of the 64 possible trinucleotide sequences located within positions −1 to −15 of every initiation site was counted by computer and plotted against distance from the initiation site (Figs. 2, 4).

## Results

### *Arrangement of λ DNA initiation sites*

Nucleotide sequences immediately upstream of initiation sites of DNA synthesis detected in vivo on both strands in a 1.5 kb region of the λ phage genome containing the replication origin (*oriλ*) (Yoda et al. 1988), are listed in Fig. 1. Arrangement of the sequences in Fig. 1A was performed on the basis of the following principles. (i) An initiation site consisting of more than three consecutive nucleotides is omitted and is listed in Fig. 1B. (ii) When an initiation site consists of two or three consecutive nucleotides, the 5′ side or the middle, respectively, is allocated as the initiation point (position +1). (iii) A leftward or rightward shift by one nucleotide is permitted in order to place the PyPu sequence at positions −12 to −11. As a result, PyPuPu sequences exist at positions −12 to −10 in 42 out of 48 sites (89%) or PyPu sequences at positions −12 to −11 in 44 out of 48 sites (92%) (Fig. 1A). Moreover, 55% (6/11) and 43% (16/37) of the PyPuPu sites are occupied by CAGs in the r-strand and in the l-strand sites, respectively, suggesting a positive correlation between the presence of the triplet CAG and the start of RNA primers with a chain length of around 11 residues.

### *Triplet analysis of λ phage DNA initiation sites*

To identify features characteristic of nucleotide sequences upstream of every initiation site of DNA synthesis more objectively, a statistical survey of the frequency of the 64 trinucleotide permutations was carried out at every trinucleotide frame from positions −1 to −15 of the initiation sites listed in Fig. 1A and B. The number of each triplet found at each nucleotide position was plotted as a function of distance from the initiation site (Fig. 2). As shown in Fig. 2, the frequency of the CAG sequence shows a sharp peak at position −11 (the number indicates the position of the center nucleotide of a triplet) with both l- and r-strand initiation sites. Furthermore, distinct peaks are found with AGA at −10 and GCA at −12. As shown in Fig. 1, 11 out of 12 AGAs come from CAGA and 11 out of 13 GCAs from GCAG. Therefore both of these trinucleotide sequences appear together with a CAG sequence at position −11. A peak of GGC was found at −13 for the r-strand but not for the l-strand. Essentially the same results were obtained when the analysis was performed only with the sites listed in Fig. 1A (data not shown). These results clearly show the parallelism between a

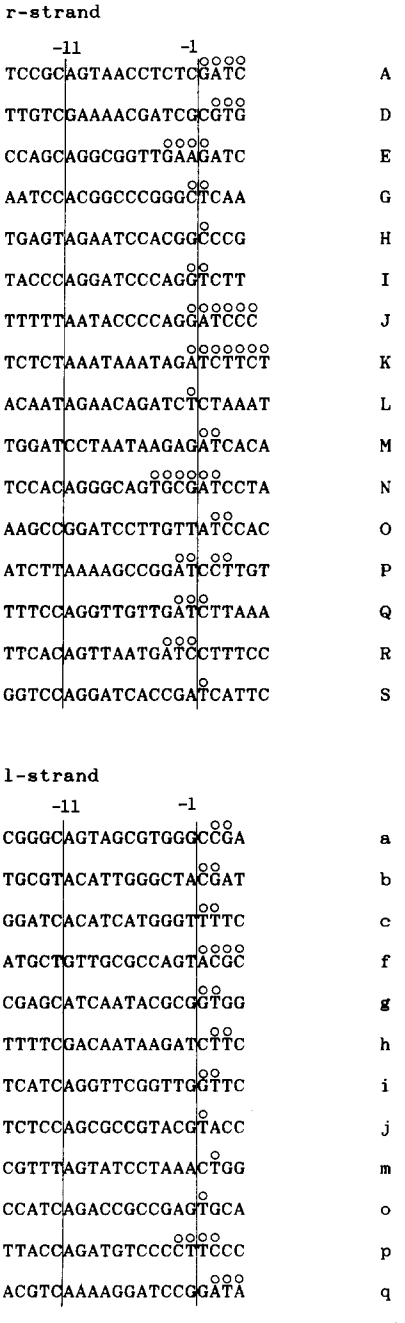


**Fig. 3.** Arrangement of the transition sites from primer RNA to *Escherichia coli* DNA synthesis in vivo near the *oriC* region. The transition sites are marked with *open circles* above the sequence and are assigned letters according to the data of Kohara et al. (1985)

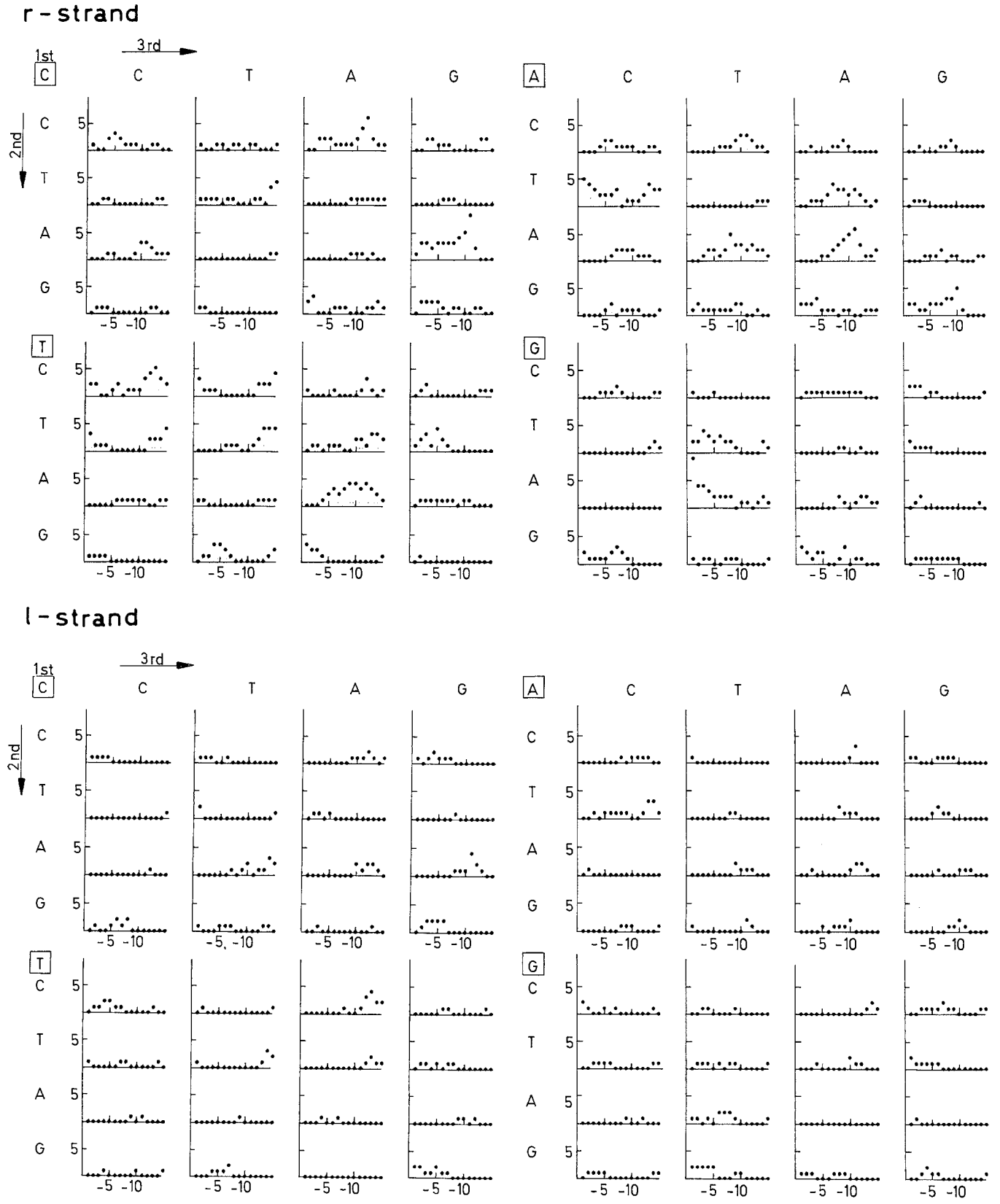


**Fig. 4.** Application of the triplet analysis to the mapping data of *E. coli* DNA synthesis in vivo shown in Fig. 3. Further details are described in the Materials and methods and the legend to Fig. 2

CAG sequence at an average position −11 and the initiation of DNA synthesis in the λ phage origin region in vivo.

### *Triplet analysis of* E. coli *DNA initiation sites*

Initiation sites of DNA synthesis in vivo detected by Kohara et al. (1985) are listed in Fig. 3. By sliding the position of the deoxynucleotide at RNA to DNA transition sites a few nucleotides in the plus or minus directions, the presence in high frequency of the CAG sequence around position −11 became obvious; 42% (5/12) and 44% (7/16) in l-strand and r-strand, respectively. This was confirmed statistically by the triplet analysis (Fig. 4). A peak of CAG sequence is detected at position −11 for both l- and r-strands. Other triplet peaks found in the r-strand, AGG at −10 and CCA at −12, are derived from the sequence CAGG (5 out of 5) and CCAG (4 out of 5), respectively. These results from both the λ phage and *E. coli* data indicate that a CAG sequence is present in high frequency at positions −12 to −10, the starting sites of primer RNA with a chain length of around 11 nucleotides. A high GAT peak observed at position −1 for the r-strand reflects the high frequency of GATC sites in the *oriC* region, proximal to which DNA synthesis initiates at high frequency, as previously pointed out by Kohara et al. (1985). However, statistical analysis fails to show this tendency for the l-strand.

## Discussion

In this paper we have demonstrated that in the *E. coli* and λ phage in vivo systems, initiation of DNA synthesis strongly correlates with a CAG sequence located 11 nucleotides upstream of the DNA start site. Based on comparison of the present results with the structural properties of primer RNA molecules isolated from *E. coli* Okazaki fragments (Kitani et al. 1985) and with the sequence specificity of DnaG primase in solo DnaG primase reactions as discussed in the following paragraph, we conclude that in the *E. coli* and λ phage in vivo systems *E. coli* primase preferentially recognizes a 3′GTC 5′ sequence in a template strand and synthesizes an 11 ± 1 nucleotide primer RNA. The areas mapped (1.5 kb for λ phage and 1.4 kb for *E. coli*) were considerably larger than the *ori* regions (163 bp for λ phage and 245 bp for *E. coli*) and no conspicuous differences in nucleotide compositions of the upstream regions of the initiation sites were found between the *ori* regions and the rest. Thus, the results of the triplet analyses may be attributed to the function of *E. coli* primase in the priming machinery for discontinuous DNA replication. *E. coli* primase may also recognize trinucleotide sequences on a template strand for primer synthesis, as does T7 or T4 primase, but in a far looser way so that we can only define the characteristics of the recognition sequence by statistical analysis.

In Fig. 5A, DNA sequences are shown at which primer RNA synthesis by the DnaG primase solo reaction is known to start; this reaction requires SSB but not DnaB protein. N. Nomura et al. (personal communication) have recently isolated the DNA segments that support SS (single-stranded DNA) to RF synthesis in solo DnaG primase reactions, where primer RNA synthesis starts at the unique T residue in the 3′GTC 5′ sequence as indicated in Fig. 5A (lines c to e) and the size of primer RNA of the RNA-DNA molecules is 12 to 14 nucleotides (Masai et al. 1990). A 3′GTC 5′ sequence was also found 10 nucleotides upstream of the initiation site (RNA-DNA transition site) of leading strand DNA synthesis of the R100 plasmid in vitro (Fig. 5B) (Miyazaki et al. 1988). Masai and Arai (1989) have proposed that leading strand synthesis in R1 plasmid, a plasmid related to R100, is primed by DnaG primase alone. A start site for leading strand DNA synthesis in the R1 plasmid has been located within a region of 6 nucleotides as shown in Fig. 5B and the size of the primer RNA of RNA-DNA molecules is approximately 8 nucleotides (Masai and Arai 1989). Recently Hiasa et al. (1989, 1990) have demonstrated by mutational analysis of the G4 origin that a 3′GTC 5′ sequence is critically important for the origin function, suggesting that the trinucleotide sequence is necessary for priming in vivo and in vitro. These results raise the possibility that in the primase solo reaction, DnaG primase recognizes a specific 3′GTC 5′ sequence on a template strand as a priming site and synthesizes a primer RNA of approximately 11 nucleotides.

It has been reported that *E. coli* primase synthesizes a primer RNA which is 10 to 60 nucleotides long, without apparent sequence specificity, at multiple sites on a single-stranded DNA template, with the aid of DnaB protein in the absence of SSB (general priming system)

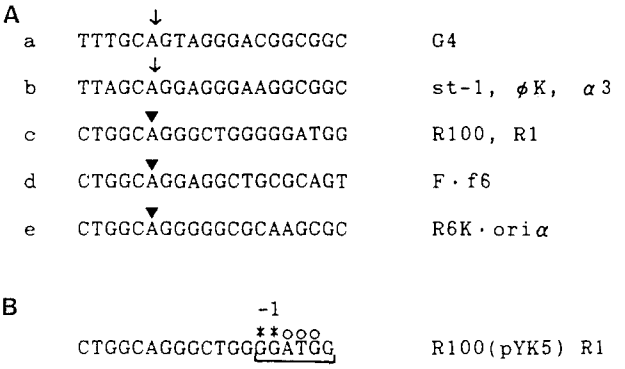


**Fig. 5A and B.** DNA sequences from various phages and plasmids where *E. coli* primase primes DNA synthesis. The complementary sequence of each template strand is given for comparison with the mapping data shown in Figs. 1 and 3. The source of each sequence is indicated at the right. **A** The start site of RNA molecules (*arrows*) (Bouche et al. 1987; Fiddes et al. 1978; Sims and Dressler 1978) and the major start site of primer RNA of RNA-DNA molecules (*filled arrowheads*) (Masai et al. 1990) are indicated. **B** The sequence containing the initiation sites of leading strand DNA synthesis in vitro of R100 and R1 plasmids, which is located about 400 bp downstream of the *ori* sequence. The nucleotide sequence around the initiation site of the leading strand is exactly the same for both plasmids. The bases determined as 5′ termini of the leading strand DNA are indicated by *open circles* and those detected as ribonucleotides are indicated by *asterisks* for R100 (Miyazaki et al. 1988); the area defined as the 5′ terminus of the leading strand DNA for R1 (Masai and Arai 1989) is *bracketed*

**Table 1.** Sequence specificity of primase from various sources

| Source | Gene for primase | Associated protein | Recognition sequence (3′→5′) | Size of primer RNA |
|---|---|---|---|---|
| T7 | gene 4 | P gene 4 (56 kDa) | CTG (PyPuPu) | 4∼5 (mainly 4) |
| T4 | gene 61 | P gene 41 | $T^{T}_{C}G$ (PyPuPu) | 4∼5 (mainly 5) |
| G4, st-1, $\phi$K, $\alpha$3 | dnaG | | GTC (PuPyPy) | |
| $\lambda$, *E. coli* | dnaG | DnaB | GPyPy (PuPyPy) | 11±1 |
| Mammal (SV40) | | | PuPy (PuPy(Py)) | ∼8 |

(Arai and Kornberg 1979, 1981b). Re-examination of their data in the light of the information that *E. coli* primase synthesizes 11±1 nucleotide RNA primers in vivo, we found that primers about 11 nucleotides in length were indeed preferentially synthesized, and synthesis of longer sized primers was selectively suppressed by the addition of SSB (Fig. 2 of Arai and Kornberg 1979). In solo primase reaction, primers about 11 nucleotides in length were often synthesized as a major component among longer sized primers (Masai et al. 1990; see also footnote of Sims et al. 1979). These data suggest that the *E. coli* primase might preferentially synthesize RNA primers approximately 11 nucleotides in length. When priming reactions by DnaG primase with DnaB protein were coupled with DNA synthesis in the $\phi$X174 primosome system, which was proposed as a model for discontinuous DNA synthesis in *E. coli* chromosome replication (Arai and Kornberg 1981c; Arai et al. 1981), the primer RNA became shorter and heterogeneous in length, 1 to 9 nucleotides (Ogawa et al. 1983).

The 5′ terminal dinucleotide sequences of RNA-DNA molecules containing mono- and dinucleotide primers were exclusively pppAdG and pppAG (Ogawa et al. 1983). Primers longer than dinucleotides were composed of 81% pppAPu (25% pppAG and 56% pppAA) (Ogawa et al. 1983), while those of *E. coli* primers made in vivo were composed of 66% pppAPu (41% pppAG and 25% pppAA) (Kitani et al. 1985). Therefore the site selection preference of DnaG primase in the primosome system (3′TPy5′ on a template strand) is in accordance with that in *E. coli* discontinuous DNA synthesis in vivo. A possible interpretation of these results is that DnaG primase associated with DnaB protein acquires a less stringent sequence specificity for priming at multiple sites on a template strand in the primosome system as well as in the general priming system. The transition sites from RNA to DNA in the primosome system appear very frequently and the transition points can occur anywhere from 4 up to 14 nucleotides from the putative primer initiation site (Ogawa et al. 1983), which may reflect the heterogeneous size of primer RNAs made in the primosome system.

Two conditions must be fulfilled in order to detect a peak of CAG distribution at position −11, namely recognition of the 3′GTC 5′ sequence by *E. coli* primase and synthesis of primer RNA of unique size (11±1 nucleotides). Initiation sites for $\lambda$ phage DNA synthesis in vitro were mapped around the $\lambda$ *ori* region by Tsurimoto and Matsubara (1984). The initiation sites A, V, Y, and Z shown in Fig. 4 of Tsurimoto and Matsubara (1984) correspond to the in vivo sites 4, 26, 31, and 32 shown in Fig. 4 of Yoda et al. (1988), respectively. CAG sequences are detected 9 to 12 nulceotides upstream of 6 sites (A, B, G, I, Y, and Z) out of 10 major initiation sites. However when triplet analysis was performed on the initiation sites shown in Fig. 4 of Tsurimoto and Matsubara (1984) and those located by Tsurimoto using different templates containing $\lambda$ *ori* sequence (personal communication), no peak was found in the distribution of CAG (data not shown). The features of the initiation sites of $\lambda$ DNA synthesis in vitro were that they were mostly located at all positions from 3 to 8 nucleotides from the putative RNA primer start as in the $\phi$X174 primosome system. Therefore the failure to detect a CAG peak at position −11 may be mainly due to the heterogeneous size of primers. The $\lambda$ in vitro system (Tsurimoto and Matsubara 1984) and the $\phi$X174 primosome system (Ogawa et al. 1983) seem to lack some factor required for synthesis of primer RNA of a unique size.

It has been proposed that DnaB protein interacts with *E. coli* primase and facilitates priming at multiple sites on the template DNA by altering DNA secondary structures for primase action (Arai and Kornberg 1981a). The present results suggest that DnaB protein may also release the specificity of the triplet recognition of *E. coli* primase. The consensus recognition sequence of *E. coli* primase can be summarized as 3′PuPyPy 5′ (or 3′GPyPy 5′). Hay et al. (Hay and DePamphilis 1982; Hay et al. 1984) have shown that the initiation sites of primer RNA synthesis in vivo around SV40 *ori* sequence consist mainly of 3′PuPy 5′ sequence (about 80%) on the template strand and primer RNA starts with pppPu using the 5′ Py of the template strand. Furthermore 3′PuPyPy 5′ sequence occupied 50% of the total initiation sites of primer RNA, so the third Py of the sequence might also be part of the recognition sequence for the primase (Hay et al. 1984). Therefore it is possible that the trinucleotide sequence may also be recognized by eukaryotic primase, although the sequence specificity seems to be less stringent than that of *E. coli* primase. Table 1 summarizes recognition sequences of primases from various sources. Interestingly, the recognition sequence of *E. coli* primase (3′PuPyPy 5′) more closely resembles that of eukaryotic primase (3′PuPy(Py) 5′) than that of T7 or T4 primase (3′PyPyPu 5′).

*Acknowledgements.* We thank Dr. A. Kornberg for encouragement in the writing of the manuscript. We are grateful to Dr. H. Honda for computer search, Dr. H. Masukata for useful discussion and Dr. C.C. Richardson for critical reading of the manuscript. We also thank M. Hirano for preparing the typescript. This work was supported by Grants-in-Aid for Special Project Research from the Ministry of Education, Science, and Culture of Japan.

## References


Arai K, Kornberg A (1979) A general priming system employing only *dnaB* protein and primase for DNA replication. Proc Natl Acad Sci USA 76:4308–4312

Arai K, Kornberg A (1981a) Mechanism of *dnaB* protein action. III. Allosteric role of ATP in the alteration of DNA structure by *dnaB* protein in priming replication. J Biol Chem 256:5260–5266

Arai K, Kornberg A (1981b) Mechanism of *dnaB* protein action. IV. General priming of DNA replication by *dnaB* protein and primase compared with RNA polymerase. J Biol Chem 256:5267–5272

Arai K, Kornberg A (1981c) Unique primed start of phage $\phi$X174 DNA replication and mobility of the primosome in a direction opposite chain synthesis. Proc Natl Acad Sci USA 78:69–73

Arai K, Low RL, Kornberg A (1981) Movement and site selection for priming by the primosome in phage $\phi$X174 DNA replication. Proc Natl Acad Sci USA 78:701–711

Bouche J-P, Rowen L, Kornberg A (1978) The RNA primer synthesized by primase to initiate phage G4 DNA replication. J Biol Chem 253:765–769

Bramhill D, Kornberg A (1988) Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. Cell 52:743–755

Cha T-A, Alberts B (1986) Studies of the DNA helicase-RNA primase unit from bacteriophage T4. J Biol Chem 261:7001–7010

DePamphilis ML, Wassarman PM (1980) Replication of eukaryotic chromosomes: A close-up of the replication fork. Annu Rev Biochem 49:627–666

Dodson M, Roberts J, McMacken R, Echols H (1985) Specialized nucleoprotein structures at the origin of replication of bacteriophage $\lambda$: Complexes with $\lambda$O protein and with $\lambda$O, $\lambda$P, and *Escherichia coli* DnaB proteins. Proc Natl Acad Sci USA 82:4678–4682

Dodson M, Echols H, Wickner S, Alfano C, Kojo M-W, Gomes B, LeBowitz J, Roberts JD, McMacken R (1986) Specialized nucleoprotein structures at the origin of replication of bacteriophage $\lambda$: Localized unwinding of duplex DNA by a six-protein reaction. Proc Natl Acad Sci USA 83:7638–7642

Fiddes JC, Barrell BG, Godson GN (1978) Nucleotide sequences of the separate origins of synthesis of bacteriophage G4 viral and complementary DNA strands. Proc Natl Acad Sci USA 75:1081–1085

Fujiyama A, Kohara Y, Okazaki T (1981) Initiation sites for discontinuous DNA synthesis of bacteriophage T7. Proc Natl Acad Sci USA 78:903–907

Hay RT, DePamphilis ML (1982) Initiation of SV40 DNA replication in vivo: Location and structure of 5′ ends of DNA synthesized in the *ori* region. Cell 28:767–779

Hay RT, Hendrickson EA, DePamphilis ML (1984) Sequence specificity for the initiation of RNA-primed simian virus 40 DNA synthesis in vivo. J Mol Biol 175:131–157

Hiasa H, Sakai H, Tanaka K, Honda Y, Komano T, Godson N (1989) Mutational analysis of the primer RNA template region in the replication origin (*ori*) of bacteriophage G4: priming signal recognition by *E. coli* primase. Gene 84:9–16

Hiasa H, Sakai H, Komano T, Godson GN (1990) Structural features of the priming signal recognized by primase: mutational analysis of the phage G4 origin of complementary DNA strand synthesis. Nucleic Acids Res 18:4825–4831

Hinton DM, Nossal NG (1987) Bacteriophage T4 DNA primase-helicase. J Biol Chem 262:10873–10878

Kaguni JM, Kornberg A (1984) Replication initiated at the origin (*oriC*) of the *E. coli* chromosome reconstituted with purified enzymes. Cell 38:183–190

Kitani T, Yoda K, Ogawa T, Okazaki T (1985) Evidence that discontinuous DNA replication in *Escherichia coli* is primed by approximately 10 to 12 residues of RNA starting with a purine. J Mol Biol 184:45–52

Kohara Y, Tohdoh N, Jiang X-W, Okazaki T (1985) The distribution and properties of RNA primed initiation sites of DNA snythesis at the replication origin of *Escherichia coli* chromosome. Nucleic Acids Res 13:6847–6866

Kornberg A (1974) DNA replication. Freeman, San Francisco

Kornberg A (1982) Supplement to DNA Replication. Freeman, San Francisco

Kurosawa Y, Okazaki T (1979) Structure of the RNA portion of the RNA-linked DNA pieces in bacteriophage T4-infected *Escherichia coli* cells. J Mol Biol 135:841–861

Masai H, Arai K (1989) Leading strand synthesis of R1 plasmid replication in vitro is primed by primase alone at a specific site downstream of *oriR*. J Biol Chem 264:8082–8090

Masai H, Nomura N, Kubota Y, Arai K (1990) Roles of $\phi$X174 type primosome- and G4 type primase-dependent primings in initiation of lagging and leading strand syntheses of DNA replication. J Biol Chem 265:15124–15133

Miyazaki C, Kawai Y, Ohtsubo H, Ohtsubo E (1988) Unidirectional replication of plasmid R100. J Mol Biol 204:331–343

Nossal NG, Hinton DM (1987) Bacteriophage T4 DNA primase-helicase. J Biol Chem 262:10879–10885

Ogawa T, Okazaki T (1979) RNA-linked nascent DNA pieces in phage T7-infected *Escherichia coli*. III. Detection of intact primer RNA. Nucleic Acids Res 7:1621–1633

Ogawa T, Okazaki T (1980) Discontinuous DNA replication. Annu Rev Biochem 49:421–457

Ogawa T, Arai K, Kornberg A (1983) Site selection and structure of DNA-linked RNA primers synthesized by the primosome in phage $\phi$X174 DNA replication in vitro. J Biol Chem 258:13353–13358

Ogawa T, Baker TA, van der Ende A, Kornberg A (1985) Initiation of enzymatic replication at the origin of the *Escherichia coli* chromosome: contributions of RNA polymerase and primase. Proc Natl Acad Sci USA 82:3562–3566

Romano LJ, Richardson CC (1979) Characterization of the ribonucleic acid primers and the deoxyribonucleic acid product synthesized by the DNA polymerase and gene 4 protein of bacteriophage T7. J Biol Chem 254:10483–10489

Scherzinger E, Lanka E, Morelli G, Seiffert D, Yuki A (1977) Bacteriophage T7 induced DNA-priming protein: A novel enzyme involved in DNA replication. Eur J Biochem 72:543–558

Seki T, Okazaki T (1979) RNA-linked nascent DNA pieces in phage T7-infected *Escherichia coli*. II. Primary structure of the RNA portion. Nucleic Acids Res 7:1603–1619

Sims J, Dressler D (1978) Site-specific initiation of a DNA fragment: Nucleotide sequence of the bacteriophage G4 negative-strand initiation site. Proc Natl Acad sci USA 75:3094–3098

Sims J, Capon D, Dressler D (1979) *dnaG* (Primase)-dependent origins of DNA replication. J Biol Chem 254:12615–12628

Sugimoto K, Miyasaka T, Fujiyama A, Kohara Y, Okazaki T (1988) Change in priming sites for discontinuous DNA synthesis between the monomeric and concatemeric stages of phage T7 replication. Mol Gen Genet 211:400–406

Tabor S, Richardson CC (1981) Template recognition sequence for RNA primer synthesis by gene 4 protein of bacteriophage T7. Proc Natl Acad Sci USA 78:205–209

Tsurimoto T, Matsubara K (1984) Multiple initiation sites of DNA replication flanking the origin region of $\lambda$dv genome. Proc Natl Acad Sci USA 81:7402–7406

van der Ende A, Baker TA, Ogawa T, Kornberg A (1985) Initiation of enzymatic replication at the origin of the *Escherichia coli* chromosome: Primase as the sole priming enzyme. Proc Natl Acad Sci USA 82:3954–3958

Yoda K, Yasuda H, Jiang X-W, Okazaki T (1988) RNA-primed initiation sites of DNA replication in the origin region of bacteriophage $\lambda$ genome. Nucleic Acids Res 16:6531–6546


Communicated by M. Sekiguchi