Volume 246, number 1,2, 21–24 FEB 06934 March 1989

Identification of single-strand initiation signals in the *terC* region of the *Escherichia coli* chromosome

Hiroshi Hiasa, Hiroshi Sakai and Tohru Komano

Laboratory of Biochemistry, Department of Agricultural Chemistry, Kyoto University, Kyoto 606, Japan

Received 4 January 1989

On the basis of clear-plaque formation, we detected initiation signals in the *terC* region of the *Escherichia coli* chromosome. At least two single-strand initiation signals were identified from the *terC* region. The nucleotide sequences of these two signals were determined. Sequence homologies, variations of the consensus sequence of n′ protein recognition sites, 5′-GAAGCGG-3′, were found within these signals. A novel conserved sequence was also found within these signals. Their initiation activities were measured both by the infection growth assay and by the ability to convert the single-stranded DNA to the duplex replicative form DNA in vivo.

DNA replication; Replication origin; Phage vector, M13; Chromosome region, *terC*

1. INTRODUCTION

In the replication of the *Escherichia coli* (*E. coli*) chromosome, in addition to DNA chain initiation occurring at *oriC*, repeated initiations are also associated with fork movement to make primer RNAs for synthesis of the lagging strand. Since melting of a duplex to a single-stranded form may occur before DNA chain initiation, DNA chain initiations seem to take place on ss-DNA templates. Some efforts have been made to identify the primosome assembly sites in the *oriC* region of the *E. coli* chromosome. However, these efforts were in vain because no DNA sequences within the *oriC* region or its close vicinity support initiation of DNA synthesis [1,2]. In contrast, some n′ protein recognition sequences which promote DNA initiation are located within and clockwise to the *asnA* gene [2,3]. The *terC* region of the *E. coli* chromosome seems interesting for the study of initiation signals because at least one of the fixed initiation sites is mapped in the *terC* region.

It has been proposed [4] that DNA chain initiations on single-stranded templates are directed by specific sequences in the template strand. These determinants are called *ssi* signals. On the ss-DNA of ϕX174 [5] and ColE1 plasmid [6–10], some *ssi* signals, which contain specific DNA sequences directing the assembly of the primosome [11–13], have been identified. To select *ssi* signals, M13ΔE101 [14], a M13 deletion mutant, and its derivatives were used. With these origin-probe vectors, several *ssi* signals have been characterized from phages [4,15,16], plasmids [7,8,17–19], and bacterial chromosomes [20].

We describe here the identification of *ssi* signals in the *terC* region of the *E. coli* chromosome. At least two *ssi* signals, designated *sitA* and *sitB*, were identified.

Correspondence address: T. Komano, Laboratory of Biochemistry, Department of Agricultural Chemistry, Kyoto University, Kyoto 606, Japan

Abbreviations: RF, replicative form DNA; ss, single-stranded DNA; *ssi*, single-strand initiation of DNA synthesis; nt, nucleotide(s); pfu, plaque forming units; moi, multiplicity of infection

2. EXPERIMENTAL

M13Δ*lac*183 is an *ssi* signal-probe vector derived from the filamentous phage M13 (Nomura, N., personal communica-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

tion). It forms small turbid plaques because a part of the complementary DNA strand origin is deleted. Upon insertion of an *ssi* signal into M13Δ*lac*183 or its equivalent, the recombinant phage does form large clear plaques [18,19]. This is the primary criterion for *ssi* signals. *E. coli* chromosomal DNA, prepared from plasmid pBS4, was digested by several restriction endonucleases and then ligated into M13Δ*lac*183. After transfection to *E. coli* RL108(21), clones that formed clear plaques were selected. Phage growth was measured by the method described in [18,19]. The accumulation of phage RFI DNA in *E. coli* RL108 cells was estimated by the method of Lambert et al. [22].

3. RESULTS

Plasmid pBS4 carries the chromosomal DNA segment from the *terC* region, a 11.95-kbp *Pst*I-restriction fragment [23]. Plasmid DNA, prepared by alkaline extraction procedures [24], was treated with a restriction endonuclease, *Pst*I. After agarose gel electrophoresis using Sea Plaque agarose (FMC BioProducts), the chromosomal DNA was recovered from the gel and then cleaved with several restriction endonucleases, followed by ligation into M13Δ*lac*183 with T_4 DNA ligase. Transfection to $CaCl_2$-treated *E. coli* RL108 cells was carried out as in [24] and positive clones were isolated from clear plaques. Two of them, designated *sitA* and *sitB*, were used for further studies. By DNA-DNA hybridization, both *sitA* and *sitB* were demonstrated to be located in the *E. coli* chromosome-borne segment of pBS4 (not shown).

(A)

sitA

CTTGTTAACCCTGAACTGGATTTAGTCCGCCAG 33

sitB

CGCACGTTTAAACAAAAGAATATTTTCTCCTGGACGATCGCTTTAGGTAA 50

TCTCGCCCTCGTAACGCCGTGCCCACTCCGCCATGTAATCGATATAGTCG 100

GCAGTAAAAGCCACTTCGACTTCAGCCAG 129

(B)

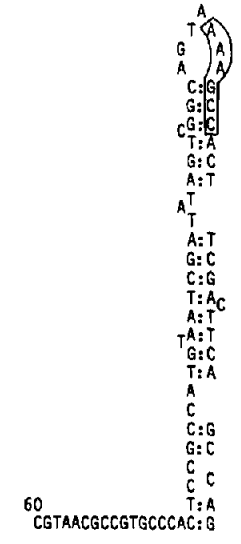

Fig.1. (A) Nucleotide sequences of the two *sit* signals. Sequences homologous to the consensus sequence, 5′-GAAGCGG-3′, and a novel conserved nucleotide sequence between the two *sit* signals are underlined with thin and thick lines, respectively. (B) Possible stem-loop structure of *sitB* ($\Delta G = -16.5$ kcal/mol, according to [29]). The number represented on the 5′-end of the structure corresponds to the nucleotide number shown in (A). The nucleotide sequence conserved in *sitB* and phage G4 origin of complementary DNA strand synthesis is boxed.

Nucleotide sequences were determined by the chain termination method [25]. As shown in fig.1A, *sitA* is composed of 33 nt and *sitB* of 129 nt. Sequence homologies, the variations of the 5′-GAAGCGG-3′ sequence, were found (fig.1A). This consensus sequence, 5′-GAAGCGG-3′, has been observed in the n′ protein recognition sites in φX174 DNA [26] and in the n′-dependent initiation signals found in plasmids [1,3]. Judging from the nucleotide sequences, a stable stem-loop structure was predicted in *sitB* (fig.1B).

Both *sitA* and *sitB* stimulated the growth of re-

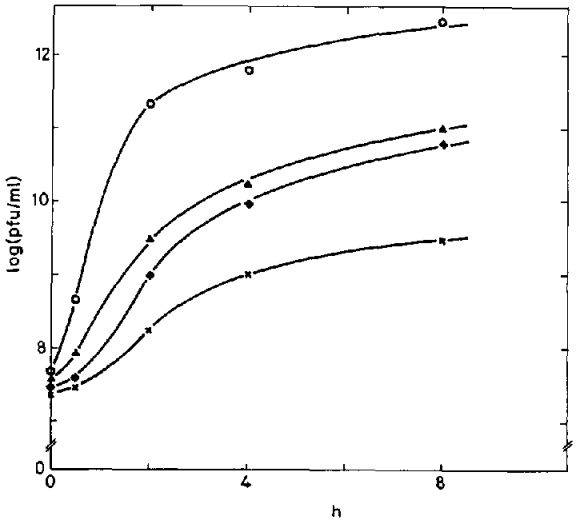

Fig.2. Infection growth curves. The phage yields are represented as a function of time (in h) after infection. (○) M13; (×) M13Δ*lac*183; (▲) M13Δ*lac*183/*sitA*; (◆) M13Δ*lac*183/*sitB*.

Table 1

Estimation of functional activities of the *sit* signals in vivo

	Infection growth assay[a]		Accumulation of RFI[b]	
	0 h	8 h	5 min	60 min
M13Δ*lac*183	1	1.5×10^2	1	4.8
M13Δ*lac*183/*sitA*	1	1.6×10^3	1	13.1
M13Δ*lac*183/*sitB*	1	1.3×10^3	1	11.2
M13	1	6.6×10^4	1	43.6

[a] Relative phage growth represented by the ratio to the value at 0 h

[b] Relative accumulation of RFI DNA represented by the ratio to the value at 5 min

combinant phages at least 10-fold higher than that of M13Δ*lac*183 (fig.2 and table 1). To monitor the effects of these *ssi* signals on DNA replication in vivo more directly, the accumulation of RFI DNA of the recombinant phages was measured. As summarized in table 1, both *sitA* and *sitB* promoted the accumulation of RFI DNA. These results are consistent with those of the infection growth assay. Both results indicate that both *sitA* and *sitB*, isolated from the *terC* region, can function as *ssi* signals on ss-DNA templates.

4. DISCUSSION

Two *ssi* signals, *sitA* and *sitB*, are isolated and mapped on the *E. coli* chromosome-borne segment of pBS4 (around 31.2 min). The origin-probe vector efficiently detects *ssi* signals from the chromosomal DNA. Isolation of the *sit* signals, *ssi* signals from the *terC* region, is based on clear-plaque formation. This plaque morphology is consistent with the results obtained both from the infection growth assay and from the synthesis of RFI DNA in vivo (table 1).

The nucleotide sequences of two *sit* signals (fig.1A) have some interesting features. First, homologies to the consensus sequence of the n′ protein recognition site [26], 5′-GAAGCGG-3′, are observed within both of the *sit* signal sequences. This type of sequence homology has been found in the initiation signals from plasmids [1] and the *E. coli* chromosome [3]. These observations suggest that these *sit* signals can function as n′-dependent primosome assembly sites, as do other *ssi* signals from plasmids and phages. A novel homology is found within *sitA* and *sitB* (fig.1A). The consensus sequence shows 16 residues out of 22. This high degree of homology, over 70%, may be significant for the functional activity.

A stem-loop structure has been predicted in many *ssi* signals [1,3,8,18,19]. Judging from the nucleotide sequence of *sitB*, a stable stem-loop structure is possible as shown in fig.1B. The secondary structure might be related to the functional activity of *sitB*. In the case of *sitA*, however, no stable stem-loop structure can be predicted, though the functional activities of *sitA* and *sitB* are almost the same (fig.2 and table 1). This can be explained in two alternative ways: (i) the *sitA* sequence is part of the fully functional *ssi* signal containing a stable stem-loop structure; (ii) the stem-loop structure is not always essential for an *ssi* signal. The latter might be the case. It is of interest to note that the loop in the stem-loop of *sitB* shown in fig.1B has the same sequence, 5′-AAAGCC-3′, as that found in the loop of stem-loop I of the phage G4 origin of complementary DNA strand synthesis. However, in loop I of phage G4 origin, base substitution and insertion do not affect the functional activity seriously (unpublished). This indicates that all the specific bases in this sequence are not essential for the functional activity.

Existence of *ssi* signals in the *terC* region implies possible functions for these signals. They might be the DNA sequences that direct the assembly of a primosome, which may function for, most probably, the synthesis of Okazaki fragments [27] or for the reinitiation in the ss-DNA region when replication is terminated prematurely. A more interesting possibility is that *sit* signals might be related to DNA replication in the absence of de novo protein synthesis. The same possibility has been suggested [2] for the function of n′ protein recognition sequences located within and clockwise to the *asnA* gene. As suggested for these sequences, the *sit* signals may be involved in the stable DNA replication, because the fixed initiation sites have been mapped in the *terC* region [28] close to the *sit* signals. To identify the relevant DNA-initiating machinery and the biological function of the *sit* signals, protein requirements for these signals to function remain to be studied.

Acknowledgements: We thank Dr Nobuo Nomura of Nippon Veterinary and Zootechnical College, Tokyo for providing the

phage vectors M13Δ*lac*183 and *E. coli* RL108. We are also grateful to Dr Jean-Pierre Bouché of Centre National de la Recherche Scientifique, Toulouse for his gift of plasmid pBS4.

REFERENCES

[1] Van der Ende, A., Teertstra, R., Van der Avoort, H.G.A.M. and Weisbeek, P.J. (1983) Nucleic Acids Res. 11, 4957–4975.
[2] Seufert, W. and Messer, W. (1986) EMBO J. 5, 3401–3406.
[3] Stuitje, A.R., Weisbeek, P.J. and Meijer, M. (1984) Nucleic Acids Res. 12, 3321–3332.
[4] Ray, D.S., Hines, J.C., Kim, M.H., Imber, R. and Nomura, N. (1982) Gene 18, 231–238.
[5] Shlomai, J. and Kornberg, A. (1980) Proc. Natl. Acad. Sci. USA 77, 799–803.
[6] Nomura, N. and Ray, D.S. (1980) Proc. Natl. Acad. Sci. USA 77, 6566–6570.
[7] Nomura, N., Low, R.L. and Ray, D.S. (1982) Proc. Natl. Acad. Sci. USA 79, 3153–3157.
[8] Nomura, N., Low, R.L. and Ray, D.S. (1982) Gene 18, 239–246.
[9] Zipursky, S.L. and Marians, K.J. (1980) Proc. Natl. Acad. Sci. USA 77, 6521–6525.
[10] Zipursky, S.L. and Marians, K.J. (1981) Proc. Natl. Acad. Sci. USA 78, 6111–6115.
[11] Arai, K. and Kornberg, A. (1981) Proc. Natl. Acad. Sci. USA 78, 69–73.
[12] Arai, K., Low, R.L. and Kornberg, A. (1981) Proc. Natl. Acad. Sci. USA 78, 707–711.
[13] Low, R.L., Arai, K. and Kornberg, A. (1981) Proc. Natl. Acad. Sci. USA 78, 1436–1440.
[14] Kim, M.H., Hines, J.C. and Ray, D.S. (1981) Proc. Natl. Acad. Sci. USA 78, 6784–6788.
[15] Lambert, P.F., Waring, D.A., Wells, R.D. and Reznikoff, W.S. (1986) J. Virol. 58, 450–458.
[16] Strathearn, M.D., Low, R.L. and Ray, D.S. (1984) J. Virol. 49, 178–182.
[17] Imber, R., Low, R.L. and Ray, D.S. (1983) Proc. Natl. Acad. Sci. USA 80, 7132–7136.
[18] Bahk, J.D., Sakai, H. and Komano, T. (1988) Gene 65, 93–99.
[19] Honda, Y., Sakai, H. and Komano, T. (1988) Gene 68, 221–228.
[20] Ray, D.S., Cleary, J.M., Hines, J.C., Kim, M.H., Strathearn, M., Kaguni, L.S. and Nomura, N. (1981) in: The Initiation of DNA Replication (Ray, D.S. ed.) pp. 169–193, Academic Press, New York.
[21] Hines, J.C. and Ray, D.S. (1980) Gene 11, 207–218.
[22] Lambert, P.F., Kawashima, E. and Reznikoff, W.S. (1987) Gene 53, 257–264.
[23] Béjar, S. and Bouché, J.P. (1983) J. Bacteriol. 153, 604–609.
[24] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
[25] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[26] Matthes, M., Weisbeek, P.J. and Denhardt, D.T. (1982) J. Virol. 42, 301–305.
[27] Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K. and Sugino, A. (1968) Proc. Natl. Acad. Sci. USA 59, 598–605.
[28] de Massy, B., Fayet, O. and Kogoma, T. (1984) J. Mol. Biol. 178, 227–236.
[29] Tinoco, I., jr, Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40–41.