Archivum Immunologiae et Therapiae Experimentalis, 2002, **50**, 393–398
PL ISSN 0004-069X

Review

Streptomyces and *Escherichia coli*, Model Organisms for the Analysis of the Initiation of Bacterial Chromosome Replication

WALTER MESSER[1]* and JOLANTA ZAKRZEWSKA-CZERWIŃSKA[2]

[1] Max-Planck-Institut für molekulare Genetik, D-14195 Berlin, Germany, [2] Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland

Abstract. *Streptomyces coelicolor* A3(2) and *Escherichia coli* are quite different in their life-style and the structures of their genomes. *Streptomyces* exhibit complex multicellular development including formation of multigenomic hyphae during growth. These organisms possess a large linear (8.7 Mb) and GC-rich (~72%) chromosome. The genome sequence of *S. coelicolor* has just been completed. The difference between *E. coli* and *Streptomyces* making them an excellent model organisms for a comparison of their replication modes. In this review, we compare initiation of chromosome replication in both organisms. Their replication origins are different in size, but both have DnaA boxes – a binding motifs for initiator DnaA protein. The two DnaA proteins have practically the same biochemical properties. Many aspects of the control of initiation seem to be similar. A comparison of the two systems thus allows us to define those aspects of replication initiation that are universally used in the eubacterial kingdom.

Key words: DnaA; *E. coli*; *oriC*; *Streptomyces coelicolor*.

Introduction

Streptomyces coelicolor A3(2) and *Escherichia coli* are quite different in their life-styles and the structures of their genomes. The differences and the fact that both genomes are completely sequenced make them ideal for a comparison of their replication modes.

E. coli is a Gram-negative, enterobacterial symbiont. Some strains can be pathogenic. It is the principal "guinea-pig" of microbiologists and has been the subject of intensive research for more than 50 years. It has a circular chromosome, like most bacteria, with an average G+C content and a size of 4.6 Mb. Replication proceeds bidirectionally from a single and unique replication origin, *oriC*, to the terminus region of approximately 200 kb on the opposite side of the circle. This region is bordered by unidirectional protein binding sites, *ter*, which allow replication forks to enter the region but not to leave it. At the end of replication, two intertwined circles result that have to be resolved before segregation.

Streptomyces are free-living mycelial Gram-positive soil bacteria known for their ability to differentiate. They produce a large number of valuable secondary metabolites, including >500 different antibiotics. *S. coelicolor* A3(2) possesses a large (8.67 Mb) linear

* Correspondence to: Prof. Walter Messer, Max-Planck-Institut für molekulare Genetik, Ihnestrasse 73, D-14195 Berlin, Germany, tel.: +49 30 81 73 788, fax: +49 30 8413 1385, e-mail: messer@molgen.mpg.de

and GC-rich (~72%) chromosome[17]. A single replication origin is located approximately in the middle of the chromosome. Replication proceeds bidirectionally towards the chromosomal ends. These form telomer-like structures with multiple palindromic sequences and a terminal protein attached to the 5'-ends[8]. Replication stops within the telomers, about 150–200 bp before the actual end. The recessed 3'-end is then filled using the terminal protein as a primer[1]. Most housekeeping genes of *Streptomyces* are located in a central core region around *oriC*. This counteracts a high genetic instability due to high recombination frequencies at the chromosome ends. The function of *oriC* is independent of the actual genome structure. The chromosomal origin can be cloned as a circular, autonomously replicating minichromosome[36].

Replication Origins

Replication origins of different bacteria are different in size. They all contain short, conserved sequences which are essential for the *oriC* function: nonpalindromic 9 bp sequences, called DnaA boxes, and AT-rich regions[23, 35]. Spacer regions that vary in nucleotide composition and length separate these conserved sequences. The initiator protein DnaA binds specifically to the DnaA boxes. DnaA boxes have the consensus sequence 5'-TT$^{A}/_{T}$TNCACA-3' or 5'-TTGTNCACA-3'. We will discuss later the different DnaA boxes.

The replication origins of *E. coli* and *Streptomyces* are especially different (Fig. 1). *E. coli oriC* has a size of 260 bp and contains 5 DnaA boxes. An AT-rich region in the left part of *oriC*, 56 bp total, is organized as an AT-cluster and 3 homologous 13-mers. The structure is quite rigid, distances between these sites must not be modified, whereas point mutations, e.g. in the DnaA boxes in the right half of *oriC*, are permissible[15].

Streptomyces oriC has 19 DnaA boxes in 600 bp[11, 36]. All of them are essential, and mutations that change the distance or the sequence are not allowed. The DnaA boxes are grouped into two clusters separated by a spacer (approximately 130 bp). The consensus sequence of the *Streptomyces* DnaA box in *oriC* is 5'--TTGTCCACA-3', which differs at the third position from the *E. coli* DnaA box. In contrast to *E. coli*, the *Streptomyces oriC* region does not contain typical AT--rich repeats adjacent to clusters of DnaA boxes. However, the *Streptomyces oriC* contains 5 short AT-rich sequences that are distributed between DnaA boxes[11].

Steps in the Initiation at *E. coli oriC*

Most of our knowledge of the sequence of events during initiation comes from *E. coli*. The establishment of the *in vitro* initiation system by the Kornberg laboratory opened the way to identifying the factors involved in the process[14]. No comparable system is available for other bacteria, except for *Bacillus subtilis* to some extent[25].

The *in vitro* initiation system requires a supercoiled *oriC* plasmid as a template. The initiator DnaA binds to its 5 binding sites, resulting in the "initial complex". Only the ATP-complexed form of DnaA is active in initiation since it recognizes an additional sequence, the 6-mer ATP-DnaA box with a consensus sequence 5'-AGATCT or a close match. ATP-DnaA boxes are especially abundant in the AT-rich region, resulting in a complex in which this region is unwound. The single--stranded DNA is stabilized by ATP-DnaA[30]. The structure is called the "open complex"[14]. A threshold level of ATP-DnaA is required for unwinding, making this the rate-limiting process in initiation.

The bubble is the entry site for the replicative helicase DnaB. Two double-hexamers of DnaB and the helicase loader protein DnaC, one double-hexamer for each replication direction, are positioned by DnaA into the loop[4, 5]. DnaC leaves the complex, accompanied by ATP hydrolysis. This activates the helicase activity of DnaB[33]. DnaB helicase expands the single-stranded region for the entry of DnaG primase and other proteins required to form a replication fork. Primase synthesizes two leading strand primers that are extended by DNA polymerase III holoenzyme. The β-subunit of polymerase III, the sliding clamp, activates the intrinsic ATPase activity of DnaA[12]. Inactivation of DnaA by ATP hydrolysis prevents further initiations. The *E. coli* initiation cycle is schematically shown in Fig. 2.

A cell-free system that replicates *oriC* plasmids of

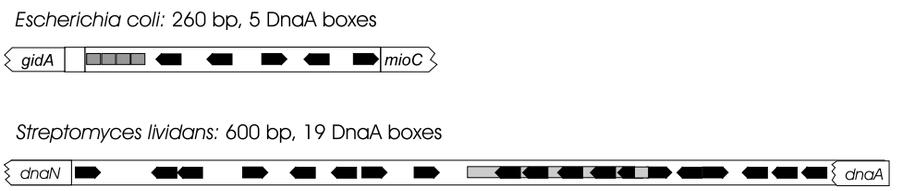

Fig. 1. Schematic view of *oriC* from *E. coli* and *S. lividans*. DnaA binding sites are black pointed boxes, AT-rich regions are shown in grey

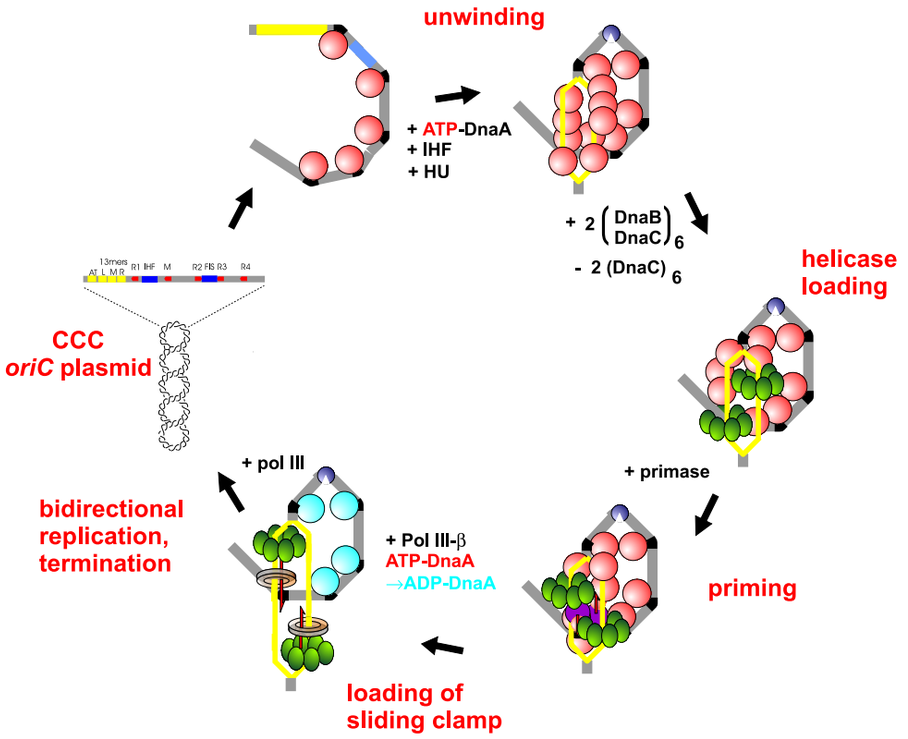

Fig. 2. Steps in the initiation at *E. coli oriC*

Streptomyces has not been established so far. Nevertheless, much circumstantial evidence suggests that the sequence of events is very similar to that in *E. coli*. Especially the regulation and the biochemical properties of the initiator DnaA are very similar in the two organisms (see below). This also holds for the function of the replication origins as sites of recognition and unwinding by DnaA protein.

The unwinding of an AT-rich region and the subsequent loading of the replicative helicase are the decisive events in the cell cycles of all organisms, and they are basically similar. In eukaryotes, the equivalent of DnaA is the 6-subunit origin-recognition complex (ORC), which is also responsible for the initial unwinding. Like DnaA, ORC binds to double-stranded and single-stranded DNA, and the change in binding specificities is associated with an ATP/ADP switch[16]. The role of the replicative helicase is fulfilled by the mini-chromosome maintenance (MCM) proteins, and the Cdc6 and Cdt1 proteins correspond to the helicase-loader DnaC.

Biochemical Properties of DnaA Proteins

DnaA proteins of all bacterial species are highly homologous. On the basis of sequence homology they were subdivided into 4 domains, which were later shown to correspond to functional domains[6, 22, 28, 32]. The N-terminal domain 1 (~80 amino acids) mediates protein-protein interactions, DnaA oligomerization and interaction with DnaB helicase[29]. Oligomerization via domain 1 has been shown using several techniques, both for *E. coli*[34] and for *Streptomyces*[9].

In contrast to the highly conserved domains 3 and 4, domain 2 appears to be less evolutionarily conserved. The length of the domain 2 varies from 48 amino acids in *Helicobacter pylori*[37] to 247 amino acids in *S. coelicolor* A3(2). In *E. coli* it is 77 amino acids long, and can be deleted without loss of function. Therefore it is likely to form a flexible linker. In different *Streptomyces sp.*, domain 2 is between 216 and 247 amino acids long. Since this much longer domain has been conserved in evolution, it may have an additional function that has so far not been discovered.

Domain 3 (244 amino acids) is very well conserved between different DnaA proteins. It contains a nucleotide-binding site of the Walker A and B type and a low-activity ATPase function. Domain 3 contains, in addition, a second oligomerization site. This has been shown for *Streptomyces* by kinetic analysis[20]and by gel retardation[9], and for *E. coli* by binding kinetics[21].

Domain 4 (94 amino acids in *E. coli* and 131 in

Streptomyces) is also highly homologous among DnaA proteins. It contains the DNA binding site. Isolated domain 4 from *E. coli*[26] or from *Streptomyces*[18, 19] interacts specifically with DnaA boxes. This domain contains an unusual binding motif, consisting of two α-helices and a basic loop between them. Mutation of basic amino acids in the loop impairs DNA binding[2, 18, 19].

The binding sites for DnaA protein are 5'-TT^{A}/$_{T}$TNCACA-3' for *E. coli* and 5'-TTGTCCACA-3' or 5'-TTATCCACA for *Streptomyces.* All these boxes represent high-affinity sites, e.g. are "strong" boxes. Both proteins can bind as monomers to the strong boxes[21]. *Streptomyces* DnaA can also bind as a dimer to a single strong box[9]. DnaA boxes with one or two mismatches to the consensus sequence of strong DnaA boxes bind DnaA protein as well; however, cooperativity between DnaA molecules bound to adjacent weak boxes is required[20, 21, 31].

Control of Initiation

Initiation control has to ensure two conditions: initiation must occur at the correct time in the cell cycle, and any one origin must initiate once and only once per generation. Recent studies revealed that DnaA protein is the best and only candidate molecule controlling replication initiation: replication is initiated when the concentration of free DnaA exceeds a characteristic threshold level. DnaA protein is autoregulated, both in *E. coli*[24] and in *Streptomyces*[10]. This seems to be contradictory to its function as a regulatory molecule. However, due to the high affinity of DnaA for its targets, free DnaA is not found in cells except when the number of DnaA molecules exceeds the number of DnaA boxes[7]. The number of DnaA boxes increases during the cell cycle due to replication. This means that the number and location of DnaA boxes on the chromosome, outside of *oriC*, is important, and the actual regulatory circuit is the dynamic ratio of ATP-DnaA over DnaA boxes.

In the *E. coli* genome, strong DnaA boxes are distributed rather uniformly over the chromosome. This is also true for 5 regions with an especially high DnaA affinity[27]. However, one of these high-affinity sites located close to *oriC*, *datA*, can bind an exceptionally high number of DnaA molecules, about 8 times more than the *oriC* region. It serves as a sink for DnaA, and the number of *datA* copies is crucial for replication initiation[13].

In *S. coelicolor* A3(2), all but 3 of the 51 strong DnaA boxes are located in the core region of the chromosome and they particularly are clustered around the *oriC* region (Fig. 3). Although it is not known whether there is a *datA*-like site in *Streptomyces*, the cluster of high-affinity sites in the center of the chromosome might serve such a regulatory role.

Titration of DnaA by high-affinity sites is also a mechanism to prevent premature initiation of origins that have already fired. The number of such sites increases immediately after initiation due to replication and titrates DnaA, which is then no longer available for initiation.

A second mechanism is the inactivitation of ATP--DnaA by ATP hydrolysis at the end of the initiation cycle when the sliding clamp is loaded to the unwound *oriC*, as described above. By analogy, this is also a likely mechanism for *Streptomyces*, since also here ATP-DnaA is the active form for initiation, and *Streptomyces* DnaA has a similar ATPase activity to *E. coli* DnaA[19].

E. coli has an additional safeguard against reinitiation not available to *Streptomyces. E. coli oriC* contains a high number of GATC sequences, recognition sites for the Dam methyltransferase. Newly replicated *oriC* stays in the hemi-methylated form for about 1/3 of the

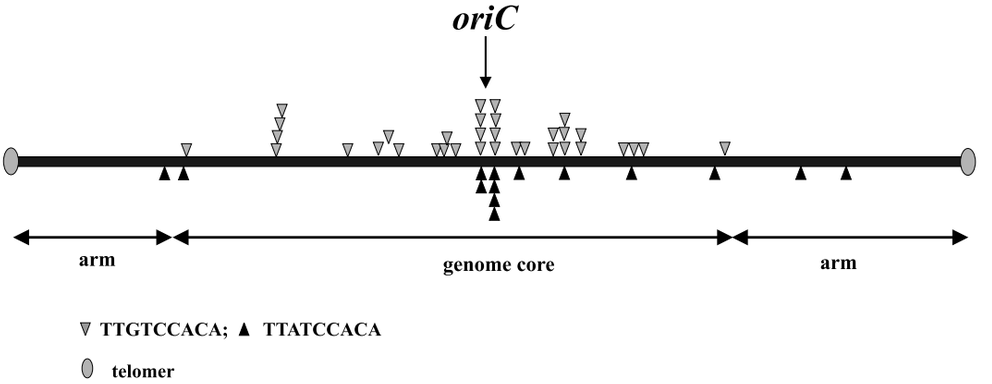

Fig. 3. Distribution of "strong" DnaA boxes in the *Streptomyces coelicolor* A3(2) chromosome

cell cycle before being fully methylated by Dam, while the other GATC sites located outside of the *oriC* region are already remethylated within 2 min. The result is that *E. coli* can tell a new origin from an old one. The new, hemi-methylated one is exempt from initiation due to sequestering to the membrane[3]. Since the *oriC* region is subject to sequestration, the control of initiation in *E. coli* does not need to be very stringent. Consequently, minichromosomes have a copy number of around 10 per chromosome equivalent. *Streptomyces* and most other bacteria do not have a comparable methylation system. Therefore *Streptomyces* must control the initiation much more tightly. Consequently, minichromosomes have a copy number of only around 1 per chromosome[36].

Conclusions

Although *Streptomyces* and *E. coli* are very different in their cell physiology, many aspects of initiation are remarkably similar. Their replication origins are different in size, but both have DnaA boxes, and the mechanism of initiation is very similar. The two DnaA proteins have practically the same biochemical properties. Many aspects of the control of initiation seem to be similar. A comparison of the two systems thus allows us to define those aspects of replication initiation that are universally used in the eubacterial kingdom.

Acknowledgment. This work was supported by the following grants: SFB344/A9, Me659/6-1, 436 POL 113/82/0 of the Deutsche Forschungsgemeinschaft, Alexander von Humboldt Fellowship IV-POL 1063505 STP, Fonds der chemischen Industrie, Germany, and the State Committee for Scientific Research (KBN), 3 P04A 079 22, Poland.

References

1. BAO L. and COHEN S. N. (2001): Terminal proteins essential for the replication of linear plasmids and chromosomes in *Streptomyces.* Genes Dev., **15**, 1518–1527.
2. BLAESING F., WEIGEL C. and MESSER W. (2000): Analysis of the DNA binding domain of *Escherichia coli* DnaA protein. Mol. Microbiol., **36**, 557–569.
3. CAMPBELL J. L. and KLECKNER N. (1990): *E. coli oriC* and the *dnaA* gene promoter are sequestered from *dam* methyltransferase following the passage of the chromosomal replication fork. Cell, **62**, 967–979.
4. CARR K. M. and KAGUNI J. M. (2001): Stoichiometry of DnaA and DnaB protein in initiation at the *Escherichia coli* chromosomal origin. J. Biol. Chem., **276**, 44919–44925.
5. FANG L. H., DAVEY M. J. and O'DONNELL M. (1999): Replisome assembly at *oriC*, the replication origin of *E. coli*, reveals an explanation for initiation sites outside an origin. Mol. Cell, **4**, 541–553.
6. FUJITA M. Q., YOSHIKAWA H. and OGASAWARA N. (1990): Structure of the *dnaA* region of *Micrococcus luteus*: conservation and variations among eubacteria. Gene, **93**, 73–78.
7. HANSEN F. G., CHRISTENSEN B. B. and ATLUNG T. (1991): The initiator titration model: computer simulation of chromosome and minichromosome control. Res. Microbiol., **142**, 161–167.
8. HUANG C.-H., LIN Y.-S., YANG Y.-L., HUANG S. and CHEN C. (1998): The telomeres of *Streptomyces* chromosomes contain conserved palindromic sequences with potential to form complex secondary structures. Mol. Microbiol., **28**, 905–916.
9. JAKIMOWICZ D., MAJKA J., KONOPA G., WEGRZYN G., MESSER W., SCHREMPF H. and ZAKRZEWSKA-CZERWINSKA J. (2000): Architecture of the *Streptomyces lividans* DnaA protein-replication origin complexes. J. Mol. Biol., **298**, 351–364.
10. JAKIMOWICZ D., MAJKA J., LIS B., KONOPA G., WEGRZYN G., MESSER W., SCHREMPF H. and ZAKRZEWSKA-CZERWINSKA J. (2000): Structure and regulation of the *dnaA* promoter region in three *Streptomyces* species. Mol. Gen. Genet., **262**, 1093–1102.
11. JAKIMOWICZ D., MAJKA J., MESSER W., SPECK C., FERNANDEZ M., MARTIN M. C., SANCHEZ J., SCHAUWECKER F., KELLER U., SCHREMPF H. and ZAKRZEWSKA-CZERWINSKA J. (1998): Structural elements of the *Streptomyces oriC* region and their interactions with the DnaA protein. Microbiology, **144**, 1281–1290.
12. KATAYAMA T., KUBOTA T., KUROKAWA K., CROOKE E. and SEKIMIZU K. (1998): The initiator function of DnaA protein is negatively regulated by the sliding clamp of the *E. coli* chromosomal replicase. Cell, **94**, 61–71.
13. KITAGAWA R., OZAKI T., MORIYA S. and OGAWA T. (1998): Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for *Escherichia coli* DnaA protein. Genes Dev., **12**, 3032–3043.
14. KORNBERG A. and BAKER T. A. (1992): DNA replication. W. H. Freeman and Company, New York.
15. LANGER U., RICHTER S., ROTH A., WEIGEL C. and MESSER W. (1996): A comprehensive set of DnaA box mutations in the replication origin, *oriC*, of *Escherichia coli*. Mol. Microbiol., **21**, 301–311.
16. LEE D. G. and BELL S. P. (2000): ATPase switches controlling DNA replication initiation. Curr. Opin. Cell Biol., **12**, 280–285.
17. LIN Y. S., KIESER H. M., HOPWOOD D. A. and CHEN C. W. (1993): The chromosomal DNA of *Streptomyces lividans* 66 is linear. Mol. Microbiol., **10**, 923–933.
18. MAJKA J., JAKIMOWICZ D., MESSER W., SCHREMPF H., LISOWSKI M. and ZAKRZEWSKA-CZERWINSKA J. (1999): Interactions of the *Streptomyces lividans* initiator protein DnaA with its target. Eur. J. Biochem., **260**, 325–335.
19. MAJKA J., MESSER W., SCHREMPF H. and ZAKRZEWSKA-CZERWINSKA J. (1997): Purification and characterization of the *Streptomyces lividans* initiator protein DnaA. J. Bacteriol., **179**, 2426–2432.
20. MAJKA J., ZAKRZEWSKA-CZERWINSKA J. and MESSER W. (2001): Sequence recognition, cooperative interaction and dimerization of the initiator protein DnaA of *Streptomyces.* J. Biol. Chem., **276**, 6243–6252.
21. MESSER W., BLAESING F., JAKIMOWICZ D., MAJKA J., NARDMANN J., SCHAPER S., SEITZ H., SPECK C., WEIGEL C., WEGRZYN G., WELZECK M. and ZAKRZEWSKA-CZERWINSKA J. (2001): Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading. Biochimie, **83**, 1–9.
22. MESSER W., BLAESING F., MAJKA J., NARDMANN J., SCHAPER S., SCHMIDT A., SEITZ H., SPECK C., TÜNGLER D., WEGRZYN G.,

WEIGEL C., WELZECK M. and ZAKRZEWSKA-CZERWINSKA J. (1999): Functional domains of DnaA proteins. Biochimie, **81**, 819–825.

23. MESSER W. and WEIGEL C. (1996): Initation of chromosome replication. In NEIDHARDT F. C., CURTISS R. III, INGRAHAM J., LIN E. C. C., LOW K. B., MAGASANIK B., REZNIKOFF W. S., RILEY M., SCHAECHTER M. and UMBARGER H. E. (eds.): *Escherichia coli* and *Salmonella*, cellular and molecular biology. ASM, Washington, D. C., 1579–1601.
24. MESSER W. and WEIGEL C. (1997): DnaA initiator – also a transcription factor. Mol. Microbiol., **24**, 1–6.
25. MORIYA S., FIRSHEIN W., YOSHIKAWA H. and OGASAWARA N. (1994): Replication of a *Bacillus subtilis oriC* plasmid *in vitro*. Mol. Microbiol., **12**, 469–478.
26. ROTH A. and MESSER W. (1995): The DNA binding domain of the initiator protein DnaA. EMBO J., **14**, 2106–2111.
27. ROTH A. and MESSER W. (1998): High-affinity binding sites for the initiator protein DnaA on the chromosome of *Escherichia coli*. Mol. Microbiol., **28**, 395–401.
28. SCHAPER S. and MESSER W. (1997): Prediction of the structure of the replication initiator protein DnaA. Proteins: Structure Function Genetics, **28**, 1–9.
29. SEITZ H., WEIGEL C. and MESSER W. (2000): The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*. Mol. Microbiol., **37**, 1270–1279.
30. SPECK C. and MESSER W. (2001): Mechanism of origin unwinding: sequential binding of DnaA initiator protein to double-stranded and single-stranded DNA in the AT-rich region of the replication origin. EMBO J., **20**, 1469–1476.
31. SPECK C., WEIGEL C. and MESSER W. (1999): ATP- and ADP-DnaA protein, a molecular switch in gene regulation. EMBO J., **18**, 6169–6176.
32. SUTTON M. D. and KAGUNI J. M. (1997): The *Escherichia coli dnaA* gene: four functional domains. J. Mol. Biol., **274**, 546–561.
33. WAHLE E., LASKEN R. S. and KORNBERG A. (1989): The dnaB-dnaC replication protein complex of *Escherichia coli*. II. Role of the complex in mobilizing dnaB functions. J. Biol. Chem., **264**, 2469–2475.
34. WEIGEL C., SCHMIDT A., SEITZ H., TUENGLER D., WELZECK M. and MESSER W. (1999): The N-terminus promotes oligomerisation of the *Escherichia coli* initiator protein DnaA. Mol. Microbiol., **34**, 53–66.
35. YOSHIKAWA H. and OGASAWARA N. (1991): Structure and function of DnaA and the DnaA-box in eubacteria: evolutionary relationships of bacterial replication origins. Mol. Microbiol., **5**, 2589–2597.
36. ZAKRZEWSKA-CZERWINSKA J. and SCHREMPF H. (1992): Characterization of an autonomously replicating region from the *Streptomyces lividans* chromosome. J. Bacteriol., **174**, 2688–2693.
37. ZAWILAK A., CEBRAT S., MACKIEWICZ P., KROL-HULEWICZ A., JAKIMOWICZ D., MESSER W., GOSCINIAK G. and ZAKRZEWSKA-CZERWINSKA J. (2001): Identification of a putative chromosomal replication origin from *Helicobacter pylori* and its interaction with the initiator protein DnaA. Nucleic Acids Res., **29**, 2251–2259.

Received in June 2002
Accepted in July 2002