Volume 283, number 2, 263–266 FEBS 09772 June 1991

ADONIS 001457939100481?

Nucleotide sequence of the replication region of the marine *Rhodobacter* plasmid pRD31

Tadashi Matsunaga, Hideaki Mihashita, Masato Miyake and J. Grant Burgess

Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan

Received 4 April 1991

The minimum region required for replication of the marine *Rhodobacter* endogenous plasmid pRD31 has been sequenced. This region is located on a 1367-bp *Hin*cII–*Pst*I restriction fragment and is 62% rich in G·C base pairs. A region homologous to the bacteriophage P1 *repA* promoter, which overlaps two inverted repeats, has been identified. Plasmids with mutations in this 1367-bp region could not replicate in marine *Rhodobacter* hosts. This is the first identified replication origin of a photosynthetic bacterium.

Marine *Rhodobacter*; Plasmid; Replication region; Sequence

1. INTRODUCTION

Although endogenous plasmids have been discovered in many species of photosynthetic bacteria their precise function remains obscure [1–7]. In addition, until now, no sequence infomation has been available to improve this situation. The mechanisms of replication and segregation of these plasmids are unknown, despite evidence that perturbations in the plasmid profile can remove photosynthetic ability [3,5,8]. The characterization and further analysis of plasmids from marine photosynthetic bacteria is of particular interest since these species have been found to contain numerous small plasmids less than 15 kb in size [6,7] whereas all of the plasmids detected so far in freshwater photosynthetic bacteria are greater than 40 kb in size [1–5,9].

Plasmid pRD31 is a 3.1-kb endogenous plasmid which was isolated from a marine photosynthetic bacterium *Rhodobacter* sp. NKPB 043402 [10]. The replication region of this plasmid was previously localised to a 1.36-kb *Hin*cII–*Pst*I restriction fragment and was used to construct a series of expression vectors which were stably maintained in several species of marine photosynthetic bacteria [10]. We describe here DNA sequencing of this minimal replicon using series of the *Hin*cII–*Pst*I fragment cloned into the phagemids pUC118 and pUC119 [11]. The sequence shows characteristics of an origin of DNA replication and its organization has been compared with replication origins from several other plasmids. This is the first report of an origin of replication from a photosynthetic bacterium at the molecular level and is significant for the study of plasmid replication in general and for an understanding of the plasmid biology of photosynthetic bacteria in particular.

Correspondence address: T. Matsunaga, Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan. Fax: (81) (432) 85 7713

2. MATERIALS AND METHODS

Marine *Rhodobacter* strain NKPB043402 was grown photosynthetically in RCBV medium [12] supplemented with 3% NaCl. The endogenous plasmid pRD31 [10] was purified using the method of Van den Hondel et al. [13]. Restriction endonucleases were purchased from Toyobo, Tokyo, Japan. Molecular biological procedures were carried out according to Sambrook et al. [14]. Inserts for sequencing were cloned into the phagemids pUC118 and pUC119 [11], which were purchased from Takara Shuzo, Kyoto, Japan. A deletion series for sequencing was prepared using a Takara Kilo Sequence Deletion kit, Takara Shuzo, Kyoto, Japan. This kit uses an exonuclease/mung bean nuclease protocol [15]. Sequencing reactions were carried out using the Sequenase version 2.0 sequencing kit (United States Biochemical, Cleveland, Ohio) according to the manufacturer's instructions [14]. The resulting sequence information was analyzed using DNASIS Software, Hitachi Software Engineering, Tokyo, Japan.

3. RESULTS

The nucleotide sequence of the 1367-bp *Hin*cII–*Pst*I fragment containing the replication region was determined (Fig. 1). This sequence is 62% G-C rich. Important features include a 10-bp sequence CGTCAGGGGT which forms 2 direct repeats starting at bases 4 and 26. The GGGG motif is then repeated twice starting at nucleotides 43 and 55. The region between bases 181 and 232 is 77% A-T rich. The following 60 bp (nucleotides 233 to 292) is 45% A-T rich while the preceding 60 bp (nucleotides 121 to 180) has an A-T content of only 22% (or 78% G-C). Also present is a region between basees 247 and 282 where four AAAA sequences occur with a spacing of 6–8 bp producing an A+T rich anti-

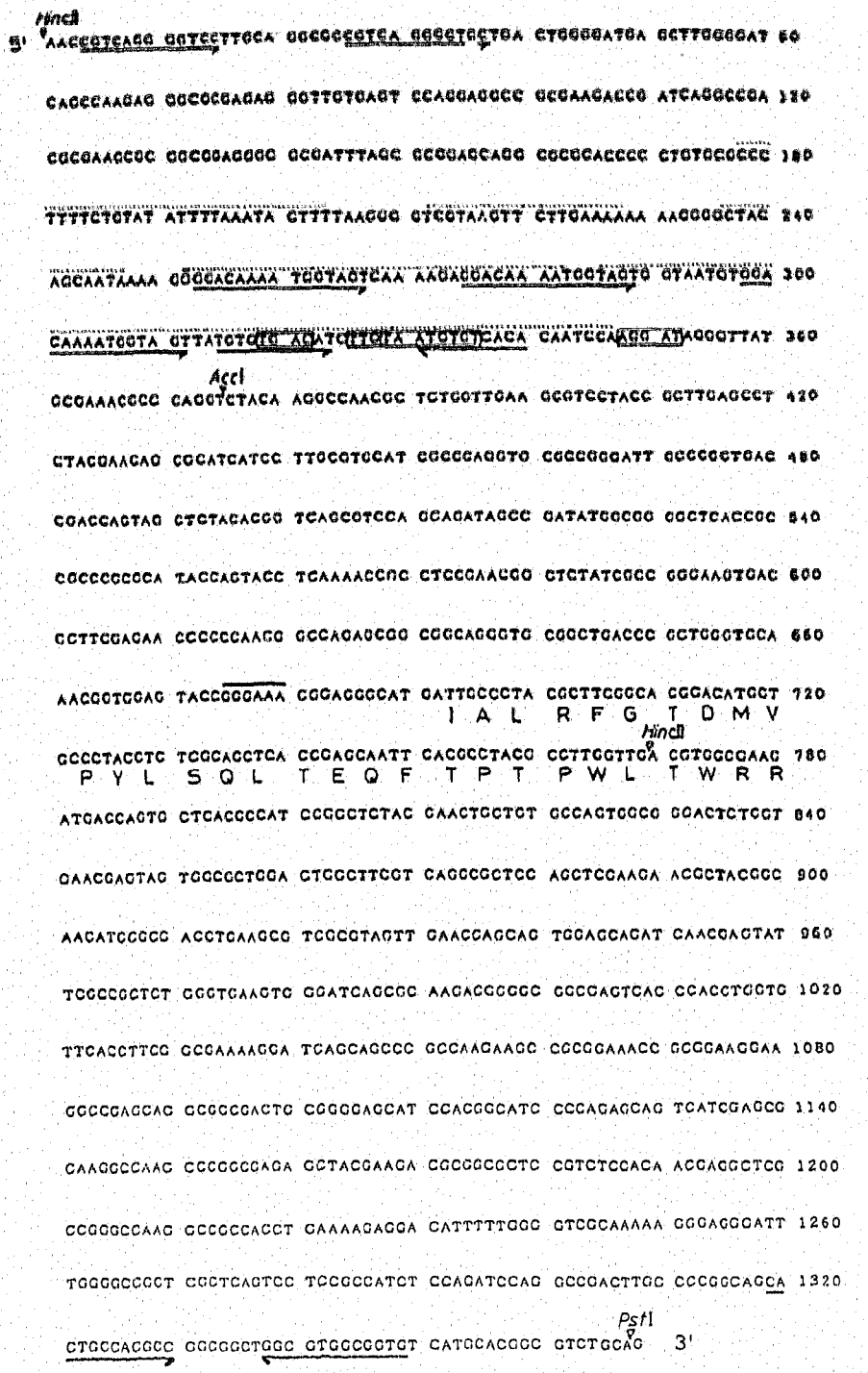

Fig. 1. Nucleotide sequence of the 1367 bp *Hin*cII–*Pst*I restriction fragment of pRD31 containing the replication origin. Inverted and direct repeats are underlined with bold arrows. Regions of A-T rich DNA (>60%) are overlined with a dotted line. Restriction sites are marked with an open arrow. Regions homologous to bacteriophage P1 *rep*A promoter are boxed. The putative ribosome binding site is overlined with a solid bar.

bent DNA structure [16]. On the opposite strand three AAAA sequences occur with a similar periodicity. Eight consecutive A residues occur starting at position 225. Three perfectly conserved 15-bp direct repeats (GGACAAAATGGTAGT) are underlined with an arrow in Fig. 1. At positions 315–324 and 331–340 there is a pair of inverted repeats which are able to form a potential stem and loop structure with a G of -9.3 kcal/mol. This loop has a stem length of 12 bp, a loop size of 4 bases (Fig. 3) and is extremely stable with

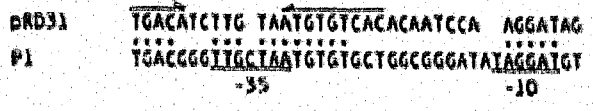

Fig. 2. Homology between the -10 and -35 regions of the P1 *rep*A promoter [21] and the putative pRD31 promoter downstream of the three direct repeats. The inverted repeats capable of forming a hairpin loop are shown by arrows.

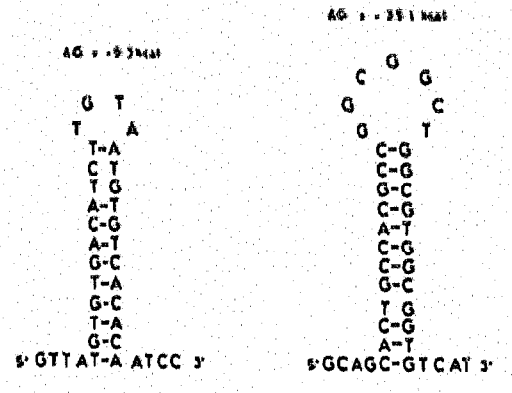

Fig. 3. Possible formation of intrastrand hairpin loop regions beween bases 315-342 and 1319-1349.

respect to its free energy of formation. The hairpin loop passes across the −35 region of a putative promoter at positions 319-352 which shows good homology with the bacteriophage P1 *rep*A promoter [17] as shown in Fig. 2. The sequence was examined for the existence of open reading frames. One open reading frame, ORF30 (beginning at base 599) has a characteristic ribosome binding site and lies downstream of the putative promoter (Fig. 1). However, to determine which open reading frames are actually expressed in vivo, further work will be necessary. ORF30 bore no significant homology to any of the known replication proteins after computer analysis. Additional gene functions required for replication may be supplied *in trans* from the chromosome. An additional hairpin loop structure may also form from the inverted repeats which occur between bases 1319 and 1349 (Figs. 1 and 3).

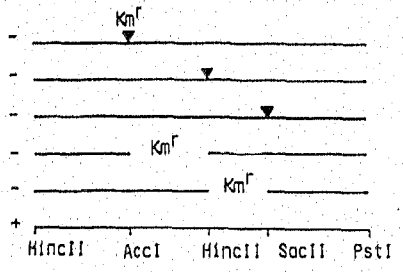

Fig. 4. Deletion and insertion mutagenesis of the 1367-bp *Hin*cII-*Pst*I replication region. For clarity, only the replication region is shown. Insertion of the kanamycin-resistance cassette from *Tn*903 is indicated by an arrow and the letters Km^r. Deleted regions are left as gaps in the solid line. Restriction sites are marked below the diagram. Ability of the resulting constructs to replicate in marine *Rhodobacter* NKPB0021 is indicated by a + for replication and a - for no replication ability.

Insertion mutagenesis was carried out by cloning the kanamycin-resistance gene cartridge from pUC4K into the *Hin*cII-*Pst*I fragment at three different positions. In addition, two restriction fragements *Acc*I-*Hin*cII and *Hin*cII-*Sac*II were deleted from this region (Fig. 4). The ability of these fragments to confer replication ability in the marine strain *Rhodobacter* sp. NKB0021 upon pUC19 was shown using the method previously described [10]. None of the mutagenised plasmids were able to replicate in *Rb.* NKPB0021.

4. DISCUSSION

The nucleotide sequence of a 1367-bp *Hin*cII-*Pst*I restriction fragment from the cryptic 3.1-kb plasmid pRD31 has been determined. This region functions as the pRD31 replication origin as shown previously [10] and confirmed in this work by deletion and insertion mutagenesis. Upstream of the pRD31 origin lie two 12 bp tandem repeats (Fig. 1). These repeats may serve as the binding site for an initiator protein. In addition, there appears to be a *dna*A box at bases 357-365 which has a 7/9 match to the *dna*A box consensus [18]. The G-C content of the replication region (62%) is high when compared to the origins of ColE1 [19], R6K [20] and the *E. coli* chromosomal replication origin [21] but similar to that reported for the broad host range plasmids RK2 [22] and RSF1010 [23] which both have G-C contents of 61%. The grouping of A-T rich and G-C rich regions described in the results section, where the A-T rich region is preceded by a G-C rich region, has also been observed in the replication region of RK2 [20] although in this plasmid the G-C rich region was downstream rather than upstream of the A-T rich sequence. Three A-T rich tandem repeats are present (Fig. 1). It has been proposed that A-T rich direct repeat motifs at origins of replication are successively melted by an initiator protein to form an open complex [24]. Plasmid replication begins by opening of the DNA molecule at this region. Three 13-mer A-T rich tandem repeats also occur at the *E. coli* chromosomal origin [24]. The putative promoter which shares homology with the P1 *rep*A promoter lies between positions 319 and 352 where a hairpin loop structure could be formed from the inverted repeat sequences in this region. This may provide a regulating mechanism, formation of the hairpin structure controlling binding of the RNA polymerase and/or the initiation protein. The potential hairpin loop between bases 1319-1349 may serve as a transcription terminator (Fig. 3). The consecutive AAAA motifs starting and bases 247, 257, 269, 279 and 302 form an antibent DNA structure, which, together with the eight consecutive adenine residues (225-232), resulting in local DNA bending [25], have also been

shown to be conserved in a number of other prokaryotic and eukaryotic replication origins [16]. Eckdahl and Anderson [16] demonstrated that the primary sites of initiator protein binding lie in areas of DNA bending while the direct repeats lie in regions of anti-bent DNA. This is consistent with our model of initiator protein binding adjacent to the direct repeats, nucleotides 253-312 in Fig. 1, and with the general model of plasmid replication originally proposed by Bramhill and Kornberg [24].

We have presented the structure of the replication origin of a cryptic, high copy number 3.1-kb plasmid from a marine photosynthetic bacteria. This DNA region was shown to be required for plasmid replication in the marine *Rhodobacter* sp. NKPB0021 and has also been shown to function in other freshwater purple non-sulphur photosynthetic bacteria (manuscript in preparation). Further studies are required to clarify molecular events which occur during replication and to determine the exact role of the repeated sequences which may define the regions responsible for incompatibility.

Acknowledgement: J.G.B. would like to thank the Japan Society for the Promotion of Science (J.S.P.S.) for financial support.

REFERENCES

[1] Suyama, Y. and Gibson, J. (1966) Biochem. Biophys. Res. Commun. 24, 549-553.
[2] Gibson, K.D. and Niederman, R.A. (1970) Arch. Biochem. Biophys. 141, 694-704.
[3] Saunders, V.A., Saunders, J.R. and Bennet, P.M. (1976) J. Bacteriol. 125, 1180-1187.
[4] Hu, N.T. and Marrs, B.L. (1979) Arch. Microbiol. 121, 61-69.
[5] Kuhl, S.A., Nix, D.W. and Yoch, D.C. (1983) J. Bacteriol. 156, 737-742.
[6] Matsunaga, T., Matsunaga, N., Tsubaki, K. and Tanaka, T. (1986) J. Bacteriol. 168, 460-463.
[7] Harashima, K. (1989) In: Aerobic photosynthetic bacteria, Harashima, K., (Shiba, T. and Murata, N., eds), pp. 149-154, Japan Scientific Societies Press, Tokyo and Springer, Berlin.
[8] Nano, F.E. and Kaplan, S. (1984) J. Bacteriol. 158, 1094-1103.
[9] Fornari, C.S., Watkins, M. and Kaplan, S. (1984) Plasmid 11, 39-47.
[10] Matsunaga, T., Tsubaki, K., Miyashita, H. and Burgess, J.G. (1991) Plasmid, in press.
[11] Viera, J. and Messing, J. (1987) Methods Enzymol. 153, 3-11.
[12] Beatty, J.T. and Gest, H. (1981) Arch. Microbiol. 129, 335-340.
[13] Van den Hondel, C.A.M.J.J., Keegstra, W., Borrias, W.E. and Van Arkel (1979) Plasmid 2, 323-333.
[14] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning, A laboratory manual, 2nd Edition, Cold Spring Harbour Laboratory Press, New York.
[15] Henikoff, S. (1984) Gene 28, 351-359.
[16] Eckdahl, T.T. and Anderson, J.N. (1990) Nucleic Acids Res. 18, 1609-1612.
[17] Abeles, A.L., Reaves, L.D. and Austin, S.J. (1989) J. Bacteriol. 171, 43-52.
[18] Fuller, R.S. and Kornberg, A. (1983) Proc. Natl. Acad. Sci. USA 80, 5817-5821.
[19] Oka, A., Nomura, N., Morita, M., Sugisaki, H., Sugimoto, K. and Takanami, M. (1979) Mol. Gen. Genet. 172, 151-159.
[20] Stalker, D.M., Kolter, R. and Helinski, D.R. (1979) Proc. Natl. Acad. Sci. USA 76, 1150-1154.
[21] Meijer, M., Beck, E., Hansen, F.G., Bergmans, H., Messer, W., von Meyenburg, K. and Schaller, H. (1979) Proc. Natl. Acad. Sci. USA 76, 580-584.
[22] Stalker, D.M., Thomas, C.M. and Helinski, D.R. (1981) Mol. Gen. Genet. 181, 8-12.
[23] Scholz, P., Haring, V., Wittman-Liebold, B., Ashman, K., Bagdasarian, M. and Scherzinger, E. (1989) Gene 75, 271-288.
[24] Bramhill, D. and Kornberg, A. (1988) Cell 54, 915-918.
[25] Koo, H.S., Wu, H.M. and Crothers, D.M. (1986) Nature 320, 501-506.