The Nucleotide Sequence of the Replication Origin $\beta$ of the Plasmid R6K*

Mark Shon‡, Joseph Germino, and Deepak Bastia§

From the Department of Microbiology and Immunology, Duke University Medical Center, Durham, North Carolina 27710

We have identified by molecular cloning a region of 283 base pairs of the HindIII 2 fragment of R6K which corresponds to the region of the replication origin $\beta$. This 283 base-pair DNA fragment, when present contiguously with the structural gene for the replication initiation protein of R6K, encoded in the HindIII 9-15 and part of HindIII 2 restriction fragments, will support the replication of a plasmid chimera containing the pBR322 replicon in a pol A1s host at the nonpermissive temperature. The nucleotide sequence of the region of replication origin $\beta$ has been determined. The nucleotide sequence has some homology with the ori $\gamma$ region of R6K; it has a 15-base-pair homology with the replication origin of Escherichia coli.

The mechanism of initiation of DNA replication, particularly Cairns type replication, mediated by replication initiation protein(s) at specific origins of replication is a major unsolved problem in molecular biology. Any experimental approach to the solution of this problem would require information regarding the structure of the replication origin, structure, and enzymatic role of the initiator protein and the nature of the DNA-protein interactions at the replication origin, between the initiator protein and the origin sequence.

The plasmid R6K ($M_c = 26 \times 10^6$) which carries resistance to ampicillin and streptomycin (1) is an attractive system for the study of initiation and termination of DNA replication. The R6K chromosome codes for its own initiation protein (2) and in a soluble in vitro DNA replication system exogenously added R6K DNA can be replicated and the replication has an obligatory requirement for the initiator protein which is coded by a stretch of sequence contained in the HindIII fragments 9-15 (3). The chromosome of R6K contains at least 3 origins of replication called $\alpha$, $\beta$, and $\gamma$ (4-8). The replication origin $\alpha$ is the most frequently used origin in vivo. In contrast, all three origins are used with equal frequency in vitro (7).

The replication in vitro is bidirectional and terminates at a unique nucleotide sequence called the terminus of replication (9). The replication terminus is active in vitro (10).

Unlike the origins $\alpha$ and $\beta$, the replication origin $\gamma$ can be complemented in trans by the replication initiator protein (2). $\alpha$ and $\beta$ origins are active only when present contiguously with the structural gene for the replication initiator protein (5).

While the location and nucleotide sequence of the replication origin $\gamma$ is known (11), the precise location and the nucleotide sequence of the ori $\alpha$ and ori $\beta$ was, hitherto, unknown.

In this communication we have narrowed down the localization of the ori $\beta$ and demonstrate that a stretch of 283 bp of DNA of the HindIII 2 fragment must be present contiguously with the HindIII 9 and 15 fragments, to form a functional replicon. The criterion used for defining a functional replicon is the ability to drive the replication of a recombinant DNA clone contained in a ColE1 type vector in the absence of active DNA polymerase I. We also report the nucleotide sequence of the region of ori $\beta$ and delineate regions of homology with the ori $\gamma$ of R6K and ori $\gamma$ of Escherichia coli.

**MATERIALS AND METHODS**

**Bacterial and Plasmid Strains**—The E. coli strain HB101 (pro, leu, thi, lacY, strp, rK, end, recA) and E. coli K12, pol A12 (pol Ats) were obtained from Dr. Roy Curtiss, III (University of Alabama, Birmingham). The plasmid pG3 was constructed by cloning the HindIII 9-15-2 segment of R6K into HindIII-cleaved pBR322. The plasmid pG4 and pG5 were constructed as described in the text. The cloning vectors M13 mp5, M13 mp7, and M13 mp9 were obtained from Dr. J. Messing through BRL, Gaithersburg, MD.

**Enzymes and Biochemicals**—The enzymes Cla I (Boehringer), Bal 31, Ace I, Bsa I, Eco RI and HindIII were purchased from a commercial source (New England Biolabs). T4 DNA ligase was prepared as previously published (12). T7 gene 6 exonuclease was prepared as described before (13).

**DNA Sequencing Procedure**—All DNA sequencing was performed by the dideoxy procedure of Sanger et al. (14) with minor variations which have been described in detail (9).

**RESULTS AND DISCUSSION**

**Molecular Cloning and Localization of the Replication Origin $\beta$**—On the basis of electron microscopic of replication intermediates of the R6K derivative RJHC 1040, Crossa et al. (5, 6) had approximately localized the position of the ori $\beta$ on the 4900-bp HindIII 2 fragment, near the HindIII 15-HindIII 2 junction. Since electron microscopic localization can be subject to an error of as much as $\pm 10\%$, it was necessary to map more precisely the location of the ori $\beta$ by subcloning experiments. Previous cloning experiments had demonstrated that the HindIII fragments 15-9-4 and 2-15-9 could form autonomously replicating minichromosomes; the ori $\alpha$ and ori $\beta$ were located in the HindIII 4 and HindIII 2 fragments, respectively (5). All attempts to replicate HindIII 2 or HindIII 4 fragments separately by providing the replication initiation protein in trans had failed (2, 5). Thus, ori $\alpha$ and ori $\beta$, unlike ori $\gamma$, were functional only when present contiguously with the HindIII 9-15 region which codes for a functional replication initiator protein (2, 3). A subsequent report suggested that the reading frame of the replication initiation protein extended into the HindIII 2 fragment (15). Since the nucleotide se-

* This work was supported by United States Public Health Service Grants GM 24357 and 1P01-CA30246-01. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address, Genex Corporation, Bethesda, MD 20814.

§ To whom all correspondence should be addressed.

1 The abbreviation used is: bp, base pair.
sequence of the initiator protein of R6K is not published, we have re-examined and confirmed this observation by performing gene fusion experiments described later.

In view of the fact that ori \( \beta \) could not be physically dissociated from the coding region of the replication initiation protein, we attempted to narrow down the boundary of the origin of replication in HindIII 2 fragment by performing subcloning experiments in the following manner.

**Construction of pJG4 and pJG5**—The miniplasmids pJG4 and pJG5 were constructed from pJG3 as shown in Fig. 1. The plasmid pJG3 contained the HindIII 9-15-2 region of R6K cloned into the HindIII site of pBR322. Inspection of the restriction map of the plasmid indicated that the HindIII 2 fragment contained a single Cla I site and a second Cla I site was located in the pBR322 sequence near the single Eco RI site. The plasmid pJG3 was digested to completion with Cla I and the fragment corresponding to the R6K ori \( \beta \) replicon was purified by agarose gel electrophoresis and ligated to the plasmid vector pBR322 which was also linearized by digestion with Cla I. Ampicillin- and tetracycline-resistant clones were isolated. The clones were then checked for the presence of a functional ori \( \beta \) by transforming the DNA obtained from a mixture of these clones into *E. coli*, pol A12 (pol \( \alpha \)ts) which is a temperature-sensitive mutant of DNA polymerase I. Previous work (16) had demonstrated that pBR322 and any other ColI-related plasmid had an absolute requirement for DNA polymerase I and, therefore, these plasmids could not be maintained in the pol \( \alpha \)ts strain (Fig. 1). The plasmid R6K and its derivatives, in contrast, did not have an absolute requirement for DNA polymerase I and, therefore, plasmid chimeras containing pBR322 and R6K replicons could be maintained in a pol \( \alpha \)ts host at 42 °C (5).

The clone pJG4 which contained the HindIII 9-15 region and part of HindIII 2 (up to the single Cla I site; Fig. 1) was viable in the pol \( \alpha \)ts host even after repeated subculturings at the nonpermissive temperature. Thus, demonstrating that the sequence of R6K present in the clone contained a functional replication origin.

We further narrowed down the location of ori \( \beta \) by constructing the derivative pJG5 in the following way. The plasmid pJG4 contains three Acc I sites, one of which is located in HindIII 2 subfragment of R6K (Fig. 1). Inspection of the DNA sequence of the Acc I sites located in the pBR322 region of the plasmid (17) with that of the Acc I site present in the R6K part of the plasmid pJG4 (DNA sequence is presented later) indicated that the sticky end generated by Acc I in the R6K region was complementary to that of the one generated at the Acc I site which was located nearest to the replication origin of pBR322 (Fig. 1).

The DNA of pJG4 was digested to completion with Acc I and the appropriate subfragment was purified by agarose gel electrophoresis and circularized by ligation. The ligation products were transformed into the pol \( \alpha \)ts host and ampicillin-resistant transformatants were selected at 30 °C. The temperature was shifted to 42 °C and those clones which survived repeated subculturing at 42 °C were examined by restriction analysis. One of the subclones thus obtained was named pJG5 and had the structure shown in Fig. 1. The structure of pJG5 as revealed by restriction enzyme digestion is shown in Fig. 2. The viability of the pJG5 plasmid in the pol \( \alpha \)ts host demonstrated that one end of the ori region was defined by the Acc I site (Fig. 1). DNA sequence analysis presented in a later part of this report demonstrated that the coding sequence of the presumptive initiator protein of R6K ended at a point which was 88 bp away from the HindIII 15-2 junction. Thus, the experiments described above assign an upper limit to one boundary of the functional ori \( \beta \). Since ori \( \beta \) cannot be physically dissociated in a functional form from the structural gene of the replication initiation protein, it is difficult to determine more accurately at this time the location of the other boundary of the ori \( \beta \) domain. Therefore, provisionally, the end of the HindIII 9 fragment which is located close to the Eco RI site of pJG5 (Fig. 1) could be regarded as the other boundary. The nucleotide sequence of this region of HindIII 9 fragment was previously published (11).

A point which merits comment is the validity of using the ability to drive the replication of a ColEI-based recombinant

---

**Fig. 1. Construction of minireplicons of R6K.** The plasmid pJG3 contains the HindIII fragments 9, 15, and 2 cloned into the HindIII site of pBR322. The plasmid pJG3 was digested with Cla I and the appropriate fragment was ligated to Cla I cut pBR322. Selection for Tet and Amp resistance and maintenance in a pol \( \alpha \)ts host at 42 °C yielded the plasmid pJG4. The DNA of pJG4 was digested to completion with Acc I and the digested DNA circularized by ligation and transformed into the pol \( \alpha \)ts host at 30 °C. Selection for amp resistance at 42 °C yielded the plasmid pJG5 containing a functional ori \( \beta \).

**Fig. 2. Restriction analysis of pJG5.** The plasmid DNA was digested with HindIII (A) and HindIII and Acc I (A) and electrophoresed in a 1% agarose slab gel. The ori \( \beta \)-containing 350-bp fragment and the fragments HindIII 9 and HindIII 15 are resolved.
We also cloned the fragment HindIII 2 into M13 mp5 (8) and sequenced the region of ori β by using various restriction fragments from the region as internal primers. The fragment HindIII 15 and various subfragments of HindIII 2 were also cloned into M13 mp7 and sequenced by using a synthetic pentadecamer as the universal primer (24). The sequence obtained by the exonuclease method confirmed the sequence obtained by the M13 cloning procedure. Thus, both strands of the region of ori β were sequenced and the sequences were internally consistent. The DNA sequence of the region of ori β including the COOH-terminal region of the presumptive initiation protein is presented in Fig. 5.

**Determination of the Reading Frame of the Replication Initiation Protein by Gene Fusion Experiments**—Since one boundary of the ori β sequence may turn out to be the COOH-terminal end of the replication initiation protein, we performed fusion of the initiator protein cistron with the sequence specifying the α polypeptide segment of β-galactosidase in order to answer the following two questions: (a) which is the correct reading frame of the initiator protein as determined independently of the DNA sequence data? (b) does the open reading frame of the protein extend from HindIII 9 through HindIII 15 into HindIII 2? The putative cistron of the initiator protein contains two recognition sites for the restriction enzyme Dde I (Fig. 3), one of which is located in the HindIII 9 segment and the other in HindIII 2. A partial Dde I digest obtained with DNA polymerase I and ligated to HindII cut M13 mp7 DNA. The recombinant clones were selected on x gal plates. Several blue plaques (β-galactosidase-producing) were isolated and the single-stranded DNA templates were sequenced by the dideoxy chain extension procedure. The sequence of the fusion region is shown below.

**Characteristics of the Ori β Sequence**—Examination of the DNA clone in the absence of DNA polymerase I as the criterion for identifying replication origins. This method has been used by other laboratories to identify and dissect replication origins of phage λ and *E. coli* (18-21). Timmis et al. (22) had reported that subcloning of the plasmid R6-5 on the basis of the above mentioned criteria yielded 3 separate minireplicons only one of which was used in vivo in *E. coli*. This observation suggests that cloning in pBR322 vectors into pol Afs host at the nonpermissive temperature can reveal “silent” replication origins of phage X and been used by other laboratories to identify and dissect replicons only one of which was used in vivo in *E. coli*. This basis of the above mentioned criteria yielded 3 separate minireplicons. However, in the case of pJG5, the replication origin corresponds to the ori β location which was previously identified by studying the replication in vivo of R6K derivatives (5). Furthermore, the 283-bp fragment containing the region of ori β has an absolute dependence on the initiator protein of R6K for its replication. Therefore, it is highly unlikely that the cloning procedure has yielded a replicon which is anything other than the ori β which is active in vivo in the R6K chromosome.

**Strategy for Sequencing of Ori β**—We have sequenced both strands of the region of the replication origin in the following manner. The plasmid pWL3, which contains the HindIII 2 fragment of R6K cloned into the single HindIII site of pBR313 (see Ref. 23) was linearized by cutting at either the Eco RI or the *Bam* HI site and the linear DNA was extensively digested with the gene 6 exonuclease of the phage T7 (13) to yield single-stranded DNA templates. The 29-bp Eco RI-HindIII fragment of the vector pBR322 provided a convenient universal primer for the dideoxy sequencing procedure (9, 13). A representative autoradiogram of the DNA sequencing gel is shown in Fig. 4.

**Fig. 4. A representative autoradiogram of a DNA sequence gel of the ori region obtained by the T7 exonuclease method.** The sequence corresponds to the coordinates 123–301 in Fig. 5.
Table I
Location of chain-terminating codons in the region of Ori β

<table>
<thead>
<tr>
<th>Reading frame</th>
<th>Sense strand*</th>
<th>Anti-sense strand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (UGA)</td>
<td>2</td>
</tr>
<tr>
<td>Sense strand</td>
<td>88 (UGA)</td>
<td>170 (UAA)</td>
</tr>
<tr>
<td></td>
<td>170 (UAA)</td>
<td>30 (UGA)</td>
</tr>
<tr>
<td></td>
<td>224 (UAG)</td>
<td>54 (UGA)</td>
</tr>
<tr>
<td></td>
<td>266 (UAG)</td>
<td>57 (UGA)</td>
</tr>
<tr>
<td></td>
<td>302 (UAG)</td>
<td>81 (UAA)</td>
</tr>
<tr>
<td></td>
<td>320 (UAA)</td>
<td>99 (UAA)</td>
</tr>
<tr>
<td></td>
<td>362 (UAA)</td>
<td>180 (UAA)</td>
</tr>
<tr>
<td></td>
<td>195 (UAA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>207 (UAA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>345 (UAG)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>251 (UAA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>119 (UGA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>107 (UAA)</td>
<td></td>
</tr>
</tbody>
</table>

* Sense strand refers to the strand shown in Fig. 5.

Fig. 6. The potential hairpin structures present in the ori β sequence.

Table II
Homology between Ori β and Ori γ and Ori β and Ori c

<table>
<thead>
<tr>
<th></th>
<th>Ori β</th>
<th>Ori γ</th>
<th>Ori β</th>
<th>Ori γ</th>
<th>Ori β</th>
<th>Ori γ</th>
<th>Ori β</th>
<th>Ori γ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>36</td>
<td>3</td>
<td>75</td>
<td>3</td>
<td>75</td>
<td>3</td>
<td>170</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>AGCTTTTGCAGGTATT</td>
<td>AACCGCGGAGGTGA</td>
<td>AACCGCGGAGGTGA</td>
<td>AACCGCGGAGGTGA</td>
<td>AACCGCGGAGGTGA</td>
<td>AACCGCGGAGGTGA</td>
<td>AACCGCGGAGGTGA</td>
<td>AACCGCGGAGGTGA</td>
</tr>
<tr>
<td></td>
<td>-12</td>
<td>-41</td>
<td>90</td>
<td>-41</td>
<td>-88</td>
<td>-41</td>
<td>183</td>
<td>-41</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>179</td>
<td>179</td>
<td>179</td>
<td>179</td>
<td>179</td>
<td>179</td>
<td>179</td>
<td>179</td>
</tr>
<tr>
<td></td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>

* * * denotes regions of non-homology. Coordinates for Ori β from Fig. 5, Ori γ from Ref. 11, and Ori C from Ref. 20.

Fig. 7. The regions of homologies between the ori β and ori γ. The coordinates of ori β are as shown in Fig. 5. The coordinates of ori γ and ori c refer to the data published in Refs. 11 and 20, respectively. The three regions of potential hairpin structures (see Fig. 6) are shown on the map.
overlaps the entire coding sequence of the entire initiator protein since all known prokaryotic replication origins are confined to a continuous stretch of just a few hundred base pairs (11, 18–21). It is more likely that a process such as transcription starting from a promoter in HindIII 9 and extending into ori β is required for origin activation. Transcriptional activation of replication origin of λ has been reported (30). Whether a similar situation applies to ori β of R6K must await further experiments.

Acknowledgments—We thank Dr. Jack Keene for his interest in the work and Elizabeth Eubanks and Sharon L. Shipman for typing the manuscript.

REFERENCES