Gene, 70 (1988) 399–403
Elsevier

GEN 02609

A new revision of the sequence of plasmid pBR322

(Plasmid; *rop* gene; expression vector; translation initiation; sequence correction; conserved sequences; pMB1)

Ned Watson

Molecular Biology Research, The Upjohn Company, Kalamazoo, MI 49007 (U.S.A.)

Received 15 March 1988
Accepted 15 May 1988
Received by publisher 28 June 1988

SUMMARY

A revised sequence in the region immediately upstream from the *rop* gene of pBR322 is reported. Two base pairs in the accepted sequence do not exist in the plasmid DNA. Specifically, a TA base pair is missing at sequence coordinate 1893 [Sutcliffe, Cold Spring Harbor Symp. Quant. Biol. 43 (1979) 77–90] and an AT base pair is missing at position 1915, giving a total size for pBR322 of 4361 bp. These changes are in a potential translation initiation sequence and probably reflect errors in the original sequence rather than recent evolution of the plasmid.

INTRODUCTION

Plasmid pBR322 (Bolivar et al., 1977) was one of the first multipurpose cloning vectors constructed for use in *Escherichia coli*. This plasmid and its derivatives have been employed for a variety of purposes including cloning, selection and expression of recombinant molecules, studying elements involved in gene expression, construction of shuttle vectors, nucleotide sequencing vectors, and as a model system for studies of prokaryotic plasmid replication (see review, Balbas et al., 1986). pBR322 was derived from the ColE1-type plasmid pMB1 (Betlach et al., 1976) and shares the same type of replication control mechanism as ColE1 and its relatives (see reviews, Tomizawa, 1978; Davison, 1984; Polisky, 1986).

In addition, pBR322 was one of the first plasmids to be completely sequenced (Sutcliffe, 1979), demonstrating the power of the then new nucleotide sequencing techniques. Besides the small size and relative simplicity of this plasmid, knowledge of the nucleotide sequence added considerably to the utility and extensive use of pBR322.

While sequencing derivatives of pBR322 constructed in this lab, several differences from the accepted pBR322 sequence were observed. These differences could reflect a polymorphism among different isolates of pBR322 or could be the result of errors in the accepted sequence. Several errors in the

Correspondence to: E.B. Watson, Molecular Biology Research, 7242/209/7, The Upjohn Company, 301 Henrietta Street, Kalamazoo, MI 49007 (U.S.A.) Tel. (616)384-9305.

Abbreviations: A, adenosine; Ap, ampicillin; C, cytidine; bp, base pair(s); d, deoxyribo; Δ, deletion; G, guanosine; nt, nucleotide(s); *ori*, origin of DNA replication; p, plasmid; ^{R}, resistance; T, thymidine; Tc, tetracycline.

original sequence of pBR322 have been reported in the last few years and appropriate revisions have been made (Peden, 1983; Backman and Boyer, 1983; Lathe et al., 1984; Heusterspreute and Davison, 1984). In order to differentiate between these possibilities, five different sources of pBR322 were sequenced in the region in question and all were missing the same two base pairs.

EXPERIMENTAL AND DISCUSSION

(a) Strains and plasmids

Strains containing pBR322 were obtained from the Upjohn Culture Collection or purchased from the American Type Culture Collection and are indicated in Table I. A commercially purified preparation of pBR322 plasmid DNA was purchased from Sigma Chemical Co. (St. Louis, MO).

(b) Sequencing

Plasmid DNA was prepared by the method of Birnboim and Doly (1979) and sequenced by the dideoxy chain termination method (Sanger et al., 1977; Sanger and Coulson, 1978), as modified by Zagursky et al. (1985), using avian myeloblastosis virus (AMV) reverse transcriptase (Seikagaku American Inc.). To sequence the upper strand (clockwise direction on the pBR322 map) an oligodeoxynucleotide primer, NW5, with the sequence dCCCGTATCGTGAGCATCC, which hybridizes to pBR322 nt 1834 to 1851 (Sutcliffe, 1979), was used. To sequence the opposite strand, primer NW10 with the sequence dGTCTGGCTTCTGATAAAGCGG, which hybridizes to pBR322 nt 1971 to 1951 (Sutcliffe, 1979), was employed.

(c) Sequence of the region upstream of the *rop* gene of pBR322

Derivatives of pBR322 sequenced in this lab had two bp missing in the region immediately upstream of the *rop* gene (Cesareni et al., 1982), also called *rom* (Som and Tomizawa, 1983), in the area of a potential ribosome-binding site. The Rop protein is plasmid encoded and has been implicated in controlling replication of ColE1-type plasmids (Cesareni et al., 1984; Tomizawa, 1986; for review see Polisky, 1986).

Both strands of pBR322 plasmid DNA from the five different sources shown in Table I were sequenced in this region using the primers described above. All of these isolates are missing an AT and a TA bp, which were reported in the original sequence. Part of a gel showing the sequence of this region is shown in Fig. 1. The location of the missing bp in relation to the *rop* coding region (Cesareni et al., 1982) is presented in Fig. 2a and 2b (region II). The previously reported revisions of the pBR322 sequence are also shown in Fig. 2a and 2b (regions I and III).

TABLE I

Sources of pBR322 sequenced in present study

Culture number	Source	*E. coli* host strain	Comments
UC6576	Upjohn Culture Collection	RR1[a]	Original source: B. O'Malley, Baylor University.
UC6962	Upjohn Culture Collection	CSH50[b]	Deposited by: J. Manis, The Upjohn Company.
ATCC31344	American Type Culture Collection	RR1	Deposited by: Genentech Inc., Patent strain.
ATCC37017	American Type Culture Collection	RR1	Deposited by: H.W. Boyer, University of California, San Francisco.
	Sigma Chemical Company	RR1	Commercial preparation, Product Number D9893.

[a] Strain RR1 genotype: F$^-$ *hsdS*20 ($r_B^- m_B^-$) *ara*-14 *leuB*6 *proA*2 *lacY*1 *galK*2 *rpsL*20 *xyl*-5 *mtl*-1 *supE*44 λ^- (Bolivar et al., 1977).

[b] Strain CSH50 genotype: F$^-$ *ara* Δ(*lac-pro*) *strA thi* (Miller, 1972).

(d) Conclusions and discussion

The results presented here show two differences in the sequence of pBR322. A TA base pair is missing at sequence coordinate 1893 (Sutcliffe, 1979) and an AT base pair is missing at nt position 1915. The accepted sequence of pBR322 contains two previous revisions of the original sequence obtained by Sutcliffe (1979) which are included in the current nucleotide sequence data bases (e.g., the GenBank and EMBO sequence libraries). The first revision was demonstrated by two groups (Peden, 1983; Backman and Boyer, 1983) by nucleotide sequencing and is shown in Fig. 2, designated as region I. This correction was the addition of a CG base pair in the coding sequence Tc^R gene and established the complete correct open reading frame for that gene. The other revision was demonstrated by restriction endonuclease analysis (Lathe et al., 1984) and by nucleotide sequencing (Heusterspreute and Davison, 1984) and is shown in Fig. 2, designated as region III. This correction changes an AT to a TA, resulting in an additional *Alu*I cleavage site in the region encoding the replication primer (RNA II) transcript.

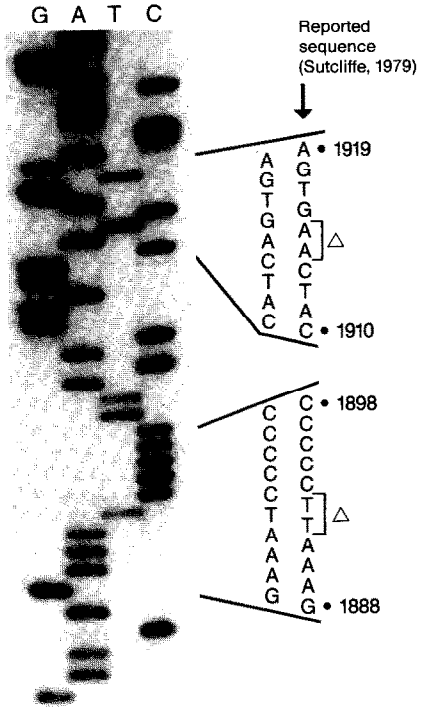


Fig. 1. Dideoxy sequence analysis of pBR322 in the region upstream from the *rop* gene. The sequence of purified plasmid DNA (Sigma Chemical Co.) using primer NW5 is shown. The sequence deduced from the gel is shown at the immediate right and the sequence reported by Sutcliffe (1979) is shown at the far right (see downward arrow). The positions of the differences in the sequences are indicated by the brackets and Δ symbols at the right. Refer to EXPERIMENTAL AND DISCUSSION, sections **a** and **b**, for further details.

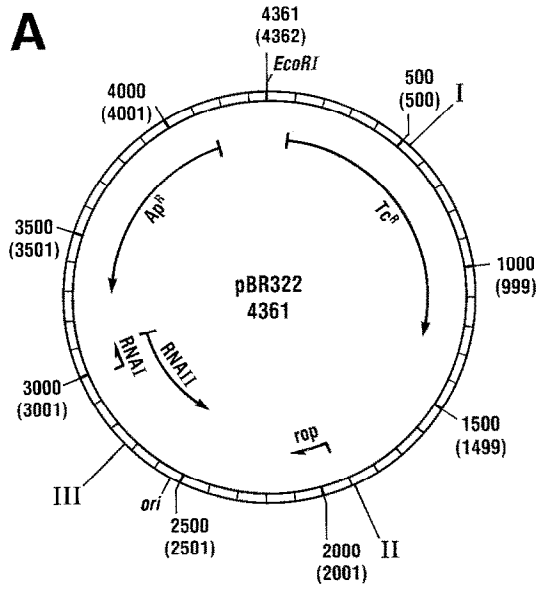
The work presented here suggests that two bp in the reported sequence upstream from *rop* do not exist in the DNA (see region II, Fig. 2). These differences neither create nor destroy sites for known restriction endonucleases. The bp indicated were missing in the sequence obtained from both strands of pBR322 DNA from five different sources. The same two bp were reported missing in pBR322 sequence by Nugent et al. (1986). Moreover, as reported here, these bp were not found in the pBR322 plasmid DNA from the original American Type Culture Collection strain deposited by H.W. Boyer, which is the same pBR322 sequenced by Sutcliffe (1979). Therefore, it is likely that these differences from the accepted sequence are due to errors in the original reported sequence, rather than the result of evolution of the plasmid DNA.

The same two bp do not appear in the sequence upstream from the *rop* gene of the related plasmid, ColE1 (Chan et al., 1985). This region is highly conserved at the DNA level between these two plasmids (Cesareni et al., 1982). The revised sequence of pBR322 presented here indicates that the region preceding the coding sequence for the *rop* protein of these plasmids is more highly conserved than previously reported. One of the sequence changes reported here decreases the spacing between a Shine-Dalgarno sequence (GGAGG) and the initiation codon (GTG) of the *rop* gene from six to five, while the other change upstream of the Shine-Dalgarno is also in a region that may be important in translation initiation.

The revisions presented here change the accepted pBR322 map coordinates as indicated in Fig. 2 and result in a total size of 4361 bp for pBR322.

ACKNOWLEDGEMENTS

I would like to thank Kathy Hiestand and Nancy Burr for help in preparing the manuscript and table. I am very grateful to Nicole Theriault's laboratory for



B

I Sutcliffe (1979) (521) G C A G G C C C G T (530)
Peden (1983) G C A G G C C C C G T
521 531

II Sutcliffe (1979) (1890) A A T T C C C C C T T A C A C G G A G G C A T C A A G T G A C C (1921)
This work and A A T C C C C C T T A C A C G G A G G C A T C A G T G A C C
Nugent et al. 1891 1920
(1986)

III Sutcliffe (1979) (2727) T C A A T G C T C A (2736)
Lathe et al. (1984) T C A T A G C T C A
2726 2735

Fig. 2. Revisions in the sequence of pBR322. (**A**) Locations of the revisions shown in relation to the map of pBR322. Region I indicates the revision reported by Peden (1983) and Backman and Boyer (1983). Region II indicates the revision reported in this work. Region III indicates the revision reported by Lathe et al. (1984) and Heusterspreute and Davison (1984). The size of the plasmid after these revisions is 4361 bp. The old coordinates of Sutcliffe (1979) are shown in parentheses. (**B**) The nucleotide sequence in the areas of the revisions (the upper strand is shown). In each case (regions I, II and III) the original sequence and map coordinates reported by Sutcliffe (1979) are shown on the upper line (in parentheses) and the revised sequence and map coordinates are shown on the lower line. In region I, the additional C reported in the revised sequence is underlined. In region II, the TT to T and AA to A revisions are overlined. The consensus Shine-Dalgarno sequence and the GTG start codon of the *rop* gene are boxed. In region III the position of the AT to TA revision is overlined.

the synthesis of oligodeoxynucleotide primers and to David Siemieniak and Peter Harris for patient help with computer analysis. My special thanks go to Eric Olson for critical reading of the manuscript, stimulating discussions, encouragement and support.

REFERENCES

Backman, K. and Boyer, H.W.: Tetracycline resistance determined by pBR322 is mediated by one polypeptide. Gene 26 (1983) 197–203.

Balbas, P., Soberon, X., Merino, E., Zurita, M., Lomeli, H., Valle, F., Flores, N. and Bolivar, F.: Plasmid vector pBR322 and its special-purpose derivatives-a review. Gene 50 (1986) 3–40.

Betlach, M., Hershfield, V., Chow, L., Brown, W., Goodman, H.M. and Boyer, H.W.: A restriction endonuclease analysis of the bacterial plasmid controlling the *Eco*RI restriction and modification of DNA. Fed. Proc. 35 (1976) 2037–2043.

Birnboim, H.C. and Doly, J.: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7 (1979) 1513–1523.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S.: Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2 (1977) 95–113.

Cesareni, G., Muesing, M.A. and Polisky, B.: Control of ColE1 DNA replication: the *rop* gene product negatively affects transcription from the replication primer promoter. Proc. Natl. Acad. Sci. USA 79 (1982) 6313–6317.

Cesareni, G., Cornelissen, M., Lacatena, R.M. and Castagnoli, L.: Control of pMB1 replication: inhibition of primer formation by Rop requires RNA I. EMBO J. 3 (1984) 1365–1369.

Chan, P.T., Ohmori, H., Tomizawa, J. and Lebowitz, J.: Nucleotide sequence and gene organization of ColE1 DNA. J. Biol. Chem. 260 (1985) 8925–8935.

Davison, J.: Mechanism of control of DNA replication and incompatibility in ColE1-type plasmids – a review. Gene 28 (1984) 1–15.

Heusterspreute, M. and Davison, J.: Restriction site bank vectors, II. DNA sequence analysis of plasmid pJRD158. DNA 3 (1984) 259–268.

Lathe, R., Kieny, M.P., Skory, S. and Lecocq, J.-P.: Linker tailing: unphosphorylated linker oligonucleotides for joining DNA termini. DNA 3 (1984) 173–182.

Miller, J.H.: Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972.

Nugent, M.E., Smith, T.J. and Tacon, W.C.A.: Characterization and incompatibility properties of ROM$^-$ derivatives of pBR322-based plasmids. J. Gen. Microbiol. 132 (1986) 1021–1026.

Peden, K.W.C.: Revised sequence of the tetracycline-resistance gene of pBR322. Gene 22 (1983) 277–280.

Polisky, B.: Replication control of the ColE1-type plasmids. In Reznikoff, W. and Gold, L. (Eds.), Maximizing Gene Expression. Butterworth Publishers, Stoneham, MA, 1986, pp. 143–170.

Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

Sanger, F. and Coulson, A.R.: The use of thin acrylamide gels for DNA sequencing. FEBS Lett. 87 (1978) 107–110.

Som, T. and Tomizawa, J.: Regulatory regions of ColE1 that are involved in determination of plasmid copy number. Proc. Natl. Acad. Sci. USA 80 (1983) 3232–3236.

Sutcliffe, J.G.: Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43 (1979) 77–90.

Tomizawa, J.: Replication of colicin E1 plasmid DNA in vitro. In Molineaux, I. and Kohiyama, J. (Eds.), DNA Synthesis: Present and Future. Plenum Press, New York, NY, 1978, pp. 797–826.

Tomizawa, J.: Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. Cell 47 (1986) 89–97.

Zagursky, R.J., Baumeister, K., Lomax, N. and Berman, M.L.: Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Techn. 2 (1985) 89–94.

Communicated by J.L. Slightom.