Nucleotide sequence of a secondary attachment site for bacteriophage lambda on the Escherichia coli chromosome

Éva Csordás-Tóth, Imre Boros and Pál Venetianer

Institute of Biochemistry, Biological Research Center, Hungarian Academy of Sciences, 6701 Szeged, P.O.B. 521, Hungary

Received 15 August 1979

ABSTRACT

The nucleotide sequence of a secondary attachment site for bacteriophage lambda was determined in a region near the *rrnB* gene at 88 min on the *E. coli* chromosome. The sequence has a 8 base pair interrupted homology GCT TTTTA to the common core of the primary attachment site (*attB*) and the corresponding phage sequence (*attP*). The site of crossover during integration lies probably between nucleotides -3 and +1. The flanking regions have no obvious homology to the arms of either *attP* or *attB*.

INTRODUCTION

Bacteriophage lambda can integrate at a unique site on the *E. coli* chromosome according to the mechanism proposed by Campbell¹. The nucleotide sequence of this site (*attB*) as well as the corresponding sequence (*attP*) on the phage DNA have been determined². Confirming predictions made on the basis of genetic data, the two sequences were found to contain a 15 base pair long complete homology, termed "common core" and widely different flanking regions, termed "arms". The two structures are usually designated POP' and BOB', the two ends of the prophage BOP' and POB'. If the *attB* site is deleted lambda can integrate at several other sites into the bacterial chromosome. This requires the same factors as the normal integration but its frequency is much lower. The number of such secondary sites is limited and their distribution is nonrandom³. It has been proposed that these sites are partially homologous with the common core of the primary *att* site⁴. In order to test this prediction and learn more about the mechanism of the secondary integration we determined and here report the nucleotide

sequence of a secondary att site.

The specialized transducing phage λ rif^d18 had been constructed by Kirschbaum and Konrad⁵ as indicated above, that is by forcing λ to integrate at a secondary site, at 88 min on the E. coli chromosome. The phage contains a large bacterial insertion to the left from the att site. Earlier we reported the molecular cloning of a fragment from this phage, containing the phage-bacterial junction (Δ OP')⁶. Recently we cloned a 13.5kB BamHI fragment of the bacterial chromosome homologous to the bacterial part of the λ rif^d18 phage (Boros et al., manuscript in preparation). This fragment must contain the entire secondary attachment site Δ OA'. The two relevant sequences from these clones has been determined and compared with the published attP sequence (POP').

MATERIALS AND METHODS

Methods of cloning, plasmid preparation, restriction endonuclease digestion, gel electrophoresis were as described earlier⁶. Restriction endonucleases: AluI, BspI, MboII, MspI, PstI, BamHI and HindIII were prepared in this laboratory according to established protocols (for review see 7.). Large DNA fragments were purified by preparative agarose gel electrophoresis, using 2x15 cm 1% gel columns.

Fragments were isolated from agarose according to⁸. Smaller fragments were separated on 8% acrylamide gels and eluted by 1M NaCl. Polynucleotide kinase was prepared by the method of Panet et al.⁹.

Preparation of [γ -³²P]-ATP and sequencing was carried out by the method of Maxam and Gilbert¹⁰.

RESULTS AND DISCUSSION

The cloning and partial restriction mapping of the 7 kB BamHI fragment of λ rif^d18 containing the phage-bacterial junction had been described elsewhere⁶. The sequencing strategy is summarized on Fig.1. Briefly, from the recombinant plasmid 2/12 a 1.6 kB HindIII-BamHI fragment was isolated and mapped with several restriction endonucleases. Fig.1. also shows which fragments were sequenced at the right end where

Fig.1. Physical map of the λ rif^{d18} phage, the cloned fragment, and the sequenced region

the *att* site is located. Fig.2. shows the autoradiogram of a sequencing gel. Fig.3. is the 365 BP long sequence determined, including the junction (Δ OP'). 80% of this sequence had been determined in both directions. Where this was not possible, the sequencing was repeated several times. The POP' sequence had been determined in two laboratories^{2,11}. As to be expected the right part of our sequence agrees with these results. (At site 18 and 35 the sequences of Landy and Ross² and that of Davies et al.¹¹ differ. In both cases our results agree with those of Landy and Ross.) In assigning amino acids to the right end of P' we follow Davies et al.¹¹ who deduced the carboxyl-terminal end of the *int* protein sequence from their nucleotide sequence. As our fragment is longer than theirs, the amino acid sequence is also longer here by 14 residues.

The schematic structure of the recombinant plasmid pBK 17 is shown on Fig.4. Its construction will be described elsewhere, the insertion is a 13.5 kB long *Bam*HI fragment of the bacterial chromosome containing the *rrnB* gene¹².

The unique *Pst*I site in this insert corresponds to the *Pst*I site at -52 of the previously determined sequence and the entire insert to the right from this site is homologous to the λ rif^{d18} fragment cloned in plasmid 2/12. This had been proven both by

Fig.2. Autoradiogram of a sequencing gel. It shows the sequence of the common core region of the attachment site.

restriction mapping and electron microscopic heteroduplex analysis /data not shown/. 120 base pairs to the left from the PstI site is an AluI site which is not present in the previous sequence. Thus the sought for $\Delta OA'$ sequence must be located on this 120 BP PstI-AluI subfragment. Plasmid pBK 17 was digested with PstI, labelled at the 5' ends with [γ - 32 P]-ATP and polynucleotide kinase, digested with AluI and the 120 BP fragment

Fig.3. Sequence of the right end of the fragment cloned from λ rif^d18.

The boxed region corresponds to the common core. The right arm sequence agrees with that reported by Landy and Ross² but differs from that reported by Daniels et al.¹¹ at sites doubly underlined. The arrows indicate repeats. The site of homology with ISI is also indicated.

Fig. 4. Physical map of the recombinant plasmid pBK 17. Lower thick bar: vector plasmid pBR322. Solid upper line: 13.5 kB BamHI fragment of the *E. coli* chromosome. Thick bars: sequences corresponding to 16S and 23S rRNA in the *rrnB* gene.

was isolated. In this case the sequence was determined only in one direction. As expected, the PstI end of this fragment proved to be identical with the sequence between -52 and +1 on Fig. 3. while the AluI end, /from +2/ region was entirely different. Fig. 5. shows the common cores and their flanking regions.

Thus the secondary attachment site is probably an eight base pair interrupted region of homology with base pairs -7 - -5 and -3 - +1 of the common core of the primary attachment site. If one assumes a simple recombination event between homologous regions, not involving branch migration and mismatch repair, then it must take place between base pairs -3 and +1. The right end left arms have no obvious homology to either the phage or the primary bacterial arms /P,P',B,B'/ . The short direct repeat flanking the core in the λ rif^d18 was not found in pBK 17 thus its presence in Δ OP' might be fortuitous.

While this manuscript was in preparation two papers have been published on the sequencing of secondary att sites in the

		-10	0	10
Wild-type	/POP' /	ggttcaGCTTTTTTATACTA	Agttgg	
λ rif ^d 18	/ Δ OP' /	ggtttgGCTaTTTTTATACTA	Agttgg	
pBK 17	/ Δ OA' /	ggtttgGCTaTTTTTAccacgactgtc		

Fig. 5. Comparison of common cores and adjacent sequences

galT¹³ and trpC¹⁴ genes. Our conclusion agrees with the authors of these papers in that the sequence of the arms is probably unimportant and secondary attachment sites are characterized by partial homology to the common core of the primary attachment sites. Bidwell and Landy¹³ report a 8 bp interrupted homology just as in our case, while Christie and Platt¹⁴ find six base pairs of continuous homology. The former authors localize the crossover site between bases +4 and +5 while in our case it must be between -3 and +1. Our result can be reconciled with the placement between -2 and +3 proposed by Christie and Platt¹³. However even if the simplifying assumptions common to these speculations are correct it is not necessary to assume that the relative position of the crossover site should be the same in the case of all secondary attachment sites.

REFERENCES

- 1 Campbell, A.M. (1962) Adv. Genet. 11, 101-145
- 2 Landy, A. and Ross, W. (1977) Science 197, 1147-1160
- 3 Shimada, K., Weisberg, R.A. and Gottesmann, M.E. (1972) J. Mol. Biol. 63, 483-503
- 4 Shulman, M.J., Mizuchi, K. and Gottesmann, M.E. (1976) Virology 72, 13-22
- 5 Kirschbaum, J.B. and Konrad, E.B. (1973) J. Bacteriol. 116, 517-526
- 6 Kiss, A., Sain, B., Kiss, I., Boros, I., Udvardy, A. and Venetianer, P. (1978) Gene 4, 137-152
- 7 Roberts, R.J. (1978) Gene 4, 183-193
- 8 Smith, H.O. and Birnstiel, M.L. (1976) Nucl. Acids. Res. 3, 2387-2398
- 9 Panet, A., Van de Sande, J.H., Loewen, P.C., Khorana, H.G., Raae, A.J., Lillehaug, J.R. and Kleppe, K. (1973) Biochemistry 12, 5045-5050
- 10 Maxam, A.M. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. USA 74, 560-564
- 11 Davies, R.W., Schreier, P.H. and Büchel, D.E. (1977) Nature 270, 757-759
- 12 Boros, I., Kiss, A. and Venetianer, P. (1979) Nucl. Acids Res. 6, 1817-1830
- 13 Bidwell, K. and Landy, A. (1979) Cell 16, 397-406
- 14 Christie, G.E. and Platt, T. (1979) Cell 16, 407-413