Mol Gen Genet (1987) 210:578–580

MGG

Short communications

Illegitimate recombination in *Bacillus subtilis*: nucleotide sequences at recombinant DNA junctions

Vladimir I. Bashkirov, Fuat K. Khasanov, and Alexander A. Prozorov
N.I. Vavilov Institute of General Genetics, USSR Academy of Sciences, Gubkin St 3, Moscow B-333 GSP-1 117809, USSR

Summary. The illegitimate integration of plasmid pGG20 (the hybrid between *Staphylococcus aureus* plasmid pE194 and *Escherichia coli* plasmid pBR322) into the *Bacillus subtilis* chromosome was studied. It was found that nucleotide sequences of both parental plasmids could be involved in this process. The recombinant DNA junctions between plasmid pGG20 and the chromosome were cloned and their nucleotide sequences were determined. The site of recombination located on the pBR322 moiety carried a short region (8 bp) homologous with the site on the chromosome. The nucleotide sequences of the pE194 recombination sites did not share homology with chromosomal sequences involved in the integration process. Two different pathways of illegitimate recombination in *B. subtilis* are suggested.

Key words: Bacillus subtilis – pGG20 plasmid – Illegitimate recombination – Topoisomerase

Illegitimate recombination between nonhomologous DNA sequences is quite common in various bacteria. The mechanisms of illegitimate recombination in most cases are unknown, although in several systems the involvement of DNA topoisomerases has been established (Ikeda et al. 1984; Ikeda 1986). An interesting system of illegitimate recombination in vivo comprising integration of *Staphylococcus aureus* plasmid pE194 into the *Bacillus subtilis* chromosome has been developed by Hofemeister et al. (1983). The authors found that the erythromycin resistance plasmid pE194, which is temperature-sensitive for replication, was able to integrate spontaneously at multiple sites on the chromosome with a low frequency (3×10^{-8}). They suggested that pE194 integrated by a Campbell-like reciprocal recombination event. Despite the low frequency of integration, the temperature-sensitive replication of the plasmid allowed selection of the integrants at a restrictive temperature (52° C) on erythromycin containing plates. It has been shown that pE194 integration is independent of the functioning of the principal gene of *B. subtilis* for homologous recombination – *recE* (Hofemeister et al. 1983). We have shown that pE194 integration is independent of mutation in the majority of *B. subtilis* rec-genes, except *rec149* mutation which decreases the frequency of integration (Khasanov et al. 1985). Hofemeister et al. (1983) have mapped three regions on pE194 involved in recombination with the chromosome. Two additional sites of integration on this plasmid were mapped in our laboratory (Khasanov et al. 1985; this communications. Sites for illegitimate integration into the *B. subtilis* chromosome were found on another *S. aureus* plasmid, pC194 and on the *Escherichia coli* plasmid, pBR322 (Khasanov et al. 1985). Prozorov et al. (1985) reported that sites on wheat DNA cloned in *B. subtilis* could also be involved in this process.

The objective of this investigation was to determine the nucleotide sequences on plasmid and chromosomal DNAs involved in illegitimate recombination. For this purpose we used the system of integration of the plasmid pGG20. This plasmid consists of pE194 and pBR322 replicons fused at their unique *Pst*I sites (Khasanov et al. 1985). Replication of this plasmid in *B. subtilis* is temperature-sensitive due to the origin of replication from pE194.

Using selection for erythromycin resistance, 4 independent integrants of this plasmid were obtained and analysed. In strains b2 and b4 the pGG20 insertion occurred due to sites in pBR322. These sites have been mapped earlier between nucleotides 3612–3757 (b2) and nucleotides 3329–3410 (b4, Khasanov et al. 1985), according to the sequence map of pBR322 (Sutcliffe 1978). In strains b1 and b5, sequences with unknown co-ordinates located on the pE194 moiety of pGG20 were involved in integrational recombination.

It was necessary to clone the chromosome-plasmid junctions to study these four cases of integration and we took advantage of the ability of pGG20 to replicate in *E. coli*. Chromosomal DNA from the integrants was cleaved by restriction endonucleases having no target sites within pGG20 and ligated at DNA concentrations less than 10 μg/ml to direct preferential formation of circular monomers. The ligation mixture was used for *E. coli* HB101 transformation. Recircularization of the DNA fragments composed of the complete pGG20 replicon and adjacent chromosomal sequences could have resulted in the appearance of hybrid plasmids in transformants. Selection was performed for tetracycline resistance. *Bgl*II restriction of b2 and b5 DNAs resulted in the isolation of two plasmids, pGG20-b2 and pGG20-b5. *Sma*I cleavage of b1 DNA resulted in the formation of a hybrid plasmid pGG20-b1. Repeated efforts to obtain a hybrid plasmid by b4 DNA restriction with *Bgl*II, *Sma*I and several other enzymes were unsuccessful. It became possible to isolate the pGG20-b4 plasmid only when *Eco*RI, which cleaves pGG20 in the region of tetracycline resistance gene promoter, was used. This plasmid carried only one recombinant DNA junction, while each of

Offprint requests to: V.I. Bashkirov

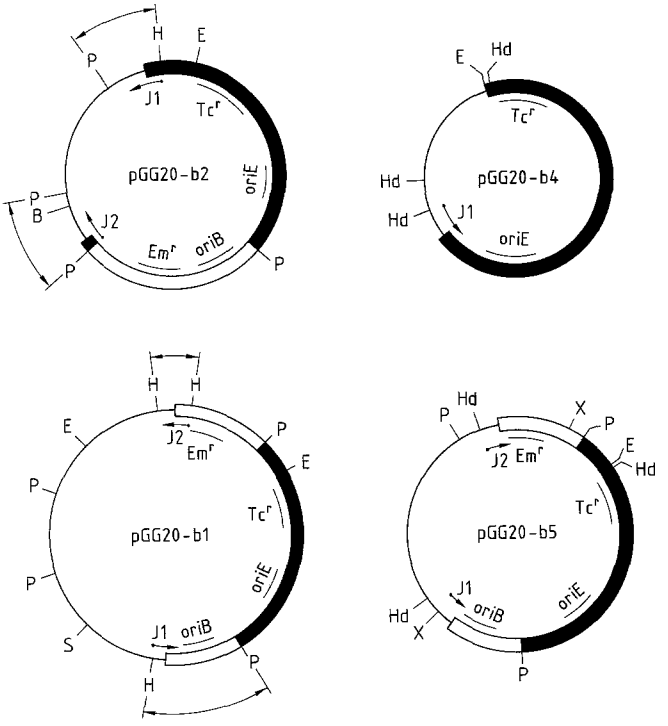

Fig. 1. Recombinant plasmids containing plasmid-chromosome junctions. The position of DNA junctions (J1 and J2) was identified by digestion of plasmid DNA with various restriction endonucleases (Boehringer-Mannheim) and agarose gel electrophoresis with subsequent blotting and hybridization with ^{32}P-labelled pGG20. Junctions were identified as labelled fragments which did not co-migrate with any fragment from pGG20. In the cases indicated in the outer segment of the circle, these fragments were cloned in the pUC19 polylinker, prior to sequencing by a modified Maxam and Gilbert (1977) procedure (Chuvpilo and Kravchenko 1984). The direction of sequencing is indicated by an *arrow*. In the plasmid maps, the symbols identifying the location of the genes and restriction enzyme sites are as follows: *oriB*, origin of pE194 replication; *oriE*, origin of pBR322 replication; Tc^r, tetracycline resistance; Em^r, erythromycin resistance; E, *Eco*RI; H, *Hin*dII; Hd, *Hin*dIII; B, *Bgl*II; P, *Pst*I; S, *Sma*I; X, *Xba*I. Only those *Hin*dII sites which were used for cloning are shown for pGG20-b1 and pGG20-b2. Chromosomal DNA is shown as a *light line*, pBR322 as a *closed bar* and pE194 as an *open bar*

the plasmids pGG20-b1, pGG20-b2 and pGG20-b5 had two DNA junctions. The restriction maps of the plasmids are shown in Fig. 1.

It is worth mentioning that hybrid plasmids carrying adjacent chromosomal regions were found in one-third of all transformants. Other transformants carried plasmids identical to pGG20 as seen from the results of their restriction mapping (data not shown). They could have arisen by recircularization of tandemly repeated copies of pGG20 (as in case of *B. subtilis* transformation via multimeric forms of plasmid DNA; Canosi et al. 1978). The possibility of plasmid amplification after illegitimate integration into the *B. subtilis* chromosome was confirmed by Khasanov et al. (1987).

The 1.25 kb *Pst*I fragment, including the recombinant junction J2 from pGG20-b2, was cloned in the *Pst*I site of the pUC19 polylinker to allow nucleotide sequencing. The recombinant junction J1 from pGG20-b2 was cloned on a *Hin*dII/*Pst*I fragment in *Sma*I/*Pst*I digested pUC19 (Fig. 1). Plasmid-chromosome junctions J1 and J2 from pGG20-b1 were cloned on a 1.5 kb *Hin*dII/*Pst*I and on a 0.7 kb *Hin*dII fragment respectively, using either *Sma*I/*Pst*I- or *Sma*I-cleaved pUC19 and then were sequenced in the direction shown in Fig. 1. The single recombinant DNA junction from pGG20-b4 was sequenced without subcloning as indicated by the arrow in Fig. 1. DNA junctions J1 and J2 from PGG20-b5 were sequenced beginning from *Xba*I (J1) or *Hin*dIII (J2) sites in the direction shown in Fig. 1.

The results of nucleotide sequence determination of recombinant DNA junctions (Fig. 2a) indicated that the b2 integrant arose by recombination between short (8 bp) regions of homology (indicated by asterisks in Fig. 2a). Moreover, there is an additional region of homology of 6 bp near the junction (indicated by asterisks, see Fig. 2a). The sequence of the single DNA junction of strain b4 allows us to suggest that the recombination occurred leftward from the first nucleotide of the chromosomal DNA sequence (indicated by carets in Fig. 2a). The region of 5 bp homology separated by one nucleotide from the junction point is shown by asterisks. In the case of b2, duplication of the short region of homology occurred during integration, as in the Campbell-like reciprocal recombination mechanism suggested by Hofemeister et al. (1983). Although the crossover point in strain b4 is uncertain, we cannot exclude the possibility of duplication in the region of recombination.

Unlike integration of pGG20 using sites on the pBR322 moiety, cointegration by means of the sequences in pE194 (strains b1 and b5) occurred irrespective of either any homologies in recombination sites (b1) or where there was coincidence of only two nucleotides (b5; see Fig. 2a). The recombinant junctions and the parental sites on pE194 differ greatly from those on pBR322, but have reasonable homology with each other (Fig. 2b). This indicates that the integrants b1 and b5 have arisen by the same mechanism, involving plasmid recombination sites without significant homology to the chromosomal ones.

Integration of plasmids into the *B. subtilis* chromosome involving short homologies (strain b2) may be explained by slippage error in the replication model proposed for deletions (Efstratiadis et al. 1980; Albertini et al. 1982). Thus, if the end of a replicating strand could mispair with a partially complementary sequence in a gap of another DNA molecule, this might result in cointegrate formation, and the additional homology adjacent to the junction would serve to "anchor" the mispaired strand in the gap, as has been suggested by Marvo et al. (1983) for intermolecular recombination.

Another type of cointegration without homology between recombination sites (strains b1 and b5) cannot be explained by this model. This type of illegitimate recombination is probably mediated by an unknown recombination enzyme of *B. subtilis*. It was shown by Ikeda et al. (1982; 1984) that illegitimate recombination between pBR322 and phage λ in *E. coli* could be mediated by DNA gyrase or T4 DNA topoisomerase (Ikeda 1986) and that recombination sites frequently had no homologies between each other. We can only speculate about the involvement of *B. subtilis* DNA gyrase in the illegitimate recombination described here, since DNA gyrase recognition and cleavage sequences are unknown. We think therefore, that an unknown *B. subtilis* "recombinase" (which could be a type I or type II DNA topoisomerase or another DNA cutting enzyme) may be involved in the formation of recombinants b1 and b5.

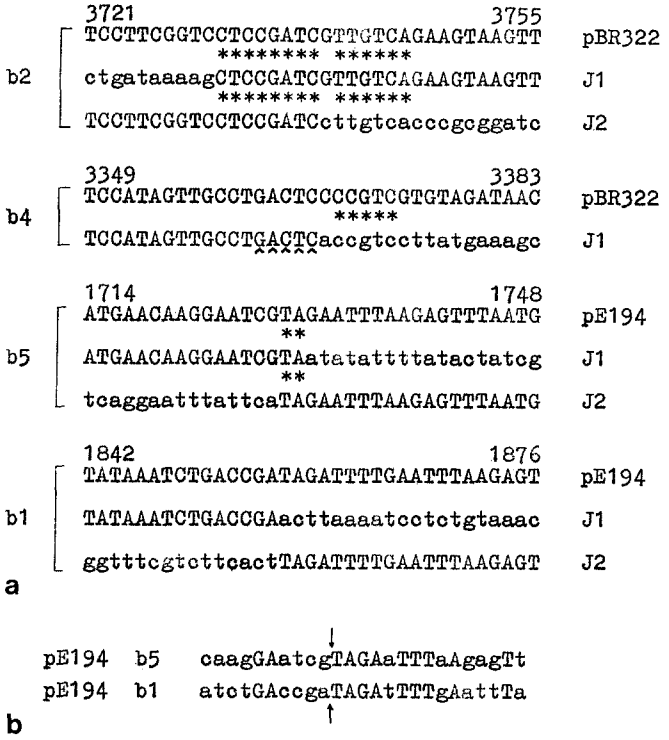

Fig. 2a, b. Nucleotide sequences involved in recombination. **a** Nucleotide sequences of plasmid-chromosome junctions in strains b1, b2, b4 and b5. Sequences at the DNA junction of the recombinants and the sites of recombination on pBR322 and pE194 prior to recombination are given. Chromosomal and plasmid sequences are in *lower-* and *upper-case type*, respectively. The homologies at crossover points and the additional homologous regions are indicated by *asterisks*. The sequences of pBR322 are taken from Sutcliffe (1978). The nucleotide sequence of pE194 is taken from Horinouchi and Weisblum (1982) except for b1 (our data), where we have found 5′-TAAGAG at position 1870, instead of 5′-TAGG. The locations of the sequences in pBR322 and pE194 are given above the first and last base pair in each sequence. The crossover point in strain b4 is uncertain and is indicated by *carets*. **b** Comparison of recombination sites on pE194. The sequences of one strand are given in the 5′→3′ direction. Common nucleotides are given in *upper case* and those which are not homologous with *lower case type*. The *arrows* identify the presumed point of crossover

As for strain b4, we cannot attribute it to any of the mechanisms discussed above. b4 might belong to the first class (as b2), but in the absence of the second junction this cannot be ascertained.

It is possible that the product of the *rec149* gene could be responsible for some illegitimate recombination events in *B. subtilis*, since this mutation decreases the frequency of integration. Recently, we have found that illegitimate recombination in our system is increased by 2 orders of magnitude during the development of competence of *rec*$^+$ and *recE4* cells (unpublished data). This suggests that hypothetical "recombinase" can be activated simultaneously with the development of competence, as has been observed for the *B. subtilis* SOS-like (SOB) system (Yasbin 1977). However, this illegitimate recombination machinery of *B. subtilis* cells differs from the competence-related SOB system, since the latter can be inhibited by host *recE4* deficiency (Love and Yasbin 1984).

Thus, our results indicate the existence of at least two different mechanisms of illegitimate recombination resulting in integration of plasmid replicons into the *B. subtilis* chromosome. One mechanisms possibly involves replication slippage and the other breakage and reunion involving perhaps a sequence-specific enzymatic machinery.

Acknowledgements. We would like to thank Mrs. I. Leonova for technical assistance.

References

Albertini AM, Hofer M, Calos MP, Miller JH (1982) On the formation of spontaneous deletions: The importance of short sequence homologies in the generation of large deletions. Cell 29:319–328

Canosi U, Morelli G, Trautner TA (1978) The relationship between molecular structure and transformation efficiency of some *S. aureus* plasmids isolated from *B. subtilis*. Mol Gen Genet 166:259–267

Chuvpilo SA, Kravchenko VV (1984) A simple and rapid method for sequencing DNA. FEBS Lett 179:34–36

Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, Forget BG, Weissman SM, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) The structure and evolution of human β-globin gene family. Cell 21:653–668

Hofemeister J, Israeli-Reches M, Dubnau D (1983) Integration of plasmid pE194 at multiple sites on the *Bacillus subtilis* chromosome. Mol Gen Genet 189:58–68

Horinouchi S, Weisblum B (1982) Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide and streptogramin type B antibiotics. J Bacteriol 150:804–814

Ikeda H (1986) Illegitimate recombination mediated by T4 DNA topoisomerase in vitro. Recombinants between phage and plasmid DNA molecules. Mol Gen Genet 202:518–520

Ikeda H, Aoki K, Naito A (1982) Illegitimate recombination mediated in vitro by DNA gyrase of *Escherichia coli*. Structure of recombinant DNA molecules. Proc Natl Acad Sci USA 79:3724–3728

Ikeda H, Kawasaki I, Gellert M (1984) Mechanisms of illegitimate recombination: Common sites for recombination and cleavage mediated by *E. coli* DNA gyrase. Mol Gen Genet 196:546–549

Khasanov FK, Bashkirov VI, Prozorov AA (1985) Study of the integration of some plasmids into the *Bacillus subtilis* chromosome. Genetika (in Russian) 21:1618–1626

Khasanov FK, Bashkirov VI, Prozorov AA (1987) Amplification of a plasmid DNA inserted into the *Bacillus subtilis* chromosome. Genetika (in Russian) 23:14–20

Love PE, Yasbin RE (1984) Genetic characterization of the inducible SOS-like system of *Bacillus subtilis*. J Bacteriol 160:910–920

Marvo SL, King SR, Jaskunas SR (1983) Role of short regions of homology in intermolecular illegitimate recombination events. Proc Natl Acad Sci USA 80:2452–2456

Maxam A, Gilbert W (1977) A new method for sequencing DNA. Proc Natl Acad Sci USA 74:560–564

Prozorov AA, Bashkirov VI, Khasanov FK, Glumova EF, Yrich VY (1985) Insertion of eukaryotic DNA into the *Bacillus subtilis* genome by means of a temperature-sensitive plasmid vector. Gene 34:39–46

Sutcliffe JG (1978) Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp Quant Biol 43:77–90

Yasbin RE (1977) DNA repair in *Bacillus subtilis*. I. The presence of an inducible system. Mol Gen Genet 153:211–218

Communicated by W. Arber

Received January 12, 1987 / May 29, 1987