Deletions and DNA rearrangements within the transposable DNA element IS2. A model for the creation of palindromic DNA by DNA repair synthesis

Jürgen Besemer*, Gabriele Görtz and Daniel Charlier

Institut für Genetik der Universität Köln, D-5000 Köln 41, Weyertal 121, GFR

Received 25 September 1980

ABSTRACT

Three derivatives of mutant galOP-308::IS2-I of Escherichia coli were characterized by DNA sequence analysis. Deletions and DNA sequence rearrangements were observed which apparently were initiated at short A-T rich inverted repeats within IS2. Two of the mutants carried newly synthesized DNA sequences which were inverted copies of already existing IS2 sequences. Thus long stretches with twofold symmetry were formed. It is discussed whether these inverted repeats were formed by DNA repair synthesis which was initiated at the A-T rich palindromes of IS2.

INTRODUCTION

Double stranded DNA which reacts with proteins frequently is characterized by regions of twofold symmetry, e.g. operators and replication origins. One may ask how such sequences were created in evolution. From the study of revertants of the insertion mutant galOP-308 of E. coli Ghosal and Saedler (1) and Ghosal et al. (2) proposed that long palindromic sequences may be created by spontaneous denaturation of DNA strands within the replication fork and two or more template slippages of the replication enzymes. From our recent analysis of similar revertants of mutant galOP-308 we would like to modify this model in postulating that DNA repair synthesis is sufficient to explain the creation of palindromic DNA sequences.

METHODS AND MATERIALS

Methods of bacteria and bacteriophage manipulations were according to Miller (3).

Isolation of mutants were described by Delius et al. (4).

Galactokinase was measured as described previously (5).

Preparation of DNA for sequence analysis: Ø80dgal phages of the appropriate mutant genotype were prepared, DNA was extracted and restricted with

restriction endonuclease HindIII. A HindIII fragment of approximately 1400 base pairs length carried the DNA of interest: part of the first gene of the gal-operon, galE, and part of the IS2 element (fig. 1). The HindIII fragment mixtures of the mutant phage DNAs were used in separate experiments for cloning into the HindIII site of plasmid pBR322. pBR322 plasmids with an integrated galE-IS2 DNA fragment were identified in minilysates, the DNA of which was restricted with HindIII and run on agarose gels as described by Birnboim and Doly (6). The galE-IS2 fragments of the different mutants were then further restricted with HpaII + TaqI and HhaI + HpaII enzymes respectively for sequence analysis as outlined in fig. 3.

DNA sequence analysis was performed as described by Maxam and Gilbert (7).

Restriction enzymes were purchased from Boehringer (Mannheim), polynucleotide kinase from P.L. Biochemicals (Milwaukee) and γ-32P ATP from the Radiochemical Center (Amersham).

RESULTS

In mutant galOP-308 an IS2 element in orientation I is integrated within the galactose-leader sequence between the galactose promotor and galE, the first structural gene of this operon. Due to the strong polarity of IS2 all three genes of the gal-operon are inactivated almost completely (5).

We isolated highly constitutive and unstable revertants of mutant galOP-308 on a Ø80dgal phage. Reversion to the constitutive gal^+ phenotype was accompanied by deletion of most of the strongly polar IS2 sequence and adjacent sequences, including the gal-promoter-operator region (4, 8). The revertants were unstable and segregated gal^- cells as well as gal^+ cells of slightly different phenotype as compared to the parental gal^+ revertant. In strain 308^c-1-0 (fig. 1) a deletion had fused the galactose genes and a residual part of IS2-I to an IS2-II which was a constituent of the Ø80dgal phage we used, and which was derived originally from the F-episome (9). DNA heteroduplex analysis revealed that small sequence rearrangements at the deletion endpoint within IS2 caused the segregation phenomenon (8). One gal^+ and one gal^- segregant from revertant 308^c-1-0 were analyzed in more detail (fig. 2).

The two gal^+ strains differ in their capacity to synthesize galactose enzymes. Whereas the parental gal^+ revertant 308^c-1-0 synthesized 200% of galactokinase compared to 100% of a fully expressed gal^+ wild-type culture, its gal^+ segregant 308^c-1-1 synthesized 100%. The gal^- segregant 308^c-1-2- synthesized only 5%.

We sequenced the region of the deletion endpoint of the parental gal^+ re-

Fig. 1: Structures of Ø80dgal-308 and revertant 308^{C}-1-0. Ø80 sequences are drawn as straight lines, bacterial chromosomal sequences as wavy lines, IS2 sequences are indicated by heavy straight lines and sequences derived from the F-episome as dashed lines. Deleted sequences are indicated by boxes. O, P, E, T, K refer to operator and promoter of the gal-operon and the structural genes for UDPG epimerase, uridyltransferase and galactokinase respectively.

vertant and the two segregants and the result is presented in fig. 3:

In the parental gal^{+} revertant a deletion which had its one endpoint at nucleotide 1263 of IS2-I of mutant galOP-308 had fused the galactose genes to the adjacent IS2-II, leaving this sequence intact.

The sequences of the gal^{+} as well as of the gal^{-} segregants differ from the parental sequences in

i) that in both cases additional and identical 13 nucleotides of the residual IS2-I sequence were deleted, but leaving the adjacent IS2-II sequence

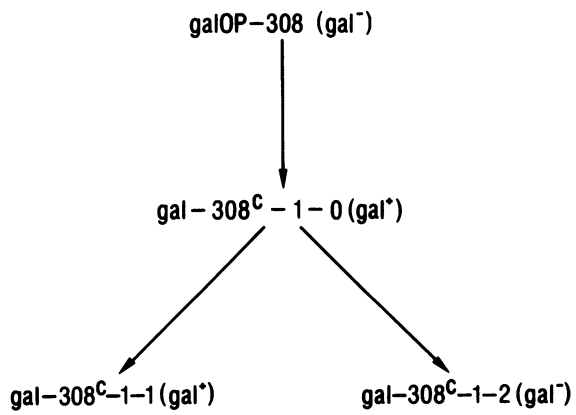

Fig. 2: Derivation of gal^{+} and gal^{-} strains from mutant galOP-308 on the Ø80dgal phage.

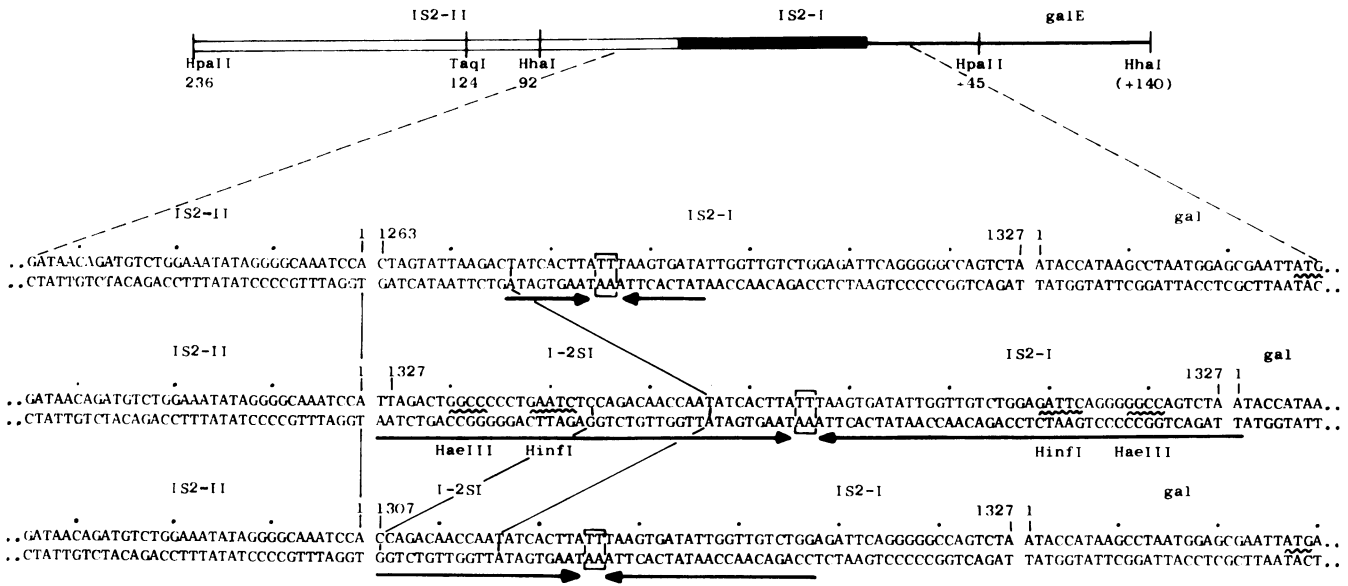

Fig. 3: DNA sequence at the deletion endpoints within IS2-I of the parental gal$^+$ revertant 308^c-1-0 (second line), its gal$^+$ segregant 308^c-1-1 (third line) and its gal$^-$ segregant 308^c-1-2 (fourth line). Nucleotide numbers of IS2 and galE are according to Ghosal et al. (10). Arrows designate regions of twofold symmetry. The endpoints of homologous sequences of the three strains are connected by thin lines. The restriction fragment used for DNA sequencing is shown in the first line. A Hpa II site in galE and a Hha I site in IS2-II were used respectively for sequence analysis of the two DNA strands.

again intact, and

ii) that new sequences were synthesized which in both cases are inverted copies of the residual part of IS2-I. In the gal$^+$ segregant this inverted copy is longer than in the gal$^-$ segregant, it comprises the complete length of the residual IS2-I sequence plus the first nucleotide of the galactose-leader sequence, and is 42 base pairs long. In the gal$^-$ segregant the inverted copy is equivalent to only 1/2 of the residual part of IS2-I, and is 21 base pairs long. Both segregants thus are characterized by long regions of complete two-fold symmetry which are not present in the parental DNA.

DISCUSSION

The deletion which caused the reversion from the gal$^-$ phenotype of mutant galOP-308 to the gal$^+$ phenotype of revertant 308^c-1-0- has removed most of the strongly polar IS2-I sequence and has fused the gal-operon to an adjacent IS2-II sequence. DNA sequence analysis (fig. 3) reveals that the deletion most probably did not occur at random, but preferred sites:

i) The deletion endpoint within IS2-I at nucleotide 1263 is in an A-T rich palindromic sequence of IS2 (10). For reasons discussed below we think it to be likely that this A-T rich palindrome between nucleotides 1250 and 1268 of IS2 can denature to form a "Gierer" tree which might serve as a substrate for an endonuclease and thereby initiate the deletion. The same A-T rich palindrome of IS2 apparently acted as the initiation signal for the sequence rearrangements of the mini-insertion IS2-6 (1).

ii) The other end of the deletion is at the first nucleotide of the adjacent IS2-II. Deletions terminating at the very ends of insertion sequences have been described repeatedly and are generally thought to be the result of a transposition event of the insertion element (11, 12). We do not think, however, that the deletions occurring in our mutants were created simultaneously with a transposition of the IS2 element. We rather believe that the deletogeneous process which is initiated at an A-T rich palindromic sequence outside of IS2-II stops preferentially at the very terminus of IS2-II. Alternatively, the deletion starts at the first nucleotide of IS2 and stops preferentially at the A-T rich palindrome outside of IS2-II. The finding that deletions adjacent to the terminus of IS2 can be created even if only a small part of IS2 is present (13), which most probably cannot be transposed, support this interpretations.

The segregation of gal$^-$ and gal$^+$ cells from the parental gal$^+$ revertant apparently results from two processes: deletion and synthesis of DNA. We

think that these two processes were initiated simultaneously by one event:

The sequence of the parental gal$^+$ revertant 308^c-1-0 carries already a rather long and A-T rich palindrome near the deletion endpoint between nucleotides 1276 and 1295 of IS2-I (fig. 3). We assume that the initial step both for deletion and for synthesis of DNA was a nick at some site beyond the axis of symmetry of this palindrome, followed by denaturation and folding back of DNA (fig. 4). The folded back DNA could serve as a primer for DNA repair synthesis to the left with respect to fig. 4, at the same time nucleotides were deleted to right. This latter process terminates preferentially at the terminal nucleotide of IS2-II. In a final step the newly formed DNA ends were ligated. After one round of replication and segregation the mutant DNA structures would be stable.

An interesting variation of this interpretation was put forward by one of the referees of this paper: To account for the fact that the deletion pro-

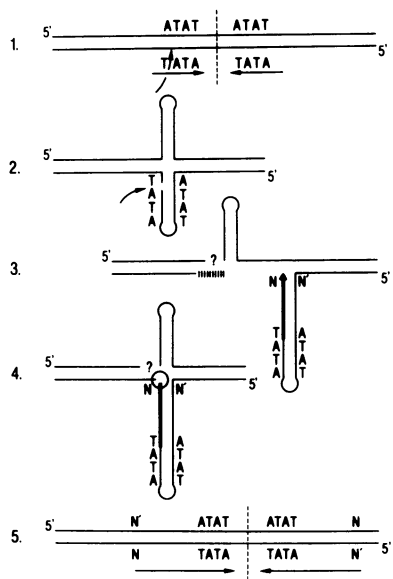

Fig. 4: Model for the creation of long palindromic DNA by repair synthesis.
1. Nicking of one DNA strand within the region of twofold symmetry.
2. Denaturation and folding back of DNA.
3. DNA synthesis by strand displacement and deletion of DNA (the question mark in steps 3. and 4. stands for our ignorance of the deletion mechanism and of the fate of the complementary nonnicked DNA strand).
4. Ligation of the newly formed DNA ends.
5. Stabilization of the long palindromic region by replication of the lower DNA strand.

cess in mutants 308^{c}-1-1 and 308^{c}-1-2 exactly terminates at the first nucleotide of IS2-II, it was proposed that the initial nicking event occurs at the terminus of IS2-II by a site specific endonuclease. The 13 nucleotides between the end of IS2-II and the beginning of the region of twofold symmetry are then trimmed back by the 3'exonuclease activity of DNA polymerase I, before the same enzyme enlarges the hairpin duplex by displacement synthesis. This variant model has the advantage for needing only one enzyme to explain deletion and synthesis of DNA, but assumes the initiation event to be a very specific one which is not necessary to understand the model.

Three observations support the model for the creation of long palindromic DNA by repair synthesis:

i) The palindromic sequence of IS2 at position 1276-1295 can indeed denature even without introducing a nick. The HinfI and HaeIII restriction fragments respectively of revertant 308^{c}-1-1 (fig. 3, third line) which are complete palindromes and contain the A-T rich palindrome of IS2 as their central part, denature completely by branch migration at $37^{o}C$ to give two pieces of folded back DNA with half the length of the original DNA piece (14).

ii) Unstable revertants (gal^{+} strains which segregate gal^{-} cells with high frequency) of mutant galOP-308 can be induced by mutagens like 1-methyl-3-nitro-1-nitrosoguanidine and UV. This indicates that DNA repair mechanisms may be involved in the creation of this class of revertants. This conclusion is strengthened by the observation that the fraction of unstable revertants among the gal^{+} revertants of mutant galOP-308 decreases by an order of magnitude if an uvrB cell, which lacks an endonuclease that recognizes DNA mispairing, was mutagenized by UV (15). The participation of enzymes of the DNA excision repair mechanism was also observed in the segregation of gal^{-} cells from unstable revertants of the galOP-3::IS2-I mutant of E. coli (16).

iii) The sequence rearrangements within IS2 which have been described by Ghosal and Saedler (1) and Ghosal et al. (2) can be explained by the DNA repair mechanism outlined above. Since in these revertants deletion of DNA is not observed, DNA synthesis must have been terminated by a template slippage of DNA polymerase. This mechanism was proposed by the authors to occur within the replication fork during replication of the IS2 sequence. Strand switches, however, are known only for polymerase I and not for polymerase III, the replication enzyme (17-20). This again supports our model which needs only polymerase I as the DNA synthesizing enzyme.

The DNA sequence where according to our hypothesis the series of events that lead to the formation of long palindromic DNA is initiated, is in itself

a rather long A-T rich palindrome. The existence of such palindromes in IS2 might explain why sequence rearrangements are found relatively frequently in IS2. However, the principle we envisage as a basis for the creation of long DNA sequences with twofold symmetry does not necessarily need such a long palindrome to start with. Our hypothesis of the events is in fact an extension of Streisinger's model for the creation of frameshift mutations. In this model small direct and A-T rich repeats are thought to initiate mispairing between complementary DNA strands. Mispairing would then be stabilized by DNA nicking and DNA repair synthesis (21). Here we propose that the same can occur in single stranded DNA at small inverted and A-T rich sequence repeats with the consequence that long palindromes are synthesized. The mechanism for the creation of DNA with long regions of twofold symmetry therefore might be general and not restricted to the IS2 element. Such a mechanism may well have played a role in the evolution of genetic signals and the evolution of oligomeric DNA binding proteins.

The promoter at which transcription of the galactose genes in our revertants is initiated has not been mapped up to now. It has been postulated that IS2 carries a strong promoter in orientation II (22). However, our data are not compatible with this hypothesis: Although in all three strains derived from mutant galOP-308 (fig. 2) the galactose genes are fused to the same IS2-II, only in two of them the galactose genes are transcribed efficiently enough to give a gal^+ phenotype. Fusion of genes to IS2-II therefore seems not to be sufficient to activate them. One may argue that the IS2-II elements differ from each other in the three strains by mutations. Indeed, IS2 elements with promoter activity both in orientation I (1, 2) and in orientation II (8) are known. They are characterized by more complex sequence rearrangements within IS2, e.g. secondary gal^+ revertants of the gal^- segregant 308^c-1-2 (fig. 2) were caused by duplication of the IS2-II sequence and by integration of another transposable DNA element into IS2-II respectively. However, by using different restriction enzymes and nuclease S1 treatment of heteroduplex DNA no differences between the IS2-II sequences of the two gal^+ and the gal^- strains of fig. 2 were seen (our unpublished experiments). Whether IS2-II serves as a promoter or not therefore may not rely on sequences of IS2, but on sequences surrounding its integration site on the chromosome. Data which support this assumption will be discussed elsewhere (23).

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft through

SFB 74 and the European Organization of Molecular Biology. D. Ch. was holder of a long term EMBO fellowship in Cologne. We thank one of the referees of this paper for letting us know his explanation for the creation of palindromic DNA.

*Correspondence should be addressed to Jürgen Besemer,
Sandoz Forschungsinstitut, Brunner Strasse 59, A-1235 Wien, Austria.

REFERENCES

1. Ghosal, D., Saedler, H. (1978) Nature 275, 611-617.
2. Ghosal, D., Gross, J., Saedler, H. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 1193-1196.
3. Miller, J.H.(1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y. USA.
4. Delius, H., Charlier, D., Besemer, J. (1980) Molec. Gen. Genet., 179, 391-397.
5. Besemer, J., Herpers, M. (1977) Molec. Gen. Genet. 151, 295-304.
6. Birnboim, H.C., Doly, J. (1979) Nucl.Acids Res. 7, 1513-1523.
7. Maxam, D., Gilbert, W. (1977) Proc. Natl Acad. Sci. USA 74, 560-564.
8. Charlier, D., Besemer, J., manuscript in preparation.
9. Besemer, J., Kubau, D. (1976) Molec. Gen. Genet. 148, 79-92.
10. Ghosal, D., Sommer, H., Saedler, H. (1979) Nucl.Acids Res. 6, 1111-1122.
11. Shapiro, J. (1979) Proc. Natl Acad. Sci. USA 76, 1933-1937.
12. Arthur, A., Sherratt, D. (1979) Molec. Gen. Genet. 175, 267-274.
13. Ahmed, A., Scraba, D. (1978) Molec. Gen. Genet. 163, 189-196.
14. Besemer, J., Görtz, G., manuscript in preparation.
15. Besemer, J., unpublished experiments.
16. Ahmed, A., Johansen, E. (1978) J. Mol. Biol. 121, 269-281.
17. Schildkraut, C.L., Richardson, Ch.C., Kornberg, A. (1964) J. Mol. Biol. 9, 24-45.
18. Masamune, Y., Richardson, Ch.C. (1971) J. Biol. Chem. 246, 2692-2701.
19. Bolivar, F., Betlach, M.C., Heyneker, H.L., Shine, J., Rodrigues, S.L., Boyer, H.W. (1977) Proc. Natl Acad. Sci. USA 74, 5265-5269.
20. Backman, K., Betlach, M., Boyer H.W. Yanofsky, S. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 69-76.
21. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., Inouye, M. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 77-84.
22. Saedler, H., Reif, H-J., Hu, S., Davidson, N. (1974) Molec. Gen. Genet. 132, 265-289.
23. Charlier, D., Besemer, J., manuscript in preparation.