From: EXTRACHROMOSOMAL ELEMENTS IN LOWER EUKARYOTES
Edited by Reed B. Wickner, Alan Hinnebusch,
Alan M. Lambowitz, I. C. Gunsalus and Alexander Hollaender
(Plenum Publishing Corporation, 1986)

SITE-SPECIFIC RECOMBINATION PROMOTED IN VITRO BY THE FLP PROTEIN OF THE YEAST TWO-MICRON PLASMID

Julie F. Senecoff, Robert C. Bruckner, Leslie Meyer-Leon,
Cynthia A. Gates, Elizabeth Wood, Scott W. Umlauf,
Janet M. Attwood, and Michael M. Cox

Department of Biochemistry
University of Wisconsin
Madison, Wisconsin 53706

INTRODUCTION

The yeast 2-micron plasmid is a 6,318-bp circular DNA molecule occurring in many yeast strains. It is generally present at 50 to 100 copies per cell and has been sequenced in its entirety (11). The plasmid encodes at least four protein products (Ref. 6; Volkert et al., this Volume), which are involved primarily in plasmid maintenance. One of these proteins, designated FLP, is the central component in a plasmid-encoded, site-specific recombination system. The sequence of the open reading frame (ORF) encoding this protein provides a predicted molecular weight of slightly less than 48,000 Da. This recombinase acts at a pair of sites located within the 599-bp inverted repeats. Recombination results in the inversion of the DNA between the sites, giving two forms of the plasmid, termed A and B (see Fig. 1). Isolated plasmid DNA is found to be a 50:50 mixture of the two forms. Inversion has been shown to be part of a copy-number amplification system (Volkert et al., this Volume). The general features of the plasmid and its maintenance systems appear to be shared by plasmids isolated from a variety of related organisms (A. Toh-e et al., this Volume).

Site-specific recombination systems provide a unique opportunity to study virtually every aspect of the chemistry involved in nucleic acid metabolism and protein-nucleic acid interactions. In each case, one enzyme or enzyme system recognizes and binds to specific sequences, juxtaposes two of these sequences, and finally breaks and reforms phosphodiester bonds in a pattern that results in specific topographical and topological changes in the DNA. Extensive work has been carried out on a variety of prokaryotic site-specific recombination systems in vitro (4,5,12,14,16,18,23). These studies have served, in part, to highlight the great variation to be found from one system to the next in terms of mechanism, specificity, requirements, etc. These facts, coupled to the important role site-specific recombination plays in a variety of biological processes in higher cells, provided ample motivation to extend these studies to eukaryotes. The recombination system encoded by the 2-micron plasmid represents the most

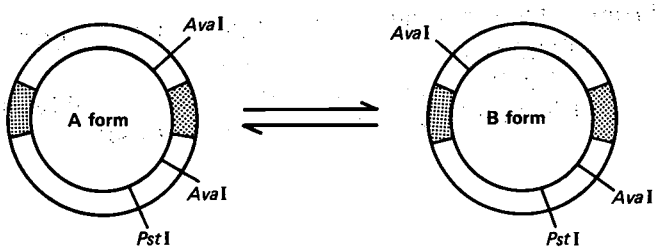

Fig. 1. Two forms of the 2-micron plasmid resulting from site-specific recombination. (From Ref. 8, with permission.)

accessible eukaryotic site-specific recombination system presently available for biochemical analysis.

To avoid problems related to proteolysis often encountered in yeast extracts, and to provide a potentially enriched source of FLP protein, several laboratories have cloned and expressed the FLP gene in Escherichia coli (8,13,22). This effectively transfers the entire recombination system to the bacteria. If appropriate substrates are provided in vivo, the FLP protein has been shown to be the only yeast protein required to promote this site-specific recombination event in E. coli (8). The properties and requirements of the reaction in bacteria are consistent with observations made in yeast (8).

Expression in E. coli led rapidly to the establishment of in vitro systems for this recombination event and elucidation of the general properties of the reaction (17,19). The FLP protein promotes all three types of site-specific recombination in vitro--deletion, inversion, or integration--depending upon the substrate employed. Reaction requirements are simple, including only a buffer and the appropriate ionic strength. Divalent cations and high energy co-factors are not needed. Both linear and supercoiled DNA molecules with recombination sites serve as substrates. These systems have been employed in the studies described below. Partially purified FLP protein and assays employed in this work have been described previously (7,20,21).

RESULTS AND DISCUSSION

The FLP Recombination Site

To date, the area of greatest progress has been the characterization of the recombination site recognized by the FLP protein. This site was shown to involve no more than 60 bp of the 599-bp repeats in early work carried out in vivo in yeast by Broach and colleagues (6). This work also demonstrated that destruction of an XbaI restriction site within this region abolished recombination site function. Further characterization has been carried out primarily in vitro.

Minimal size of the recombination site. The most prominent DNA structural features found in the 60-bp region defined by Broach et al. (6) are illustrated in Fig. 2. Two inverted 13-bp repeats are found, and these are

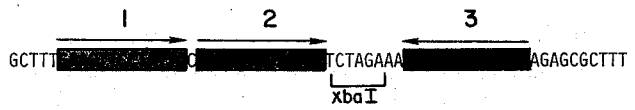

Fig. 2. Structural features of the FLP recombination site. (From Ref. 21, with permission.)

separated by an 8-bp spacer. A third copy of this 13-bp repeat is immediately adjacent, in tandem with and separated by 1 bp from one of the other two repeats. The XbaI site is present within the spacer.

The third 13-bp repeat (designated "1" in Fig. 2) has been shown to be unnecessary for FLP-mediated recombination in vitro (2,21). This repeat is also not required in vivo, although some modest effects on the reaction have been observed (13). As noted below, FLP protein binds to this dispensable repeat and protects it from methylation and nuclease digestion (2,7). The function of this sequence is still unclear.

Recombination in vitro also does not require all 13 bp of the other two repeats. Two base pairs can be deleted from either end without affecting any known contact points with FLP protein (7). An additional base pair can be deleted on either side without affecting the efficiency of recombination. Further deletion results in decreased levels of recombination. Detectable recombination is observed with deletions of 5 or 6 bp (9,21). A deletion of 8 bp abolishes the reaction (21). The minimal site that is fully functional is therefore 28 bp in length, including the internal 10 bp of each of the two 13-bp repeats. Detectable recombination requires at least the central 20 bp of this sequence. A summary of the features of this recombination site is provided in Fig. 3.

The crossover points. The FLP protein cleaves this site at the boundaries of the 8-bp spacer, as shown in Fig. 3 (2,21). Staggered cuts are made, and the protein becomes covalently linked to the DNA via a 3' phosphate. The resulting 8-nucleotide overhangs terminate at a 5' hydroxyl group. Gronostajski and Sadowski (10) have demonstrated that the protein is covalently linked at a tyrosine residue.

The spacer: Size and homology requirements. Before the recombination site was defined, it was clear that addition or deletion of 4 bp at the XbaI site abolished recombination (6,8). This suggested that the size of

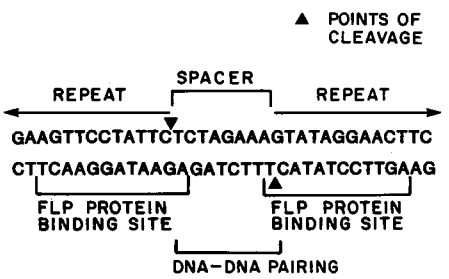

Fig. 3. Function of sequences within the minimal FLP recombination site.

the spacer was important. A series of spacer mutations was constructed by J. Senecoff (20,21), in which the spacer was increased or decreased in size by 1 or 2 bp. The efficiency of recombination was reduced but not abolished when spacer size was altered by ±1 bp. This was especially apparent when mutant sites were reacted with normal sites. The efficiency of recombination was partially restored when the reaction was restricted to sites with identical mutant spacers (20,21). Restoration of reaction efficiency involved homology rather than spacer size, since no reaction could be detected between sites with different 1-bp spaced insertions (20). Increasing the size of the spacer by 2 bp abolishes recombination (20). This work confirmed that the size of the spacer is important in this recombination system. The partial restoration of reaction efficiency when identical mutants were employed further suggested that homology between the spacers of the two sites undergoing reaction might also be required.

A series of point mutations was therefore constructed to define the functions of the spacer and the adjacent repeats. A substitution mutation within the internal 6 bp of the spacer again reduced the efficiency of recombination in reactions with normal sites. In this case the efficiency of the reaction was fully restored when sites with identical mutant spacers were used (1,20). Sites with spacers singly altered at four of the positions provide identical results (S. Umlauf, unpubl. data). The sequence at these six positions is therefore irrelevant. As long as spacer size is maintained, any sequence is tolerated, indicating that FLP protein does not recognize these base pairs. The only requirement is that this region be homologous in the two sites undergoing recombination.

This homology requirement implies that DNA-DNA pairing between the spacer sequence of two sites occurs at some point during the recombination event. Similar results have been obtained in several other site-specific recombination systems (4,12,14,23). This pairing could occur either before or after DNA cleavage. If pairing occurs before cleavage, a four-strand intermediate could exist, such as that described by Kikuchi and Nash (15). Such an intermediate could function in the proper alignment of two sites during recombination. At present we favor the idea that pairing precedes cleavage, although the only evidence available is a lack of an increase in the rate of unproductive cleavage observed when sites with nonhomologous spacers are mixed.

A mutation was also obtained in the spacer position immediately adjacent to the cleavage site. Recombination was again reduced when this mutant was reacted with a normal site. When identical mutants were reacted, however, only a partial restoration of recombination efficiency was observed (20). Subsequent studies have shown that this is a position of functional overlap. This base pair is within the spacer and exhibits the homology requirement observed at the other spacer positions. It is also a contact point recognized by FLP protein, however, and restoration of homology does not entirely overcome the defect in FLP protein recognition (see below).

The pattern changes abruptly for a mutation on the repeat side of the repeat-spacer junction (20). Here no homology dependence is observed. Recombination efficiency is reduced in a reaction with a normal site, and reduced still further in reactions between two identical mutant sites. This implies that this position functions as part of the FLP protein binding site, a conclusion confirmed in studies described below.

Spacer homology and reaction directionality. The homology requirement exhibited by the central 6 bp of the spacer is also relevant in predicting the outcome of a recombination event. The FLP recombination site is asymmetric, as are the recombination sites of all characterized site-specific systems. This asymmetry is used by the recombinase to align the sites in the same orientation during recombination. This, in turn, determines the outcome or directionality of the reaction. In the FLP site, there are potentially three elements of asymmetry: the third 13-bp repeat (already shown to be unnecessary for recombination), a 1-bp mismatch in the two 13-bp repeats immediately flanking the spacer (the second base pair from the spacer), and the spacer itself. Eliminating the mismatch in the flanking repeats has no effect on the efficiency or directionality of recombination (21), so the effective determinants of asymmetry in this site lie entirely within the spacer. The DNA-DNA pairing described above, therefore, may also be responsible for the directionality of recombination.

In order to test this idea, a recombination site was constructed in which the spacer was completely symmetrical. The XbaI restriction site was moved to the center of the spacer, effectively changing the sequence at five positions. The base pairs at either end of the spacer were unaltered. This site was found to be completely functional in FLP protein-promoted site-specific recombination (20). As expected, recombination was detected only between sites with identical mutant spacers. No recombination was detected in reactions between this site and a normal one. The most important result, however, is that recombination between mutant sites no longer exhibited directionality. In intermolecular reactions, recombination products were detected which could only arise from the juxtaposition of sites, during recombination, in an alignment which is found to be forbidden in sites with asymmetric spacers. No bias was found in the formation of products from different substrate alignments. This demonstrates that DNA-DNA pairing, at whatever stage of the recombination reaction it occurs, is responsible for the directionality in this system (20). Similar results have recently been found in the cre/lox system of bacteriophage P1 (12).

The FLP protein binding sites. It is clear from the evidence described above that FLP protein does not bind specifically to at least the central 6 bp of the spacer. This leaves the flanking repeats as potential binding sites for FLP protein. At least two mutations in these repeats have been found to lower the efficiency of recombination (1,20). Recombination efficiency is not restored in these cases when two identical mutant sites are employed in the reaction. Additional evidence for FLP recognition of sequences in these repeats is derived from an examination of the effect of 1-bp insertions or deletions in the spacer on the position at which FLP protein cleaves the site. In each case the point of cleavage is at the boundary of the altered spacer, with cuts staggered by 7 bp and 9 bp for deletions and insertions, respectively. The cleavage site remains constant relative to the flanking repeats.

Building on this evidence, the binding site has been mapped in detail by R. Bruckner, utilizing a variety of techniques (7). The results are summarized in Fig. 4. Methylation protection experiments have been employed to define purine contacts. These serve to define the limits of the binding site. Purine contacts extend through a 12-bp sequence on either side of the spacer. In both cases, the first purine contact is the external base pair in the spacer (this is the position described above as an overlap position), and the site includes the adjacent 11 bp in the repeat. The guanine contacts defined by this technique are located in the major

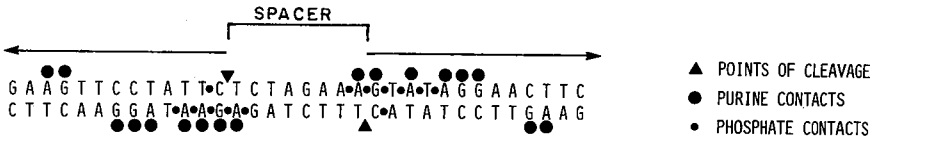

Fig. 4. Identified points of contact between the FLP protein and its recombination site.

groove of the DNA, whereas the adenine contacts are located in the minor groove. When superimposed on a helical representation of the DNA, most of these contacts (in both repeats) are found on a single face of the helix (7). Purines within the central 6 bp of the spacer are not protected from methylation when FLP protein is bound to the site (7). The spacer region, however, is protected from nuclease digestion by FLP protein binding (2). This suggests that FLP protein spans the spacer region without actually binding to it. These results, coupled to the effects of the spacer size mutations described above, and other results (see Andrews et al., this Volume), indicate an interaction between FLP protomers bound at each of the binding sites flanking the spacer. At positions immediately adjacent to the 12-bp binding sites defined above, methylation of purines is found to be enhanced as a result of FLP protein binding (7).

In the course of these experiments, it was found that the third 13-bp repeat exhibited a pattern of methylation protection in the presence of FLP protein identical to that observed for the other repeats (7). This is part of the evidence that this repeat, although clearly not required for recombination, is bound by FLP protein (2). More work is necessary to determine if this third binding site serves a function in vivo or in vitro.

The significance of the contacts identified by methylation protection can be analyzed by methylation interference experiments. Here, the substrate is partially methylated before FLP protein is added to determine if methylation at a given position interferes with recombination. If methylation at a position blocks recombination, then DNA molecules methylated at that position will be under-represented in the products. The results obtained for guanine methylation are completely consistent with the methylation protection experiments and confirm that all of the identified guanine contacts are functionally significant (7). Methylation of guanines within the central 6 bp of the spacer or outside the 12-bp binding sites has no effect on recombination. A similar set of experiments, in which phosphate groups were ethylated before FLP protein was added, was used to define the phosphate contacts within the binding site. All of the phosphate contacts were clustered around the points at which FLP protein cleaves the site (Fig. 4). When superimposed on a helical DNA representation, the phosphate contacts are found to lie on the same face of the helix as the majority of the purine contacts (7).

The interference experiments employed the recombination site that contained the symmetrical spacer described above. Use of this site simplifies a variety of experiments (7,20). By several criteria, including guanine contacts identified in methylation protection experiments, the rate at which recombination proceeds, and the point at which the site is cleaved by FLP protein, this site is identical to a normal site in terms of its interaction with FLP protein (7).

All of these results are consistent with the mutational analysis and provide a fairly detailed picture of the function of subsets of sequences within the recombination site. The central 6 bp of the spacer are involved in DNA-DNA pairing at some point during the recombination reaction. Sequence is unimportant at these positions, although homology must exist between the sites undergoing reaction. This pairing is also responsible for correct site alignment during the reaction. The external base pairs in the spacer, adjacent to the cleavage sites, also participate in this DNA-DNA pairing. These positions are also part of the FLP protein binding site, however, and mutations here have a deleterious effect on recombination efficiency. The FLP binding site also includes the 11 bp immediately adjacent to the spacer. No contacts are found in the external two base pairs of the repeats. The deletion analysis demonstrates that these two base pairs can be removed without affecting recombination efficiency. The external base pair within the defined binding site can also be deleted without apparent effects on the efficiency of recombination. Further deletions result in a progressive decline in site function. Recombination is abolished only when half of the defined binding site has been removed.

One of the most interesting aspects of this analysis is the flexibility evident in the spacer region sequence. The properties of this sequence indicate that it is possible, in principle, to construct thousands of different FLP recombination sites. Each of these would be fully functional as long as the spacer size was maintained at 8 bp. Each site with a different spacer, however, would be able to react only with an identical site. To permit future exploitation of this feature of the system, we have begun to construct sites with extensive spacer alterations. One, in which the XbaI site has been replaced with a SnaBl site (the sequence is altered at four of the internal six base pairs of the spacer), has already been shown to be functional.

The FLP Protein

Purification of the FLP protein from extracts of *E. coli* in which the protein has been expressed has been hampered by several factors. The levels of expression in this host have been lower than those achieved for a number of bacterial proteins utilizing identical promoters and transcription and translation signals. Solubility and stability problems have been encountered in fractions derived from a number of procedures. Evidence has been obtained that FLP protein is proteolytically degraded to a significant extent in *E. coli*. In addition, it has been determined that one or more proteins in *E. coli* extracts enhance the activity of FLP protein, resulting in large apparent losses in activity when FLP protein is purified away from them.

Some success has been achieved, however. Protocols resulting in active and partially purified FLP protein have been published (3,17,19). We have recently succeeded in purifying the protein to near-homogeneity (L. Meyer-Leon, C. Gates, and M.M. Cox, unpubl. data). The procedure, which employs a combination of cation exchange and affinity chromatography, can be completed in 48 hr and yields more than a milligram of pure FLP protein from 100 grams of cell paste. Efficient recombination in vitro is achieved at levels of protein approximately stoichiometric with the concentration of recombination sites in a reaction. The availability of pure FLP protein and a well-characterized recombination site should now permit a more detailed analysis of this recombination system.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health (NIH) Grant GM32335. L.M. and S.U. were supported by NIH Training Grant 5-T32 GM07215. M.M.C. is the recipient of NIH Research Career Development Award AI00599.

REFERENCES

1. Andrews, B.J., M. McLeod, J. Broach, and P.D. Sadowski (1986) Interaction of the FLP recombinase of the yeast 2-micron plasmid with mutated target sequences. Mol. Cell. Biol. (in press).
2. Andrews, B.J., G.A. Proteau, L.G. Beatty, and P.D. Sadowski (1985) The FLP recombinase of the 2-micron circle DNA of yeast: Interaction with its target sequences. Cell 40:795-803.
3. Babineau, D., D. Vetter, B.J. Andrews, R.M. Gronostajski, G.A. Proteau, L.G. Beatty, and P.D. Sadowski (1985) The FLP protein of the 2-micron plasmid of yeast: Purification of the protein from Escherichia coli cells expressing the cloned FLP gene. J. Biol. Chem. 260:12313-12319.
4. Bauer, C.E., S.D. Hesse, J.F. Gardner, and R.I. Gumport (1984) DNA interactions during bacteriophage λ site-specific recombination. Cold Spring Harbor Symp. Quant. Biol. 49:699-705.
5. Benjamin, H.W., M.M. Matzuk, M.A. Krasnow, and N.R. Cozzarelli (1985) Recombination site selection by Tn3 resolvase: Topological tests of a tracking mechanism. Cell 49:147-158.
6. Broach, J.R., V.R. Guarascio, and M. Jayaram (1982) Recombination with the yeast 2 micron plasmid is site-specific. Cell 29:227-234.
7. Bruckner, R.C., and M.M. Cox (1986) Specific contacts between the FLP protein of the yeast 2 micron plasmid and its recombination site. J. Biol. Chem. (in press).
8. Cox, M.M. (1983) The FLP protein of the yeast 2 micron plasmid: Expression of a eukaryotic genetic recombination system in Escherichia coli. Proc. Natl. Acad. Sci., USA 80:4223-4227.
9. Gronastajski, R.M., and P.D. Sadowski (1985) Determination of DNA sequences essential for FLP-mediated recombination by a novel method. J. Biol. Chem. 260:12320-12327.
10. Gronastajski, R.M., and P.D. Sadowski (1985) The FLP recombinase of the yeast 2-micron plasmid attaches covalently to DNA via a phosphotyrosyl linkage. Mol. Cell. Biol. 5:3274-3279.
11. Hartley, J.L., and J.E. Donelson (1980) Nucleotide sequence of the yeast plasmid. Nature (London) 268:860-864.
12. Hoess, R.H., A. Wierzbicki, and K.A. Abremski (1986) The role of the loxP spacer region in P1 site-specific recombination. Nucl. Acids Res. 14:2287-2300.
13. Jayaram, M. (1985) Two-micrometer circle site-specific recombination: The minimal substrate and the possible role of flanking sequences. Proc. Natl. Acad. Sci., USA 82:5875-5879.
14. Johnson, R.C., and M.I. Simon (1985) Hin-mediated site-specific recombination requires two 26 base pair recombination sites and a 60 base pair recombinational enhancer. Cell 41:784-791.
15. Kikuchi, Y., and H. Nash (1979) Nicking-closing activity associated with bacteriophage λ int gene product. Proc. Natl. Acad. Sci., USA 76:3760-3764.
16. Kitts, P., E. Richet, and H. Nash (1984) λ Integrative recombination; supercoiling, synapsis, and strand exchange. Cold Spring Harbor Symp. Quant. Biol. 49:735-744.

17. Meyer-Leon, L., J.F. Senecoff, R.C. Bruckner, and M.M. Cox (1984) Site-specific recombination promoted by the FLP protein of the yeast 2-micron plasmid *in vitro*. *Cold Spring Harbor Symp. Quant. Biol.* 49:797-804.
18. Reed, R.R., and N.D.F. Grindley (1981) Transposon-mediated site-specific recombination: A defined *in vitro* system. *Cell* 25:713-719.
19. Sadowski, P.D., D.D. Lee, B.J. Andrews, D. Babineau, L. Beatty, M.J. Morse, G.A. Proteau, and D. Vetter (1984) *In vitro* systems for the genetic recombination of the DNAs of bacteriophage T7 and yeast 2-micron circle. *Cold Spring Harbor Symp. Quant. Biol.* 49:789-796.
20. Senecoff, J.F., and M.M. Cox (1986) Directionality in FLP protein-promoted site-specific recombination is mediated by DNA-DNA pairing. *J. Biol. Chem.* 261:7380-7386.
21. Senecoff, J.F., R.C. Bruckner, and M.M. Cox (1985) The FLP recombinase of the yeast 2-micron plasmid: Characterization of its recombination site. *Proc. Natl. Acad. Sci., USA* 82:7270-7274.
22. Vetter, D., B.J. Andrews, L. Roberts-Beatty, and P.D. Sadowski (1983) Site-specific recombination of the yeast 2 micron DNA *in vitro*. *Proc. Natl. Acad. Sci., USA* 80:7284-7288.
23. Weisberg, R.A., L.W. Enquist, C. Foeller, and A. Landy (1983) Role for DNA homology in site-specific recombination. The isolation and characterization of a site affinity mutant of coliphage λ. *J. Mol. Biol.* 170:318-342.