The EMBO Journal vol.7 no.6 pp.1845 - 1852, 1988

The functional significance of DNA sequence structure in a site-specific genetic recombination reaction

Scott W.Umlauf and Michael M.Cox

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA

Communicated by P.Berg

A series of sequence changes in the spacer region of the FLP recombination target (FRT) site are presented which drastically reduce site function without affecting recognition by the FLP protein. The effects follow a pattern which indicates that two structural features of the FRT site are essential for site function: a pair of pyrimidine tracts arranged in a palindrome and a predominance of AT base pairs in the spacer. The FRT site represents a sequence that serves to facilitate unwinding of the DNA within the spacer region during recombination. The results provide a clear demonstration of a role for a DNA sequence element that is distinct from protein recognition.

Key words: DNA structure/DNA unwinding/site-specific recombination/2μ plasmid/yeast plasmid

Introduction

Recent work has revealed a number of structural variations in DNA. Among these are the well-characterized A and Z forms, which have been observed in crystal structures (Shakked *et al.*, 1981; Wang *et al.*, 1979). The bent structures formed by poly(A) sequences (Marini *et al.*, 1982; Burkhoff and Tulius, 1987; Laundon and Griffith, 1987) and the S1 sensitivity exhibited by polypurine tracts (Pulleyblank *et al.*, 1985; Kilpatrick *et al.*, 1986) are two prominent examples of local, sequence-dependent structural variations. Potential roles for these sequence elements in DNA include: facilitation of site recognition by a protein, and facilitation of site function. A role of sequence structure in protein binding has been documented in the case of histones (Drew and Travers, 1985). Similarly the binding of the bacteriophage λ O protein to the λ replication origin is dependent upon a sequence structure which facilitates DNA bending (Zahn and Blattner, 1987). Local structural variations might also serve as an initial scanning target for a protein to facilitate specific site location (Lomonossoff *et al.*, 1981; Husain *et al.*, 1987). A role for sequence structure in site function, independent of any effect on protein binding, is more difficult to demonstrate. Such a role might be anticipated for sites of replication initiation and recombination, which typically exhibit a high degree of non-random sequence structure. Drawing an unambiguous distinction between the function of a DNA site and its recognition by a protein has recently become a focus of our studies on yeast recombination.

The 2μ plasmid, found in most naturally occurring strains of *Saccharomyces cerevisiae*, encodes a site-specific recombinase that promotes an inversion of the DNA between two 599-bp inverted repeats (Broach, 1981). This inversion plays a central role in plasmid copy-number amplification (Futcher, 1986; Volkert and Broach, 1986). The FLP protein has been cloned and expressed in *Escherichia coli* and purified to near homogeneity (Cox, 1983; Vetter *et al.*, 1983; Gates *et al.*, 1987; Meyer-Leon *et al.*, 1987). This protein catalyzes site-specific recombination events with a turnover number of $\sim 0.3\ min^{-1}$ (Gates and Cox, 1988). The minimal recombination site (FRT) has been defined and contains two inverted 13-bp repeats surrounding an asymmetric 8-bp spacer, as shown in Figure 1 (Andrews *et al.*, 1985; Senecoff *et al.*, 1985). The FLP protein cleaves the site at the junctions of the repeats and the spacer and is covalently linked to the DNA via a 3′ phosphate. This cleavage leaves 8-nt overhangs with 5′-OH termini (Andrews *et al.*, 1985; Gronostajski and Sadowski, 1985). Thus, the spacer region is the site of cross-over. Methylation protection and mutagenesis experiments have defined two 12-bp FLP protein binding sites within the FRT site. Each includes 11 bp of each repeat but only the single outermost base of the spacer. The six central base pairs of the spacer are not contacted by FLP protein (Brückner and Cox, 1986; Senecoff *et al.*, 1988). Our emphasis in this report is on these 6 bp.

In vitro assays of FRT sites containing mutations at each of the central six spacer base pairs (Senecoff and Cox; 1986) suggest that the sequence of the spacer is unimportant. The spacers of two reacting FRT sites, however, must be homologous. Furthermore, the asymmetry of the wild-type spacer determines the direction of site alignment in the recombination event. When sites with symmetrical spacers are used, this directionality is lost (Senecoff and Cox, 1986).

In the process of generating new FRT sites with altered spacers we have discovered that some mutant spacer sequences can greatly influence the rate of reaction of *in vitro* recombination assays without affecting recognition of the site by FLP protein. The results provide information about general structural features of the spacer sequence which must be present in a functional FLP recombination site.

Results

Experimental design

We have generated a library of FLP recombination sites that differ in sequence in the central 6 bp of the spacer region. Since homology between reacting spacer sequences is required, plasmids containing altered spacers were reacted only with themselves in all cases. The oligonucleotide-directed



Fig. 1. Minimal FRT site.

©IRL Press Limited, Oxford, England

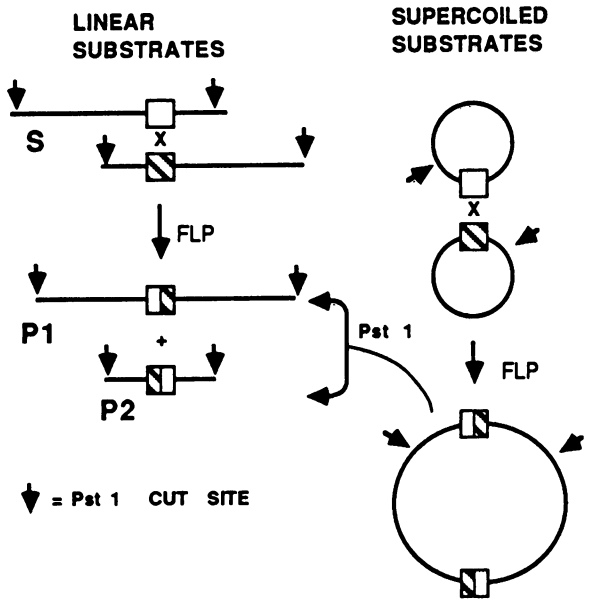


Fig. 2. *In vitro* assay for FLP protein-promoted recombination. One of the two standard recombination assays used in this study is illustrated. This approach is used for FRT sites with symmetrical spacers. Substrate DNA is linearized with the restriction enzyme *PstI*. Unique products are produced when substrates (S) are aligned in an anti-parallel orientation. The larger product is denoted P1, the smaller P2. The effect of supercoiling in the substrate can be examined by recombining circular substrates followed by digestion with *PstI*. A dimer is used to represent products in the supercoiled assay although the products formed are actually much larger circles resulting from multiple recombination events.

mutagenesis that we used permitted us to design sequences to answer specific structural questions.

The recombination assays used to compare mutant recombination sites are of two types: (i) FLP titration assays in which the concentration of FLP protein is varied and the extent of recombination at a fixed point is determined, and (ii) kinetic assays in which the extent of reaction at various time points and a fixed concentration of FLP protein is determined. A schematic diagram of these assays is shown in Figure 2.

The first assay establishes the minimum concentration of FLP protein required to produce detectable products for each substrate. A similar assay is described in detail elsewhere (Senecoff *et al.*, 1988). This was our primary method for comparing the reactivity of each of the mutant recombination sites. All sites were compared as linear substrates. In all experiments a comparison was made to an identical control assay employing previously characterized recombination sites. We compared asymmetric mutant spacers with the wild-type spacer in the plasmid pJFS36 (TCTAGAAA) and symmetric mutant spacers with the symmetric spacer in the plasmid pJFS39 (TTCTAGAA). This standard control reaction was always carried out on the same day and utilized the same reagents. Control reactions were highly reproducible and showed no differences in reactivity between pJFS36 and pJFS39, as reported previously (Senecoff and Cox, 1986). These controls provided a point of reference which allowed us to quantitate and compare the results. The reactivity of a mutant site is defined by dividing the minimum amount of FLP protein required for detectable recombination, by the minimum amount required with a wild-type sub-

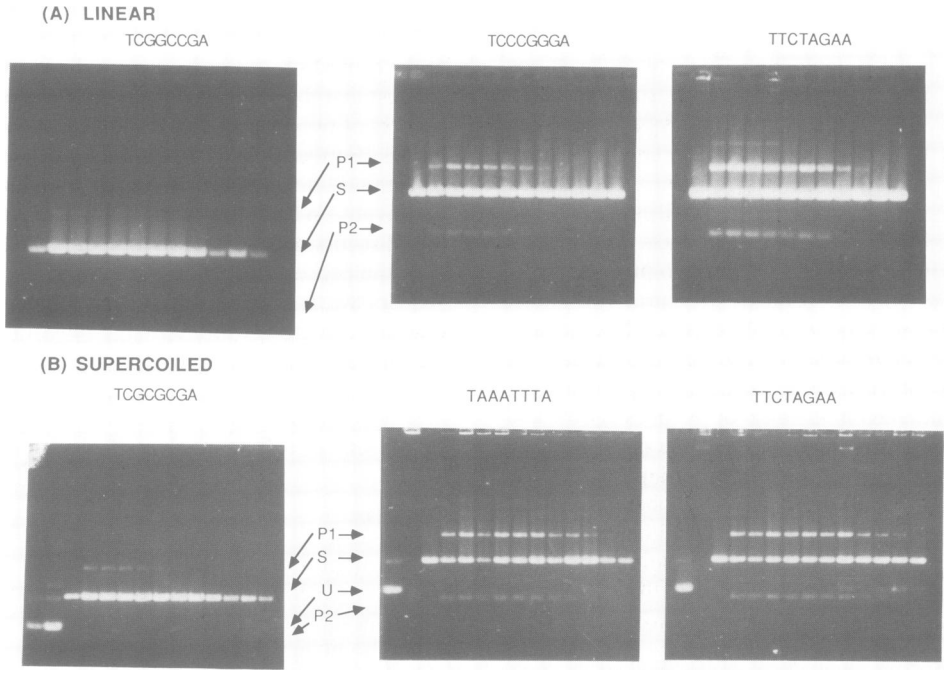


Fig. 3. Effects of spacer alterations on recombination: FLP titration assays. Reactions were carried out as described in Materials and methods, with incubation at 30°C for 60 min. DNA (recombination site) concentration was 5.0 nM in all experiments. **(A)** Linear substrates. For each set of reactions, FLP protein concentrations by lane were: **(1)** 0.0, **(2)** 80.0 nM, **(3)** 40.0 nM, **(4)** 20.0 nM, **(5)** 10.0 nM, **(6)** 4.0 nM, **(7)** 2.0 nM, **(8)** 1.0 nM, **(9)** 0.4 nM, **(10)** 0.2 nM, **(11)** 0.1 nM, **(12)** 0.05 nM. **(B)** Supercoiled substrates. DNA was digested with *PstI* following reaction with FLP protein as described in Materials and methods. For each set of reactions, the first two lanes represent samples without or with 80.0 nM FLP protein, respectively, which were not treated with *PstI*. The *PstI*-treated reactions (**lanes 3–13**) contained amounts of FLP protein shown above for **lanes 1–11** in reactions with linear substrates.

strate. A higher number, therefore, corresponds to a less reactive mutant site.

FLP titrations were also performed on all mutant sites with symmetrical spacers as supercoiled substrates. These reactions were compared to reactions with supercoiled pJFS39 as well as to assays of the same mutants as linear substrates. Reactions containing supercoiled molecules produce large, multimeric circles, which are at least as large as 11 monomer lengths as determined by partial digestion with *Pst*I (data not shown). These frequently do not enter the gel. Complete digestion with *Pst*I generates products identical to those formed in an FLP reaction with *Pst*I-generated linear substrates, as described above allowing for simple quantitation of recombination efficiency.

Rates of recombination were measured for a representative subset of mutants. These results were again compared to the rate of appropriate control reactions, and were performed on both linear and supercoiled substrates as well as at low and high concentrations of FLP protein. Relative rates of reaction were obtained by dividing the time (in minutes) required for a mutant substrate to reach a given extent of reaction, (15–35% conversion of substrates to products) by the time required for a control reaction to reach the same extent. A large number again reflects a lower relative rate of reaction.

DNA binding was measured with a gel retardation assay. This technique produces results comparable in sensitivity to nitrocellulose filter binding protocols (Garner and Revzin, 1981) and has the advantage of producing distinct complexes. Again, both FLP titrations and kinetic experiments were employed with both mutant and wild-type sites.

Effects of spacer mutations on recombination efficiency

Typical recombination assays are presented in Figure 3 and the data are summarized in Table I. Every possible base change in the central 6 bp of the spacer is represented in one or more of the altered spacer sequences generated. The data presented in Table I indicate that there is no significant relationship between a mutation in any single base of the central 6 bp of the spacer and the efficiency of recombination, with one possible exception discussed below. At least 14 of 18 possible single changes are present in sites which behave identically to wild-type in the titration assay. All of the other four mutations (T→A at position −2, pSWU3, 7, and 13; A→C at position −1, pSWU12; A→G or C at position +3, pSWU12 and 20) appear in sites which produce reductions in recombination of 5-fold or less. The only mutation which produces a significant effect on recombination on its own is the T→A change at position −2 (pSWU3). The small effect noted here may reflect an incidental steric contact with FLP protein or a DNA structural effect which is distinct from the patterns described below. A symmetrical A→T change at position +2 (pSWU4) has no effect on recombination. Single mutations were not constructed for the A→C change at position −1 or the A→G or C changes at position 3. In each case symmetrical changes at positions +1 or −3 are present in sites which react normally. There appear to be no essential sequence specific contacts in the spacer, a result consistent with the conclusions of Bruckner and Cox (1986).

The reductions in the efficiency of recombination produced by some mutant spacers with multiple alterations, however, are dramatic. For example, the mutant site plasmid pSWU14 (TGGGCCCA) has never produced detectable recombination products when used as a linear substrate. The other mutant sites listed in the lower third of Table I also produce large (20- to 50-fold) reductions in recombination efficiency. The results in Table I have led us to focus on two sequence-dependent structural elements of the spacer which, if disrupted, produce a deleterious effect on recombination. These are: (i) the 5′ polypyrimidine tracts that extend from each 13-bp repeat into the center of the spacer, and (ii) the relatively high AT content of the spacer.

Table I. Effects of spacer alterations on recombination

	Spacer	T	R
WT	TCTAGAAA	1.0	1.0
pJFS39	TTCTAGAA	1.0	1.0
pJFS40	TCTAGATA	1.0	
pSWU1	TACGTAAA	1.0	
pSWU4	TCTAGTAA	1.0	
pSWU6	TGTACAAA	1.0	2.0
pSWU8	TATATATA	1.0	1.0
pSWU9	TTGATCAA	1.0	2.0
pSWU11	TTTTAAAA	1.0	0.7
pSWU19	TGTAGAAA	1.0	
pSWU12	TCCCGGGA	2.0	6.0
pSWU18	TGTTTTCA	2.0	
pSWU3	TCAAGAAA	2.5	
pSWU7	TCATGAAA	2.5	
pSWU13	TAAATTTA	4.0	15.0
pSWU2	TGATCAAA	5.0	7.0
pSWU20	TGTAGACA	5.0	
pSWU16	TCGCGCGA	20	>100
pSWU17	TCGGCCGA	20	
pSWU10	TGCGCGCA	50	>100
pSWU14	TGGGCCCA	>400	

Summary of results in titration (T) and rate (R) assays. A larger number corresponds to a decrease in recombination efficiency relative to the unaltered site by the factor indicated. See text for details.

The first element consists of 11 pyrimidines in the first 12 bases to the left of position −2, and 12 in the 14 bases to the right of position −1 (Figure 1). The conclusion that these polypyrimidine tracts are important is based on the general observation that multiple spacer base changes, that do not alter the overall GC content, are deleterious only if they are transversions (e.g. pSWU2, 13, 20). Several rules seem to govern this effect. For the moment, sequence changes which increase GC content will not be considered. First, a significant effect is not observed unless there are transversions in at least two of the four bases at positions −2, −3, +2 and +3. Changes in positions +1 and −1 do not appear to affect the reaction. The only exception to this rule is the significant effect of the single T→A change at −2 (pSWU3) described above. We feel that this does not fit into the general pattern because a single T→G change at the same position (pSWU9) has no apparent affect. Second, at least one of the two changes must be at positions +3 or −3. A larger effect is observed if the substitution at +3 or −3 is a guanine. The sequence TTGATCAA (pSWU9) has no effect on recombination in titrations while significant effects are observed for TGATCAAA (pSWU2),

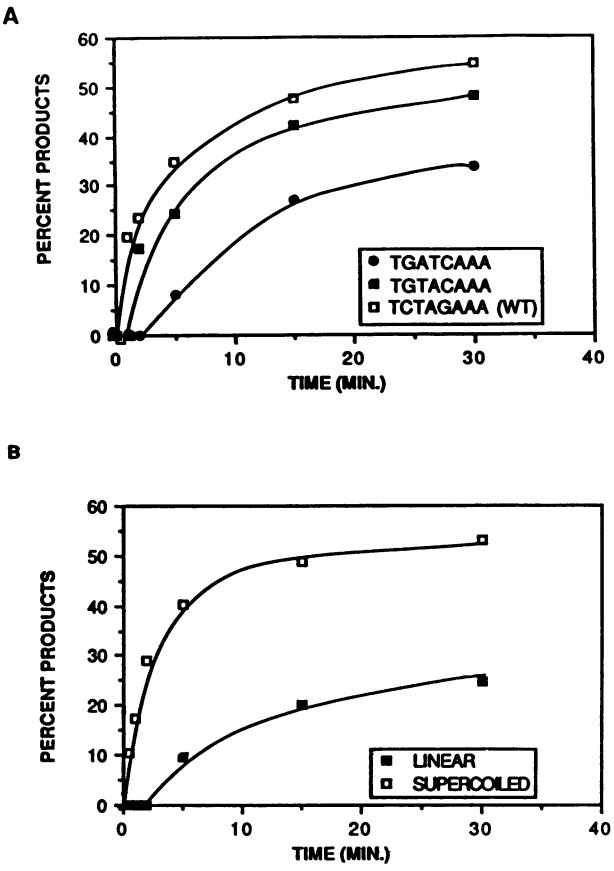


Fig. 4. Effects of spacer alterations on rates of recombination. Reactions were carried out as described in Materials and methods. (**A**) The effects of two spacer mutants, utilized as linear substrates, on rates of recombination. (**B**) The effect of supercoiling on rates of recombination is shown for the mutant TAAATTTA (pSWU13).

and TGTTTTCA (pSWU18). The largest of these effects is for pSWU2, with a C→G transversion at −3 and a T→A transversion at −2. It should also be noted that TATATATA (pSWU8) does not produce a significant effect on the reaction.

The effect of the second DNA structural element, AT content, is apparent from the data in the lower third of Table I. The four mutants which have the greatest effect on recombination efficiency have six G:C base pairs in the center of the spacer. The relative importance of the polypyrimidine tracts versus the AT content is difficult to evaluate since all of these sites contain two or more transversions at positions −2, −3, 2 and 3. The only site with six G:C base pairs that react with significant efficiency as a linear substrate is TCCCGGGA (pSWU12) which has the strongest polypyrimidine strings. Further, the only site with six A:T base pairs between positions −3 and +3 that does not recombine at the wild-type level is TAAATTTA (pSWU13) which has a maximal disruption of the polypyrimidine tracts. Finally, the most severe reduction in recombination efficiency is seen with the spacer TGGGCCCA (pSWU14) which combines maximum disruption of the polypyrimidine tracts with the highest GC content. Intermediate effects which generally follow rules outlined above, are observed with the other mutant spacers.

Time course experiments (Figure 4) reveal that the spacer alterations primarily affect the rate of recombination. The relative reaction rates of mutant substrates (both linear and supercoiled) generally parallel the results obtained with the FLP titration assays (Table I). We find that at extremely long time points the extent of recombination of relatively inefficient mutant substrates (e.g. pSWU12, pSWU13 and pSWU20) reaches the endpoint observed with the unaltered recombination site. The kinetic experiments, however, are somewhat more sensitive and suggest that the titration results underestimate the effects of many mutations by 3- to 4-fold. The sensitivity of the rate measurements permits the resolution of subtle differences between some of the mutant sites in the upper third of Table I. A single guanine substitution at position −3 is found to have a small deleterious effect on reaction rate (pSWU6) while an adenine substitution still has no effect (pSWU8). Simultaneous transversions at position −2 and +2 also exhibit a modest effect (pSWU9). All of these sites either maintain or increase the AT content relative to the wild-type site. A site which optimizes both the AT content and the polypyrimidine tracts (pSWU11) actually exhibits a small rate advantage over the wild-type site. A short lag in product formation was observed in some experiments and is evident in some time courses presented in Figure 4. This has not yet been characterized.

Titration assays carried out with supercoiled substrates indicate that supercoiled molecules react more efficiently than linear ones. All substrates with symmetrical spacers were tested as supercoiled DNA. In every case an improvement in reaction efficiency of at least 2- to 3-fold was noted relative to linear substrates (data not shown). In particular, the mutant site which decreases recombination >400-fold (pSWU14), i.e. shows no reaction as a linear substrate, exhibits a detectable level of product formation at high concentrations of FLP protein when the substrate is supercoiled (data not shown). However, the same general patterns of relative efficiency of recombination seen with linear substrates hold true at the native superhelical density of *E.coli* plasmids. Several mutant sites with symmetrical spacers (pSWU10, 12, 13 and 16) were tested to determine the effects of supercoiling on recombination rate. Each mutant reacted 2- to 5-fold more rapidly as a supercoiled substrate than as a linear one at a given FLP protein concentration. A typical comparison (with pSWU13) is presented in Figure 4B. The rate advantage of supercoiled substrates is described in detail elsewhere for wild-type FRT sites (Gates and Cox, 1988).

Effects of spacer mutations on DNA binding

We have used the technique of gel retardation of nucleic acid fragments by bound proteins in order to assign relative rates and extents of binding to a subset of our mutant sites in comparison to the wild-type. Sites evaluated include pSWU2, pSWU10, pSWU12, pSWU14 and pSWU16 as well as the sites in the standard control plasmids pJFS36 and pJFS39. These cover the range of recombination effects observed. The FLP recombination site is contained on one end of a 316-bp fragment generated by restriction digestion (see Materials and methods). This arrangement was originally chosen for ease of cutting and labelling. The anticipated protein-bound complexes resolve easily in the gel system used. These complexes are referred to as complex 1 and complex 2 in the manner of Andrews *et al.* (1987; Senecoff and Cox, 1988) and result from the binding of 1 and 2 units (either monomers or multimers) of FLP protein, respectively, to one or both FLP binding sites in the FRT site. These

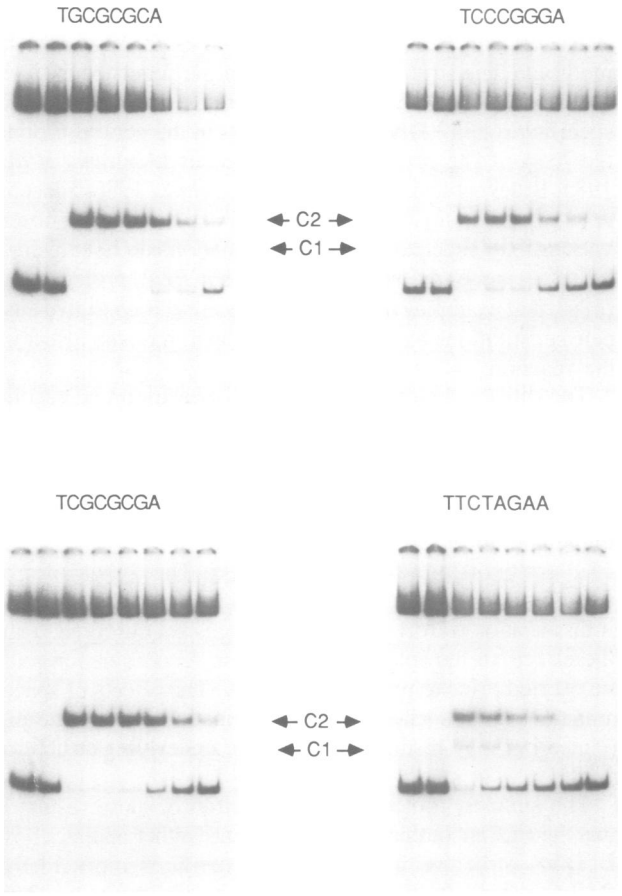


Fig. 5. Effects of spacer alterations on FLP protein binding. Reactions were carried out as described in Materials and methods. Concentration of FRT sites was 0.5 nM. Labelled fragments of **lane 1** of each set of reactions are, from top to bottom (excluding material in the loading wells) a 2505-bp non-specific control DNA fragment and a 316-bp DNA fragment containing the FRT site. FLP concentration, in nM, for each lane left to right was: 0.0, 40.0 (+SDS treatment before loading), 40.0, 20.0, 10.0, 4.0, 2.0 and 1.0 respectively. Complex 1 and complex 2 are denoted C1 and C2. See text for details.

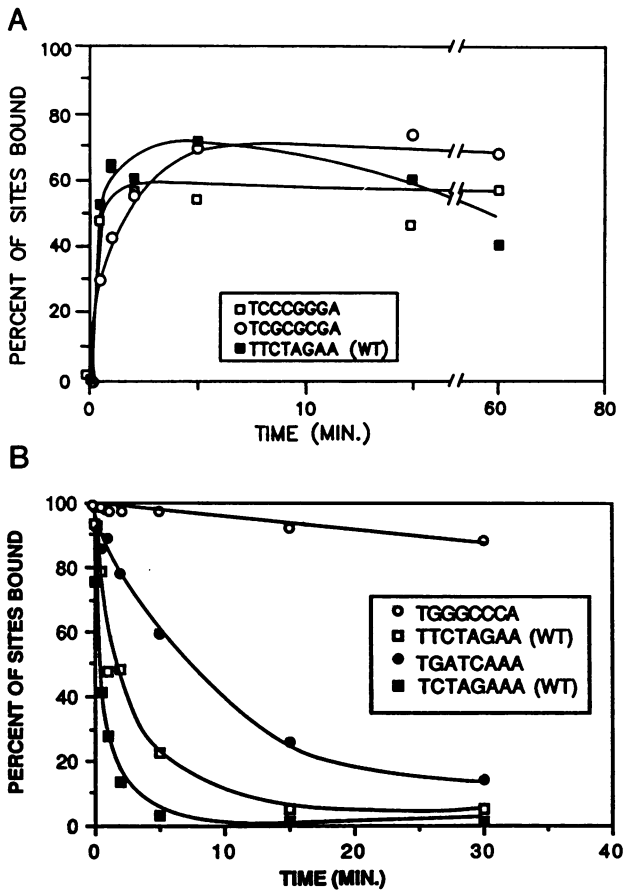


Fig. 6. Effects of spacer alterations on FLP protein binding: rates of association and dissociation. Reactions were carried out as described in Materials and methods. FLP protein concentration was 4.0 nM. DNA (FRT site) concentration was 0.5 nM. (**A**) Assocation. (**B**) Dissociation: pSWU14 DNA was added as a specific competitor at 0.0 min. See text for details.

complexes are formed in the presence of a 100-fold nucleotide excess of non-specific DNA and are sensitive to SDS (lane 2 in each panel of Figure 5). Experimental conditions were chosen to be optimal for binding; the presence of competitor DNA, the absence of polyethylene glycol (PEG) and glycerol and the low concentration of the substrate DNA preclude observable levels of recombination in these assays. A 598-bp product expected from a recombination reaction, and readily resolved in these gels, is never observed.

Typical FLP protein titrations are presented in Figure 5. At high FLP protein concentrations (80 FLP monomers per site), the fragment containing the recombination site is almost completely bound and therefore exhibits a slower migration in the gel. This is true for all sites examined. Complex 2 is in all cases more prominent than complex 1. A further result worth noting is the fact that all complexes and unbound fragments migrate nearly identically when a comparison is made between the mutants and the wild-type sites (data not shown). This implies that any DNA structural perturbations that contribute to the slower gel migration are roughly identical in all cases.

The most intriguing result is that an apparent inverse relationship exists between the efficiency of recombination and the efficiency of DNA binding. That is, mutant sites that recombine inefficiently bind at lower concentrations of FLP protein than do sites that recombine more efficiently. This is easily seen in Figure 5. Recombination sites with the spacers TGCGCGCA (pSWU10) and TCGCGCGA (pSWU16), which are reduced 50- and 20-fold in recombination respectively, are bound at an FLP protein: site ratio of 2:1. Sites with the spacer TCCCGGGA (pSWU12), reduced in recombination 2-fold, are bound at a ratio of 4:1. Binding to the pJFS39 recombination site, which recombines at wild-type efficiency, is not observed below a ratio of 8:1. A consistent pattern is seen with sites containing the spacers TGATCAAA (pSWU2) and TGGGCCCA (pSWU14) which are both bound at a ratio of 2:1, and the wild-type site in pJFS36 which is bound at a ratio of 8:1 (data not shown). Spacer mutants do not, in any case examined, produce a deleterious effect on site recognition and binding as measured by this assay.

The kinetics of FLP protein binding were assayed to estimate the relative association and dissociation rates of selected mutant substrates by comparison to wild-type. Association assays were performed in the presence of an excess of non-specific competitor DNA and were started by the addition of FLP protein. Rates of dissociation were measured in experiments in which substrates were pre-bound with FLP protein until well after equilibrium had been

established (30 min) then challenged with a 50-fold FRT site excess of specific competitor DNA (pSWU14) to start the reaction. These experiments appear in Figure 6. The association results (Figure 6A) indicate that the rate of formation of bound complexes is similar in all cases. In the presence of a high concentration of FLP protein (80 monomers per FRT site), binding is stable for at least 2 h (data not shown). However, at a lower concentration of FLP protein (eight monomers per FRT site), the fraction of bound complexes formed with more reactive spacers (TCCCGGGA, pSWU12 and TTCTAGAA, pJFS39) begins to decline over 1 or 2 h while the fraction formed with less reactive spacers (TGCGCGCA, pSWU10 and TCGCGCGA, pSWU16) remains constant for 2 h (data not shown). These results suggest that the reduction in binding with reactive substrates is not due to a slow binding process but is due to a higher dissociation rate of the complexes. The results of the dissociation assays (Figure 6B) confirm this hypothesis. In the presence of an excess of specific competitor DNA, the wild-type substrates pJFS36 and 39 exhibit a much higher dissociation rate than the mutant substrates pSWU2 and 14.

Discussion

Our principle conclusion is that the sequence structure of the FLP recombination site is critical to site function. This role is distinct from protein−DNA interactions involved in recognition. This conclusion is based on several specific observations:

(i) There are two important DNA sequence elements in the FLP recombination site that influence the rate of recombination. These are: (a) the 5′ polypyrimidine tracts that extend from within each 13-bp repeat into the center of the spacer and (b) the predominance of AT base pairs in the spacer. Significant alteration of either or both sequence elements in the spacer region has a deleterious effect on recombination site function. We find that the effects associated with disruption of these elements are position dependent. A purine (guanine) substitution at position −3 is slightly more deleterious to recombination than a substitution at position −2. These two elements are superimposed on the site; spacer sequences with a high GC content are tolerated only when the pyrimidine tracts are optimized, and the effects of transversions are minimized if purine substitutions are adenines. Great flexibility in the spacer sequence is tolerated as long as these structural elements are maintained.

(ii) Supercoiling increases the rate and efficiency of recombination for both altered and unaltered recombination sites. The relative rates of recombination, however, still reflect the sequence-dependent effects described above.

(iii) The efficiency of binding of the FLP protein to mutant sites, as measured by the gel retardation assay, is inversely related to the efficiency of recombination observed with each site. This binding efficiency appears to reflect an effect on the rate of dissociation. Within the limits of our measurements, mutations in the spacer do not affect the rate of association. FLP protein−DNA complexes formed at low FLP protein concentrations are observed to be most stable over long time periods with those mutant sites that are the least recombinogenic.

Based on a small set of spacer mutations, which nevertheless included changes at each of the central 6 bp, we previously concluded that the space sequence was irrelevant as long as homologous spacers were reacted with one another (Senecoff and Cox, 1986). This conclusion must now be modified. The structure or context of the sequence is clearly important to site function. There is no evidence that specific protein−DNA contacts exist in the central 6 bp of the spacer. The results of earlier studies (Andrews *et al.*, 1985; Bruckner and Cox, 1986; Prasad *et al.*, 1986; Senecoff *et al.*, 1988) in this regard are reinforced here by the observations that (i) deleterious spacer mutations do not adversely affect recognition of the site by FLP protein, and (ii) several examples of spacer sequences with up to five base pair substitutions have been found which have no affect on the reaction.

The direct relationship between dissociation and recombination rates might suggest that dissociation itself is the affected step. If this were true, however, the slow step in the reaction would occur after the first recombination events had already generated products. The rate of product formation though, follows the pattern outlined in Table I even when FLP protein is present in large excess. The 'burst' of product formation which would be expected for poor substrates if dissociation were kinetically limiting is not observed. Thus the effects of the spacer alterations on recombination must be exerted prior to product formation. The effects of spacer mutations on dissociation and recombination are not causally related but may result from the same underlying structural effect.

The simplest molecular interpretation of these results is that the FLP recombination site sequence either adopts a non-B DNA structure in solution, or (perhaps more likely) facilitates a conversion to some higher order structure. This structure or structural change is required for recombination. Since FLP protein association is not adversely affected by the deleterious spacer mutations, the step facilitated by the recombination site sequence must be subsequent to binding. The effects of high GC content and supercoiling suggest that the affected step involves DNA unwinding.

A role for these DNA sequence elements in unwinding is consistent with their known properties. DNA containing regions of high GC content has marked differences in base stacking and increased dynamic stability by comparison to canonical B DNA (Drew and Travers, 1985; Koudelka *et al.*, 1987). Sequences that must be unwound to effect a process in DNA metabolism, such as promoters (Drew *et al.*, 1985) and origins of replication (Bramhill and Kornberg, 1988), do so most efficiently under conditions of negative supercoiling. Such sequences are also generally AT rich. Although interpretation of the effects of the polypyrimidine string is less straightforward, the general observation that tracts of purines and pyrimidines are hypersensitive to digestion by single strand nucleases is again consistent with a role in unwinding. It has recently been demonstrated that the FRT site in the plasmid pJFS39 is hypersensitive to S1 nuclease digestion in the absence of FLP protein (S.Umlauf and S.Loh, unpublished results), reinforcing the link between sequence structure and function in this recombination site. The specific nature of structures (and structural transitions) adopted by these sequences is under investigation in several laboratories (Pulleyblank *et al.*, 1985; Kilpatrick *et al.*, 1986; Htun *et al.*, 1985). The polypyrimidine tracts in the FLP recombination site extend for 10 bp into the FLP protein binding sites. Although an extensive mutational analysis of the binding sites found no substitutions that affect recombi-

nation without affecting recognition, it is notable that transition mutations were generally well tolerated by comparison to transversions (Senecoff *et al.*, 1988).

The relative roles of the DNA and protein in this system may be aptly described in the context of the recently proposed 'handshake' theory of antibody–antigen complex formation (Colman *et al.*, 1987). The final structure of the transition state or reactive complex requires conformational changes by the DNA and probably the protein as well. A random DNA sequence is unsuitable for the required DNA transformations; thus, the energy available in site-specific binding of FLP protein alone is insufficient to cause the required DNA structural changes. The effects of the spacer alterations suggest that DNA structure contributes at least 1–2 orders of magnitude to the rate of recombination in this system. The dynamic instability of the DNA sequence of the FLP recombination site suggested by this study may be manifested by a degree of structural flexing both in solution and in protein-bound complexes. This would provide a simple explanation for the correlation between recombination and dissociation, both of which could be influenced by DNA flexing.

Recent work in related site-specific recombination systems has indicated that the recombining DNA proceeds through an intermediate resembling a 'Holliday structure' (Hoess *et al.*, 1987; Kitts and Nash, 1987; Nunes-Duby *et al.*, 1987). Recent work indicates that the same is true for the FLP recombination system (M.Jayaram, personal communication; L.Meyer-Leon, L.-C.Huang and S.Umlauf, unpublished results). It is possible that the spacer mutations considered here interfere with some step in the formation or resolution of these intermediates.

Materials and methods

Bacterial strains, enzymes and reagents

Plasmids were maintained in the *E.coli* strain HB101 and were prepared as described (Maniatis *et al.*, 1982). Plasmids pSWU1–20 were constructed as follows. Four complimentary oligonucleotides were synthesized, annealed and ligated to construct each altered recombination site. These oligonucleotides were arranged in a cassette format such that the 3′ end fragment on each side of the site was synthesized containing the sequence for a 13-bp repeat, the outermost base of the spacer and one linker sequence (either *Bam*HI or *Eco*RI). These two oligomers were used for all constructs. The complimentary strands were synthesized so that the desired spacer sequence was found in a six-base overhang at the center when the pairs of strands were annealed. Phosphate was added to the 5′ ends of the oligonucleotides that contained the spacer sequences by polynucleotide kinase prior to annealing. The oligonucleotides were annealed to the cassette strands containing the complementary repeat sequence. The pairs were then ligated into the large fragment of pXF3 digested with *Bam*HI and *Eco*RI (Maniatis *et al.*, 1982). Oligonucleotides were prepared at the DNA Synthesis Facility in the University of Wisconsin Biotechnology Center. Mutations were confirmed by DNA sequencing as described below.

The FLP protein purification scheme is described in detail elsewhere (Gates *et al.*, 1987; Meyer-Leon *et al.*, 1987). Final concentration of FLP protein stock solutions was estimated by the assay of Bradford (1976). In all cases FLP protein was diluted in a buffer containing 25 mM TAPS (pH 8.0), 1 mM EDTA and 1 M NaCl. All FLP protein used in this study was at least 80% pure.

Restriction enzymes and polynucleotide kinase were purchased from New England Biolabs. Reverse transcriptase was purchased from Life Sciences. The large fragment of T4 DNA ligase was purified as described (Davis *et al.*, 1980). [γ-^{32}P]ATP ($\geq$5000 Ci/mmol; 1 Ci = 37 GBq) was purchased from Amersham, and [α-^{32}P]dATP was purchased from New England Nuclear. TAPS [*N*-Tris (hydroxymethyl) methyl-3-aminopropanesulfonic acid] was purchased from Research Organics.

FLP recombination reaction conditions

All reactions were intermolecular, using plasmids with a single FRT site. In the case of assays using linearized substrates, plasmids were digested with either *Eco*RI or *Pst*I to produce full-length linear DNA. Some of the constructed spacers are asymmetric, like the wild-type site. These mutant sites were assayed by combining equal amounts of substrates obtained from either *Eco*RI or *Pst*I digests of the same plasmid. A reaction involving one of each of these substrates results in the creation of novel products (Senecoff *et al.*, 1985). In the case of mutants with symmetrical spacers, unique products are obtained from recombination in which all substrates are derived from digestion with the same restriction enzyme, *Pst*I in all cases reported here (Senecoff and Cox, 1986). In both cases, reactions are assumed to be taking place (with substrates in parallel alignment) that result in products identical to the starting material. These reactions are not detectable and are assumed to occur with equal frequency as those that produce novel products. Thus a maximum reaction results in only 50% conversion of substrate to detectable products.

Assays were also performed on supercoiled substrates. In this case, plasmids containing a recombination site with a symmetrical spacer were reacted with FLP protein and digested with *Pst*I after the reaction was stopped.

DNA substrate concentrations are reported in terms of total recombination sites. This is in all cases equivalent to the concentration of DNA molecules. All reactions contained 25 mM TAPS buffer (pH 8.0), 1 mM EDTA, 10% PEG (8000 mol. wt), 20% glycerol (v/v), 2.0 mg/ml bovine serum albumin (BSA), 5.0 nM of substrate DNA and 250 mM NaCl. [This amount of NaCl was chosen as a compromise in order to compare reactions of supercoiled and linear substrates. The reaction with linear substrates has an optimum of 200 mM NaCl and the reaction with supercoiled substrates has an optimum of 300 mM NaCl (data not shown). The reason for this difference is as yet unknown.] For experiments in which a single time point was required, the volume of the standard reaction mixture before the addition of FLP protein was 18 μl. Reactions were started by the addition of 2 μl of appropriately diluted FLP protein. In the case of extended time courses under a single set of conditions, reaction volumes before addition of FLP protein were either 198 μl, for linear substrates, or 234 μl for supercoiled substrates. Reactions were started by the addition of 22 or 26 μl of appropriately diluted FLP protein, respectively. Aliquots (20 μl) were removed at the indicated times and treated as described below. Amounts of FLP protein and incubation times are indicated in the figure legends. Reactions were carried out at 30°C. Standard reactions (or 20 μl aliquots from time course reactions) with linear substrates were stopped by the addition of 2 μl of 10% SDS. Reactions with supercoiled substrates were stopped by heating to 70°C for 10 min; this was found to halt reactions as quickly as treatment with SDS (data not shown). Samples were then cooled to room temperature for 20 min at which time 10 μl of 30 mM $MgCl_2$ and 20 units of *Pst*I were added. Samples were then incubated for 60 min at 37°C. Digestion was stopped by the addition of 2 μl of 10% SDS. Following these treatments samples from linear or supercoiled DNA reactions were mixed with 8 μl of GED (50% glycerol, 10 mM EDTA and 0.01% Bromphenyl blue) loading dye and were loaded on a 0.8% agarose gel. After electrophoresis, gel bands were visualized by illumination with a UV transilluminator after staining with ethidium bromide. Gels were photographed with Polaroid Type 55 film. This film produces both a negative and a positive image. Gel band intensity was quantitated with a Zeineh soft-laser scanning densitometer SL-504-XL from Biomed Instruments Inc. Corrections for intensity variation due to DNA fragment length differences and non-linearity in the scanning data were employed when appropriate as described elsewhere (Gates and Cox, 1988).

DNA binding measurements

Gel retardation assays for DNA binding were performed essentially as described (Fried and Crothers, 1981; Garner and Revzin, 1981). Plasmids containing recombination sites with mutant spacers were digested with *Sal*I and *Eco*RI and end labelled by filling in the resulting protruding ends with [α-^{32}P]dATP and the other three unlabelled nucleotides using reverse transcriptase. Substrates were labelled to $\sim 6 \times 10^6$ c.p.m./μg. Samples for titration and association assays were treated with FLP protein as indicated in the figure legends in solutions containing: 25 mM TAPS (pH 8.0), 1 mM EDTA, 250 mM NaCl, 0.5 nM of substrate DNA, 2.0 mg/ml BSA, and 50.0 μg/ml of unlabelled calf thymus DNA. Reactions were carried out at 30°C for the indicated period of time. In the case of dissociation assays, conditions were identical except for the replacement of the calf thymus DNA with 50 μg/ml of pSWU14 (2.5 μM). Final reaction volumes for FLP titration assays were 20 μl. After addition of 10 μl of loading buffer (50% glycerol, 50 mM EDTA, 0.01% Bromophenyl blue, 0.01% xylene cyanol and 2.0 mg/ml BSA) reactions were loaded onto 5% polyacrylamide gels.

The gels were run at 125–150 V in 1 × TBE (89 mM Tris-borate, 89 mM boric acid and 2 mM EDTA). Association assays were started with the addition of FLP protein and contained 200 μl final volume from which 20 μl aliquots were removed at the indicated times. Substrates in dissociation assays were pre-incubated with FLP for 30 min to ensure establishment of a binding equilibrium. Data collection was started after the addition of the specific competitor DNA and also contained 200 μl final volume. The competitor DNA, supercoiled pSWU14, was chosen from the results of the binding titration assays as the most stably bound site. The final concentration constituted a 50-fold site excess over the substrate DNA. Kinetic assay reactions were stopped by loading samples directly onto gels while running at 125–150 V. Gels were dried and subjected to autoradiography with preflashed film (Agfa X-ray film). Quantitation of bands was peformed by liquid scintillation counting of gel slices (Bio-Safe II fluor was obtained from Research Products International) on a Beckman LS 3801 counter.

DNA sequencing

Sequencing of plasmids was performed as described (Senecoff *et al.* 1988). The protocol is a variation of the dideoxy-nucleotide primer extension method of Sanger *et al.* (1977) using reverse transcriptase and a pBR322 primer, on alkaline denatured plasmids. The *Eco*RI primer was obtained from either New England Biolabs or the University of Wisconsin Biotechnology Center with identical results.

Acknowledgements

The authors would like to thank P.Rossmeissl, L.Meyer-Leon, B.F.Pugh and C.Gates for helpful discussions and critical review of the manuscript. Special thanks go to J.Senecoff and M.Sundaralingam for critical advice. This work was supported by the National Institutes of Health grant GM37835. Supplemental research support was provided by a Shaw Fellowship awarded to M.M.C. by the Milwaukee Foundation. M.M.C. is supported by National Institutes of Health Research Career Develoment Award AI00599. S.W.U. is supported by National Institutes of Health predoctoral training grant GM07215.

References

Andrews,B.J., Proteau,G.A., Beatty,L.G. and Sadowski,P.D. (1985) *Cell*, **40**, 795–803.

Andrews,B.J., Beatty,L.G. and Sadowski,P.D. (1987) *J. Mol. Biol.*, **193**, 345–358.

Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.

Bramhill,D. and Kornberg,A. (1988) *Cell*, in press.

Broach,J.R. (1981) In Strathern,J.N., Jones,E.W. and Broach,J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 445–470.

Bruckner,R.C. and Cox,M.M. (1986) *J. Biol. Chem.*, **261**, 11798–11807.

Burkhoff,A.M. and Tullius,T.D. (1987) *Cell*, **48**, 935–943.

Colman,P.M., Laver,W.G., Varghese,J.N., Baker,A.T., Tulloch,P.A., Air,G.M. and Webster,R.G. (1987) *Nature*, **326**, 358–362.

Cox,M.M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4223–4227.

Davis,R.W., Botstein,D. and Roth,J.R. (1980) *Advanced Bacterial Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 196–197.

Drew,H.R. and Travers,A.A. (1985) *J. Mol. Biol.*, **186**, 773–790.

Drew,H.R., Weeks,J.R. and Travers,A.A. (1985) *EMBO J.*, **4**, 1025–1032.

Fried,M.G. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.

Futcher,A.B. (1986) *J. Theor. Biol.*, **119**, 197–204.

Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047–3060.

Gates,C.A. and Cox,M.M. (1988) *Proc. Natl. Acad. Sci. USA*, in press.

Gates,C.A., Meyer-Leon,L., Attwood,J.M., Wood,E.A. and Cox,M.M. (1987) In Burgess,R. (ed.), *Protein Purification: Micro to Macro*. Alan R.Liss, Inc, New York, pp. 197–206.

Gronostajski,R.M. and Sadowski,P.D. (1985) *Mol. Cell. Biol.*, **5**, 3274–3279.

Hoess,R., Wierzbicki,A. and Abremski,K. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6840–6844.

Htun,H., Lund,E., Westin,G., Pettersson,U. and Dahlberg,J.E. (1985) *EMBO J.*, **4**, 1839–1845.

Husain,I., Sancar,G.B., Holbrook,S.R. and Sancar,A. (1987) *J. Biol. Chem.*, **262**, 13188–13197.

Kilpatrick,M.W., Torri,A., Kang,D.S., Engler,J.A. and Wells,R.D. (1986) *J. Biol. Chem.*, **261**, 11350–11354.

Kitts,P.A. and Nash,H.A. (1987) *Nature*, **329**, 346–348.

Koudelka,G.B., Harrison,S.C. and Ptashne,M. (1987) *Nature*, **326**, 886–888.

Laundon,C.H. and Griffith,J.D. (1987) *Biochemistry*, **26**, 3759–3762.

Lomonossoff,G.P., Butler,P.J.G. and Klug,A. (1981) *J. Mol. Biol.*, **149**, 745–760.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Marini,J.C., Levene,S.D., Crothers,D.M. and Englund,P.T. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7664–7668.

Meyer-Leon,L., Gates,C.A., Attwood,J.M., Wood,E.A. and Cox,M.M. (1987) *Nucleic Acids Res.*, **15**, 6469–6488.

Nunes-Duby,S.E., Matsumoto,L. and Landy,A. (1987) *Cell*, **50**, 779–788.

Prasad,P.V., Horensky,D., Young,L.-J. and Jayaram,M. (1986) *Mol. Cell. Biol.*, **6**, 4329–4334.

Pulleyblank,D.E., Haniford,D.B. and Morgan,A.R. (1985) *Cell*, **42**, 271–280.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.

Senecoff,J.F. and Cox,M.M. (1986) *J. Biol. Chem.*, **261**, 7380–7386.

Senecoff,J.F., Bruckner,R.C. and Cox,M.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7270–7274.

Senecoff,J.F., Rossmeissl,P.J. and Cox,M.M. (1988) *J. Mol. Biol.*, in press.

Shakked,Z., Rabinovich,D., Cruse,W.B.T., Egert,E., Kennard,O., Sala,G., Salisbury,S.A. and Viswamitra,M.A. (1981) *Proc. R. Soc. Lond.*, **213**, 479–487.

Vetter,D., Andres,B.J., Roberts-Beatty,L. and Sadowski,P.D. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7284–7288.

Volkert,F.C. and Broach,J.R. (1986) *Cell*, **46**, 541–550.

Wang,A.H.-J., Quigley,G.J., Kolpak,F.J., Crawford,J.L., van Boom,J.H., van der Marel,G. and Rich,A. (1979) *Nature*, **282**, 680–686.

Zahn,K. and Blattner,F.R. (1987) *Science*, **236**, 416–422.

Received on March 16, 1988