

MuA Transposase Separates DNA Sequence Recognition from Catalysis[†]

Ilana Goldhaber-Gordon[‡]

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Michael H. Early[‡]

Thornton School of Music, University of Southern California, Los Angeles, California 90089-0851

Tania A. Baker*

Department of Biology, Massachusetts Institute of Technology and Howard Hughes Medical Institute, Cambridge, Massachusetts 02139

Received July 31, 2003; Revised Manuscript Received October 9, 2003

ABSTRACT: Confronted with thousands of potential DNA substrates, a site-specific enzyme must restrict itself to the correct DNA sequence. The MuA transposase protein performs site-specific DNA cleavage and joining reactions, resulting in DNA transposition—a specialized form of genetic recombination. To determine how sequence information is used to restrict transposition to the proper DNA sites, we performed kinetic analyses of transposition with DNA substrates containing either wild-type transposon sequences or sequences carrying mutations in specific DNA recognition modules. As expected, mutations near the DNA cleavage site reduce the rate of cleavage; the observed effect is about 10-fold. In contrast, mutations within the MuA recognition sequences do not directly affect the DNA cleavage or joining steps of transposition. It is well established that the recognition sequences are necessary for assembly of stable, multimeric MuA–DNA complexes, and we find that recognition site mutations severely reduce both the extent and the rate of this assembly process. Yet if the MuA–DNA complexes are preassembled, the first-order rate constants for both DNA cleavage and DNA strand transfer (the joining reaction) are unaffected by the mutations. Furthermore, most of the mutant DNA molecules that are cleaved also complete DNA strand transfer. We conclude that the sequence-specific contacts within the recognition sites contribute energetically to complex assembly, but not directly to catalysis. These results contrast with studies of more orthodox enzymes, such as *EcoRI* and some other type II restriction enzymes. We propose that the strategy employed by MuA may serve as an example for how recombinases and modular restriction enzymes solve the DNA specificity problem, in that they, too, may separate substrate recognition from catalysis.

During DNA transposition, a transposase protein performs DNA cleavage and joining reactions at the ends of a transposon DNA. The transposon ends are precisely defined by a DNA sequence. Thus, transposases, like all proteins that modify DNA at a specific site, face an impressive challenge: surrounded by incorrect sequences—potential substrates chemically similar to their own—they must select their correct substrate. Little is known about the molecular mechanisms used by transposases and other recombinases to link substrate recognition to the covalent modification of DNA.

MuA transposase is one of the best understood members of the transposase/retroviral integrase protein family. The MuA protein performs the initial steps required to transpose the genome of bacteriophage Mu from its starting DNA location to a new DNA location, the DNA “target”. At each end of the Mu genome, MuA cleaves the 3′ strand of the Mu DNA away from its surrounding sequence. Then in a

one-step “strand transfer” reaction, MuA inserts this cleaved 3′ end into the target DNA. Strand transfer is a transesterification reaction, in which a 3′-oxygen within the target DNA is replaced with the 3′-OH at the end of the Mu DNA (1). A wide range of DNA sequences can serve as the target site (2, 3). In total, MuA performs four reactions to permit transposition of the Mu genome: one cleavage and one strand transfer at each end of the Mu DNA (4, 5). A single MuA subunit performs the two sequential reactions for one DNA end, using a single active site (6, 7).

Multiple DNA sequence elements define the ends of the Mu DNA and thus the sites of recombination (Figure 1A). Each DNA end contains three MuA recognition sites, which in total direct the binding of six MuA monomers to the Mu DNA. These recognition sites are 22 base pairs long, and each consists of an “inner” half (distal to the end of the genome) and an “outer” half (end proximal). MuA itself carries three DNA recognition domains, each possessing a helix–turn–helix motif. Two of these, domains I β and I γ , cooperate in binding each 22 base pair recognition sequence: I β binds the inner half, and I γ binds the outer half. A third DNA recognition domain, I α , binds a separate

[†] This work was supported by United States Public Service Health Grant GM49224 from the NIH.

* To whom correspondence should be addressed. Phone: (617) 253-3594. Fax: (617) 252-1852. E-mail: tabaker@mit.edu.

[‡] These authors contributed equally to this work.

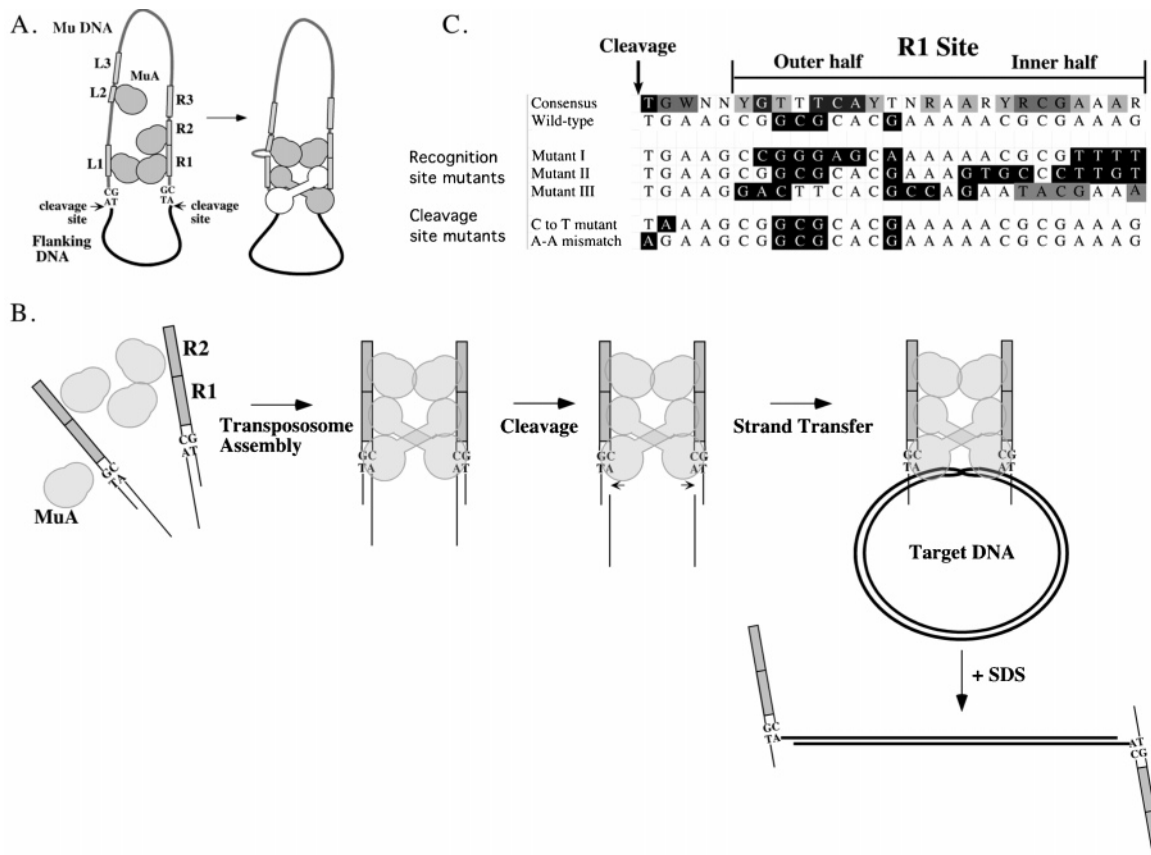


FIGURE 1: Overview of the transposition reaction and DNA substrates. Elements are not drawn to scale. (A) Transpososome assembly. The ends of the Mu DNA each contain three recognition sites, related to one another by a 22 base pair consensus sequence. The sites are named R1, R2, and R3 on the right end, and L1, L2, and L3 on the left end. Binding to these sites triggers transpososome assembly. Cleavage occurs within the transpososome, at the very end of the Mu DNA, after the sequence 5'-(T or A)CA. Here, one MuA subunit is shown in white, to highlight the crisscross structure of the transpososome. (B) Transposition of DNA fragments. (C) Summary of the fragments used in this study. These are the sequences of the uncleaved strands (so that they could be listed 5' to 3' beginning at the end of the Mu DNA). We list here only the cleavage site and R1 site sequences. Complete sequences are shown in Table 1. Each fragment also contained a second recognition site, not shown in the figure: the natural R2 sequence for the wild-type, mutant I, mutant II, and cleavage site mutant fragments, and a repeat of the mutant sequence for mutant III. The consensus sequence (top line) is derived from the naturally occurring MuA recognition sites (70). In the consensus sequence, highlighted boxes are positions at which base-specific contacts were suggested by a sequence selection study (32); darker highlights indicate stronger selection. W stands for T or A. In the fragment sequences, black boxes are positions that do not match the consensus. Note that the natural R1 site contains several black boxes: compared to the other natural sites, R1 has relatively weak affinity for MuA (70). The mutant III sequence is based on a sequence from the genome of an unrelated phage, Φ X174; it was found during a selection for sites on this phage that could be transposed at a low level by MuA (32). The gray boxes in the mutant III sequence represent positions that match the consensus sequence with a phase shift of +1 near the center of the recognition site; several experiments suggest that this fragment is contacted with this phase shift (32).

enhancer-like sequence, which stimulates transposition (8, 9). The enhancer sequence is dispensable *in vitro* and will not be discussed further in this paper. DNA cleavage occurs precisely at the end of the Mu DNA, which is five base pairs 3' to the outermost recognition site on each end of the Mu DNA (R1 or L1; see Figure 1A). Each cleavage site is marked by the sequence 5'-(T or A)CA↓.

MuA functions as part of a large protein–DNA complex called a transpososome (10, 11). At the core of the transpososome are four MuA subunits (12, 13), at least two of which bridge the ends of the Mu DNA by binding a recognition site on one end and a cleavage site on the opposite end (Figure 1A) (14, 15). Active transpososomes can also assemble on short DNA fragments containing two MuA recognition sites (Figure 1B). A MuA tetramer bridges two such fragments (16).

Additional proteins and DNA sequences interact transiently with the core transpososome. For example, the MuB protein is a DNA-binding ATPase that interacts directly with MuA.

MuB is best-known for its role in assisting MuA in choosing a target DNA site, but it also stimulates other steps of transposition (17–20).

The core transpososome is remarkably stable, resisting long incubations with competitor DNA, high temperature, or high concentrations of urea (11–13). Yet in the absence of its recognition sequences, MuA is monomeric and inactive and exchanges between DNA sites with a rapid k_{off} (11–13, 21–23). Thus, specific DNA contacts with the recognition sites must permit multimerization of MuA, causing four MuA subunits to commit to each other and to their bound DNA.

In the current study, we ask whether the specificity of the recognition site sequence also contributes directly to the chemistry of transposition. The specificity of an enzymatic reaction is often considered to be inextricably linked to the catalytic efficiency. For example, classic studies of serine proteases reveal that changing the amino acid sequence of the enzyme's substrate can change the k_{cat} value (24, 25). Among enzymes that modify DNA at a specific site, the

Table 1: Sequences of DNA Fragments Used in This Study

Fragment	Strand	Sequence
wild-type	non-cleaved	ctagtgaagcggcgacgaaaaacgcgaaagcgtttcacgataaatcgaaaaac
	cleavable	gttttcgcattatcgtgaaacgcttcgcgttttcgtgcgccgttcactagacgcttggcgtaatcggcgtaatgc
	precleaved	gttttcgcattatcgtgaaacgcttcgcgttttcgtgcgccgttcac
outer mutant	non-cleaved	ctagtgaagcggcgacgaaaaacgcgaaagcgtttcacgataaatcgaaaaac
	cleavable	gttttcgcattatcgtgaaacgcttcgcgttttcgtgcgccgttcactagacgcttggcgtaatcggcgtaatgc
inner mutant	non-cleaved	ctagtgaagcggcgacgaaaaagtgcccttgtttcacgataaatcgaaaaac
	cleavable	gttttcgcattatcgtgaaacaaggcgacgttttcgtgcgccgttcactagacgcttggcgtaatcggcgtaatgc
R1-R2 mutant	non-cleaved	ctagtgaaggacttcacgccagaatacgaagacttcacgccagaatacgaaacg
	cleavable	cgttcgtattctggcgtagaagcttcttctgattctggcgtagaagcttcactagacgcttggcgtaatcggcgtaatgc

problem of how sequence recognition is linked to catalysis has been carefully studied for several type II restriction enzymes, including *EcoRV*, *EcoRI*, and *BamHI*. These enzymes, like the serine proteases, use specific binding energy to lower the transition state for DNA cleavage (26, 27).

MuA transposase differs in many ways from a conventional enzyme, and therefore, we were interested in examining how MuA ensures that catalysis occurs only at the proper DNA sequence. MuA is modular, having a catalytic domain that folds independently from its DNA recognition domains. Furthermore, because MuA functions in a stable transpososome, transposition involves an assembly step that is distinct from both the DNA-binding step and the cleavage step. Finally, MuA does not turn-over as true catalysts do; once MuA assembles into an active transpososome, it remains in that complex until disassembled by the ATPase ClpX (28, 29).

To study the contribution of specific DNA recognition to individual steps of transposition, we constructed Mu DNA fragments with multiple mutations in a MuA recognition site (Figure 1C). By following the kinetics of each reaction step performed with these fragments (assembly, DNA cleavage, and strand transfer), we assess the contribution of the mutated regions to that step. We find that the mutant fragments are severely compromised in their ability to assemble active transpososomes. Yet once assembled, the mutant DNA substrates can be cleaved with a first-order rate constant equal to or higher than the rate constant observed for the wild-type fragments. Similar results were obtained for the strand transfer reaction. We conclude that the recognition sequences contribute directly and substantially only to the assembly stage of transposition, the commitment step for the overall reaction. These results contrast with studies of simpler proteins, in which sequence recognition contributes directly to catalysis, and serve as an example for how other DNA modification reactions that include a distinct complex-assembly step may solve the DNA specificity problem.

MATERIALS AND METHODS

Proteins and DNA. MuA (30) and MuB (31) were prepared as described. Target DNA (Φ X174 RFI) was purchased from New England Biolabs, and DNA fragments were synthesized by the MIT/HHMI biopolymers lab and gel purified. The

sequences of most of the fragments are shown in Table 1. Cleavage site mutants were identical to the wild-type fragment, except for the changes specified in the text.

Transposition Reaction. Solution conditions for all transposition reactions (assembly, cleavage, or strand transfer) were essentially as described (23). However, the divalent metal ion concentration depended on the reaction stage, as specified in the individual figures and captions. All DNA fragments used in the study were 5' labeled on the cleaved or cleavable strand with T4 polynucleotide kinase. During assembly, MuA was at 200 nM and DNA fragments were at 50 nM; these reactions were diluted 2-fold for cleavage, or 10-fold for strand transfer. During strand transfer, MuB was at 690 nM and target DNA was at 2.8 nM (or 10 ng/ μ L). (In dilution experiments, the dilutions described in the figures were in addition to the dilutions mentioned here. Thus, a 10-fold transpososome dilution for cleavage (see Figure 3D) was really a 20-fold dilution, since the standard dilution was 2-fold.) Reactions were performed at room temperature (22 ± 1 °C), except the assembly steps of cleavage or strand transfer experiments, which were performed at 30 °C. This higher temperature enhanced the assembly rate (data not shown).

Gel Analyses. Assembly and cleavage experiments were analyzed on 8% acrylamide gels containing 0.05% SDS and $0.5 \times$ TBE buffer (45 mM Tris–borate, 1 mM EDTA). Samples were heated to 40 °C before being loaded onto the gel. After electrophoresis, gels were transferred to Whatman paper, dried, and exposed to a Molecular Dynamics phosphorimaging cassette. Strand transfer reactions were analyzed on agarose gels, as described (23).

Complex Stability Analysis. To assess the stability of transpososomes (see Figure 4B), MuA and DNA fragments were incubated for 2 days at room temperature, in the presence of 10 mM CaCl_2 . Complexes were then diluted 2-fold into a solution containing 70 mM MgCl_2 . Samples were withdrawn at $t = 0, 1$, and 2 h, and unlabeled wild-type DNA fragments were added to the samples to a concentration ~ 30 -fold higher than the concentration of labeled fragments. After being incubated on ice for 5 min, each sample was loaded onto a 2% MetaPhor agarose (BioWhittaker Molecular Applications) gel. Gels were run for 3 h at 60 V in $0.5 \times$ TBE buffer and visualized as described above.

RESULTS

Mutant DNA Fragments Are Severely Compromised Transposase Substrates. To dissect the role of sequence specificity in individual steps of transposition, it was necessary to design DNA substrates whose transposon sequences were seriously compromised but that could still transpose. We designed three DNA fragment sequences with mutations in the recognition sites that dramatically reduced but did not eliminate transposition (Figure 1C). For two of the sequences, the mutations were confined to the R1 site, as previous studies suggested that mutations in R1 would be more deleterious to transposition than mutations in R2 (32). The mutant I fragment was most severely disrupted close to the cleavage site, in the region of R1 that is contacted by domain I γ of MuA, although this fragment also contained less severe substitutions in the inner part of R1. Mutant I was the most active of the three mutant fragments (see below). The mutant II fragment was disrupted in the inner half of R1. This region is normally contacted by domain I β of MuA. The sequences for both mutants I and II were based on an "anticonsensus" sequence; that is, the substituted nucleotides were those that appear most infrequently in the natural MuA recognition sites (33). Finally, the mutant III fragment contained mild mutations throughout both the R1 and R2 sites. This sequence was discovered in a selection for non-Mu sequences that can be transposed, albeit weakly, by MuA (32).

Each of the three mutant DNA substrates was, as expected, a poor transposition substrate. Under standard reaction conditions, incubation of MuA with any of these fragments did not result in detectable cleavage even after a 24 h incubation (data not shown). However, the mutant DNA molecules could participate in transposition if the reaction was separated into stages and the reaction conditions were optimized for each stage (Figure 2 and also Figures 3–6 below).

To stage the reactions, we relied primarily on divalent metal ion requirements (data not shown, but see refs 7 and 16). Transpososome assembly with precleaved DNA fragments (fragments synthesized to terminate at the cleavage site) does not require divalent metal ion to be added to the reaction mixture. Ca^{2+} supports transpososome assembly on uncleaved fragments, but it does not support DNA cleavage. A 10 mM concentration of Mg^{2+} supports all three stages: transpososome assembly, DNA cleavage, and DNA strand transfer. Higher magnesium concentrations inhibit transpososome assembly but not DNA cleavage (T. Williams and T. A. Baker, unpublished results). Strand transfer, as measured here, requires addition of target DNA.

Mutant Fragments Assemble into Transpososomes Slowly and Incompletely. To test the influence of the DNA sequence on transpososome assembly, we used DNA cleavage as a readout for successful assembly (Figure 2A). Uncleaved DNA fragments were incubated with MuA for up to 27 h without divalent metal ion. At each time point, a sample was withdrawn, and cleavage was initiated by addition of Mg^{2+} and allowed to go to completion (2 h; see Figure 3). This high Mg^{2+} concentration also inhibited further assembly, as indicated by the absence of cleaved product in the 0 time point (Figure 2B); thus, the extent of cleavage reflected the success of transpososome assembly during the first incubation.

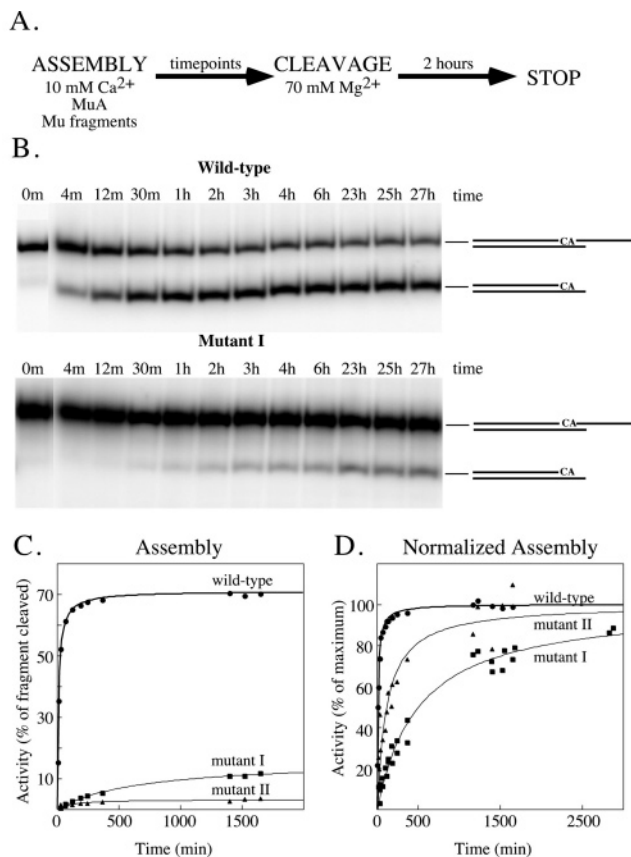


FIGURE 2: Mutant fragments assemble slowly into active complexes. (A) Summary of experimental design. This is a functional assay for assembly of active complexes. The readout is the extent of cleavage. (B) Acrylamide gel from an assembly experiment, showing bands of cleaved and uncleaved fragments. This panel and panel C are typical of three independent experiments. (C) Graph of assembly rates. This is a graph from a single assembly experiment, attained by analysis of gels such as that shown in (B). The data are fit to a second-order rate equation. (D) The relative assembly rate is slow with mutant fragments. This graph shows normalized results from three independent experiments for the wild-type and mutant I, and two experiments for mutant II. Each data point was normalized to the asymptote for its data set, again using a second-order rate equation. The normalization gives a sense of relative assembly rates, given that reactions with mutant fragments appear to reach completion with few active complexes assembled.

Both the mutant fragments I and II assembled only slowly into active complexes (Figure 2C,D). After 27 h, assembly on these mutant fragments had slowed considerably but not stopped (Figure 2D), with only 15% of the mutant I DNA or 5% of the mutant II DNA assembled into active complexes (Figure 2B,C). In contrast, assembly with wild-type sequences was complete within about 2 h, with an average of 70% of the fragments assembled into active complexes (Figure 2B,C).

Although reactions with mutant fragments had not reached completion after 27 h, the shape of the progress curves suggested that these reactions would never generate the number of active complexes observed with wild-type fragments (Figure 2C,D). We anticipated that the assembly reaction would be slow with these mutants; it was less certain that the reactions would also have a lower end point. The low end point is not simply due to loss of MuA activity, as adding additional MuA to these reactions for a second 24 h incubation prior to the initiation of cleavage did not increase the yield of cleaved product (data not shown). Rather, we

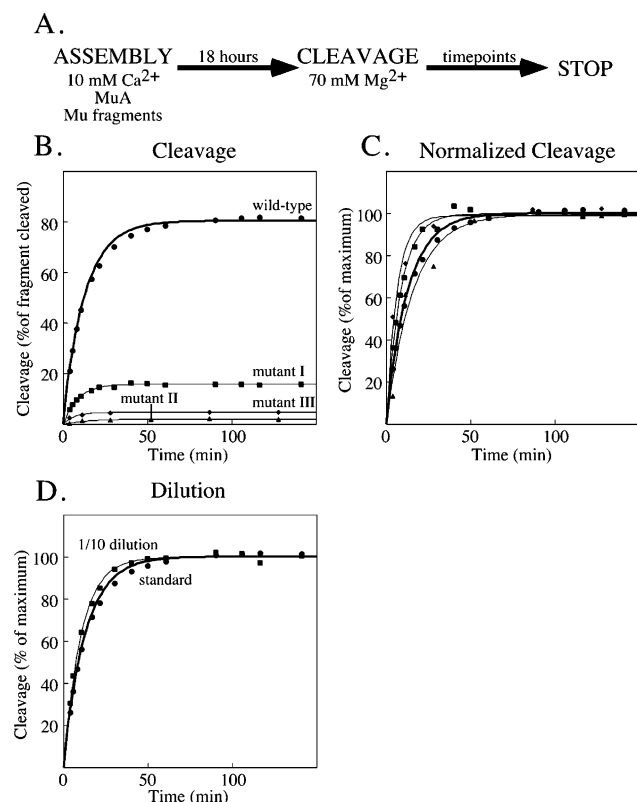


FIGURE 3: Mutant fragments are cleaved with rate constants similar to those of wild-type fragments. (A) Summary of experimental design. (B) Graph of one data set for each fragment type. Data are fit to a first-order equation. Table 2 summarizes results from multiple data sets. (C) Normalization reveals that mutant fragments are cleaved with relative rates similar to those of the wild type. These are the same data as shown in (B), but each point was normalized to the asymptote for its set. The wild-type data are highlighted with a dark line. (D) Dilution of preassembled transpososomes does not change the cleavage rate, supporting that the data fit first-order kinetics. For simplicity only the experiment with wild-type fragments is shown, but mutant I gave the same results.

suspect that the final extent is low because the mutant DNA fragments become trapped in inactive complexes, as discussed further below. Gel-retardation assays revealed that the yield of competitor-stable complexes was comparable to the yield of active complexes, whether on wild-type or mutant fragments (data not shown, but see Figure 4B below and also ref 23).

MuB Enhances the Assembly Rate on Wild-Type Sequences but Not on Mutant Sequences. The transposition-accessory protein MuB is present during transpososome assembly *in vivo*. Therefore, we also performed assembly assays in the presence of this protein. Previous experiments established that MuB directly stimulates transpososome assembly and strand transfer (18, 22, 34), and under some conditions also cleavage (13, 20). As expected, addition of MuB and ATP stimulated the rate of assembly on the wild-type sequences. Its presence, however, had no significant effect on assembly with mutant fragments (data not shown). Thus, either with or without MuB, the DNA sequence strongly contributes to the rate and efficiency of transpososome assembly.

The Recognition Site Sequence Does Not Contribute Directly to DNA Cleavage. To determine the importance of specific DNA sequence contacts to the cleavage reaction, we followed the rates of cleavage of preassembled complexes. Transpososomes were assembled for 18 h, and the

Table 2: Rate Constants for Cleavage Reactions

fragment	k^a	% fragment cleaved ^b	n^c
wild type	0.08 ± 0.01	74 ± 9	6
mutant I	0.12 ± 0.03	11 ± 5	3
mutant II	0.10 ± 0.06	2 ± 0	3
mutant III	0.15 ± 0.02	5 ± 1	2
C to T mutant	0.002 ± 0.001	62 ± 8	2
A•A mismatch	0.007 ± 0.001	45 ± 1	2

^a Rate constants were determined by fitting data like those of Figure 3B to the equation $[\text{fragment}]_{\text{cleaved}} = [\text{fragment}]_{\text{total}} - [\text{fragment}]_{\text{total}}e^{-kt}$.

^b This column is the asymptote of graphs such as the one shown in Figure 3B. ^c This column is the number of experiments.

cleavage time course was then initiated by addition of Mg²⁺ (Figure 3A). As expected for an activity of a preassembled complex, appearance of cleaved product fit well to a first-order rate equation (Figure 3B). This conclusion was supported by the results of a dilution experiment; diluting the assembly mixture an additional 10-fold when cleavage was initiated did not change the observed rate constants (Figure 3D and data not shown).

Remarkably, the first-order rate constants for each of the three mutant fragments were equal to or even slightly higher than that observed with wild-type fragments (Table 2). This result can be visualized by normalizing each data set to its asymptote and thereby comparing the relative cleavage rates for each fragment type (Figure 3C). Thus, although the absolute rates of cleavage were slow with mutant fragments (Figure 3B), the similarity in rate constants reveals that the slower reactions were due to a low concentration of active transpososomes, rather than a slow catalytic step.

MuA–DNA Complexes Are Stable during DNA Cleavage. The DNA cleavage reactions reached completion with only a small percentage of the mutant fragments cleaved (Table 2 and Figure 3B). These results are consistent with the assembly experiments, and are probably due to sequestration of the DNA fragments in inactive complexes. By fitting the cleavage data to a simple first-order scheme, we are assuming that the activity loss did not occur during the cleavage time course. Rather, we are assuming that inactivation occurred during the assembly time course or immediately upon initiation of cleavage. The only changes in reaction conditions from assembly to cleavage are the addition of Mg²⁺ and the 2-fold dilution of protein and DNA. Thus, given the lengthy assembly (18 h) and rapid cleavage (~1 h), it seems likely that the activity loss occurs primarily during the assembly stage. Nonetheless, the following experiments were designed to test for complex stability during the cleavage stage and confirm the validity of the cleavage first-order rate constants.

To test for inactivation of transpososomes during the cleavage stage, we assayed cleaved complexes for their ability to participate in strand transfer (Figure 4A). This experiment required three incubation periods: (i) complexes were assembled on uncleaved fragments (either the wild type or mutant I), (ii) Mg²⁺ was added, the complexes were allowed to complete cleavage, and a sample was taken to assay the extent of cleavage, and finally (iii) MuB and target DNA were added, and the complexes were allowed to complete strand transfer. As expected, only a small percentage of mutant DNA fragments were successfully cleaved.

A.

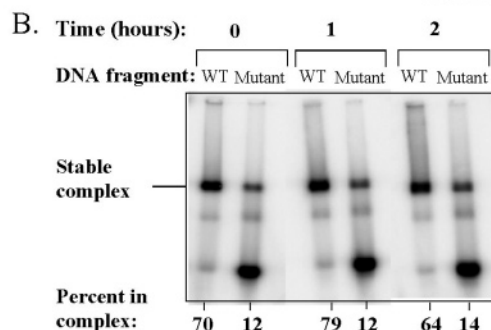
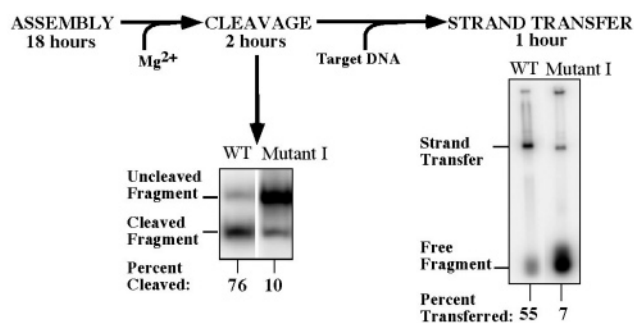


FIGURE 4: Transpososomes are stable over the course of a cleavage reaction. (A) Most complexes that are active for cleavage retain their activity for strand transfer. Complexes were assembled in the presence of 10 mM Ca^{2+} . Cleavage was initiated by adding Mg^{2+} to 70 mM. Strand transfer was initiated by adding target DNA, MuB, and ATP. (B) During the cleavage reaction, no significant changes in complex stability are visible on an agarose gel. Complexes were assembled in the presence of 10 mM Ca^{2+} . Cleavage was initiated by adding Mg^{2+} to 70 mM. Rather than assaying for cleavage, we assayed for complex stability by adding an excess of unlabeled DNA and loading samples onto a native agarose gel.

However, of those cleaved fragments, most ($69 \pm 1\%$) went on to complete strand transfer. The completion efficiency was similar with wild-type fragments: $71 \pm 12\%$ of those that were cleaved also completed strand transfer. Thus, the mutant fragments appear to partition into two groups during assembly. The majority (80–90%) of the mutant fragments partition into an inactive form. But a small fraction are incorporated into active transpososomes, and these remain active throughout the cleavage time course, as judged by their ability to participate in strand transfer.

To further test for transpososome stability during the cleavage reaction, we used a physical assay to directly quantify stable transposase–DNA complexes (Figure 4B). Complexes were assembled as if in preparation for a cleavage experiment. At time 0, the reaction was diluted 2-fold and Mg^{2+} was added, as during a cleavage experiment. At subsequent times, aliquots were removed, an excess of unlabeled DNA was added (as a competitor to trap any transposase that may dissociate from the labeled DNA), and samples were run on an agarose gel. Over a period of 3 h there was no significant decrease in the total number of visible protein–DNA complexes, indicating that the complexes were stable over the course of the cleavage reaction.

Together, the physical assay for complexes (Figure 4B), the strand transfer experiment (Figure 4A), and the dilution experiment (Figure 3D) strongly support the conclusion that cleavage occurs with first-order kinetics. This conclusion justifies our comparison of the rate constants observed for

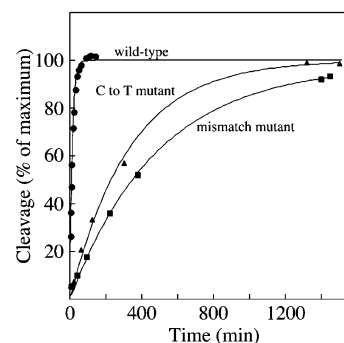


FIGURE 5: Mutations near the cleavage site slow the relative cleavage rate. These experiments follow the experimental design summarized in Figure 3A. The C to T mutant is a base pair substitution at the second position from the Mu DNA end; A•A mismatch is a single nucleotide substitution on the uncleaved strand at the first position. The results are detailed in Table 2.

mutant and wild-type DNA fragments (Table 2). We interpret the similarity in these rate constants as follows: although only a small percentage of the mutant fragments assembled into stable, active complexes, those complexes that did assemble were able to cleave DNA as well as the complexes formed with the wild-type Mu DNA sequence.

Mutations Near the Cleavage Site Slow the Rate of DNA Cleavage. The DNA recognition sites are located five base pairs away from the site of DNA cleavage. Though mutations in the recognition sites did not lower the rate constant for cleavage; mutations very near the cleavage site did. Specifically, either changing position 2 from a G•C base pair to an A•T (C to T mutant) or the introduction of a mismatched basepair (A•A mismatch) at position 1 (see DNA sequences in Figure 1B) slowed the cleavage rates more than 10-fold (Figure 5 and Table 2).

The Recognition Site Sequence Does Not Contribute Directly to Strand Transfer. To investigate the contribution of the MuA recognition site sequences to the strand transfer reaction, we used precleaved transposon fragments carrying either the wild-type or mutant DNA sequences. Assembly was conducted in the absence of divalent metal ions. Strand transfer was then initiated by diluting the assembly reaction 1 in 10 into a reaction mixture containing 70 mM MgCl_2 , target DNA, MuB, and ATP (Figure 6A). The dilution and the high Mg^{2+} concentration prevented further assembly of stable complexes (data not shown). The accumulation of strand transfer products fit well to a first-order equation (Figure 6B,C). In an experiment with wild-type fragments, diluting the assembly mixture an additional 10-fold upon initiation of strand transfer did not change the first-order rate constant (data not shown), further supporting the conclusion that strand transfer is an apparent first-order reaction.

The rate constant for strand transfer of mutant I fragments was indistinguishable from the rate constant for strand transfer with wild-type DNA sequences. As with cleavage, this similarity can be seen graphically by normalizing each data set to the asymptote for that set, and directly comparing the relative rates of strand transfer (Figure 6C). Thus, substantial mutations in the outer half of the R1 recognition site do not directly impact a fragment's ability to participate in strand transfer. Strand transfer is more difficult to quantify than cleavage, and of the mutant fragments, only mutant I gave data of sufficient quality for this analysis. The strand

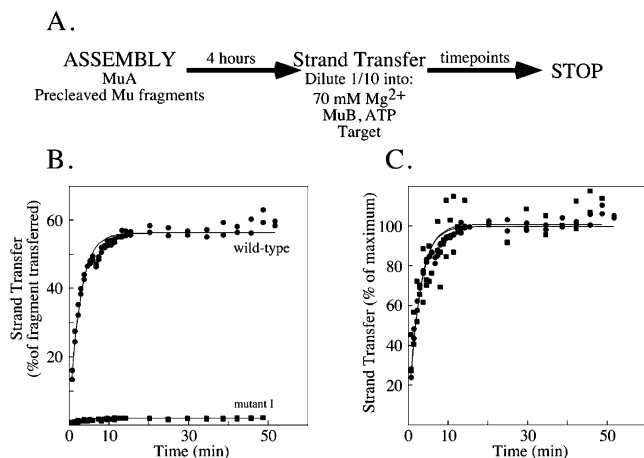


FIGURE 6: Mutant fragments participate in strand transfer with rate constant similar to that of wild-type fragments. (A) Summary of experimental design. (B) Graph of one data set for each fragment type. Each time point was analyzed on two separate agarose gels, and the data from both gels are included on the graph. Data are fit to a first-order equation. (C) Normalization reveals that mutant fragments are cleaved with relative rates similar to those of the wild type. These are the same data as shown in (B), but each point was normalized to the asymptote for its set.

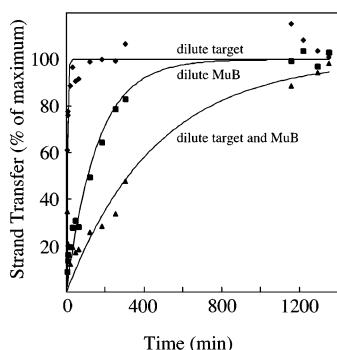


FIGURE 7: The strand transfer reaction rate is dependent on the concentration of MuB and target DNA. For simplicity we did not include a data set for the standard conditions, but the experiment in which target alone was diluted can be superimposed on the data points collected under standard conditions. By contrast, diluting MuB or target and MuB caused substantial rate reductions. The data are again normalized to the asymptote for each data set, because diluting MuB also reduces the final extent of the reaction. Thus, the absolute rates are even more dependent on MuB than are the relative rates represented here. For all experiments shown, transpososome dilutions were standard (see Figure 6A).

transfer rate constant for both this mutant and the wild-type sequence was 0.4 min^{-1} (± 0.02 for the wild type or ± 0.03 for the mutant). This value is roughly 4 times the cleavage rate constant (Table 2).

MuB Is Involved in the Rate-Determining Step for Strand Transfer. As discussed above, MuB protein did not stimulate transpososome assembly on the mutant DNA fragments. MuB also did not stimulate the rate of DNA cleavage with either the mutant or wild-type DNA fragments (data not shown). To complete this analysis of the impact of MuB, we investigated its influence on the rate of strand transfer. We found that both MuB and target DNA concentrations impacted the rate of strand transfer (Figure 7). In reactions with wild-type fragments, diluting MuB 10-fold compared to the standard conditions decreased the strand transfer rate roughly 50-fold. Diluting the target by the same amount had no effect, but when *both* MuB and target were diluted, the

rate decrease was roughly 3-fold greater than if MuB alone was diluted. One function of MuB during transposition is to bind and deliver a target DNA molecule to the transpososome (35). Our results suggest that interactions between the transpososome and a MuB–target DNA complex were rate limiting in the strand transfer experiments. The same step may well be rate limiting for strand transfer *in vivo*, since in the cell, MuB stimulation is required for robust transposition (36, 37).

DISCUSSION

MuA Recognition Sites Regulate Assembly. Once MuA transposase and the Mu DNA assemble into a transpososome, they remain in that complex until actively removed (11, 13, 27, 38). Thus, assembly of the transpososome is the first committed step in transposition. The MuA recognition sites, located at the ends of the Mu DNA, are essential for transpososome assembly, ensuring that only ends of the Mu genome assemble into transpososomes and are transposed (39). The current study presents the first kinetic analysis of transpososome assembly on mutant transposon end sequences, confirming expectations based on previous end-point-based studies. We find that mutations in the R1 recognition site dramatically reduce assembly rates. Undoubtedly, assembly in the absence of any intact MuA recognition sequences would be poorer still.

The slow assembly rate is not due to poor binding of MuA to the mutant DNA fragments. MuA has high nonspecific DNA-binding activity (40), and the mutant DNA fragments were fully occupied by MuA under our experimental conditions (unpublished results). Rather, past studies indicate that the recognition sites stimulate assembly by at least two mechanisms. Interactions with the recognition sequence induce allosteric changes in MuA that activate assembly, and the recognition sites help align the MuA subunits with respect to each other and with respect to the cleavage site DNA (23, 39, 41).

Only a small percentage of mutant DNA fragments successfully assembled into active transpososomes, even when the assembly reaction was nearly complete. Cleavage activity with these mutant fragments was not enhanced by the addition of more MuA for a second 24 h incubation prior to initiating cleavage. Thus, we could rule out two simple explanations for the low quantity of active mutant complexes: (i) MuA unfolds or aggregates during the long assembly incubation, or (ii) unlike for wild-type fragments, the disassembly rate for mutant fragments is substantial so that at equilibrium most fragments are not incorporated into complexes. If this second explanation were correct, addition of more MuA should shift the equilibrium in favor of transpososome assembly and increase the cleavage activity. We therefore conclude that the mutant DNA fragments are often incorporated into inactive complexes. We can only speculate as to the nature of these inactive complexes. They could be poorly ordered assemblies, or they could be synaptic complexes, perhaps with the DNA fragments bound in an orientation that did not permit them to be cleaved.

The MuA Recognition Site Sequence Is Not Critical during Postassembly Events. Mutations in the MuA recognition site sequences did not reduce the relative rates of DNA cleavage or of strand transfer. The absolute rates of reactions with

mutant DNA fragments were low because the concentration of active complexes at the start of each reaction was low. However, the first-order rate constants were similar for wild-type and for mutant fragments. We interpret these results as follows: although relatively few active transpososomes assembled with the mutant DNA fragments, those that successfully assembled were unimpaired during cleavage or strand transfer. These results suggest that, once a transpososome is assembled, interactions between the MuA DNA-binding domain and the DNA recognition sites do not have a significant impact on the catalytic domain.

One caveat to this conclusion is that we cannot say whether the rate-determining steps are the same here as under physiological conditions. We reported two experiments, one each for cleavage and for strand transfer, that help characterize a rate-determining step. We found that mutations in the DNA near the cleavage site slow the cleavage rates more than 10-fold (Figure 5 and Table 2). These results are consistent with the possibility that precise engagement of the DNA cleavage site in the protein's active site contributes to the observed reaction rate. Though still speculative, engagement of the cleavage site in the active site is a reasonable candidate for a physiologically relevant rate-determining step. We also found that strand transfer rates are enhanced by addition of MuB. Since *in vivo* transposition is enhanced by MuB (36, 37), this result suggests that the same step or steps may be rate determining *in vitro* as *in vivo*.

Nonetheless, it is possible that the *in vivo* rates are governed by steps different from those reflected in our experiments. In fact, cleavage reactions can be more rapid than reported here if the transposon sequences are contained within a circular plasmid (13). Thus, it is possible that recognition site mutations would slow the physiological rates of cleavage or of strand transfer, and that this slowing is masked in our experiments by a still slower step. However, if the mutations do slow a physiologically critical step, the effect must be mild for it to not have been detected here.

Interestingly, the rate constants for the cleavage step for two of the mutant fragments (mutant I and mutant III) were consistently slightly higher than for wild-type sequences. (The error in measurements with the third mutant was too great to detect subtle differences.) Mutations in the DNA recognition sites significantly destabilize transpososomes, as measured by the transpososomes' resistance to challenge with heparin (23). Our results suggest that the mutations used in this study destabilize the precleavage complexes (the "stable synaptic complex" or "type 0" complexes) more than they do the transition state for cleavage. As a result, the activation barrier for cleavage of mutant DNA fragments was actually slightly smaller than for cleavage of the wild-type DNA fragment, causing a slight enhancement in the cleavage rate.

Separation of Function between the DNA Recognition Site and Cleavage Site. In contrast to mutations in the recognition sites, mutations in the DNA near the cleavage site do lower the observed rate constant for DNA cleavage. The cleavage site and the recognition site are contacted by different domains of transposase (40, 42). In addition, the cleavage site sequence, (T or A)CA, is short compared to the 22 base pair recognition sequence, contributing very little toward identifying the ends of the Mu DNA in the context of the entire Mu genome. However, this sequence is important in

identifying the precise cleavage site within the local DNA sequence of an assembled transpososome (32).

Past studies have shown that the cleavage site sequence of the Mu DNA is important for assembly as well as for cleavage (43–45), but other transposons may not sense the cleavage site sequence during transpososome assembly. The second base pair of the Tn5 DNA has also been implicated in transpososome assembly (46). In contrast, the sequence of the first few base pairs of IS10 appears to be less important during transpososome assembly (47, 48). We speculate either that IS10 does not need to engage the cleavage site to assemble a transpososome or, more likely, that it relies on the correct base pair spacing between the recognition site and the end of the transposon DNA to ensure that the cleavage site is correctly engaged. As an interesting aside, MuA's repeated reliance on the cleavage site sequence helps explain why the cleavage site DNA has more restricted sequence requirements than any other location on the Mu DNA ends (32). Other elements (IS10, Tn5, and IS903) show stronger sequence requirements at specific locations within the recognition sites (47–51).

MuA's Mode of Sequence Recognition Diverges from Conventional Models. Our finding that the alterations in the MuA recognition site sequence do not lower the observed rate constant for catalysis is unusual. Conventional models of enzyme function, developed through studies of serine proteases, hexokinase, and other enzymes, propose that enzyme specificity and efficiency are inextricably linked (24, 25). Though an enzyme may interact nonspecifically with many different molecules, only specific binding interactions with the "true" substrate activate the enzyme's catalytic abilities. Among enzymes that interact with DNA, the best-studied restriction enzymes fit well with this classical view. Restriction enzymes bind with high affinity to noncognate as well as cognate DNA sequences (52, 53), but in several cases it has been shown that the mode of binding is unique at the cognate site. For example, *EcoRV* and *EcoRI* both severely distort the DNA at their cognate sites, and this distortion is necessary to activate DNA cleavage (54, 55). *BamHI* only mildly distorts its DNA site, but the protein itself is altered by cognate site binding as compared to noncognate binding (56). For all three of these restriction enzymes, it is believed that contacts made only in the cognate complex stabilize the transition state for cleavage, so that the free energy of interactions between the enzyme and its cognate site lowers the activation barrier to cleavage (26, 27, 57).

Our results with MuA transposase contrast with studies of conventional restriction enzymes, in that transposase's specificity for its recognition sites is not directly linked to catalytic efficiency. Once a transpososome is assembled, both the rate of cleavage and of DNA strand transfer are unaffected by heavy mutations in the recognition sequence. However, MuA transposase differs from these restriction enzymes in at least two relevant ways. (i) MuA is a modular protein (40), with distinct DNA recognition and catalytic domains. As a result, DNA cleavage occurs a considerable distance (five base pairs) outside of the DNA recognition sequence. This modularity may enable separation of function between DNA recognition and catalysis. Some restriction enzymes are also modular, but the restriction enzymes discussed above each contain a single domain, with regions

that function in sequence recognition interspersed with catalytic regions (27). (ii) Transposase performs its reactions when it is part of a stable transpososome, and it does not turn-over with each reaction cycle. Thus, transposition includes a discrete assembly phase, distinguishable from simple DNA binding. These two steps, simple binding and transpososome assembly, may be considered analogous to the first two stages of restriction cutting: simple, noncognate binding and active cognate binding. A restriction enzyme first binds DNA nonspecifically and then becomes activated via additional DNA contacts made upon binding the cognate site. Analogously, transposase first binds DNA in a non-committed fashion and then becomes activated via inter-subunit contacts made upon transpososome assembly.

MuA May Represent a Large Class of Proteins That Separate Substrate Recognition from Catalysis. Many other DNA-modifying enzymes share with MuA these traits of being modular, of requiring DNA sequence cues to trigger assembly of multimeric complexes, and of modifying DNA at a location that is distinct from the principal DNA recognition sequence. MuA is part of a large family of transposases that share these properties (58). Two families of site-specific recombinases, the serine family and the tyrosine family, are also modular proteins that function in synaptic complexes (59). In addition, there are also restriction enzymes that differ from the conventional prototypes discussed above and are more like MuA (60). For example, the type IIs enzyme, *FokI*, is modular, is monomeric in solution, and assembles into an active dimer only in the presence of its cognate DNA sequence (61, 62). The type IIf enzyme, *SfiI*, is a tetrameric restriction enzyme that synapses two cognate sequences to cleave DNA. DNA synapsis fails to occur if a noncognate sequence is bound (63).

These proteins are attractive candidates for being regulated by DNA specifically during complex assembly, in the same manner as MuA. For example, *FokI* is able to cleave DNA even if its DNA-binding domain is replaced with one from a related protein that recognizes a distinct sequence (64–67), demonstrating separation of function between sequence recognition and DNA cleavage. Similarly, chimeras between Gin invertase and ISXc5 resolvase, two related site-specific recombinases that recognize distinct sequences, can perform recombination (68). Phage λ integrase can catalyze simple DNA cleavage and joining reactions even in the absence of its specific recognition sequences (*att* sites) (69). However, the *att* sites are required for assembling a fully functional “intasome” complex (59). This observation suggests that the energy of specific binding is more important for establishing the structure of the λ integration complex than it is for performing covalent modification of DNA. λ integrase is also similar to MuA in that its full recombination pathway requires interactions between the protein and a series of long DNA recognition sequences, allowing the recombination site to be uniquely identified within a genome. In general, proteins that interact with long sequences may be particularly likely to be regulated through a protein–DNA complex assembly step, as it may be easier to evolve mechanisms for a long sequence to contribute to complex assembly than to contribute directly to catalysis of a DNA modification.

ACKNOWLEDGMENT

We thank Stephen Bell, Kiyoshi Mizuuchi, Felix Mueller-Planitz, Glenn Sanders, Robert Sauer, Frank Solomon, Kevin Travers, and especially Stephen Halford for helpful discussions and/or comments on the manuscript. T.A.B. is an employee of the Howard Hughes Medical Institute.

REFERENCES

1. Mizuuchi, K., and Adzuma, K. (1991) *Cell* 66, 129–40.
2. Mizuuchi, M., and Mizuuchi, K. (1993) *Cold Spring Harbor Symp. Quant. Biol.* 58, 515–23.
3. Haapa-Paananen, S., Rita, H., and Savilahti, H. (2002) *J. Biol. Chem.* 277, 2843–51.
4. Mizuuchi, K. (1992) *Annu. Rev. Biochem.* 61, 1011–51.
5. Chaconas, G. (1999) *Biochem. Cell Biol.* 77, 487–91.
6. Namgoong, S. Y., and Harshey, R. M. (1998) *EMBO J.* 17, 3775–85.
7. Williams, T. L., Jackson, E. L., Carritte, A., and Baker, T. A. (1999) *Genes Dev.* 13, 2725–37.
8. Leung, P. C., Teplow, D. B., and Harshey, R. M. (1989) *Nature* 338, 656–8.
9. Mizuuchi, M., and Mizuuchi, K. (1989) *Cell* 58, 399–408.
10. Craigie, R., and Mizuuchi, K. (1987) *Cell* 51, 493–501.
11. Surette, M. G., Buch, S. J., and Chaconas, G. (1987) *Cell* 49, 253–62.
12. Lavoie, B. D., Chan, B. S., Allison, R. G., and Chaconas, G. (1991) *EMBO J.* 10, 3051–9.
13. Mizuuchi, M., Baker, T. A., and Mizuuchi, K. (1992) *Cell* 70, 303–11.
14. Aldaz, H., Schuster, E., and Baker, T. A. (1996) *Cell* 85, 257–69.
15. Savilahti, H., and Mizuuchi, K. (1996) *Cell* 85, 271–80.
16. Savilahti, H., Rice, P. A., and Mizuuchi, K. (1995) *EMBO J.* 14, 4893–903.
17. Adzuma, K., and Mizuuchi, K. (1991) *J. Biol. Chem.* 266, 6159–67.
18. Baker, T. A., Mizuuchi, M., and Mizuuchi, K. (1991) *Cell* 65, 1003–13.
19. Surette, M. G., Harkness, T., and Chaconas, G. (1991) *J. Biol. Chem.* 266, 3118–24.
20. Naigamwalla, D. Z., Coros, C. J., Wu, Z., and Chaconas, G. (1998) *J. Mol. Biol.* 282, 265–74.
21. Kuo, C. F., Zou, A. H., Jayaram, M., Getzoff, E., and Harshey, R. (1991) *EMBO J.* 10, 1585–91.
22. Mizuuchi, M., Baker, T. A., and Mizuuchi, K. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 9031–5.
23. Goldhaber-Gordon, I., Williams, T. L., and Baker, T. A. (2002) *J. Biol. Chem.* 277, 7694–702.
24. Fersht, A. (1999) *Structure and Mechanism in Protein Science*, W. H. Freeman and Co., New York.
25. Kraut, D. A., Carroll, K. S., and Herschlag, D. (2003) *Annu. Rev. Biochem.* 72, 517–71.
26. Jen-Jacobson, L. (1997) *Biopolymers* 44, 153–80.
27. Pingoud, A., and Jeltsch, A. (1997) *Eur. J. Biochem.* 246, 1–22.
28. Levchenko, I., Luo, L., and Baker, T. A. (1995) *Genes Dev.* 9, 2399–408.
29. Krukltis, R., Welty, D. J., and Nakai, H. (1996) *EMBO J.* 15, 935–44.
30. Baker, T. A., Mizuuchi, M., Savilahti, H., and Mizuuchi, K. (1993) *Cell* 74, 723–33.
31. Yamauchi, M., and Baker, T. A. (1998) *EMBO J.* 17, 5509–18.
32. Goldhaber-Gordon, I., Early, M. H., Gray, M. K., and Baker, T. A. (2002) *J. Biol. Chem.* 277, 7703–12.
33. Schumacher, S., Clubb, R. T., Cai, M., Mizuuchi, K., Clore, G. M., and Gronenborn, A. M. (1997) *EMBO J.* 16, 7532–41.
34. Surette, M. G., and Chaconas, G. (1991) *J. Biol. Chem.* 266, 17306–13.
35. Adzuma, K., and Mizuuchi, K. (1988) *Cell* 53, 257–66.
36. Coelho, A., Maynard-Smith, S., and Symonds, N. (1982) *Mol. Gen. Genet.* 185, 356–62.
37. Chaconas, G., Gloor, G., Miller, J. L., Kennedy, D. L., Giddens, E. B., and Nagainis, C. R. (1984) *Cold Spring Harbor Symp. Quant. Biol.* 49, 279–84.
38. Nakai, H., and Krukltis, R. (1995) *J. Biol. Chem.* 270, 19591–8.
39. Baker, T. A., and Mizuuchi, K. (1992) *Genes Dev.* 6, 2221–32.

40. Nakayama, C., Teplow, D. B., and Harshey, R. M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 1809–13.
41. Namgoong, S. Y., Jayaram, M., Kim, K., and Harshey, R. M. (1994) *J. Mol. Biol.* 238, 514–27.
42. Davies, D. R., Goryshin, I. Y., Reznikoff, W. S., and Rayment I. (2000) *Science* 289, 77–85.
43. Coros, C. J., and Chaconas, G. (2001) *J. Mol. Biol.* 310, 299–309.
44. Lee, I., and Harshey, R. M. (2003) *J. Mol. Biol.* 330, 261–75.
45. Lee, I., and Harshey, R. M. (2001) *J. Mol. Biol.* 314, 433–44.
46. Yanagihara, K., and Mizuuchi, K. (2003) *Mol. Cell* 11, 215–24.
47. Jilk, R. A., Makris, J. C., Borchardt, L., and Reznikoff, W. S. (1993) *J. Bacteriol.* 175, 1264–71.
48. Huisman, O., Errada, P. R., Signon, L., and Kleckner, N. (1989) *EMBO J.* 8, 2101–9.
49. Haniford, D., and Kleckner, N. (1994) *EMBO J.* 13, 3401–11.
50. Derbyshire, K. M., Hwang, L., and Grindley, N. D. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 8049–53.
51. Makris, J. C., Nordmann, P. L., and Reznikoff, W. S. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 2224–8.
52. Phadnis, S. H., and Berg, D. E. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 9118–22.
53. Maxwell, A., and Halford, S. E. (1982) *Biochem. J.* 203, 93–8.
54. Engler, L. E., Welch, K. K., and Jen-Jacobson, L. (1997) *J. Mol. Biol.* 269, 82–101.
55. Kim, Y. C., Grable, J. C., Love, R., Greene, P. J., and Rosenberg, J. M. (1990) *Science* 249, 1307–9.
56. Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., and Wilson, K. S. (1993) *EMBO J.* 12, 1781–95.
57. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1995) *Science* 269, 656–63.
58. Vipond, I. B., and Halford, S. E. (1993) *Mol. Microbiol.* 9, 225–31.
59. Rice, P. A., and Baker, T. A. (2001) *Nat. Struct. Biol.* 8, 302–7.
60. Hallet, B., and Sherratt, D. J. (1997) *FEMS Microbiol. Rev.* 21, 157–78.
61. Bath, A. J., Milsom, S. E., Gormley, N. A., and Halford, S. E. (2002) *J. Biol. Chem.* 277, 4024–33.
62. Bitinaite, J., Wah, D. A., Aggarwal, A. K., and Schildkraut, I. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 10570–5.
63. Vanamee, E. S., Santagata, S., and Aggarwal, A. K. (2001) *J. Mol. Biol.* 309, 69–78.
64. Embleton, M. L., Williams, S. A., Watson, M. A., and Halford, S. E. (1999) *J. Mol. Biol.* 289, 785–97.
65. Kim, Y. G., and Chandrasegaran, S. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 883–7.
66. Huang, B., Schaeffer, C. J., Li, Q., and Tsai, M. D. (1996) *J. Protein Chem.* 15, 481–9.
67. Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156–60.
68. Kim, Y. G., Smith, J., Durgesha, M., and Chandrasegaran, S. (1998) *Biol. Chem.* 379, 489–95.
69. Schneider, F., Schwikardi, M., Muskhelishvili, G., and Droge, P. (2000) *J. Mol. Biol.* 295, 767–75.
70. Kikuchi, Y., and Nash, H. A. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 3760–4.
71. Craigie, R., Mizuuchi, M., and Mizuuchi, K. (1984) *Cell* 39, 387–94.

BI0353600