Flp catalyzes site-specific recombination in a highly sequence-specific manner despite making few direct contacts to the bases within its binding site. Sequence discrimination could take place in the binding and/or the catalytic steps. In this study, we independently measure the binding affinity and initial cleavage rate of Flp recombinase with ~20 designed alternate target DNA sequences. Our results show that Flp specificity is largely, although not entirely, imparted at the binding step and is the result of a combination of direct and indirect readout. The Flp binding site includes an A/T-rich region that displays a characteristically narrow minor groove. We find that many A → T changes are tolerated at the binding step, whereas C or G substitutions tend to decrease binding affinity. The effects of the latter can be alleviated by replacing guanine with inosine, which removes the N2 amino group that protrudes into the minor groove. Some A → T changes reduce binding affinity, due to clashing with nearby residues, reinforcing that specificity requires avoiding negative contacts as well as creating positive ones. A tracts, which can lead to unusually rigid DNA structure, are tolerated during the binding step when placed within the region where the minor groove is already narrow. However, most A tracts slow catalysis more than C or G substitutions. Understanding what kind of sequence variation is tolerated in the binding and catalytic steps helps us understand how the target DNA is recognized by Flp and will be useful in guiding the design of Flp variants with altered specificities.

Specificity in protein–DNA interactions is required for many cellular processes including transcription, DNA repair, and recombination. Proteins that modify DNA must distinguish their substrates from an overwhelming number of competing DNA sequences. In many cases, DNA-binding proteins recognize a specific DNA sequence even when there are few direct, unique contacts between the DNA bases and the amino acid side chains. Whereas it was once thought that unique hydrogen bond donor and acceptor patterns in the major groove would be the most important feature in DNA–protein recognition, further studies have revealed that other sequence-specific features of DNA often play a critical role. A large number of DNA-binding proteins have been shown to use indirect readout in their recognition of the appropriate DNA sequence, including the 434 repressor, TATA-binding protein, restriction enzymes including EcoRI, and the DNA bending proteins IHF and HU (1–6).

In this study we focus on understanding substrate recognition in a site-specific DNA recombinase. Flp is a tyrosine recombinase from Saccharomyces cerevisiae that specifically recognizes a pair of 13-bp inverted repeats and catalyzes recombination of the intervening sequence. The recombination reaction involves two rounds of DNA cleavage, strand exchange and religation (7). Flp is widely used as a biotechnology tool to knock out genes, especially in Drosophila. Efforts to design recombinases with altered specificity for biotechnology applications are ongoing (8, 9), and would be aided by further understanding of DNA recombinase recognition. There is no simple set of rules for engineering protein–DNA interfaces. A few examples of successful engineering of DNA-binding proteins exist, most of which relied on selection methods rather than rational design. Zinc finger designs have become sophisticated with the use of phage display and DNA-shuffling (10), Cre recombinase has been evolved to have different specificity (11, 12), and DNA polymerase has been converted to an RNA polymerase (13). The modular design of serine recombinases, where DNA binding and catalysis are physically separable, has made fusion proteins with transcription factor DNA binding domains replacing the DNA binding domains possible (8, 14). In the case of Flp, where DNA binding and catalysis are concentrated in one region of the protein, engineering has focused on a combination of selection and rational design (9).

Flp recognizes and cleaves DNA specifically at the Flp recognition target (FRT) site, despite the presence of only a few direct side chain to base contacts between Flp and the DNA (Fig. 1). This suggests that much of Flp-DNA recognition is driven by indirect readout, where sequence-dependent structure, flexibility, and other properties of the DNA create specificity, rather than unique contacts between amino acid side chains and DNA bases. The need for a narrow minor groove may explain the specificity for A:T base pairs in the Flp binding site near the scissile phosphate. A/T-rich regions typically have a narrow minor groove, because they are able to achieve a greater degree of propeller twist between adjacent base pairs than C/G pairs where the N2 amino group of the guanine clashes with the base in the n + 1 position on the opposite strand (15, 16). A:T base pairs are common among protein-DNA com-

* This work was supported by National Institutes of Health Grant GM058827. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom correspondence should be addressed: 929 E. 57th St., GCIS W136, Chicago, IL 60637. E-mail: price@uchicago.edu.

2 The abbreviations used are: FRT, Flp recognition target; WT, wild type; TAPS, 3-[(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)amino]-1-propanesulfonic acid; Tricine, N-tris(hydroxymethyl)methylglycine.
FIGURE 1. Flp–DNA interactions. A, a Holliday junction-bound Flp tetramer. The active monomers are yellow, the inactive monomers are blue, and the catalytic tyrosines (Tyr-343) are red. B, the DNA from a Flp tetramer, with the scissile phosphates marked by orange spheres. C, closeup view of Flp interacting with the arm highlighted in B. Green side chains are in proximity to DNA in D. A different view of the same interactions shown in C, with an A to T transversion modeled in position 3 (pink). The methyl group of T clashes with Ala-55 (green). Also shown is Met-58, which is poorly ordered but lies near the hydrophobic methyl group of T5. A was adapted from Ref. 7, using Protein Data Bank code 1FLO (28), and parts B–D made with PyMol (29), PDB 1M6X (22).
Flp Recombinase Specificity in Binding and Catalysis

**FIGURE 2. The FRT sequence.** A, the full FRT site, with arrows marking the 3 Flp-binding subites. Positions discussed in the text are numbered. B, the WT FRT suicide substrate used in binding and cleavage assays (variations are shown in Table 1). Vertical arrow marks the cleavage site, and several Flp side chains are shown next to the bases they are in close proximity to (see Fig. 1).

To gain insight into how Flp recognizes its DNA substrate, we designed the FRT DNA mutants shown in Table 1 for both binding and cleavage studies. Our analysis of the structure of the DNA in two Flp-DNA crystal structures revealed that the minor groove is unusually narrow in the A/T-rich region near the scissile phosphate (Fig. 3), and that there is an unusually large drop in the roll angle between the two consecutive A:T base pairs near the scissile phosphate.

**EXPERIMENTAL PROCEDURES**

WT Flpe and FlpeY343F (“FlpeYF”) were purified as previously described (21, 22).

*Binding Competition Assays—*PAGE-purified FRT substrates were ordered from IDT. Aliquots of Flpe were diluted in

---

**TABLE 1.** Positions discussed in the text are numbered.

| Position | Description |
|----------|-------------|
| 1–6      | WT base     |
| 1–6      | WT base     |
| 1–6      | WT base     |
| 1–6      | WT base     |

---

**FIGURE 3. Analysis of the Flp-bound DNA structure.** A, minor groove width. Note that it becomes narrow in the A/T-rich region near the scissile phosphate. The dashed line represents the average minor groove width for canonical B DNA. B, roll angles between base pairs. Both plots were made using 3DNA version 1.5 (30). Three examples of the Flp binding site were taken from structures determined in two different space groups (PDB codes 1FLO and 1M6X).
buffer containing 50 mM HEPES pH 7.1, 1 mM EDTA, 1 mM fresh dithiothreitol, 15% glycerol, 1 mM NaCl, 0.5 M (NH₄)₂SO₄, 0.2 mg/ml bovine serum albumin from fresh 10 mg/ml aliquot, and 2 g/ml salmon sperm DNA. A master mixture of 2/11003 buffer with 0.25 nM ³²P-labeled FRT DNA and 25 nM Flpe YF was mixed together. After a 30-min incubation at 30 °C, 5-l aliquots of the reaction mixture were distributed into tubes, 5 l of the appropriate cold competitor DNA were added (leading to final concentrations of 1, 9, 18, 60, 90, or 180 nM) and incubated at 30 °C for 30 min. Final reaction conditions were 5% glycerol, 0.1 mg/ml bovine serum albumin, 0.1 μg/ml salmon sperm DNA, 2 mM EDTA, 12.5 mM TAPS, pH 8, 1 mM fresh dithiothreitol, 100 mM NaCl, 50 mM (NH₄)₂SO₄, and 25 nM Flpe YF. 10-l samples were loaded onto a native 16×18-cm gel containing 10% 29:1 acrylamide/bisacrylamide (w/w), 5% glycerol, and 0.5× TBE (45 mM Tris borate, 1 mM EDTA) at 4 °C and run for 1 h at 150 V. Gels were visualized following exposure to PhosphorImager screens (Kodak) using a Bio-Rad phosphorimager. Free DNA and DNA-protein complex intensities in each lane were quantified with the volume analysis function in rectangle mode of Quantity One. Competition data were plotted with f, the ratio of fraction bound in the

FIGURE 4. Competition binding assays. A, example of a competition assay native gel, and B–E, plots of competition assays with FRT variants containing scissile phosphate position changes (B), A tracts (C), A to T transversions (D), and C or G substitutions (E). The fraction of radiolabeled WT FRT that is bound to Flpe Y343F decreases with increasing amounts of unlabeled competitor FRT DNA. Lines represent fitting the average and standard deviation of 3–6 experiments to an equation for the IC₅₀ (shown under “Experimental Procedures”).
Flp Recombinase Specificity in Binding and Catalysis

The presence of inhibitor divided by the fraction bound in the absence of inhibitor on the y axis and c, the concentration of cold competing DNA on the x axis. We used the fraction bound in the presence of 1 nM cold competitor DNA as the denominator in the ratio above, because the fraction bound in the absence of competitor DNA was artificially low, possibly because of low Flp solubility with low DNA or salt concentration. The IC\textsubscript{50} was determined by fitting the competition curves in Kaleidagraph to the following equation,

\[ f = \frac{[IC_{50}]/(c) + [IC_{50}]}{1 + \frac{[L^*]_{\text{tot}}(\phi + 2)}{2K_0(\phi + 1)} + \phi + K_c(\phi + 2)^{-1}} \]  

(Eq. 1)

where IC\textsubscript{50} is the concentration at which \( f = 0.5 \). Each IC\textsubscript{50} was fit to the average and standard deviation of at least three independent experiments, with the errors weighted in the fit. The equilibrium dissociation constant for the competitor, \( K_c \), cannot be determined from the IC\textsubscript{50} using the simple Cheng-Prusoff equation because more than 50% of the DNA is bound to Flp in the absence of competitor DNA. \( K_c \) was solved for using an equation that takes the initial fraction bound into account (23, 24),

\[ K_c = \frac{IC_{50} \times \left[ 1 + \left( \frac{[L^*]_{\text{tot}}(\phi + 2)}{2K_0(\phi + 1)} + \phi + K_c(\phi + 2)^{-1} \right) \right]^{-1}}{\text{l}} \]  

(Eq. 2)

where \([L^*]_{\text{tot}}\) represents the total amount of radiolabeled DNA present, \([L^*R]\) represents Flp-bound radiolabeled DNA, \( \phi = \frac{[L^*R]/[L^*]}{\text{L}} \), and represents the fraction bound without competitor, and the WT FRT \( K_c \) was solved for iteratively until \( K_c \) and \( K_c \) converged. Experiments were repeated a minimum of three times; error bars in Fig. 4 represent the standard deviation, and errors in Table 1 represent the standard error from the fit. Substrates with weak binding affinity were poorly determined in the range of concentrations we used for the competition assays, leading to greater errors in Table 1.

Catalysis Assays—Using the FRT substrates found to have \( K_c \)s less than \( \sim 100\text{ nM} \), catalysis assays were performed in 5% glycerol, 0.1 mg/ml bovine serum albumin, 0.1 \( \mu \text{g/ml} \) salmon sperm DNA, 2 mM EDTA, 12.5 mM TAPS, pH 8, 1 mM fresh dithiothreitol, 100 mM NaCl, 50 mM (NH\textsubscript{4})\textsubscript{2}SO\textsubscript{4} with <2 nM \( ^{32}\text{P} \)-FRT DNA and 100 nM WT Flpe. After establishing how long it took each substrate to reach 10% completion of the cleavage reaction, time points were chosen for each substrate, ranging from 10 s to 1 min, which would allow us to measure the initial rate. Reactions were stopped in 2 \( \times \) SDS buffer and run on 16% Tris Tricine mini gels with SDS buffer at 150 V for 40 min. Gels were exposed, scanned, and quantitated as described above, and the average of at least three experiments time points were fit to a straight line going through the origin with Kaleidagraph, using the errors to weight the overall fit. Error bars on \( v_c \) in Fig. 5 represent the standard error from the fit. To determine the cleavage rate constant, \( k_{\text{cleavage}} \), we divided the \( v_c \) by the fraction bound, as calculated based on the \( K_c \) with the following equation.

\[ [\text{DNA} - \text{Flp}] = (([[\text{DNA} + [\text{Flp} + K_c]]) \pm (([[\text{DNA} + [\text{Flp} + K_c])^2 - 4[[\text{DNA}][\text{Flp}]]^2/2]) \]  

(Eq. 3)

The errors were determined by numerically propagating the errors through on \( v_c \) and the fraction bound, and multiplying these errors in quadrature to obtain the error on \( k_{\text{cleavage}} \).

RESULTS

The equilibrium inhibition constants \( (K_c) \) we derived from competition binding assays are shown in Table 1, these equilibrium inhibition constants can be treated as equilibrium dissociation constants \( (K_d) \). To rule out any possible covalent species in our competition binding assays, we used a Y343F mutant, which lacks the catalytic tyrosine. In all of these experiments we used a variant of Flp, Flpe, containing four mutations that were evolved to give Flp greater stability at 37 °C (12, 25). The DNA substrates we used each contain a single Flp binding site, which is one-half of the inverted repeat found in the FRT site (Fig. 4). Using competition assays with increasing amounts of unlabeled competitor DNA and a single concentration of Flpe YF eliminated many of the problems we had with Flpe solubility and behavior. Competition assays were repeated 3–6 times, and a weighted fit of the averages with their errors was carried out using an equation for the 50% inhibition concentration (IC\textsubscript{50}) and converted to \( K_c \) values as described under “Experimental Procedures.” We also measured the catalytic efficiency of a subset of FRT variants with WT Flpe and a DNA suicide substrate containing the FRT variation of interest (Table 2). The DNA substrates used in both binding and catalysis studies have only two bases after the scissile phosphate, which are diluted into solution following cleavage during the catalysis experiments, resulting in a covalent bond between Flp and the DNA. Early cleavage time points were taken, and the average fraction DNA cleaved from 3 to 5 experiments was fit with a weight for the error to a line to obtain the initial velocity, \( v_0 \). The \( v_0 \) was converted to a cleavage rate constant by dividing by the fraction bound based on the concentration of Flp, DNA, and the \( K_c \) for that substrate (Fig. 5). The \( K_c \) for the WT FRT is 4.7 ± 1.3 nM, and the cleavage rate constant, \( k_{\text{cleavage}} \), in these conditions is 2.8 ± 0.06 \( \times \) 10\textsuperscript{3} s\textsuperscript{-1}.

Binding

Exploring the Scissile Phosphate Position—The change from C to T (FRT WT and FRT1ta) is well tolerated, leading to a very
small enhancement in affinity (Table 1 and Fig. 4B). Additionally, comparing the ΔΔG of FRT14ta and Frt4ta, which differ only by the presence of a T in the scissile phosphate position, shows almost no change. However both, G and A (FRT1gc and FRT1at) in the scissile phosphate position weaken binding significantly. The crystal structure shows that there is a hydrogen bond between Lys-82 and the N7 amino group of guanine across from the scissile phosphate position (Fig. 1).
H-bond would be preserved by a C:G → T:A change, as the A supplies a similar H bond acceptor in this position. However, C:G → A:T and C:G → G:C changes would disrupt this H-bond (Fig. 1). Interestingly, a K82M mutation can recombine both the WT FRT sequence and the C:G → G:C transversion in Frt1gc in an in vivo assay (26).

We also probed whether the scissile phosphate itself affects binding, by comparing WT sequence substrates that end with a hydroxyl and one that contains the scissile phosphate (FRT WToh and FRT WTP, Table 1). The lack of a scissile phosphate does deter binding somewhat (WT and FRT WTP, Table 1). The presence of the scissile phosphate returns binding to levels that are slightly better than WT WT with a two-base pair overhang.

Comparing A Tracts with the Sequential Changes of Which They Are Comprised—A tracts are stretches of 4–6 As or Ts, which give a rigid structure with a narrow minor groove that can result in DNA bending at their edges (27). Because A tracts do not contain pyrimidine-purine (Y-R) base steps such as T-A, which have been found to be more flexible than other dinucleotide steps, they may also be less amenable to induced fit in DNA-protein interactions. The pure stretches of As or Ts in the FRT substrates shown in Fig. 4c (as opposed to A/T mixtures), reveal that the position of the A tract within the FRT site has an enormous effect on binding. An A tract (FRT46ta), which still allows for a C in the scissile phosphate position but extends back further toward position 7 where the minor groove is wider in the Flp-DNA crystal structures (Fig. 3) drastically weakens binding, whereas the A tract positioned closer to the scissile phosphate (FRT14ta) has a much smaller effect on binding. The narrowing of the minor groove in an A tract increases in the 5'-3' direction for a stretch of As, meaning that the FRT46ta sequence may become extremely narrow where the WT sequence widens a bit with a C. Individually making the changes that add up to the A tract that destroys binding does not reveal an additive effect: the individual changes in FRT4ta and FRT6ta do not seriously affect binding, suggesting that the rigid nature of the full A tract in FRT46ta may be responsible for the decrease in binding activity. Other proteins with affinity for A/T-rich segments have struggled to bind to rigid A tracts. For example, whereas the TATA-binding protein functions well and can co-crystallize with many A → T transversions in its target sequence, it cannot function or co-crystallize with a rigid A tract (2).

### Table 2

|     | $K_a$  | Fraction bound | $v_0$  | $k_{\text{cleave}}$ |
|-----|--------|----------------|--------|---------------------|
| WT  | 4.7 ± 1.3 | 0.95 ± 0.012 | 2.7 ± 0.34 | 2.8 ± 0.06 |
| Frt1ta | 0.06 ± 0.02 | 0.99 ± 0.0002 | 3.1 ± 0.05 | 3.1 ± 0.1 |
| Frt1gc | 180 ± 2.2 | 0.36 ± 0.003 | 0.0108 ± 0.00047 | 0.03 ± 0.04 |
| Frt14ta | 17 ± 3.4 | 0.85 ± 0.025 | 0.29 ± 0.072 | 0.3 ± 0.01 |
| Frt46ta | 212 ± 5.9 | 0.32 ± 0.06 | 0.043 ± 0.0033 | 0.1 ± 0.03 |
| Frt4ta | 15 ± 2.2 | 0.87 ± 0.017 | 0.54 ± 0.0552 | 0.6 ± 0.04 |
| Frt6ta | 20 ± 2.1 | 0.83 ± 0.015 | 0.87 ± 0.014 | 1.0 ± 0.2 |
| Frt3at | 228 ± 26 | 0.30 ± 0.024 | 0.1 ± 0.017 | 0.33 ± 0.03 |
| Frt2at | 2.6 ± 1.1 | 0.97 ± 0.01 | 4.9 ± 0.28 | 5.0 ± 0.05 |
| Frt5at | 106 ± 30.2 | 0.48 ± 0.07 | 0.9 ± 0.11 | 1.9 ± 0.27 |
| Frt4cg | 49 ± 5.6 | 0.67 ± 0.025 | 0.64 ± 0.032 | 0.96 ± 0.04 |
| Frt3cg | 137 ± 0.33 | 0.42 ± 0.0006 | 0.5 ± 0.046 | 1.2 ± 0.02 |
| Frt3au | 43 ± 6.8 | 0.70 ± 0.03 | 0.9 ± 0.025 | 1.3 ± 0.07 |

### Figure 6

**Diagram showing A:T and G:C base pairs, along with the modified bases of uridine and inosine shown in A:T and G:C base pairs.** Uridine lacks the methyl group that may cause a steric clash with Ala-55 in the Frt3at substrate; a uridine-containing substrate (Frt3au) had much higher affinity for Flp. Inosine lacks the amine group of guanosine that normally prevents propeller twisting and narrowing of the minor groove. Note that the minor groove (lower) face of an I:C base pair also mimics that of an A:T base pair.

**A → T Transversions**—A → T changes preserve the minor groove hydrogen bond donor/acceptor pattern, and also the increased propeller twist that leads to a narrow minor groove. If these are the recognizable properties of A/T-rich regions, we should not expect A → T transversions to significantly affect binding. This is true for two of the four A → T transversions in the TATT sequence (FRT4ta and FRT2at), which have reasonable binding affinity. In fact, FRT2at has slightly improved binding affinity. However, the other two (FRT3at and FRT5at) have strongly reduced affinity, for position 3 even worse than a
CG substitution (Table 1). Examination of the Flp-DNA structure reveals that there are two side chains located quite close to the negatively affected A\textsuperscript{3}T transversions. In the case of FRT3at, modeling an A\textsuperscript{3}T transversion in this position reveals that the Ala-55 of Flp would be positioned very close to the methyl on the swapped T, leading to a steric clash (Fig. 1B).

To test the potential clash between Ala-55 and the methyl on the swapped T in FRT3at, we substituted uridine for thymine (FRT3au). We found that removing this methyl largely restored binding; the binding affinity of Frt3au is much stronger than for FRT3at. The binding defect of FRT5at is harder to explain. Met-58, although poorly ordered, is potentially within range to have some favorable hydrophobic interaction with the methyl of T5 (22). To test the potential interaction between Met-58 and the methyl on the swapped T in FRT3at, we substituted uridine for thymine (FRT3au). We found that removing this methyl largely restored binding; the binding affinity of Frt3au is much stronger than for FRT3at. The binding defect of FRT5at is harder to explain. Met-58, although poorly ordered, is potentially within range to have some favorable hydrophobic interaction with the methyl of T5 (22). To test the potential interaction between Met-58 and the methyl on the swapped T in FRT3at, we substituted uridine for thymine (FRT3au). We found that removing this methyl largely restored binding; the binding affinity of Frt3au is much stronger than for FRT3at.

These clashes are an important reminder that specificity comes from not only positive direct contacts between an enzyme and its substrate, but also from bad contacts that prevent the binding of the incorrect substrate. This likely explains why FRT235at4ta, with all As and Ts swapped, leads to extremely poor binding. Similarly, the poorly binding A tract FRT1235at includes the deleterious changes in FRT1at, FRT3at, and FRT5at.

Separate from our questions about Flp recognition, we noticed some differences in how FRT2at behaves in our binding and catalysis assays as compared with the full recombination assays used by Cox and co-workers (17). FRT2at is an alternative WT sequence found in one of the three Flp protomer binding sites that comprise a WT FRT site in the 2-μ plasmid of S. cerevisiae. (Note that there are only 3 known “WT” Flp binding sites, and thus we cannot do the sort of large scale comparisons possible for transcription factors such as TATA-binding protein (2).) The Cox laboratory (17) found that this alternative WT sequence hindered full recombination activity only when the change shown in FRT2at was present in both half-sites, on both sides of the spacer. Interestingly, we find that FRT2at has even stronger binding affinity than the WT FRT substrate. This implies that the defect in full recombination assays lies in other steps, such as the conformational changes involved in strand exchange and Holliday junction isomerization.

**Substituting C:G Pairs**—Replacing an A:T pair with a C:G pair, which presumably widens the minor groove through the addition of the N2 amino group of guanine, consistently weakens binding of FRT variants in positions 2–5 (Fig. 4E). To confirm that it is specifically the N2 amino group of guanine causing the decrease in binding, we chose two positions to examine with inosine:cytosine pairs, which are identical to G:C pairs but lack the N2 amino group of guanine (Fig. 6). These changes preserve the WT pattern of pyrimidine-purine steps as well. This probes whether the narrow minor groove is indeed the sequence-specific structural characteristic that Flp recognizes. We found that the I:C pairs do not disrupt binding nearly as much as the G:C pairs (FRT5ci and FRT3ci in Table 1), suggesting that the N2 amino group of guanine is responsible for a great deal of the decrease in affinity.

**Catalysis**

We chose 13 of the FRT substrates to compare the efficiency of the first cleavage step with the binding step (Table 2). Using radiolabeled DNA suicide substrates and WT Flpe, we took early time points to measure covalent cleavage product formation. It is technically difficult to measure catalysis with FRT substrates that have a low binding affinity, because it is difficult to achieve saturated binding conditions that are comparable with WT within a concentration range where Flp is soluble (<200 nM). Interestingly, the cleavage rate constants appear to...
Flp Recombinase Specificity in Binding and Catalysis

be more seriously reduced with the A tracts that can bind well (FRT14ta and FRT4ta) compared with other FRT substrates with poor binding due to C:G pair substitutions (FRT5cg, FRT4cg, and FRT3cg, see Fig. 5). Compared with WT, FRT14ta has a ~3-fold decrease in binding affinity, but a ~10-fold decrease in cleavage rate constant. FRT3cg has a ~30-fold decrease in binding but just a ~3-fold decrease in cleavage rate constant. Examining the structure of FRT DNA before and after cleavage shows that the cleaved DNA kinks at the site of the scissile phosphate (Fig. 7). This suggests that the scissile phosphate may need to be in a specific conformation that a rigid A tract would interfere with, inhibiting cleavage more than poor binding from a C/G substitution does.

Most of the differences in cleavage rate constants for the different FRT substrates are not significant on a logarithmic scale. This implies that many of our DNA sequence changes have not affected the Achilles heel of the reaction because they do not critically hinder cleavage. However, there are some changes that do significantly affect catalysis: Frt1gc, with a C → G change right at the scissile phosphate, essentially prevents cleavage from taking place. When combined with the binding data, the catalysis data reveal that whether discrimination takes place primarily in the binding or catalytic step depends on the position of the sequence change. Maintaining a narrow minor groove with a well placed A tract allows good binding, but hinders catalysis, likely due to the rigid nature of A tracts affecting the movement of the substrate in or before the transition state. Interestingly, the substitution of C:G pairs to widen the minor groove interferes with binding but not as much with catalysis. However, the C:G → G:C change in Frt1gc is devastating at both the binding and the catalytic steps.

DISCUSSION

We measured the binding equilibrium and the apparent rate of the first chemical step in the recombination reaction, so that we can independently compare how the changes in DNA sequence affect the binding and catalysis. Our results show that Flp specificity is largely imparted at the binding step. Specificity can be thought of as the ability to discriminate between substrates in the ground state, or the competition of substrates for an enzyme. Because many DNA-binding proteins do not have a catalytic function beyond binding to a particular DNA sequence, and others have unusually divorced specificity in their binding and catalytic steps, we could have found that sequence changes had no impact on binding and only affected catalysis or vice versa.

There are a few intriguing instances where Flp substrates that bind well are cleaved poorly, suggesting that the sequence requirements for optimal binding and catalysis are not identical. Some changes in the DNA sequence may affect the ability of the DNA-protein complex to reach the transition state, explaining why a rigid A tract such as FRT14ta that can bind well is a poor cleavage substrate.

Our results also show that Flp specificity is the result of a combination of direct and indirect readout, but that the details of the indirect readout are more subtle than initially thought. For example, the single C:G → T:A and A:T → T:A changes in FRT4ta and FRT6ta are only mildly deleterious, but when they are combined in FRT46ta, the effect is far more than additive. FRT46ta contains a bone fide “A tract,” whose special structure may be incompatible with that of the Flp-DNA complex. However, FRT14ta, where the A tract is moved only one position to the right, binds reasonably well. This can be rationalized by the crystal structure, where the minor groove becomes increasingly wide (and thus A tract-incompatible) toward position C7. Disrupting the A/T-rich region near the scissile phosphate through the substitution of C:G pairs worsens binding, whereas A/T transversions and replacements with inosine:cytosine pairs have much smaller impacts on binding. This suggests that Flp does indeed prefer sequences that can easily adopt a narrow minor groove in this region. Some A/T transversions hindered binding, which can be attributed to clashes with nearby Flp side chains. While focusing on how specificity is created by positive interactions it can be easy to forget that specificity also comes from avoiding negative interactions so that the incorrect substrate does not bind. In addition, whereas binding is disrupted by the substitution of C:G pairs in the A/T-rich region of the FRT site, catalysis seems to be much less hindered. Even the A tract FRT14ta, which binds similarly to WT FRT, is outperformed by FRT4cg in the catalysis assays, where the substitution of a C:G pair does disrupt binding. Flp sequence discrimination in this region near the scissile phosphate is thus imparted at both the binding and catalytic steps, depending on the change that is made.

Acknowledgments—We thank Ying Zhang Pigli and Mark Gaijjar for technical assistance and Kent Mouw for helpful discussions. Joe Pecirilli and Shohei Koide made insightful suggestions. Yuen-Ling Chan contributed electrophoresis advice and Daniel Whiteson provided statistical guidance for the data analysis.

REFERENCES

1. Mauro, S. A., Pawlowski, D., and Koudelka, G. B. (2003) J. Biol. Chem. 278, 12955–12960
2. Patikoglou, G. A., Kim, J. L., Sun, L., Yang, S. H., Kodakek, T., and Burley, S. K. (1999) Genes Dev. 13, 3217–3230
3. Wenz, C., Jeltsch, A., and Pingoud, A. (1996) J. Biol. Chem. 271, 5565–5573
4. Kurpiewski, M. R., Engler, L. E., Wozniak, L. A., Kobylnska, A., Kozikiewicz, M., Stec, W. J., and Jen-Jacobson, L. (2004) Structure 12, 1775–1788
5. Lynch, T. W., Read, E. K., Mattis, A. N., Gardner, J. F., and Rice, P. A. (2003) J. Mol. Biol. 330, 493–502
6. Swinger, K. K., and Rice, P. A. (2007) J. Mol. Biol. 365, 1005–1016
7. Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006) Annu. Rev. Biochem. 75, 567–605
8. Gordley, R. M., Smith, J. D., Graslund, T., and Barbas, C. F., 3rd (2007) J. Mol. Biol. 367, 802–813
9. Voziyanov, Y., Konieczka, J. H., Stewart, A. F., and Jayaram, M. (2003) J. Mol. Biol. 326, 65–76
10. Pabo, C. O., Peisach, E., and Grant, R. A. (2001) Annu. Rev. Biochem. 70, 313–340
11. Santoro, S. W., and Schultz, P. G. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 4185–4190
12. Buchholz, F., and Stewart, A. F. (2001) Nat. Biotechnol. 19, 1047–1052
13. Xia, G., Chen, L., Sera, T., Fa, M., Schultz, P. G., and Romesberg, F. E. (2002) Proc. Natl. Acad. Sci. U. S. A. **99**, 6597–6602

14. Akopian, A., He, J., Boocock, M. R., and Stark, W. M. (2003) Proc. Natl. Acad. Sci. U. S. A. **100**, 8688–8691

15. Fratini, A. V., Kopka, M. L., Drew, H. R., and Dickerson, R. E. (1982) *J. Biol. Chem.* **257**, 14686–14707

16. Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1995) *J. Mol. Biol.* **247**, 34–48

17. Senecoff, J. F., Rossmeissl, P. J., and Cox, M. M. (1988) *J. Mol. Biol.* **201**, 405–421

18. Fersht, A. (1998) *Structure and Mechanism in Protein Science*, 3rd Ed., W. H. Freeman and Co., New York

19. Kraut, D. A., Carroll, K. S., and Herschlag, D. (2003) *Annu. Rev. Biochem.* **72**, 517–571

20. Goldhaber-Gordon, I., Early, M. H., and Baker, T. A. (2003) *Biochemistry* **42**, 14633–14642

21. Chen, Y., and Rice, P. A. (2003) *J. Biol. Chem.* **278**, 24800–24807

22. Conway, A. B., Chen, Y., and Rice, P. A. (2003) *J. Mol. Biol.* **326**, 425–434

23. Krohn, K. A., and Link, J. M. (2003) *Nucl. Med. Biol.* **30**, 819–826

24. Munson, P. J., and Rodbard, D. (1988) *J. Recept. Res.* **8**, 533–546; Erratum (1989–1990) *J. Recept. Res.* **9**, 511

25. Buchholz, F., Angrand, P. O., and Stewart, A. F. (1998) *Nat. Biotechnol.* **16**, 657–662

26. Voziyanov, Y., Stewart, A. F., and Jayaram, M. (2002) *Nucleic Acids Res.* **30**, 1656–1663

27. Barbic, A., Zimmer, D. P., and Crothers, D. M. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2369–2373

28. Chen, Y., Narendra, U., Iype, L. E., Cox, M. M., and Rice, P. A. (2000) *Mol. Cell.* **6**, 885–897

29. DeLano, W. L. (2002) *The PyMol Molecular Graphics System*, DeLano Scientific, Palo Alto, CA

30. Lu, X. J., and Olson, W. K. (2003) *Nucleic Acids Res.* **31**, 5108–5121

**Flp Recombinase Specificity in Binding and Catalysis**