Published in final edited form as:

J Mol Biol. 2008 September 26; 382(1): 188–202. doi:10.1016/j.jmb.2008.07.010.

Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA nicking activity

Yan Niu¹, Kristen Tenney¹, Hongye Li², and Frederick S. Gimble¹

¹Department of Biochemistry, Purdue University, West Lafayette, Indiana, 47907

²Center for Genome Research, Institute of Biosciences and Technology, TAMHSC, Houston, Texas, 77030

Summary

The number of strand-specific nicking endonucleases that are currently available for laboratory procedures and applications in vivo is limited, and none is sufficiently specific to nick single target sites within complex genomes. The extreme target specificity of homing endonucleases makes them attractive candidates for engineering high specificity nicking endonucleases. I-SceI is a monomeric homing enzyme that recognizes an 18 bp asymmetric target sequence, and cleaves both DNA strands to leave 3'-overhangs four base-pairs in length. In single turnover experiments using plasmid substrates, I-SceI generates transient open circle intermediates during the conversion of supercoiled to linear DNA, indicating that the enzyme sequentially cleaves the two DNA strands. A novel hairpin substrate was used to demonstrate that although wild-type I-SceI cleaves either the top or bottom DNA strand first to generate two nicked DNA intermediates, the enzyme has a preference for cleaving the bottom strand. The kinetics data are consistent with a parallel sequential reaction mechanism. Substitution of two pseudo-symmetric residues, Lys-122 and Lys-223, markedly reduces top or bottom-strand cleavage, respectively, to generate enzymes with significant strand- and sequence-specific nicking activity. The two active sites are partially interdependent since alterations to one site affect the second. The kinetics analysis is consistent with X-ray crystal structures of I-SceI/DNA complexes that reveal a role for the lysines in establishing important solvent networks that include nucleophilic water molecules thought to attack the scissile phosphodiester bonds.

Keywords

homing endonuclease; strand-specific endonuclease; protein engineering; enzyme kinetics; active site mutagenesis

Introduction

DNA endonucleases that nick one DNA strand rather than cleave both strands have become effective tools for laboratory procedures that require the removal or synthesis of DNA on one specific strand and for studies of DNA repair in vivo. The Nt.BstNBI restriction enzyme has been used in a novel isothermal amplification method to generate short DNA oligonucleotides¹. Similarly, the Nt.CviPII nicking endonuclease has been used in combination

© 2008 Elsevier Ltd. All rights reserved.

Correspondence to: Frederick S. Gimble.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

with a DNA polymerase in a single step isothermal reaction to randomly amplify genomic DNA². Sequence- and strand-specific nicking endonucleases have also facilitated the study of DNA excision repair pathways because they can be used to generate pre-nicked DNA substrates that harbor DNA lesions at defined positions³. Some of the currently available nicking endonucleases have been engineered from naturally occurring restriction enzymes by inactivating or replacing one of two subunits that each cleave one DNA strand, or by applying selection strategies to isolate mutant variants of enzymes that nick one strand specifically^{4; 5; 6; 7}. It has been possible in some instances to isolate single subunits of heterodimeric enzymes that possess strand-specific nicking activity⁸. All of the engineered nicking enzymes interact with short recognition sequences that are usually seven or fewer base-pairs in length. The frequent occurrence of these sequences limits their application for methods that require nicking at single loci within complex genomes.

Protein engineering of homing enzymes has the potential to create sequence- and strand-specific DNA endonucleases with high target specificity. Homing endonucleases are a class of enzymes that recognize DNA target sequences >14 bp in length, and can be divided into families based on conserved sequence and structural motifs⁹. I-SceI belongs to the LAGLIDADG family and is encoded by an intron (ω) situated within the 21S mitochondrial RNA gene of *Saccharomyces cerevisiae*¹⁰. LAGLIDADG homing endonucleases exist as homodimers that recognize symmetric substrates or as pseudosymmetric monomers that interact with asymmetric recognition sequences. I-SceI is a monomeric enzyme whose overall topological structure is similar to other LAGLIDADG homing enzymes¹¹ (Figure 1). Two signature α -helices located near its center of pseudo-symmetry define the hydrophobic core of the protein. The DNA-binding interface responsible for the sequence-specific and phosphate backbone contacts the 18 bp I-SceI recognition sequence using a saddle-like structure comprised of anti-parallel β -sheets.

I-SceI cleaves its DNA recognition sequence across the minor groove at two scissile phosphates to yield a four base 3'-overhang¹⁰. This enzyme is believed to use a variant form of a two-metal cleavage mechanism initially proposed for the I-CreI homing endonuclease¹². In this scheme, three divalent metals coordinated by the scissile phosphates and by two acidic residues located at the carboxyl-termini of the LAGLIDADG helices comprise part of two overlapping active sites that each cleave one DNA strand. All enzymatic activity is abolished when the acidic side chains are substituted with non-conservative residues^{13; 14}. Only two of the three metals are observed in a X-ray crystal structure of I-SceI bound to uncleaved DNA¹¹. However, in a recent structure of I-SceI complexed to a nicked DNA substrate in which the bottom strand has been cleaved, all three metals are coordinated by Asp-44 and Asp-145 at the expected positions¹⁵. In each active site, one bound metal ion coordinates a water molecule that is correctly positioned to effect in-line hydrolytic attack of the scissile phosphate group. A likely role of these "unshared" metal ions is to activate the nucleophilic water molecule. A third metal ion located between the two unshared metals is coordinated by the carboxylate oxygens of both conserved acidic residues and by the oxygen atoms of the scissile phosphates of each DNA strand. This "shared" metal ion is believed to stabilize the phosphoanion transition state and the 3'-hydroxylate leaving group at each active site.

Kinetics studies of wild-type I-SceI indicated that cleavage of the I-SceI substrate was biphasic, consisting of a rapid initial phase of product formation in which the amount of product was proportional to the amount of enzyme present, followed by a second phase in which product formation was slower¹⁶. This kinetics profile is mainly caused by the fact that I-SceI remains tightly bound to one of the two DNA cleavage products following double strand cleavage through contacts made by the amino-terminal domain^{11; 16}. (Figure 1) Hairpin substrates were used in kinetics analyses to demonstrate that I-SceI cleaves the bottom strand faster than the top strand to generate a nicked reaction intermediate¹⁶. However, no intermediate was

detected corresponding to one with a nick in the top strand, suggesting that bottom strand cleavage was an obligatory precursor step to top strand cleavage in the reaction pathway.

In this study, we engineered an I-SceI homing enzyme derivative with site- and strand-specific DNA nicking activity. We employed a novel internally radio-labeled hairpin substrate to confirm that I-SceI cleaves the bottom DNA faster than the top strand, and we present the first evidence that the enzyme can nick the top strand first. Single turnover kinetics data using wild-type I-SceI are most consistent with a parallel-sequential reaction pathway in which either DNA strand can be cleaved first or second. We used site-directed mutagenesis to replace two pseudosymmetric lysine residues that lie at the periphery of the catalytic center and contribute to the ordering of a network of water molecules that are involved in catalysis. Replacing the lysines leads to complementary reductions in active site efficiency, yielding nicking endonucleases that selectively cleave one DNA strand or the other.

Results

Comparison of LAGLIDADG endonuclease structures to identify conserved amino acids

To identify amino acid residues that play roles in the DNA cleavage mechanism, the X-ray crystal structure of I-SceI was superposed on five other LAGLIDADG homing endonucleases whose structures have been determined in a complex with their DNA recognition sequences. Structural alignments using the conserved LAGLIDADG helices showed high levels of structural similarity ((Figure 2) r. m. s. deviation: 1.03–1.27 Å). The two conserved acidic residues located at the carboxyl-termini of the LAGLIDADG α -helices have 0.51–1.6 Å r. m. s. deviation values. D44A and D145A mutant proteins were generated, purified, and assayed for their ability to cleave supercoiled pBS I-SceI (E/H) DNA. No detectable double or single-strand cleavage activity was evident for either mutant in the presence of magnesium (data not shown). These findings are consistent with the essential role that the acidic residues play in coordinating the shared and unshared metal ions that are required for catalysis. Mutation of analogous acidic residues in other LAGLIDADG proteins also abolishes DNA cleavage activity ^{12; 13; 17; 18; 19; 20}.

The structural alignment reveals that the only other amino acids that are conserved among the five homing endonucleases are the pseudo-symmetric Lys-122 and Lys-223 residues (Figure 2). Although the entire lysine side chains do not superpose well (r.m.s. deviation 2.33–4.69 Å), the ϵ -amino groups overlap more closely. In the structure of I-SceI complexed to its uncleaved DNA recognition sequence, the Lys-223 side chain extends from α -helix 12 and hydrogen bonds to a water molecule that is coordinated to one of the metal ions and is the proposed nucleophile that initiates scission of the bottom DNA strand. Lys-122 extends from a short β -strand between α -helices 7 and 8 but does not hydrogen bond to any metal-coordinated water molecules. The structurally analogous lysines in the other homing endonucleases are found either in loops or α -helices that differ in length between the proteins. As in the case of I-SceI, these lysines are part of a basic pocket in which the side chain hydrogen bond to solvent molecules rather than directly binding to the metal ions.

DNA cleavage of a supercoiled plasmid substrate by wild-type I-SceI and mutant variants

The location of Lys-122 and Lys-223 on the periphery of either one of the two overlapping active sites suggests that unlike the conserved acidic residues that are essential for cleavage of both DNA strands, they may be involved in DNA cleavage of one strand only. I-SceI protein variants containing substitutions for Lys-122 and Lys-223 were generated by site-directed mutagenesis, and the wild-type and mutant proteins were purified. Lys-122 and Lys-223 mutant proteins contain a carboxyl-terminal his₆ extension for purification that does not affect activity (data not shown). We used the K122I and K223I proteins in this study, but the kinetics profiles

for the alanine- and methionine-substituted proteins yielded similar results (data not shown). The extended hydrophobic side chain of isoleucine is similar to that of lysine, but lacks an ϵ -amino group responsible for hydrogen bonding.

Previous studies revealed that turnover of I-SceI following DNA cleavage is extremely slow due to the fact that the protein remains tightly bound to the downstream DNA product¹⁶ (Figure 1). For the analysis presented here, single-turnover experiments were performed in which enzyme was in molar excess over substrate ($[E_0] \gg [S]$) in order to reveal the different enzyme-bound states of the DNA as it progresses along the reaction pathway. In initial experiments, a supercoiled plasmid (SC) that contains a single I-SceI recognition sequence was used as a substrate. If the enzyme sequentially cleaves either the top or bottom DNA strands, rather than cleaving them concertedly, a nicked, open circle reaction intermediate (OC) would result that would migrate slower in agarose gels than supercoiled DNA. Subsequent cleavage of the second DNA strand of the OC intermediate would generate the linear DNA product (LIN), which migrates between the SC and OC forms of the DNA. When reaction rates are rapid, it is necessary to use a quench-flow apparatus in order to initiate and terminate the enzymatic reaction by rapid mixing. Under the conditions described in this report, the slow progress of the I-SceI-mediated reaction permitted manual manipulation to be used.

The equilibrium dissociation constant for the interaction between I-SceI and a 54 bp linear duplex containing the I-SceI recognition sequence, as measured by electrophoretic mobility shift assays, is 0.8 nM²¹. In the DNA cleavage assays reported here, 2.5 nM plasmid and 100 nM protein concentrations were used, resulting in > 99% of the DNA being present as the enzyme-bound complex. In order to ascertain whether the initial bimolecular interaction between I-SceI and the DNA is the rate-limiting step of the reaction pathway, the reaction rate was measured either by adding $MgCl_2$ to a mixture of DNA and I-SceI or by adding I-SceI to a mixture of $MgCl_2$ and DNA. The reaction rate was similar regardless of the order of addition (data not shown). If I-SceI binding to the DNA were the rate-limited step, a lag phase prior to DNA cleavage would have been evident when the reactions were initiated by adding I-SceI to a mixture of $MgCl_2$ and DNA, but not when $MgCl_2$ was added to a mixture of DNA and I-SceI. In our experiments, the reactions were initiated by adding $MgCl_2$ to a pre-equilibrated mixture of DNA and I-SceI.

The amount of the SC plasmid declines exponentially when it is incubated with I-SceI, initially resulting in the accumulation of the OC form (Figure 3). A subsequent decrease of the OC form occurs concomitantly with the appearance and increase of the LIN product, which is only generated after a lag period of several seconds. Taken together, this pattern is consistent with a sequential reaction pathway in which I-SceI cleaves one DNA strand to yield the OC intermediate and then the second strand. Thus, under single turnover conditions, the reaction pathway can be represented by the following scheme that involves two consecutive cleavage steps:

where ES, EO and EL are the enzyme-bound forms of the supercoiled, open circle and linear DNA, respectively and k_a and k_b are the apparent rate constants for the first and second steps of the pathway.

DNA cleavage of the plasmid substrate by the K122I and K223I variants was also examined under single-turnover conditions with enzyme in excess of the plasmid pBS I-SceI (E/H). In reactions using either variant protein, exponential decrease of the SC DNA form occurred faster than observed for the wild-type protein. The SC form was completely converted to other forms

by wild-type I-SceI after 20 minutes, but was totally converted by the K122I and K223I proteins after 5 minutes. This indicates that the rate of cleavage of the first strand by the mutant proteins occurs faster than by wild-type I-SceI. However, this experiment does not reveal whether one or both strands are cleaved faster. In contrast, the accumulation of the LIN product form, which is complete for wild-type I-SceI after 20–30 minutes, required several hours for completion by the two mutant proteins. This observation suggests that cleavage of the second strand by the mutant proteins takes longer than the wild-type protein. The most distinguishing feature of the reactions catalyzed by the mutant proteins is the large accumulation of the OC intermediate forms. Unlike the reaction catalyzed by wild-type I-SceI, in which all three forms of DNA were present midway through the reaction, all of the SC DNA is converted to the OC form by the mutant proteins prior to appearance of the LIN product. The eventual conversion of the SC DNA to the LIN product indicates that neither lysine residue is essential for double-strand cleavage activity. However, the large accumulation of the OC form and the longer time required for production of the linear DNA indicates that the mutant enzymes cleave the second DNA strand significantly slower than wild-type I-SceI.

DNA target site specificity of the K122I and K223I mutant proteins

The amino acid substitutions introduced into the K122I and K223I proteins may affect their DNA binding specificities as well as their DNA cleavage activities. In particular, increases in entropy associated with the burial of the hydrophobic isoleucine side-chains at the DNA-protein interface may foster DNA binding at non-cognate sites, thereby resulting in relaxed DNA cleavage specificity. We tested this possibility by assaying the DNA cleavage activities of the wild-type and mutant proteins using DNA substrates containing point mutations that reduce or eliminate cleavage by the wild-type protein²². Two positions that were tested are distant from Lys-122 and Lys-223 and include the G/C₋₄ base-pair, which makes base-specific contacts to Asn-192, and the T/A₋₇ base-pair, which makes a base specific contact to Asn-152¹¹. Also assayed was the C/G₊₃ base-pair, which is in close proximity to Lys-122, and makes base-specific contacts to Glu-61 and Arg-88. No nicked or linearized products were generated by wild-type I-SceI after incubation for one hour with pBS I-SceI (E/H) containing G₋₄A, or C₊₃T, mutations in the I-SceI recognition sequence, and low levels of linearized DNA plasmid were generated using the T₋₇A mutant target (Figure 4). These results are consistent with the predicted cleavage properties of these mutant substrates by I-SceI based on a comprehensive specificity profiling study²². The K122I and K223I proteins yield no cleavage products using the G₋₄A, or C₊₃T substrates and low levels of nicked and linearized DNA using the T₋₇A mutant substrate. The comparable levels of cleavage activity by the wild-type and mutant proteins suggests that their ability to discriminate among target sites containing mutations at several different positions is similar.

Creation of a hairpin substrate for thermodynamic and kinetics analyses

Plasmid DNA substrates similar to pBS I-SceI (E/H) have been used to determine the apparent and intrinsic rate constants for first and second strand DNA cleavage by restriction enzymes²³. The amounts of SC, OC and LIN forms of plasmid DNA cannot be accurately quantified from gels stained with ethidium bromide due to the fact that each DNA form binds different amounts of the dye, making it necessary to internally radiolabel the SC substrate and to measure DNA amounts by scintillation counting after extraction from the gel.

In order to readily identify the properties of the nicked reaction intermediates and to facilitate determination of reaction rate constants, we constructed a novel 94 bp hairpin substrate containing a single I-SceI recognition sequence (Figure 5). This substrate was assembled by ligating three ³²P-end-labeled synthetic oligonucleotides and by purifying the resulting hairpin DNA. This substrate is modeled after a hairpin substrate reported previously that was labeled at either the 5' - or 3' - end, and was used to demonstrate the existence of a nicked intermediate

in the I-SceI pathway¹⁶. The advantage of the substrate that is internally radiolabeled at two locations and 5'-end-labeled is that it permits the amounts of the substrate, all nicked intermediates and all reaction products to be determined simultaneously using denaturing gel electrophoresis and imaging analysis.

Binding of wild-type I-SceI and point mutants to intact, nicked and product DNA

The DNA binding affinities of wild-type I-SceI and the K122I and K223I mutant proteins for different DNA substrates were measured by a filter binding assay (see Materials and Methods) either in the presence or absence of divalent Ca^{2+} metal ions (Table 1). Substitution of Ca^{2+} for Mg^{2+} abolishes DNA cleavage activity by I-SceI¹¹. There is structural evidence that steric crowding of Ca^{2+} ions prevents their full occupancy at the active sites, leading to inactivation of the enzyme^{12; 15}. In the absence of divalent metal ion, the 14.5 nM dissociation constant (K_d) measured by filter binding for the interaction of I-SceI and the 94 base-pair hairpin duplex is higher than the 0.8 nM value measured by electrophoretic mobility shift assays for I-SceI and a 54 bp linear duplex²¹. This difference may be due to the different methods used to measure the dissociation constants, the different forms of DNA used, or the different methods used to measure protein concentrations. The dissociation constant is dependent on the concentration of Ca^{2+} since it decreases by >6-fold to 2.3 nM when Ca^{2+} is present. Comparable increases in binding were observed for I-CreI when Ca^{2+} was included in the binding reaction¹². The presence of divalent metal ions may increase DNA binding by partially neutralizing the electrostatic repulsion between the active site acidic residues and the DNA phosphate backbone.

I-SceI binding to hairpin DNAs that contain nicks in either the top or bottom DNA strands is similar to that of the intact hairpin (Table 1). These hairpin intermediates were assembled by annealing hairpin and linear oligonucleotides as described in Materials and Methods. I-SceI makes most of the same DNA contacts when it is bound to uncleaved DNA as when it is bound to duplexes that contain nicks in either the top or bottom DNA strands at the scissile bonds¹⁵. We also confirm the previously reported observation¹⁶ that I-SceI binds as tightly to the downstream DNA product as to the intact hairpin substrate (3.3 nM vs. 2.3 nM) but binds weakly to the upstream DNA product ($K_d=175$ nM). The extremely slow turnover of I-SceI in the DNA cleavage reactions may be the consequence of the large number of DNA contacts it makes to the tight binding product.

Isoleucine substitution at Lys-122 or Lys-223 has little effect on the binding of the proteins to the hairpin substrate either in the presence or absence of divalent metal ion (Table 1). Each of these lysines makes a water-mediated contact to the phosphate backbone in the crystal structure, and their loss may contribute to the slightly elevated equilibrium dissociation constants, but equilibrium dissociation values are within the levels of variation for this type of assay. Thus, substitution of the lysines with isoleucine does not result in dramatic changes in the binding affinity to the ground state.

Single turnover DNA cleavage of the hairpin substrate by wild-type I-SceI

Single turnover reactions of wild-type I-SceI using the hairpin substrate were performed under similar conditions as the experiments using the plasmid substrates. In all of the experiments described here, the single turnover reactions were started by adding MgCl_2 to the enzyme-DNA mixture. In the single turnover experiments, the amount of the intact hairpin substrate declined exponentially over time, giving rise to the transient appearance of two nicked DNA species whose amounts first increase then decrease during the course of the reaction (Figure 6A, 6B). A greater amount of the 73 bp intermediate with a nick in the bottom strand is generated than the 67 bp intermediate with a nick in the top strand. Furthermore, the 21 bp DNA linear product appears prior to the 27 bp product. These observations confirm the earlier

study that detected the bottom DNA strand intermediate using a hairpin substrate that was radiolabeled only at the 5'-end¹⁶. However, in this study, we use the internally labeled hairpin substrate to observe the intermediate nicked on the top DNA strand for the first time. Presumably, its low abundance prevented its detection previously. The 46 bp hairpin product and the 27 bp and 21 bp species, which comprise the linear cleavage product, begin to accumulate to detectable levels by 30 seconds (Figure 6A, 6B). We also assayed I-SceI-mediated cleavage of a 5'-endlabeled synthetic 94-mer hairpin substrate, and determined that the cleavage rate, as indicated by the rate of appearance of the 27 bp product, was similar to that of the assembled hairpin substrate (data not shown). Different levels of radio-labeling at each of the three phosphate positions in the hairpin substrate does not account for the differences in the rate of appearance of the different fragments because they are similarly labeled (< 8% difference, data not shown).

The observation that both nicked intermediates are produced by I-SceI is consistent with a parallel-sequential reaction pathway for double-strand cleavage rather than a pathway in which obligatory cleavage of one specific strand occurs prior to that of the second DNA strand. The single-turnover reaction can be depicted according to the following scheme:

(2)

where k_1 and k_4 are the rate constants for first and second-strand cleavage of the bottom strand; k_2 and k_3 are the rate constants for first and second-strand cleavage of the top strand; ES is the enzyme-bound substrate DNA; EN_{top} and EN_{bot} are the nicked top and bottom strand DNA intermediates; and EL_1 and L_2 are the two products. To obtain values for k_1 – k_4 , the experimental DNA cleavage data were fit to the rate equations for this scheme by numerical integration (see Materials and Methods). The computation failed to converge to a solution using the wild-type data, which is often the case when the rate constants have similar values²⁴. Therefore, values of k_3 and k_4 were independently obtained by assembling DNA species from synthetic oligonucleotides that correspond to the nicked intermediates, and by measuring the cleavage rate of the EN_{top} and EN_{bot} species in single-turnover experiments (data not shown). The experimental data were fit to a single exponential to obtain values of 0.091 s^{-1} and 0.053 s^{-1} , respectively, for top and bottom strand cleavage rate constants (Table 2). Values of k_1 and k_2 were obtained from the experimental data by fitting rate equations describing the formation of the hairpin product, the conversion of hairpin substrate, and the transient accumulation of nicked bottom DNA strand intermediate using fixed values of k_3 and k_4 (see Supplemental materials). Consistent with the larger accumulation of the nicked bottom DNA strand intermediate, k_1 is over 3-fold higher than k_2 (0.0062 s^{-1} vs. 0.0018 s^{-1}). The curve fit indicates that the nicked top DNA strand intermediate accumulates slower and to a lesser degree than the nicked bottom DNA strand intermediate, and its low amount was not quantitated. Interestingly, the rate of second-strand cleavage is higher than first-strand cleavage for both the top and bottom DNA strands ($k_1=0.0062\text{ s}^{-1}$ vs. $k_4=0.053\text{ s}^{-1}$ and $k_2=0.0018\text{ s}^{-1}$ vs.

$k_3=0.091\text{ s}^{-1}$). A low amount of I-SceI-independent radiolysis at the position of the labeled phosphates (~10%) may account for the deviation of the curve fits from the experimental data (see Materials and Methods). Alternatively, it cannot be ruled out that additional catalytic steps are part of the reaction pathway that are not accounted for by the model shown in Equation 2.

Single turnover DNA cleavage of the hairpin substrate by mutant I-SceI proteins

Single turnover reactions using the K122I and K223I mutant proteins were performed manually with an excess of enzyme relative to the hairpin DNA substrate (see Materials and Methods). An exponential decrease in the amount of the hairpin substrate was observed for both mutants (Figure 6A, 6C, 6D). Interestingly, this decrease is not mirrored by an increase in the amount of the hairpin product, as in the case of the wild-type I-SceI, but instead by large accumulations of the nicked top or bottom DNA strand intermediates for the K223I and K122I proteins, respectively. After the nicked intermediates peak in amount, their levels decline slowly, while the amount of the hairpin product DNA gradually increases. This kinetics profile is entirely consistent with rapid DNA cleavage of one DNA strand followed by very slow cleavage of the second strand. Moreover, the direct observation of the different nicked DNA intermediates allows us to conclude that lysine-122 is important for top DNA strand cleavage while lysine-223 is important for bottom DNA strand cleavage, and that neither residue is essential. This same conclusion is consistent with the results of the plasmid DNA cleavage experiments (Figure 3).

For both the K122I and K223I proteins, all four reaction rate constants were obtained by fitting the experimental data to rate equations describing the formation of the hairpin product, the disappearance of the hairpin substrate and the transient appearance of the nicked intermediate species (Table 2 and Supplemental materials). The k_3 and k_4 values obtained from cleavage reactions using pre-nicked substrates were similar to those obtained by fitting the data to the full model (data not shown). Comparison of the rate constants of the mutant proteins with those of wild-type I-SceI indicates that the value of k_3 for K122I is over 300-fold lower while the k_4 value of K223I is over 1000-fold lower. These markedly lower rate constant values can account for the large accumulation of the nicked top or bottom DNA strand intermediate species for the K223I and K122I proteins, respectively. By contrast, there is little or no decrease apparent in the K122I k_2 and K223I k_1 values. Thus, substitution of the lysines severely reduces second strand cleavage of one of the two DNA strands, but has less effect on first strand cleavage. The K122I k_1 value, which describes first-strand cleavage of the bottom strand, is nearly three-fold higher than that of wild-type I-SceI, while the K223I k_2 value is nearly five-fold higher. These higher values account for the more rapid conversion of the supercoiled plasmid DNA to nicked open circle intermediates by the mutant proteins (Figure 3).

Discussion

Structural asymmetry underlies DNA binding and catalytic asymmetry

A comparison of the X-ray crystal structures of several LAGLIDADG homing endonucleases reveals a variety of tertiary and/or quaternary structures of differing levels of two-fold symmetry. At one end of the symmetry spectrum are the enzymes such as I-CreI and I-MsoI that exist as homodimers^{25; 26}. The two identical subunits within these enzymes can bind to identical half-sites within palindromic recognition sequences, although their naturally occurring target sequences lack perfect symmetry. I-CeuI differs from I-CreI and I-MsoI since it is a homodimeric enzyme that prefers to cleave its asymmetric target, which is only 36% palindromic, rather than either of two palindromic recognition sequences that can be constructed from the two half-sites²⁷. Less symmetric are the monomeric homing endonucleases such as I-SceI and I-AniI whose two domains exhibit pseudo-two fold symmetry that recognize and cleave asymmetric recognition sequence^{28; 29}. The homing endonuclease

domains situated within inteins are also pseudo-symmetric and cleave asymmetric target sequences. One scenario for the evolution of homing endonucleases posits that ancestral homodimeric enzymes gave rise to monomeric variants through gene duplication and gene fusion. Once homing enzymes became monomeric and were no longer constrained to interact with perfectly symmetric DNA recognition sequences, they acquired greater genetic mobility as a result of their increased flexibility in recognizing alternative target sites³⁰.

Evolution of structural asymmetry in homing endonucleases would be expected to be accompanied by the evolution of asymmetric DNA binding and DNA cleavage. Interestingly, asymmetric DNA binding is observed even in some homodimeric enzymes. Examination of the I-CeuI crystal structure indicates that the basis for its preference for an asymmetric target is that the side chains of symmetrically related identical residues assume different rotameric torsion angles and vary their use of bridging water molecules in order to contact non-identical base pairs at symmetric positions in the substrate²⁷. This flexibility in allowing recognition of asymmetric targets may have facilitated the divergence of homing enzymes, ultimately leading to their acquiring new target site specificities²⁷. Once homing enzymes evolved to become monomers, significant divergence of the two domains would be expected to result in each acquiring a different DNA binding affinity. As shown in this report and elsewhere, I-SceI binds tightly to the downstream cleavage product following catalysis, but not to the upstream product¹⁶. The X-ray structure reveals that the basis for this tight binding is the larger number of DNA contacts made by the amino-terminal domain of the protein (38 contacts of 56) to this region of the substrate¹¹. Asymmetric binding to the recognition site is not restricted to I-SceI since it is also reported for the intein-encoded PI-TfuII which, like I-SceI, exhibits product inhibition due to tight binding to one of its DNA cleavage products³¹.

By using an internally radio-labeled hairpin substrate, we demonstrated that I-SceI cleaves either DNA strand first or second through a parallel-sequential reaction mechanism. The data did not fit well to a simplified model in which the rate constants for first and second strand cleavage of the top or bottom strand were made equal (i.e. where $k_1=k_4$ and $k_2=k_3$) (data not shown). The EcoRI restriction endonuclease is a symmetric homodimer that uses a parallel sequential reaction mechanism in which the four rate constants (k_1 - k_4) are similar³². I-SceI cleaves the DNA strands at different rates as a result of structural differences between the two active sites. We proposed that the faster rate of bottom strand cleavage is due to the asymmetric distribution of metal ions¹⁵. The structure of I-SceI bound to uncleaved DNA represents a complex that is poised to cleave the bottom DNA strand since it shows that only one active site, that which cleaves the bottom strand, contains a metal (Ca^{2+} 1) capable of activating a nucleophilic water for this cleavage²⁹ (Figure 7A). By contrast, in a crystal structure of I-SceI bound to a nicked substrate in which the bottom strand is already cleaved (Figure 7B), a second unshared metal (Metal 4) and a putative nucleophilic water (W60) are present that are correctly positioned to aid in top strand cleavage¹⁵. Thus, we speculate that top strand cleavage occurs more rapidly once the bottom strand has already been cut, as shown here, because of the presence of a suitable catalytic metal ion. It is possible that the generation of the free hydroxyl and phosphate groups following scission of one strand facilitates cleavage of the other for LAGLIDADG enzymes. The existence of the previously undetected intermediate with a nick in the top strand indicates that bottom strand cleavage is not an obligatory step in the reaction, but how first strand cleavage of the top strand occurs is unclear.

The observed occupancies of the metals in the different crystal structures may reflect different metal binding affinities for the active sites, which could lead to preferential cleavage of one DNA strand. Other instances of asymmetric DNA cleavage by homing enzymes reported for I-CeuI³³, I-CpaII³⁴, I-SceIII³⁵ and PI-TfuI³⁶ might be accounted for by different metal binding affinities of the two active sites. I-CpaII cleaves only one DNA strand at low Mg^{2+} concentrations, but at higher concentrations, it is able to cleave the second strand as well³⁴.

Similarly, PI-TfuI only cleaves one DNA strand when Mg^{2+} is used as a co-factor, but in the presence of Mn^{2+} , which is a “soft” metal that tolerates imprecise ligand geometry relative to Mg^{2+} ,³⁷ it cleaves both strands. As a result of its marked preference for cleaving one DNA strand, PI-TfuI, like I-SceI, generates nicked intermediates during the reaction, which indicates a sequential strand cleavage mechanism³¹. Kinetics studies using phosphorothioate analogues demonstrated that the enzyme cleaves either strand first, although top strand cleavage requires Mn^{2+} .³⁶

Mutation of peripheral amino acids partially inactivates homing enzyme active sites

Mutation of the pseudo-symmetric lysines leads to reciprocal effects on the cleavage reaction since the K122I derivative cleaves the top strand slower than wild-type I-SceI while the K223I variant cleaves the bottom strand slower. Hydrogen bonding of Lys-223 to a nucleophilic water molecule implicates a direct role for this residue in helping to order the solvent molecules surrounding one active site¹¹ (Figure 7A). Although Lys-122 is hydrogen bonded to solvent in the nicked bottom DNA structure (Figure 7B), it is not hydrogen bonded to W60, the putative nucleophile. The K122 side chain is retracted ~2.5 Å from the catalytic center relative to the structure with uncleaved DNA, perhaps due to crystal packing forces. Therefore, it cannot be ruled out that in the active complex, the side chain is actually closer to metal 4 and plays a more direct role in positioning W60. K98 and K98' in the homodimeric I-CreI, which are analogous to K122 and K223 in I-SceI, help coordinate a network of solvent molecules that bond to the nucleophilic waters¹². Mutation of these residues does not affect the cleavage rate, but dramatically reduces the enzyme affinity for the uncleaved substrate^{12; 18}. Mutation of the K122 and K223 analogues in PI-SceI, K301 and K403, affects the reaction differently since a K301A mutant is unable to cleave the DNA whereas a K403A exhibits reduced activity^{38; 39}. It has been suggested that the weak conservation of the lysines among LAGLIDADG enzymes and the different phenotypes displayed by the lysine mutants is consistent with their contribution to the electrostatic environment surrounding the active sites¹². It has been shown for EcoRV that the efficiency of phosphoryl transfer is strongly influenced by changes to this environment.³⁷

The degree of coupling between the active sites can be inferred by examination of the different rate constants. If the two active sites are independent, perturbation of one active site would not be expected to affect the function of the second active site. A K122I mutant exhibits a >350-fold reduction in k_3 relative to the wild-type enzyme, but more similar k_1 and k_4 values. Conversely, proteins containing a K223I mutation display reduced bottom strand cleavage (>1000-fold decrease in the k_4 value) but markedly smaller decreases in top strand cleavage (~20-fold decrease in k_3). Thus, although the active sites are in close proximity and share a metal ion, the levels of cross-communication are only moderate. The two active sites in PI-SceI are also partially interdependent because mutation of residues proximal to the catalytic center, including one of the lysine analogues, leads to reduced cleavage of both DNA strands, with a larger reduction occurring in the active site in which the residue resides⁴⁰. Communication between active sites is indicated for PI-TfuI as well because when top strand cleavage is inhibited by the presence of a phosphorothioate analogue, bottom strand cleavage requires that Mn^{2+} be used as a co-factor rather than Mg^{2+} .³⁶

Engineering nicking enzymes from homing endonucleases

There has been significant success in engineering restriction endonucleases that specifically nick one DNA strand or the other, but it may be difficult to extend the same strategies to homing enzymes. Monomeric variants of the AlwI enzyme that nick DNA were created from the dimeric enzyme by introducing mutations that prevented dimerization⁵. For homodimeric homing enzymes such as I-CreI, monomeric enzymes created by reducing dimerization will be completely inactive since a shared metal required for cleavage of both strands is coordinated

using side chains contributed by each subunit. One class of restriction enzyme forms a heterodimer in which each subunit specifically cleaves one DNA strand, and it has been possible to engineer strand-specific enzymes by inactivating one subunit or the other^{7; 23}. This technique has been successfully applied to the restriction endonuclease I-BbvI, which contains two non-identical subunits that each cleave one DNA strand within an asymmetric DNA sequence²³. Other restriction enzymes are monomers that contain two active sites that can be separately mutated to create nicking enzymes⁶. In the case of SapI, it is thought that some of these mutations alter acidic residues that coordinate essential metal ions at one active site⁴. Our results indicate that complete inactivation of one active site in I-SceI by mutation of the acidic residues that coordinate the metal ions is not possible since these side chains ligate a shared metal required by both sites. However, we were successful in engineering I-SceI derivatives that nick DNA by introducing lysine mutations at the periphery of either active site. The K223I derivative exhibited more nicking activity than the K122I mutant, and may prove useful in applications that require the rapid formation of nicked DNA, on the order of minutes, that have half-lives on the order of several hours. Further suppression of the residual double-strand cleavage activity is still required to obtain homing enzymes that exclusively nick DNA. We envision that engineering these enzymes may require the isolation of suitable nicking specific variants from randomized libraries using genetic selections⁶.

Materials and Methods

Materials

Synthetic oligonucleotides that were used for cloning of I-SceI recognition sites, for mutagenesis, for PCR, as DNA substrates and as size standards were purchased from Integrated DNA Technologies, Inc. All DNA polymerases, restriction and DNA-modifying enzymes were obtained from New England Biolabs, Inc. Cobalt metal affinity resin (TALON) was obtained from BD Biosciences Clontech. SP-sepharose, heparin agarose and PD-10 desalting columns were purchased from GE Healthcare. Nitrocellulose filters (BA-85) were obtained from Schleicher and Schull. All other chemicals were of reagent grade and were obtained from commercial sources.

Media and bacterial strains

Luria-Bertani (LB) and LB agar media were prepared as specified by the manufacturer (Difco). Superior broth was prepared as specified by the manufacturer (US Biological). The *E. coli* strain BL21 (DE3) ($F^- ompT gal[dcn] [lon] hsdS_B (r_B^-, m_B^-)$; an *E. coli* B strain with λ DE3, a prophage carrying the T7 RNA polymerase gene) was used as the I-SceI expression strain.

Plasmid construction and mutagenesis

Plasmid pBS I-SceI (E/H) was constructed by annealing two complementary oligonucleotides that contain the I-SceI recognition sequence and by ligating them to pBluescript II (Stratagene) digested with *EcoRI* and *HindIII*. Derivatives of pBS I-SceI (E/H) containing the T₋₇A, G₋₄A or C₊₃T mutations were generated by similar methods. Plasmid DNA was purified by density-gradient centrifugation.

Synthetic oligonucleotides complementary to the N- and C-termini of I-SceI were used to generate a PCR product that encodes the wild-type enzyme fused to a C-terminal his₆ tag by a serine linker. Plasmid pET15b-I-SceI (Ctermhis) was constructed by subcloning this I-SceI ORF between the *NdeI* and *BamHI* sites of pET15b I-SceI, which is a pET15b (Novagen) derivative that expresses native I-SceI. Overlapping PCR mutagenesis using oligonucleotides complementary to the N- and C-terminal regions of the I-SceI gene and two mutagenic oligonucleotides (K223I: 5'-GCAGAATGTACATCATCTGGGGGATGAGGTACGGTTT-3' and 5'-

GATCAAACCGTACCTCATCCCCAGATGATGTACATTCTGCCG-3'; K122I: 5'-CGTTAACAACAAAATTACCATCCCCG-3' and 5'-CGGGATGGTAATTTTGTGTTAACG-3') was used to generate PCR fragments encoding the K223I and K122I mutant alleles, which were subcloned into the *Nde*I and *Bam*HI sites of pET15b I-SceI. All constructs were verified using automated DNA sequencing.

Protein purification

The pET15b I-SceI (Ctermhis) derivatives encoding the wild-type and mutant I-SceI alleles were transformed into *E. coli* strain BL21 (DE3) for protein expression. Cells were grown in Superior broth (US Biological) at 37°C to mid-log phase and were induced with 0.25 mM IPTG after being cooled to 18°C. Following overnight growth at 18°C, the cells were harvested and lysed by sonication in buffer containing 100 mM Tris pH 8.0, 200 mM KCl, 5% glycerol, 1.4 mM β -mercaptoethanol (BME) and 1 mM phenylmethylsulfonyl fluoride. Insoluble material was removed by centrifugation and the soluble extract was incubated with TALON metal affinity resin in sonication buffer according to the manufacturer's instructions. Bound I-SceI was eluted using an imidazole buffer and the buffer of the eluted I-SceI was immediately changed to 10 mM HEPES pH 8.0, 100 mM (Ethylenedinitrilo)Tetraacetic acid (EDTA), 5% glycerol, 1.4 mM BME and 100 mM KCl using a PD-10 desalting column. The protein was further purified by chromatography using SP-sepharose, and dialyzed overnight in 20 mM KPO₄ (pH 7.6), 0.1 mM EDTA, 5% glycerol, 1.4 mM BME and 80 mM KCl. The eluted protein was applied to a heparin agarose column equilibrated in the same buffer and purified using a KCl gradient. Fractions containing purified I-SceI were pooled, dialyzed against 20 mM KPO₄ (pH 7.4), 0.2 mM EDTA, 2 mM dithiothreitol and 200 mM NaCl and diluted 1:1 with glycerol for storage at -20°C. Protein concentrations were determined by UV absorbance using an extinction coefficient (ϵ_{280})=5.30 $\times$ 10⁴/M/cm.

Construction of DNA substrates

A hairpin substrate radiolabeled at three positions was prepared for kinetics experiments by sequential ligation reactions. Each of three synthetic oligonucleotides, 5'-GGTGAAACACTCGAACGCTAGGGATAACAGGGTAATA-3' (oligonucleotide 1), 5'-TAGCCAATCCAATGATTGGCTATATTACCCTGTTATCCCT-3' (oligonucleotide 2) and 5'-AGCGTTCGAGTGTTC-3' (oligonucleotide 3) were individually 5'-end-labeled using T4 polynucleotide kinase and γ^{32} P-ATP for 30 minutes at 37°C according to the manufacturer's instructions. Unlabeled ATP (0.5 mM) was added for a 10 minute incubation at the same temperature and the reactions were then terminated by incubation at 65°C for 20 minutes. Aliquots of radiolabeled oligonucleotides 1 and 2 were combined at 65°C for 5 minutes, cooled slowly to 25°C and ligated 16 hrs at 16°C using T4 DNA ligase. An aliquot of radiolabeled oligonucleotide 3 was added to the ligation mixture, which was heated to 65°C and allowed to cool slowly to 25°C. The oligonucleotide was ligated by adding T4 DNA ligase and by incubating the DNA at 16°C for 2.5 hours. The resulting 94 bp hairpin substrate was subsequently purified by denaturing electrophoresis using a 10% polyacrylamide gel. After I-SceI-mediated digestion, comparison of the 94 bp hairpin substrate, the 73 bp nicked reaction intermediate and the linear 27 bp reaction product with synthetic oligonucleotide standards reveals that two extra base-pairs were added to the 5'-end of the 94-mer during the labeling and synthesis procedure by an unknown mechanism. Digestion of the 94 bp hairpin substrate with the restriction endonuclease BfaI, which has a target site (5'-CTAG-3') that overlaps the 5' border of the I-SceI recognition sequence, and resolution of the products on a denaturing polyacrylamide gel revealed that the top strand of the linear duplex product (5'-GGTGAAACACTCGAACGC-3') was two bases longer than expected while the bottom strand of the product (5'-TAGCGTTCGAGTGTTC-3) was the expected size. Thus, given the labeling scheme, it is most likely that the additional bases are located at the 5'-end of the hairpin DNA and were added during the ligation.

Synthetic oligonucleotides used as size standards or as substrates included a 94-mer that corresponds to the intact hairpin substrate (5'-GGTGAACACTCGAACGCTAGGGATAACAGGGTAATATAGCCAATCCAATGATTGGCTATATTACCCTGTTATCCCTAGCGTTCGAGTGTTC-3'), a 73-mer (termed the nicked bottom DNA strand hairpin, 5'-GGTGAACACTCGAACGCTAGGGATAACAGGGTAATATAGCCAATCCAATGATTGGCTATATTACCCTGTTATCCCTAGCGTTCGAGTGTTC-3') and a 67-mer (termed the nicked top strand hairpin, 5'-CAGGGTAATATAGCCAATCCAATGATTGGCTATATTACCCTGTTATCCCTAGCGTTCGAGTGTTC-3') that correspond to the hairpin products resulting from partial I-SceI cleavage of the hairpin substrate at the bottom or top cleavage sites, respectively, a 46-mer (termed the hairpin product, 5'-CAGGGTAATATAGCCAATCCAATGATTGGCTATATTACCCTGTTATCCCTAGCGTTCGAGTGTTC-3') that corresponds to the hairpin product resulting from complete digestion of the intact hairpin substrate, and a 27-mer (termed the top linear DNA, 5'-GGTGAACACTCGAACGCTAGGGATAA-3') and a 21-mer (termed the bottom linear DNA, 5'-CCCTAGCGTTCGAGTGTTC-3') that correspond to the top and bottom strands, respectively, comprising the duplex product that results from digestion of the intact hairpin. To prepare size standards or enzyme substrates, these oligonucleotides were radiolabeled at their 5'-termini using $\gamma^{32}\text{P}$ -ATP and T4 DNA kinase and further purified by electrophoresis on a 10% urea-polyacrylamide gel. A radiolabeled duplex product identical to the upstream cleavage product of the hairpin substrate was assembled for filter-binding assays by combining the radiolabeled top and bottom linear DNAs at 65°C for 20 minutes and by slowly cooling to 25°C. The duplex was further purified by electrophoresis on a 15% native polyacrylamide gel. A hairpin substrate in which the top I-SceI cleavage site has been nicked (the nicked top strand intermediate) was assembled for filter-binding experiments by combining 5'-radiolabeled top linear DNA with nicked top strand hairpin that had been phosphorylated with ATP and T4 DNA kinase. These were heated at 65°C for 5 minutes, allowed to cool slowly to 25°C, and the nicked species was purified by electrophoresis on a 12% native gel. For kinetics experiments, this substrate was prepared using the same oligonucleotides except that the hairpin oligonucleotide was 5'-radiolabeled instead. The hairpin substrate containing a nick at the bottom I-SceI cleavage site was assembled by combining phosphorylated bottom linear DNA with 5'-radiolabeled nicked bottom DNA strand hairpin.

Filter binding assays

Nitrocellulose filter binding assays were performed on a vacuum manifold (Hoeffer). The radiolabeled binding substrates included the 94-mer intact hairpin, the 46-mer hairpin cleavage product, the duplex cleavage product (comprised of the bottom and top linear DNAs), and the hairpin substrates containing nicks at either the top or bottom cleavage sites. Samples of wild-type or mutant I-SceI were diluted to different concentrations in binding buffer (10 mM Tris-Cl (pH 8.0), 10 mM NaCl, 2.5 mM DTT and 20 $\mu\text{g/ml}$ bovine serum albumin (BSA)) and were incubated in the same buffer with radio-labeled substrates for 10 minutes at room temperature (22°–24°C). For assays performed in the presence of Ca^{2+} , CaCl_2 (2 mM) was included in the binding buffer. The samples were passed through 25 mm BA85 nitrocellulose filters that had been presoaked in binding buffer and were washed twice with the same buffer. The filters were placed in a liquid scintillation vial containing 2.5 ml of liquid scintillation fluid. The binding data was fit to a single-site binding isotherm ²⁴ given by nonlinear regression analysis using Kaleidagraph software (Synergy Software, Inc.).

Kinetics analysis of cleavage rates

The kinetics of plasmid DNA cleavage were analyzed using pBSI-SceI (E/H) containing the wild-type I-SceI recognition site or pBSI-SceI (E/H) derivatives containing single point mutations within the site. Reaction mixtures included cleavage buffer (10 mM Tris-HCl (pH

8.8), 1 mM DTT, and 0.1 mg/ml BSA) pBS-I-SceI (E/H) DNA (2.5 nM), I-SceI (100 nM), and 15 mM MgCl₂. Reactions were initiated by the addition of MgCl₂, and were allowed to proceed for chosen time intervals at 25°C before they were terminated by the addition of EDTA (100 mM). Supercoiled, open-circle and linear DNAs were resolved by electrophoresis on a 0.8% agarose gel in Tris-borate EDTA buffer (TBE). The gels were stained with ethidium bromide and fluorescent images were obtained from a transilluminator using a Kodak EDAS 290 imager.

Single-turnover reactions of the hairpin substrates were carried out using the intact hairpin substrate that had been labeled at three positions, singly end-labeled intact hairpin substrate or the nicked intermediate hairpin substrates. Reaction mixtures (20 µl) included cleavage buffer, I-SceI (100 nM), and DNA substrate (~20,000 cpm, ~5 nM), and were initiated by the addition of MgCl₂ (15 mM). Reactions were incubated at 25°C for chosen intervals of time and were terminated by addition of EDTA (100 mM). The reaction products were resolved by electrophoresis on a 10% polyacrylamide gel containing 7 M urea. Dried gels were exposed to storage phosphor screens and quantified using a Molecular Dynamics phosphorimager and ImageQuant software (Molecular Dynamics).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Moon-Soo Kim, Rakesh Joshi, Carmen Moure and Florante Quioco for valuable discussions and comments. Particular thanks to Barbara Golden for suggesting the use of an internally labeled hairpin substrate for the kinetics analysis. This work was funded by a National Science Foundation grant (MCB-0321550) and a National Institutes of Health grant (GM 070553), and a Welch Foundation grant (BE-1452) to FSG.

References

1. Van Ness J, Van Ness LK, Galas DJ. Isothermal reactions for the amplification of oligonucleotides. *Proc Natl Acad Sci U S A* 2003;100:4504–4509. [PubMed: 12679520]
2. Chan SH, Zhu Z, Van Etten JL, Xu SY. Cloning of CviPII nicking and modification system from chlorella virus NYs-1 and application of Nt.CviPII in random DNA amplification. *Nucleic Acids Res* 2004;32:6187–6199. [PubMed: 15570069]
3. Wang H, Hays JB. Mismatch repair in human nuclear extracts. Quantitative analyses of excision of nicked circular mismatched DNA substrates, constructed by a new technique employing synthetic oligonucleotides. *J Biol Chem* 2002;277:26136–26142. [PubMed: 12006560]
4. Samuelson JC, Zhu Z, Xu SY. The isolation of strand-specific nicking endonucleases from a randomized SapI expression library. *Nucleic Acids Res* 2004;32:3661–3671. [PubMed: 15247348]
5. Xu Y, Lunnen KD, Kong H. Engineering a nicking endonuclease N.AlwI by domain swapping. *Proc Natl Acad Sci U S A* 2001;98:12990–12995. [PubMed: 11687651]
6. Zhu Z, Samuelson JC, Zhou J, Dore A, Xu SY. Engineering strand-specific DNA nicking enzymes from the type IIS restriction endonucleases BsaI, BsmBI, and BsmAI. *J Mol Biol* 2004;337:573–583. [PubMed: 15019778]
7. Heiter DF, Lunnen KD, Wilson GG. Site-specific DNA-nicking mutants of the heterodimeric restriction endonuclease R.BbvCI. *J Mol Biol* 2005;348:631–640. [PubMed: 15826660]
8. Morgan R, Calvet C, Demeter M, Agra R, Kong H. Characterization of the specific DNA nicking activity of restriction endonuclease N.BstNBI. *Biol Chem* 2000;381:1123–1125. [PubMed: 11154070]
9. Stoddard BL. Homing endonuclease structure and function. *Q Rev Biophys* 2006;38:49–95. [PubMed: 16336743]

10. Colleaux L, D'Auriol L, Betermier M, Cottarel G, Jacquier A, Galibert F, Dujon B. Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *E. coli* as a specific double strand endonuclease. *Cell* 1986;44:521–533. [PubMed: 3004738]
11. Moure CM, Gimble FS, Quioco FA. The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. *J Mol Biol* 2003;334:685–695. [PubMed: 14636596]
12. Chevalier B, Sussman D, Otis C, Noel AJ, Turmel M, Lemieux C, Stephens K, Monnat RJ Jr, Stoddard BL. Metal-dependent DNA cleavage mechanism of the I-CreI LAGLIDADG homing endonuclease. *Biochemistry* 2004;43:14015–14026. [PubMed: 15518550]
13. Gimble FS, Stephens BW. Substitutions in conserved dodecapeptide motifs that uncouple the DNA binding and DNA cleavage activities of PI-SceI endonuclease. *J Biol Chem* 1995;270:5849–5856. [PubMed: 7890714]
14. Schöttler S, Wende W, Pingoud V, Pingoud A. Identification of Asp218 and Asp326 as the principal Mg²⁺ binding ligands of the homing endonuclease PI-SceI. *Biochemistry* 2000;39:15895–15900. [PubMed: 11123916]
15. Moure CM, Gimble FS, Quioco FA. Crystal structures of I-SceI complexed to nicked DNA substrates: snapshots of intermediates along the DNA cleavage reaction pathway. *Nucleic Acids Res* 2008;36:3287–3296. [PubMed: 18424798]
16. Perrin A, Buckle M, Dujon B. Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions. *EMBO J* 1993;12:2939–2947. [PubMed: 8335007]
17. Lykke-Andersen J, Garrett RA, J K. Mapping metal ions at the catalytic centres of two intron-encoded endonucleases. *EMBO J* 1997;16:3272–3281. [PubMed: 9214642]
18. Seligman LM, Stephens KM, Savage JH, Monnat RJ Jr. Genetic analysis of the *Chlamydomonas reinhardtii* I-CreI mobile intron homing system in *Escherichia coli*. *Genetics* 1997;147:1653–1664. [PubMed: 9409828]
19. Turmel M, Otis C, Côté V, Lemieux C. Evolutionarily conserved and functionally important residues in the I-CeuI homing endonuclease. *Nucleic Acids Res* 1997;25:2610–2619. [PubMed: 9185572]
20. Komori K, Ichianagi K, Morikawa K, Ishino Y. PI-PfuI and PI-PfuII, intein-coded homing endonucleases from *Pyrococcus furiosus*. II. Characterization of the binding and cleavage abilities by site-directed mutagenesis. *Nucleic Acids Res* 1999;27:4175–4182. [PubMed: 10518608]
21. Beylot B, Spassky A. Chemical probing shows that the intron-encoded endonuclease I-SceI distorts DNA through binding in monomeric form to its homing site. *J Biol Chem* 2001;276:25243–25253. [PubMed: 11279183]
22. Doyon JB, Pattanayak V, Meyer CB, Liu DR. Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc* 2006;128:2477–2484. [PubMed: 16478204]
23. Bellamy SR, Milsom SE, Scott DJ, Daniels LE, Wilson GG, Halford SE. Cleavage of individual DNA strands by the different subunits of the heterodimeric restriction endonuclease BbvCI. *J Mol Biol* 2005;348:641–653. [PubMed: 15826661]
24. Connolly, BA.; Liu, H-H.; Parry, D.; Engler, LE.; Kurpiewski, MR.; Jen-Jacobson, L. Assay of restriction endonucleases using oligonucleotides. In: Moss, T., editor. *DNA-protein interactions: Principles and protocols*. Vol. 2nd edit.. Totowa, NJ: Humana Press, Inc.; 2001. p. 465-490.
25. Chevalier B, Turmel M, Lemieux C, Monnat RJ, Stoddard BL. Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J Mol Biol* 2003;329:253–269. [PubMed: 12758074]
26. Chevalier BS, Monnat RJ Jr, Stoddard BL. The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites. *Nat. Struct. Biol* 2001;8:312–316. [PubMed: 11276249]
27. Spiegel PC, Chevalier B, Sussman D, Turmel M, Lemieux C, Stoddard BL. The structure of I-CeuI homing endonuclease: evolving asymmetric DNA recognition from a symmetric protein scaffold. *Structure* 2006;14:869–880. [PubMed: 16698548]
28. Bolduc JM, Spiegel PC, Chatterjee P, Brady KL, Downing ME, Caprara MG, Waring RB, Stoddard BL. Structural and biochemical analyses of DNA and RNA binding by a bifunctional homing endonuclease and group I intron splicing factor. *Genes Dev* 2003;17:2875–2888. [PubMed: 14633971]

29. Moure CM, Gimble FS, Quioco FA. Crystal structure of the intein homing endonuclease PI-*SceI* bound to its recognition sequence. *Nat Struct Biol* 2002;9:764–770. [PubMed: 12219083]
30. Gimble FS. Invasion of a multitude of genetic niches by homing endonuclease genes. *FEMS Microbiol Lett* 2000;185:99–107. [PubMed: 10754232]
31. Saves I, Ozanne V, Dietrich J, Masson JM. Inteins of *Thermococcus fumicolans* DNA polymerase are endonucleases with distinct enzymatic behaviors. *J Biol Chem* 2000;275:2335–2341. [PubMed: 10644683]
32. Lesser DR, Kurpiewski MR, Jen-Jacobson L. The energetic basis of specificity in the *EcoRI* endonuclease-DNA interaction. *Science* 1990;250:776–785. [PubMed: 2237428]
33. Marshall P, Lemieux C. The I-*CeuI* endonuclease recognizes a sequence of 19 base pairs and preferentially cleaves the coding strand of the *Chlamydomonas moewusii* chloroplast large subunit rRNA gene. *Nucleic Acids Res* 1992;20:6401–6407. [PubMed: 1475201]
34. Turmel M, Mercier J-P, Cote V, Otis C, Lemieux C. The site-specific DNA endonuclease encoded by a group I intron in the *Chlamydomonas pallidostigmatica* chloroplast small subunit rRNA gene introduces a single-strand break at low concentrations of Mg^{2+} . *Nucleic Acids Res* 1995;23:2519–2525. [PubMed: 7630730]
35. Schapira M, Desdouets C, Jacq C, Perea J. I-*SceIII* an intron-encoded DNA endonuclease from yeast mitochondria. Asymmetrical DNA binding properties and cleavage reaction. *Nucleic Acids Res* 1993;21:3683–3689. [PubMed: 8367285]
36. Thion L, Laurine E, Erard M, Burlet-Schiltz O, Monsarrat B, Masson JM, Saves I. The two-step cleavage activity of PI-TfuI intein endonuclease demonstrated by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J Biol Chem* 2002;277:45442–45450. [PubMed: 12239210]
37. Horton NC, Otey C, Lusetti S, Sam MD, Kohn J, Martin AM, Ananthnarayan V, Perona JJ. Electrostatic contributions to site specific DNA cleavage by EcoRV endonuclease. *Biochemistry* 2002;41:10754–10763. [PubMed: 12196013]
38. He Z, Crist M, Yen H-C, Duan X, Quioco FA, Gimble FS. Amino acid residues in both the protein splicing and endonuclease domains of the PI-*SceI* intein mediate DNA binding. *J Biol Chem* 1998;273:4607–4615. [PubMed: 9468518]
39. Gimble FS, Hu D, Duan X, Quioco FA. Identification of Lys403 in the PI-*SceI* homing endonuclease as part of a symmetric catalytic center. *J Biol Chem* 1998;273:30524–30529. [PubMed: 9804821]
40. Christ F, Schoettler S, Wende W, Steuer S, Pingoud A, Pingoud V. The monomeric homing endonuclease PI-*SceI* has two catalytic centres for cleavage of the two strands of its DNA substrate. *Embo J* 1999;18:6908–6916. [PubMed: 10601013]
41. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera-Visualization system for exploratory research and analysis. *J. Comput. Chem* 2004;25:1605–1612. [PubMed: 15264254]

Figure 1. The 2.25 Å X-ray crystal structure of I-SceI bound to a 24 base-pair uncleaved duplex containing the 18 bp recognition sequence (1R7M11)
 (A) Ribbon representations depicting the secondary structure elements of the amino-terminal (colored cyan) and carboxyl-terminal (colored violet) domains of I-SceI complexed to DNA where the regions corresponding to the upstream and downstream products are colored green and blue, respectively. The positions of the phosphates at the scissile bonds are colored orange on each DNA strand. Side chains of the conserved acidic residues Asp-44 and Asp-145 and those of Lys-122 and Lys-223 are indicated. (B) As in (a) but rotated 90° about the y-axis. The figure was generated using Chimera⁴¹.

Figure 2.

Superposed LAGLIDADG-helices and neighboring lysine residues of I-SceI (residues D44, D145, K122, K223 colored green; 1R7M¹¹), I-CreI (residues D20, D20', K98, K98' colored red; 1G9Z²⁶), I-CeuI (residues E66, E66', K153, K153' colored yellow; 2EX5²⁷), I-AniI (residues D16, E148, K94 and K227 colored purple; 1P8K²⁸), I-MsoI (residues D22, D22', K104, K104' colored orange; 1M5X²⁵), and PI-SceI (residues D218, D326, K301, K403 colored cyan; 1LWS²⁹). Protein backbone residues are depicted in ribbon form while the side chains of the conserved acidic and lysine residues are shown as sticks. The figure was generated using PyMol.

Figure 3. Time course of cleavage of a plasmid substrate containing the I-SceI recognition sequence by wild-type I-SceI and the K122I and K223I variants

(A) DNA cleavage reactions were initiated by adding MgCl_2 (15 mM) to mixtures of plasmid pBSI-SceI (E/H) (2.5 nM) and purified I-SceI (100 nM) and were allowed to incubate for chosen lengths of time at 25°C before being terminated by the addition of EDTA (100 mM). Samples were subjected to electrophoresis on a 0.8% gel. A small amount of the OC form of the DNA is present at the zero time point because OC DNA is not entirely removed by equilibrium density centrifugation of the SC DNA. Arrows indicate the positions of open-circle (OC); linear (LIN), and supercoiled (SC) DNA. (B) Cleavage results for the K122I and K223I derivatives. The reaction was sampled at different time points than in the experiment shown in (a).

(a)

(b)

Figure 4. DNA cleavage of plasmid substrates containing wild-type or mutant I-SceI recognition sequences by wild-type I-SceI and the K122I and K223I variants

(A) Supercoiled plasmid substrates (2.5 nM) containing the wild-type, T₋₇A G₋₄A, or C₊₃T mutations were incubated with purified wild-type, K122I or K223I I-SceI (100 nM) for one hour at 25°C, and the reactions were terminated by the addition of EDTA (100 mM). The DNA substrates and products were resolved by electrophoresis on a 0.8% gel. Arrows indicate the positions of open-circle (OC); linear (LIN), and supercoiled (SC) DNA. NP indicates that no protein was added to the reaction mixture. (B) A diagram depicting the I-SceI recognition sequence. The base-pairs that were substituted by mutation are colored orange. The I-SceI cleavage sites are indicated by arrows on each DNA strand and the regions of the sequence corresponding to the upstream and downstream cleavage products are colored green and blue, respectively.

no cleavage	94
complete cleavage	46 — 27 — 21
top-strand cleavage	67 — 27
bottom-strand cleavage	73 — 21

Figure 5. Structure of an internally radiolabeled hairpin DNA substrate

The top diagram depicts the predicted secondary structure and sequence of the hairpin substrate. The positions of the radiolabeled phosphates are indicated by asterisks. The 18 bp I-SceI recognition sequence is enclosed in the outlined box, and the scissile phosphodiester bonds are indicated by arrows. The 46 bp hairpin product is colored green, and the top and bottom linear DNAs, which comprise the second cleavage product, are colored red and blue, respectively. Below the diagram is shown a table that depicts the different DNA species that are resolved on a denaturing gel that result from I-SceI-mediated cleavage. The DNAs are color-coded similarly to the DNA diagram.

(a)

Figure 6. Time course of single turnover cleavage reactions of DNA hairpin substrate with wild-type and mutant I-SceI proteins

(A) Wild-type, K122I, and K223I I-SceI proteins (100 nM) were combined with radiolabeled intact hairpin substrate (~5 nM), and reactions were initiated by addition of MgCl_2 (15 mM). Aliquots were removed at different times, EDTA (100 mM) was added to terminate the reactions, and the samples were analyzed by electrophoresis on a 10% denaturing polyacrylamide gel. The different substrate and product species are shown to the left. The open and closed arrows indicate the positions of the nicked bottom DNA strand hairpin and the nicked top strand hairpin products, respectively, generated by wild-type I-SceI. Low levels of radiolysis products are apparent at the zero time point which give rise to product species during

the course of the reaction (see Materials and Methods). The intact hairpin substrate, the nicked bottom DNA strand hairpin, and the top linear DNA migrate as species that are two base-pairs longer than their expected sizes due to modification at the 5'-end of the DNA. This modification has no effect on the cleavage activity of the molecule (see Materials and Methods). (B) Graphical representation of wild-type I-SceI single turnover reactions. Denaturing gels containing resolved reaction products were dried and scanned using a phosphorimager to determine the percentage of intact hairpin substrate, open squares; nicked bottom DNA strand hairpin, filled circles; nicked top strand hairpin, filled diamonds; and hairpin product, open triangles. Each data point is the mean from three separate experiments and the bars represent the standard errors. The lines indicate the best fit of the data to the model shown in equation (2) using rate constant values taken from Table 2. These values were calculated from the substrate, hairpin product, and reaction intermediate data (see Materials and Methods) using the rate equations given in Supplemental Materials. The inset diagrams show the reaction data at early time points. (C) As in panel B using K122I protein. (D) As in panel B using K223I protein.

Figure 7. Location of Lys-122 and Lys-223 in X-ray crystal structures of wild-type I-SceI bound to uncleaved (1R7M)¹¹ and nicked bottom DNA strand substrates (3COX)¹⁵

(A) I-SceI uncleaved DNA complex. The putative nucleophilic water molecule W25 that is hydrogen bonded to Lys-223 is labeled. (B) I-SceI nicked bottom DNA strand complex. The putative nucleophilic water molecule W60 involved in top strand cleavage is labeled. Ca²⁺ ions are depicted as purple spheres; the Na⁺ as a yellow sphere, and solvent molecules as cyan spheres. The top DNA strand is colored yellow, the bottom DNA strand is colored pink, and the scissile phosphates are colored gray.

Table 1

Thermodynamic Parameters

I-SceI	DNA	K_D (nM)	
		$-Ca^{2+}$ ^a	$+Ca^{2+}$ (2 mM)
WT	Hairpin substrate	14.5 ± 5.7	2.3 ± 0.4
	Duplex linear product	ND	175 ± 48
	Hairpin product	ND	3.3 ± 0.6
	Nicked top DNA strand substrate	19.4 ± 0.7	4.2 ± 1.1
	Nicked bottom DNA strand substrate	13.2 ± 2.1	3.2 ± 0.2
K122I	Hairpin substrate	17.3 ± 0.4	7.3 ± 0.4
K223I	Hairpin substrate	15.4 ± 0.3	7.4 ± 1.3

^aEquilibrium association constants (mean $\pm$ SD; n=3 to 5) were obtained by direct measurement by a filter binding assay. ND, not determined.

Table 2

Reaction rate constants

I-SceI	$k_1 \text{ (s}^{-1}\text{)}^a$	$k_2 \text{ (s}^{-1}\text{)}$	$k_3 \text{ (s}^{-1}\text{)}$	$k_4 \text{ (s}^{-1}\text{)}$
Wild-type	0.0062±0.0021	0.0018±0.0006	0.091±0.009	0.053±0.005
K122I	0.018±0.002	0.0024±0.0006	0.00025±0.00002	0.016±0.009
K223I	0.0018±0.0004	0.0090±0.0020	0.0046±0.0013	0.00004±7E-06

^aReaction rate constants determined for a parallel-sequential reaction pathway (see Equation 2). These were obtained by fitting the experimental data describing the formation of the hairpin product, the disappearance of the substrate and the transient formation of reaction intermediates to their respective rate expressions (see Supplemental materials). Values are the mean of three experiments (± SD).