J|A|C|S
ARTICLES
Published on Web 01/31/2006

Directed Evolution and Substrate Specificity Profile of Homing Endonuclease I-SceI

Jeffrey B. Doyon,[‡] Vikram Pattanayak,[‡] Carissa B. Meyer,[‡] and David R. Liu*[,†,‡]

Contribution from the Howard Hughes Medical Institute and the Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Received November 11, 2005; E-mail: drliu@fas.harvard.edu

Abstract: The laboratory evolution of enzymes with tailor-made DNA cleavage specificities would represent new tools for manipulating genomes and may enhance our understanding of sequence-specific DNA recognition by nucleases. Below we describe the development and successful application of an efficient in vivo positive and negative selection system that applies evolutionary pressure either to favor the cleavage of a desired target sequence or to disfavor the cleavage of nontarget sequences. We also applied a previously described in vitro selection method to reveal the comprehensive substrate specificity profile of the wild-type I-SceI homing endonuclease. Together these tools were used to successfully evolve mutant I-SceI homing endonucleases with altered DNA cleavage specificities. The most highly evolved enzyme cleaves the target mutant DNA sequence with a selectivity that is comparable to wild-type I-SceI's preference for its cognate substrate.

Introduction

Homing endonucleases recognize and cleave long DNA sequences (14−40 bp) making them ideally suited for the site-specific manipulation of genomic DNA.[1] These enzymes have recently been used for genome engineering in diverse systems including bacterial,[2] yeast,[3] plant,[4] and mammalian cells.[5] Because the sequences recognized by known homing endonucleases represent only a minute fraction of all possible 14−40 bp sequences, the development and application of general methods to create homing endonuclease variants with tailor-made DNA cleavage specificities would significantly enhance our ability to manipulate genomes. In addition, determining the ways in which a homing endonuclease can be altered to cleave a novel DNA substrate may reveal the importance of specific contacts and provide information about the ability of the nuclease−DNA interface to be evolved.

Toward these goals, researchers have attempted to generate homing endonucleases with novel DNA specificity using methods that enable both large[6−8] and small[9−11] changes in the recognition sequence. Stoddard and co-workers created a chimeric enzyme by fusing two monomers from different homodimeric homing endonucleases and remodeling the interface between the heterologous domains.[6] The resulting chimeric homing endonuclease cleaved a novel DNA sequence comprising the juxtaposed half-sites recognized by the original enzymes. The scope of this approach is limited to DNA sequences constructed from substrates of natural homing endonucleases. Seligman and co-workers recently used a screen that linked the cleavage of a recognition sequence by a homing endonuclease to the loss of kanamycin resistance or lacZ activity.[9,11] Screening for variants of I-CreI that could cleave a new sequence differing at one position in the half-site of the recognition sequence resulted in the isolation of enzymes with altered in vivo and in vitro specificities.

The direct in vivo selection for homing endonuclease activity is an attractive strategy for generating nucleases with altered substrate preferences. Ideally, such an approach would enable both the positive selection for cleavage of a desired target DNA sequence as well as the negative selection against cleavage of the wild-type (or any nontarget) DNA substrate. Previously we reported the first steps toward the development of a positive selection system for homing endonuclease activity by placing substrate sequences within a plasmid expressing a toxic protein.[12] Although endonuclease-catalyzed DNA cleavage in this first-generation selection system resulted in cell viability changes sufficient to distinguish nucleases with varying activi-

[†] Howard Hughes Medical Institute.
[‡] Department of Chemistry and Chemical Biology.

(1) Chevalier, B. S.; Stoddard, B. L. *Nucleic Acids Res.* **2001**, *29*, (18), 3757−3774.
(2) Kang, Y. S.; Durfee, T.; Glasner, J. D.; Qiu, Y.; Frisch, D.; Winterberg, K. M.; Blattner, F. *J. Bacteriol.* **2004**, *186*, (15), 4921−4930.
(3) Storici, F.; Durham, C. L.; Gordenin, D. A.; Resnick, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, (25), 14994−14999.
(4) Tzfira, T.; Frankman, L. R.; Vaidya, M.; Citovsky, V. *Plant Physiol.* **2003**, *133*, (3), 1011−1023.
(5) Miller, D. G.; Petek, L. M.; Russell, D. W. *Mol. Cell. Biol.* **2003**, *23*, (10), 3550−3557.
(6) Chevalier, B. S.; Kortemme, T.; Chadsey, M. S.; Baker, D.; Monnat, R. J.; Stoddard, B. L. *Mol. Cell* **2002**, *10*, (4), 895−905.
(7) Steuer, S.; Pingoud, V.; Pingoud, A.; Wende, W. *ChemBioChem* **2004**, *5*, (2), 206−213.
(8) Epinat, J. C.; Arnould, S.; Chames, P.; Rochaix, P.; Desfontaines, D.; Puzin, C.; Patin, A.; Zanghellini, A.; Paques, F.; Lacroix, E. *Nucleic Acids Res.* **2003**, *31*, (11), 2952−2962.
(9) Seligman, L. M.; Chevalier, B. S.; Chadsey, M. S.; Edwards, S. T.; Savage, J. H.; Veillet, A. L. *Nucleic Acids Res.* **2002**, *30*, (17), 3870−3879.
(10) Gimble, F. S.; Moure, C. M.; Posey, K. L. *J. Mol. Biol.* **2003**, *334*, (5), 993−1008.
(11) Sussman, D.; Chadsey, M.; Fauce, S.; Engel, A.; Bruett, A.; Monnat, R.; Stoddard, B. L.; Seligman, L. M. *J. Mol. Biol.* **2004**, *342*, (1), 31−41.
(12) Gruen, M.; Chang, K.; Serbanescu, I.; Liu, D. R. *Nucleic Acids Res.* **2002**, *30*, (7), e29.

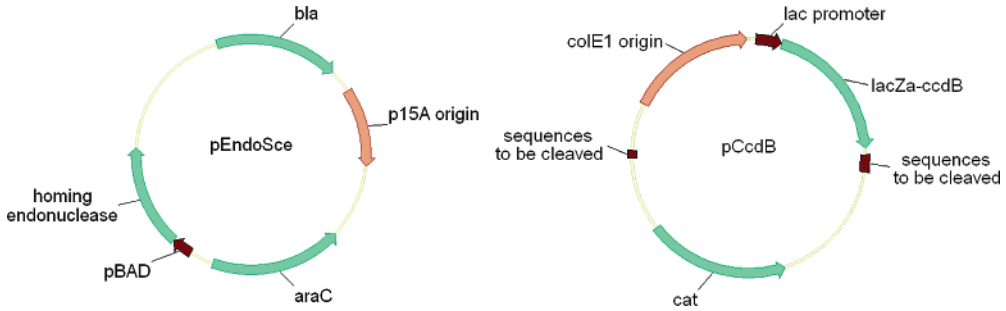

Figure 1. Two-plasmid in vivo selection system for homing endonuclease activity. pEndoSce is a low-copy plasmid that expresses I-SceI or a library of I-SceI variants under control of the *araBAD* promoter. pCcdB is a high-copy plasmid containing multiple copies of the desired or undesired recognition sequence distributed between two locations. This plasmid also expresses the toxic protein CcdB under control of the *lac* promoter. In the presence of IPTG, nucleases capable of cleaving pCcdB enable cells to survive (positive selection). pCcdB also expresses chloramphenicol acetyltransferase (Cat). In the presence of chloramphenicol and in the absence of IPTG, nucleases capable of cleaving pCcdB induce cell death (negative selection).

ties, the background survival rate from this system was too high to enable its use in directed evolution experiments.[12]

Here we report the development and application of an effective two-plasmid selection system that selects directly for homing endonuclease cleavage of target DNA substrates, or against the cleavage of undesired substrates, depending on the identity of a small-molecule inducer added to the growth media. We also used a selection-based method[13] to determine the comprehensive in vitro substrate specificity of I-SceI. Together, these selection systems were used to evolve an I-SceI variant that prefers to cleave a novel DNA sequence over the wild-type I-SceI substrate with a degree of specificity comparable to that of the wild-type I-SceI enzyme. Our findings validate an approach for the directed evolution of homing endonucleases with altered DNA cleavage specificity and also implicate specific residues involved in the recognition of the −7 position of the substrate DNA by I-SceI.

Results

Development of Positive and Negative In Vivo Selections for Endonuclease Activity. We previously reported a selection system for homing endonuclease activity that used the destruction of a plasmid expressing the toxic protein barnase to link DNA cleavage to cell survival.[12] If a cell expresses a homing endonuclease capable of cleaving the toxic plasmid, the resulting linearized DNA is presumably destroyed by the endogenous RecBCD nuclease, enabling cell survival. Due to the extreme toxicity of barnase,[14] its expression required additional control by the insertion of nonsense codons and the coexpression of suppressor tRNAs. The complexity of this system resulted in a high background survival rate that precluded its use for the efficient evolution of novel homing endonuclease specificity.

To simplify the positive selection system, we replaced barnase with the less toxic "control of cell death B" protein (CcdB), an *Escherichia coli* gyrase poison.[15] When expressed from a high-copy plasmid (pCcdB, Figure 1), this protein allowed only low rates of background survival without requiring additional gene expression control elements (see Experimental Section, below, for details). The homing endonuclease I-SceI was expressed on a low-copy plasmid (pEndoSce) using the araBAD promoter (Figure 1).

To test the efficiency of this selection system, pEndoSce was transformed into competent cells harboring pCcdB-wt1, a CcdB expression plasmid containing one copy of the wild-type I-SceI recognition sequence. A significant fraction (10−20%) of the resulting transformants survived upon IPTG induction of CcdB expression in rich media at 37 °C, consistent with the cleavage and destruction of pCcdB-wt1. In contrast, when the inactive I-SceI Asp 44 Ser mutant was used instead of wild-type I-SceI, only 1 in 1×10^3 colonies survived.

To further reduce the background survival rate, the positive selection was evaluated in different growth media and at different temperatures. In minimal media at 30 °C, wild-type pEndoSce rescued cells harboring pCcdB-wt1 at a 20−40% survival rate upon IPTG induction of CcdB expression. The inactive I-SceI Asp 44 Ser mutant under these selection conditions exhibited a greatly reduced survival rate of <1 in 5×10^4. These results suggest that the positive selection system can distinguish active and inactive homing endonucleases with an efficiency that may enable the directed evolution of endonucleases with the ability to cleave novel DNA substrates.

Many homing endonucleases are known to be product inhibited.[16] In our selection system, product inhibition could allow CcdB expression even after the plasmid is cleaved by an active endonuclease since RecBCD nuclease-mediated destruction of linear DNA requires an accessible 3′ DNA terminus.[17] Two additional copies of the I-SceI recognition sequence were therefore installed into pCcdB at a new location, along with three more copies adjacent to the original sequence, resulting in pCcdB-wt6 (Figure 1). We reasoned that the presence of distal cleavage sites within the toxic plasmid would increase the likelihood of survival in the presence of an active endonuclease. These changes increased the survival rate of cells harboring pCcdB-wt6 when transformed with pEndoSce to ∼100% without increasing the rate of undesirable background survival (Table 1). Independently, a similar positive selection system was recently reported by Chen and Zhao[18] that exhibits properties consistent with the above results.

Since homing endonucleases with broadened specificity cannot be distinguished from those nucleases with altered specificity solely on the basis of a positive selection, we extended the above two-plasmid system to also enable a negative

(13) Argast, G. M.; Stephens, K. M.; Emond, M. J.; Monnat, R. J. *J. Mol. Biol.* **1998**, *280*, (3), 345−353.
(14) Liu, D. R.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, (9), 4780−4785.
(15) Critchlow, S. E.; Odea, M. H.; Howells, A. J.; Couturier, M.; Gellert, M.; Maxwell, A. *J. Mol. Biol.* **1997**, *273*, (4), 826−839.
(16) Guhan, N.; Muniyappa, K. *Crit. Rev. Biochem. Mol. Biol.* **2003**, *38*, (3), 199−248.
(17) Yu, M.; Souaya, J.; Julin, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, (3), 981−986.
(18) Chen, Z.; Zhao, H. *Nucleic Acids Res.* **2005**, *33*, (18), e154.

Table 1. In Vivo Activity of Wild-Type I-SceI and Evolved Variants

enzyme	wild-type site[a]	T(−7)C site[a]
wild-type I-SceI	100%	1%
evolved mutant 1-6 (Asn 152 Lys)	84%	50%
evolved mutant 2-4 (Trp 149 Gly, Asp 150 Cys, Asn 152 Lys)	2%	100%

[a] Values indicate the percent survival in the positive selection using pCcdB-wt6 and pCcdB-mut6 to test for cleavage of the wild-type and T(−7)C sites, respectively. The data represent one trial to maximize consistency between clones but are representative of the trends observed with repeated measurement.

selection against enzymes that cleave undesired (e.g., wild-type) recognition sequences. To achieve this goal using the same two plasmids, undesired cleavage sites were introduced into pCcdB in place of the desired sites used in the positive selection. In the absence of IPTG, pCcdB does not express the toxic protein CcdB but continues to express chloramphenicol acetyltransferase (Cat), a protein that confers resistance to the antibiotic chloramphenicol (Figure 1). When cells are treated with chloramphenicol but not exposed to IPTG, endonuclease-mediated cleavage of the undesired sites destroys Cat expression and therefore renders cells unable to survive in the presence of chloramphenicol.

The negative selection was also validated using I-SceI. When a plasmid encoding wild-type I-SceI was transformed into bacteria containing pCcdB-wt6, the resulting bacteria survived at a very low rate (<1 in 7500) in the presence of chloramphenicol and in the absence of IPTG. However, transformation of a plasmid encoding the inactive I-SceI Asp 44 Ser mutant into the same cells resulted in 100% survival. Taken together, the above systems represent both a positive selection for endonuclease cleavage of a desired DNA sequence and a negative selection against cleavage of an undesired sequence.

Comprehensive Profiling of I-SceI Substrate Specificity In Vitro. Before applying our selection system to the evolution of I-SceI variants with novel cleavage specificities, we determined the specificity of the wild-type endonuclease in a comprehensive manner. We used the method of Monnat and co-workers[13] in which a partially randomized library of substrate sequences is incorporated into a plasmid and incubated in vitro with the endonuclease. Cleaved plasmids are separated from uncleaved plasmids by electrophoresis, circularized by ligation, and then transformed into *E. coli* to create a new pool of substrates for a subsequent round of selection or for analysis by DNA sequencing.

We created a large library of I-SceI substrate variants (in plasmid pCcdB-lib) consisting of 5.5×10^6 transformants. The library was designed to have a 72% probability of containing a wild-type nucleotide (Figure 2) at each of the 18 substrate positions, corresponding to a theoretical average of five mutations per recognition sequence. Sequence analysis of 287 clones from the library before selection revealed an actual probability of 74% ± 4% of any substrate position containing a wild-type nucleotide and an average of 4.7 mutations per substrate in the pre-digested library.

We subjected pCcdB-lib to three rounds of selection for sequences that can be cleaved by wild-type I-SceI. Sequence analysis of 597 members of the resulting post-selection library and comparison with the composition of the pre-selection library revealed a comprehensive profile of the substrate sequences that can be cleaved by wild-type I-SceI (Figure 2). As expected, the wild-type substrate sequence was the most common sequence to emerge from the selected library (137 of 597 sequenced clones, or 23%). Cleavable substrates in the selected library contained an average of 1.1 mutations compared to 4.7 in the pre-selection library. Half of the selected library contained a substrate sequence with one mutation (304/597, 51%), 20% contained two mutations (120/597), 5% contained three mutations (31/597), 0.7% contained four mutations (4/597), and 0.2% contained five mutations (1/597).

Based on the in vitro sequence profiling of I-SceI (Figure 2), thymine is strongly preferred at position −7 in the recognition sequence where it represents 85% of the cleavable sequences (Figure 2). Of 597 selected substrate sequences, A, G, and C appear at position −7 at normalized frequencies of only 7%, 6%, and 3%, respectively. Based on the recently reported structure of I-SceI,[19] only a single contact involving the side chain of Asn 152 is made between the enzyme and the base pair at position −7 of the substrate, making this interaction an ideal starting point to investigate the protein−DNA interface of I-SceI using a focused library of I-SceI mutants.

Evolution of I-SceI Variants with Altered DNA Cleavage Specificities. The specificity profile revealed that C was the least likely base to be present in a cleavable substrate at position −7 (Figure 2). To verify in vivo the importance of this position, wild-type I-SceI was expressed in combination with pCcdB-mut6 (containing six copies of the mutant T(−7)C recognition site), resulting in a low survival rate of 1% (Table 1). Based on the above in vitro and in vivo results, we chose to use our positive and negative selection system to evolve I-SceI variants

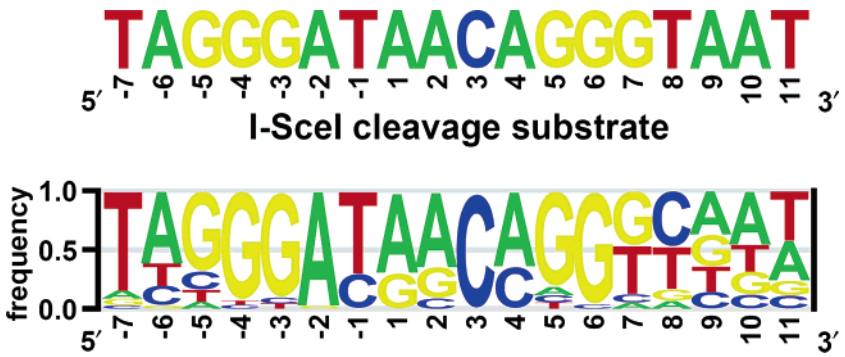

Figure 2. Sequence preference of wild-type I-SceI. Top: 18-base pair wild-type I-SceI substrate sequence is shown. Substrate cleavage occurs between A(2) and C(3) on the sense strand and between C(−3) and T(−2) on the antisense strand. Bottom: The DNA cleavage specificity of wild-type I-SceI is represented as a sequence logo. For each position the height of a base indicates the frequency of that base among the set of all sequences that are cleaved by the enzyme. Frequencies were normalized to account for the statistical bias favoring the wild-type cleavage sequence in the preselection library.

that cleave the mutant T(−7)C recognition sequence efficiently and preferentially over the wild-type substrate.

A small round 1 library of I-SceI variants was created by randomizing position 152 and selected as described above for the ability to cleave the T(−7)C recognition sequence in pCcdB-mut3. Clone 1-6 (Asn 152 Lys) exhibited robust survival when challenged with pCcdB-mut3, although this clone also exhibited significant in vivo activity when challenged with pCcdB-wt6 (Table 1). These results therefore identified clone 1-6 (Asn 152 Lys) as a mutant with broadened specificity that represents a starting point for further rounds of evolution.

Clone 1-6 was diversified into a round 2 library in which neighboring positions 149, 150, and 151 were each randomly mutated to one of 20 amino acids and position 153 was mutated to one of 13 amino acids. The resulting 4.2×10^6-membered library (7×10^6 transformants) was selected for cleavage of the mutant T(−7)C recognition sequence in pCcdB-mut3, yielding 230 surviving clones (corresponding to a survival rate of 1 in 3×10^4). To disfavor mutants with broadened, rather than altered, DNA cleavage specificities, these clones were subjected *en masse* to a negative selection against cleavage of the wild-type I-SceI substrate as described above. Surviving clones were subjected to a final positive selection (without additional diversification) for cleavage of the T(−7)C sequence. Clone 2-4 (Trp 149 Gly, Asp 150 Cys, Asn 152 Lys) was able to survive both the positive and negative selections upon rescreening (Table 1). When assayed in vivo using our selection system, clone 2-4 exhibited a 50-fold higher survival rate when cleaving the T(−7)C substrate compared with the wild-type substrate (Table 1). A third round of positive and negative selection did not yield further improvements in cleavage specificity.

To dissect the role of each of the two new mutations acquired in round 2, the mutations at positions 149 and 150 were individually reverted, and the resulting clones were assayed for activity in vivo against both the wild-type and T(−7)C DNA sequences. Reversion of the Trp 149 Gly mutation resulted in activity and broad specificity similar to those of clone 1-6, while reversion of the Asp 150 Cys mutation resulted in no activity toward either substrate (data not shown). These results indicate that both mutations acquired in round 2 are necessary for the preference of clone 2-4 for the T(−7)C substrate.

In Vitro Characterization of Evolved Homing Endonucleases. To characterize the activities of the evolved I-SceI variants in vitro, we purified evolved endonucleases 1-6 and 2-4 as well as the wild-type I-SceI enzyme using immobilized metal affinity chromatography.[20] An in vitro competition assay was used to compare the cleavage proficiency of each enzyme with the wild-type and T(−7)C substrates.[21] Briefly, a DNA plasmid containing one copy of the wild-type substrate and one copy of the T(−7)C substrate was digested with varying amounts of enzyme (see Supporting Information). The amount of enzyme required to cleave 50% of each site was compared to reveal the relative in vitro cleavage efficiencies (Table 2).

The substrate specificities of the evolved homing endonucleases measured by this assay are consistent with the results of the in vivo assay, although the magnitudes of the in vivo preferences (defined by survival rates) are larger than those measured in vitro (Table 1 and Table 2). Consistent with the results in vivo, in vitro assays indicate that clone 1-6 is an endonuclease with broadened specificity that is proficient at cleaving both the wild-type and T(−7)C substrates. In contrast, clone 2-4 exhibits a 3.7-fold in vitro preference for the T(−7)C recognition sequence over the wild-type substrate. This preference is slightly greater than the observed 3.1-fold in vitro preference of wild-type I-SceI for cleavage of the wild-type site over the T(−7)C mutant sequence (Table 2).

To determine if clone 2-4 acquired the ability to cleave purines at position −7, in vitro cleavage assays were repeated with the T(−7)A and T(−7)G substrates. Clone 2-4 has 3.2-fold and 6.6-fold preferences for T(−7)C over T(−7)A and T(−7)G, respectively (Table 3). Encouragingly, the selection for novel DNA specificity in clone 2-4 did not generate an endonuclease with broadened specificity even though a negative selection was not applied against purines at position −7.

To further probe the basis of the evolved substrate specificities, we measured the affinity of each enzyme for the wild-type and T(−7)C substrates by fluorescence polarization under conditions that are known to prevent substrate cleavage but that allow DNA binding (see Supporting Information).[21] The observed K_d of the wild-type I-SceI enzyme and the wild-type DNA substrate (9.7 nM) is in good agreement with the previously reported value derived from gel shift assays (8.4 nM), suggesting that the presence of the fluorophore does not significantly perturb binding affinity (Table 4).[22] During the

Table 2. In Vitro Cleavage Efficiency and Specificity Changes for Wild-Type I-SceI and Evolved Mutants

enzyme	relative cleavage efficiency[a] of T(−7)C[50]/wild-type[50]	specificity shift[b]	specificity broadening[c]
wild-type I-SceI	3.1	—	—
evolved mutant 1-6 (Asn 152 Lys)	1.8	1.7	5.6
evolved mutant 2-4 (Trp 149 Gly, Asp 150 Cys, Asn 152 Lys)	0.27	11.5	0.84

[a] Wild-type I-SceI and evolved clones 1-6 and 2-4 were incubated with the wild-type and T(−7)C recognition sequences simultaneously. The concentration of enzyme necessary to cleave 50% of each site was determined (see Supporting Information). The ratio of these concentrations represent relative cleavage efficiency.[11] [b] The specificity shift is determined by dividing the relative cleavage efficiency of the wild-type enzyme by the relative cleavage efficiency of each evolved enzyme. [c] Specificity broadening is defined as the product of the relative cleavage efficiency of the wild-type enzyme and evolved enzyme. A specificity broadening of 1 indicates no broadening, a value greater than 1 indicates broader specificity, and a value less than 1 indicates a tighter specificity.[11]

Table 3. In Vitro Cleavage Specificity of Evolved Mutant 2-4

base at position (−7) in the substrate	relative cleavage efficiency[a]
C	1
T (wild-type)	0.27
A	0.31
G	0.15

[a] Evolved clone 2-4 was incubated with T(−7)C and either T(−7)A or T(−7)G recognition sequences simultaneously. The concentration of enzyme necessary to cleave 50% of each site was determined (see Supporting Information). The ratio of these concentrations represent relative cleavage efficiency,[11] which was then combined with data from Table 2 for comparison and normalized such that the cleavage efficiency of T(−7)C equals 1.

(19) Moure, C. M.; Gimble, F. S.; Quiocho, F. A. *J. Mol. Biol.* **2003**, *334*, (4), 685−695.
(20) Hochuli, E.; Dobeli, H.; Schacher, A. *J. Chromatogr.* **1987**, *411*, 177−184.
(21) Monteilhet, C.; Perrin, A.; Thierry, A.; Colleaux, L.; Dujon, B. *Nucleic Acids Res.* **1990**, *18*, (6), 1407−1413.
(22) Beylot, B.; Spassky, A. *J. Biol. Chem.* **2001**, *276*, (27), 25243−25253.

Table 4. Dissociation Constants for Wild-Type I-SceI, 1-6, and 2-4 Variants with Wild-Type and T(−7)C Recognition Sequences

enzyme	K_d, wild-type recognition sequence (nM)[a]	K_d, T(−7)C recognition sequence (nM)[a]
wild-type I−SceI	9.7 ± 1.8	13.4 ± 7.7
evolved mutant 1-6	11.4 ± 1.2	11.6 ± 2.7
evolved mutant 2-4	41.3 ± 5.1	12.4 ± 5.3

[a] The data represent the average and standard deviation of three independent measurements of binding affinity using fluorescence polarization of fluorescein-labeled recognition sequences.

course of two rounds of evolution, the characterized mutants show a decreasing affinity for the wild-type sequence while the affinity for the T(−7)C sequence remains constant (Table 4). With the exception of the wild-type enzyme, the difference between the binding affinities for the two substrates largely mirrors the difference between the relative cleavage efficiencies (Tables 2 and 4).

Discussion

Previously developed methods to evolve DNA-binding proteins with tailor-made substrate specificities have provided valuable insights into the mechanisms of DNA recognition. Extensive efforts to evolve zinc finger domains, for example, have led to the development of recognition modules for most three-base triplets and a detailed understanding of many of the key interactions at a molecular level.[23,24] In contrast, successful methods to evolve DNA-modifying enzymes have only begun to be reported. For example, Voziyanov, Schultz, and their respective co-workers reported the use of fluorescent activated cell sorting to evolve interactions between recombinases and novel DNA substrates.[25,26] More recently, Griffiths and co-workers have used in vitro compartmentalization to alter the DNA specificity of *Hae*III methyltransferase.[27] These efforts have highlighted the value of effective selection methods to enable the engineering of enzymes with novel DNA specificity.

In the present work we have developed an efficient in vivo selection system for homing endonuclease activity that enables the researcher to exert evolutionary pressure to favor cleavage of desired substrates and also to disfavor the cleavage of non-target DNA sequences. The use of this system to generate multiple endonuclease variants with tailor-made DNA specificities, in combination with structural analysis, may enhance our understanding of mechanisms of sequence-specific DNA recognition and cleavage. In addition, such efforts may enable the creation of new nucleases for the site-specific manipulation of long DNA sequences or complex mixtures of DNA.

To apply this selection system to the evolution of I-SceI variants with novel DNA specificities, it was first necessary to understand the comprehensive specificity of the wild-type enzyme. I-SceI has previously been characterized by Dujon and co-workers[28] as one of the most specific homing endonucleases known. They created single-base mutants at each position and assessed the efficiency of their cleavage by the wild-type enzyme in a qualitative manner. Our study rigorously quantifies the sequence preference of this enzyme at each position of its target site.

Our substrate profiling results generally agree with the findings of Dujon and co-workers[28] and indicate that I-SceI is highly specific for its wild-type cleavage site. The wild-type sequence was enriched 52-fold after three rounds of selection for cleavage, growing from a predicted 0.44% prevalence in the pre-enrichment library to an observed 23% representation in the enriched library. The next most prevalent sequence, which contains the single mutation G(7)T, represents only 3.4% of the enriched library (Figure 2). The specificity of the enzyme can also be measured by our observation that I-SceI only tolerates an average of 1.1 mutations in its 18-base site among cleavable substrates.

I-SceI shows little specificity at the 3′ end of its cleavage site. Indeed, the T(8)C transition is slightly favored (47% C, 36% wild-type), and non-wild-type bases are favored over the wild-type bases (55% to 45%) for positions 9-11. The lack of specificity at the 3′ end of the substrate is also consistent with the structure of I-SceI,[19] which suggests only one direct protein−DNA contact in this region. Most of the enzyme−DNA contacts in this region involve phosphate interactions that are more likely to enhance affinity rather than specificity.

The enzyme exhibits a strong preference (≥85% of the sites that can be cleaved) for wild-type bases at positions −7, −4, −3, −2, 3, and 6. In particular, wild-type I-SceI almost exclusively requires adenine at position −2 (97% of cleavable sites), cytosine at position 3 (98%), and guanine at position 6 (97%). Based on the available crystal structure of the I-SceI bound to its recognition sequence, Arg 88 and Arg 48 form bidentate hydrogen bonds to G(3) on the antisense strand and G(6) on the sense strand, respectively. This canonical mode of interaction of $N_{\eta 1}$ and $N_{\eta 2}$ of arginine with O^6 and N^7 of guanine is one of the most abundant specific contacts observed in recent surveys of protein−DNA contacts, suggesting that it is likely an important specificity-determining interaction. The high degree of specificity at position (−2) may be explained by the hydrogen bond between Lys 193 and the O^4 carbonyl oxygen of T(−2) on the antisense strand, another frequently observed interaction among protein−DNA interfaces.[29]

On the basis of the structural information suggesting that the high degree of specificity we observed for I-SceI's cleavage of T(−7) arises from a single Asn 152-T(−7) interaction, I-SceI variants were selected for their ability to cleave the T(−7)C recognition sequence preferentially over the wild-type sequence. After two rounds of evolution involving both positive and negative selection, clone 2-4 was shown to significantly prefer the T(−7)C sequence over the wild-type substrate both in vivo and in vitro. The 3.7-fold in vitro preference of mutant 2-4 to cleave the T(−7)C substrate instead of the wild-type substrate is slightly greater than the wild-type enzyme's 3.1-fold preference for the wild-type sequence over the T(−7)C sequence.

Despite oversampling in our round 2 library (1.7 times the theoretical complexity), clone 2-4 was the sole survivor. Theoretically, this same enzyme should have been represented at least 16 times through degenerate codon usage. Diversity may

(23) Beerli, R. R.; Barbas, C. F. *Nat. Biotechnol.* **2002**, *20*, (2), 135−141.
(24) Pabo, C. O.; Peisach, E.; Grant, R. A. *Ann. Rev. Biochem.* **2001**, *70*, 313−340.
(25) Voziyanov, Y.; Konieczka, J. H.; Stewart, A. F.; Jayaram, M. *J. Mol. Biol.* **2003**, *326*, (1), 65−76.
(26) Baldwin, E. P.; Martin, S. S.; Abel, J.; Gelato, K. A.; Kim, H.; Schultz, P. G.; Santoro, S. W. *Chem. Biol.* **2003**, *10*, (11), 1085−1094.
(27) Cohen, H. M.; Tawfik, D. S.; Griffiths, A. D. *Protein Eng. Des. Sel.* **2004**, *17*, (1), 3−11.
(28) Colleaux, L.; Dauriol, L.; Galibert, F.; Dujon, B. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, (16), 6022−6026.
(29) Lejeune, D.; Delsaux, N.; Charloteaux, B.; Thomas, A.; Brasseur, R. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, (2), 258−271.

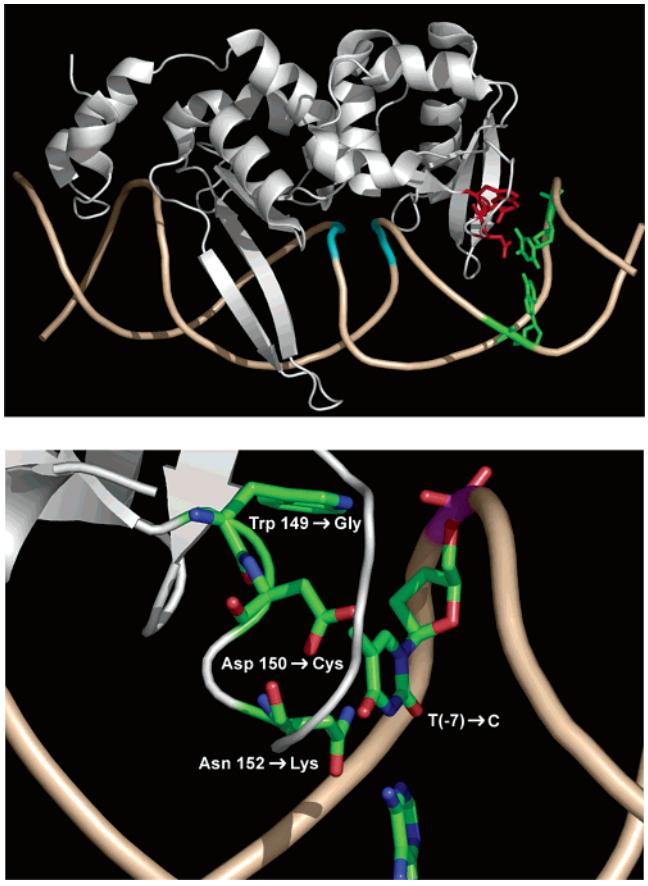

Figure 3. Mutations in evolved enzyme 2-4. (Top) Crystal structure of I-SceI[19] (gray) bound to the wild-type recognition sequence (tan) with the location of the three amino acids (red) mutated in evolved clone 2-4. The two scissile bonds are shown in cyan, and the first base pair of the recognition sequence (−7) is shown in green. (Bottom) Close-up view of I-SceI indicating the residues that are mutated in clone 2-4. This figure was created using PyMOL[30] and the coordinates of Protein Data Bank structure 1R7M.[19]

have been lost transitioning between the positive and negative selections where the larger background colonies can dominate by cell mass the smaller true survivors. A second round of enrichment in the positive selection may be helpful to amplify the true survivors at the expense of background survivors.

The recently solved structure of I-SceI bound to its recognition sequence[19] provides the opportunity to analyze the basis of the evolved mutants' altered specificity (Figure 3). In the wild-type structure, the side-chain carboxamide of Asn 152 is predicted to form a single hydrogen bond with the O^4 carbonyl oxygen of T(−7) in the major groove. Wild-type I-SceI does not recognize the T(−7)C substrate as efficiently as the wild-type substrate presumably because the exocyclic amine of cytosine cannot accept a hydrogen bond from the side-chain amide nitrogen of Asn 152. The carboxamide may not be able to rotate to accommodate this new substrate because it is predicted to make an additional hydrogen bond to Asp 150. Clone 1-6 (Asn 152 Lys) may still be able to make hydrogen bonds to thymine in the wild-type recognition sequence and also to cytosine (or to the guanosine nucleotide on the opposite strand) in the T(−7)C sequence (Figure 3). This possibility may contribute to the observed ability of clone 1-6 to cleave the wild-type and mutant substrate with comparable efficiency.

Clone 2-4 (Trp 149 Gly, Asp 150 Cys, Asn 152 Lys) exhibits a change in specificity in which the T(−7)C recognition sequence is preferred over the wild-type substrate. Our observations that both round 2 mutations are necessary to effect this change in specificity together with their adjacency in the primary sequence of the enzyme suggests that these substitutions together induce a conformational change that accommodates the T(−7)C substrate better than the wild-type substrate. Indeed, the Trp 149 Gly mutation represents a 115 $Å^3$ decrease in protein volume at the protein−DNA interface (Figure 3).[31] This loss in volume could be accommodated by a significant rearrangement of the protein but perhaps more likely is filled, at least in part, with water molecules that may participate in enhancing interactions with the T(−7)C substrate.[32] While these hypotheses are consistent with the structure of I-SceI bound to its recognition sequence,[19] further studies are required to fully dissect the structural basis of the specificity change for clone 2-4.

Since reverting Trp 149 in mutant 2-4 to the wild-type glycine results in a loss of specificity, mutations Asp 150 Cys and Asn 152 Lys are necessary but not sufficient to explain the observed specificity change in clone 2-4. The creation and application of a negative selection system to disfavor the wild-type sequence was a necessary step in the evolution of clone 2-4 as its round 1 precursor was able to survive the positive selection at a high survival rate despite its undesirably broad specificity. The use of a negative selection was also required to achieve altered, as opposed to broadened specificity during the evolution of recombinase enzymes.[25,33] In contrast, previous efforts to generate homing endonucleases with novel cleavage specificities have not applied negative selection, with varying results.[9,11] Although negative selections are not always necessary to generate nucleases with altered specificities, our results demonstrate that a negative selection can be a simple and powerful tool for applying evolutionary pressure against enzymes with broadened specificities.

Homing endonuclease activities as assayed in vivo and in vitro have previously been reported to differ significantly.[9,10] We observed similar differences in our system, as survival rates in vivo between wild-type and T(−7)C mutant substrates typically differed more than cleavage efficiencies measured in vitro. Intrinsic differences between the requirements for cell survival and the nature of in vitro competition assays may explain these observations. If nuclease cleavage of the T(−7)C site falls below a threshold of activity needed to reduce CcdB expression to sub-lethal levels, the in vivo assay will not detect cleavage even if significant cleavage is taking place. Given that we are able to detect a range of survival frequencies in vivo (Table 1), however, it is unlikely that this explanation is the sole cause of in vivo−in vitro differences. Previous studies have also observed higher apparent specificities in vivo,[9,10] where conditions more closely mimic a natural setting in which the enzymes must find their cognate sequence from among a vast excess of noncognate genomic DNA—an excess that is reduced under typical in vitro assay conditions.

Our most evolved enzyme possesses a 3.7-fold preference for the T(−7)C site over the wild-type site. This degree of

(30) DeLano, W. L. *The PyMOL User's Manual*, ed.; DeLano Scientific: San Carlos, CA, 2004.
(31) Creighton, T. E. *Proteins*, 2nd ed.; W. H. Freeman and Company: New York, 1993.
(32) Jayaram, B.; Jain, T. *Ann. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 343−361.
(33) Santoro, S. W.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, (7), 4185−4190.

specificity is very similar to the specificity of wild-type I-SceI for the wild-type site over T(−7)C (Table 2.) As mentioned, I-SceI is known to be one of the most specific homing endonucleases, and as a result, it is often used for genome engineering where cleavage of a noncognate site would result in a lethal double-stranded chromosome break.[2−5] Especially in light of differences between in vitro and in vivo assays, the observed specificity of these and similar I-SceI variants may be sufficient for their use in vivo to target natural or engineered noncognate sequences within genomic DNA.

Conclusions

We report the development of a strategy for the directed evolution of homing endonucleases with altered substrate specificities and its application to evolve mutant I-SceI enzymes that prefer to cleave a mutant DNA sequence over the wild-type substrate. Our approach combines a positive selection for target DNA cleavage with a negative selection against cleavage of nontarget DNA sequences. The resulting two-plasmid selection system is simple to use, should be applicable to other endonucleases, and is efficient enough to evaluate large libraries ($>10^8$ transformants) using only two rounds of enrichment. We also used a selection-based system to comprehensively profile the substrate specificity of the I-SceI homing endonuclease. These results suggest the feasibility of extensively exploring the plasticity of nuclease−DNA interfaces and may also facilitate the creation of additional homing endonuclease enzymes with tailor-made DNA cleavage specificities.

Experimental Section

General Reagents and Methods. DNA oligonucleotides were purchased from Sigma-Genosys or Integrated DNA Technologies or synthesized on an ABI Expedite 8909 DNA synthesizer (PerSeptive Biosystems). Restriction enzymes, Vent DNA polymerase, T7 DNA polymerase, and T4 DNA ligase, were purchased from New England BioLabs. *Taq* DNA polymerase was purchased from Promega.

Plasmid Preparation. Plasmid pCcdB was constructed by introducing *Aat*II and *Asc*I sites into the plasmid pBar2-I-ScaI-site (trimer)[12] using the PCR primers 5′-GCTATAGGCGCGCCTTTACAAAAATCAGATAAGCATGCC and 5′-GCATTAGACGTCCAGGCACATTATGCATCGATGATAAGC. A fragment including the *lac* promoter driving expression of a *lacZ*α−*ccdB* fusion was amplified from pZero-2 (Invitrogen), introducing *Aat*II and *Asc*I sites using the PCR primers 5′-GCATTAGACGTCGAGTTACAACAGTCCGCACCGCTGTCC and 5′-GCATTAGACGTCGCGCAACGCAATTAATGTGAGTTAGC. The ColE1 origin was amplified from pUC19 using the PCR primers 5′-TGTCTAATCGATCTGGCGTTTTTCCATAGGCTCC and 5′-TGTCTACCGCGGAGACAGATCGCTGAGATAGG and introduced into pCcdB using *Sac*II and *Cla*I sites. *Afl*III and *Bgl*II sites were introduced into pCcdB using the PCR primers 5′-GGCGCGCCTTTACAAAAATCAGATAAACGTGTCTGAGATCAAACTGTCACATTGAGGTGC and 5′-GCACTAGTTATTACCAGAGATCTTCGGATCAAACGCCATGAGCG with a modified version of the multiple Quickchange protocol[34] that lacks the final PCR step. These restriction sites were used to introduce one copy of wild-type I-SceI site (TAGGGATAACAGGGTAAT) for pCcdB-wt1. Two copies of wild-type sites were also introduced into the *Nhe*I/*Sac*II fragment of this plasmid along with three copies adjacent to the original copy to create pCcdB-wt6. For pCcdB-mut3, two copies of T(−7)C sites (CAGGGATAACAGGGTAAT) were introduced between the *Nhe*I and *Sac*II sites and a single copy was introduced into the *Afl*III/*Bgl*II fragment. Three additional copies of T(−7)C sites were introduced into the *Afl*III/*Bgl*II fragment to generate pCcdB-mut6.

Plasmid pEndoSce was constructed by introducing *Sph*I and *Not*I sites into pACYC177 using the PCR primers 5′-GCTACAGCATGCCGACAACGCAGACCGTTCCGTGGC and 5′-GATCTAAGCGGCCGCTCAGAATTGGTTAATTGGTTGTAACACTGG. The *araC* and pBAD promoter were amplified from pT2Bgl[12] using the PCR primers 5′-GCTACAGCATGCGCATAATGTGCCTGTCAAATGGACG and 5′-GCGGACGTGCCATGGGTATATCTCCTTCTTAAAGTTAAACAAAATTC, creating *Sph*I and *Nco*I sites. The DNA encoding I-SceI endonuclease was amplified by PCR from pSupE-I-SceI[12] using 5′-AGCTCCATGGCAATGAAAAACATCAAAAAAAACCAGG (primer 1) and 5′-TATCAAATGCGGCCGCTTATTATTTCAGGAAAGTTTCGGAGG (primer 2). A silent *Eco*RI site for library cloning was introduced into I-SceI using PCR primer 1, primer 2, 5′-GCAGTCGAATTCTACCAACAAATCGATCGTAC, and 5′-GCAGTCGAATTCTTGTTGTAATCCCATTTACCACC. The catalytically inactive mutant of I-SceI (containing the Asp 44 Ser mutation) was amplified from pSupE-I-SceI-D44S[12] with primers 1 and 2 and subcloned into pEndoSce as a *Nco*I/*Not*I fragment.

Library Construction. The round 1 library was constructed by PCR using the primers 5′- CTTTAAGAAGGAGATATACCCATGG (primer 3) and 5′-GTTGGTAGAATTCTTNNNGTAATCCCATTTACCACC and cloned into the large *Nco*I/*Eco*RI fragment of pEndoSce. The round 2 library was created using primer 3 together with 5′-GTTGGTAGAATTCNNTTTNNNNNNNNNTTTACCACCATCATCCATGAACC (where N = equal mixtures of all four nucleotides) and cloned into the *Nco*I/*Eco*RI backbone of pEndoSce. The mutations at positions 149 and 150 from clone 2-4 were individually reverted back to wild-type amino acid using primer 3 together with 5′-CTTGGTACAATTCTTTTTAT-AACACCATTTACCACCATCATCC and 5′-CTTGGTACAATTCTT-TTTATAATCGCCTTTACCACCATCATCC, respectively. The amplified fragments were cloned as *Nco*I/*Eco*RI fragments into pEndoSce.

Positive Selection. Individual pEndo plasmids or libraries were transformed by electroporation into DH12S *E. coli* competent cells (Invitrogen) harboring pCcdB-mut3. Cells were recovered in 2xYT containing 0.02% arabinose for 30 min at 37 °C. Carbenicillin (Cb) was added to final concentration of 100 μg/mL and the cells were transferred to 30 °C for 3.5 h. The cells were washed twice with 1x M9 salts and then plated on both nonselective plates (2xYT + Cb) and selective plates (1x M9 salts, 1% glycerol, 0.8% Cas amino acids, 1 mM $MgSO_4$, 1 mM $CaCl_2$, 2 μg/mL thiamine, 100 μg/mL Cb, 0.02% arabinose). Selective plates were grown for 3 days at 30 °C.

Round 2 Selection. The surviving clones from the positive selection of the saturation mutagenesis library at positions 149, 150, 151, and 153 were streaked onto Cb and Cb/chloramphenicol (Cam) plates to screen for clones without the CcdB encoding plasmid. The clones that survived on Cb only were combined, the genes encoding the endonucleases were amplified with primer 3 and 5′-GCATTGCTGCAGGCATCGTG, and cloned into the *Nco*I/*Pst*I fragment of pEndoSce.

The resulting library was transformed into pCcdB-wt6 and processed exactly as the positive selection, except that plates were supplemented with Cam at 40 μg/mL and did not contain IPTG. The plates were grown for 2 days at 30 °C. The surviving colonies were combined, and genes encoding endonucleases were amplified using primer 3 together with 5′-CCAATTCTGAGCGGCCGC (primer 4) and cloned into the *Nco*I/*Not*I fragment of pEndoSce. The resulting library was subjected to a final positive selection as described above.

Protein Purification. A C-terminal His_6-tag was added to the genes encoding wild-type I-SceI, 1-6, and 2-4 by PCR using primer 3 together with 5′-CGTCCAGCTGCGGCCGCTTCCTGATGATGAT-GATGAT-GATGTTTCAGGAAAGTTTCGGAGATAG. The resulting PCR products were cloned into the large *Nco*I/*Not*I fragment of pBluescript II SK (+)-Nco.[12] Plasmids were transformed into DH12S competent cells and grown to $OD^{600} = 0.5$. IPTG (1 mM) was added, and the cells

(34) Sawano, A.; Miyawaki, A. *Nucleic Acids Res.* **2000**, *28*, (16), e78.

were grown for 4 h at 30 °C. Cells were harvested by centrifugation and resuspended in 40 mL of buffer A (200 mM Tris pH 8.0, 200 mM KCl, 1 mM PMSF, 2 mM BME). Lysozyme (156 μg/mL) was added for 10 min at room temperature, and the mixture was sonicated for three 1 min cycles (10 s on, 4 s off, ×6) using the microtip of a Sonic Dismembrator 550 (Fisher). Cellular debris was removed by centrifugation at 4 °C, and the supernatant was filtered through a 0.2 μm membrane. Ni-NTA agarose resin (100−300 μL, Qiagen) was added to the supernatant and agitated for 1 h or overnight at 4 °C. The mixture was loaded into a column, washed with 10 mL of buffer A containing 15 mM imidazole, and eluted with 1.0 mL of buffer A containing 200 mM imidazole. Fractions containing homing endonuclease, identified by PAGE analysis, were pooled and dialyzed overnight against storage buffer: 10 mM Tris (pH 8.0), 100 mM NaCl, 1 mM DTT, 50% glycerol. Aliquots were stored at −20 °C. Relative protein concentrations were determined by Western blot using a polyclonal horseradish peroxidase-conjugated 6× His antibody (Abcam). The determination of one absolute protein concentration was accomplished by densitometry using BSA standards analyzed by SDS-PAGE and staining with Gelcode blue stain (Pierce) and was used to calculate the protein concentrations of the remaining proteins. Quantifications were performed in triplicate.

Library Construction for Substrate Specificity Assay. I-SceI substrate specificity experiments were based on the method described by Monnat and co-workers.[13] A starting plasmid library of partially randomized I-SceI homing sites was created using a 54-base oligonucleotide designed (5′-GCAAGCGGTGCTAGC**TAGGGATAACAGGGTAAT**CCGCGGCGATGTACGGCAATC) so that any given position in the homing site (shown in bold) would have a 72% probability of containing the wild-type base and a 9.3% probability of containing each of the other three non-wild-type bases. The partially randomized positions were created by incorporating pools of hand-mixed phosphoramidites with different pools for each base. The oligonucleotide was made double-stranded by primer extension with T7 DNA polymerase and the primer 5′-GATTGCCGTACATCGCCGCGG. The resulting double-stranded partially randomized I-SceI substrate library was digested with *Nhe*I and *Sac*II and ligated into the *Nhe*I/*Sac*II fragment of pCcdB to generate pCcdB-lib.

Sequential Enrichment of Cleavable I-SceI Substrates. Three rounds of in vitro sequential enrichment were performed on pCcdB-lib. For each round, 275 ng of the plasmid library was digested in a 20 μL reaction with purified wild-type I-SceI in I-SceI buffer (New England BioLabs) and BSA (100 μg/mL) for 30 min at 37 °C. The amount of I-SceI added was experimentally determined to be the amount necessary to completely cleave a plasmid containing one copy of the wild-type I-SceI sequence in the same assay. Cleaved, linearized plasmids were separated from uncleaved plasmids by gel electrophoresis in loading buffer with 1% SDS at 1.4 V/cm overnight on a 1% TAE agarose gel. Finally, gel purified linear plasmid libraries were recircularized with T4 DNA ligase (20 units) for 1 h at 16 °C and transformed into DH12S *E. coli* competent cells. Plasmid DNA was isolated and used as a substrate for the next round of cleavage and enrichment.

Individual clones of the starting library (287 clones) and the thrice-enriched library (597 clones) were sequenced. Base frequencies at each site, f_b^s, were calculated and renormalized to approximate the results of enrichment on an unbiased homing site library with all possible 18-mer sequences using the following formula:[35]

$$f_b^s = \frac{\frac{f_b}{p_b}}{\sum_{\text{bases}} \frac{f_b}{p_b}}$$

where f_b is the observed frequency of base b at a given position in the enriched library and p_b is the frequency of base b at a given position in the starting library. Sequence logos were created by inputting f_b^s values into the ENOLOGOS program.[36]

In Vitro Cleavage Assays. One copy of the wild-type and T(−7)C recognition sequences were cloned between the *Nhe*I/*Sac*II and *Afl*III/*Bgl*II sites of pCcdB, respectively, to yield pSites. One copy of the T(−7)G and T(−7)A were cloned between the *Nhe*I/*Sac*II sites of pSites to yield pSites-G and pSites-A, respectively. Eight aliquots of 50 μL containing cleavage buffer (10 mM Tris pH 8.0, 10 mM $MgCl_2$, 1 mM DTT), 100 ng of pSites, pSites-A, or pSites-G, and 100 μg/mL BSA was preequilibrated at 30 °C for 10 min. Purified enzyme was added in 2-fold dilution to each of the tubes and incubated at 30 °C for 20 min. Reactions were stopped by adding EDTA (50 mM) and freezing the samples on dry ice. Samples were extracted with 1:1 phenol/chloroform (EMD Biosciences) and then exchanged into water using Centrisep columns (Princeton Separations). Samples were then digested to completion with excess *Not*I to distinguish between the cleavage of the wild-type and T(−7)C sites. Samples were analyzed using agarose gel electrophoresis and stained with ethidium bromide. The extent of cleavage was quantified by densitometry using an Eagle Eye II densitometer (Stratagene). The amount of enzyme necessary to cleave 50% of each site was determined in KALEIDAGRAPH (See Supporting Information). Each assay was performed in triplicate.

Fluorescence Polarization Binding Assays. Fluorescein-labeled oligonucleotides were purchased from Sigma Genosys. The fluorescein label was incorporated into the recognition sequence using a modified thymine base (FLDT). The wild-type substrate was GCTTATATTACCCTG[FLDT]TATCCCTAGCGTAA, and the T(−7)C substrate was GCTTATATTACCCTG[FLDT]TATCCCTGGCGTAA. Each oligonucleotide was slowly annealed to its complement. The binding conditions were 10 mM Tris pH 8.8, 10 mM $CaCl_2$, 1 mM DTT, 1.5 nM labeled substrate, and a constant concentration of enzyme storage buffer. Increasing concentrations of enzyme were added to the reaction and equilibrated at room temperature for 30−60 min. Fluorescence polarization measurements were taken using an Analyst AD (Molecular Devices) and analyzed using KALEIDAGRAPH as described elsewhere.[37] For each enzyme/substrate pair, multiple measurements were taken, and the experiments with the three best data fits were averaged.

Acknowledgment. We thank Kathy Chang and Jon Thomas for assistance with the initial cloning of the positive selection system. Gregory Gatto provided valuable insight for the fluorescence polarization measurements. Keith O'Neill at the Broad Institute provided sequencing for the substrate specificityassays. This work was supported by the National Institutes of Health/National Institute of General Medical Sciences (R01GM-065400) and the Howard Hughes Medical Institute. J.B.D. gratefully acknowledges a National Defense Science and Engineering Graduate Research Fellowship.

Supporting Information Available: Representative example of the in vitro DNA cleavage assay. This material is available free of charge via the Internet at http://pubs.acs.org.

JA057519L

(35) Stormo, G. D. *Methods Enzymol.* **1991**, *208*, 458−468.
(36) Workman, C. T.; Yin, Y. T.; Corcoran, D. L.; Ideker, T.; Stormo, G. D.; Benos, P. V. *Nucleic Acids Res.* **2005**, *33*, W389−W392.
(37) Michel, S. L. J.; Guerrerio, A. L.; Berg, J. M. *Biochemistry* **2003**, *42*, 4626−4630.