Journal of Protein Chemistry, Vol. 15, No. 5, 1996

Splase: A New Class IIS Zinc-Finger Restriction Endonuclease with Specificity for Sp1 Binding Sites

Baohua Huang,[1] Christopher J. Schaeffer,[1] Qunhui Li,[2] and Ming-Daw Tsai[1–4]

Received January 27, 1996

A new restriction endonuclease, named Splase, was constructed by genetically fusing the DNA-cleavage domain of the restriction endonuclease *Fok*I with the zinc-finger DNA-binding domain of the transcription factor Sp1. The resulting protein was expressed in *Escherichia coli.*, partially purified, and shown to selectively digest plasmid DNA harboring consensus Sp1 sites. Splase was also shown to selectively digest the long terminal repeat of the HIV-1 DNA at Sp1 sites. Splase recognizes a 10-bp DNA sequence and hydrolyzes phosphodiester bonds upstream of the binding sequence. The binding specificity of Splase makes this a "rare cutter" restriction enzyme which could be valuable in creating large DNA fragments for genome sequencing projects. The result also presents the opportunity to create other restriction enzymes by altering the binding specificity of the zinc-finger recognition helix.

KEY WORDS: Endonuclease; zinc finger; *Fok*I; restriction enzyme; HIV-1.

1. INTRODUCTION

Restriction modification systems are composed of enzymes which site-specifically methylate or cleave DNA. These systems can be grouped into three distinct classes. Class I systems are composed of enzymes containing three subunits each separately controlling recognition, modification, and restriction. These enzymes cleave DNA more than 1000 bp away from the recognition site and require ATP for phosphodiester hydrolysis. Class II systems consist of two dimeric enzymes; one for modification (methyltransferase) and one for restriction (endonuclease). Each enzyme binds the same sequence but catalyzes different reactions. Class II restriction endonucleases bind and cleave symmetric sequences within the DNA recognition sequence. Class III enzymes are rare and are not currently used in molecular biology.

The class IIS restriction modification system is a subset of class II systems. It binds an asymmetric sequence as a monomer (Skowron *et al.*, 1993) and cleaves DNA a short distance downstream of the recognition sequence. One example of a class IIS restriction endonuclease is *Fok*I from *Flavobacterium okanokoites* (Sugisaki and Kanazawa, 1981). *Fok*I has been shown to bind an asymmetric DNA sequence and cleave DNA 9 bp downstream on the 5′ strand and 13 bp on the 3′ strand (Yonezawa and Sugiura, 1994). Further studies have shown that in the presence of substrate the protein can be digested by trypsin to give two distinct fragments (Li *et al.*, 1992). Upon purification, these two fragments retain their original functions of DNA binding and DNA cleavage. Kim and Chandrasegran reported the construction of the first chimeric restriction

[1] Ohio State Biochemistry Program, Ohio State University, Columbus, Ohio 43210.
[2] Department of Chemistry, Ohio State University, Columbus, Ohio 43210.
[3] Department of Biochemistry, Ohio State University, Columbus, Ohio 43210.
[4] To whom correspondence should be addressed. E-mail: mtsai@chemistry.ohio-state.edu.

0277-8033/96/0700-0481$09.50/0

endonuclease, Ubx-F_N (Kim and Chandrasegran, 1994), by fusing the DNA cleavage domain of *Fok*I to the helix-turn-helix DNA-binding domain of the *Drosophila* Ultrabithorax protein. This work demonstrated the feasibility of creating chimeric restriction endonucleases with *Fok*I.

Transcription factors bind specific DNA sequences upstream of the transcription initiation site. The classical Cys_2His_2 zinc fingers are one class of transcription factors (Pelham and Brown, 1980). The zinc-finger domain is composed of two β-sheets followed by an α-helix. The cysteine residues of the β-sheet and the histidine residues of the α-helix coordinate a zinc atom that appears to hold the DNA recognition helix in position to accurately recognize specific DNA sequences (Pavletich and Pabo, 1991).

Sp1 (Kadonaga *et al.*, 1987) is a well-studied representative of the classical zinc-finger family of transcription factors. The DNA recognition domain of Sp1 consists of three consecutive zinc fingers which bind a consensus DNA sequence [5′-G(T)GGGCGGG(A)G(A)C(T)]3′] (Kriwacki *et al.*, 1992). The 92 amino acids which comprise this domain have been recently shown to bind this sequence with a similar affinity to that of the naturally occurring protein ($K_d \approx 10$ nM). Sp1 has also been shown to be involved in the regulation of HIV-1 gene expression (Jones *et al.*, 1986).

We describe here the construction of a new rare cutter restriction endonuclease (Splase) containing a zinc-finger DNA-binding domain. This protein recognizes the consensus binding sequence for the transcription factor Sp1 and cleaves DNA upstream of the binding sequence. The zinc fingers of Sp1 were chosen for this study due to the presence of three high-affinity Sp1 binding sites located in the long terminal repeat (LTR) of HIV-1 (Jones *et al.*, 1986). This protein preferentially digests plasmid DNA containing Sp1 binding sites and also recognizes the Sp1 binding sites found in the LTR of HIV-1.

2. EXPERIMENTAL PROCEDURES

2.1. Construction of the Splase Gene

The zinc fingers of Sp1 were amplified using pSp1-516C (Kadonaga *et al.*, 1987) as a template. pSp1-516C was a generous gift of Dr. Robert Tjian and contains the coding sequence for the transcription factor Sp1. The following primers were used in 30 rounds of PCR [denaturation (D) = 94°C, 2 min, annealing (A) = 60°C, 2 min, extension (E) = 72°C, 2 min: 5′ primer 5′-GTC CAT GGC TAA AAA GAA CAG CAT ATT TGC CAC-3′ and the 3′ primer 3′-CTG GGT GGT CTT ATT CTT CCA TGG CC-3′ with 100 ng of template DNA, 20 pmol of each primer, 4 mmol of each dNTP, 1 unit of Ultima polymerase (Perkin Elmer), Ultima polymerase buffer, and 100 nmol of $MgCl_2$ in a 100 μl total volume. The resulting fragment was digested with *Nco*I and *Kpn*I and cloned into pTrc99a (Pharmacia) to generate pTrcSp1. *Flavobacterium okeanokoites* (American Tissue Type Culture Center) genomic DNA was isolated as previously described (Sugisaki and Kanazawa, 1981). The 609-bp fragment encoding the nuclease domain of *Fok*I was amplified from *Flavobacterium okeanokoites* genomic DNA using 30 cycles of PCR (D = 94°C, 2 min, A = 55°C, 2 min E = 72°C, 3 min). The same concentrations of reactants were used for this amplification as were used above, except 300 ng of genomic DNA was substituted as template. The primers used in the amplification were the 5′ primer 5′-CGG GTA CCT AAT CGT GGT GTG ACT AAG-3′ and the 3′ primer 5′-TTA TTG CCG CTC TAT TTG AAA ATT CCT AGG CG-3′. The resulting fragment was first digested with *Kpn*I and *Bam*HI and then ligated into pTrcSp1, generating pTrcSplase. One additional round of PCR was used to add a DNA sequence encoding six histidine residues to the N-terminus. The primers used for PCR are the 5′ primer 5′GTC CAT GGC TCA TCA CCA TCA CCA TCA CAA AAA GAA ACA GCA TAT TTG CCA C-3′ and the 3′ primer 5′-TTA TTG CCG CTC TAT TTG AAA ATT CCT AGG CG-3′. pTrcSplase was used as template for this reaction. The resulting fragment was digested with *Nco*I and *Bam*HI and cloned into pTrc99a creating pTrcHSplase.

2.2. Expression

pTrcHSplase was transformed into *E. coli* JM105 cells (Pharmacia) for expression. Cells harboring the pTrcHSplase gene were grown in 2.5 L of rich medium (19 g of yeast extract, 12 g of tryptone, and 10 mmol of $MgSO_4$ per liter) containing 200 μg/ml ampicillin at 37°C to $A_{600\,nm} = 0.6$. The flasks were then cooled on ice to

25°C, induced with 0.4 mM IPTG, and further grown for 16 h at 25°C. The cells were harvested by centrifugation at 6000 × *g* for 5 min in a Beckman JA-10 rotor.

2.3. Purification

The harvested cells (18 g) were resuspended in 54 ml of binding buffer (6 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, *p*H 7.9 at 4°C), sonicated 6 × 1 min on ice using a Vibra Cell Ultrasonic Processor 300-W model (Sonics and Materials), and centrifuged at 15,000 × *g* for 25 min at 4°C to remove the cellular debris. The supernatant was adjusted to *p*H 7.9 and loaded onto a 5-ml His-bind (Novagen) column preequilibrated with binding buffer. After washing with 10 volumes of binding buffer followed with 10 volumes of wash buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, *p*H 7.9 at 4°C), the protein was eluted with 20 ml of elution buffer (0.5 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl, *p*H 7.9 at 4°C). The enzyme-containing fractions were loaded onto a Sephadex G-25 column, and the buffer was exchanged with enzyme buffer (10 mM Tris-HCl, 50 mM NaCl, 5 mM DTT, *p*H 7.7). Then 50-μl aliquots of the protein solution were subsequently placed into 1.5-ml microfuge tubes, frozen in liquid nitrogen, and stored at −70°C.

2.4. Construction of Substrates for HSplase

A plasmid carrying a single Sp1 site, pUCSp1, was prepared by inserting a 26-mer DNA fragment containing a single Sp1 site into the polylinker site of pUC19 between the *Eco*RI site and the *Bam*HI site.

5′-AATTCGCCGGGGCGGGGCTTCTGCAG-3′
3′-GCGGCCCCGCCCCGAAGACGTCCTAG-5′

The inserted DNA fragment contained the Sp1 binding site with its flanking sequence from the human metallothionein IIa gene (Kriwacki *et al.*, 1992). Other pUC19 derivatives with multiple copies of Sp1 sites were constructed by inserting one to four copies of a 24-mer DNA fragment containing a single Sp1 site into pUCSp1 at the *Xba*I site. The resulting plasmids contain between two and five copies of the Sp1 site and are designated as pUC2Sp1 through pUC5Sp1.

The pUC18-derived plasmid carrying the LTR sequence of HIV-1, pUC-BENN-CAT (Gendelman *et al.*, 1986), was obtained from NIH.

2.5. Activity Assay

Plasmid DNA, pUC19, and pUC35Sp1 were initially linearized with *Alw*NI and separated on a 1% agarose gel, and the 2.7-kb linearized plasmid gel was purified using Geneclean (Bio101). The resulting DNA fragment was then subjected to digestion by HSplase. pUC-BENN-CAT was first linearized with *Bam*HI, gel purified, and then digested with HSplase. A total of 0.3 pmol of substrate DNA was digested by 30 fmol of HSplase in a 20-μl reaction mixture containing 10 mM Tris-HCl, 50 mM NaCl, 0.1 mM $ZnSO_4$, 25 mM DTT, 100 μg/ml BSA, and 2 mM $MgCl_2$ at 37°C. After 1 hr the reaction mixture was removed and the products separated on a 0.7% agarose gel. The products were visualized by ethidium bromide staining.

A complementary experiment which examined the effect of HSplase on supercoiled DNA was performed by first incubating the DNA substrate (0.3 pmol of pUC5Sp1) in the presence of 30 fmol HSplase. The products were separated on a 0.7% agarose gel and the 2.7-kb linearized fragment gel-purified. The 2.7-kb fragment was then digested with *Alw*NI to generate fragments which correspond to the digestion position of HSplase.

2.6. Mapping the Site of Cleavage for HSplase

Oligonucleotides containing three Sp1 sites were designed: (G-rich strand) 5′-AAA CGA CGG CCA GTG AAT TCG CCG GGG CGG GGC GAT CCT CTA GGC CGG GGC GGG GCT TCT GCA CTA GGC CGG GGC GGG GCT TCT GCA CTA GGC CGG GGC GGG GCT TCT GCA CTA-3′, (C-rich strand) 5′-GGT AGT GCA GAA GCC CCG CCC CGG CCT AGT GCA GAA GCC CCG CCC CGG CCT AGA GGA TCG CCC CGC CCC GGC GAA TTC ACT GGC CGT CGT-5′. Each oligonucleotide was independently 5′-phosphorylated using T4 polynucleotide kinase (New England Biolabs) and γ-^{32}P-ATP. The kinase reactions were incubated at 37°C for 1 hr and loaded onto a Biospin-6 (BioRad) spin column to remove unincorporated nucleotides. Each oligonucleotide was eluted into 45 μl of water to which a tenfold excess of both the G-rich and

C-rich unlabeled oligonucleotides were added. The oligonucleotides were denatured at 90°C for 2 min, removed, and allowed to cool slowly to room temperature for efficient annealing. Each of the specifically end-labeled DNA fragments was used as a separate substrate in the activity assay described above. Upon completion of the digestion, sequencing loading dye was added to each of the samples. The samples were denatured at 90°C for 5 min, placed on ice briefly, and then loaded onto a 12% denaturing PAGE gel (7 M urea). Simultaneously Maxam and Gilbert G + A sequencing reactions were performed (Maxam and Gilbert, 1980). These reactions were loaded alongside the corresponding digestion reactions in order to determine the site of cleavage. The relative positions of the resulting products were determined using autoradiography. Fragments generated during the sequencing reactions corresponded to the length of the product generated by HSplase digestion. Thus, by reading the sequence, the base which corresponded to the site of cleavage could be determined.

2.7. Electrophoretic Mobility Shift Assay

A 63-mer oligonucleotide containing a single Sp1 site 5′-CCG GAG TCA GAA TTC GAA GAC TTG CCG CCG GGG CGG GGC TTC TGC AAT CTG CAG GCC AGC TGT-3′ and its complementary strand 5′-AAT TAC AGC TGG CCT GCA GAT TGC AGA AGC CCC GCC CCG GCG GCA AGT CTT CGA ATT CTG ACT-3′ were annealed and both strands were 5′-end-labeled with γ-^{32}P-ATP using T4 polynucleotide kinase. The radiolabeled DNA fragment was loaded onto a Biospin-6 column to remove unincorporated γ-^{32}P-ATP and to change the buffer to binding buffer [10 mM Tris-HCl, *p*H 8.5, 50 mM NaCl, 100 μM $ZnSO_4$, 5 mM DTT, 8% glycerol, and 200 ng/μl poly dA/dT (Pharmacia)]. Labeled DNA (20 fmol) was incubated with increasing amounts of HSplase (0, 50, 100, and 200 fmol) at 25°C for 20 min and loaded onto a 5% native PAGE gel (30:1 polyacrylamide:bisacrylamide) in 0.25 × Tris borate buffer (23 mM Tris-borate). The same procedure was followed for the 90-mer oligonucleotide used previously in the mapping of the site of cleavage.

The protein concentration throughout these experiments was estimated from protein run on 12.5% SDS-PAGE.

3. RESULTS

3.1. Construction of the Splase Gene

The first two residues of Splase are methionine and valine. The next 92 amino acids are derived from the zinc-finger motif of Sp1, except that G92 is changed to valine for the ease of gene fusion. The C-terminal 203 amino acids of Splase are derived from the C-terminal cleavage domain of *Fok*I. The DNA fragment coding for the zinc-finger domain of Sp1 was generated by PCR from the template plasmid pSp1-516C. The fragment was then digested with *Nco*I/*Kpn*I and ligated into pTrc99a to give pTrcSp1. The DNA fragment coding for the cleavage domain of *Fok*I (Li *et al.*, 1992) was also generated by PCR; here the template was genomic DNA isolated from *Flavobacterium okeanokoites* (Sugisaki and Kanazawa, 1981). The DNA fragment was digested with *Kpn*I/*Bam*HI and ligated into pTrcSp1. The resultant plasmid, pTrcSplase, fused the DNA sequence coding for the DNA-binding domain of Sp1 to that coding for the DNA-cleavage domain of *Fok*I (Li *et al.*, 1992). In addition, codons for six consecutive histidines were added to the Splase gene to obtain the HSplase gene, which facilitated purification of Splase by metal chelating chromatography (Hochuli *et al.*, 1987). The final expression plasmid is shown in Fig. 1.

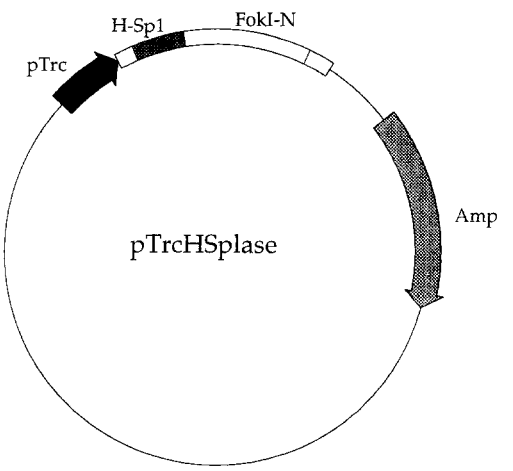

Fig. 1. pTrcHSplase, the expression plasmid used to express HSplase. This plasmid was derived from pTrc99a (Pharmacia). The plasmid contains the HSplase gene cloned into the *Nco*I and *Bam*HI sites located within the multiple cloning site, downstream of the Trc promoter. The plasmid contains an ampicillin gene for the purpose of selection of transformants containing pTrcHSplase.

3.2. Expression and Purifiction of HSplase

Expression of the protein at 37°C resulted in the formation of inclusion bodies. The expressed protein was refolded and purified to homogeneity. This protein showed binding to Sp1 sites in the absence of Mg^{2+}; however, in the presence of Mg^{2+} the DNA was degraded nonspecifically (data not shown). When the temperature was lowered to 25°C prior to induction, a small fraction of the protein was expressed as a soluble protein. This fraction was partially purified using a His-bind column producing a typical yield of 5–10 mg (~60% pure) from a 2.5-L culture. The HSplase thus obtained is used for the rest of the studies described below, and is referred to interchangeably with Splase.

In addition to the pTrc vector, both Splase and HSplase genes were also expressed with the pTO-N (Deng *et al.*, 1990) vector (results not shown).

3.3. Restriction Digestion of DNA by HSplase

The plasmid DNA of pUC19 and pUC5Sp1 were initially linearized by the restriction enzyme *Alw*NI, gel purified, and digested with HSplase. Detectable cleavage was not observed in the lanes containing the linearized pUC19 DNA (Fig. 2, lanes 2 and 3), while pUC5Sp1 was digested into two fragments with the expected sizes of 1.9 and 0.8 kb (lanes 4 and 5). The sizes of the two fragments were consistent with DNA cleavage near the Sp1 sites. More importantly, the *Bam*HI-linearized pUC-BENN-CAT carrying the HIV-1 LTR sequence with three consecutive Sp1 sites was also cut specifically by HSplase. As expected, the cleavage of the *Bam*HI-linearized pUC-BENN-CAT by HSplase near the Sp1 sites generated two fragments of 4.2 and 1.7 kb (lanes 6 and 7). The plasmid pUC3Sp1 was also digested using the same procedure. However, reactions involving pUCSp1 and pUC2Sp1 did not result in detectable digestion at the Sp1 sites (data not shown). In order to test if HSplase was binding substrates with only one binding site, band shift experiments were performed as described in a later section.

In a complementary experiment, pUC5Sp1 was first linearized with HSplase, the 2.7-kb fragment purified, and the product digested with *Alw*NI. As shown in Fig. 3 the 2.7-kb linear DNA was cut into two fragments of 1.9 and 0.8 kb. This experiment showed the effective cleavage of supercoiled DNA

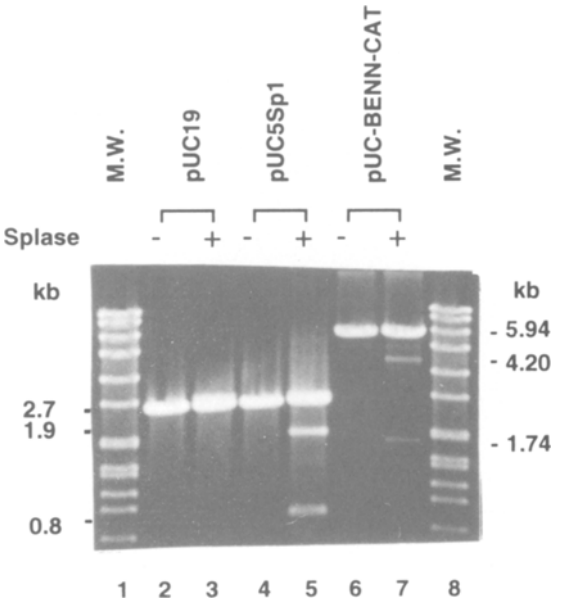

Fig. 2. Digestion of linearized DNA by HSplase. The pUC19 and pUC5Sp1 plasmid DNAs were first linearized by digestion with *Alw*NI and pUC-BENN-CAT DNA was linearized by *Bam*HI. The linearized DNAs were then used as substrates for Splase digestion. The reaction mixtures (total volume = 10 μl) contained 10 mM Tris-HCl, 0.75 μg of plasmid DNA, 2 mM $MgCl_2$, 25 mM DTT, 0.1 mM $ZnCl_2$, 100 μg/ml BSA, and 10 ng of HSplase. The reaction was incubated at 37°C for 3 hr and subjected to 0.7% agarose gel electrophoresis. Lanes 1 and 8, molecular weight marker (marker VII from Boehringer Mannheim); lane 2, pUC19 + *Alw*NI; lane 3, pUC19 + *Alw*NI + HSplase; lane 4, pUC5Sp1 + *Alw*NI; lane 5, pUC5Sp1 + *Alw*NI + HSplase; lane 6, pUC-BENN-CAT + *Bam*HI; lane 7, pUC-BENN-CAT + *Bam*HI + HSplase.

by HSplase. The Mg^{2+} concentrations for each of the previously described experiments were optimized by digesting substrate in a reaction buffer containing between 0.1 and 10 mM Mg^{2+} (data not shown). The optimal concentration of 2 mM Mg^{2+} optimized catalysis while minimizing nonspecific cleavage.

3.4. Mapping the Site of Cleavage for HSplase

A 90-mer oligonucleotide containing 3 Sp1 binding sites (G-rich strand) was selectively 5′-end-labeled and annealed with the corresponding unlabeled strand (C-rich strand). The fragment was digested with HSplase and the products separated alongside the G + A Maxam and Gilbert chemical sequencing reactions on a 12% denaturing (7 M urea) PAGE gel. The same protocol was used for the C-rich strand and the results detected by

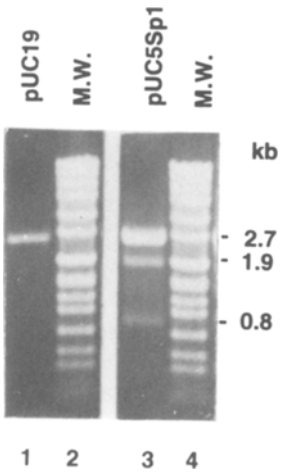

Fig. 3. Linearization of supercoiled DNA by HSplase. The reaction conditions were similar to those of Fig. 2. The amounts of plasmid DNA used were 0.5 μg for pUC19 and 1 μg for pUC35Sp1. The reaction mixture was incubated at 37°C for 2 hr. The linearized DNA band at 2.7 kb was purified and then incubated with 10 units of *Alw*NI in the buffer supplied by the manufacturer at 37°C for 1 hr and analyzed by 0.8% agarose gel electrophoresis. Lanes 1 and 2: from pUC19 as a control; lanes 3 and 4 from pUC5Sp1.

autoradiography. The cleavage products were compared with the DNA sequence and the sites of cleavage determined. HSplase appears to cleave the DNA substrate six times on the G-rich strand between bases 4 and 10, 5′ to the second binding site (Sp1-2). The C-rich strand has three major cleavage sites after the 2nd, 3rd, and 13th bases 3′ to the Sp1 binding sequence. There are also minor cleavages after bases 10, 11, and 12, 3′ to the sp1 binding sequence. The results are summarized in Fig. 4. The cleavages observed on the C-rich strand were significantly more intense than the cleavages observed on the G-rich strand. The same cleavage pattern was observed upstream of the other two Sp1 binding sites; however, the cleavage pattern observed upstream of the first Sp1 binding site (Sp1-1) was very weak, and the cleavage observed upstream of the third Sp1 binding site (Sp1-3) was masked by nonspecific degradation products.

3.5. Electrophoretic Mobility Shift Assay

The mobility shift assays were performed in the absence of Mg^{2+} to prevent nuclease digestion of the probes. Both the 63-mer and 90-mer oligonucleotides were annealed with their respective complementary strands and 5′-end-labeled. These DNA fragments were incubated with increasing amounts of HSplase, and the ability of HSplase to bind was assessed. In the presence of an ~10 fold excess of HSplase both the 63-mer and the 90-mer are shifted. The 63-mer contains only one Sp1 site and was clearly shifted. The 90-mer contains three Sp1 sites and three shifted bands were clearly observed, even though the third band is very faint. The results of this experiment are shown in Fig. 5.

4. DISCUSSION

We report here the construction of a rare cutter restriction endonuclease, Splase, which has been designed to digest DNA containing the Sp1 binding sequence. There were two original goals: (i) to design a new restriction enzyme that can digest the DNA of HIV prior to its integration into the human genome; (ii) to design a rare cutter restriction enzyme and to develop a general approach to designing new restriction enzymes. In the following sections we discuss the results and evaluate the two goals.

4.1. Splase as a Potential Anti-HIV Agent

The goal was to generate an enzyme which could recognize the Sp1 binding sites located on the LTR of HIV-1 and which could cleave the LTR prior to the integration. Splase has been shown to cleave the Sp1 sites located on the LTR of HIV-1

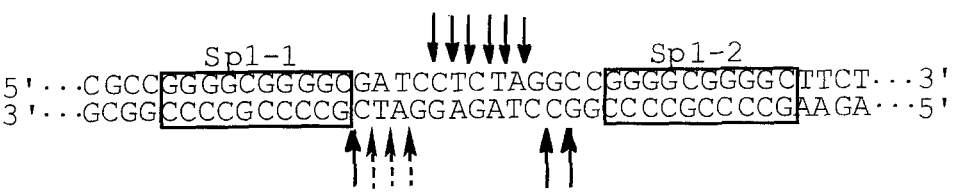

Fig. 4. Diagram of cleavage specificity of HSplase. The arrows correspond to cleavage sites observed on autoradiogram. The arrows designate the intensity of the band observed by autoradiography. The long solid arrows show the positions of the most intense cleavage sites. The dashed arrows indicate the positions of relatively weaker sites of cleavage.

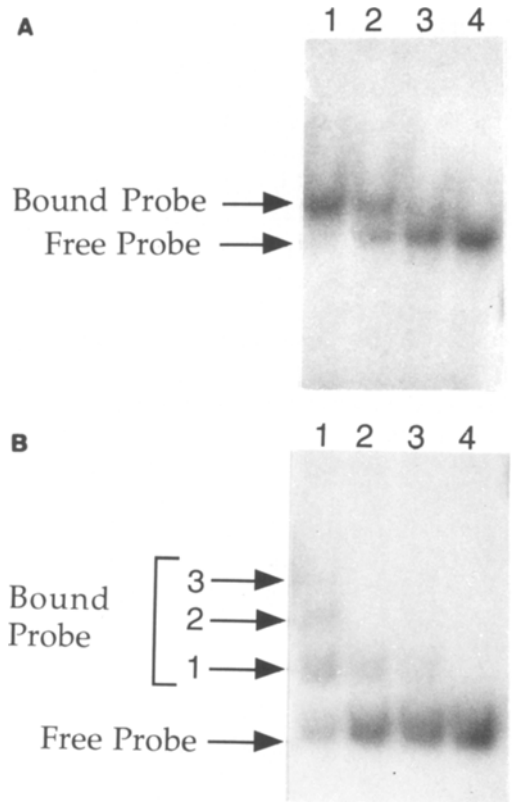

Fig. 5. (a) An electrophoretic mobility shift assay for a 63-mer probe containing only one Sp1 binding sequence. Twenty fmol of DNA probe was used in each lane. Lane 1, DNA probe + 200 fmol HSplase. Lane 2, DNA probe + 100 fmol HSplase. Lane 3, DNA probe + 50 fmol HSplase. Lane 4, DNA probe incubated in the absence of HSplase. (b) An electrophoretic mobility shift assay for a 90-mer containing three Sp1 sites. Lane 1, 20 fmol of DNA probe + 200 fmol HSplase. Lane 2, DNA probe + 100 fmol HSplase. Lane 3, DNA probe + 50 fmol HSplase. Lane 4, DNA probe in the absence of HSplase.

(5′-GA/GGGCGT/GGGC-3′). Thus the enzyme, if expressed *in vivo,* may cleave the LTR. One possible concern is that this enzyme could also cleave the host genomic DNA. This probably will not occur in this case because the host DNA is separated from the cytoplasm by a nuclear membrane. Protein traffic across this membrane is controlled by the presence of a nuclear localization signal. The nuclear localization signal for Sp1 is located N-terminal to the residues used in the construction of Splase. This would suggest that this enzyme is no longer able to transgress the nuclear membrane. Therefore, the only time the expression of this enzyme could pose a problem would be during cell division. The cell, however, has host nucleases which would create the same problem. Therefore, the cell most likely has mechanisms to deal with this problem. Experiments designed to examine the feasibility of expression of Splase in mammalian cells and the long-term survival of cells containing Splase are being pursued.

Our results also indicate that Splase digests DNA containing three or more Sp1 binding sites. Reactions involving DNA containing less than three Sp1 binding sites have not resulted in detectable product formation. However, the results of mobility shift experiments indicate that DNA fragments containing only one Sp1 binding site were bound by HSplase. A possible explanation is that the presence of three Sp1 binding sites within the DNA, or binding of three enzyme molecules to the DNA, may be required for the DNA to assume a conformation accessible to the nuclease site of HSplase. Although the requirement of three consecutive Sp1 sites may not be desirable if Splase is to be used as a molecular biology tool, it makes the enzyme more specific to the HIV DNA.

4.2. Generation of New Restriction Enzymes

Splase is a rare cutter restriction enzyme due to the large 10-bp recognition sequence conveyed by the sp1 zinc fingers. This could be very useful in molecular biology by generating large DNA fragments for genomic sequencing. The approach can be extended to other zinc fingers with different specificity. Zinc fingers have been shown to be modular proteins (Choo and Klug, 1994). A combinatorial library of peptides corresponding to the recognition helix of the zinc finger motif has been created. This library has shown that if specific residues are used in the −1, 3, and 6 positions, the specificity of the zinc finger could be changed. Using the combinatorial library as a guide, a restriction enzyme could be generated for a very diverse set of recognition sequences. Restriction enzymes with very large recognition sequences could be generated, making the possibility for endless numbers of new restriction enzymes.

4.3. Limitations

There are, however, imperfections in the properties of Splase. Due to the lack of corresponding studies with the natural restriction enzyme *Fok*I, it is not totally clear whether the limitations in the properties of Splase are also inherent properties of *Fok*I. Further studies on both *Fok*I and Splase are required to understand

the nature of the following limitations: (i) Under our experimental conditions Splase cleaves DNA containing three or more Sp1 binding sites. This makes the recognition sequence over 30 bp, much like the intron-encoded restriction endonucleases (Perlman and Butow, 1989). Although this property increases its specificity toward the HIV DNA as indicated above, it could limit the use of the enzyme as a molecular biology tool. (ii) As shown in Figs. 2 and 3, Splase does not completely digest the substrate into products. There are two possible explanations for this observation: the enzyme may form a protein–DNA complex which is very stable, and in the time frame of the experiment the protein may not dissociate, or the half-life of the enzyme may be shorter than the time it takes to completely digest the substrate. Further experiments will be designed to elucidate the true reason for this observation. (iii) Splase cleaves the two strands of DNA differently and cleaves at multiple sites. Nonspecific cleavage at the DNA termini can also be detected. The issue of specificity is further addressed below.

4.4. Cleavage Sites and Specificity

The cleavage pattern shown in Fig. 4 depicts the results observed by autoradiography. The cleavage pattern observed on the G-rich strand was significantly weaker than the cleavage observed on the C-rich strand. This may suggest that the enzyme-active site is optimized for cleavage of one strand over the other. On the other hand, the fact that the two strands are cleaved differently raises the question of whether Splase (and presumably *Fok*I also) possesses a single active site or two active sites. In any case, the mechanism of *Fok*I is likely to be unique and interesting and should be investigated in detail.

4.5. Comparison with Other Artificial Restriction Enzymes

Although several artificial enzymes with specific DNA cleavage activity have been reported, most of them are not direct gene products (Pei and Schultz, 1994; Pan *et al.*, 1994; Chen and Sigman, 1987, 1988; Chen *et al.*, 1993; Coray *et al.*, 1989*a,b*; Sutton *et al.*, 1993). The only exception, to our knowledge, is Ubx-F_N, which recognizes and digests DNA containing a single Ubx site. Ubx-F_N is able to recognize and cleave DNA with only one Ubx site. Ubx-F_N does, however, show nonspecific nuclease activity unless low Mg^{2+} concentrations and high DNA/enzyme molar ratio are used.

After the submission of this manuscript a similar work was published by Kim *et al.* (1996). This paper describes the construction of a similar zinc-finger restriction enzyme with a slightly different DNA sequence specificity.

ACKNOWLEDGMENTS

We are indebted to Dr. R. Tjian for pSp1-516C used as PCR template for the zinc-finger domain of Sp1, and to the NIH AIDS Research and Reference Reagent Program for pUC-BENN-CAT. The work was supported in part by NIG grants GM30327, GM41788, and GM43268.

REFERENCES

Chen, C., and Sigman, D. S. (1987). *Science* **237,** 1197–1201.
Chen, C., and Sigman, D. S. (1988). *J. Am. Chem. Soc.* **110,** 6570.
Chen, C., Gorin, M., and Sigman, D. S. (1993). *Proc. Natl. Acad. Sci. USA* **90,** 4206–4210.
Choo, Y., and Klug, A. (1994). *Proc. Natl. Acad. Sci. USA* **91,** 11163–11172.
Coray, D. R., Pei, D., and Schultz, P. G. (1989*a*). *Biochemistry* **28,** 8277–8286.
Coray, D. R., Pei, D., and Schultz, P. G. (1989*b*). *J. Am. Chem. Soc.* **111,** 8523–8525.
Deng, T., Noel, J. P., and Tsai, M.-D. (1990) *Gene* **93,** 229–234.
Gendelman, H. E., Phelps, W., Feigenbaum, L., Ostrove, J. M., Adachi, A., Howley, P. M., Khoury, G., Ginsberg, H. S., and Martin, M. A. (1986). *Proc. Natl. Acad. Sci. USA* **83,** 9759–9763.
Hochuli, E., Debeli, H., and Schacher, A. (1987). *J. Chromatogr.* **411,** 177–184.
Jones, K. A., Kadonaga, J. T., Luciw, P. A., and Tjian, R. (1986). *Science* **232,** 755–759.
Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. T. (1987). *Cell* **51,** 1079–1090.
Kim, Y.-G., and Chandrasegran, S. (1994). *Proc. Natl. Acad. Sci. USA* **91,** 883–887.
Kim, Y., Cha, J., and Chandrasegaran, S. (1996). *Proc. Natl. Acad. Sci. USA* **93,** 1156–1160.
Kriwacki, R. W., Schultz, S. C., Steitz, T. A., and Caradonna, J. P. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 9759–9763.
Li, L., and Chandrasegran, S. (1993). *Proc. Natl. Acad. Sci. USA* **90,** 2764–2768.
Li, L., Wu, L. P., and Chandrasegran, S. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 4275–4279.
Maxam, A. M., and Gilbert, W. (1980). *Meth. Enzymol.* **65,** 499–560.
Pan, Q. C., Landgraf, R., and Sigman, D. S. (1994). *Mol. Microbiol.* **12**(3), 335–342.
Pavletich, N., and Pabo, C. (1991). *Science* **252,** 809–817.
Pei, D., and Schultz, P. G. (1994). *J. Am. Chem. Soc.* **112,** 4579–4580.

Pelham, H., and Brown, D. (1980). *Proc. Natl. Acad. Sci. USA* **77,** 4170–4174.
Perlman, P. S., and Butow, R. A. (1989) *Science* **246,** 1106–1109.
Skowron, P., Kaczorowski, T., Tucholski, J., and Podhajska, A. (1993). *Gene* **125,** 1–10.
Sugisaki, H., and Kanazawa, S. (1981). *Gene* **16,** 73–78.
Sutton, C., Maxumder, A., Chen, C., and Sigman, D. S. (1993). *Biochemistry* **32,** 4225–4230.
Yonezawa, A., and Sugiura, Y. (1994). *Biochim. Biophys. Acta* **1219,** 369–379.