Photocross-linking of the Homing Endonuclease PI-SceI to Its Recognition Sequence

(Received for publication, September 16, 1998, and in revised form, January 15, 1999)

Vera Pingoud‡‡, Hubert Thole¶, Frauke Christ‡‡, Wolfgang Grindl‡, Wolfgang Wende‡, and Alfred Pingoud‡

From the ‡‡Institut für Biochemie, Fachbereich Biologie, Justus-Liebig-Universität, Heinrich-Buff-Ring 58, D-35392 Gießen and §Zentrum Kinderheilkunde, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, D-30623 Hannover, Germany

PI-SceI is an intein-encoded protein that belongs to the LAGLIDADG family of homing endonucleases. According to the crystal structure and mutational studies, this endonuclease consists of two domains, one responsible for protein splicing, the other for DNA cleavage, and both presumably for DNA binding. To define the DNA binding site of PI-SceI, photocross-linking was used to identify amino acid residues in contact with DNA. Sixty-three double-stranded oligodeoxynucleotides comprising the minimal recognition sequence and containing single 5-iodopyrimidine substitutions in almost all positions of the recognition sequence were synthesized and irradiated in the presence of PI-SceI with a helium/cadmium laser (325 nm). The best cross-linking yield (approximately 30%) was obtained with an oligodeoxynucleotide with a 5-iododeoxyuridine at position +9 in the bottom strand. The subsequent analysis showed that cross-linking had occurred with amino acid His-333, 6 amino acids after the second LAGLIDADG motif. With the H333A variant of PI-SceI or in the presence of excess unmodified oligodeoxynucleotide, no cross-linking was observed, indicating the specificity of the cross-linking reaction. Chemical modification of His residues in PI-SceI by diethylpyrocarbonate leads to a substantial reduction in the binding and cleavage activity of PI-SceI. This inactivation can be suppressed by substrate binding. This result further supports the finding that at least one His residue is in close contact to the DNA. Based on these and published results, conclusions are drawn regarding the DNA binding site of PI-SceI.

Homing endonucleases are a fascinating new class of enzymes that cleave DNA with very high specificity within an extended recognition site and, thereby, in vivo initiate a double strand break repair that may lead to the insertion of the sequence coding for the homing endonuclease into an allele that it lacks (for reviews, see Refs. 1 and 2). They have been found in prokaryotes and eukaryotes as well as in archaea bacteria and are encoded by introns or inteins (for review, see Ref. 3). The largest group is characterized by the presence of one or two copies of a conserved dodecapeptide sequence, the LAGLIDADG motif (4). PI-SceI, a homing endonuclease from yeast, belongs to this group and occurs as an intein within the vacuolar H\^+-ATPase, from which it is spliced in an autocatalytic reaction (5).

The mature protein recognizes an extraordinarily long DNA sequence of 35–45 bp,^1 bends the DNA, and cleaves the substrate to produce a 4-bp 3’ overhang (5, 6, 7). The molecular details of DNA recognition and cleavage by PI-SceI are unclear, in spite of the fact that the crystal structure of PI-SceI is known (8). According to the structure analysis and mutational studies (9), PI-SceI is composed of two domains, one responsible for protein splicing (domain I) and one for DNA cleavage (domain II), which are connected by two peptide segments. The structure of the elongated domain I consists almost entirely of \(\beta\)-sheets, whereas the compact domain II is an almost equal mixture of \(\alpha\)-helices and \(\beta\)-strands. Domain II is built up from two substructures that are related by local 2-fold symmetry about an axis between the two LAGLIDADG sequences. The domain II of PI-SceI is structurally very similar to the homodimeric homing endonuclease I-CreI, which contains one LAGLIDADG motif per subunit and lacks the protein splicing domain. As shown by the crystal structure analysis (10), the LAGLIDADG motifs in I-CreI form part of the dimer interface while simultaneously positioning one of the conserved Asp residues adjacent to the scissile phosphates. These residues may function to coordinate Mg\(^{2+}\) and thereby help to attack the DNA; substitution of these residues abolishes the endonuclease activity of I-CreI (11). Mutation of the analogous residues Asp-218 and Asp-326 in PI-SceI also destroys the nucleolytic activity of this enzyme (12), whereas substrate binding of the mutated PI-SceI is not affected, suggesting that these Asp residues are involved in catalysis. Similar results were obtained with I-SceII (13), I-DmoI (14), I-CeuI (15), I-PortI (14), and PI-TleI (16), all members of the LAGLIDADG family of homing endonucleases.

All homing endonucleases have long recognition sequences (15–45 bp) that can tolerate variation of the sequence, as shown for PI-SceI (6) and other homing endonucleases, e.g. I-CreI (17), I-DmoI (18), I-PortI (19), I-PpoI (17), and I-TecI (20). Footprinting studies with PI-SceI (6), I-DmoI (18), and I-TecII (21) and their substrates show that they are involved in both major and minor groove interactions. After cleavage, PI-SceI (6, ^1 The abbreviations used are: bp, base pair(s); CL, cross-link; DEPC, diethylpyrocarbonate; ds, double-stranded; 5-IdU, 5-iododeoxuridine; 5-IdC, 5-iododeoxycytidine; PAGE, polyacrylamide gel electrophoresis; TBE, Tris borate/EDTA.

* This work has been supported by grants from the Deutsche Forschungsgemeinschaft (PI 122/13-1), the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie, and the Fonds der Chemischen Industrie. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This paper is dedicated to Professor Dr. Günter Maass on the occasion of his 65th birthday.

¶ To whom correspondence should be addressed: Justus-Liebig-Universität Gießen, Institut für Biochemie, FB 15, Heinrich-Buff-Ring 58, D-35392 Gießen, Germany. Tel.: +49 641 99–35409; Fax: +49 641 99–35402; E-mail: vera.pingoud@chemie.bio.uni-giessen.de.

‡ Recipient of a scholarship of the Graduiertenförderung des Landes Hessen.
Photocross-linking of PI-SceI

7), I-SceI (22), F-SceII (23), I-TevI (24), and I-TevII (21) remain bound to one of the two cleavage products that may be required for the subsequent recombination event which completes the homing reaction. The genetically engineered domain DI of PI-SceI binds specifically and with similar affinity as full-length PI-SceI to DNA containing the PI-SceI recognition site as well as to one of the two cleavage products (9), demonstrating that domain I is not only involved in protein splicing but also in DNA binding. In contrast, the genetically engineered domain II of PI-SceI is not able to bind DNA with strong affinity, suggesting that domain I is responsible for a decisive part of the contacts between PI-SceI and its substrate (9). A similar two-domain structure with a catalytic domain and a DNA binding domain has been proposed for I-TevI (25).

The long recognition site of PI-SceI makes it an attractive model for studying the mechanism of DNA sequence recognition by proteins but makes it difficult to model the DNA into the structure of PI-SceI, in particular as it is known that both domains, which may be linked in a flexible manner, are involved in DNA binding (9, 26, 27). In the study presented here, we tried to identify residues of PI-SceI in close contact to the recognition site by a photocross-linking technique using 5-iodine-substituted pyrimidines (5-IdU and 5-IdC). We show that PI-SceI can be cross-linked via His-333 to a 5-Iododeoxyuridine (5-IdU) residue located in position +9 of the bottom strand, i.e., the right half of the PI-SceI recognition sequence. Substitution of the cross-linked amino acid His-333 by Ala results in a PI-SceI mutant that is not significantly impaired in its ability to bind and cleave DNA. However, no photocross-linking could be observed with the H333A mutant, demonstrating that the region around His-333 is in close contact with the DNA. With this information and knowing where the active site is located, it is possible to present a model that describes the approximate location of the DNA binding site in the structure of PI-SceI.

MATERIALS AND METHODS

Purification of Recombinant His6-tagged PI-SceI—His6-tagged PI-SceI was expressed in Escherichia coli and purified as described by Wenne et al. (7).

Synthesis of Iododeoxyuridine- and Iododeoxycytidine-substituted Oligodeoxynucleotides—Oligodeoxynucleotides were chemically synthesized by automated β-cyanoethylphosphoramidite DNA synthesis using 5-IdU-β-cyanoethylphosphoramidites (Glen Research) on a Cyclone plus DNA synthesizer (Millipore) or obtained by INTERACTIVA. To reduce possible deiodination of the 5-IdU and 5-IdC, the final deprotection step was carried out at ambient temperature for 24 h, as suggested by the manufacturer. We used either the minimal full-length recognition sequence (cf. Fig. 1) or the right-half cleavage product comprising the upper strand 5′-GGGAGAGAGTATAGAATGGCAGAGTC-T3′ (31-mer) and the lower strand 5′-GATCAGCTCTTCGCCCATTA-TTACCTCTTCTCGCGAC-3′ (39-mer). Some oligodeoxynucleotides were labeled at their 5′-terminus using T4 polynucleotide kinase and γ[32P]ATP.

Photocross-linking of ds Oligodeoxynucleotides Carrying a Single Iodopyrimidine Substitution and PI-SceI—For analytical scale photocross-linking, approximately 10 μmol PI-SceI was preincubated with 10 μmol radioactively labeled ds oligodeoxynucleotide monosubstituted with 5-IdU or 5-IdC at various positions of the top and bottom strand of either the full-length recognition site or the right-half cleavage product (cf. Fig. 1) in buffer P (10 mM Tris/HCl, pH 8.5, 100 mM KCl, 2.5 mM EDTA) for 30 min at ambient temperature in a volume of 50 μl. Photocross-linking was carried out with a 40 milliwatt helium/cadmium laser emitting at 325 nm (Laser 2000). The total irradiation time was usually 2 h; in kinetic experiments, 0–3 h. Samples of 2.5 μl were withdrawn before and after cross-linking and analyzed on a 15% (w/v) SDS-polyacrylamide gel. Gels were silver-stained and dried, and radioactive bands were visualized by autoradiography with intensifying screens or by using an imager. For preparative isolation of the cross-linked PI-SceI oligodeoxynucleotide complex (see below), the analytical scale was increased 10-fold.

Photocross-linking of ds Oligodeoxynucleotide T + 9 and Nicked PI-SceI—PI-SceI (20 μmol) was digested under limiting conditions in 50 mM Tris/HCl, pH 8.0 with trypsin at a substrate:protease ratio of 500:1 (w/w) at ambient temperature, similar to that described recently (27). After 2 h of incubation, the reaction was terminated by the addition of 5 mM phenylmethylsulfonyl fluoride. The cross-linking reaction with nicked PI-SceI (10 μmol) was performed in the presence of radioactively labeled ds oligodeoxynucleotide T + 9 (10 μmol right-half cleavage product) for 2 h in buffer P. 5-μl aliquots withdrawn before and after irradiation were analyzed by electrophoresis on a 15% (w/v) SDS-polyacrylamide gel with subsequent silver-staining and autoradiography. Sequencing of the cross-linked-T-terminus tryptic fragment was performed as described recently (27).

Purification of Recombinant His6-tagged PI-SceI—The cross-linked PI-SceI-oligodeoxynucleotide T + 9 (right-half cleavage product) complex was purified from unreacted PI-SceI by anion exchange chromatography on a Mono Q column (HR5/5, Amersham Pharmacia Biotech). After irradiation, the reaction mixture was incubated in the presence of 2 μl urea for 5 min at 60 °C and directly applied to the column. The elution buffers were used: Buffer A, 50 mM Tris/ HCl, pH 8.0 and B, 50 mM Tris/HCl, pH 8.0 with 1 M NaCl. The gradient applied was 0–80% B in 40 min. The flow-rate was 1.0 ml/min. The elution was monitored by measuring the absorbance at 260 nm. Fractions of 1 ml were collected, and aliquots were analyzed by electrophoresis on 15% (w/v) SDS-polyacrylamide gel.

Photocross-linking of the Cross-linked PI-SceI-Oligodeoxynucleotide T + 9 Complex—The cross-linked PI-SceI-oligodeoxynucleotide T + 9 (right-half cleavage product) complex was purified by preparative anion exchange chromatography on a Mono Q column. After irradiation, the reaction mixture was incubated in the presence of 2 μl urea for 5 min at 60 °C and directly applied to the column. The elution buffers were used: Buffer A, 50 mM Tris/HCl, pH 8.0, 1 M NaCl, and Buffer B, 1 M NaCl. The gradient applied was 0–80% B in 40 min. The flow-rate was 1.0 ml/min. Fractions of 1 ml were collected, and aliquots were analyzed by electrophoresis on a 15% (w/v) SDS-polyacrylamide gel.

Protease Digestions of the Cross-linked PI-SceI-Oligodeoxynucleotide T + 9 Complex—The purified cross-linked complex of the cross-linked PI-SceI-Oligodeoxynucleotide T + 9 was incubated in the presence of 2 μl urea for 5 min at 60 °C and directly applied to the column. The elution buffers were used: Buffer A, 50 mM Tris/HCl, pH 8.0, 1 M NaCl, and Buffer B, 1 M NaCl. The gradient applied was 0–80% B in 40 min. The flow-rate was 1.0 ml/min. Fractions of 1 ml were collected, and aliquots were analyzed by electrophoresis on a 15% (w/v) SDS-polyacrylamide gel.

Identification of the Cross-link Site in the Cross-linked PI-SceI-Oligodeoxynucleotide T + 9 Complex—5 mmol of PI-SceI were incubated with an equivalent amount of ds oligodeoxynucleotide T + 9 (right-half cleavage product) for 2 h at 325 nm. The DNA in the reaction mixture was subsequently radioactively labeled with γ[32P]ATP and T4 polynucleotide kinase. The buffer was adjusted to 50 mM Tris/HCl, pH 8.0, 1 mM CaCl2, and 40 μl/mg trypsin was added. The digestion was performed for 2 h at 37 °C. To test the progression of trypsin and chymotrypsin were added to give a final concentration of 2, 5, 10, 20, 40, and 200 μM. Proteinase K and subtilisin were added to give a final concentration of 2 μg/ml and 20 μg/ml. The digestions were performed for 16 h at 37 °C. The reactions were terminated by precipitation with 0.1 volume of 1 M sodium acetate, pH 6.8, and 2 volumes of ethanol. The samples were dissolved in 20 μl of 6 M urea, 0.025% (w/v) bromphenol blue, and 0.025% xylene cyanol and subjected to electrophoresis on a 12% (w/v) polyacrylamide gel containing 0.5% TBE and 2 μl urea. Radioactive bands were visualized by autoradiography.
a pulsed liquid phase sequenator Model 477A (Applied Biosystems) with a 120A on-line high performance liquid chromatography system according to Thole et al. (28). 50 pmol of the material were amenable to sequencing.

Site-directed Mutagenesis, Purification, and Characterization of the H333A Variant of PI-SceI—Site-directed mutagenesis of the PI-SceI gene was performed by a polymerase chain reaction-based technique (29) using the primers 5′-GATTTGATCTGATGAGGCGCGAT-3′ and 5′-AGACTCTCAATCTCTCCATGACGCGAC-3′ (X, 5′-IDU) and 5′-GATTTGATCTGATGAGGCGCGAT-3′ and 5′-GCTGATCCAGGTTCCTGCTG-3’ (Y) and 5′-GATTTGATCTGATGAGGCGCGAT-3′ and 5′-GCTGATCAGGTTCCTGCTG-3’ and the unmodified 311-bp fragment as template. In the absence of primers 5′-IDU, both DNA fragments were annealed with the 311-bp DNA by incubation at 95 °C for 10 min and, after cooling to 37 °C, ligated with T4 DNA ligase (AGS). The 32P-labeled ligation product was purified by electrophoresis on a 6% (w/v) polyacrylamide gel. The 311-bp substrate modified in position T + 9 of the recognition site was incubated with 50 nM PI-SceI and irradiated for 1.5 h at 325 nm as described above. Analytical gel shift experiments were performed in the absence and presence of the 56-bp competitor oligodeoxynucleotide F (7). Preparative gel shift experiments of the PI-SceI/511-bp T + 9 complex were carried out before and after photocross-linking. Bands corresponding to the upper and lower complex were excised, eluted in 50 mM Tris/HCl, pH 8.0, precipitated, and analyzed on a 6% (w/v) polyacrylamide gel containing 1% (w/v) SDS.

Modification of Histidine Residues in PI-SceI by Diethylpyrocarbonate (DEPC)—PI-SceI was dialyzed against a buffer consisting of 30 mM inorganic sodium phosphate, pH 6.4, 150 mM NaCl, and 20 mM dithiothreitol. PI-SceI (6 μg) in the absence and presence of equimolar amounts of a 62-bp substrate (oligo(deoxy)nucleotide G (7)) was treated with DEPC at final concentrations of 0.25, 0.5, 1, 2.5, 5, and 10 mM for 30 min at ambient temperature. To remove the ethoxyformyl residue from the histidine residues, an aliquot of the DEPC-treated mixture was acidified by NaH2PO4 to reach a pH value of 6.25. Hydroxylamine from the histidine residues, an aliquot of the DEPC-treated mixture was incubated for 16 h at 4 °C. Circular dichroism spectra of 17 monomeric PI-SceI before and after incubation with DEPC (3 mM) were measured in a buffer consisting of 30 mM inorganic sodium phosphate, pH 6.4, and 150 mM NaCl on a JASCO J-710 spectrophotometer at ambient temperature.

Electrophoretic Mobility Shift Assay—For electrophoretic mobility shift assays, PI-SceI was mixed with 10,000 cpm of 32P-labeled 311-mer polymerase chain reaction product containing the PI-SceI recognition sequence in a total volume of 10 μl of binding buffer (10 mM Tris/HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA, 0.05% (v/v) nonfat dry milk, 5% (v/v) glycerol, 10 mM dithiothreitol, 0.1 μg of poly(dI-dC)) (27). After electrophoresis, the gels were dried, and bands were visualized and quantified in an imager.

RESULTS

To obtain detailed topological information about specific contacts between PI-SceI and its DNA substrate, we performed photocross-linking experiments using either 5-IDU or 5-IdC monosubstituted ds (oligo)deoxynucleotides comprising the recognition sequence for PI-SceI (Fig. 1). Photocross-linking of halogenated pyrimidines has been used successfully to identify contacts in DNA- and RNA-protein complexes (30–34). 5-IDU is an almost perfect analogue of thymidine (35) and particularly useful for the study of DNA-binding proteins using oligodeoxynucleotides with a T → 5-IDU substitution in defined positions. It is a photoactivatable zero-length cross-linker that has the advantage that cross-linking to regions of a protein not involved in DNA binding is minimized and that irradiation with long wavelength UV light (325 nm) does not lead to the excitation of other nucleic acid and protein chromophores. Mechanistic studies of the 5-IDU chromophore relevant to its use in nucleoprotein photocross-linking have been performed by Norris et al. (36). The site-specific 5-IDU- or 5-IdC-mediated photocross-linking method does not depend on information regarding the structure of the protein or the structure of the protein-DNA complex. Efficient cross-linking requires the close proximity of the modified base and a reactive amino acid. Preferential targets are Phe, Tyr, Trp, His, and Met residues (35, 37). In addition, the cross-link yield depends on a suitable orientation of the reacting groups. With aromatic amino acid side chains as acceptor, the yield of the cross-linking reaction is significantly enhanced when a π–π stacking interaction between the base analog and an aromatic amino acid residue is possible (36).

Analytical Photocross-linking of PI-SceI-DNA Complexes—We used synthetic oligodeoxynucleotides comprising either the minimal full-length recognition sequence of 36 bp (Fig. 1) or containing the right-half cleavage product with varying yields. A cross-link in very good yields is observed with an oligodeoxynucleotide substituted in position T + 9 of the bottom strand, and two cross-links in good yield are detected with oligodeoxynucleotides modified in position G + 4 and A + 5 of the upper strand, whereas cross-link yields are found for several additional positions.

FIG. 1. Oligodeoxynucleotides used in the photocross-linking reactions. Double-stranded oligodeoxynucleotides comprising either the full-length PI-SceI recognition sequence (in capital letters) or the right-half cleavage product were monosubstituted with 5-IDU in the cases of T, A, or G or 5-IdC in the case of C. The positions (numbered as proposed by Gimble and Wang (6)) that were substituted are indicated by circles or arrows. Circles denote positions that did not give rise to a photocross-link, and arrows denote photoreactive substitutions that lead to a photocross-link product with varying yields. A cross-link in very good yields is observed with an oligodeoxynucleotide substituted in position T + 9 of the bottom strand, and two cross-links in good yield are detected with oligodeoxynucleotides modified in position G + 4 and A + 5 of the upper strand, whereas cross-link yields are found for several additional positions.
product with DNA substituted at the other positions as shown in Fig. 1 indicates that either a suitable acceptor amino acid is not available at these positions in sufficient proximity or the stereochemical requirements for cross-linking are not fulfilled. Only the cross-link to position +9 was investigated further, because the yield was high enough to ensure that after purification of the cross-linked complex and its proteolytic degradation, sufficient amounts of a peptide/oligonucleotide adduct would be available for sequencing.

Substitution of 5-IdU in position +9 of the bottom strand results in an apparently homogenous cross-linked species as judged by SDS-PAGE (Fig. 2). The time course of irradiation of the PI-SceI-oligonucleotide T + 9 complex is shown on a silver-stained gel (Fig. 2A) and on its autoradiogram (Fig. 2B). The yield of this photocross-linked complex was routinely about 20–30% after 2 h of irradiation and could not be further increased by longer irradiation. Omission of PI-SceI or incubation without irradiation failed to yield any cross-linked material, nor could unmodified DNA be cross-linked to PI-SceI to a significant extent.

Analytical Photocross-linking of Nicked PI-SceI-DNA Complexes—To find out whether the cross-link position is localized in the C- or N-terminal half of the endonuclease, we performed photocross-linking with nicked PI-SceI that was generated by limited tryptic digestion under native conditions (27), with the ds oligodeoxynucleotide substituted by 5-IdU in position +9 of the bottom strand. Trypsin cleaves PI-SceI preferentially after Arg-277 between β-strand 16 and α-helix 6, separating the two conserved LAGLIDADG motifs. The two resulting halves comprise residues 1–277 and 278–454, which remain associated under native conditions but dissociate upon treatment with SDS. DNA binding and cleavage experiments had shown that nicked PI-SceI binds to substrate DNA with essentially the same affinity as intact PI-SceI but is devoid of DNA cleavage activity, possibly because it is not capable of inducing the same strong distortion in its substrate DNA as uncleaved PI-SceI, i.e. a bend of 75° (27). Photocross-linking of native and nicked PI-SceI is compared in Fig. 3. As shown in lane 2, cross-linked PI-SceI migrates on an SDS-polyacrylamide gel corresponding to a molecular species that is by 20-kDa larger than uncross-linked PI-SceI. If the complex between nicked PI-SceI and ds oligodeoxynucleotide T + 9 is irradiated and the reaction mixture analyzed by SDS-PAGE, a cross-linked protein fragment that migrates as a 40-kDa species (lane 4) is observed. Based on these results, this species can only be interpreted as a covalent adduct between the oligodeoxynucleotide modified in position +9 of the bottom strand and the 20-kDa C-terminal half of PI-SceI. This was confirmed by sequencing; the analyzed sequence was NNLNTNEPLWDAIVG, which corresponds to the N terminus of the cross-linked C-terminal peptide comprising residues 278–454.

Preparative Photocross-linking of PI-SceI-DNA Complexes, Isolation, and Proteolytic Digestion of Cross-linked PI-SceI—After preparative cross-linking of PI-SceI and oligodeoxynucleotide T + 9 (right-half cleavage product), the resulting mixture was resolved using anion-exchange chromatography. The eluted fractions were analyzed by SDS-PAGE, and the covalently-linked protein-DNA complex was identified as the peak eluting at 500 mM NaCl immediately before the free DNA (Fig. 4). Peak fractions contained the PI-SceI-oligonucleotide T + 9 complex with >95% purity.

To find out which protease would be most suitable to obtain a small and defined fragment of CL PI-SceI, extensive proteolytic digestions were performed with aliquots of the radioactive labeled cross-linked PI-SceI-oligonucleotide T + 9 complex in the presence of increasing amounts of trypsin, chymotrypsin, proteinase K, and subtilisin at 37 °C overnight. The degree of digestion was analyzed by polyacrylamide gel electrophoresis in the presence of urea and visualized by autoradiography (Fig. 5). The treatment of the photocross-linked PI-SceI-oligonucleotide T + 9 complex with specific and unspecific proteases converted most of the cross-linked PI-SceI-DNA complex to small cross-linked peptides. The chymotryptic degradation was chosen for further analysis, because it produced an apparently defined end product that, because of the specificity of chymotrypsin, is likely to be homogenous. (We found out later that apparently homogenous end products obtained with protease K upon sequencing turned out not to have a defined N-terminal sequence).

Identification of the Cross-link Position in the Cross-linked Chymotryptic Peptide—To identify the amino acid residue of the homing endonuclease PI-SceI attached to the 5-IdU residue in position +9 of the bottom strand of ds oligodeoxynucleotide T + 9, a preparative cross-linking experiment with equimolar amounts of PI-SceI and ds oligodeoxynucleotide T + 9 was carried out. After irradiation and radioactive labeling, the re-
tocrss-linking is at residue His-333. This cross-link site resides in domain II, following the second LAGLIDADG motif, which is essential for catalysis.

Loss of Photocross-linking Activity in the H333A Mutant Protein—Based on the identification of His-333 as the amino acid covalently attached to the recognition site, the H333A variant of PI-SceI was produced. Both the binding and cleavage properties of the PI-SceI mutant were characterized and compared with the wild type enzyme. The H333A protein was shown in gel shift experiments to bind to a polymerase chain reaction-generated 311 bp substrate with the same affinity as wild type PI-SceI. In these experiments, two complexes can be observed, an upper complex and a lower complex with low and high mobility, respectively. The two complexes differ by the degree of DNA bending as shown for wild type PI-SceI before (6, 7). Compared with wild type PI-SceI, the lower complex is slightly less populated with the H333A mutant. Cleavage properties of the H333A, however, are the same as for the wild type PI-SceI, as shown independently by He et al. (26). As expected, H333A was inactive in producing a cross-link with the T + 9-substituted oligodeoxynucleotide (Fig. 7), indicating that His-333 is indeed responsible for the cross-link to ds oligodeoxynucleotide T + 9.

Absence of Effects of Cross-linking on Formation of Upper and Lower Complex—To find out whether cross-linking of DNA and PI-SceI interferes with the formation of the upper or lower complex observed in gel electrophoretic shift experiments with PI-SceI (6, 7), we generated a 311-bp substrate containing a 5-IdU substitution in position +9 of the bottom strand. This substrate was subjected to a photocross-linking reaction with PI-SceI. In an analytical gel shift experiment, it could be demonstrated that cross-linking had occurred, because a band shift is observed even in the presence of an excess of a specific oligodeoxynucleotide competitor added to the reaction mixture before loading the gel (data not shown). Under these conditions, the radioactively labeled 311-bp substrate T + 9 in the

**Fig. 5. Digestion of the purified photocross-linked PI-SceI-DNA complex with various proteases.** The isolated PI-SceI-DNA complex was digested in a volume of 50 μl of buffer (50 mM Tris/HCl, pH 8.0, 1 mM CaCl₂) either without or with increasing amounts of the indicated protease (trypsin, chymotrypsin, proteinase K (PK) or subtilisin (Sub)) for 16 h at 37 °C. After ethanol precipitation, the digests were analyzed by 2 x urea, 12% polyacrylamide gel electrophoresis in 0.5 x TBE followed by autoradiography. The lane designated ds oligo shows the positions of the two strands comprising ds oligodeoxynucleotide T + 9 (right-half cleavage product). The radioactivity in the two strands is different because of the lower efficiency of labeling of the recessed 5'-end of the upper strand. The lower strand is involved in the cross-link, whereas the upper strand is released upon addition of urea. Small amounts of the lower strand seen on the gel may be a contaminant carried over from the purification of CL PI-SceI. Fig. 4 shows that there is no base-line separation of CL PI-SceI and free oligodeoxynucleotide.

**Fig. 4. Purification of the cross-linked PI-SceI-DNA complex by anion-exchange chromatography.** 5 nmol PI-SceI were irradiated in the presence of ds oligodeoxynucleotide T + 9 (5 nmol, right-half cleavage product). The reaction mixture was applied to a Mono Q column that was eluted by an increasing NaCl gradient from 0 to 0.8 M. Individual fractions of 1 ml were collected, and 2.5-μl aliquots were analyzed on a 12% (w/v) SDS-polyacrylamide gel. Free, uncrss-linked PI-SceI elutes in fractions 4 and 5 as monitored at 280 nm in a parallel chromatogram. Lane S shows size markers, lane 1 shows the reaction mixture before, and lane 2 shows the reactions mixture after irradiation for 2 h at 325 nm. Lanes 3, 4, 5, and 6 show aliquots of the fractions 29, 30, 31, and 32 after chromatography. The gel was silver-stained. The positions on the gel corresponding to CL PI-SceI and PI-SceI are indicated.
Photocross-linking of PI-SceI

Modification of His Residues by Diethylpyrocarbonate—To test whether His residues of PI-SceI are involved in DNA binding and cleavage, a group-specific chemical modification of PI-SceI was performed using DEPC. The ethoxyformylation was carried out with an increasing amount of DEPC in a 2–80-fold excess over His residues. As shown in Fig. 8, the modification of His residues in the PI-SceI molecule has a dramatic effect on DNA binding. A 20-fold molar excess of DEPC over His residues completely abolishes DNA binding by the endonuclease. In the presence of equimolar amounts of DNA containing the recognition site, the modification of His residues is slowed down, indicating that these amino acids are protected from modification by the DNA and, therefore, might be involved in a specific protein-DNA interaction. The ethoxyformylation of His residues is in part reversible by treatment with hydroxylamine, as shown by restoration of the binding activity. This means that the DEPC treatment does not lead to an unspecific denaturation of the enzyme but is a specific effect of the ethoxyformylation of His residues. Experiments in which the cleavage activity of DEPC-modified and -demodified PI-SceI was measured produced nearly the same results, indicating that an impaired DNA binding correlates with a reduced DNA cleavage activity. The cleavage activity is restored by hydroxylamine as well. In fact, circular dichroism analysis demonstrates that in the presence of 3 mM DEPC (10-fold molar excess of DEPC over His residues), the secondary structure composition of PI-SceI is not significantly different from that in the absence of DEPC (Fig. 8B). This result suggests that at least one His residue must be involved in DNA binding and cleavage or located at the protein-DNA interface such that its chemical modification interferes with DNA binding and cleavage.

DISCUSSION

PI-SceI recognizes an extremely long asymmetrical sequence of more than 30 bp in length, as shown by a primer extension analysis (5), footprinting studies (12), and cleavage assays with substrates of different length (7). Specific binding is associated with strong bending into the major groove (6, 7). The center of bending is located at position +7, i.e. 5 and 9 nucleotides downstream of the sites of cleavage in the upper and lower strands, respectively (6, 7). Despite the fact that the crystal structure of PI-SceI is known (8), the size of the recognition site makes it difficult to imagine how the DNA is bound. The authors of the crystal structure analysis proposed a docking model based on several criteria.

(i) The scissile phosphodiester bonds must be close to the two Asp residues of the putative active site in domain II, Asp-218 and Asp-326.

(ii) The center of the cleavage site has to be positioned such that the left and right parts are in contact with the two symmetry-related β-sheets 7 and 9, respectively, of domain II.

(iii) The DNA backbone must be close to the positively charged residues following the LAGLIDADG helices, at the interface between domains I and II and in the extended region of domain I.

(iv) The bend of the DNA induced by specific complex formation has to be implemented into the docking model.

The published model fulfills these criteria. Its characteristic feature is that the DNA follows the concave contour of the gel and analyzed by SDS-PAGE. Cross-linked DNA was shown to be present in both complexes after irradiation (data not shown). This result suggests that His-333 is in close proximity to T + 9 in the bottom strand in both the upper and the lower complex or that the cross-linked PI-SceI/311-bp T + 9 adduct is in equilibrium between the two conformations, characteristic for the upper and the lower complex.

Photocross-linking of H333A to ds oligodeoxynucleotide T + 9. Wild type PI-SceI, 5 μM (lanes 1 and 2) or the H333A mutant, 5 μM (lanes 3 and 4) were incubated with 5 μM ds oligodeoxynucleotide (right-half cleavage product) containing 5-IdU in position +9 of the bottom strand for 30 min at ambient temperature. The mixtures were irradiated for 2 h. Aliquots before (lanes 1 and 3) and after irradiation (lanes 2 and 4) were analyzed by electrophoresis on a 12% (w/v) SDS-polyacrylamide gel that was silver-stained. Lane S shows size markers.

PI-SceI-DNA complex that had not been irradiated is replaced by unlabeled oligodeoxynucleotide F (7) and therefore does not show a band shift. In addition, a preparative gel shift experiment was performed with PI-SceI before and after cross-linking. Both, upper and lower complexes were extracted from the...
PI-SceI structure. According to this model, domain II interacts with about 14 bp, covering about 8 bp upstream and 6 bp downstream of the cleavage site. The additional 16 or more bp on the right side extend to domain I. It is conceivable that alternative docking models might satisfy the criteria given by Duan et al. (8), and it is clear that it would be very useful to know points of contacts between the protein and the DNA to test alternative models. Presumably, the most straightforward approach to determine such points of contact is to produce cross-links between PI-SceI and its substrate. We have chosen to use 5-iodopyrimidine-substituted DNA for our cross-linking study because 5-iodouracil and 5-iodocytosine are excellent chromophores to achieve photocross-linking of nucleoprotein complexes in high yield (35, 38, 31). In these studies up to 95% of the modified nucleic acid could be cross-linked. A recent example in which photocross-linking of 5-IdU-substituted double-stranded DNA to protein has been reported is for Thermus aquaticus MutS, which is cross-linked with a Phe residue close to the N terminus to a 5-IdU-substituted heteroduplex DNA (33).

To find out which positions when substituted by 5-iodopyrimidines in the recognition sequence of PI-SceI give rise to cross-links in good yield, we have synthesized 63 oligodeoxynucleotides monosubstituted with 5-IdU (for T, G, and A) and 5-IdC (for C) in almost all positions of the recognition sequence. These modified oligodeoxynucleotides were hybridized with their complementary unmodified sequence, incubated with PI-SceI in the absence of Mg$^{2+}$ (which is not required for specific binding), and irradiated. 12 of 32 positions tested in the upper strand and 3 of 31 positions tested in the lower strand proved to be reactive in photocross-linking, albeit with different yields (Fig. 1). The reactive positions thus identified must be in close contact to the protein in the PI-SceI-DNA complex and in the vicinity of a reactive amino acid residue; depending on the extent to which these conditions are fulfilled, different cross-linking yields are obtained. Our cross-linking results indicate that reactive positions are found preferentially in the upper strand with a cluster around the cleavage site that is in good agreement with the results of methylation interference experiments (6). The most reactive positions in the upper strand are G + 4 and A + 5, 2 and 3 bp downstream of the site of cleavage. The guanine residue was shown to be protected by PI-SceI in a methylation interference experiment (6). The lower strand contains only few reactive positions. The most reactive position of all positions tested in the lower and upper strand, however, was found in the lower strand at T + 9, 11 bp downstream of the cleavage site. This region contains a run of pyrimidine residues and, thus, was not informative in methylation experiments; it is, however, a region shown to be in close contact with the protein by hydroxyl radical footprint (12) and ethylation interference experiments (6). In addition, photocross-linking experiments in which the phosphate group linking C + 6/T + 7 and T + 7/T + 8, respectively, in the lower strand was substituted by a phosphorothioate group and coupled with p-azidophenacylbromide produced cross-links in good yield,2 confirming that this region is in close contact with the protein.

We have selected position T + 9 in the lower strand for the identification of a point of contact between PI-SceI and its DNA substrate, as this position proved to be the most reactive of all those tested. Furthermore, it is located approximately one helix turn away from the site of cleavage and, therefore, promised to give useful information regarding the location and orientation of the DNA substrate in that part of the DNA binding site, which is far away from the putative catalytic centers around Asp-218 and Asp-326.

Cross-linking experiments with nicked PI-SceI (27) and an oligodeoxynucleotide modified at position T + 9 demonstrated that cross-linking had occurred to the C-terminal half of PI-SceI (residues 278–454). Peptide sequencing of an isolated peptide oligodeoxynucleotide adduct from the chymotryptic digestion of the cross-linked PI-SceI-DNA complex showed that the cross-link involves His-333. In the PI-SceI structure, this residue, located between β-strands 19 and 20, protrudes from the protein surface into the cleft, which is generated by the junc-

---

2 F. Christ, unpublished information.
Our results demonstrate that this His-333 is in touch with the DNA in the upper and lower complex. If the scenario is correct that PI-SceI binds DNA first to form a complex (lower complex) in which the DNA is bent by 45° and then undergoes a conformational change to give a catalytically competent complex (upper complex) in which the DNA is bent by 75° (7), then the implication is that the region around His-333 remains in contact with DNA during the conformational change or does not interfere with it.

Conclusions—The main purpose of the cross-linking study presented here has been to find out how the DNA might be located with respect to the enzyme in the specific PI-SceI-DNA complex. Given the structural similarity between PI-SceI (8) and I-CreI (10), there is no reasonable doubt that the DNA with its scissile phosphodiester bonds is close to the Asp residues at the ends of the LAGLIDADG motifs. This is one region that interacts with the DNA. Another region has now been defined by a zero-length cross-link to be around His-333. If conformational changes are not excluded, involving for example loops connecting β-strands 21 and 22 as well as 15 and 16 and possibly also a subdomain movement to better position β-sheet 7, then this means that the DNA will be located in a groove defined by β-sheets 7 and 9 in domain II, will leave domain II at His-333, cross the interdomain cleft, and make contact with domain I via β-sheet 6, α-helix 1, and interconnecting loops. In contrast to the docking model proposed by Duan et al. (8) in which the bent DNA follows the concave contour of PI-SceI, in our model the DNA takes up a convex curvature (Fig. 9). The comprehensive mutational analysis carried out by He et al. (26) to support the docking model of Duan et al. (8) is also in good agreement with our model as are the hydroxyl radical footprint and ethylation interference experiments (6, 12). It must be pointed out that our model resembles with respect to domain II the docking model proposed for I-CreI (10), a homodimeric homing endonuclease with one LAGLIDADG motif per subunit. This similarity suggests that PI-SceI, a monomeric protein with a pseudosymmetrical catalytic domain containing two LAGLIDADG motifs like I-CreI, has two catalytic sites that cooperate in cleaving the two strands of the duplex substrate in one binding event (7). In this respect, PI-SceI presumably also resembles I-PpoI, the only homing endonuclease for which a co-crystal structure has been solved and for which it is understood how the DNA is bound in the enzyme-substrate and enzyme-product complex (41).

Acknowledgments—We thank Ms. Anja Wahl for excellent technical assistance and Drs. H. Sklenar and M. Zacharias for supplying the computer program CURVES.

REFERENCES
1. Belfort, M., and Roberts, R. J. (1997) Nucleic Acids Res. 25, 3379–3388
2. Cucuio, M. J., and Belfort, M. (1996) Cell 84, 9–12
3. Lambowitz, A. M., and Belfort, M. (1993) Annu. Rev. Biochem. 62, 567–622
4. Belfort, M., and Perlman, P. S. (1995) J. Biol. Chem. 270, 30237–30249
5. Grindl, W., and Thornier, J. (1993) J. Biol. Chem. 268, 21844–21853
6. Gimble, F. S., and Wang, J. (1996) J. Mol. Biol. 263, 163–180
7. Wende, W., Grindl, W., Christ, F., Pingoud, A., and Pingoud, V. (1996) Nucleic Acids Res. 24, 4123–4132
8. Duan, X., Gimble, F. S., and Quiocio, F. (1997) Cell 89, 555–564
9. Grindl, W., Wende, W., Pingoud, A., and Pingoud, V. (1998) Nucleic Acids Res. 26, 1857–1862
10. Heath, P. J., Stephens, K. M., Monnat, R., and Stoddard, B. L. (1997) Nat. Struct. Biol. 4, 468–476
11. Seligman, L. M., Stephens, K. M., Savage, J. H., and Monnat, R. J. (1997) Genetics 147, 1653–1664
12. Gimble, F. S., and Stephens, W. B. (1995) J. Biol. Chem. 270, 5849–5856
13. Henke, R. M., Butow, R. A., and Perlman, P. S. (1995) EMBO J. 14, 5094–5099
14. Lykke-Andersen, J., Barrett, R. A., and Kjems, J. (1997) EMBO J. 16, 3272–3281
15. Turnel, M., Otis, C., Cote, V., and Lemieux, C. (1997) Nucleic Acids Res. 25, 2610–2619
16. Hodges, R. A., Perler, B. F., Noren, C. J., and Jack, W. R. (1992) Nucleic Acids Res. 20, 6153–6157
17. Argast, M. G., Stephens, K. M., Emond, M. J., and Monnat, R. J. (1998) J. Mol. Biol. 280, 345–353
Photocross-linking of PI-SceI

10243

18. Ågaard, C., Awayez, M., and Garrett, R. (1997) Nucleic Acids Res. 25, 1523–1530
19. Lykke-Andersen, J., Thi-Ngoc, H. P., and Garrett, R. A. (1994) Nucleic Acids Res. 22, 4583–4590
20. Mueller, J. E., Smith, D., Bryk, M., and Belfort, M. (1995) EMBO J. 14, 5724–5735
21. Loizos, N., Silva, G. H., and Belfort, M. (1996) J. Mol. Biol. 255, 412–424
22. Perrin, A., Buckle, M., and Dujon, B. (1993) EMBO J. 12, 2939–2947
23. Jin, Y., Binkowski, G., Simon, L. D., and Norris, D. (1997) J. Biol. Chem. 272, 7352–7359
24. Mueller, J. E., Smith, D., and Belfort, M. (1996) Genes Dev. 10, 2158–2166
25. Derbyshire, V., Kowalski, J. C., Danzerau, J. T., Hauer, C. R., and Belfort, M. (1997) J. Mol. Biol. 265, 494–506
26. He, Z., Crist, M., Yen, H., Duan, X., Quiocho, F. A., and Gimble, F. S. (1998) J. Biol. Chem. 273, 4607–4615
27. Pingoud, V., Grindl, W., Wende, W., Thole, H., and Pingoud, A. (1998) Biochemistry 37, 8233–8243
28. Thole, H. H., Mascherl, I., and Jungblut, P. W. (1995) Eur. J. Biochem. 231, 510–516
29. Ito, W., Ishiguro, H., and Kurowsawa, Y. (1991) Gene 102, 67–70
30. Blatter, E. E., Elbright, Y. W., and Elbright, R. H. (1992) Nature 359, 650–652
31. Meisenheimer, K. M., Meisenheimer, P. L., Willis, M. C., and Koch, T. H. (1996) Nucleic Acids Res. 24, 981–982
32. Wang, Y., and Adzuma, K. (1996) Biochemistry 35, 3563–3571
33. Malkov, V. A., Biwas, I., Camerin-Otero, R. D., and Hoieh, P. (1997) J. Biol. Chem. 272, 23811–23817
34. Jenkins, T. M., Esposito, D., Engelman, A., and Craigie, R. (1997) EMBO J. 16, 6849–6859
35. Willis, M. C., Hicke, B. J., Uhlenbeck, O. C., Cech, T. R., and Koch, T. H. (1993) Science 262, 1255–1257
36. Norris, C. L., Meisenheimer, P. L., and Koch, T. H. (1996) J. Am. Chem. Soc. 118, 5796–5803
37. Wong, D. L., Pavlovich, J. G., and Reich, N. O. (1998) Nucleic Acids Res. 26, 645–649
38. Stump, W. T., and Hall, K. B. (1995) RNA (N.Y.) 1, 55–63
39. Dalgaard, J. Z., Klar, A. J., Moser, M. J., Holley, W. R., Chatterjee, A., and Mian, I. S. (1997) Nucleic Acids Res. 25, 4626–4638
40. Lykke-Andersen, J., Garrett, R. A., and Kjems, J. (1996) Nucleic Acids Res. 24, 3982–3989
41. Flick, K. E., Jurica, M. S., Monnat, R. J., and Stoddard, B. L. (1998) Nature 394, 96–101