Sau3AI, a Monomeric Type II Restriction Endonuclease That Dimerizes on the DNA and Thereby Induces DNA Loops*

Received for publication, February 23, 2001, and in revised form, March 23, 2001
Published, JBC Papers in Press, April 20, 2001, DOI 10.1074/jbc.M101694200

Peter Friedhoff†, Rudi Lurz§, Gerhild Lüder§, and Alfred Pingoud¶

From the †Institut für Biochemie, Justus-Liebig-Universität Giessen, D-35392 Giessen, Germany and the §Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73, D-14195 Berlin-Dahlem, Germany

Here, we report that Sau3AI, an unusually large type II restriction enzyme with sequence homology to the mismatch repair protein MutH, is a monomeric enzyme as shown by gel filtration and ultracentrifugation. Structural similarities in the N- and C-terminal halves of the protein suggest that Sau3AI is a pseudo-dimer, i.e. a polypeptide with two similar domains. Since Sau3AI displays a nonlinear dependence of cleavage activity on enzyme concentration and a strong preference for substrates with two recognition sites over those with only one, it is likely that the functionally active form of Sau3AI is a dimer of a pseudo-dimer. Indeed, electron microscopy studies demonstrate that two distant recognition sites are brought together through DNA looping induced by the simultaneous binding of two Sau3AI molecules to the DNA. We suggest that the dimeric form of Sau3AI supplies two DNA-binding sites, one that is associated with the catalytic center and one that serves as an effector site.

Orthodox type II restriction endonucleases are homodimeric enzymes that cleave their substrates within or directly adjacent to their recognition sequence (for reviews, see Refs. 1 and 2). It became increasingly apparent over the last years that several variations to this theme exist, namely in the form of the monomeric type II restriction endonucleases that dimerize on the DNA substrate via their catalytic domains, e.g. FokI (3, 4); the dimeric type II enzymes, e.g. EcoRII and NaeI, which require binding to a second recognition site for cleavage (5, 6); dimeric restriction endonucleases like SgrAI, which also require two recognition sites for activity and tetramerize on the DNA (7); and tetrameric type II restriction endonucleases like SfiI, Cfr10I, and NgoMIV (8–10), which interact with two recognition sites, loop out the DNA, and cleave the two sites in a concerted manner. Structural information is available for FokI (4, 11), NaeI (12), Cfr10I (9), and NgoMIV (10), in addition to structural information on the orthodox homodimeric type II restriction endonucleases BamHI, BglII, BsoBI, EcoRI, EcoRV, MunI, and PvuII (for review, see Ref. 13). The typical restriction endonuclease fold is also found in endonucleases involved in DNA repair (MutH (14) and Vsr endonuclease (15)), recombination (T7 endonuclease I (16) and λ-exonuclease (17)), and transposition (TnsA (18)).

It was interesting to note that the type II restriction endonuclease Sau3AI shares sequence homology with the DNA mismatch repair endonuclease MutH (14). Both Sau3AI and MutH recognize the same DNA sequence (GATC); but whereas MutH nicks the unmethylated strand of a hemimethylated or unmethylated GATC site, Sau3AI cleaves both strands of GATC sites regardless of the methylation status (19). MutH is a monomer, both in solution and in the crystal, which makes sense because it only nick the DNA (14, 20). In vivo, it shows only catalytic activity after mismatch-dependent activation by MutS and MutL (21), whereas in vitro, the requirement for MutS and MutL is less pronounced (22, 23). In contrast to MutH, however, Sau3AI is an active enzyme by itself. The quaternary structure of Sau3AI is not known, but it is likely that it is a functional dimer, as most type II restriction endonucleases that have to catalyze a double-strand cut. Sau3AI is an unusually large restriction endonuclease. It possesses an additional 270 residues C-terminal to the presumptive catalytic N-terminal domain which is homologous in sequence to MutH (14), resulting in its being twice the size of MutH. Therefore, the question arises concerning the function of this additional C-terminal domain. It has been speculated (14) that the role of these residues in Sau3AI may be the functional equivalent to that of MutL and MutS in the case of MutH, such that Sau3AI is permanently activated.

Since there is only a small amount of information regarding the biochemistry of Sau3AI, we therefore started to investigate some of the properties of Sau3AI, namely its quaternary structure and its mode of DNA cleavage. Here, we provide evidence that Sau3AI defines a new subtype of type II restriction endonucleases, being a monomer in solution that, to become active, has to dimerize in the presence of DNA, probably by binding to two recognition sites, which is accompanied by DNA looping. We conclude this from the following findings. First, we show that Sau3AI, unlike most other type II restriction endonucleases, is a monomer in solution. Second, the rate of DNA cleavage by Sau3AI is not linearly proportional to the protein concentration, suggesting a requirement for cooperative binding of Sau3AI to DNA to afford cleavage. Third, substrates containing more than one GATC site are cleaved by an order of magnitude more quickly than substrates with a single GATC site, suggesting either that a monomer of Sau3AI is interacting with two GATC sites or, more likely, that Sau3AI dimerizes on the DNA, in response to binding to two GATC sites, a conclusion that is strengthened by electron microscopy, which shows that Sau3AI induces loops on DNA with two GATC sites.

EXPERIMENTAL PROCEDURES

Oligodeoxynucleotides and Enzymes—Oligodeoxynucleotides were obtained from MWG Biotech AG (Ebersberg, Germany). Tag DNA polymerase was purchased from Promega (Mannheim, Germany). Sau3AI was obtained from Roche Molecular Biochemicals (Mannheim). Protein purity was checked by SDS-polyacrylamide gel electrophoresis and was >95%. Sau3AI concentration was determined by UV spectroscopy using...
the theoretical extinction coefficient at 280 nm of 85,600 M$^{-1}$ cm$^{-1}$ based on the amino acid sequence of Sau3AI (24).

**Fold Recognition and Sequence Alignment—**Protein sequences were obtained from the NCBI Entrez server. In a search for structural homologs of the C-terminal domains of Sau3AI, LlaKR2I, and Sth368I, we used fold recognition methods based on sequence-derived predictions. For this purpose, we employed the metaserver Pcons (25), which uses severalfold recognition programs, namely the Fold and Function Assignment System FFAS (26), the 3D-PSSM server (27), Gen-THREADER (28), FUGUE, INBGU (29), Sam-T99 (30), and Protein Data Bank BLAST, in addition to the secondary structure prediction servers PSIPred (31) and Jpred (32). Protein sequence alignments of the N-terminal domains of Sau3AI, LlaKR2I, and Sth368I with MutH protein sequences and alignment of the C-terminal domains of Sau3AI, LlaKR2I, and Sth368I were performed using the ClustalW program (33, 34). The alignment of the C-terminal domain of LlaKR2I with Escherichia coli MutH is based on the results of the fold prediction server Pcons. The alignments were combined and analyzed further using the program GeneDoc (35).

**Sucrose Density Gradient Centrifugation—**Sau3AI (100 μl, 400 nM) was analyzed on a 4.4-ml linear sucrose density gradient from 8 to 38% (w/v) sucrose in 20 mM HEPES-KOH, pH 8.0, 50 mM KCl, and 0.1 mM EDTA at 4 °C in an SW 60Ti rotor using a Beckman L-60 ultracentrifuge at 60,000 × g for 16 h. Samples of 200 μl were fractionated from the bottom of the tube using an Amersham Pharmacia Biotech F10 pump and analyzed for restriction endonuclease activity using pUC8 as a substrate (see below). The following marker proteins were used: β-amylase (Mᵦ = 200,000, sₑₒₜ = 8.9 S), alcohol dehydrogenase (Mᵦ = 155,000, sₑₒₜ = 7.9), bovine serum albumin (Mᵦ = 66,000, sₑₒₜ = 4.3 S), ovalbumin (Mᵦ = 42,000, sₑₒₜ = 3.5), carbonic anhydrase (Mᵦ = 29,000, sₑₒₜ = 2.8), and cytochrome c (Mᵦ = 12,000, sₑₒₜ = 2.1).

**Gel Filtration—**Gel filtration was performed with Sau3AI (100 μl, 400 nM) on a Superdex 200™ column (Amersham Pharmacia Biotech) equilibrated with 10 mM Tris-HCl, pH 7.5, 10 mM CaCl₂, 150 mM NaCl, and 5% (v/v) glycerol at 22 °C with a flow rate of 1 ml/min using a Merck-Hitachi Model 6200A high pressure liquid chromatography apparatus with a Model 655 photometer (set at 280 nm) and a Model D-2500 Chromato-Integrator. The same standard proteins were used as described for sucrose density gradient centrifugation.

**Plasmid DNA Cleavage Assay—**The concentration of pUC8 plasmid DNA isolated from the dam-negative strain JM110 was determined spectrophotometrically at 260 nm with an extinction coefficient of 3.6 × 10$^4$ M$^{-1}$ cm$^{-1}$. Cleavage reactions were performed at 37 °C with 20 μg/ml pUC8 (11 μm) in 33 mM Tris-HCl, pH 7.9, 10 mM magnesium acetate, 96 mM potassium acetate, 0.5 mM dithiothreitol, 0.05 mM/mg bovine serum albumin, and Sau3AI at the concentrations indicated. Cleavage reactions were stopped by removing 10-μl aliquots from the reaction mixture and adding 0.2 volume of gel loading buffer (250 mM EDTA, pH 8.0, 25% [v/v] sucrose, 0.1% [w/v] bromophenol blue, and 0.1% [w/v] xylene cyanol). The reaction products were analyzed by 12% agarose gel electrophoresis. The ethidium bromide-stained gels were analyzed with a video documentation system (INTAS, Göttingen, Germany). The intensities of the DNA bands were quantified using the program TINA (Version 2.07d). Initial rates were calculated from the disappearance of the supercoiled form of the plasmid and the appearance of the linear form of the plasmid, respectively.

**PCR1 Product Cleavage Assay—**Substrates containing one or two Sau3AI recognition sequences were produced by PCR using Taq DNA polymerase. Either plasmid pET15b-XhoI (a variant of pET15b in which three GATC sites at positions 491, 497, and 501 were mutated and a new XhoI site at position 491 was introduced) or pHisPl-SceI-N (36) was used as a template to obtain PCR products of varying length and site distance using the primers indicated in Table I. PCR purification was performed with the PCR purification kit from QiAGEN Inc. (Hilden, Germany). Each substrate (548 and 529 bp in length, respectively) at a concentration of 20 nM was incubated for 10 min at 37 °C with the indicated concentrations of Sau3AI in the same buffer as used in the plasmid DNA cleavage assay. At defined time intervals, samples were removed from the reaction mixture. The reaction was stopped by addition of 0.2 volume of gel loading buffer, and the reaction products were analyzed on 6% polyacrylamide gels. After staining with ethidium bromide, the gels were analyzed as described above for the plasmid DNA cleavage assay.

**Transmission Electron Microscopy—**DNA fragments of 1204 or 906 bp in length containing one or two Sau3AI recognition sequences, respectively, were produced by PCR (Table I). The DNA fragments were purified by gel extraction using NucleoSpin® Extract 2-in-1 (Macherey Nagel, Düren, Germany). DNA and protein were incubated at 37 °C for 10 min in a 10-μl reaction volume containing 5 mM DNA and various concentrations of Sau3AI ranging from 0 to 100 nM (molar ratios of enzyme to DNA between 1:1 and 20:1) in 10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, and 2 mM CaCl₂. Complexes were fixed with 0.2% (v/v) glutaraldehyde for 10 min at 37 °C and, after 3-fold dilution in 10 mM triethanolamine chloride, pH 7.5, and 10 mM MgCl₂, adsorbed to freshly cleaved mica as described (37). Micrographs were taken using a Philips CM100 electron microscope at 100 kV and a BioCam CCD camera (Tietz Video and Image Processing Systems GmbH, Gauting, Germany). To determine contour length, measurements were carried out on projections of 35-mm negatives using a digitizer (LM4, Brühl, Nuremberg, Germany). Histograms were generated as described (38).

### RESULTS

**Sau3AI Is a Non-typical Type II Restriction Endonuclease—**An analysis of the molecular weight and quaternary structure of type II restriction endonucleases revealed that Sau3AI has an unusual high molecular weight for a typical type II restriction endonuclease. Based on its amino acid sequence, the Mᵦ of 56,468 for Sau3AI is well above the average molecular weight of an orthodox dimeric type II restriction enzyme (excluding type IIs restriction endonucleases), which is ~31,000 ± 6000 (mean ± S.D.). Of 152 restriction endonucleases (Rebase 01/25/2001) classified as type II, 143 have Mᵦ values below 45,000. Type IIs restriction endonucleases, a subclass of type II enzymes, have Mᵦ values ranging from 54,000 to 116,000, with a median of 52,000. For one type IIs restriction endonuclease, FokI (Mᵦ = 66,215), it has been shown that it is monomeric in solution, but has to dimerize on the DNA to form an active endonuclease (3). Other larger than average type II restriction endonucleases are type IId enzymes, e.g. EcoRHI, which is a homodimer (Mᵦ = 2 × 45,608) in solution and has two binding sites for DNA, one of which is only catalytically competent (39–42). Some other type II restriction endonucleases with an unusually high molecular weight have been shown to be monomeric (at least in the absence of substrate), e.g. CcrI with an Mᵦ of 65,300 (43) and BsuRI with an Mᵦ of 66,300 (44), for which only little biochemical information is available.

**Sau3AI Is a Monomer in Solution—**To examine the quaternary...
Sau3AI Induces DNA loops

FIG. 1. Ultracentrifugation and gel filtration analyses of Sau3AI. A, ultracentrifugation analysis of Sau3AI on a sucrose density gradient. The standard curve of \( S_{20, w} \) versus fraction number, shown in the inset, was derived from the peak positions of the marker proteins (see “Experimental Procedures”) after fractionating (200-µl aliquots) the sucrose gradient and analysis by SDS-polyacrylamide gel electrophoresis and BCA assay. In the case of Sau3AI, fractions were tested for pUC8 cleavage activity (see inset). The peak position of Sau3AI activity is indicated by a square. Numbers indicate the marker proteins as follows: 1, β-amylase; 2, bovine serum albumin; 4, ovalbumin; 5, carbonic anhydrase; 6, cytochrome c. B, gel filtration analysis of Sau3AI on Superdex 200. The standard curve of log \( M_r \) versus \( V_0 \), shown in the inset, was derived from the elution profiles of the protein standards (see “Experimental Procedures”), with \( V_0 \) corresponding to the peak elution volume of the protein and \( V_e \) representing the void volume of the column determined with blue dextran 2,000,000. The peak position of Sau3AI is indicated by a square. Numbers indicate the marker proteins as described for a and, in addition: 2, alcohol dehydrogenase.

The native structure of Sau3AI, we subjected Sau3AI to both gel filtration and sucrose density gradient centrifugation analyses. Both methods can be used to determine the native molecular weight of a protein. As shown in Fig. 1, Sau3AI sedimented between bovine serum albumin (\( M_r = 66,000 \)) and ovalbumin (\( M_r = 42,000 \)). On a Superdex 200 gel filtration column, Sau3AI again eluted between bovine serum albumin and ovalbumin. Our results therefore suggest that, in the absence of DNA, Sau3AI, whose predicted \( M_r \) is 56,468, is as a monomer in solution. This raises the question of how a monomeric enzyme can catalyze a concerted double-strand cut, which Sau3AI does (Fig. 2). Two possibilities must be considered. 1) Sau3AI has two active sites/monomer, or 2) Sau3AI dimerizes in the presence of DNA.

Sau3AI Is a Pseudo-dimer of Non-identical Domains—The N-terminal domain of Sau3AI has been reported to share sequence homology with MutH (14). Since the catalytically important residues of MutH are conserved in the N-terminal domain of Sau3AI (Fig. 3), it is likely that this domain of Sau3AI is responsible for DNA cleavage. The function of the C-terminal domain of Sau3AI is not known. In a search for a function of the C-terminal domain comprising ~270 amino acid residues, we looked for sequence homologs. The only other sequence homologs of the C-terminal domain of Sau3AI found are the C-terminal domains of LlaKR21 and Sth368I, putative type II restriction endonucleases that have sequence homology to both the N- and C-terminal domains of Sau3AI (Table II). PSI-BLAST searches (45) with the C-terminal domain of Sau3AI, LlaKR21, or Sth368I failed to identify significant sequence similarity to any other protein in the data base. Therefore, we searched for structural homologs of the C-terminal domains of Sau3AI, LlaKR21, and Sth368I using the metaserver Peona, which employs severalfold recognition programs (see “Experimental Procedures”). Remarkably, the top hit for all three C-terminal domains was MutH (Table III). The pattern of secondary structures predicted for the C-terminal domains of Sau3AI, LlaKR21, and Sth368I agreed with the experimentally determined structure of MutH (data not shown), further supporting the threading results. Moreover, inspection of the sequence alignment using the N- and C-terminal domains of Sau3AI, LlaKR21, and Sth368I revealed three C-terminal domains was MutH (Table III). The pattern of

FIG. 2. Time course of supercoiled plasmid DNA cleavage by Sau3AI. pUC8 DNA (11 nM) was incubated with 0.9 nM Sau3AI. Aliquots were withdrawn at the times indicated, and the reaction products were analyzed by agarose gel electrophoresis. The supercoiled (oc) substrate was cleaved to give the linear (li) DNA without the accumulation of the open circle (oc) form.

Sau3AI, LlaKR21, and Sth368I and all available MutH protein sequences revealed that almost all of the conserved residues are located in the structural core and/or around the active site of the crystal structure of MutH (14). This led us to the conclusion that Sau3AI can be considered to be a pseudo-dimer, i.e. a polypeptide with two structurally similar domains. However, since only some of the catalytically important active-site residues (which are all present in the presumptive active site of the N-terminal domain) can be found in the C-terminal domain of Sau3AI (or LlaKR21 and Sth368I), it is unlikely that the C-terminal domain contains a functional active site.

Sau3AI Exhibits Cooperativity in DNA Cleavage—There are only few examples of monomeric endonucleases that are capable of making a specific double-strand cut in one binding event, e.g. PI-SceI, which harbors two catalytic centers (46). On the other hand, there are now examples of type IIs (e.g. FokI) and II (e.g. SgrAI) restriction endonucleases known, which are monomeric (dimeric) in the absence of DNA, but dimerize (tetramerize) in the presence of DNA to achieve catalytic activity (3, 47). In these cases, a nonlinear dependence of catalytic activity on protein concentration has been reported. Therefore, the rate of DNA cleavage catalyzed by Sau3AI at various concentrations was determined to establish the relationship between the initial velocity of the reaction and the enzyme concentration. Supercoiled pUC8 plasmid DNA containing 15 Sau3AI recognition sites was used as a substrate. Sau3AI restriction endonuclease was added to the reaction mixture, and the extent of cleavage was measured by withdrawing aliquots at defined time intervals and analyzing the amount of supercoiled plasmid DNA remaining. Typical results of such cleavage assays are shown in Fig. 4A. Initial velocities (\( v_0 \)) were calculated for any given concentration of Sau3AI as described under “Experimental Procedures.” In the \( v_0 \) versus Sau3AI concentration plot, a nonlinear dependence was observed (Fig. 4B). The fact that, at low Sau3AI concentrations, the initial velocity of the reaction is not directly proportional to the enzyme concentration suggests that the Sau3AI-catalyzed reaction is higher than first order with respect to the concentration of Sau3AI. The nonlinear relationship of \( v_0 \) versus Sau3AI concentration is best explained by a cooperative binding of Sau3AI molecules to the DNA substrate. An alternative explanation for a sigmoidal dependence of the initial rate of DNA cleavage on the enzyme concentration could be that the enzyme is being inactivated at low concentrations. Although this cannot be excluded, we regard it as unlikely because the DNA
cleavage assay was carried out in the presence of 0.05 mg/ml bovine serum albumin. It must be emphasized that the intermediates with only one recognition site left (Fig. 4, I-1) were cleaved only very slowly, an observation typical for type IIe restriction enzymes, which require two sites for efficient DNA cleavage.

**Sau3AI Requires Two Sites for Optimal Activity** — The cooperative interaction of Sau3AI with DNA could be explained by different mechanisms. For instance, it could be that the active site is formed only by dimerization of the N-terminal domains of Sau3AI with DNA.
of two Sau3AI molecules, as observed for FokI endonuclease (3). It also could be that Sau3AI dimerizes in the presence of DNA and, to do this, requires that two recognition sites are occupied simultaneously. Since it is known that several restriction enzymes, namely the type Ile (e.g. EcoRII and NaeI) and IIf (e.g. SfiI, Cfr10I, and NgoMIV) enzymes, require more than one recognition site for optimal cleavage activity, we addressed this possibility for Sau3AI. For this purpose, PCR substrates containing one or two Sau3AI recognition sites (Table I) were incubated in the presence of varying concentrations of Sau3AI. As shown in Fig. 5, the substrate with two sites was cleaved up to five times faster (depending on the Sau3AI concentration) than the corresponding substrate with only one site. This is not a consequence of flanking sequence preferences, as site A in the two-site substrate corresponds to site A in the one-site substrate. The cleavage reaction for the substrate with two sites shows an accumulation of intermediates resulting from cleavage at only one site, suggesting that, in common with type Ile enzymes, but in contrast to type IIf enzymes, the cleavage of the two sites does not happen in a concerted manner.

**Sau3AI Induces Loops on DNA Containing Two Recognition Sequences**—The results of the cleavage experiments with substrates containing one or two sites demonstrate that efficient cleavage requires the cooperation of two Sau3AI molecules bound to the two sites. To demonstrate that Sau3AI interacts simultaneously with the two recognition sites of a DNA containing two such sites at a distance, an electron microscopic analysis of the binding of Sau3AI to DNA containing either one or two recognition sites was carried out. A PCR product (908 bp) harboring two Sau3AI recognition sites at nucleotides 447 and 719 (Table I) was used in this experiment. A 1204-bp PCR product with a single Sau3AI recognition site at nucleotide 302 (Table I) served as a control. The Sau3AI complexes with DNA were prepared in the presence of Ca2+, which supports formation of a specific complex, but does not allow for cleavage (data not shown), similarly as described for EcoRV (48). Electron micrographs of the Sau3AI-DNA complexes indeed demonstrated the formation of a specific complex at the recognition site (Fig. 6). The distribution of the lengths of the “arms” emerging from the protein-DNA complex shows two distinct peaks corresponding to the position of the Sau3AI recognition sites on the DNA substrate. Moreover, in the case of the substrate containing two Sau3AI recognition sites, frequent formation of loops was observed. Measurement of the loop size and the length of the arms emerging from the loop showed a close correlation to the positions of the Sau3AI recognition sites, indicating that the loops are formed only via Sau3AI bound to its recognition sites. Thus, electron microscopy experiments demonstrate that Sau3AI can bind to two distant recognition sites present on the same DNA molecule simultaneously and thereby induces a loop. This finding is consistent with the assembly of a Sau3AI dimer on the DNA when two DNA-binding sites are present on the DNA substrate.

**DISCUSSION**

The family of type II restriction endonucleases can be subdivided into several subfamilies according to various criteria (13). Based on available sequence, biochemical, and biophysical information, the by far largest group comprises the orthodox type II restriction endonucleases, which are homodimeric enzymes, but in contrast to type IIf enzymes, the cleavage of the two sites does not happen in a concerted manner.

**Sau3AI Induces DNA loops by Sau3AI.** A, time course of pUC8 cleavage at three different concentrations of Sau3AI. pUC8 (11 nM) was cleaved with 6, 3, or 1.5 nM Sau3AI in 33 mm Tris acetate, 10 mM magnesium acetate, 66 mM potassium acetate, 0.5 mM dithiothreitol, and 0.05 mg/ml bovine serum albumin, pH 7.9. After the indicated times, 10-μl aliquots were withdrawn from the reaction mixture, and the reaction was terminated by adding 2 μl of stop solution. The cleavage products were analyzed by 1% agarose gel electrophoresis. The two largest final cleavage products (P) are 932 and 585 bp in length. The largest intermediates with only one Sau3AI site left (I-1) are 1075 and 728 bp in length. Note that the intermediates were cleaved only very slowly. B, dependence of the initial rate of DNA cleavage on Sau3AI concentrations. The amount of linearized plasmid DNA was determined as described under “Experimental Procedures” from the cleavage pattern shown. sc, supercoiled substrate; li, linear DNA; oc, open circle DNA.
IIe enzymes; they also interact with two copies of their recognition sequence, which they cleave (in contrast to the type IIe enzymes) in a concerted manner (for review, see Ref. 50), e.g. SfiI (8), Cfr10I (9), and NgoMIV (10). Here, we describe a unique restriction enzyme that has features in common with the type IIe and IIf enzymes, namely Sau3AI. It deserves special attention because it is a sequence homolog of the monomeric mismatch repair endonuclease MutH, which, when triggered by mismatch recognition by MutS and stimulated by MutL, nicks DNA in the unmethylated strand at a hemimethylated GATC site (for review, see Refs. 51 and 52). Like MutH, Sau3AI recognizes GATC sites; but in contrast to MutH, Sau3AI cleaves both strands of the duplex. Sau3AI attracted our attention not only for its similarity to MutH, but also because of its unusually high molecular weight. Based on our sequence analysis, Sau3AI is a two-domain protein whose N-terminal domain displays sequence homology to MutH (as noticed before (14)) and presumably harbors the active center, characterized by the (P)D . . . (D/E)K motif (see below). Here we show, based on fold recognition analysis, that the C-terminal domain may adopt a similar fold as the N-terminal domain. Thus, Sau3AI can be regarded as a pseudodimeric protein, i.e. a polypeptide with two similar domains. Only one of these domains, the N-terminal domain, which shows the highest sequence similarity to MutH (Table II), seems to have a functional catalytic center with a bona fide (P)D . . . (D/E)XXK motif, typical for many type II restriction endonucleases (53–55), whereas the other domain of Sau3AI has only a cryptic catalytic center, in which the second carboxylate (D/E) is substituted by Val. The function of the C-terminal domain is not known; it might be involved in DNA binding and/or in protein dimerization. Our results discussed below suggest that the C-terminal domain is involved in DNA binding, which does not exclude an involvement also in dimerization.

Both gel filtration and ultracentrifugation analyses demonstrate that Sau3AI is a monomeric enzyme in solution, which raises the question of how a monomeric enzyme can catalyze a

**Fig. 5.** Comparison of the cleavage of PCR fragments containing one or two Sau3AI recognition sites. Each substrate (20 nM) was incubated for 10 min at 37 °C with the indicated concentration of Sau3AI, and the concentration dependence of the cleavage reaction was analyzed by gel electrophoresis. A, two-site substrate with cleavage sites at positions 85 and 357. Note that the two sites, A and B, were cleaved at almost the same rate. B, one-site substrate with a cleavage site at position 255. Site A in the one-site substrate has the same flanking sequence context as site A in the two-site substrate. Fragments produced upon cleavage of the two- and one-site substrates are indicated by their length in base pairs to the right of each gel. C, a product versus concentration plot. D, reaction with the two-site substrate; E, reaction with the one-site substrate.
double-strand cut. There is a precedent for this: the homing endonuclease PI-ScI is a monomeric enzyme (56) that, however, cleaves the two strands of its extended recognition sequence in a concerted manner (36). It was shown that this enzyme harbors two catalytic centers (46). Given the structural similarities of the N- and C-terminal domains of Sau3AI (see above), this possibility must be considered here. However, the absence of a functional catalytic motif in the C-terminal domain (based on sequence considerations) argues against this possibility. Another alternative could be that Sau3AI dimerizes on the DNA and only then is able to cleave the DNA.

The dependence of the cleavage rate on enzyme concentration is nonlinear and strongly suggestive of a cooperative interaction between Sau3AI monomers in the presence of a DNA substrate. In this respect, Sau3AI resembles the type IIs restriction enzyme FokI, which is a monomer in solution and has to dimerize on the DNA for cleavage (3). That Sau3AI requires the interaction with two recognition sites for effective DNA cleavage is demonstrated by the strong preference of Sau3AI for DNA substrates with two sites over those with only one site. This is typical for type Ile enzymes, like EcoRII (5) and NaeI (6), but has been observed also for the type II enzymes Cfr10I (9), NgoMIV (10), and SfiI and SgrAI (47) as well as the type IIs enzyme FokI. Like type Ile enzymes, but different from type II enzymes, Sau3AI does not cleave two-site substrates in a concerted manner. Rather, it cleaves the first site of a two-site substrate quickly and the remaining site as slowly as a one-site substrate. This can be interpreted to mean that, as proposed for the type Ile enzyme mechanism, the second site functions as an effector site, whose occupation by the enzyme is required for efficient cleavage to occur. This stimulation is optimal when the effector-binding site is on the same molecule as the site to be cleaved (57). Sau3AI shares the property of being inactive as a monomer with the type IIs enzyme FokI, which is converted to an active enzyme by dimerization in the presence of a DNA substrate.

Based on our results, we suggest that Sau3AI is best classified as a variant of the type Ile subfamily of restriction enzymes that differs from the more typical members of this subfamily, like EcoRII and NaeI, by a monomer-dimer equilibrium that is shifted toward the monomer state in the absence of DNA. Otherwise, it behaves like a type Ile enzyme in requiring an effector site and inducing loops on DNA with more than one Sau3AI recognition site. We propose (but it requires further experiments to prove the detail of this conjecture) that a Sau3AI dimer has two DNA-binding sites, the catalytic binding site formed by the N-terminal domains of two subunits and the effector-binding site formed by the C-terminal domains. The structural similarities of the N- and C-terminal domains suggest that one evolved out of the other, presumably after a gene duplication event that converted an originally orthodox dimeric type II enzyme into a tetrameric type IIf enzyme. Mutations in the C-terminal domain that inactivated the catalytic center but rendered the DNA recognition function of the C-terminal domain largely unaffected turned Sau3AI into a type Ile enzyme. As a matter of fact, Sau3AI resembles the EcoRII V258SN variant, which (due to the amino acid substitution) is monomeric, but can dimerize in the presence of DNA, albeit at lower efficiency, and presumably therefore has only little activity (58). It is noteworthy that the type Ile enzyme NaeI is the outcome of a different evolutionary scenario: as shown by the crystal structure analysis, NaeI is composed of two very different domains, an N-terminal catalytic domain (which is also responsible for dimerization) with a typical restriction enzyme fold and a C-terminal domain with a strong structural resemblance to the catabolite activator protein, which functions as the effector domain (12). Most likely, this domain (the “endodomain”), which in similar form is present in type IA and II topoisomerases, was acquired by the “endodomain” through a gene fusion...
event. A similar evolutionary scenario has been discussed for the evolution of the type IIIs enzyme FokI, which is composed of a cleavage domain and a DNA recognition domain consisting of three catabolite activator protein-like subdomains (11), one of which, together with the cleavage domain, is involved in dimerization. It has been asked whether type IeII's enzymes need at least two sites for fast DNA cleavage. A possible answer is that this is one way of making interaction more accurate, by checking the substrate at two separate (identical) sites before proceeding with the catalysis (59).

As was pointed out above, MutH shares sequence homology with Sau3AI. It therefore appears to be likely that these two endonucleases, which recognize the same DNA sequence, are evolutionarily related and have a common ancestor, possibly an orthodox type II restriction enzyme, which had to be trans-evolutionarily related and have a common ancestor, possibly an endonuclease, which recognize the same DNA sequence, are evolutionarily related and have a common ancestor, possibly an orthodox type II restriction enzyme, which had to be trans-formed into a dimer of a pseudo-dimer to yield Sau3AI or into a monomer to become a constituent of the MutHLS system.

Acknowledgments—We are obligated to Dr. J. M. Bujnicki for pointing out to us that SaI36SI shares sequence homology with Sau3AI and for comments on the manuscript. We thank Drs. V. Pingoud and A. Jeltsch for helpful discussions and critical reading of the manuscript. The expert technical assistance of Ina Steindorf is gratefully acknowledged.

REFERENCES

1. Roberts, R. J., and Halford, S. E. (1993) in Nucleases (Linn, S. M., Lloyd, R. S., and Robette, R. J., eds) pp. 35–88, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
2. Pingoud, A., and Jeltsch, A. (1997) Eur. J. Biochem. 246, 1–22
3. Bitinaite, J., Wah, D. A., Aggarwal, A. K., and Schildkraut, I. (1996) Proc. Natl. Acad. Sci. U. S. A. 95, 10570–10575
4. Wah, D. A., Bitinaite, J., Schildkraut, I., and Aggarwal, A. K. (1996) Proc. Natl. Acad. Sci. U. S. A. 95, 10564–10569
5. Kruger, D. H., Barcak, G. J., Reuter, M., and Smith, H. O. (1988) Nucleic Acids Res. 16, 3997–4008
6. Conrad, M., and Topal, M. D. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 9707–9711
7. Bieleck, D. T., Daniels, L. E., Bath, A. J., and Halford, S. E. (1999) J. Biol. Chem. 274, 36379–36386
8. Westeitl, L. M., and Halford, S. E. (1998) J. Mol. Biol. 281, 433–444
9. Siksnys, V., Skirgaila, R., Sasnauskas, G., Urbanke, C., Cherny, D., Grazulis, S., and Huber, R. (1999) J. Mol. Biol. 291, 1105–1118
10. Deibert, M., Grazulis, S., Sasnauskas, G., Siksnys, V., and Huber, R. (2000) Nat. Struct. Biol. 7, 792–799
11. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1997) Nature 388, 97–109
12. Hui, Q., Colandene, J. D., Chen, Y., Luo, F., Zhao, Y., Topal, M. D., and Ke, H. (2000) EMBO J. 19, 3110–3118
13. Pingoud, A., and Jeltsch, A. (2001) Nucleic Acids Res., in press
14. Ban, C., and Yang, W. (1998) EMBO J. 17, 1528–1534
15. Tsutakawa, S. E., Jingami, H., and Morikawa, K. (1999) Cell 99, 615–623
16. Hadden, J. M., Convery, M. A., Declais, A. C., Lilley, D. M., and Phillips, S. E. (2001) Nat. Struct. Biol. 8, 62–67
17. Kovall, R., and Matthews, B. W. (1997) Science 277, 1824–1827
18. Hickman, A. B., Li, Y., Mathew, S. V., May, E. W., Craig, N. L., and Dyda, F. (2000) Mol. Cell 5, 1025–1034
19. Kessler, C., and Manta, V. (1996) Gene (Amst.) 92, 1–248
20. Welsh, K. M., Lu, A. L., Clark, S., and Modrich, P. (1987) J. Biol. Chem. 262, 15624–15629
21. Modrich, P. (1991) Annu. Rev. Genet. 25, 229–253
22. Ban, C., Junop, M., and Yang, W. (1999) Cell 97, 85–97
23. Hall, M. C., and Matson, S. W. (1999) J. Biol. Chem. 274, 1306–1312
24. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) Protein Sci. 4, 2411–2423
25. Bujnicki, J. M., Ellefson, A., Fischer, D., and Rychlewski, L. (2001) Protein Sci. 10, 352–361
26. Rychlewski, L., Jaroszewski, L., Li, W., and Godzik, A. (2000) Protein Sci. 9, 222–241
27. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) J. Mol. Biol. 299, 499–520
28. Jones, D. T. (1999) J. Mol. Biol. 287, 797–815
29. Fischer, D. (2000) Proc. Symp. Bioocomput. 119–130
30. Karplus, K., Barrett, C., Cline, M., Diekhans, M., Grate, L., and Hughey, R. (1999) Proteins Struct. Funct. Genet. 3, suppl. 121–125
31. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) Bioinformatics 16, 494–455
32. Cuff, J. A., and Barton, G. J. (2000) Proteins Struct. Funct. Genet. 40, 502–511
33. Lloyd, A. (1997) embnet.news 4, 2
34. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
35. Nicholas, K. B., Nicholas, H. B. J., and Deerfield, D. W. I. (1997) embnet.news 4, 14
36. Wende, W., Grindl, W., Christ, F., Pingoud, A., and Pingoud, V. (1996) Nucleic Acids Res. 24, 4123–4132
37. Spiess, R., and Lurz, R. (1998) Methods Microbiol. 20, 293–323
38. Weigel, C., Schmidt, A., Backert, S., Lurz, R., and Messer, W. (1997) EMBO J. 16, 6574–6583
39. Reuter, M., Kupper, D., Meisel, A., Schroeder, C., and Krueger, D. H. (1998) J. Biol. Chem. 273, 8284–8300
40. Kupper, D., Reuter, M., and Krueger, D. H. (1995) Gene (Amst.) 157, 97–98
41. Petruska, O. V., Karpova, E. A., Gromova, E. S., and Guschlbauer, W. (1994) Biochem. Biophys. Res. Commun. 198, 885–890
42. Gabbara, S., and Bhagwat, A. S. (1992) J. Biol. Chem. 267, 16823–16830
43. Syddall, R., and Stachow, C. (1985) Biochim. Biophys. Acta 825, 236–243
44. Bron, S., and Horz, W. (1980) Methods Enzymol. 65, 112–132
45. Ailschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
46. Christ, F., Schottelt, S., Wende, W., Steuer, S., Pingoud, A., and Pingoud, V. (1999) EMBO J. 18, 6908–6916
47. Bieleck, D. T., and Halford, S. E. (1999) Mol. Microbiol. 31, 1243–1254
48. Vijay, L. B., and Halford, S. E. (1995) Biochemistry 34, 1113–1119
49. Krug, D. H., Kupper, D., Meisel, A., Tierlich, M., Reuter, M., and Schroeder, C. (1995) Gene (Amst.) 157, 165
50. Halford, S. E., Bieleck, D. T., Stanford, N. P., Williams, S. A., Milsen, S. E., Gormley, N. A., Watson, M. A., Bath, A. J., Embleton, M. L., Gowers, D. M., Daniels, L. E., Parry, S. H., and Szeluckin, M. D. (1999) Biochem. Soc. Trans. 27, 696–699