Kinetic Analysis of the Coordinated Interaction of SgrAI Restriction Endonuclease with Different DNA Targets*

Received for publication, May 2, 2003, and in revised form, June 17, 2003 Published, JBC Papers in Press, July 8, 2003, DOI 10.1074/jbc.M304603200

Kamini Hingorani-Varma and Jurate Bitinaite‡

From New England Biolabs Inc., Beverly, Massachusetts 01915

SgrAI restriction endonuclease cooperatively interacts and cleaves two target sites that include both the canonical sites, CPuCCGGPyG, and the secondary sites, CPuCCGGPy(A/T/C). It has been observed that the cleaved canonical sites stimulate SgrAI cleavage at the secondary sites. Equilibrium binding studies show that SgrAI binds to its canonical sites with a high affinity ($K_a = 4.8 \times 10^{10}$ M$^{-1}$) and that it has a 15-fold lower affinity for the cleaved canonical sites and a 30-fold lower affinity for the secondary sites. Steady-state kinetic reveals substrate cooperativity for SgrAI cleavage on both canonical and secondary sites. The specificity of SgrAI for the secondary site CACCGGCT, as measured by $k_{cat}/K$ is about 500-fold lower than that for the canonical site CACCGGCG, but this difference is reduced to 10-fold in the presence of the cleaved canonical sites. The efficiency of canonical site cleavage also increases by 3-fold when the cleaved canonical sites are present in the reaction. Furthermore, the substrate cooperativity for SgrAI cleavage is abolished for both types of sites in the presence of cleaved canonical sites. These results indicate that target site cleavage occurs via a coordinated interaction of two SgrAI protein subunits, where the subunit bound to the cleaved site stimulates the cleavage of the uncut site bound by the other subunit. The free subunits of SgrAI have the flexibility to bind different target sites and, consequently, assemble into various catalytically active complexes, which differ in their catalytic efficiencies.

Restriction endonucleases serve as useful models to study the molecular basis of specificity in the interaction of proteins with DNA. The prototypical homodimeric Type IIP endonuclease recognizes a symmetrical palindromic sequence, 4–8 bp long and in the presence of Mg$^{2+}$ specifically cleaves within the recognition site (1, 2). These enzymes, including EcoRI, EcoRV, BamHI, and PvuII are highly sequence-specific with a very strong ability to discriminate between specific and nonspecific sites, both at the level of DNA binding (recognition) as well as DNA cleavage (catalysis) (3–10). For example, the equilibrium-association constant of EcoRV for specific DNA is $0.2–4 \times 10^{10}$ M$^{-1}$, whereas for nonspecific DNA it is $2–7 \times 10^{6}$ M$^{-1}$ (7). EcoRV binding to the star sites, which differ from the canonical site by one incorrect base pair is about $5–7 \times 10^{6}$ M$^{-1}$, i.e. this affinity is closer to nonspecific binding as compared with specific binding (6, 7). Similarly, EcoRI (3–5), BamHI (8), and PvuII (9) show comparable patterns of discrimination against the nonspecific or star sites. These enzymes also discriminate effectively at the level of catalysis as seen in their inability to cleave at nonspecific sites, and a very reduced capability to cleave at star sites usually observed under extreme reaction conditions (4–7, 10).

Lately more attention has been focused on restriction endonucleases that differ from the orthodox enzymes by their structural organization and mode of action. The Type IIS endonuclease FokI recognizes asymmetric sites and cleaves at fixed distances away from its recognition site. Although FokI binds DNA as a monomer (11, 12), two DNA:FokI complexes must interact for cleavage to occur (13, 14), implying a general mechanism by which the monomeric Type IIS restriction endonucleases cleave their DNA targets (15). A number of endonucleases have been found to interact with two copies of their recognition sequence before they can cleave DNA (16–21). For example, Type IIE endonucleases such as NaeI and EcoRII are homodimeric proteins that have two distinct DNA binding clefs (18–20) and thus are capable of binding two copies of their recognition sequence per protein dimer. However, only one of these recognition sequences is cleaved while the other serves as an allosteric effector (21–28). Studies with NaeI reveal that the recognition sequences flanked by G/C-rich regions are preferred by the activator site, whereas the catalytic site prefers A/T-rich regions (22, 25). In the case of EcoRII, it has been shown that some non-cleaveable or pre-digested recognition sequences can be bound at the activator site and stimulate the cleavage of refractory sites (27).

The Type IIF class of restriction endonucleases, which include SfiI, NgoMIV, Cfr10I, and SgrAI, also interact with two copies of their recognition sequence, but unlike NaeI and EcoRII, they cleave both sites simultaneously (29–32). Except for SgrAI, these endonucleases exist as tetramers of identical subunits and carry two active sites per tetramer. The tetrameric form seems to be the catalytically functional form, because a point mutation that converts the Cfr101I endonuclease into a dimer also results in almost complete loss of activity (31).

The SgrAI restriction endonuclease, which recognizes a degenerate octanucleotide sequence 5’-CPuCCGGPyG (33), also cleaves two copies of its recognition sequence simultaneously (32), placing it in the general category of SfiI-like endonucleases. However, in contrast to the SfiI-like enzymes, SgrAI was found to exist as a homodimer in a solution (34). Furthermore, the SgrAI endonuclease has been found to cleave effectively at a subset of less-specific sequences, 5’-CPuCCGGPy(A/T/C) and 5’-CPuCCGGGG, referred to as secondary sites (35). DNA termini generated by cleaving at the SgrAI canonical site are a pre-requisite for efficient cleavage at secondary sites. To gain a better understanding of the reaction pathway of SgrAI, we...
have performed a detailed kinetic analysis of SgrAI binding and cleavage at various recognition sites.

**EXPERIMENTAL PROCEDURES**

*Enzymes and Chemicals—* All enzymes, plasmid vectors, strains, and dNTPs were obtained from New England Biolabs unless stated otherwise. Labeled nucleotide triphosphates were purchased from PerkinElmer Life Sciences. Novex TBE gels were purchased from Invitrogen. Oligodeoxyribonucleotides were synthesized by the phosphoramidite method by the Organic Synthesis Division of New England Biolabs. SgrAI restriction endonuclease was purified from a recombinant Escherichia coli strain carrying the cloned sgrAIR gene (New England Biolabs). All concentrations of SgrAI given here refer to the dimeric form of the protein with two subunits of $M_r$, 37,831.

*Substrates for DNA Binding and Cleavage Analysis—* 32P-Labeled double-stranded (ds) 30-mer and 80-mer oligonucleotide duplexes containing the canonical sites (CGCCGGCG, CACCGGCCG, and CACCGGTG), secondary site (CGCCGGCT), and nonspecific site (TCTAGCAG) were used as substrates in the gel retardation experiments. These oligonucleotide duplexes were generated using adaptations of a method described previously (36). Two overlapping single-stranded (ss) oligonucleotides, 5 pmol each (shown below as underlined) were annealed to each other by heating at 85 °C for 5 min, followed by an incubation for 30 min at 37 °C and 30 min at room temperature. The annealed oligonucleotide duplexes were made fully double-stranded by filling-in with DNA polymerase I Large (Klenow) fragment in the presence of 6.6 pmol of [γ-32P]dATP and 0.33 nmol each of unlabeled dGTP, dCTP, and dTTP in a 30-μl reaction containing the EcoPol reaction buffer (10 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 1 mM DTT). Subsequent to this, 0.33 nmol of unlabeled dATP was added, and the reaction was incubated for an additional 30 min. Unincorporated label and cold dNTPs were removed using Quick Spin™ columns as per the manufacturer’s instructions (Roche Applied Science). Incorporation of the label was confirmed by 10% (w/v) polyacrylamide gel electrophoresis followed by autoradiography. Representative 30-mer and 80-mer ds oligonucleotide duplexes with the canonical CGCCGGCG site (indicated in boldface) are shown as follows:

$$5'$$-GTTGTTGTGCACGGCCGGCAGCACACACCA$$3'$$

$$3'$$-CAGCAGATGCTATCCGGCTGATTGCTGACGCC$$5'$$

The remaining 30-mer and 80-mer oligonucleotide duplexes were designed and generated in the same way as indicated above, except that they varied in the SgrAI recognition site: CACCGGCCG, CACCGTG, CGCCGGCT, or TCTAGCAG. The sequence of the oligonucleotide duplex flanked by the SgrAI-specific termini (indicated in boldface) is as follows:

$$5'$$-GTGTGTTGTCACGGCCGGCAGCACACACCA$$3'$$

$$3'$$-CAGCAGATGCTATCCGGCTGATTGCTGACGCC$$5'$$

This duplex (10 pmol/μl) was labeled using 10 units of T4 polynucleotide kinase, in the presence of 6.6 pmol of [γ-32P]ATP in polynucleotide kinase buffer (70 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 5 mM DTT), in a 30-μl reaction volume for 30 min at room temperature, followed by a 15-min incubation at 37 °C. Unincorporated label was removed using Quick Spin™ columns as per the manufacturer’s instructions (Roche Applied Science). Incorporation of the label was confirmed by 10% (w/v) polyacrylamide gel electrophoresis followed by autoradiography. Unlabeled oligonucleotide duplexes used in the competition assays were generated and purified essentially as above, except that unlabeled dNTPs were used.

For cleavage assays, 80-bp-long DNA fragments containing the canonical sites (CGCCGGCG, CACCGGCCG, or CACCGTG) and second-

![Image](http://example.com/image.png)

**DNA Binding by Gel Retardation Assay**—The gel retardation method (37) was used to determine the apparent association constant ($K_a$) of SgrAI to its canonical sites (CGCCGGCG, CACCGGCCG, or CACCGTG). 32P-Labeled ds oligonucleotides (0.5 μM final concentration) was incubated with 1 μl of a range of serially diluted SgrAI protein (0–80 μM final concentration) in a 10-μl reaction, containing binding buffer (20 mM Tris-acetate, pH 7.9, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, 10% glycerol) for 30 min at 25 °C. The binding...
reaction was immediately loaded onto a 10% polyacrylamide gel that had been pre-run at 110 V for 30 min in TBE buffer (45 mM Tris, pH 8.1, 45 mM boric acid, 2 mM EDTA). The protein-DNA complexes were separated from unbound DNA by gel electrophoresis at 110 V for 1 h at 25 °C. After performing electrophoresis, the gels were fixed in solution containing 70:3:7:20 (v/v) water/glycerol/acytic acid/methanol for 10 min at 25 °C, dried at 65 °C for 2 h, and exposed to X-Omat (Kodak) x-ray film. The autoradiograph was scanned, and the intensity of the bands was analyzed by Image software (version 1.61, National Institutes of Health). The data were plotted using KaleidaGraph (version 3.08) as \[ \frac{[S_{\text{free}}]}{[S_{\text{bound}}]} \] versus \[ [S_{\text{bound}}] \] (3, 38). Extrapolation of the plot yields an intercept on the abscissa equal to \( K_a \) (1 + \( [S_{\text{initial}}]/[S_{\text{bound}}] \)).

The values for \( K_a \) were obtained from three independent experiments and are shown as mean ± S.D.

### Table 1

| Canonical sequence | L2 | K<sub>n</sub>  
|-------------------|----|----------------|
| CGCCGGCG          | 80 | 4.4(±0.4) \( \times 10^{10} \) |
| GC GCCGGCG        | 30 | 4.0(±0.3) \( \times 10^{10} \) |
| CACCCGGC          | 80 | 5.7(±0.3) \( \times 10^{10} \) |
| GTGGCCGC          | 30 | 7.7(±0.9) \( \times 10^{10} \) |
| CACCGGTG          | 80 | 6.2(±0.7) \( \times 10^{10} \) |
| GTGGCCGAC         | 30 | 5.3(±0.3) \( \times 10^{10} \) |

*The degenerate positions of the canonical sequences occupied by different base pairs are marked in boldface.

Fig. 2. Binding of SgrAI to canonical site (A) and secondary site (B) assayed by the equilibrium competition method. Competition binding reactions were performed in the presence of 5 mM Ca<sup>2+</sup> and 10 μM of 32P-labeled 80-mer oligonucleotide duplex with the canonical CGCCGGCGCT site (\( S_{\text{initial}} \)), 20 μM of SgrAI and various concentrations of unlabeled competitor duplex with either canonical site CGCCGGCGC (A) or secondary site CGCCGGCT (B). In A, the inset contains the following lanes: lane 1, no SgrAI, no competitor; lane 2, no competitor; lane 3–6, 50, 100, 150, 200 μM of competitor; \( S_{\text{in}} \), bound substrate; \( S_{\text{f}} \), free substrate. In B, the inset contains the following lanes: lane 1, no SgrAI, no competitor; lane 2, no competitor; lane 3–6, 0.5, 1.25, 2.0, 2.75 nm of competitor; \( S_{\text{in}} \), bound substrate; \( S_{\text{f}} \), free substrate. The data were plotted as [competitor DNA] versus 1/[ES], where [ES] = \( [S_{\text{f}}] + 1/[S_{\text{in}}]/[S_{\text{bound}}] \). Extrapolation of the plot yields an intercept on the abscissa equal to \( K_a \).  

where \( v \) and \( V_{\text{max}} \) are the initial and maximal initial SgrAI reaction velocities, \([S]\) is the concentration of the substrate, \( K \) is the substrate dissociation constant, and \( n \) is the Hill coefficient. Equation 1 was rearranged to give Equation 2.

### Equation 1

\[
\frac{v}{V_{\text{max}}} = \frac{[S]}{K^n + [S]^n} 
\]

### Equation 2

\[
\log\left(\frac{v}{V_{\text{max}} - v}\right) = n \log[S] - n \log K 
\]
TABLE II
Equilibrium association constants of SgrAI binding to the oligonucleotide duplexes with canonical and secondary sequences and to duplexes with SgrAI-specific termini by competition gel retardation analysis in the presence of Ca\(^{2+}\)

The association constants for competitors (K\(_a\)) were determined as shown in Fig. 2. The values for K\(_a\) and K\(_d\), were obtained from three independent experiments and are shown as mean ± S.D.

| Recognition sequence | K\(_a\) \(\times 10^{10}\) | K\(_d\) \(\times 10^{10}\) | Difference\(^a\) |
|----------------------|----------------|----------------|----------------|
| Canonical site       | 4.0(±0.4) | 4.4(±0.2) | 1 |
| CGCGGGCG | GCGCGCGG |
| Secondary site\(^b\) | 1.4(±0.4) | 2.8(±0.3) | 30 |
| CACCGGCT | GTGGCCGA |
| Duplex with SgrAI-specific termini | 2.8(±0.3) | 2.8(±0.3) | 15 |
| ...CG   | ...GCGGGCC-5' |
| Non-specific site | 0.9(±0.1) | 1.4(±0.1) | 510 |
| TCTAGCAG | AGATCGTC |

\(^a\) The -fold difference refers to the difference in K\(_a\) values relative to the K\(_a\) of the canonical site.

\(^b\) The 8th position of the secondary sequence, occupied by a variable base pair, is marked in boldface. 

The effect of oligonucleotide duplexes flanked by SgrAI-specific or nonspecific termini on the reaction velocities at the canonical site, CACCGGGCG, and the secondary site, CACCGGCT, was tested. The sequence of the oligonucleotide duplex flanked by the SgrAI-specific termini is the same as shown above. The sequence of the oligonucleotide duplex flanked by the nonspecific termini was identical, except that it carried 5'-GCC as extensions. The cleavage reactions in the presence of the oligonucleotide duplex were performed as described above, except the oligonucleotide duplex was added at a 50-fold molar excess over the substrate with the canonical site and at a 4-fold molar excess over the substrate with the secondary site (35).

RESULTS
SgrAI Specifically Binds to Canonical, Secondary, and Cleaved Canonical Sites—The binding of SgrAI restriction endonuclease for all three canonical sites was determined using direct-binding gel retardation assays. Experiments were first conducted in the absence of the metal cofactor Mg\(^{2+}\) to prevent cleavage of the oligonucleotide duplexes, but no specific SgrAI-DNA complex was observed (data not shown). Studies with EcoRV and PvuII have shown that Ca\(^{2+}\) stabilizes site-specific DNA binding, while not supporting catalysis (7, 40, 41). Thus, SgrAI retardation experiments were carried out in the presence of either 2 or 5 mM Ca\(^{2+}\). In the presence of 5 mM Ca\(^{2+}\), SgrAI formed a complex with an oligonucleotide duplex containing the canonical site. Control experiments indicated that the observed complex was sequence-specific, because in the presence of 5 mM Ca\(^{2+}\) there was no specific complex formation with an oligonucleotide duplex lacking the canonical site at all tested concentrations of the SgrAI protein (data not shown).

Representative gel-retardation results of SgrAI binding to the CGCGGGCG site are shown in Fig. 1A. The equilibrium binding constants (K\(_a\)) on all three canonical sites were determined from the direct-binding measurements as shown in Fig. 1B. The calculated K\(_a\) values are listed in Table I. To determine any dependence of the length of the oligonucleotide duplex toward the affinity of SgrAI for the specific site, the oligonucleotides were also varied in length (30-mer versus 80-mer). For the given lengths of the oligonucleotide duplexes tested, the binding affinity of SgrAI endonuclease was found to be independent of their length. The obtained K\(_a\) values for all three canonical sites were similar (4.0–7.7 × 10\(^{10}\) M\(^{-1}\)). Gel retardation assays were also performed with duplexes containing the
The turnover number \( (k_{\text{cat}}) \), substrate dissociation constant \((K_s)\), and Hill coefficient \((n)\) were obtained from two independent determinations as shown in Fig. 3. The values are shown as mean ± range.

| Recognition sequence | \( k_{\text{cat}} \) \( s^{-1} \) | \( K_s \) \( M \) | \( k_{\text{cat}}/K_s \) \( s^{-1} M^{-1} \) | \( n \) |
|----------------------|--------------------------|-----------------|-------------------|---|
| Canonical CGCCGGCG | 4.2(±0.7) \times 10^{-3} | 8.0(±1.0) \times 10^{-10} | 5.2(±0.2) \times 10^{-6} | 2.4(±0.2) |
| Canonical GGCGCCGC | 6.8(±0.4) \times 10^{-2} | 6.7(±0.3) \times 10^{-9} | 10.4(±0.1) \times 10^{-6} | 3.0(±0.2) |
| Canonical CACCGGCCG | 4.0(±0.3) \times 10^{-2} | 4.0(±0.3) \times 10^{-9} | 10.0(±0.5) \times 10^{-6} | 3.1(±0.1) |
| Canonical GTGGCCGC | 3.0(±0.5) \times 10^{-4} | 9.6(±0.4) \times 10^{-9} | 3.0(±0.4) \times 10^{-4} | 3.2(±0.2) |
| Secondary GCGGGCT | 2.0(±0.3) \times 10^{-4} | 7.9(±0.3) \times 10^{-9} | 2.7(±0.5) \times 10^{-4} | 2.8(±0.3) |
| Secondary GCGGGCAG | 3.0(±0.5) \times 10^{-4} | 9.6(±0.4) \times 10^{-9} | 3.0(±0.4) \times 10^{-4} | 3.2(±0.2) |
| Secondary CACCGGCT | 4.0(±0.3) \times 10^{-2} | 4.0(±0.3) \times 10^{-9} | 10.0(±0.5) \times 10^{-6} | 3.1(±0.1) |
| Secondary GTGGCCGA | 6.8(±0.4) \times 10^{-2} | 6.7(±0.3) \times 10^{-9} | 10.4(±0.1) \times 10^{-6} | 3.0(±0.2) |

Secondary site (CGCCGGCT), nonspecific site (TCTAGCAG), and an oligonucleotide duplex flanked by the SgrAI-specific termini with ss extensions of 5’CCGG. However, SgrAI binding to these substrates was too weak to permit measurements of \( K_s \) values by the direct-equilibrium binding method.

Equilibrium-competition assays were then performed to quantify the affinity of SgrAI for the weaker binding sites. The apparent inhibition of SgrAI binding to its canonical site, CACCGGGCG, was measured in the presence of increasing concentrations of the respective competitor duplex, i.e., either in the presence of duplex with a secondary site (GCACGGCT), nonspecific site (TCTAGCAG), or duplex flanked by SgrAI-specific termini. The disappearance of the specific complex was monitored, and the data were plotted to determine the equilibrium-association values for competitor \( (K_{\text{in}}) \) (Fig. 2). As a control, the competitor duplex with the same canonical site as in the substrate duplex was used to determine the \( K_{\text{in}} \) value (Fig. 2A), thereby allowing the comparison of binding parameters obtained by the direct-binding and equilibrium-competition methods. The measured \( K_{\text{in}} \) and \( K_{\text{ib}} \) values (4.0 \times 10^{-10} M^{-1} and 4.4 \times 10^{-10} M^{-1}, respectively) are virtually identical indicating that equilibrium-competition may be used for weaker binding substrates.

Fig. 2B shows the results of a typical equilibrium-competition assay obtained with the secondary CGCCGGCT site as a competitor. The measured \( K_{\text{in}} \) values for the recognition sites tested are reported in Table II. As expected, the \( K_{\text{in}} \) value for the secondary site (1.4 \times 10^{-9} M^{-1}) is on average 30-fold lower than that for the canonical site (4.4 \times 10^{-10} M^{-1}). Nevertheless, the SgrAI affinity for this secondary site is about 15-fold higher than nonspecific binding, when compared with the \( K_{\text{in}} \) value for the nonspecific oligonucleotide duplex (0.9 \times 10^{-8} M^{-1}), which in turn is ~500-fold lower than that for the canonical site.

The \( K_{\text{ib}} \) value of 2.8 \times 10^{-8} M^{-1} for the oligonucleotide duplex flanked by the SgrAI-specific termini shows that SgrAI has high affinity for these duplexes, which is only about 15-fold lower compared with the uncut canonical site and is ~2-fold higher relative to that for the secondary site. The oligonucleotide duplexes carrying self-complementary ss extensions of 5’CCGG are most likely to anneal to each other, thus forming an SgrAI canonical site. However, the re-associated site differs from the uncut canonical site in two respects: it has two truncated phosphodiester bonds and it also lacks the two 5’ terminal phosphate groups, because non-phosphorylated oligonucleotides were used in this study. The missing 5’ phosphate groups are the ones that derive from the scissile phosphodiester bonds in the catalytic site of the SgrAI-DNA complex.

However, as revealed by the binding data, neither the truncated phosphodiester bonds nor the 5’ phosphate groups seem to be of a great importance for SgrAI binding. These results are in agreement with our previous work (35) that shows that the dephosphorylated canonical site cleavage products are capable of assisting in secondary site cleavage as those flanked with 5’ phosphates.

**DNA Cleavage by SgrAI Reveals Substrate Cooperativity for Canonical and Secondary Sites**—Kinetic parameters for SgrAI cleavage were determined under steady-state reaction conditions with Mg\(^{2+}\) as the activating metal. The initial velocity \( (v) \) was calculated for a given concentration of substrate by measuring the increase in the concentration of product as a function of time. The representative time-course profiles for the canonical CACCGGGC site and a secondary CACCGGCT site at 4 nM substrate concentration are shown in Fig. 3A. A non-linear dependency of catalytic activity of SgrAI relative to the substrate concentration was observed as demonstrated by the sigmoidal nature of the velocity curves (Fig. 3, A and B), implying cooperative binding of SgrAI to DNA substrate. Substrate dissociation constant \( (K_s) \), turnover number \( (k_{\text{cat}}) \) and Hill coefficient \( (n) \) were determined by Hill plots as described under “Experimental Procedures” and from the data shown in Fig. 3 (B and C). The measured kinetic parameters for all three canonical sites and two secondary sites are reported in Table III. The comparison of \( K_s \) and \( k_{\text{cat}} \) obtained with different canonical sites reveals that SgrAI kinetics on CGCCGGCG site is different from the kinetics on the remaining two canonical sites. The \( k_{\text{cat}} \) for the CGCCGGCG site was ~10-fold lower than the \( k_{\text{cat}} \) for either CACCGGGC or CACCGGCT sites. On the other hand, the apparent substrate dissociation constant \( (K_s) \) for the CACCGGCG and CACCGGCT sites were 10-fold higher than that seen for the CGCCGGCG site, suggesting that SgrAI has about 10-fold lower affinity for the first two sites relative to the CGCCGGCG site. Nonetheless, despite the significant differences for \( K_s \) and \( k_{\text{cat}} \), the values of the apparent second order rate constant \( (k_{\text{cat}}/K_s) \) are within a 2-fold range for all three sites (Table III). The apparent second order rate constants \( (k_{\text{cat}}/K_s) \) for the secondary sites are lower than those for the canonical sites by about 200- 500-fold (Table III). The Hill coefficient, \( n \), for either of the three canonical sites or the two secondary sites varied between 2 and 3.5, indicating positive cooperativity of substrate binding by SgrAI (Table III).

**SgrAI Cleavage at Target Sites Is Stimulated by Canonical Site Cleavage Products**—Previous studies have shown that...
The molar concentrations of substrate with canonical CACCGGGCG site and the specific (or nonspecific) duplex were maintained at a constant ratio of 1:50. The excess of specific duplex is necessary due to the lower binding affinity of SgrAI for the cleaved site versus the uncut canonical site. In the presence of the nonspecific duplex, a sigmoidal velocity curve (curve 2 in Fig. 4A) was obtained, which is very similar to that obtained in the absence of oligonucleotide duplex (curve 1 in Fig. 4A). In the presence of specific duplex, however, the cleavage efficiency increased at all ranges of substrate concentrations (curve 3 in Fig. 4A). Thus at low substrate concentrations, the initial velocity increased by 7- to 21-fold with the maximum effect seen at 3 nM substrate and 150 nM specific duplex, whereas at higher than 5 nM substrate concentrations, the increase was about 4- to 5-fold. The apparent second order rate constant ($k_{cat}/K$) in the presence of the specific duplex was also higher by 3.5-fold as shown in Table IV. Furthermore, the sigmoidal nature of the velocity curve was abolished in the presence of specific duplex as seen by the reduction in the Hill coefficient from 3 to 1.1 (Fig. 4B).

To evaluate the effect of the specific duplex on the cleavage rates at a secondary site (CACCAGGTCT), the molar ratio between the substrate and the specific duplex was maintained at 1:4, because at higher concentrations of specific duplex, an inhibitory effect on secondary site cleavage was previously observed (35). In the presence of the specific duplex, a nonlinear increase in the initial velocities of secondary site cleavage was encountered at all ranges of substrate concentrations. At low secondary site concentrations (1–5 nM), up to 100-fold increases in the initial velocities were detected, whereas at higher substrate concentrations the increase was about 50-fold. Thus the overall activation effect of the specific duplex is significantly stronger for secondary site cleavage than that observed for the canonical site cleavage. Accordingly, in the presence of the activator duplex the apparent second order rate constant ($k_{cat}/K$) for the secondary site was higher by 50-fold and the Hill coefficient was also reduced to 1.2, indicating an absence in substrate cooperativity (Table IV). Taken together, the data indicate that the cleaved canonical sites behave as allosteric activators for either the canonical CACCAGGTCT site or secondary CACCAGGTCT site cleavage by SgrAI.

**DISCUSSION**

SgrAI restriction endonuclease reveals an unusual substrate specificity, because it is capable of cleaving 17 different sequences, which include three canonical sites, CGCCAGGCG, CACCAGGTCT, and CACCAGGTG, and 14 secondary sites, including the sequences CPuCCGGPyATCTC and CPuCCGGG (33, 35). The SgrAI cleavage at secondary sites is a self-stimulated process evoked by reaction products generated from canonical site cleavage. Consequently, only substrates carrying canonical site(s) are cleaved efficiently at secondary sites. These observations suggest that SgrAI is capable of specifically interacting with three types of DNA sequences: the canonical sequence, the secondary sequence that differs from the canonical sequence by 1 bp, and the cleaved sequence that is the product of canonical sequence cleavage.

In this study, we have demonstrated using gel retardation studies that SgrAI specifically binds to all three types of specific sequences (Tables I and II). Overall, our results indicate that the binding affinity of SgrAI to either the canonical sites or to the secondary sites is in a range typically seen with other Type II restriction endonucleases to their canonical sequences (3–7). SgrAI also shows strong binding affinity to the oligonucleotide duplex that mimics the cleaved canonical site; the binding affinity is only 15-fold weaker relative to the uncut canonical site. SgrAI binding to secondary sites is 30-fold.
Specific duplex is flanked by the SgrAI-specific termini with as extensions of 5'-CCGG, which mimic the canonical site cleavage products. Nonspecific duplex is the same as the above-mentioned duplex, except it contains as extensions of 5'-GGCC. The turnover number ($k_{cat}$), substrate dissociation constant ($K$), and Hill coefficient ($n$) were obtained from two independent determinations as shown in Fig. 4. The values are shown as mean ± range.

| Recognition  | Oligonucleotide duplex | $k_{cat}$ | $K$ | $k_{cat}/K$ | $n$ |
|--------------|------------------------|-----------|----|-------------|-----|
| Canonical    | CACCGGCG               | 6.8 ± 0.4 | 6.7 ± 0.3 | 10.4 ± 0.1 | 3.0 ± 0.2 |
| GTGGCCGC     |                        | 8.4 ± 0.6 | 6.2 ± 0.1 | 13.4 ± 0.4 | 4.0 ± 0.5 |
| Canonical    | Nonspecific duplex     | 2.7 ± 0.3 | 6.2 ± 0.2 | 4.4 ± 0.4  | 1.1 ± 0.1 |
| GTGGCCGC     |                        | 2.0 ± 0.3 | 7.9 ± 0.3 | 2.7 ± 0.5  | 2.8 ± 0.3 |
| Canonical    | Specific duplex        | 5.8 ± 0.2 | 5.0 ± 0.2 | 1.2 ± 0.1  | 1.2 ± 0.2 |
| GTGGCCGA     |                        | 6.0 ± 0.3 | 6.7 ± 0.3 | 10.4 ± 0.3 | 3.0 ± 0.3 |

weaker relative to canonical sites (Table II), nevertheless, it still appears to be significantly stronger than the binding of the Type IIP restriction endonucleases to their star sites, which on average is 2–4 orders of magnitude weaker than canonical site binding (4–8).

Enzymes that associate into higher complexes for catalysis generally show a positive cooperativity effect. For example, the cooperativity for DNA cleavage has been previously observed for SfiI restriction endonuclease, which is a tetrameric enzyme that cleaves two copies of its recognition sequence simultaneously (42). A similar effect was also observed for NaeI and EcoRII restriction endonucleases, both of which are dimeric enzymes that interact with two copies of their recognition sequence but cleave at only one sequence (22, 28), and with FokI restriction endonuclease that binds DNA as a monomer but assembles into a dimeric complex for catalysis (13). Steady-state kinetic analysis for cleavage of either canonical or secondary sites by SgrAI also resulted in sigmoidal velocity curves indicating substrate cooperativity for SgrAI, and confirming our previous hypothesis that SgrAI requires the coordinated interaction with two sites for efficient catalysis (35). The values of the apparent second order rate constant ($k_{cat}/K$) are within a 2-fold range for all three canonical sites. Because the $k_{cat}/K$ value indicates the relative specificity of an enzyme toward its substrate regardless of reaction mechanism (43), it may be concluded that SgrAI exhibits similar specificity for all canonical sites. Similarly, SgrAI specificity for the secondary sites was found to be lower by about two orders of magnitude compared with the canonical sites. Such a low specificity to the secondary sites could be attributed to the trans configuration of secondary sites, which represents the least favorable substrate for SgrAI (35).

Most intriguing is the ability of SgrAI to make use of the cleaved canonical site in assisting in the cleavage of the next target, whether it is the canonical or the secondary site, which is achieved by a coordinated interaction between the two sites. It has been shown that all endonucleases requiring two sites for catalysis cleave plasmid substrates with a single site less efficiently than those with two sites. The cleavage efficiency of solitary sites, however, may be increased by providing a second site on a short oligonucleotide duplex in trans (22, 27, 42). SgrAI did not follow this pattern, because the canonical sites provided on oligonucleotide duplexes inhibited the cleavage of a plasmid substrate containing a single canonical site (34). However, a different outcome was observed in our study, wherein an oligonucleotide duplex having termini that mimic the cleaved canonical site was used instead of a duplex with the uncut canonical site. We have demonstrated that such specific duplexes have a stimulatory effect on the cleavage efficiency of SgrAI at both canonical and secondary sites. An especially high stimulatory effect (up to 100-fold) was observed for the secondary site. Thus, the cleaved canonical site behaves as an allosteric activator, where binding to the activator molecule increases the efficiency of catalysis. In the presence of this activator, substrate cooperativity is abolished, also suggesting that the catalytically active SgrAIdNA complex contains one protein subunit bound to the cleaved DNA molecule (activator), whereas the second subunit is bound to the uncut DNA molecule (substrate).

Some distant parallels may be drawn between SgrAI and EcoRII endonuclease, which belongs to a different sub-type of restriction endonucleases (2). As predicted from the amino acid sequence comparison, the putative catalytic site organization of EcoRII and SgrAI is similar (10, 44). An EcoRII complex that is bound to two copies of its recognition sequence, only one sequence is cleaved, whereas the other sequence serves as an allosteric activator (23, 24, 26–28). The products of the EcoRII reaction (cleaved sites) may also serve as an allosteric activator by stimulating the cleavage activity of EcoRII at the refractory sites (27). In this respect, SgrAI resembles EcoRII, because it also can use the reaction products to assist in cleavage at other sites; however, there is a major difference between the two enzymes. It has been shown that removal of the N-terminal domain of EcoRII, which binds to the activator molecule, results in a catalytically efficient enzyme that no longer requires two sites for catalysis (20). In contrast to EcoRII, the catalytically active SgrAI complex consists of two homodimeric subunits, both of which carry the binding and catalytic functions.

The restriction endonucleases NgoMIV (GCGGC) and Cfr101 (PuCCGGPy), which recognize and cleave DNA sequences that are subsets of the SgrAI recognition site (CPuCCGPy), have also been extensively studied (30, 31). Both NgoMIV and Cfr101 are tetramers of identical subunits. The tetrameric complex contains two active sites, one active site per homodimeric subunit, capable of interacting with two copies of a specific DNA, both of which are cleaved simultaneously (30). In contrast to NgoMIV and Cfr101, SgrAI has been found to exist as a homodimer in a solution (34). Hence, to coordinate...
the cleavage at two sites, two SgrAI dimers have to associate into a complex. It has been pointed out that the SgrAI dimer bound to a solitary site may have some, albeit low, activity suggesting that the SgrAI dimer may act as an orthodox dimeric restriction enzyme (34). Although the SgrAI dimer possesses a full set of catalytic residues sufficient to form a catalytic site that may result in slow cleavage, most likely, only after the association of two SgrAI dimers into a tetrameric complex, can the catalytic sites become optimally aligned for maximum activity. Our data indicate that the two SgrAI dimers in the active complex are not necessarily functionally identical, because in the complex one dimer bound to the uncleaved recognition site may have a catalytic function while the other dimer bound to the cleaved site may have a stimulatory function. This further implies that, in the SgrAI complex, catalysis may take place in one subunit at a time and the requisite conformational geometry for effective catalysis may be provided by the other subunit bound to either the uncleaved DNA or the cleaved DNA. Furthermore, the catalytically active tetrameric complexes of SgrAI may be bound to either two uncleaved canonical sequences or two uncleaved secondary sequences, or one cleaved canonical sequence in combination with one uncleaved non-canonical sequence or uncut secondary sequence.

All of these observations suggest that SgrAI is a new type of restriction endonuclease, albeit having evolutionary ties with both the tetrameric Cfr101I-like endonucleases and the dimeric EcoRII-like endonucleases. Most likely, all these endonucleases derived from a common ancestor but, during evolution, have either acquired/lost structural domains or evolved different protein-protein interfaces thus resulting in enzymes that cleave DNA by different kinetic pathways. We have recently cloned and overexpressed the PsuNI restriction endonuclease (1), which is an isochozyme of SgrAI. Our initial results demonstrate that PsuNI endonuclease, although revealing a 33% amino acid sequence identity with SgrAI, lacks the ability to cleave DNA at the secondary sites. We are currently performing mutagenesis studies on SgrAI and PsuNI endonucleases along with an in-depth biochemical analysis of PsuNI, because this will offer more insights into their structural and functional organization and shed light on the biological significance of these endonucleases.

Acknowledgments—We are very grateful to Steve Halford for sharing unpublished results and advise and Mike Dalton for purifying homologous SgrAI protein. We also thank Chris Noren for many insightful discussions, Rich Roberts and Elizabeth Raleigh for critical reading of the manuscript, and Don Comb for support of this research.

REFERENCES

1. Roberts, R. J., and Macelis, D. (2001) Nucleic Acids Res. 29, 268–269
2. Roberts, R. J., Belfort, M., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S. Kh., Dryden, D. T., Dyvig, K., Firman, K., Gromova, E. S., Gumpert, R. I., Halford, S. E., Hattman, S., Heitman, J., Hornby, D. P., Januzlatis, A., Jeltsch, A., Josephsen, J., Kiss, A., Klenhammer, T. R., Kobayashi, I., Kong, H., Kruger, D. H., Lack, S., Marinus, M. G., Miyahara, M., Morgan, R. D., Murray, N. E., Nagaraj, V., Pickarowicz, A., Pingoud, A., Raleigh, E., Rao, D. N., Reich, N., Repin, V. E., Selker, E. U., Shaw, P. C., Stein, D. C., Stoddard, B. L., Styhalski, W., Trautner, T. A., Van Etten, J. L., Vitor, J. M., Wilson, G. G., and Xu, S. Y. (2003) Nucleic Acids Res. 31, 1805–1812
3. Terry, B. J., Jack, W. E., Rubin, R. A., and Modrich, P. (1983) J. Biol. Chem. 258, 9820–9825
4. Lesser, D. R., Kupweski, M. R., and Jen-Jacobsen, L. (1990) Science 250, 766–768
5. Thellung, V., Alves, J., Fliss, A., Maass, G., and Pingoud, A. (1990) Biochemistry 29, 4862–4869
6. Alves, J., Selent, U., and Wolffe, H. (1996) Biochemistry 34, 11191–11197
7. Engler, L. E., Welch, K. K., and Jen-Jacobsen, L. (1997) J. Mol. Biol. 269, 82–101
8. Engler, L. E., Sapienza, P., Dorner, L. F., Kucera, R., Schildkraut, I., and Jen-Jacobsen, L. (2001) J. Mol. Biol. 307, 615–636
9. Conlan, L. H., and Dupureur, C. M. (2002) Biochemistry 41, 14848–14855
10. Pingoud, A., and Jeltsch, A. (2001) Nucleic Acids Res. 29, 3705–3727
11. Kozaroski, T., Skowron, P., and Podjaska, A. J. (1989) Gene (Amst.) 80, 269–276
12. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1997) Nature 398, 97–100
13. Bitinaite, J., Wah, D. A., Aggarwal, A. K., and Schildkraut, I. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 10570–10575
14. Vanamee, E. S., Santagata, S., and Aggarwal, A. (2001) J. Mol. Biol. 309, 615–636
15. Bath, A., Milsom, S. E., Gormley, A. A., and Halford, S. E. (2002) J. Biol. Chem. 277, 4024–4033
16. Ford, S. E., Bilcock, D. T., Stanford, N. P., Williams, S. A., Milsom, S. E., Gomley, N. A., Watson, M. A., Bath, A. J., Emberton, M. L., Growes, D. M., Daniels, L. E., Parry, S. H., and Szczelkun, M. D. (1999) Biochem. Soc. Trans. 27, 696–699
17. Halford, S. E. (2001) Biochem. Soc. Trans. 29, 363–373
18. Colandene, J. D., and Topal, M. D. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 3531–3536
19. Huai, Q., Colandene, J. D., Topal, M. D., and Ke, H. (2001) Nat. Struct. Biol. 8, 665–669
20. Mucke, M., Grelle, G., Behlke, J., Kraft, R., Kruger, D. H., and Reuter, M. (2002) EMBO J. 21, 5282–5286
21. Bickle, T. A., and Kruger, D. H. (1993) Microbiol. Rev. 57, 434–450
22. Yang, C. C., and Topal, M. D. (1992) Biochemistry 31, 9657–9664
23. Petrauskene, O. V., Kubareva, E., Gromova, R. S., and Shabarova, Z. A. (1992) J. Biol. Chem. 267, 6515–6522
24. Petrauskene, O. V., Babkina, O. V., Tashlitsky, V. N., Kazanov, G. M., and Gromova, E. S. (1998) FEBS Lett. 425, 29–34
25. Conrad, M., and Topal, M. D. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9707–9711
26. Kruger, D. H., Barcak, G. J., Reuter, M., and Smith, H. O. (1988) Nucleic Acids Res. 16, 4997–4908
27. Pein, C. D., Reuter, M., Meisel, A., Cech, D., and Kruger, D. H. (1991) Nucleic Acids Res. 19, 5139–5142
28. Gabbara, S., and Bhagwat, A. S. (1992) J. Biol. Chem. 267, 18623–18630
29. Nokels, T. J., Szczelkun, M. D., Wentzell, L. M., and Halford, S. E. (1998) J. Mol. Biol. 281, 419–432
30. Diebert, M., Grazulis, S., Sanausauska, G., Siksnys, V., and Huber, R. (2000) J. Mol. Biol. 297, 792–799
31. Siksnys, V., Skirgaila, R., Siksnys, G., Urbanke, C., Cherny, D., Grazulis, S., and Huber, R. (1999) J. Mol. Biol. 291, 1105–1118
32. Bilcock, D. T., Daniels, L. E., Bath, A., and Halford, S. E. (1999) J. Biol. Chem. 274, 36379–36386
33. Tautz, N., Kaluza, K., Jarch, M., Schmitz, G. G., and Kessler, C. (1990) Nucleic Acids Res. 18, 3087
34. Daniels, L. E., Wool, K. M., Scott, D. J., and Halford, S. E. (2003) J. Mol. Biol. 327, 579–591
35. Ikauka, J., Rossi, J. J., and Wallace, B. R. (1984) Annu. Rev. Biochem. 53, 323–356
36. Friedman, M., and Crothers, D. M. (1981) Nucleic Acids Res. 9, 6555–6525
37. Whitaker, R. D., Dorner, L. F., and Schildkraut, I. (1999) J. Mol. Biol. 285, 1525–1536