Nonidentical DNA-Binding Sites of Endonuclease *Nae*I Recognize Different Families of Sequences Flanking the Recognition Site

Charles C. Yang and Michael D. Topal*

UNC Lineberger Comprehensive Cancer Center and Department of Pathology, University of North Carolina Medical School at Chapel Hill, Chapel Hill, North Carolina 27599-7295

Received April 3, 1992; Revised Manuscript Received July 20, 1992

ABSTRACT: *Nae*I endonuclease uses a two-site binding mechanism to cleave substrate DNA: reaction-rate studies imply that occupancy of the second DNA site causes an allosteric change in the protein that enables DNA cleavage at the first site [Conrad, M., & Topal, M. D. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86,* 9707–9711]. Measurements of relative binding affinities for 14-base-pair DNA fragments containing the *Nae*I recognition sequence GCCGGC and various flanking sequences showed that the two DNA-binding sites are not identical. G·C-rich flanking sequences were preferred by the activator binding site, whereas A·T-rich flanking sequences were preferred by the substrate binding site: **GGGTGCCGGCAGGG** was preferred 8-fold more by the activator site but 14-fold less by the substrate site than **TTTCGCCGGCGTTT**. Substitution of pyrimidine or 7-deazapurine for purine immediately 3′ to GCCGGC reduced DNA affinity for only the activator site by up to 26-fold, implying that the activator DNA-binding site requires N-7 base contacts immediately flanking GCCGGC. The implications of nonidentical DNA-binding sites, one of which binds a specific DNA site to allosterically activate the other, are discussed.

Endonucleases *Bsp*MI, *Hpa*II, *Nae*I, *Nar*I, *Sac*II (Conrad & Topal, 1989; Oller et al., 1991), and *Eco*RII (Krüger et al., 1988) cleave some of their DNA substrates very poorly. The addition of exogenous DNAs containing readily cleavable cognate recognition sites activates the cleavage of these "resistant" substrates. For example, M13 double-stranded DNA contains a unique *Nae*I site that is resistant to cleavage. The addition of pBR322 or SV40 DNAs, which contain *Nae*I sensitive sites, to the reaction mixture results in the rapid cleavage of the M13 DNA (Conrad & Topal, 1989; Oller et al., 1991). Similarly, the addition of either pBR322 or λ DNA, which contain *Eco*RII-sensitive sites, activates *Eco*RII cleavage of resistant substrate sites in bacteriophage T3 and T7 DNA (Krüger et al., 1988). These results led the respective authors to propose a two-site model for enzyme-mediated DNA cleavage by these endonucleases.

Reaction rate studies showed that the enzymes *Bsp*MI, *Hpa*II, *Nae*I, *Nar*I, and *Sac*II (Oller et al., 1991) can be divided into two groups with respect to how two-site binding achieves DNA cleavage. Measurements of the effect of increasing substrate concentrations on initial reaction rates of DNA cleavage under conditions of a constant amount of activator DNA showed that the binding of one DNA site by *Hpa*II, *Nar*I, and *Sac*II increases the binding affinity (decreases the K_m) of the enzyme for the second DNA site (without affecting its k_{cat}), resulting in substrate cleavage. On the other hand, binding of a second DNA site by *Nae*I and *Bsp*MI enables cleavage of the first bound site (increases its k_{cat} without affecting its K_m) (Conrad & Topal, 1989; Oller et al., 1991).

Direct electron microscopic (EM) visualization of the interaction of *Nae*I with DNA also indicates multiple binding sites on the enzyme: *Nae*I interaction with pBR322 DNA, which contains multiple GCCGGC recognition sequences, showed loop formation with up to four DNA contacts per protein molecule bound (Topal et al., 1991). In agreement with the reaction rate measurements, these EM results imply two binding sites per protein monomer, assuming the protein binds its palindromic DNA recognition site as a dimer, as done by many [reviewed in Modrich and Roberts (1985)].

Thus, the existence of multiple binding sites on *Nae*I is well supported. To determine whether these sites are identical or not, we studied the *Nae*I cleavage of bacteriophage M13mp18 double-stranded DNA. This DNA contains a unique GCCGGC *Nae*I recognition site that is resistant to cleavage (Conrad & Topal, 1989). We measured the ability of DNA fragments containing the GCCGGC recognition sequence and various flanking sequences to activate the *Nae*I cleavage reaction. As expected, these DNA fragments were found both to activate *Nae*I cleavage of double-stranded (ds) bacteriophage M13mp18 DNA and to compete with this substrate for the *Nae*I substrate binding site. The ability to bind to both sites simultaneously enabled us to resolve the intrinsic dissociation constants, K_A and K_i, for binding of the DNA fragments to the activator and inhibitor (substrate) binding sites respectively. The relative affinities of various short DNA fragments for the two endonuclease *Nae*I DNA-binding sites demonstrated that these sites are nonidentical. The implications of these features and their possible relevance to protein–DNA recognition are discussed.

MATERIALS AND METHODS

Materials. Oligodeoxynucleotides were synthesized by machine (Applied Biosystems Model 380A), purified by gel electrophoresis, eluted from appropriate gel slices, precipitated with ethanol, phenol extracted, precipitated with ethanol again, and desalted prior to use. Distilled deionized water was used throughout. Endonuclease *Nae*I was purchased from New England Biolabs and characterized after further purification by DEAE-cellulose chromatography: the *Nae*I purity was ≥70% as judged by silver-stained SDS–polyacrylamide gel electrophoresis (PAGE) of the protein after purification. All additional reagents were purchased from commercial sources and used without further preparation.

Cleavage Reactions. Restriction enzyme cleavage reactions were performed in 10 μL containing 0.15 μg of M13mp18 replicative form I DNA (Boehringer-Mannheim) to give a

0006-2960/92/0431-9657$03.00/0

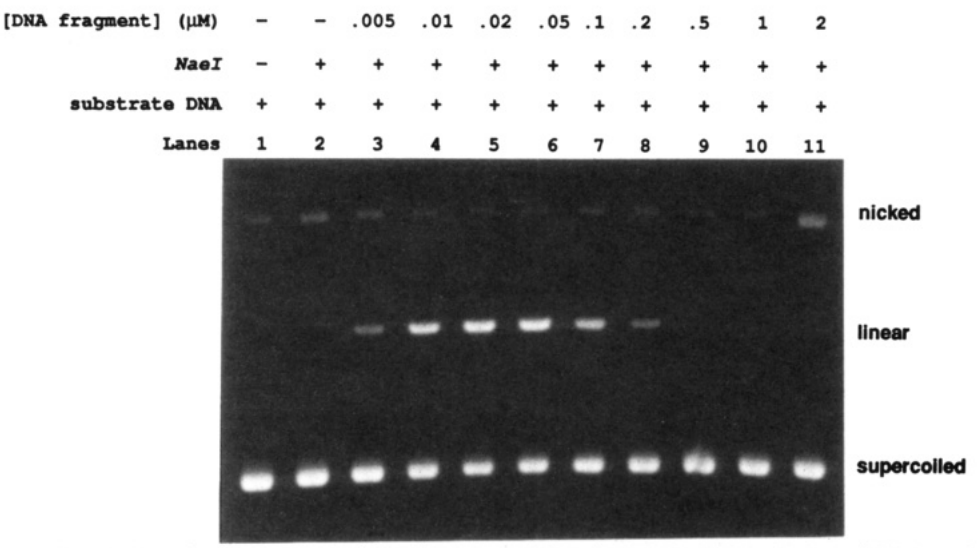

FIGURE 1: Activation and inhibition of *Nae*I cleavage of its unique recognition site in M13 DNA. 0.15 μg of M13mp18 replicative form I DNA was incubated with 0.4 unit of *Nae*I and the indicated final concentration of DNA activator fragment in 10 μL of 20 mM NaCl, 10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol, and bovine serum albumin at 0.1 μg/mL concentration. Lane 1, M13 DNA only. Lane 2, M13 DNA with *Nae*I and no fragment. Lanes 3–11, *Nae*I cleavage of M13 DNA in the presence of indicated concentrations of fragment **8**.

final concentration of 3.1 nM and 0.4 unit of *Nae*I in 20 mM NaCl, 10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol, and bovine serum albumin at 0.1 mg/mL concentration, unless otherwise indicated. One unit of *Nae*I activity is that amount of enzyme required to completely digest 1 μg of Adeno-2 DNA in 1 h at 37 °C in 50 μL of the above reaction buffer. The molar ratio of M13 DNA substrate to *Nae*I in the reactions was ≥20 based on a dimer molecular mass of 70 kDa, 70% enzyme purity, and the optical density of the protein banded by SDS–PAGE along with known amounts of bovine syrum albumin. Reaction mixtures were incubated at 37 °C. Reactions were terminated by heat inactivation of the enzyme at 65 °C for 10 min, which reduced *Nae*I cleavage activity to below our level of detection, and then kept at 4 °C until analysis; reaction products were separated on 1% agarose gels.

Quantitation of NaeI Cleavage of DNA. Cleavage of M13mp18 RFI DNA was quantitated by densitometric analysis of negatives from photographs of the ethidium bromide-stained gels taken under UV light, as described previously (Conrad & Topal, 1989). Optical density was linear with concentration over the range of concentrations and photographic conditions chosen, as shown by a standard curve. Results were corrected for differences in ethidium bromide staining of linear, relaxed, and supercoiled DNA. Velocities of M13 DNA cleavage reactions were determined from experiments performed under initial rate conditions in the linear region of the cleavage reaction (≤20 min).

Melting Points of the DNA Fragments. To certify that activators remained double-stranded under the reaction conditions employed, DNA melting curves were determined for several of the activators including those predicted to contain the most destabilizing substitutions. T_m of 14-mer DNA fragments were determined by adding 6 μg of each of the two 14-base oligomers (to give double-stranded DNA fragment) into 1.0 mL of *Nae*I reaction buffer. Reaction buffer contained 20 mM NaCl, 10 mM Tris-HCl (pH 8.0), and 10 mM $MgCl_2$. The hyperchromic absorbance change at 260 nm was followed using a Perkin-Elmer 320 UV–vis spectrophotometer connected to a digital temperature controller set to increase the temperature from 20 to 85 °C at 0.5 °C/min. T_m values were determined from first-order derivative plots of absorbance vs temperature.

DATA ANALYSIS

Effects of Various DNA Fragments on the NaeI Cleavage Reaction. The effects of 14-base-pair DNA fragments (**1–13** listed in Tables I and II) on *Nae*I cleavage of M13 DNA substrate were measured. These DNA fragments contain GCCGGC and flanking sequences with either base-pair substitutions or with DNA base analogs. Short DNA fragments containing the GCCGGC *Nae*I recognition sequence activate *Nae*I cleavage of M13mp18 double-strand DNA (Conrad & Topal, 1989). Under appropriate conditions, these fragments also inhibit M13 DNA cleavage because they contain the recognition sequence. These properties allowed us to measure the effects of sequences flanking the recognition site on binding to both the activator and the substrate DNA-binding sites. The results of such a cleavage reaction are illustrated for fragment **8**, which both activates and inhibits strongly, making it easy to visualize both of these effects (Figure 1).

Reaction Scheme for Activated Cleavage of DNA by NaeI. A reaction scheme (Figure 2) for the interaction of activator DNA fragment (A) and substrate M13 DNA (S) with *Nae*I (-E-) that incorporates the properties of DNA cleavage by *Nae*I is shown in Figure 2. These properties are briefly outlined as follows: occupancy of the substrate site and the activator site leads directly to cleavage; initial rate studies imply that occupation of the activator site causes an allosteric change in the protein (O↔□) that enables cleavage at the other site (Conrad & Topal, 1989). Activator, by virtue of its *Nae*I recognition sequence, occupies either the substrate site (-EA) or the activator site (AE-): occupation of the activator site cleaves either S (AES) or A (AEA) occupying the other site. Activator occupation of the substrate site (-EA) inhibits S cleavage (Figure 1). A is cleaved upon occupation of the other site (AEA); this bimolecular interaction predicts a sigmoidal velocity curve.

The two-site model predicts that S also binds to the activator site. The affinity of S for the activator site must be relatively low, however, since S concentrations up to K_m do not result

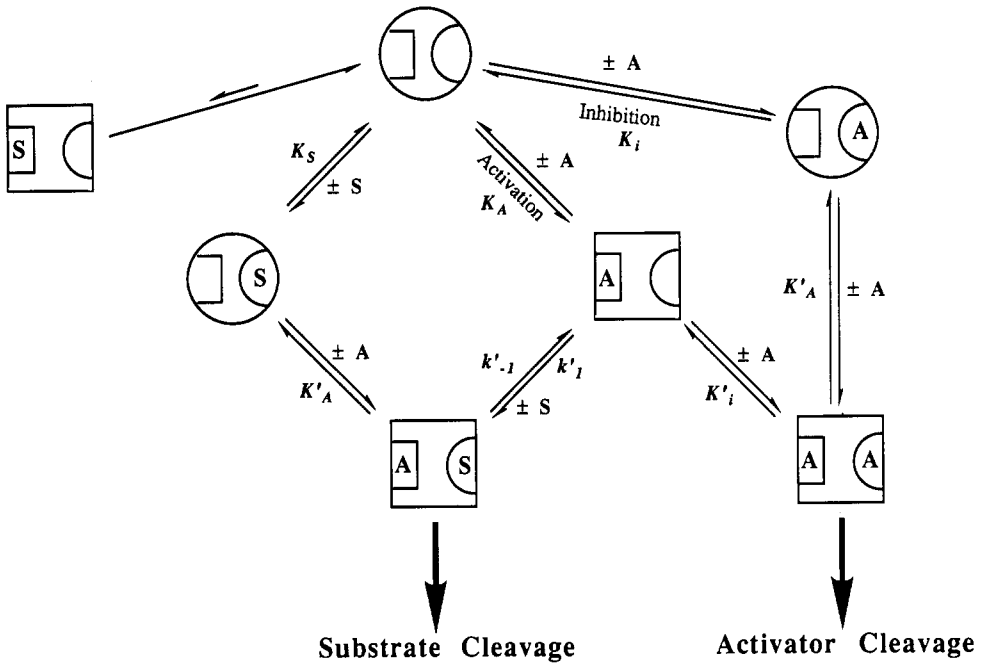

FIGURE 2: Reaction scheme for the activation and inhibition of M13 DNA cleavage by *Nae*I. The circle and square represent unactivated and activated forms of *Nae*I, respectively. Semicircular and square enzyme binding sites represent substrate and activator binding sites, respectively. M13 DNA substrate and DNA activator fragment are represented by S and A, respectively; constants are defined by eqs 1–4. See text for a description of the reaction properties and evidence to support this scheme.

in significant levels of cleavage (Conrad & Topal, 1989). It could be argued that S can occupy the activator site without activating cleavage. This possibility is eliminated by the shape of the v_{obs} versus [S] curves in the presence of A [Conrad and Topal (1989) and results not shown]: binding of "nonactivating" substrate to the activator site would inhibit cleavage at high [S].

[A] does not affect the apparent K_m for cleavage of M13 DNA (Conrad & Topal, 1989), so $K_S = K'_S$. Therefore, since $K_S K'_A = K'_S K_A$, then $K_A = K'_A$, and by similar arguments, $K_i = K'_i$, where K_A and K_i are the intrinsic dissociation constants for the activator and substrate sites, (AE-) and (-EA), respectively. The scheme is independent of *Nae*I subunit interactions, but the palindromic recognition site and the EM results imply that *Nae*I interacts with DNA as a multimer.

Rate Equation for Activated Cleavage of DNA by NaeI. Based on the scheme in Figure 2, an equation was derived using steady-state reaction assumptions [see Segal (1975) for a discussion of these assumptions], relating reaction velocity (v_{obs}) to [A] and [S], which enables determination of K_A and K_i:

$$K_S = [\text{-E-}][S]/[\text{-ES}] = [\text{AE-}][S]/[AES] \quad (1)$$

$$K_m = K_S + k_c/k_1 \quad (2)$$

where, k_1 and k_c are the association and cleavage rate constants, respectively.

$$K_A = [\text{-E-}][A]/[\text{AE-}] = [\text{-ES}][A]/[AES] \quad (3)$$

$$K_i = [\text{-E-}][A]/[\text{-EA}] = [\text{AE-}][A]/[AEA] \quad (4)$$

$$[E_t] = [AEA] + [\text{-EA}] + [AES] + [\text{AE-}] + [\text{-ES}] + [\text{-E-}] \quad (5)$$

$$v_{obs} = k_p[AES] \quad (6)$$

$$v_{obs}/V_{max} = [AES]/[E_t] \quad (7)$$

Substituting from eq 4 for $[E_t]$ gives

$$v_{obs}/V_{max} = ([\text{AE-}][S]/K_m)/([\text{AE-}][A]/K_i + [\text{-E-}][A]/K_i + [\text{AE-}][S]/K_m + [E][A]/K_A + [\text{-E-}][S]/K_m + [\text{-E-}]) \quad (8)$$

Substituting from eqs 1–3 gives

$$v_{obs}/V_{max} = ([A][S]/K_m K_A)/([A]^2/K_A K_i + [A]/K_i + [A][S]/K_A K_m + [A]/K_A + [S]/K_m + 1) \quad (9)$$

Simplifying and rearranging, we obtain

$$v_{obs}/V_{max} = [S]/[(1 + K_A/[A])(K_m(1 + [A]/K_i) + [S])] \quad (10)$$

By comparison of eq 10 with the standard form of the Henri–Michaelis–Menten equation:

$$v_{obs}/V_{max} = [S]/(K_m + [S]) \quad (11)$$

it can be seen that

$$K_m^{app} = K_m(1 + [A]/K_i) \quad (12)$$

and

$$V_{max}^{app} = V_{max}/(1 + K_A/[A]) \quad (13)$$

Equation 10 at high [A] to K_A ratio reduces to the standard form of the velocity equation in the presence of competitive inhibitor:

$$v_{obs}/V_{max} = [S]/[(K_m(1 + [A]/K_i) + [S])] \quad (14)$$

K_m and V_{max} values used in eq 10 were determined from a double-reciprocal plot of initial reaction rate measurements in the presence of a constant amount of activator as described in Conrad and Topal (1989) (but using reaction conditions

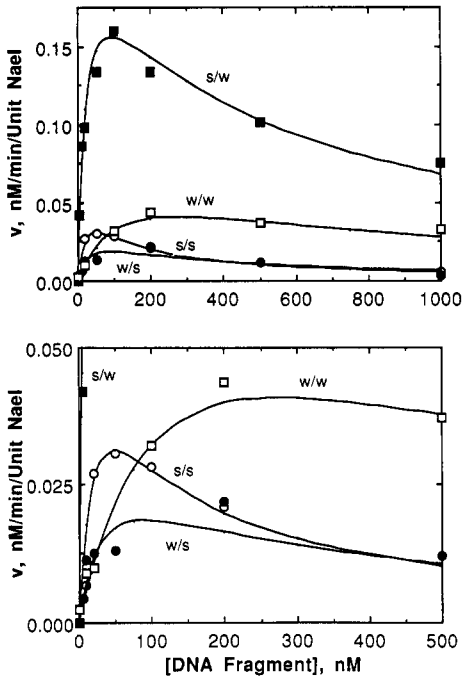

FIGURE 3: Effects of DNA fragments on velocities of the *Nae*I cleavage of M13 DNA. Top: Activation of cleavage was by DNA fragments **2** (■), **6** (□), **8** (○), and **11** (●). Fragment **2** strongly activates but weakly inhibits cleavage (s/w); fragment **6** weakly activates and weakly inhibits cleavage (w/w); fragment **8** is strong at doing both (s/s); and fragment **11** weakly activates but strongly inhibits cleavage (w/s). Reaction velocities were determined as described in Materials and Methods. The curves shown are best-fit from eq 10 as described in Data Analysis and gave the K_A and K_i listed in Tables I and II. Bottom: The bottom panel expands the region of the top figure from 0 to 500 nM DNA fragment and from 0 to 0.05 nM M13 DNA cleaved per minute per unit of *Nae*I.

described above). The values determined for K_m and V_{max} were 7 nmol of M13 ds DNA per liter and 0.75 nmol of M13 ds DNA cleaved per liter per minute per unit of enzyme, respectively. V_{max} was assumed to be independent of flanking sequence effects on reaction velocity and is defined as the maximum velocity for fully activated enzyme at saturating substrate concentrations.

K_A and K_i values were determined iteratively from eq 10 using Cricket Graph (Cricket Software, Malvern, PA) to generate families of curves representing different ranges of values of K_A and K_i. K_A and K_i values listed in Tables I and II entered in eq 10 gave curves that best fit (as judged by eye) the experimental data points. The error ranges of K_A and K_i values were determined from multiple independent determinations and are indicated in the tables. Small variations in individual data points do not significantly affect the K_A and K_i values, which were found to be more sensitive to the overall shape of the curve (see Results).

RESULTS

DNA Fragment Effects on NaeI Cleavage Reaction. Densitometric analysis of photographic negatives of electrophoresed cleavage products (e.g., see Figure 1) allowed quantitation of the effects of DNA fragment concentration on reaction velocity (Figure 3). Figure 3 illustrates these effects for fragment **2**, which strongly activates but weakly inhibits cleavage (s/w); fragment **6**, which weakly activates and weakly inhibits cleavage (w/w); fragment **8**, which is strong at doing both (s/s); and fragment **11**, which weakly activates but strongly inhibits cleavage (w/s). The differences among each fragment's ability to activate and inhibit are clearly illustrated. For example, the two extremes of behavior are s/w and w/s: in the former a high velocity of cleavage is observed over a wide range of fragment **2** concentration because of the inability of fragment **2** to inhibit the highly activated enzyme; at the other extreme fragment **11** strongly inhibits the already low cleavage activity. In Figure 3, the curves were calculated from eq 10 as described in Data Analysis: the K_A and K_i values that, when substituted in eq 10, define the shape of the velocity curves that best-fit the experimental points are reported in Tables I and II.

Table I: DNA Fragments and Values of K_A and K_i for Activation of M13 DNA Cleavage by *Nae*I: Effect of Nearest-Neighbor Base-Pair Modifications on Activation and Inhibition[a]

	DNA fragment	K_A, nM	K_i, nM	T_m, °C
1	GGGCGCCGGCGGGG CCCGCGGCCGCCCC	15 ± 3 (3)	83 ± 14	nd
2	GGGTGCCGGCAGGG CCCACGGCCGTCCC	17 ± 3 (3)	310 ± 40	70 ± 1
3	GGGCGCCGGCGGTG CCCGCGGCCGCCAC	45 ± 8 (2)	55 ± 10	nd
4	GGGGGCCGGCCGTG CCCCCGGCCGGCAC	470 ● 80 (2)	140 ± 55	nd
5	GGGUGCCGGCAGGG CCCACGGCCGUCCC	23 ● 5 (2)	250 ± 50	nd
6	GGGCGCCGGC*GGG CCC*CGGCCGCCCC	180 ± 60 (2)	100 ± 30	68 ± 1
7	GGGTGCCGGC#GGG CCC#CGGCCGTCCC	440 ± 50 (3)	127 ± 25	nd

[a] The *Nae*I recognition site is underlined. (*) 7-Deazaguanine. (#) 7-deazaadenine. Bases differing from those in DNA fragment **1** are represented in bold type. The values for K_A and K_i are given as mean ± SD for n (the number of independent determinations of both K_A and K_i indicated in parentheses) = 3 and mean ± range for n = 2; n refers to both K_A and K_i. T_m values are mean (n = 2) ± range.

Table II: DNA Fragments and Values of K_A and K_i for Activation of M13 DNA Cleavage by *Nae*I: Effects of T·A for G·C Substitutions on Activation and Inhibition[a]

	DNA fragment	K_A, nM	K_i, nM	T_m, °C
8	GGGCGCCGGCGTGG CCCGCGGCCGCACC	26 ± 5 (2)	33 ± 7	nd
9	GGGCGCCGGCGTTG CCCGCGGCCGCAAC	41 ± 8 (3)	10 ± 2	nd
10	GTTCGCCGGCTTTC CAAGCGGCCGAAAG	83 ± 11 (3)	97 ± 25	65 ± 1
11	TTTCGCCGGCGTTT AAAGCGGCCGCAAA	137 ± 30 (3)	22 ± 3	nd
12	TTTTGCCGGCATTT AAAACGGCCGTAAA	210 ± 50 (2)	85 ± 20	61 ± 1
13[b]	GGGCGCCGGGGGGG CCCGCGGCCCCCCC			

[a] The *Nae*I recognition site is underlined. Bases differing from those in DNA fragment **1** are represented in bold type. Error ranges are defined in Table I. [b] No activation of M13 DNA cleavage was observed for fragment **13** at concentrations up to 2 μM.

Effects of Base-Pair Substitutions on Values of K_A and K_i. Since both the activator DNA fragment and the M13 substrate DNA contain an identical core recognition sequence, the observed differences among K_A and K_i values for the DNA fragments tested (Tables I and II) must be due to the DNA sequences flanking the DNA recognition site. The DNA fragments listed in Table I were used to examine the effects on K_A and K_i of base-pair substitutions immediately 5′ and 3′ to the recognition sequence. Values determined for K_A show that DNA fragments possessing a purine immediately 3′ and a pyrimidine immediately 5′ to the *Nae*I recognition sequence have the highest relative affinities for the *Nae*I

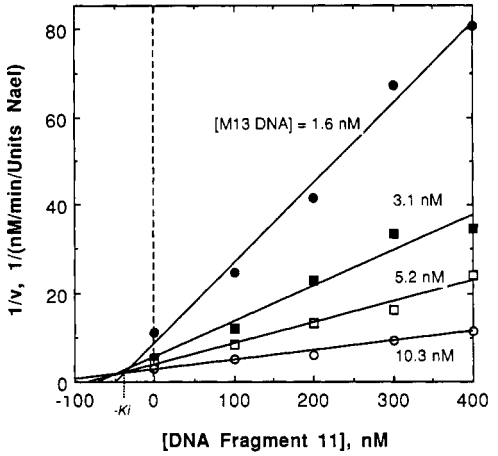

FIGURE 4: Dixon plot showing competitive inhibition of *Nae*I cleavage of M13mp18 RF DNA by DNA fragment **11**. Reaction conditions are described in Materials and Methods. *Nae*I was activated by the addition of 100 nM DNA fragment **2** to each reaction mixture. The lines intersect at $-K_i$.

activator site (fragments **1** and **2**). Substitution of a pyrimidine immediately 3′ and a purine immediately 5′ to the *Nae*I recognition sequence caused an 10-fold increase in K_A (fragments **3** and **4**).

Base analogs of A, G, and T were substituted in the nearest-neighbor positions to determine if base recognition was responsible for the effects of substituting pyrimidine for purine at these positions. Substitution of the 7-deaza analog for A and G had a large effect on K_A but only small effect on K_i; substitution of U for T had no significant effect on K_A and K_i values (Table I).

We also examined the effect on binding to the activator site of base substitutions further removed from the recognition sequence within a 14-base-pair DNA fragment (Table II). T·A for G·C substitutions were tested because they maintain the exocyclic carbonyl and amino group interactions, but eliminate the imidazole ring contacts. A single T·A for G·C substitution two bases removed from the *Nae*I recognition site (fragment **8**) resulted in a 2-fold increase in K_A, but a 2- to 3-fold decrease in K_i. A single T·A for G·C substitution three bases removed (fragment **3** in Table I) from the recognition site increased K_A 3-fold. Double substitution at these positions (fragment **9**) resulted in a combination of these effects: relative binding affinity for the substrate site was increased, and the relative binding affinity for the activator site was decreased. Values for K_A progressively increased as the number of T·A base pairs increased (see fragments **8–12**). Substitution within the recognition sequence completely eliminated activation and inhibition (fragment **13**).

Competition between Substrate and Activator for the Substrate Site. The reaction scheme (Figure 2) assumes that competition between substrate and activator for the substrate site is responsible for the inhibition of cleavage at high activator concentrations. Competition between the M13 DNA substrate and DNA fragment **11** for binding to the enzyme substrate site is shown by a Dixon plot (Figure 4). The cleavage of M13 DNA substrate is at maximal velocity with fragment **2** at 100 nM concentration; fragment **11** competitively inhibits cleavage. The lines of the Dixon plot intersect at [A] = $-K_i$. K_i was determined to be 35 nM, and V_{max} was determined to be 0.50 nmol of M13 DNA cleaved per liter of reaction mixture per minute per unit of enzyme. Also, an estimate of K_m = 9 nM was made from a $1/v$ versus $1/[S]$ plot made from the values in Figure 4 at [I] = 0. These values compare well with the values of K_m (7 nM) and V_{max} [0.75 nM/(min·unit)],

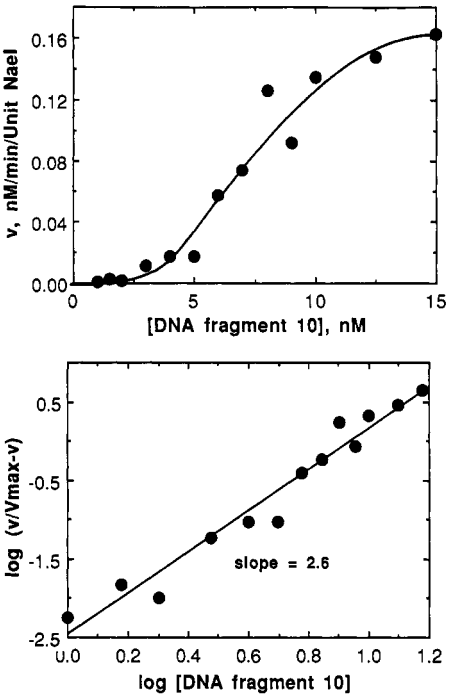

FIGURE 5: Effect of increasing activator concentration on activator cleavage. Reaction conditions were the same as described in Figure 1 except no M13 DNA was present. The activator was fragment **10**, the sequence of which is identical to the DNA sequence surrounding the *Nae*I recognition site in M13 DNA. The top panel shows the velocity dependence on concentration. The bottom panel shows a Hill plot of the data; n_{app} = slope; V_{max} was estimated to be 0.17 nM/(min·unit) at high [A], from the top panel.

determined from initial rate measurements by double-reciprocal plot (see Data Analysis), and with the value for K_i for fragment **11** (22 nM) determined from eq 10 (Table II).

Self-Activation. The reaction scheme, because it is based on a two-site model, predicts a sigmoidal velocity curve for the cleavage of any self-activating DNA (AEA). The initial cleavage velocity of fragment **10**, which is identical to the DNA sequence surrounding the *Nae*I recognition site in M13 DNA, was studied at increasing fragment **10** concentrations (Figure 5). The velocity curve was sigmoidal and a Hill plot of the velocity data is linear; the slope of the Hill plot gives n_{app} = 2.6; the velocity curve for a slightly longer version of DNA fragment **10**, 23 base pairs, was also sigmoidal with n_{app} = 2.0 (not shown).

DNA Fragment Concentrations. Accurate values for K_A and K_i depend on a knowledge of the concentration of DNA fragment in the cleavage reaction and lack of significant cleavage of the DNA fragments during the reaction. The presence of denatured DNA fragment would cause overestimation of the values of K_A and K_i. The relative affinities of DNA fragment for the activator and substrate sites would be correct, but comparisons between fragments would be affected. Therefore, the T_ms of fragments **2**, **6**, **10**, and **12** were measured to determine their stability under reaction conditions. The results (Figure 6) show that the measured T_ms were between 60 and 70 °C with no significant melting at 37 °C under our reaction conditions; among these DNA fragments are those containing the most destabilizing substitutions. Thus, the comparisons are valid among all the DNA fragments listed in Tables I and II.

Different rates of cleavage for the different activator DNA fragments could also lead to actual concentrations of fragment quite different from the expected concentrations. Initial reaction rate conditions, however, were used for all our

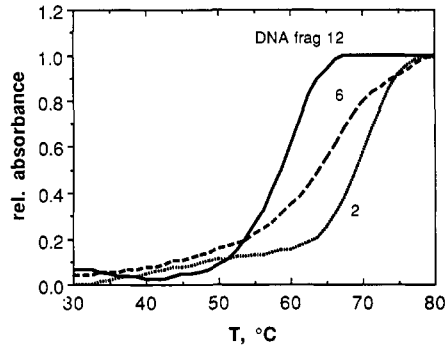

FIGURE 6: Melting curves of example 14-base-pair DNA fragments shown in Tables I and II. Samples were prepared for melting as described in Materials and Methods. T_m values were determined from first-order derivative plots of absorbance vs temperature at which the rate of increase in absorbance at $\lambda = 260$ nm was greatest (inflection point of the melting curve). Starting absorbances were all 0.34. The T_ms determined are reported in Tables I and II. The melting curve for fragment **10** was omitted for clarity. Relative absorbance, the absorbance curves normalized so that maximum absorbance was set equal to 1.0, is shown for visual comparison. The actual increases in absorbance ($\lambda = 260$ nm) with temperature for fragments **2**, **6**, **10**, and **12** were 16%, 12%, 11%, and 14% of the starting absorbance, respectively.

measurements to ensure that the concentration of DNA fragments and M13 DNA substrate did not significantly change over the reaction incubation times. Also, example DNA fragments were radiolabeled, and their cleavage was observed by gel analysis of cleavage products over the time course of our experiments; the maximum amount of cleavage observed was approximately 3% of starting material, under our reaction conditions. In addition, fragment **2** was resynthesized to contain phosphorothioate linkages at the scissile bonds: these linkages were found to be resistant to *Nae*I cleavage (Conrad & Topal, 1992) so that the DNA fragment persisted during the *Nae*I cleavage reaction. The K_A and K_i values determined with this fragment (17 nM and 250 nM, respectively) are in reasonable agreement with those values reported in Table I for fragment **2**.

DISCUSSION

Extended Recognition of Flanking Sequences. Seeman et al. (1976) showed by model building that all four possible base pairs at any sequence position may be distinguished by hydrogen bonding between protein side chains and nucleotides within the DNA major groove. This observation is supported by the crystal structure of the *Eco*RI–substrate DNA complex (McClarin et al., 1986; Kim et al., 1990) and by the crystal structures of repressor–operator complexes (Anderson et al., 1987; Jordan & Pabo, 1988). With hydrogen bonds from one amino acid interacting with more than one base pair, six-member ring contacts involving exocyclic groups and imidizole ring contacts at N-7 were characterized by the crystal structures. Substitution of T for G and C for A conserves the protein–exocyclic DNA contacts, but eliminates the N-7 contacts. Substitution of C for G and T for A eliminates both the exocyclic and N-7 contacts (Seeman et al., 1976). Our observation of the effect on K_A of pyrimidine for purine substitutions 3′ to the recognition sequence are explained by these findings. Our results show that substitution of C for either G or A at this position significantly increases the value of K_A; this can be explained by a reduction in the activator's affinity for the activator site because of loss of the N-7 contact common to both substitutions. Similar arguments suggest that the enzyme substrate site recognizes exocyclic contacts at this sequence position.

To test this hypothesis, 7-deazaA and 7-deazaG, which substitute a carbon in place of the nitrogen at N-7, were substituted for A and G, respectively, at sites immediately 3′ to the recognition site for *Nae*I (fragments **6** and **7**). Base-analog substitutions within the recognition site of restriction enzymes have been successfully used to probe the protein–DNA contacts employed by restriction enzymes to find their recognition sequences among the large background of other DNA sequences in the cell (Modrich & Rubin, 1977; Berkner & Folk, 1979; Dwyer-Hallquist et al., 1982; Bodnar et al., 1983; Brennan et al., 1986; Jiricny & Martin, 1986; Jiricny et al., 1986; McLaughlin et al., 1987; Lesser et al., 1990; Voigt & Topal, 1990). Our studies use base-analog substitution to look for specific contacts of protein with neighbors nearest the recognition sequence of *Nae*I.

Measurement of K_A and K_i for the DNA fragments containing deaza analogs confirms the importance of N-7 at the nearest-neighbor positions: decreased affinity of the enzyme activator site for the DNA fragment relative to its unmodified analog is evident. There was no significant effect of the 7-deaza analogs on K_i. Substitution of U for the T paired opposite the 3′ A, on the other hand, resulted in little change in K_A and indicates that the 5-methyl group from thymine does not significantly affect base recognition by *Nae*I at this sequence position (fragment **5**).

The ability of pyrimidine and 7-deaza analog substitutions for the 3′ neighboring purine to greatly increase K_A suggests that there is direct sequence recognition at the activator site of *Nae*I endonuclease that extends beyond the six-base-pair recognition sequence: the extended recognition involves purine/pyrimidine discrimination at the positions immediately adjacent to the recognition site. Whereas these results do not rule out alteration in the structure of the six-base-pair core recognition site because of these changes in the flanking sequences, the subtlety of the deaza analog substitutions relative to the magnitude of their effects on K_A argues for extended recognition to the base pairs immediately neighboring the recognition sequence. Differences in stability between the DNA fragments containing 7-deaza analogs and their unmodified counterparts did not contribute to these K_A differences: stability differences between these fragments were shown by T_m measurements to be negligible. DNA conformation and flexibility can have significant effects on protein binding [Lesser et al. (1990), and see Travers (1989) for a review]; base pairs beyond the nearest neighbors affected *Nae*I binding to DNA, but this study did not address the basis for these effects.

The effect of bases outside the recognition sequence on *Nae*I binding and cleavage are consistent with site-directed mutagenesis, chemical protection, and X-ray crystallography studies of the interaction of *Eco*RI with DNA. These studies demonstrated that bases outside the *Eco*RI recognition sequence are contacted by *Eco*RI (Lu et al., 1981; McClarin et al., 1986). These contacts may account for the up to 10-fold variation in *Eco*RI cleavage of the five recognition sites in λ DNA (Thomas & Davis, 1974) and the five recognition sites in adenovirus DNA (Forsblum et al., 1976). In addition, Taylor and Halford (1992) found that *Eco*RV cleavage of sequences that differ from its recognition sequence by one base pair depends on flanking sequences. Although generally cleaved up to 10^6-fold slower than the recognition sequence, some noncognate sites are cleaved 50-fold faster than others: faster cleavage at the preferred noncognate sites was caused by the DNA sequences in the 4 base pairs from either side of the site. For *Eco*RV the DNA flanking sequence effects were

attributed to the relative flexibility of the noncognate recognition sites.

Resistance of M13 DNA to NaeI Cleavage. A possible explanation for the resistance of M13 DNA substrate to *Nae*I cleavage comes both from the lower affinity of M13 DNA for the *Nae*I activator and substrate sites and from the reaction scheme (Figure 2). The lower affinity is evident from the higher K_A and K_i values for fragment **10** compared to those for fragment **8**. The sequence of fragment **10** is identical to that of M13 DNA surrounding the *Nae*I recognition site. And, according to the reaction scheme (Figure 2), self-activated cleavage of either substrate (SES) or activator (AEA) is second order and so the velocity curve for cleavage of activator is sigmoidal (Figure 5). The half-life for a second-order reaction with a single reactant is given by $kt_{1/2} = 1/[A_0]$: the half-life of the reactant depends on its concentration. On the other hand, to activate cleavage of substrate (AES), activator concentrations are generally used at about 25–100 nM, making for a pseudo-first-order reaction with respect to S. The half-life for a first-order reaction is given by $kt_{1/2} = \ln 2$: the half-life of the reactant is independent of its initial concentration. Thus, the lower affinity of resistant substrates for the activator site of *Nae*I coupled with the second-order rate constant and the low concentration of natural DNA used in most cleavage reactions (3–5 nM) preclude significant amounts of cleavage.

Nonidentical Enzyme Binding Sites. Our results demonstrate that native *Nae*I endonuclease contains nonidentical binding sites that prefer different classes of flanking sequences. Both sites require a core *Nae*I recognition sequence for binding; however, we found additional sequence recognition extending into the sequences flanking the recognition sites. The enzyme activator site requires a pure N-7 3′ to the core sequence; the absence of this site reduces binding by up to 26-fold (see fragments **2** and **7**). The enzyme activator site demonstrates up to a 14-fold preference for G·C-rich flanking sequences, whereas the enzyme substrate site demonstrates up to a 14-fold preference for A·T-rich flanking sequences (see fragments **1**, **2**, **11**, and **12**). These preferences were strikingly apparent in the binding of fragments **2** and **11** at the substrate and activator sites. **GGGTGCCGGCAGGG** was preferred 8-fold more by the enzyme activator site but 14-fold less by the enzyme substrate site than **TTTTGCCGGCGTTT**.

Thus, *Nae*I contains nonidentical sites that recognize different families of sequences. One of these sites binds activator DNA, and the other binds substrate DNA. Why is a complex control mechanism needed by an enzyme whose only purpose is allegedly to recognize and cleave foreign DNA? One possibility is that the requirement for binding two DNA sequences before cleavage occurs ensures against chance cleavage of imperfect recognition sequences—an event that kills the host: the probability of recognizing two sites is the product of the probabilities of recognizing either site alone. If increasing the accuracy of cleavage is the rational for two DNA-binding sites, however, then two identical sites would be expected. Of course, binding site differences could be an artifact of the evolution of *Nae*I: a defense against invading DNA may have been recruited from an existing nuclease activity in the cell.

The rational for the interesting features of *Nae*I in which nonidentical enzyme binding sites bind different and distant DNA sequences to effect protein function is under study. These features are not likely to be unique to the restriction enzymes. Other proteins that process DNA require binding of multiple DNA sites. The mechanisms that govern site-specific recombination (Gellert & Nash, 1987; Moitoso de Vargas et al., 1989; Hughes et al., 1990), initiation of DNA replication (Dodson et al., 1985), and control of transcription of the eukaryotic genome (Knight & Botchan, 1991; Li et al., 1991; Mastrangelo et al., 1991; Su et al., 1991) require multiple DNA–protein interactions to achieve precise site specificity (Echols, 1986). The λ Integrase protein, like *Nae*I, contains DNA-binding sites that recognize two different sequence families. For the site-specific recombinases (Hughes et al., 1990) and type II topoisomerases (Maxwell & Gellert, 1984; Corbett et al., 1992) apparently all enzyme sites must be filled before DNA cleavage occurs; this requirement may be to avoid unproductive cleavage of the DNA. This avoidance of unproductive cleavage could be related to the feature that we speculate enhances the accuracy with which *Nae*I differentiates foreign from host DNA.

ACKNOWLEDGMENT

We thank Terri Maness for performing some of the activator titrations.

REFERENCES

Anderson, J. E., Ptashne, M., & Harrison, S. C. (1987) *Nature 326*, 846–852.
Berkner, K., & Folk, W. R. (1979) *J. Biol. Chem. 254*, 2551–2560.
Bodnar, J. W., Zempsky, W., Warder, D., Bergson, C., & Ward, D. C. (1983) *J. Biol. Chem. 258*, 15206–15213.
Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986) *J. Biol. Chem. 261*, 7270–7278.
Conrad, M., & Topal, M. D. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9707–9711.
Conrad, M., & Topal, M. D. (1992) *Nucleic Acids Res.* (in press).
Corbett, A. H., Zechiedrich, E. L., & Osheroff, N. (1992) *J. Biol. Chem. 267*, 683–686.
Dodson, M., Roberts, J., McMacken, R., & Echols, H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4678–4682.
Dwyer-Hallquist, P., Kezdy, F. J., & Agarwal, K. L. (1982) *Biochemistry 21*, 4693–4700.
Echols, H. (1986) *Science 233*, 1050–1056.
Forsblom, S., Rigler, R., Ehrenberg, M., Pettersson, U., & Philipson, L. (1976) *Nucleic Acids Res. 3*, 3255–3269.
Gellert, M., & Nash, H. (1987) *Nature 325*, 401–404.
Jiricny, J., & Martin, D. (1986) *Nucleic Acids Res. 14*, 1943–1949.
Jiricny, J., Wood, S. G., Martin, D., & Ubasawa, A. (1986) *Nucleic Acids Res. 14*, 6579–6590.
Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–899.
Kim, Y., Grable, J. C., Love, R., Greene, P. J., & Rosenberg, J. M. (1990) *Science 249*, 1307–1309.
Knight, J. D., Li, R., & Botchan, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3204–3208.
Krüger, D. H., Barcak, G. J., Reuter, M., & Smith, H. O. (1988) *Nucleic Acids Res. 16*, 3997–4008.
Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776–786.
Li, R., Knight, J. D., Jackson, S. P., Tjian, R., & Botchan, M. R. (1991) *Cell 65*, 493–505.
Lu, A.-L., Jack, W. E., & Modrich, P. (1981) *J. Biol. Chem. 256*, 13200–13206.
Mastrangelo, I. A., Courey, A. J., Wall, J. S., Jackson, S. P., & Hough, P. V. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 5670–5674.
Maxwell, A., & Gellert, M. (1984) *J. Biol. Chem. 259*, 14472–14480.
McClarin, J. A., Fredrick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1986) *Science 234*, 1526–1541.

McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238–7245.
Modrich, P., & Rubin, R. A. (1977) *J. Biol. Chem. 252*, 7273–7278.
Modrich, P., & Roberts, R. J. (1985) in *Nucleases* (Linn, S. M., & Roberts, R. J., Eds.) pp 109–154, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Moitoso de Vargas, L., Pargellis, C. A., Hasan, N. M., Bushman, E. W., & Landy, A. (1988) *Cell 54*, 923–929.
Moitoso de Vargas, L., Kim, S., & Landy, A. (1989) *Science 244*, 1457–1461.
Oller, A. R., Vanden Broek, W., Conrad, M., & Topal, M. D. (1991) *Biochemistry 30*, 2543–2549.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804–808.
Segel, I. H. (1975) *Enzyme Kinetics*, pp 125–129 and 400, Wiley, New York.
Su, W., Jackson, S., Tjian, R., & Echols, H. (1991) *Genes Dev. 5*, 820–826.
Taylor, J. D., & Halford, S. E. (1992) *Biochemistry 31*, 90–97.
Thomas, M., & Davis, R. W. (1975) *J. Mol. Biol. 91*, 315–328.
Topal, M. D., Thresher, R. J., Conrad, M., & Griffith, J. (1991) *Biochemistry 30*, 2006–2010.
Travers, A. A. (1989) *Annu. Rev. Biochem. 58*, 427–452.
Voigt, J. M., & Topal, M. D. (1990) *Biochemistry 29*, 1632–1637.