Alteration of Sequence Specificity of the Type II Restriction Endonuclease HincII through an Indirect Readout Mechanism*§

Hemant K. Joshi‡, Christopher Etzkorn‡, Lorentz Chatwell‡, Jurate Bitinaite§, and Nancy C. Horton‡1

From the 1Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, Arizona 85721 and §New England Biolabs Inc., Beverly, Massachusetts 01915

The functional and structural consequences of a mutation of the DNA intercalating residue of HincII, Q138F, are presented. Modeling has suggested that the DNA intercalation by Gln-138 results in DNA distortions potentially used by HincII in indirect readout of its cognate DNA, GTYRAC (Y = C or T, R = A or G) (Horton, N. C., Dorner, L. F., and Perona, J. J. (2002) Nat. Struct. Biol. 9, 42–47). Kinetic data presented here indicate that the mutation of glutamine 138 to phenylalanine (Q138F) results in a change in sequence specificity at the center two base pairs of the cognate recognition site. We show that the preference of HincII for cutting, but not binding, the three cognate sites differing in the center two base pairs has been altered by the mutation Q138F. Five new crystal structures are presented including Q138F HincII bound to GTTTAAC and GTCGAC both with and without Ca2+ as well as the structure of wild type HincII bound to GTTTAAC. The Q138F HincII/DNA structures show conformational changes in the protein, bound DNA, and at the protein-DNA interface, consistent with the formation of adaptive complexes. Analysis of these structures and the effect of Ca2+ binding on the protein-DNA interface illuminates the origin of the altered specificity by the mutation Q138F in the HincII enzyme.

High resolution x-ray crystal structures of protein-DNA complexes have revealed a rational basis for sequence specificity in the form of direct contacts between amino acid side chains and DNA bases in the major groove (1, 2). However, cases have been reported where specificity of a protein for a particular base pair (or base pairs) within its recognition sequence exists even when no direct contacts are seen between the protein and DNA (3–11). In these cases the specificity exhibited by the DNA-binding protein appears to be manifested by indirect means. This “indirect readout” of the DNA sequence may involve water-mediated contacts, contacts to the sugar-phosphate backbone, or the utilization of sequence-dependent DNA structure and/or deformability (3, 12).

The crystal structure of the type II restriction endonuclease HincII reveals distortion from canonical B-form of its bound cognate DNA (13). HincII recognizes and cleaves the sequence GTYYAC, having degeneracy in its specificity for the center two nucleotides by allowing the presence of any pyrimidine-purine, i.e. TA, TG, CA, or CG. Sequence-specific contacts between the protein and DNA can explain this specificity. The contact to the center YR step (CG in the original crystal structure (13)) consists of a hydrogen bond from Asn-141 to the N7 of the purine base. Such a contact can occur when either G or A, but not C or T, occur in the fourth position of the recognition sequence. No contacts are made to the pyrimidine base. Typically, the high degree of specificity achieved by type II restriction endonucleases is due in part to the fact that contacts occur to both bases of each base pair. The question remains of whether or not one hydrogen bond to each purine of the center base pairs can explain the HincII specificity against sequences containing GTYYAC, GTRYAC, or GTRRAC.

The crystal structure of wild type HincII bound to one of the three cognate DNA sequences, GTCGAC, shows that HincII distorts the bound DNA at three loci, two as a result of intercalation of the Gln-138 side chain into the DNA duplex just outside of its recognition sequence (Fig. 1A, light brown space-filling atoms = the recognition sequence, dark brown space-filling atoms = flanking DNA), and the third site at the center YR base pair (Fig. 1, A, red = Y (C or T), and blue = R (A or G), B, right). A hypothesis has been put forth that links the distortion of DNA at the center CG to indirect readout of this sequence by HincII. Furthermore, modeling of the DNA distortion revealed a connection between the Gln-138 intercalation and the center base pair distortion three base pairs away (13).

The studies herein describe the characterization of HincII structure and function when glutamine 138 has been substituted with phenylalanine, Q138F. Preliminary velocity measurements showed that Q138F HincII favors TA over CG at the center YR step to a greater extent than wild type. Subsequent DNA binding assays and single turnover cleavage rate measurements revealed that the alteration in specificity toward cognate sequences containing either CG or TA at the center step occurs at the cleavage and not the binding step. The structural studies of Q138F HincII bound to cognate DNA containing either CG

*This work was supported by National Institutes of Health Grant 1 R01 GM066805-01A1 (to N. C. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§The on-line version of this article (available at http://www.jbc.org) contains supplemental material.

The atomic coordinates and structure factors (code 2GIJ, 2GIH, 2GII, 2GIJ, and 2GIE) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1 To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, AZ 85721. Tel.: 520-626-3828; Fax: 520-626-9204; E-mail: nhorton@u.arizona.edu.
Indirect Readout of DNA by HincII Endonuclease

Experimental Procedures

Mutagenesis—To introduce Q138F mutation into position 414 nucleotides of the hincIIR gene, a novel technique referred to as USER (uracil-specific excision reagent) Friendly Mutagenesis Method was employed. The hincIIR gene was amplified as two overlapping PCR fragments using primer pairs, Pstart + P1 and P2 + Pstop; Pstart = 5'-TATACATATGAGTTTCAAAAACCTATTTCACGG; P1, 5'-AGCTGATTUACTAATATCTTGTGACATC; P2, 5'-AAATCCAGUTTGCACCAATATTATTTACGTACATATAAATTAGC; Pstop, 5'-TGAAACCGCTCTTCCGGAAATATATATTTCTTAAATGGCTTAAC.

The overlapping primers, P1 and P2, were selected in the vicinity of targeted position 414 nucleotides of the hincII gene to allow the introduction of a desired codon change (CAA to TTG) in the primer P2 sequence (marked as bold). Each overlapping primer contained one deoxyuridine residue that flanked the overlap region on the 3' side. 420- and 380-bp fragments of hincIIR gene were amplified with AmpliTaq DNA polymerase (Applied Biosystems) following the manufacturer’s recommendations as follows: initial denaturation at 94 °C for 5 min, denaturation at 94 °C for 10 s, annealing at 55 °C for 1 min, extension at 72 °C for 20 s, final extension at 72 °C for 5 min, for 25 number. PCR products were purified by phenol-chloroform extraction and alcohol precipitation and dissolved in 50 μl of Tris-EDTA buffer (10 mM Tris, pH 7.5, 1 mM EDTA). For a 50-μl assembly reaction, 20 pmol (20 μl) of each PCR fragment were mixed with 5 μl of 10× T4 DNA ligase buffer, 1 μl of 10 mM dNTPs, and 1 μl (5 units) of DNA polymerase I large (Klenow) fragment. The reaction was incubated for 10 min at 37 °C to remove unpaired nucleotides from the 3'-termini of PCR fragments added by TaqDNA polymerase. The Klenow Fragment was then inactivated by incubating for 20 min at 80 °C. 1 μl (1 unit) of USER enzyme (New England Biolabs) was added to the reaction and incubated for 15 min at 37 °C.

The USER enzyme is a mixture of two DNA glycosylases, uracil DNA glycosylase and Endo-VIII DNA glycosylase/AP lyase. Uracil DNA glycosylase catalyzes the excision of a uracil base, forming an abasic (apyrimidinic) site (18, 19), whereas lyase activity of Endo-VIII breaks the phosphodiester backbone at the 3' and 5' sides of the abasic site (20, 21). After phosphodiester bond breakage, the terminal 8-mer oligonucleotides dissociate, leaving PCR products flanked with complementary 3' single-stranded extenstions (Fig. 1). Next, 1 μl (400 units) of T4 DNA ligase was added, and fragments were ligated for 15 min at room temperature. T4 DNA ligase was inactivated by incubating for 20 min at 80 °C. Ligated 780-bp PCR product carrying the complete hincIIIR/Q138F gene was then digested with NdeI and SapI restriction endonucleases and subcloned into NdeI-SapI-cleaved dephosphorylated pTXB1 vector (New England Biolabs). The new construct, pLC/HincIIIQ138F, was introduced into Escherichia coli ER2566 strain carrying the pLysS and pSYX20/HincIIIM plasmids.

2 J. Bitinaite, manuscript in preparation.
Protein Purification—Purification of wild type and mutant HincII was performed as described (22). Purified enzyme was dialyzed exhaustively against buffer A (50 mM bistris propane, pH 7.5, at room temperature, 140 mM NaCl, 1 mM DTT) made 50% in glycerol, flash-frozen in liquid nitrogen, and stored at −80 °C until just before use. Purity has been estimated from overloaded SDS-PAGE (stained with Coomassie Blue) as greater than 99.6 and 98.5% for Q138F and wild type HincII, respectively.

Preparation of DNA Substrates—Plasmids containing a single HincII cleavage site, GTCGAC, GTTGAC/GTCAAC, or GTTAAAC were prepared using Qiagen miniprep kits (Qiagen) using pUC19, pACYC177, or a modified pUC19 with GTTAAAC in place of the SalI/HincII site, respectively. Each plasmid was cleaved with either BsrFI (both pUC19 plasmids) or XhoI (pACYC177) before the multiple turnover assays such that the HincII site was located approximately in the center of the linear DNA.

Both 6-hexachlorofluorescein (HEX)-labeled and non-HEX-labeled oligonucleotides were obtained from commercial synthetic sources and purified using C18 reverse phase high performance liquid chromatography (23). The concentration of each single strand was then measured spectrophotometrically, with extinction coefficients calculated from standard values for the nucleotides (24) and 32.2 mM cm−1 (at 260 nm) for the HEX molecule. Equimolar amounts of each strand (typically 0.5–2 mM each) were mixed and annealed by heating to 90 °C for 10 min followed by slow-cooling to 4 °C over 4–5 h in an Eppindorf thermocycler. Radiolabeled oligonucleotides were prepared using [γ-32P]ATP (MP Biomedicals, Irvine, CA) and T4 polynucleotide kinase (Fermentas, Hercules, CA). Excess [32P]ATP was removed using Micro Bio-spin 6 columns P-polyacrylamide P-6 gel (Bio-Rad).

DNA substrates are shown below, with the recognition sequence in bold, center step underlined, and cleavage site marked by the following sequences.

CG 5′-GGG CCG GTC ↓ GAC CAA C-3′
3′-CCC GGC

TA 5′-GGG CCG GTT ↓ CTG GTT G-5′
3′-CCC GGC CAA ↓ TTG GTT G-5′

SEQUENCES 1–4

Multiple Turnover Velocity Measurements—Linearized plasmid DNA (2 nM) was incubated at 37 °C with 0.6 nM concentrations of wild type HincII or Q138F HincII in 80 μl of reaction buffer (50 mM Tris-HCl, pH 7.9, at 25 °C, 10 mM MgCl2, 100 mM NaCl, 1 mM DTT). 10-μl samples were withdrawn at the indicated times from the reaction, quenched by the addition of 5 μl of quench solution (60 mM EDTA, 0.2% SDS, 50% glycerol, 0.02% bromphenol blue), and subjected to electrophoresis on 1% agarose-gel containing ethidium bromide (3 ng/ml) (Fig. 2A). The density of each band was quantitated from the fluorescence of ethidium bromide using NIH ImageJ Version 1.31 program. The product (cleaved DNA) concentration was then calculated using a calibration curve constructed from the known concentrations of pUC19 DNA cleaved with BsrFI and HincII. The rates of cleavage, or multiple turnover velocities, were measured as the slope of product concentration versus time of incubation (Fig. 2, B and C).

The abbreviations used are: DTT, dithiothreitol; HEX, hexachlorofluorescein; r.m.s.d., root mean square deviation(s); wt, wild type; CSPS, cross-strand purine stack.

FIGURE 2. Multiple turnover studies of cleavage by wild type or mutant HincII (0. 6 nM) and DNA containing the cognate sequence (2 nM initially). A, representative gel showing the time course for wild type HincII reaction on GTTAAAC substrate used to measure the rate of product formation with time in a multiple turnover assay. Uncleaved substrate and cleaved product DNA are marked. The length of incubation in minutes before quenching is shown above each lane. B, a plot of cleaved product formed with time for wild type HincII and the three cognate substrates: GTTAAC (squares), GTTGAC/GTCAAC (circles), GTTAAAC (triangles). C, as in B but with Q138F HincII.

23854 JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 281 • NUMBER 33 • AUGUST 18, 2006
Binding Assays—The dissociation constant $K_D$ was measured against the 5'-$^{32}$P-end-labeled 16-bp DNA containing the relevant sequence by a gel shift assay (10, 15, 25–27). The binding buffer was composed of mixing buffer B with buffer C until the conductivity (19.00 millisiemens/cm) of buffer A was achieved. Buffer compositions were as follows: buffer B, 50 mM bistris propane (pH 7.5 at 25 °C), 500 mM NaCl, 5 mM CaCl$_2$, 1 mM DTT; buffer C, 50 mM bis-tris propane (pH 7.5 at 25 °C), 5 mM CaCl$_2$, 1 mM DTT. Protein concentration (in terms of monomers) was measured spectrophotometrically using the extinction coefficient $\epsilon = 36,840$ cm$^{-1}$M$^{-1}$ at 280 nm. The calculated concentration was then adjusted for molarity of dimers. Protein and DNA were incubated at 37 °C for 30 min, rapidly mixed with prewarmed loading buffer composed of binding buffer and 20% glycerol, and immediately loaded onto a 10% native PAGE gel undergoing electrophoresis. The running buffer was composed of 89 mM Tris, 89 mM boric acid, 5 mM CaCl$_2$, pH 8.3. The gels were run at 190 V and 4 °C with recirculation of the running buffer. Gels were autoradiographed using a phosphorimage plate (GE Healthcare) (Fig. 3A). The data were fit by nonlinear regression (Fig. 3B) to Equation 1 (derived from the quadratic solution to the equilibrium association binding equation).

$$A = A_{\text{min}} + (A_{\text{max}} - A_{\text{min}})[(P_T + O_T + K_D) - [(P_T + O_T + K_D)^2 - (4P_TO_T)]^{1/2}]/(2O_T) \quad \text{(Eq. 1)}$$

where $A$ represents the intensity of the protein bound DNA band. Both $A_{\text{max}}$ and $A_{\text{min}}$ are fit in addition to $K_D$. The DNA concentration, $O_T$, was held constant, and $P_T$ is the total protein concentration.

![Image of gel and graph](image-url)
Indirect Readout of DNA by HincII Endonuclease

concentration, is the independent variable, varying with each measurement of Δ. Each reported K_{D} is the average of at least three independent measurements.

Single Turnover Reactions—HEX or 32P-labeled DNA containing the cognate sequence GTCGAC or GTTAAAC at concentrations ranging from 0.62 to 77 nM were incubated with wild type or Q138F HincII at a concentration of 8–5000-fold molar excess of protein to DNA. The reaction conditions included a buffer prepared by mixing buffer D with buffer E until the conductivity of buffer A’ was achieved. Buffer concentrations were as follows: buffer D, 50 mM bis-tris propane (pH 7.5 at 25 °C), 10 mM MgCl₂, 0.1 mg/ml bovine serum albumin, 1 mM DTT; buffer E, 50 mM bis-tris propane (pH 7.5 at 25 °C), 10 mM MgCl₂, 0.1 mg/ml bovine serum albumin, 1 mM DTT, 500 mM NaCl; buffer A’, 50 mM bis-tris propane (pH 7.5 at 25 °C), 0.1 mg/ml bovine serum albumin, 1 mM DTT, 140 mM NaCl. This procedure for preparing the reaction buffer was used to ensure constant ionic strength between various binding and reaction buffers having different concentrations of Ca^{2+} or Mg^{2+}. Initiation of the reaction was performed by mixing preheated (at 37 °C) enzyme and DNA followed by incubation of the reaction at 37 °C. At various indicated times after initiation, 10 μl of the reaction were removed, and the reaction was quenched by the addition of 10 μl of 4 M urea, 50 mM EDTA, 80% formamide. The quenched reaction aliquots were then separated on a 20% acrylamide denaturing PAGE gel (containing 8M urea) in 1× TBE (89 mM Tris, 89 mM boric acid, 10 mM EDTA, pH 8.3). A Typhoon scanner (GE Healthcare) was used to detect the HEX-DNA within the gels or the bands on a phosphorimage plate after exposure to the gel containing 32P-labeled DNA. The product bands were then placed on 1 ml of the precipitating solution. Crystals of mutant HincII were prepared by cocrystallization ner identical to those containing CG at the center step (22).

Single turnover cleavage rate constants were determined for each strand (Fig. 3A). The single turnover cleavage rate constants are the independent variable, varying with each measurement of Δ. Each reported K_{D} is the average of at least three independent measurements. The cleavage site is off-center in the 16-base pair substrate, and the products of cleavage of both strands are distinguishable on the denaturing PAGE gel (Fig. 3C). The single turnover cleavage rate constants were determined for each strand (Fig. 3D).

Crystallization, Diffraction Data Collection, and Structure Refinement—Crystals of wild type HincII bound to cognate DNA containing TA at the center step were prepared in a manner identical to those containing CG at the center step (22). Crystals of mutant HincII were prepared by cocryocrystallization with two different recognition sequences, GTTAAC or GTCGAC. Both DNA substrates were 14 base pairs in length and 5’GAC. Both DNA substrates were 14 base pairs in length and Crystals of mutant HincII were prepared by cocrystallization ner identical to those containing CG at the center step (22).

Crystals of wild type HincII bound to cognate DNA-containing TA, the previously solved isomorphous structure of HincII bound to cognate DNA-containing CG was used beginning with rigid body refinement as implemented in CNS (30) followed by iterative model rebuilding using XtalView (31), simulated annealing, temperature factor, and positional refinement using CNS (30). All structures of Q138F HincII bound to DNA with and without Ca^{2+} are nonisomorphous with the previously solved wild type HincII/cognate DNA structure. Three (HQ9, HQ65, HQ61; Table 3) have space group P_{2}₁ symmetry and are isomorphous. The binding of Ca^{2+} by the Q138F/CG crystals changes the space group symmetry from P_{2}₁ to P_{4} in data set HQ51. The HQ9 (Q138F/CG) structure was solved by molecular replacement using a monomer from the wild type DNA-bound form (1KC6, wtHincII/CG) as a search model and the program CNS (30). The same solution was used as starting models for HQ61 and HQ65. The HQ51 structure was solved independently also with the wild type HincII bound to CG DNA (1KC6) monomer and MOLEP (32). Models were then refined using CNS (30) and rigid body refinement, simulated annealing, temperature factor refinement, and manual model building using XtalView (31).

Analysis of DNA Structure—Root mean square deviation (r.m.s.d.) were calculated with LSQKAB (33).

Analysis of Protein Conformation—r.m.s.d. were calculated using LSQKAB (33). Parts of the protein molecules which move as rigid bodies were determined using the difference distance matrix method implemented in DDMP. Rotations and translations of parts of a structure relative to another were calculated using LSQKAB (33). Rotation angles describing the opening up or closing down of one dimer relative to another are calculated by first superimposing one subunit using all Ca atoms, then calculating the rotation angle necessary to align the second subunit.

RESULTS

Multiple Turnover Velocity Measurements—An example of a gel used to measure the amount of cleaved DNA as a function of the length of incubation with HincII is shown in Fig. 2A. The plots of product versus time for wild type HincII or Q138F

4 Difference distance matrix plots were produce using the DDMP program from the Center for Structural Biology at Yale University, New Haven, CT.
HincII with each of the three cognate sequences are shown in Fig. 2, B and C, respectively, and the resulting rates are provided in Table 1. Multiple rounds of cleavage take place as evidenced by the linear production of cleaved DNA with time. From the appearance of a single cleavage in each plasmid bearing one of the three cognate HincII recognition sites, it is clear that the endonuclease activity of Q138F HincII is specific for the HincII recognition sites. However, Q138F HincII is reduced in overall cleavage activity and presents an altered preference for the sequence at the center step of the six-base pair recognition sequence. Although wild type HincII prefers TA(TG/CA):CG at the center step on the order of 6:2:1, the mutant, Q138F HincII, prefers the same sequences on the order 36:8.1. Thus, the mutant prefers TA relative to CG 6-fold more than does the wild type HincII protein.

**Binding Measurements**—The measurement of binding affinities of the two proteins, wild type and Q138F HincII, for the cognate sequences containing either TA or CG at their center step was performed using the gel shift assay. The divalent cation Ca$^{2+}$ was included in the incubation and running buffer because numerous studies have shown that this ion inhibits cleavage while increasing the affinity and specificity of the enzyme for its cognate DNA (14, 15, 35). Fig. 3A shows an example of a scan of one gel shift assay, and Fig. 3B shows the fit of the data from Fig. 3A used to determine the $K_D$. The measured $K_D$ values for wild type and Q138F HincII with the cognate sequences containing CG or TA are shown in Table 2, as are the standard deviations determined from three independent binding measurements.

Wild type HincII binds to its cognate DNA containing TA or CG at the center step with a $K_D$ of 0.58–1.14 nm, respectively, indicating a 2-fold preference for the TA substrate. Q138F HincII bound the cognate sequences with roughly the same affinity as the wild type HincII. Therefore, the substitution of phenylalanine for glutamine at the DNA intercalating residue did not appreciably affect either the affinity or sequence preference (TA versus CG) of HincII. In contrast, no binding of CG DNA (1 nm) was detected using the identical gel shift assay without Ca$^{2+}$ in the binding or running buffers with either wild type or Q138F HincII up to a concentration of $1 \times 10^{-6}$ M (supplemental data).

**Single Turnover Cleavage Rate Constant Measurements**—The single turnover DNA cleavage rate measurements were performed with HEX- or $^{32}$P-labeled DNA. To assess any possible effect of the HEX label on HincII activity, a single turnover cleavage rate measurement was performed on $^{32}$P-labeled DNA identical to the HEX-labeled DNA-containing CG at the center step and the Q138F enzyme. Identical rates were observed (data not shown).

The single turnover DNA cleavage rate constants, $k_{\text{cleave}}$, were determined in the presence of an 11–5000-fold molar excess of protein over DNA and either 8.1 $\times 10^{-7}$ M (wild type HincII reactions) or 3.0 $\times 10^{-6}$ M (Q138F HincII reactions) protein. Given the dissociation constants determined above (in the presence of Ca$^{2+}$), these conditions ensure $\geq$99% binding of the DNA (if the binding constants with Mg$^{2+}$ are similar to those with Ca$^{2+}$), with only a single round of cleavage. In all cases the quantified product at specific time points followed a single exponential, indicating the detection of a single step. For all single turnover measurements both with wild type or Q138F HincII, the product release step was not of interest, and the quenching of the reaction with the denaturant urea before electrophoreses ensured that all product was measured. In this way the product release step did not contribute to the $k_{\text{cleave}}$ measurement.

The single turnover cleavage rate constants are given in Table 2 for both wild type and Q138F HincII, with cognate DNA containing either CG or TA at the center step. The reactions were initiated by the mixing of HincII enzyme with DNA at time 0, both in buffer containing Mg$^{2+}$. The mutation impairs the single turnover cleavage rate constant by 37–39-fold in the case of the TA substrate and 84–128-fold in the case of the CG substrate. Although there is a 1.6–2.5-fold difference found between $k_{\text{cleave}}$ of wild type HincII and these two cognate substrates, 5.6-fold differences occur with the mutant.

The data of Table 2 show that HincII exhibits sequence preferences at both the binding and cleavage steps. A specificity constant, $K_D$, was measured for each enzyme/substrate pair, also shown in Table 2. Using this parameter, wild type HincII prefers TA over CG by 3:1. The mutation, Q138F, designed to test the connection between DNA intercalation and sequence specificity, alters this preference to 14:1, differing from the wild type preference by nearly 5-fold (Fig. 3E).

**Overall Structure of Wild Type HincII with TA (wtHincII/TA)**—The structure of wild type HincII bound to cognate DNA containing TA at the center step (GTCTACGTC/ATTCGTTCAT) has been solved to 2.6 A in the $I_22_12_1$ space group (Table 3) and is isomorphous with wild type HincII bound to cognate DNA containing CG (GTCCCTAGC) (13). Like the isomorphous structure, the asymmetric unit contains two protein dimers and two duplexes of DNA. Only two protein residues could not be located in the electron density maps, both occurring at the carboxyl terminus and only in chains B and D. However, all 13 nucleotides of all four DNA strands were visible in the maps. The model also includes 292

---

**Table 1**

Multiple turnover velocity measurements

| Enzyme     | Substrate | Sequence | Velocity $^a$ | Ratio within series | Relative ratio $^b$ |
|------------|-----------|----------|---------------|---------------------|---------------------|
| Wild type HincII | CG/C | GTAGAC | 0.038 ($\pm 0.003$) | 1 |  
| Wild type HincII | TG/C | GCTTGAC | 0.069 ($\pm 0.011$) | 2 |  
| Wild type HincII | TA/TA | GTTTAC | 0.233 ($\pm 0.041$) | 6 |  
| Q138F HincII | CG/C | GTAGAC | 0.004 ($\pm 0.001$) | 1 |  
| Q138F HincII | TG/C | GCTTGAC | 0.032 ($\pm 0.003$) | 8 |  
| Q138F HincII | TA/TA | GTTTAC | 0.143 ($\pm 0.029$) | 36 |  

$^a$ Enzyme velocities were determined as shown in Fig. 2 from at least three independent experiments and are shown as the mean $\pm$ S.D.

$^b$ The relative ratio is calculated from (ratio within series for Q138F HincII)/(ratio within series for wild type HincII).

---

**Indirect Readout of DNA by HincII Endonuclease**
TABLE 2
Dissociation and single turnover cleavage rate constants

| Enzyme       | Substrate | Sequence            | $K_D$ | $k_{on}$ | $k_{cleave}$ | S.D. | $k_{on}$/S.D. | $k_{cleave}$/S.D. |
|--------------|-----------|---------------------|-------|----------|-------------|------|--------------|------------------|
| Wild type HincII | CG/CG     | GTGCCG              | 1.14  | 0.23     | 7.71 x 10^{-2} | 1.61 x 10^{-2} | 6.8 x 10^{-2} |
| Wild type HincII | TA/TA     | GTTTAAC             | 0.58  | 0.07     | 1.25 x 10^{-1} | 0.10 x 10^{-1} | 2.1 x 10^{-1} |
| Q138F HincII  | CG/CG     | GTGCCG              | 1.03  | 0.09     | 6.01 x 10^{-4} | 0.99 x 10^{-4} | 5.8 x 10^{-4} |
| Q138F HincII  | TA/TA     | GTTTAAC             | 0.42  | 0.06     | 3.36 x 10^{-3} | 0.27 x 10^{-3} | 8.0 x 10^{-6} |

$^a$ Using the upper strand $k_{on}$.

$^b$ The total ratio is calculated from (ratio within series for Q138F HincII)/(ratio within series for wild type HincII).

$^c$ Concentration of protein = 8.1 x 10^{-7} M; concentration of HEX-DNA = 7.7 x 10^{-8} M.

$^d$ Concentration of protein = 3.0 x 10^{-4} M; concentration of 32P-labeled DNA = 1.2 x 10^{-4} M.

$^e$ Concentration of protein = 3.0 x 10^{-4} M; concentration of 32P-labeled DNA = 6.4 x 10^{-10} M.

TABLE 3
Crystallographic statistics

|                | Q138F HincII/GTTAAC | Q138F HincII/GTGAC/Ca^{2+} | Q138F HincII/GTTAAC/Ca^{2+} | Wild type HincII/GTTAAC |
|----------------|----------------------|-----------------------------|-----------------------------|------------------------|
| Protein data bank code | 2GIG                 | 2GH                         | 2GII                        | 2GIE                   |
| Code           | HQ9                  | HQ51                        | HQ65                        | H56                    |
| Space group    | P2_1                 | I4_1                        | P2_1                        | I2_1,2_1               |
| Cell           | 50.38 Å              | 127.86 Å                    | 50.38 Å                     | 50.46 Å                |
|                | 90.99 Å              | 127.86 Å                    | 90.04 Å                     | 91.77 Å                |
|                | 67.20 Å              | 81.23 Å                     | 68.01 Å                     | 66.54 Å                |
| Resolution     | 1.83 Å               | 2.3 Å                       | 2.1 Å                       | 2.6 Å                  |
| Total observations | 167.967              | 199.662                     | 478.673                     | 537.523                |
| Unique observations | 47.047              | 28.688                      | 32.583                      | 45.628                 |
| % Complete     | 90.7% (90.7%)        | 98.9% (100.0%)              | 99.9% (99.6%)               | 99.7% (100%)           |
| %sigma         | 9.1 (2.0)            | 38.9 (2.8)                  | 24.78 (2.0)                 | 31.5 (2.0)             |
| Multiplicity   | 3.6                  | 7.0                         | 1.9                         | 3.7                    |
| $R_{merge}^a$  | 4.6% (31.8%)         | 8.6% (70.3%)                | 5.1% (54.5%)                | 4.8% (52.5%)           |
| $R_{merge}^b$  | 18.8%                | 19.7%                       | 23.9%                       | 23.3%                  |
| $R_{free}^c$   | 22.2%                | 28.0%                       | 28.2%                       | 26.3%                  |
| Overall B factor (Å^2, Wilson plot) | 26.7 | 65.4 | 50.0 | 36.4 |
| r.m.s.d. bonds | 0.006                | 0.008                       | 0.007                       | 0.006                  |
| r.m.s.d. angles | 1.30                 | 1.44                        | 1.42                        | 1.19                   |
| Number of waters | 466                 | 365                         | 381                         | 330                    |
| Number of Ca^{2+} | 0                     | 2                           | 0                           | 2                     |
| Protein residues | A2–A23              | A2–A258                     | A2–A19                      | A2–A22                 |
|                | A31–A257             | B2–B22                      | A35–A175                    | A35–A258               |
|                | B2–A22              | B29–B258                   | A180–A258                   | B2–B18                 |
|                | B37–B257            | B2–B16                      | B37–B174                    | B2–D257               |
|                |                     | B35–B100                    | B177–B257                   | B2–B18                 |
|                |                     | B104–B174                  | B182–R256                   | C2–C258               |
| Nucleotides    | E1–E14              | F1–F14                      | E1–E14                      | F1–F14                 |
|                | E1–E14              | F1–F14                      | E1–E14                      | F1–F14                 |
|                | G1–G14              | H1–H14                      |

$^a$ $R_{merge} = \frac{\sum_{hk} |F_{calc} - F_{obs}|}{\sum_{hk} F_{obs}}$, where $F_{calc}$ is the observed and calculated structure factor amplitude for all reflections. The sum is carried out over the 95% of the observed reflections that are used in refinement.

$^b$ $R_{merge} = \frac{\sum_{hk} |F_{calc} - F_{obs}|}{\sum_{hk} F_{calc}}$, where $F_{calc}$ is the observed and calculated structure factor amplitude for all reflections. The sum is carried out over the 95% of the observed reflections that are used in refinement.

$^c$ $R_{free}$ refers to the $R$ factor for the test reflection set (5% of the total observed) that was excluded from refinement.

Indirect Readout of DNA by HincII Endonuclease

The structures of Q138F HincII bound to two cognate DNA containing CG at the center step and that containing TA—The structures of Q138F HincII bound to GTCGAC (Q138F/TA), and Q138F HincII Bound to GTGCAG (Q138F/CG)—The structures of Q138F HincII bound to two of the three cognate DNA sequences (containing TA or CG at the center step of GTYRAC) have been solved to 2.30 and

water molecules and has been refined to an $R_{cryst}$ of 22.6% and $R_{free}$ of 29.9%. The overall tertiary structure is relatively unchanged due to the substitution of the center step sequence, having a r.m.s.d. for each monomer (Ca atoms of residues 2–257) when compared with wild type HincII bound to CG DNA (wtHincII/CG) of 0.43 and 0.59 Å. Similarly, the conformation of the DNA is little changed, having a r.m.s.d. of 0.31 Å compared with the wtHincII/CG structure using all atoms of the six base pairs of the recognition sequence. The center YR = TA step of the cognate DNA contains the same unusual conformation seen in the structure of wtHincII/CG (Fig. 1B, right). Looking at quaternary conformation, the r.m.s.d. when using the whole protein dimer for the superposition onto wtHincII/CG, is 0.53 Å, and that of the bound DNA using this superposition is 0.33 Å, showing little shifting of the DNA relative to the protein dimer. These numbers show that there is little difference between the structure of wild type HincII bound to cognate DNA containing CG at the center step and that containing TA.
1.83 Å, respectively. These structures are isomorphous with each other in space group P2₁, but are nonisomorphous with the two cognate DNA-bound wild type HincII structures, which form crystals of the I2₁2₁2₁ space group. The DNA used in the Q138F HincII crystals is longer (14 nucleotides per strand) than that in the wild type HincII structure (13 nucleotides per strand). Crystals of Q138F HincII with the 13 nucleotide DNA did not diffract, and therefore, other crystal forms were sought using different DNA lengths (11–15 nucleotides per strand), resulting in the identification of the P₂₁ form. The DNA of the protein–DNA complex is often found to make lattice contacts, and therefore, it is not unusual for different crystal forms to be found by using DNA of differing lengths (22, 36). The asymmetric unit of the Q138F/cognate DNA structures contains only a single dimer and a single duplex of DNA. The Q138F/TA and Q138F/CG structures have been refined to 

\[ \text{R}_{	ext{cryst}} \] of 23.2 and 17.7% and 

\[ \text{R}_{	ext{free}} \] of 28.9 and 22.0%, respectively.

Table 4 shows r.m.s.d. of the Q138F/DNA structures after superposition with wtHincII/CG/Ca²⁺ (the Ca²⁺-bound form of the wild type HincII bound to CG DNA was used in these calculations for consistency, and Ca²⁺ binding changed the wtHincII/CG structure little (16)). The coordinate error (using the Luzzati method (37) as implemented in CNS (30)) for these structures is 0.4 and 0.2 Å; therefore, these large r.m.s.d. indicate the presence of significant conformational changes within the structures. High r.m.s.d. are evident at the monomeric as well as at the dimeric level, indicating a change both within as well as between the two monomers of the dimers. The r.m.s.d. of both subunits A and B as well as in the dimer are greater in the structure with TA (Q138F/TA) than that with CG (Q138F/CG). Significant conformational differences are also evident in the bound DNA. Judging by the lower r.m.s.d., it appears that the core 6 base pairs of the CG DNA bound to Q138F HincII is more like the CG DNA bound to wild type HincII than the TA DNA bound to Q138F HincII, as is apparent in viewing the superposition (see Fig. 5A, blue is Q138F/CG, gray is wtHincII/CG/Ca²⁺, red is Q138F/TA). Because very little difference is found in the wild type HincII-bound DNA whether CG or TA or with or without Ca²⁺, the larger r.m.s.d. in the DNA of the Q138F/TA structure is not an inherent property of the DNA sequence alone or due to the bound Ca²⁺. We can also look at the r.m.s.d. of the DNA when the protein dimer is used for the superposition onto wtHincII/CG/Ca²⁺, which gives additional information as to how differently (relative to wtHincII/CG/Ca²⁺) the DNA sits in the Q138F HincII dimer, and find that it is significant in both mutant structures. However, it is much greater for the core six base pairs of TA bound to Q138F HincII than for the CG DNA. Again, the larger numbers indicate a mis-positioning of the DNA within the dimer that is greater in Q138F/TA than Q138F/CG. Thus, the high r.m.s.d. in these two structures, Q138F/TA and Q138F/CG, indicates numerous differences (from wild type) in conformation, and in addition, in all measures of r.m.s.d., the Q138F/TA structure contains the greatest deviations from wild type HincII bound to cognate DNA.

**Overall Structure; Q138F HincII Bound to GTTAAC with Ca²⁺ (Q138F/TA/Ca²⁺), and Q138F HincII Bound to GTC-GAC with Ca²⁺ (Q138F/CG/Ca²⁺)—** Structures of Q138F HincII bound to DNA and Ca²⁺ were prepared by soaking the Q138F/DNA crystals in 50 mM CaCl₂ for 5 h. The Q138F/TA/

| Structure          | Monomer A | Monomer B | Whole dimer DNA | Protein-DNA interface |
|--------------------|-----------|-----------|-----------------|-----------------------|
| Q138F/TA           | 0.81 Å    | 1.22 Å    | 2.02 Å          | 1.15 Å 1.48 Å        |
| Q138F/TA/Ca²⁺      | 0.60 Å    | 1.09 Å    | 1.45 Å          | 0.67 Å 0.75 Å        |
| Q138F/CG           | 0.53 Å    | 1.09 Å    | 1.72 Å          | 0.62 Å 1.17 Å        |
| Q138F/CG/Ca²⁺      | 0.47 Å    | 1.08 Å    | 1.14 Å          | 0.57 Å 0.64 Å        |

* a C₀ of residues 2–20 and 35–258 of this subunit.
* b C₀ of residues 2–20 and 35–258 of both subunits.
* All atoms of the core 6-bp recognition sequence.
* r.m.s.d. of the 6-bp recognition sequence after superposition with residues 31–257 of both subunits of the protein dimer.

**FIGURE 4.** Stereo diagram of simulated annealing omit electron difference density (Fo − Fc) around the active site of subunit B of the Q138F HincII/TA/Ca²⁺ structure (A) and or subunit B of the Q138F/CG/Ca²⁺ structure (B), contoured at 3σ (gray) and 6σ (blue).
Ca\(^{2+}\) structure is isomorphous with the unsoaked crystals, whereas calcium ion binding within the Q138F/GC/Ca\(^{2+}\) structure changed the space group symmetry from P2\(_1\) to I4\(_2\). The structures have been refined to 1.95 and 2.3 Å with R\(_{\text{cryst}}\) of 23.8 and 19.0% and R\(_{\text{free}}\) of 27.7% and 28.3% for Q138F/TA/Ca\(^{2+}\) and Q138F/GC/Ca\(^{2+}\), respectively. Representative omit electron density is shown for both structures in Fig. 4.

The r.m.s.d. of different parts of these structures when compared with the wtHincII/GC/Ca\(^{2+}\) structure is given in Table 4. The r.m.s.d. of residues in each monomer have either changed very little or decreased in response to binding Ca\(^{2+}\) in both mutant HincII structures, Q138F/TG/Ca\(^{2+}\) and Q138F/GC/Ca\(^{2+}\). Significantly, the r.m.s.d. of the whole dimer has decreased significantly in response to binding Ca\(^{2+}\) (Table 4, compare the r.m.s.d. of whole dimer of Q138F/TG/Ca\(^{2+}\) to Q138F/TG and Q138F/TG/Ca\(^{2+}\) to Q138F/TG). Thus, on first inspection, the effect of Ca\(^{2+}\) binding appears to have brought the protein structure, especially that of the protein dimer (last column, Table 4), closer to that of wild type. In addition, the structure containing CG at the center step, Q138F/TG/Ca\(^{2+}\), is closer to the wild type structure at both the monomer and dimer level. Looking at the DNA, the r.m.s.d. of all atoms of the six base pairs of the recognition sequence has decreased in the mutant structure bound to TA in response to binding Ca\(^{2+}\) but has increased in the structure with CG on binding Ca\(^{2+}\). Thus, the effect of Ca\(^{2+}\) binding to the mutant protein/DNA complexes has had opposite effects on the two different DNA sequences, making the TA DNA appear more like that bound to wild type HincII but making the CG DNA look less like it (Fig. 5, A–D). Looking at the way the DNA sits in the protein dimer, namely the r.m.s.d. of the DNA after superposition of the protein dimer, these values have decreased relative to the corresponding structures without Ca\(^{2+}\); therefore, the Ca\(^{2+}\) has had the effect of positioning both DNA sequences within the Q138F HincII dimer much more similarly to that within wild type HincII (see Fig. 9). In summary, we find that Ca\(^{2+}\) binding has had the effect of bringing both the dimer conformation and the positioning of the DNA within the dimer much closer to that found in wtHincII/GC/Ca\(^{2+}\) in the mutant structures with either DNA sequence. In addition, it appears to bring the core six base pair recognition sequence of DNA of the TA DNA closer to that of the DNA bound to the wild type enzyme. The dimer conformation still deviates more from that of wild type with TA DNA than with CG, even with bound Ca\(^{2+}\). The r.m.s.d. of the DNA itself and of its positioning also shows that it is still greater in the mutant structures bound to TA DNA than CG DNA (i.e. in Q138F/TG/Ca\(^{2+}\) compared with Q138F/GC/Ca\(^{2+}\)). These observations will be important below because it appears that the distortion of the protein dimer as a result of loss of the center step base stacking in the TA DNA allows for better recovery of the alignment of the active sites in Q138F/TG/Ca\(^{2+}\) as compared with those in Q138F/GC/Ca\(^{2+}\).

**DNA Structure**—Comparison of the structures of the six base pair recognition site of DNA from the Q138F/DNA structures to that bound to wtHincII/GC/Ca\(^{2+}\) is shown in Fig. 5. The superposition of these base pairs from the mutant structures without Ca\(^{2+}\) onto that from wtHincII/GC/Ca\(^{2+}\) is shown in Fig. 5A. It is quite apparent here that the DNA from Q138F/TG (red) possesses much greater deviation from that bound to wtHincII/GC/Ca\(^{2+}\) (gray) than that from Q138F/GC (dark blue). The deviations are especially apparent in the center two base pairs. It appears that stress on the protein-DNA complex as a result of the mutation produced distortion of the TA DNA much more than the CG DNA. Fig. 5B shows the superposition of the DNA from the Ca\(^{2+}\)-bound mutant structures, Q138F/GC/Ca\(^{2+}\) (light blue) and Q138F/TG/Ca\(^{2+}\) (pink) onto that of wtHincII/GC/Ca\(^{2+}\) (gray). The TA DNA now resembles that bound to wild type much more closely than without Ca\(^{2+}\) (Fig. 5A, red). Thus, Ca\(^{2+}\) “corrects” the DNA structure as well. Fig. 5, C–D, shows the r.m.s.d. for atoms of each nucleotide when all atoms are used for the superposition. Fig. 5C shows the r.m.s.d. of only backbone atoms, whereas Fig. 5D shows the r.m.s.d. for only the base atoms. The DNA backbone r.m.s.d. (Fig. 5C) shows a systematic trend of increasing differences from nucleotides 8 to 10 in structures Q138F/GC (dark blue), Q138F/GC/Ca\(^{2+}\) (light blue), Q138F/TG/Ca\(^{2+}\) (pink) and increasing differences from nucleotides 5 to 9, then falling slightly at nucleotide 10, in Q138F/TG (red). The r.m.s.d. are much larger at nearly all nucleotides in the Q138F/TG structure (red, Fig. 5, C–D). The intercalation site and the location of the mutation Q138F is between nucleotides 10 and 11. Thus, the larger deviations in the DNA backbone structure near nucleotide 10 appear to be a result of the mutation and changes in contacts at the protein-DNA interface (described in more detail below). The r.m.s.d. of the base atoms (Fig. 5D) show a similar pattern with respect to each structure; however, the greatest deviations in Q138F/TG (red) occur at the center two base pairs (nucleotides 7–8), as seen in the superposition (Fig. 5A).

The unusual distortion of cognate DNA bound to wild type HincII at the center two base pairs (GTYRAC) is shown in Fig. 1B, right. For comparison, the same DNA, but with B-form conformation, is shown in Fig. 1B, left. When the DNA is bound to wild type HincII, the pyrimidine base of the center step, Cyt 7, is unstacked from its 3’ neighbor, Gua 8. In addition the center step purines from opposite strands, Gua 8 and Gua 8’, are shifted toward one another, increasing their stacking surface area. This type of conformation is known as a cross-strand purine stack (CSPS). The CSPS was identified in the original structure and may play a role in indirect readout by HincII. Furthermore, modeling studies suggest a connection between DNA intercalation by glutamine 138 (green, Fig. 1A) and the CSPS (13). The conformations of the CG and TA DNA when bound to Q138F HincII are shown in Fig. 6, A–D. The CG DNA bound to wild type HincII is shown in black. The superpositions were done with the base atoms of one of the two purine bases (Pur 8, Fig. 6, A–D). The view is down the helical axis to illustrate the amount of base stacking by the two center step purines (Pur 8 and Pur 8’, Fig. 6, A–D). The DNA closest to that bound to wild type is the CG DNA of Q138F/CX (Fig. 6C), and Ca\(^{2+}\) binding to this complex appears to actually decrease the amount of CSPS (Fig. 6D). The center step purines do not stack at all in Q138F/TG (Fig. 6A) but come closer when Ca\(^{2+}\) is bound (Fig. 6B), although the stacking is still worse than in Q138F/CX (Fig. 6B).

**Local Conformational Effects of the Mutation**—Both tertiary and quaternary conformational changes relative to wild type
HincII bound to cognate DNA are found to varying degrees in the four structures of Q138F HincII with cognate DNA with and without bound Ca\textsuperscript{2+}. This is surprising, because residue 138 extends away from the protein in both the unliganded (38), and DNA-bound structures of wild type HincII (13). These changes differ depending on the sequence of the bound DNA as well as on the presence or absence of Ca\textsuperscript{2+}. A t the r.m.s.d. level, the changes are greatest when comparing whole dimers to that of wild type and especially in the absence of bound Ca\textsuperscript{2+}. The subunits are opened, widening the DNA binding domain. The increased rotation relative to the wtHincII/CG/Ca\textsuperscript{2+} of one subunit relative to the other is 10.9° and 9.2° in Q138F/TA and Q138F/CG, respectively, without bound Ca\textsuperscript{2+}. The binding of Ca\textsuperscript{2+} causes the Q138F HincII dimer to close up around the bound DNA, reducing the subunit rotations (relative to wild type) to 7.0° and 4.0°, respectively. The locations and magnitudes of differences in conformation between two structures can be viewed using

FIGURE 5. Superposition of core 6-bp DNA of wild type HincII/CG/Ca\textsuperscript{2+} (gray) and Q138F/DNA (colored). A, stereo diagram, Q138F/TA (red) and Q138F/CG (blue). B, stereo diagram, Q138F/TA/Ca\textsuperscript{2+} (pink) and Q138F/CG/Ca\textsuperscript{2+} (cyan). C, r.m.s.d. of backbone atoms of DNA with wild type HincII/CG/Ca\textsuperscript{2+} using all atoms of core 6 bp of both DNA strands, Q138F/TA (red), Q138F/TA/Ca\textsuperscript{2+} (pink), Q138F/CG (blue), and Q138F/CG/Ca\textsuperscript{2+} (cyan). D, r.m.s.d. of base atoms of DNA with wild type HincII/CG/Ca\textsuperscript{2+} using all atoms of core 6-bp both DNA strands, Q138F/TA (red), Q138F/TA/Ca\textsuperscript{2+} (pink), Q138F/CG (blue), and Q138F/CG/Ca\textsuperscript{2+} (cyan). Pyr, a pyrimidine; Pur, a purine.
difference distance plots (Fig. 7). These figures highlight changes in the protein conformation relative to the wtHincII/CG/Ca²⁺ structure. They are calculated by first measuring the distances between all Cα atoms within the wtHincII/CG/Ca²⁺ structure and measuring the same distances within a Q138F/DNA structure. These distances are then

FIGURE 6. Stereo diagrams of superpositions using base atoms of purine 8 of wild type HincII/CG/Ca²⁺ (black) showing the central two base pairs of the recognition sequence GTYRAC (underlined) of both strands. A, Q138F/TA (red). B, Q138F/TA/Ca²⁺ (pink). C, Q138F/CG (blue). D, Q138F/CG/Ca²⁺ (cyan). E, Q138F/CG/Ca²⁺ (cyan), Q138F/TA/Ca²⁺ (pink), and wild type HincII/CG/Ca²⁺ (gray) showing the propagation of differences in stacking with the center step purines to the phosphate position of Ade 9. A residue of the active site, Asp-127, as well as the adjacent residue implicated in potential steric conflicts with the DNA, Thr-130, are also shown. Py, a pyrimidine; Pur, a purine.
compared. Large differences in Ca-Ca distances indicate a change in the protein conformation within the mutant structure relative to wild type. The differences in Ca-Ca distances are plotted using colored boxes inside a matrix where the residue number of the first Ca of the pair is specified along the horizontal, and the residue number of the second Ca is specified along the vertical. Blue indicates an increase in the distance between selected Ca atoms relative to the wild type structure, whereas red indicates a decrease. The upper left and lower right quadrants refer to changes in Ca-Ca distances within the same monomeric subunits, subunits A or B, respectively, whereas the upper right and lower left quadrants refer to changes in Ca-Ca distances within the same monomeric subunits, subunits A or B, respectively.
rants show changes in the Ca-Cα distances of the two different subunits in the dimer, A and B and in B and A, respectively.

The difference distance plots of Fig. 7 indicate that the Q138F mutation has caused disruptions in the protein structure that depend on the sequence of the bound DNA, which is further changed by the binding of Ca²⁺. To better understand the origin of these changes, we looked closely at the protein-DNA interface at the site of mutation (Fig. 8). It is apparent from these superpositions that relative to wild type (gray, Fig. 8, A–D), the backbone around residue 138 is pulled back away from the DNA in the Q138F/DNA structures. This appears to allow better stacking between the phenylalanine side chain and the two neighboring cytosines, Cyt 10 and Cyt 11.

Probably as a response to the shifting of the protein main chain at 138, the protein main chain at the neighboring residue Ala-137 is flipped, such that the side chain of Ala-137 no longer points toward the DNA where it had made a van der Waals contact to the deoxyribose of Cyt 10 in the wtHincII/CG/Ca²⁺ structure (Fig. 8). The flipping of Ala-137 occurs in both subunits of all DNA-bound Q138F HincII structures. The similarities within the different subunits of the different DNA-bound structures ends here, with some structures showing a recovery of the protein trace to follow that of wild type, other subunits shift with a structural portion active as a rigid domain, whereas still others shift smaller segments of each subunit. Rotation of a greater portion of the subunit (i.e. rigid body rotation versus local distortion) is more likely to occur in complexes without bound Ca²⁺.

To understand why distortions occur in the structures of Q138F HincII bound to DNA, with and without Ca²⁺, and also why they involve different segments of the protein in each, we took the DNA structure from each Q138F/DNA structure, superimposed it onto the bound DNA of wtHincII/CG/Ca²⁺, and asked where steric conflicts occur. The most consistent conflicts occurred between the side chain of Thr-130 and the phosphate of Ade 9. Interestingly, this did not occur in the DNA bound to the subunits of Q138F HincII that show less subunit rotation relative to wild type (Table 5). The reason for this appears to be a conformation taken on by the phosphate of Ade 9 (Table 5). When the gamma (C5′-C4′) angle of the Ade 9 phosphate is close to 170°, the phosphate appears twisted, or rotated, relative to the wild type conformation (gamma angle of 30°–40° or 90°, Figs. 8 and 9). This rotation moves the nonesterified oxygens of the Ade 9 phosphate away from the side chain of Thr-130. The twisting of the phosphate is present only in subunits with 3.2° or less of rotation (Table 5), which is always the subunit arbitrarily assigned as A in these structures.

To understand why the conformation of the phosphate of Ade 9 should
be affected, we looked closely at the DNA around the site of intercalation. The change in contacts to the DNA backbone, namely the loss of Ala-137 to Cyt 10 and gain of S136 to Cyt 11, appear to be responsible for twisting the sugar-phosphate DNA backbone (Fig. 8, A–D). The base planes of the outer two base pairs of the recognition sequence (i.e. GTYRAC) are unmoved relative to the intercalating segment (Figs. 5B and 8), and the positions of the phosphorus atoms of Cyt 11 and Cyt 10 are not significantly shifted; however, the sugar rings of Cyt 11 and Cyt 10 as well as Ade 9 are rotated, causing a shift and rotation of the Ade 9 phosphate (Figs. 8 and 9). This would result in a steric conflict between this phosphate and Thr-130 unless one of two things happens; either 1) segments of the protein subunit shift to increase the distance between Thr-130 and Ade 9 phosphate, or 2) the phosphate of Ade 9 rotates to move the offending phosphate oxygen out of the way of the Thr-130 side chain.

**Active Site Conformations**—To understand how the binding of Ca\(^{2+}\) has appeared to alter the conformations of the protein, DNA, and the protein-DNA interface and to investigate the origins of the reduced cleavage activity by the mutant enzyme, we examined the vicinity of the Ca\(^{2+}\) in the active site of Q138F HincII structures, each active site binds one Ca\(^{2+}\) ion, similar to that found in wild type HincII/CG/Ca\(^{2+}\) (Figs. 4 and 9, C–D). Although two divalent cations are expected to bind in the case of Mg\(^{2+}\) and as observed with Mn\(^{2+}\) (17), Ca\(^{2+}\) binds in only one of the two positions seen to be occupied by Mn\(^{2+}\) in both wild type and Q138F HincII.

The Ca\(^{2+}\) bound in the active site bridges the protein and DNA. It is ligated to both active site aspartate residues, Asp-114 and Asp-127, as well as to the scissile phosphate (the phosphate to be cleaved) found between the central pyrimidine-purine step. For this reason the Ca\(^{2+}\)-bound structures appear to resist displacement of the protein near these residues to maintain these contacts. In addition, the efficiency of catalysis of DNA cleavage will be dependent on the precise positioning of any groups within the active site, and therefore, mispositioning of important active site side chains, the divalent cation, and/or the DNA can result in a decrease in the rate of DNA cleavage.

We wanted to characterize exactly how much, if any, mispositioning occurred within the active sites of the subunits of the various structures. Fig. 9 shows each active site superimposed onto wild type HincII bound to CG DNA with Ca\(^{2+}\) (wtHincII/CG/Ca\(^{2+}\)) (16), which caused very little change in the protein or DNA structure. In the case of the DNA-bound Q138F HincII structures, each active site binds one Ca\(^{2+}\) ion, similar to that found in wtHincII/CG/Ca\(^{2+}\) (Figs. 4 and 9, C–D). Although two divalent cations are expected to bind in the case of Mg\(^{2+}\) and as observed with Mn\(^{2+}\) (17), Ca\(^{2+}\) binds in only one of the two positions seen to be occupied by Mn\(^{2+}\) in both wild type and Q138F HincII.
Indirect Readout of DNA by HincII Endonuclease

Table 5: Relationship between subunit rotation relative to wild type HincII and Ade 9 phosphate conformation

| Structure       | Subunit | Rotation of Ade 9 phosphate | Rotation of protein subunit |
|-----------------|---------|----------------------------|-----------------------------|
| Q138F/CG        | B       | No (γ = 42°)               | 16.1°                       |
| Q138F/TG        | B       | No (γ = 92°)               | 10.5°                       |
| Q138F/CG/Ca²⁺   | B       | No (γ = 92°)               | 9.3°                        |
| Q138F/TG/TA     | A       | No (γ = 91°)               | 7.5°                        |
| Q138F/TG/TA/CA²⁺| B       | No (γ = 32°)               | 6.6°                        |
| Q138F/CG/CA²⁺   | A       | Yes (γ = 171°)             | 3.2°                        |
| Q138F/CG        | A       | Yes (γ = 168°)             | 2.7°                        |
| Q138F/TG/TA/CA²⁺| A       | Yes (γ = 170°)             | 2.7°                        |
| Wild type HincII/TG | A | 0.8° | |
| Wild type HincII/TA | B | 0.7° | |

*The rotation angle was calculated by first superimposing the 6 base pairs of the DNA recognition sequence of the listed structure onto that of wild type HincII bound to CG and Ca²⁺, then calculating the rotation required to bring residues 126–132 and 167–175 of the listed structure onto the same residues of wild type HincII/CG/CA²⁺.

Out bound Ca²⁺ (Fig. 9, A and B) than with bound Ca²⁺ (Fig. 9, C and D). On the whole mispositioning is greater in Q138F/CG/CA²⁺ than in Q138F/TG/TA/CA²⁺. This is significant as the single turnover rate of cleavage is depressed to a greater degree with CG than with TA at the center step of the recognition sequence, GTYRC.

Mispositioning in the active site is a consequence of distortions of the protein and alterations in the protein-DNA interface. The protein distortion appears to be linked to the clash between the shifted phosphate of Ade 9 and the side chain of Thr-130, resulting in less protein deformation including intercalation, and protein conformational changes to better align active site groups. Alternatively, the mutation may also alter the rates of other steps in the reaction pathway preceding cleavage, which include binding, translocating, DNA distortion including intercalation, and protein conformational changes (see below).

Why Does the Q138F Mutation Reduce the Single Turnover Cleavage Rate of Cognate DNA by HincII?—The positioning of important groups in the active sites of these complexes is misaligned relative to the wild type HincII structure bound to cognate DNA (and Ca²⁺). The misalignment is worse without bound Ca²⁺, indicating that the binding of Ca²⁺ serves to better align these groups. This is in contrast to the wild type HincII structures, where little difference is seen with and without bound Ca²⁺ (16). The requirement of realignment to bind Ca²⁺ in the mutant structures may decrease the overall affinity for Ca²⁺. In addition, the misalignment is expected to decrease the rate of DNA cleavage through either a higher energy transition state for cleavage or a requirement for conformational changes to better align active site groups. Alternatively, the mutation may also alter the rates of other steps in the reaction pathway preceding cleavage, which include binding, translocating, DNA distortion including intercalation, and protein conformational changes (see below).

Why Is the DNA Binding Affinity of HincII for Cognate DNA Altered Very Little by the Q138F Mutation?—When a modification is introduced into an endonuclease at the protein-DNA interface, particularly if distant from the active site, it is expected, due to the similarity in structure, that the mutation will destabilize both the ground and transition states to the same extent (40). In such cases the effect of the destabilizing modification will be exhibited at the binding, rather than cleavage step of the reaction pathway. However, in many cases it is known that such modifications can lead to conformational readjustments in the ground state protein-DNA complex to minimize the energetic cost of that modification (10, 34, 40). Such structural adjustments result in a less than expected effect on the binding affinity, and the phenomenon has been termed adaptation (40). Because of the requirement for a precise arrangement of active site groups, the transition state is more structurally stringent.
than the ground state complex, and thus, the destabilization caused by the modification cannot be minimized by structural adaptation. With the destabilized transition state and stabilized ground state, effects will be seen on the cleavage, rather than binding step of the reaction pathway. Thus, structural adaptation is consistent with the results obtained here for the Q138F mutant of HincII. The mutation had nearly no effect on the binding affinity of HincII for cognate DNA but had a 16–214-fold effect on the single turnover rate constant for cleavage of DNA. We find that there are many changes in the protein, DNA, and protein-DNA interface structure, consistent with structural adaptation. These changes must work to minimize the detrimental effects of the mutation on the DNA binding affinity.

Why Does the Q138F Mutation of HincII Reduce the Single Turnover Rate of Cleavage More for GTCGAC than for GTTAC?—The mutation Q138F of HincII affected the overall and relative multiple turnover rates of the enzyme for its cognate sequences. However, the mutation did not affect the magnitude or specificity of binding of the enzyme for these substrates but did affect the single turnover rate of cleavage. Before any structural interpretations can be made of this result, we must first consider the molecular step or steps measured by the single turnover rate of cleavage.

The single turnover rate of cleavage as implemented here is a measurement meant to isolate a single round of DNA cleavage up to and including the actual chemical cleavage of DNA. The product release step is not measured because all product, whether bound to the enzyme or released into solution, is measured in the assay. When the reaction is initiated by mixing enzyme and DNA, both in buffer containing Mg2+ and H1001, steps including binding, DNA bending, DNA intercalation, and protein conformational changes as well as the actual cleavage step contribute to the observed rate of product formation. In many cases, and especially at higher concentrations of enzyme, the actual binding step can be very fast relative to later steps. Alternatively, the reaction can be initiated by the addition of Mg2+ to premixed enzyme and DNA. If the enzyme binds to the recog-
Indirect Readout of DNA by HincII Endonuclease

...tion sequence strongly in the absence of the divalent cation, then the pathway to DNA cleavage will involve binding of the Mg$^{2+}$ as well as the DNA hydrolysis. However, no binding to cognate DNA containing CG at the center step was detected in gel shift assays without Ca$^{2+}$ with either Q138F or wild type HincII. The DNA concentration was 1 nM in these experiments, and enzyme levels up to 1 μM were tested, indicating that the equilibrium dissociation constant would be higher than 1 μM. Thus, binding of the DNA by the HincII enzymes before the addition of Mg$^{2+}$ would not occur to any great extent in the single turnover assays using 0.8–3.0 μM HincII. The fact that both wild type and Q138F HincII have been crystallized with cognate DNA in the absence of divalent cation indicates that binding does take place under the crystallization conditions; however, the enzyme and DNA concentrations are very high (90 and 180 μM), and solution conditions are radically different (25% polyethylene glycol) from the reaction buffer.

Both for the mutant and for wild type HincII, the single turnover cleavage rate constants were found to be independent of enzyme concentration (in the range of 1 × 10$^{-7}$ to 3 × 10$^{-3}$ M). In addition, we found very little dependence on Mg$^{2+}$ concentration (in the range of 10–50 mM, with buffer adjusted to constant ionic strength); therefore, protein binding to DNA and Mg$^{2+}$ binding to the protein-DNA complex is probably not rate-limiting. The data fit well to a single exponential, implying a single rate-limiting step. We do not have any information to indicate whether this step is Mg$^{2+}$-dependent target site location, Mg$^{2+}$-dependent DNA distortion, Mg$^{2+}$-dependent protein conformational changes, or the actual (Mg$^{2+}$-dependent) rate of DNA cleavage. Further studies will be needed to isolate these steps.

From the structural results, it is apparent that although the protein and DNA conformations overall are more distorted in the Q138F/TA structures, the active sites in the presence of Ca$^{2+}$ are better aligned in the Q138F/TA/Ca$^{2+}$ structure than in the Q138F/CG/Ca$^{2+}$ structure. The effect of Mg$^{2+}$ binding on the structures is likely to be similar to that of Ca$^{2+}$ binding. If the increased depression of the single turnover rate of cleavage with the CG substrate relative to the TA as a result of the mutation is due to effects on the rate of chemical cleavage, then the greater misalignment in the Q138F/CG/Ca$^{2+}$ structure is consistent with this result. In addition, if it is due to a slower conformational change on the part of the Q138F/CG/Ca$^{2+}$ structure, then the structural observations are also consistent with this result, as we find much more flexibility in terms of structural deviations from the wild type in the Q138F/TA and Q138F/TA/Ca$^{2+}$ structure than in Q138F/CG and Q138F/CG/Ca$^{2+}$. Although the structures do not directly report on conformational dynamics, much greater conformational freedom is suggested in the Q138F/TA/Ca$^{2+}$ and Q138F/TA structures by the greater differences in structure from wild type. Both of these effects seem to originate in the ability of the DNA bases to take on different stacking around the center TA step (Fig. 6, A, B, and E), by either allowing the multiple conformations themselves, or by allowing multiple structural modifications to compensate for the protein distortions that otherwise lead to poor active site alignment.

There is little distinction between these two arguments if one believes that the cleavage reaction transition state is purely stringent. This is because the actual reaction rate of the chemical cleavage event would be the same in both cases, since the active site arrangement by definition must be identical and the difference in single turnover cleavage rate would be due to the differences in the rate of attaining that stringent conformation. That is, the rates must differ due to a difference in the energy of a transition state for attaining a near attack conformation (the ground state reaction just before reaction), such as protein or DNA structural rearrangement. The energy of this transition state could be dependent on the degree of misalignment and/or the flexibility of the protein-DNA complex. Of course, it is entirely possible that the actual cleavage reaction (with bond making and breaking) proceeds through transition states of differing energies due to differences in the ability of the enzyme to stabilize the reaction transition state. Further studies into the kinetics of the cleavage reactions by the Q138F HincII and its various substrates will distinguish between these possibilities.

Summary and Conclusion—A DNA intercalating residue of HincII was mutated from glutamine to phenylalanine, Q138F, to test the connection between intercalation of DNA by this residue and indirect readout of the center YR step of the cognate recognition sequence GTYRAC. Multiple turnover velocity measurements showed a 6-fold decrease in preference for cognate DNA containing CG at the center step than for that containing TA, as a result of the mutation. In contrast, the binding affinity measurements of the wild type and mutant enzyme to the two cognate DNA sequences showed little difference between wild type and the mutant enzyme. Single turnover cleavage measurements showed a change of 3.5-fold in ratio of the rate constant of cleavage of cognate DNA containing TA relative to CG. The single turnover cleavage rate constant is also reduced 16–214-fold by the mutation and reduced more for cognate DNA containing CG than that containing TA at the center step. A high resolution crystal structure of wild type HincII bound to TA DNA showed no significant differences from that containing CG. High resolution crystal structures of the mutant HincII, Q138F, bound to DNA showed marked changes in conformations at the tertiary and quaternary (i.e. dimeric) level as well as changes in DNA conformation and the protein-DNA interface, which were dependent on the center DNA step sequence and the presence or absence of bound Ca$^{2+}$. The perturbations in structure were traced from the site of mutation showing a change in position of residue 138 in the intercalated DNA. This shift results in a marked shift in the position of the side chain of the adjacent residue, Ala-137, which may explain the observed change in a deoxyribose conformation normally contacted by Ala-137 in the wild type structures. The altered sugar conformation propagates changes in the sugar-phosphate backbone to adjacent nucleotides, leading to a steric conflict between the O5’ oxygen of the phosphate of Ade 9, just adjacent to the scissile phosphate, and the side chain of Thr-130. The conflict is avoided in different ways in different subunits of the different structures. In three of the four
structures, one subunit shows a rotation about the γ angle (at C5’-C4’) of the sugar ring, moving the phosphate away from Thr-130. The other subunits of these dimeric structures have residue 130 shifted away from the bound DNA by varying degrees. The binding of Ca²⁺ at the active site very near to this point of steric collision acts to favor a conformation closer to that seen in the wild type HincII DNA-bound structures. The realignment of the protein-DNA interface at the active site as a result of Ca²⁺ binding is more successful when TA is found at the center step of the recognition sequence than with CG. This difference in the ability of Ca²⁺ to correct misalignment appears to be a result of an alteration in the stacking of the two center step adenine bases with each other as well as with the adjacent adenines (Ade 9) of the recognition sequence (i.e. GTTAAC). The recognition sequence containing CG at the center step, namely GTCGAC, retains the conformation seen for cognate DNA bound to wild type HincII, and consequently, the active site is more poorly aligned, presumably leading to a decreased rate of DNA cleavage. Hence, the conformation of the center step DNA seen in the original wild type HincII/cognate DNA structure, the cross-strand purine stack or CSPS, which is suspected of contributing to indirect readout of the sequence by HincII, appears to become a deficit to the recognition by the mutant Q138F HincII enzyme.

Acknowledgments—Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University, on behalf of the United States Department of Energy, Office of Basic Energy Sciences. The Stanford Synchrotron Radiation Laboratory Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and by the National Center for Research Resources, Biomedical Technology Program and the NIGMS, National Institutes of Health.

REFERENCES

1. Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 804–808
2. Luscombe, N. M., Austin, S. E., Berman, H. M., and Thornton, J. M. (2000) Genome Biology http://genomeweb.org/2000/1/1/reviews/001
3. Otwinski, Z., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) Nature 335, 321–329
4. Lesser, D. R., Kurpiewski, M. R., and Jen-Jacobson, L. (1990) Science 250, 776–786
5. Koudelka, G. B., Harrison, S. C., and Ptashne, M. (1987) Nature 326, 886–888
6. Bareket-Samish, A., Cohen, I., and Haran, T. E. (2000) J. Mol. Biol. 299, 965–977
7. Schumacher, M. A., Hurlbut, B. K., and Brennan, R. G. (2001) Nature 409, 215–219
8. Garvie, C. W., and Phillips, S. E. (2000) Structure Fold. Des. 8, 905–914
9. Bareket-Samish, A., Cohen, I., and Haran, T. E. (1998) J. Mol. Biol. 277, 1071–1080
10. Martin, A. M., Sam, M. D., Reich, N. O., and Perona, J. J. (1999) Nat. Struct. Biol. 6, 269–277
11. Chen, S., Gunasekera, A., Zhang, X., Kunkel, T. A., Ebright, R. H., and Berman, H. M. (2001) J. Mol. Biol. 314, 75–82
12. Gromiha, M. M., Siebers, J. G., Selvaraj, S., Kono, H., and Sarai, A. (2005) Gene (Amst.) 364, 108–113
13. Horton, N. C., Dorner, L. F., and Perona, J. J. (2002) Nat. Struct. Biol. 9, 42–47
14. Viboud, I. B., and Halford, S. E. (1995) Biochemistry 34, 1113–1119
15. Martin, A. M., Horton, N. C., Lusetti, S., Reich, N. O., and Perona, J. J. (1999) Biochemistry 38, 8430–8439
16. Etzkorn, C., and Horton, N. C. (2004) Biochemistry 43, 13256–13270
17. Etzkorn, C., and Horton, N. C. (2004) J. Mol. Biol. 343, 833–849
18. Lindahl, T., Liangquist, S., Siegert, W., Nyberg, B., and Sperrens, B. (1977) J. Biol. Chem. 252, 3286–3294
19. Lindahl, T. (1982) Annu. Rev. Biochem. 51, 61–87
20. Melamede, R. J., Hatahet, Z., Kow, Y. W., Ide, H., and Wallace, S. S. (1994) Biochemistry 33, 1255–1264
21. Jiang, D., Hatahet, Z., Melamede, R. J., Kow, Y. W., and Wallace, S. S. (1997) J. Biol. Chem. 272, 32320–32329
22. Horton, N. C., Dorner, L. F., Schildkraut, I., and Perona, J. J. (1999) Acta Crystallogr. Sect. D 55, 1943–1945
23. Aggarwal, A. K. (1990) Methods: A Companion to Methods Enzymol., 1, 83–90
24. Fasman, G. D. (1975) CRC Handbook of Biochemistry and Molecular Biology, 3rd Ed., p. 589, CRC Press, Inc., Cleveland, OH
25. Fried, M., and Crothers, D. M. (1981) Nucleic Acids Res. 9, 6505–6525
26. Taylor, J. D., Badcoe, I. G., Clarke, A. R., and Halford, S. E. (1991) Biochemistry 30, 8743–8753
27. Thiéling, V., Selent, U., Köhler, E., Wolfes, H., Pieper, U., Geiger, R., Urbanke, C., Winkler, F. K., and Pinkow, A. (1991) Biochemistry 30, 6416–6422
28. Leslie, A. G. (1999) Acta Crystallogr. Sect. D 55, 1696–1702
29. CCP4 (1994) Acta Crystallogr. Sect. D 50, 760–763
30. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kukszewski, J.I., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D 54, 905–911
31. McRee, D. E. (1999) J. Struct. Biol. 125, 156–165
32. Vagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
33. Kabsch, W. (1976) Acta Crystallogr. A 32, 922–923
34. Jen-Jacobson, L., Engler, L. E., Lesser, D. R., Kurpiewski, M. R., Yee, C., and McVerry, B. (1996) EMBO J. 15, 2870–2882
35. Reid, S. L., Parry, D., Liu, H. H., and Connolly, B. A. (2001) Biochemistry 40, 2484–2494
36. Jordan, S. R., Whitcombe, T. V., Berg, J. M., and Pabo, C. (1985) Science 230, 1383–1385
37. Moore, G., Choo, Y., and Klug, A. (1995) Proc. Natl. Acad. Sci. U. S. A. 98, 1432–1436
38. Little, E. J., and Horton, N. C. (2005) J. Mol. Biol. 351, 76–88
39. Pingoud, A., Fuxreiter, M., Pingoud, V., and Wende, W. (2005) Cell. Mol. Life Sci. 62, 685–707
40. Jen-Jacobson, L. (1997) Biopolymers 44, 153–180