Published in final edited form as:

J Mol Biol. 2008 October 31; 383(1): 186–204. doi:10.1016/j.jmb.2008.08.032.

DNA Distortion and Specificity in a Sequence Specific Endonuclease

Andrea C. Babic¹, Elizabeth J. Little¹, Veena M. Manohar¹, Jurate Bitinaite², and Nancy C. Horton^{1,*}

¹Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, AZ, 85721

²New England Biolabs Inc., Beverly, MA 01915

Summary

Five new structures of the Q138F HincII enzyme bound to a total of three different DNA sequences and three different metal ions (Ca²⁺, Mg²⁺, and Mn²⁺) are presented. While previous structures were produced from soaking Ca²⁺ into pre-formed Q138F HincII/DNA crystals, the new structures are derived from cocrystallization with Ca²⁺, Mg²⁺, or Mn²⁺. The Mn²⁺ bound structure provides the first view of a product complex of Q138F HincII with cleaved DNA. Binding studies and a crystal structure show how Ca²⁺ allows trapping of a Q138F HincII complex with noncognate DNA in a catalytically incompetent conformation. Many Q138F HincII/DNA structures show asymmetry, despite the binding of a symmetric substrate by a symmetric enzyme. The various complexes are fit into a model describing the different conformations of the DNA bound enzyme, and show how DNA conformational energetics determine DNA cleavage rates by the Q138F HincII enzyme.

Many fundamentally important functions within cells are mediated by proteins which recognize and bind to specific DNA sequences. Studies of sequence specific DNA binding proteins indicate that specific target sequences are recognized using both direct and indirect readout¹. Direct readout involves direct contacts such as hydrogen bonds and van der Waals interactions between the protein and the DNA base functional groups, mostly in the major groove. Indirect readout is defined as sequence specificity occurring in the absence of such direct contacts, and has been proposed to involve contacts mediated by water² or other small molecules³, as well as distortions of DNA that can distinguish different sequences energetically^{4–13} (Little, *et al.*, *submitted for publication*).

Restriction endonucleases are a class of enzymes found in bacteria which cleave DNA at specific target sequences and are thought to protect the bacterial cell from phage infection¹⁴. The target sites in the host DNA are methylated by a methyltransferase, and thereby protected from the cleavage activity of the endonuclease. Cleavage at sequences other than the target site can result in damage and possibly death to the bacterial cell, as only the target site sequence is protected by the activity of the cognate methyltransferase enzyme. Therefore, restriction

*Corresponding author, NCH: Telephone: 520-626-3828 FAX: 520-626-9288 nhorton@u.arizona.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 3E3Y, 3E40, 3E41, 3E42, 3E43, 3E44, 3E45.

endonucleases have evolved very stringent DNA sequence specificity, and are an excellent model system for the study of sequence specific DNA recognition.

The type II restriction endonuclease HincII recognizes and cleaves the sequence GTY|RAC, (where Y indicates either pyrimidine, and R either purine), cutting between the center YR step (indicated by |). (Fig. 1(a)–(b)). HincII utilizes both direct and indirect readout to recognize its target DNA¹⁵ (Little *et al.*, *submitted for publication*). A HincII side chain, Gln 138, intercalates into the bound DNA just outside of the 6 bp recognition sequence (cyan, Fig. 1 (a)–(b)). Binding and intercalation by HincII induces distortions in the DNA, including bending, unwinding, and a shifting of the base planes into the minor groove. One result of the distortion is an unstacking of the pyrimidine and purine bases within the same strand at the YR step, as well as the formation of a different type of stacking – the cross strand purine stack (CSPS), where the purine bases of each strand are shifted toward each other, and exhibit an increased stacking surface area relative to B form DNA (Fig. 1(c)).

We hypothesized that the CSPS was used by HincII in indirect readout¹⁶. To elucidate the role of the DNA intercalation in DNA distortion, the HincII mutant Q138F was made and characterized¹⁵. This enzyme exhibits altered specificity at the YR center step, being reduced in DNA cleavage rate more with the sequence CG than with TA.

Structures of Q138F bound to different DNA sequences provide insights into the origin of the altered specificity¹⁵. The conformation of Q138F HincII/DNA complexes is different with different DNA sequences (TA or CG at the center step). HincII requires Mg²⁺ for DNA cleavage, but Ca²⁺ is often used to mark the Mg²⁺ binding site, while inhibiting cleavage. Ca²⁺ was soaked into the preformed Q138F HincII/DNA crystals, and the resulting structures show marked changes that differ with the different DNA sequences. HincII is a dimer, and the structures of Q138F HincII/DNA are asymmetric, where one active site of each shows a common conformation (color, Fig. 1(d)). The other active site of each structure shows large displacements of the bound DNA relative to that found in the wild type HincII/DNA structure. The displacements differ depending on DNA sequence, with DNA containing TA at the center step less displaced than that with CG. Since positioning in the active site is important for catalysis, the better positioning of the DNA with TA at the center step is consistent with the faster cleavage of that sequence by Q138F HincII. The ability to position the DNA better appears related to the DNA structure at the differing base pairs. In the case of the TA center step, the two purines form the CSPS but with far less overlap of the base planes than found in the wild type HincII/DNA structure. The CSPS of the CG DNA bound to Q138F HincII shows greater overlap, closer in conformation to that of the wild type HincII/DNA structure. Hence, a relationship between the disruption of the CSPS, and the ability to organize into a more optimal conformation in the active site (i.e. more like that found in the wild type HincII/DNA structure), was proposed¹⁵.

Building on this work, we sought to characterize the Q138F HincII enzyme with cognate DNA sequences in the presence of active metal ions such as Mg²⁺ and Mn²⁺. In addition, to investigate if the asymmetry in the structures determined previously with Ca²⁺ was due to restrictions imparted by lattice forces on the conformational changes that occur upon Ca²⁺ binding, (since Ca²⁺ was soaked into preformed enzyme/DNA crystals), we also determined the structures of Q138F with DNA containing TA or CG at the center step cocrystallized with Ca²⁺. The new structures reveal connections between the DNA structure and active site arrangement that further illustrate the origin of the altered specificity by the Q138F mutation. In addition, binding measurements show that wild type and Q138F HincII bind a noncognate DNA only 300 fold more weakly in the presence of Ca²⁺, however, are not capable of cleaving this sequence in the presence of Mg²⁺, even at micromolar concentrations of enzyme. The

structure of Q138F HincII with the noncognate DNA and Ca^{2+} show how an inactive conformation is trapped with Ca^{2+} .

Results

Structures of Q138F HincII with cognate DNA cocrystallized with Ca^{2+}

Q138F HincII was cocrystallized with 5 mM CaCl_2 and 14 bp duplex DNA containing the HincII cognate sequence, GTYRAC, with either TA or CG at the center YR step, (TA DNA and CG DNA respectively). Diffraction data sets from two different crystals with each DNA sequence were collected, and treated separately throughout and structure refinement. The resolutions of the data vary from 2.10 Å to 2.73 Å, and the structures have been refined to R_{cryst} and R_{free} values in the ranges of 19.6–20.5% and 24.8–28.6%, respectively (Table 1). Each asymmetric unit contains one dimer of the Q138F HincII enzyme (chains A and B), bound to one duplex of DNA (chains E and F), with one Ca^{2+} identified in each of the two active sites by peak height ($8-13\sigma$) in simulated annealing omit Fo-Fc maps. The structures containing Q138F HincII bound to the same DNA sequence were found to be very similar, with RMSD using residues 50–257 of both A and B subunits of 0.23 Å and 0.25 Å. The RMSD is 0.60 Å and 0.73 Å for structures having different bound DNA sequences, which reduces to 0.30 Å and 0.40 Å if only individual subunits are used in the comparison, implying that much of the difference between the structures derives from the inter-subunit relationship in the dimer.

Figure 2(a)–(c) compares the active site conformations and DNA structure found in the Q138F/CG/ Ca^{2+} (cocrystal) structure (dark blue), to those found in the previously analyzed crystals of Q138F/CG/ Ca^{2+} (soak) (light blue, PDB ID 2GIH). The cocrystal structures are derived from crystals where Ca^{2+} was present during crystallization, whereas the soak structures are derived from crystals of Q138F HincII bound to DNA which were soaked in 50 mM CaCl_2 ¹⁵. Selected portions of the structure of wild type HincII bound to CG DNA and Ca^{2+} (wtHincII/CG/ Ca^{2+} , grey, PDB ID 1TW8)¹⁷ are shown in grey. Panels (a) and (b) of Figure 2 (and Panels (a) and (b), and (d) and (e) of Fig. 2–Fig. 5) use a superposition of the alpha carbon atoms of the active site residues, Asp 114, Asp 127, and Lys 129 of either subunit A or B, respectively, onto those of subunit A of wtHincII/CG/ Ca^{2+} , to determine the position of the bound DNA and metal ion(s) relative to the enzyme active site. (Note: The subunit nomenclature of Q138F/CG/ Ca^{2+} (soak) is reversed relative to those of the other Q138F/DNA structures, as it was determined in a different space group. Subunit A of Q138F/CG/ Ca^{2+} (soak) is most like subunit B of all other Q138F/DNA structures, and subunit B of Q138F/CG/ Ca^{2+} (soak) is most like subunit A of all other Q138F/DNA structures). Only the sugar-phosphate backbone atoms of the nucleotides 7–9 are shown in Fig. 2(a)–(b), with the phosphate at the scissile phosphodiester bond between Pyr 7 and Pur 8 (SP) and the phosphate 3' from the SP, between Pur 8 and Ade 9 (3'P) identified. The position of a water molecule ligated to the Ca^{2+} of wtHincII/CG/ Ca^{2+} (grey, Fig. 2(a)–(b)), marking the position of the putative nucleophile of the DNA cleavage reaction, is shown as a small pink sphere, with its direction of attack onto the SP of wtHincII/CG/ Ca^{2+} shown as a pink arrow. The Ca^{2+} ions are shown as spheres of the same color as their parent structure. The positions of the SP in the two Q138F/CG/ Ca^{2+} structures (dark and light blue) are closer to that of wtHincII/CG/ Ca^{2+} (grey) in one subunit (Fig. 2(a)), than the other (Fig. 2(b)). The Ca^{2+} ligates to an oxygen of the SP in the active sites of all three structures (dashed lines, Fig. 2(a)–(b)), except for the A subunit of Q138F/CG/ Ca^{2+} (soak) (light blue, Fig. 2(b)). In subunit A, the 3'P of Q138F/CG/ Ca^{2+} (cocrystal) (dark blue, Fig. 2(a)), is rotated relative to that found in wtHincII/CG/ Ca^{2+} (grey, Fig. 2(a)), closely approaches the site of the putative nucleophile (pink sphere labeled nuc, Fig. 2(a)), and ligates the bound Ca^{2+} ion in Q138F/CG/ Ca^{2+} (cocrystal) (dark blue dashes, Fig. 2(a), Ca^{2+} -O distance = 2.8 Å). Such ligation is not seen in the wtHincII/CG/ Ca^{2+} (grey, Fig. 2(a)), subunit B of either Q138F/CG/ Ca^{2+} (cocrystal) structure (dark blue, Fig. 2(b)), or either subunit of Q138F/CG/ Ca^{2+} (soak)

(light blue, Fig. 2(a)). The bound DNA is less displaced, relative to that found in wtHincII/CG/ Ca^{2+} (grey, Fig. 2(b)), in the Ca^{2+} cocrystal (dark blue, Fig. 2(b)) than in the Ca^{2+} soak (light blue, Fig. 2(b)).

Panel (c) of Figure 2 (and panels (c) and (f) of Fig. 2–Fig. 5) uses a superposition of the nine atoms of the ring of one of the two center step purine bases (at nucleotide 8, Fig. 1(b)), in order to compare the degree of overlap in the CSPS. The structure of Q138F HincII and CG DNA derived from cocrystallization with Ca^{2+} (dark blue, Fig. 2(c)), shows a CSPS more like that in the DNA bound to wild type HincII (grey, Fig. 2(c)), having 91% the buried surface area (Table 2). The structure derived from soaking in Ca^{2+} (light blue, Fig. 2(c)), has only 79% of the buried surface area of the CSPS present in the wild type structure.

Figure 2(d)–(f) shows a similar comparison of Q138F/TA/ Ca^{2+} structures derived from cocrystallization (red) with or soaking (pink) crystals in Ca^{2+} ¹⁵, with wtHincII/CG/ Ca^{2+} shown in grey. The positions of the SP groups of the DNA bound to the mutant structures (red, pink, Fig. 2(d)–(e)), with the exception of that of subunit B of the soak (pink, Fig. 2(e)), are very close to that found in the DNA bound to wild type HincII (grey, Fig. 2(d)–(e)). The 3'P is rotated and ligates the bound Ca^{2+} in all Q138F/TA/ Ca^{2+} structures (pink and red dashes, Fig. 2(d)–(e)) with the exception of subunit B of Q138F/TA/ Ca^{2+} (soak) (pink, Fig. 2(d)). The oxygen of the 3'P, which ligates the Ca^{2+} , overlaps with the nucleophile binding site (dark blue sphere labeled nuc, Fig. 2(d)–(e)) (O- Ca^{2+} distance of 0.6 Å). The conformation of the DNA bound to the active sites is identical in the two subunits of Q138F/TA/ Ca^{2+} (cocrystal) (red, Fig. 2(d)–(e)) and very close to that of subunit A of Q138F/TA/ Ca^{2+} (soak) (pink, Fig. 2(d)). As seen before¹⁵, overlap of the center step purine bases (the CSPS) is less in the Q138F/TA/ Ca^{2+} structures (red, pink, Fig. 2(f)), than in the wild type (grey, Fig. 2(f)) or Q138F/CG/ Ca^{2+} structures (dark and light blue, Fig. 2(c)). However, the amount of surface area buried in the CSPS is slightly better in the mutant structure derived from cocrystallization (red, Fig. 2(f)), with 85% of that found in the structure wild type HincII bound to TA DNA (PDB ID 2GIE, Table 2), than that derived from soaking (pink, Fig. 2(f)), with only 79%.

Structures of Q138F HincII cocrystallized with TA DNA and Mg^{2+}

Q138F HincII was crystallized with TA DNA and in the presence of 2.5 mM Mg^{2+} . The diffraction data extend to 2.75 Å, and the model refined to an R_{cryst} of 19.3% and R_{free} of 26.9% (Table 1). The DNA is uncleaved, with one Mg^{2+} ion modeled into each active site in the M1 site (orange sphere, Fig. 3(a)–(b)), where M1 refers to one of the two metal ion binding sites of the two metal ion mechanism¹⁸. Mg^{2+} cannot be identified by scattering alone, as it contains the same number of electrons as water. Instead, Mg^{2+} can be identified by its shorter ligand distances than those between water molecules, and preference for octahedral geometry with 6 ligands, rather than tetrahedral with 4 hydrogen bonding partners, as between water molecules. However, neither of these criteria is sufficient to unambiguously identify the M1 peak as Mg^{2+} in the Q138F/TA/ Mg^{2+} structure. Yet, in support of Mg^{2+} binding in the Q138F/DNA structure, the DNA backbone is pulled deeper into the active site (orange, Fig. 3(a)–(b)) than found in the structure of Q138F HincII with the same DNA and no metal ions¹⁵ (brown, Fig. 3(a)–(b), PDB ID 2GII). The DNA conformation in the two structures is also very different, as seen in Fig. 3(c). Therefore the M1 peak has tentatively been assigned as Mg^{2+} (and is identified in the PDB file as residue and atom type SOL for solvent, to indicate the ambiguity in its identification as water or Mg^{2+}).

The positioning of the DNA in the active site is similar in the two subunits of Q138F/TA/ Mg^{2+} (orange, Fig. 3(a)–(b)), with the rotated conformation of the 3'P, and close approach to the atom at M1 (orange sphere, Fig. 3(a)–(b)), as well as the putative nucleophile binding site (blue sphere labeled nuc, Fig. 3(a)–(b)). The distance between the closest oxygen of the 3'P and the putative nucleophile is 1.0 Å and 1.3 Å, in subunits A and B, respectively, which more

distant than that found in Q138F/TA/Ca²⁺ (soak) (red, Fig. 3(a)–(b), Ca²⁺-O distance = 0.6 Å). The oxygen of the 3'P is 2.6 Å from the putative Mg²⁺ in both active sites, suggesting direct ligation (orange dashes, Fig. 3(a)–(b)), however the putative Mg²⁺ is 4.1 Å and 3.4 Å from the nearest nonbridging oxygen of the SP in subunits A and B, respectively, suggesting absence of ligation. This is unusual as the M1 metal ion ligates the SP, but not the 3'P, in the wild type structures (grey, Fig. 3(a)–(b)). The M1 ion also typically ligates the SP in most Q138F/DNA structures. The SP is more displaced from that in the wild type structure (grey, Fig. 3(d)–(e)) and the Ca²⁺ bound mutant structure, Q138F/TA/Ca²⁺ (cocrystal) (red, Fig. 3(d)–(e)).

The overlap of the center step adenine bases in the CSPA is much poorer in Q138F/TA/Mg²⁺ (orange, Fig. 3(c)), than wtHincII/CG/Ca²⁺ (grey, Fig. 3(c)), with only 72% of the buried surface area (Table 2). The degree of contact in the CSPA of Q138F/TA/Mg²⁺ (orange, Fig. 3(f)) is less even than that found in the Ca²⁺ bound mutant structure (red, Fig. 3(f)).

Structures of Q138F HincII cocrystallized with TA DNA and Mn²⁺

Q138F HincII was also cocrystallized in the presence of 2.5 mM MnCl₂ and the TA DNA. The data extend to 2.5 Å, and the structure is refined to an R_{cryst} of 20.2% and an R_{free} of 25.7% (Table 1). Two Mn²⁺ ions have been identified in each active site by peak height in simulated annealing omit Fo-Fc maps (6–8σ), and the DNA is cleaved between the Thy 7 and Ade 8.

Figure 4(a)–(c) compares the Q138F/TA/Mn²⁺ (purple) structure with those of wtHincII/CG/Mn²⁺ (black) and wtHincII/CG/Ca²⁺ (grey). The DNA is also cleaved in the wtHincII/CG/Mn²⁺ (black, Fig. 4(a)–(b)), and up to four Mn²⁺ have been found in the crystallographically independent active sites¹⁸. The M1 and M2 Mn²⁺ sites of wtHincII/CG/Mn²⁺ (black, Fig. 4(a)–(b)) are positioned most closely with those found in Q138F/TA/Mn²⁺ (purple spheres, Fig. 4(a)–(b)). The positioning of the DNA is similar in the active sites of subunits A and B of Q138F/TA/Mn²⁺ (purple, Fig. 4(a)–(b)), and shows the rotation of the 3'P, similar to that found in other Q138F/DNA structures. However, the oxygen of the 3'P does not approach the M1 metal site as closely (Mn²⁺-O distance = 3.5 Å) as found in Q138F/TA/Ca²⁺ (red, Fig. 5(a)–(b)) (Ca²⁺-O distance = 2.4–2.5 Å). The position and rotation of the 3'P resembles that in Q138F/TA/Mg²⁺ (orange, Fig. 4(d)–(e)) more than that of Q138F/TA/Ca²⁺ (red, Fig. 5(a)–(b)), and is 1.3–1.4 Å from the putative nucleophile (blue sphere, Fig. 4(a)–(b)).

The degree of overlap of the center step purines is shown in Fig. 4(c). The overlap of the center step purine bases (the CSPA) is far less in Q138F/TA/Mn²⁺ (purple, Fig. 4(c)), than wtHincII/CG/Ca²⁺ and wtHincII/CG/Mn²⁺ (grey and black, respectively, Fig. 4(c)). Since the CSPA is the same in both wild type structures, regardless of whether or not the DNA is cleaved, the loss of CSPA in Q138F/TA/Mn²⁺ cannot be due to DNA cleavage alone. The loss of the CSPA in Q138F/TA/Mn²⁺ is greater than the losses in any other Q138F/DNA structure (with metal ions bound) having only 69% of the buried surface area of that found in the wild type structure (Table 2). Comparison of Q138F/TA/Mn²⁺ (purple, Fig. 4(f), 5(c)) with that of Q138F/TA/Mg²⁺ (orange, Fig. 4(f)) or Q138F/TA/Ca²⁺ (red, Fig. 5(c)), shows a much greater disruption of the CSPA in Q138F/TA/Mn²⁺ (purple, Fig. 4(f), 5(c)).

Structure of Q138F HincII with noncognate GC DNA and Ca²⁺

Q138F HincII was also cocrystallized with the noncognate GC DNA and 5 mM Ca²⁺. The data extend to 2.78 Å, and the structure is refined to an R_{cryst} of 20.0% and an R_{free} of 28.2% (Table 1). Figure 5(d)–(f) compares the active site arrangements and CSPA of Q138F/GC/Ca²⁺ (green, Fig. 5(d)–(e)), with those of Q138F/TA/Ca²⁺ (red, Fig. 5(d)–(e)) and wtHincII/CG/Ca²⁺ (grey, Fig. 5(d)–(e)). The DNA backbone sits in the two active sites very much like that of Q138F/TA/Ca²⁺ (red, Fig. 5(d)–(e)). A single Ca²⁺ was identified by peak height in a simulated

annealing omit Fo-Fc map (12σ), located at the M1 site of each active site. The 3'P is twisted and ligates the Ca^{2+} in subunit A (green, Fig. 5(d)) (Ca^{2+} -O distance=2.1 Å), just as in Q138F/TA/ Ca^{2+} (red, Fig. 5(d)), but is more twisted in subunit B, allowing both oxygen atoms of the 3'P to ligate the Ca^{2+} (green, Fig. 5(e)) (Ca^{2+} -O distances of 2.9 Å and 3.2 Å for the two different oxygen atoms). A CSPS cannot form with this DNA, as position 8 is occupied by a cytosine base, which is too small to reach across the duplex (green, Fig. 5(f)).

Binding and cleavage of noncognate GC DNA

Figure 6 shows the results of gel shift and fluorescence polarization DNA binding assays of wild type and Q138F HincII for a 16 bp oligonucleotide containing either the abasic site DNA, or the noncognate GC DNA sequence, in the presence of 5 mM CaCl_2 . The weak binding of wild type and Q138F HincII to the GC DNA made the measurement of direct binding using fluorescence polarization difficult, due to the large quantity of enzyme needed to reach 100% binding in the 2 ml cuvette (*data not shown*). In addition, the very strong binding of wild type and Q138F HincII to cognate DNA in the presence of CaCl_2 ¹⁵, also made direct binding measurements by fluorescence polarization difficult as the subnanomolar concentrations of HEX-DNA that would be necessary for a binding site titration did not provide sufficient signal. Therefore, DNA containing abasic sites within the cognate sequence was designed to weaken the binding affinity sufficiently to allow for the titration to be performed at concentrations of enzyme and DNA that give a complete binding isotherm.

Fitting of the amount of unshifted DNA in gel shift experiments (three replicates) with wild type HincII and noncognate GC DNA to the binding equilibrium equation results in a K_D (equilibrium dissociation constant) of 335 ± 63 nM (Fig. 6(a)). Similar measurements with Q138F HincII (three replicates) result in a K_D of 323 ± 134 nM. This equilibrium dissociation constant was also measured using a fluorescence polarization assay. The measurement utilized a competition between unlabeled noncognate GC DNA and a HEX fluorophore labeled duplex DNA containing an abasic site on each strand in the HincII recognition sequence. Figure 6(b) shows the change in polarization of the emission from a HEX fluorophore upon addition of Q138F HincII enzyme. The data can be fit with a binding isotherm characterized by a K_D of 101 ± 4 nM and $n = 1$. Figure 6(c) shows a competition binding experiment using the same HEX labeled abasic site duplex at 1 nM, but with a constant amount of Q138F HincII (221 nM), and increasing amounts of an unlabeled duplex containing the HincII noncognate GC DNA sequence. Using the previously measured binding constant for the HEX-DNA with abasic sites, the equilibrium dissociation of the noncognate GC DNA with Q138F HincII can be calculated to be 322 ± 87 nM. Therefore, the two measurements of binding affinity, the gel shift and the fluorescence polarization assay, give very similar numbers: 323 ± 134 nM for the gel shift measured three times, and 322 nM measured using the fluorescence assay in the presence of 5 mM CaCl_2 .

Binding measurements of wild type and Q138F HincII to the abasic site and noncognate GC DNA in the presence of 20 mM Mg^{2+} indicate very weak binding (*data not shown*). The equilibrium dissociation constants could not be measured reliably with the fluorescence polarization assay due to the high concentrations of enzyme necessary to fully bind the DNA, however they can be estimated as 6 to ≥ 10 fold weaker than binding in the presence of 5 mM Ca^{2+} . Using the gel shift assay, no binding of the noncognate GC DNA was seen with 1 nM DNA and concentrations of wild type or Q138F HincII up to 1 μM in the presence of 20 mM Mg^{2+} (*data not shown*), indicating that binding of the noncognate GC DNA is at least three fold weaker in the presence of 20 mM Mg^{2+} than in the presence of 5 mM Ca^{2+} . Single turnover DNA cleavage assays were performed with the noncognate GC DNA, however no specific cleavage of the DNA was seen with either 3 μM wild type or 3 μM Q138F HincII (*data not shown*).

Discussion

Previous knowledge of HincII structure-function relationships

Wild type HincII is a dimeric type II restriction endonuclease that binds and cleaves duplex DNA containing the sequence GTYRAC, in a Mg^{2+} dependent reaction. The DNA is cleaved between the center YR base step, with water as the nucleophile that attacks the SP phosphorus in-line with the O3' leaving group (Fig. 7(a)). The structure of HincII bound to GTCGAC shows that the enzyme makes direct, sequence specific contacts to all bases of the recognition sequence, with the exception of the center step pyrimidine (underlined in GTCGAC)¹⁶. The conformation of HincII bound to cognate DNA does not change when bound to Ca^{2+} ions¹⁷ (mimicking the binding of Mg^{2+}), or DNA containing other YR sequences, such as TA¹⁵. HincII distorts the bound DNA by bending, intercalating, and shifting the base planes of the recognition sequence into the minor groove. The center YR step adopts an unusual conformation, with unstacking of the pyrimidine base from the purine on the same strand, and an increase in stacking of the purine bases across the two strands (Fig. 1(c)). The latter structure is known as a cross-strand purine stack (CSPS). The missing contact to the pyrimidine, and the unusual DNA conformation, led us to propose the involvement of the CSPS in indirect readout of DNA sequence by HincII¹⁶. Recently we have shown that HincII does recognize the center YR base pairs using indirect readout, as removal of all direct contacts results in no loss of sequence specificity (Little, *et al.*, *submitted for publication*).

The proposed catalytic mechanism for the cleavage of DNA by HincII is shown schematically in Figure 7(a). This mechanistic model was derived from crystal structures of wild type HincII bound to cognate DNA and either Ca^{2+} ¹⁷, Mg^{2+} , or Mn^{2+} ¹⁸. The natural cofactor *in vivo* is thought to be Mg^{2+} , however, Mn^{2+} will also confer cleavage activity to HincII¹⁸. Structures solved from HincII/DNA crystals soaked in either of these ions shows the conformation after DNA cleavage has taken place. In contrast, Ca^{2+} inhibits the cleavage reaction allowing capture of the HincII/DNA conformation prior to DNA cleavage¹⁷. Only a single Ca^{2+} is found in each active site of the HincII dimer, while several binding sites are found for Mn^{2+} . (Due to the resolution limit of the diffraction data, the binding sites for Mg^{2+} could not be unambiguously located.) Two of the Mn^{2+} binding sites are common to all crystallographically independent active sites, and have been used to propose a two metal ion mechanism¹⁸. The additional Mn^{2+} bound in each active site may have resulted from the high concentrations (50 mM) of $MnCl_2$ used in preparing the crystals. Mechanistic models for the Mg^{2+} dependent enzymatic cleavage of DNA have been proposed based on x-ray crystal structures and biochemical studies^{19–22}. The models contain similarities and differences that can be understood using the framework of the two-metal ion mechanism shown in Fig. 7(a). In the two metal ion mechanism, both metal ions, M1 and M2, are ligated to SP in order to stabilize the transition state (Fig. 7(a)). One metal ion, M1, also activates the nucleophile (nuc, Fig. 7(a)) for attack on the phosphorus, by lowering its pKa inducing deprotonation to hydroxide, which is more nucleophilic than water. The second ion, M2, stabilizes the O3' leaving group by direct ligation, a catalytic interaction since the negative charge on the O3' resulting from P-O bond cleavage would be relatively unfavorable. Differences in the various proposed mechanistic models include the number of metal ions, where 1, 2 and 3 have been proposed, as well as the presence or absence of the ligations from the metal ions to the phosphate and 3' oxygen atoms. Interactions with the 3'P and the lysine side chain, as well as the identity of the residue in the position of the lysine shown in Fig. 7(a), also differ in the various proposed mechanisms. These differences may represent real variations in the mechanisms utilized by different enzymes, or derive from artifacts of the biochemical and structural studies.

Restriction endonucleases typically contain an active site motif with two acidic residues and a third residue, often lysine²³. The acidic residues (Asp 114 and Asp 127 in HincII) ligate the metal ions, while the lysine (Lys 129 in HincII) side chain interacts with the SP and the

nucleophile (Fig. 7(a)). In several type II restriction endonuclease structures, including HincII, the nucleophile is also hydrogen bonded to the phosphate 3' of the SP (the 3'P). Substitution of the 3'P in the HincII recognition sequence inhibits cleavage by HincII¹⁷ suggesting a catalytic role for the group. Deletion of the active site lysine^{24; 25} in EcoRV, a structural relative of HincII, greatly diminishes its catalytic activity as well, indicating that the presence of the lysine is also important for optimal cleavage activity. In addition, mutations that cause mispositioning of these or other groups, including the metal ions, in the active site of EcoRV, have been found to also reduce DNA cleavage rates^{24; 26–28}. These studies suggest that the optimal endonuclease activity depends upon the formation of the appropriate protein-DNA interface, allowing for the optimal arrangement of the SP, as well as the 3'P, relative to the metal ions and lysine side chain. This principle is important in understanding how DNA structure plays a role in enzyme specificity, and we refer to the conformation found in the wild type HincII/DNA structures as the optimal active site alignment.

The Q138F mutation was designed to alter the HincII induced DNA distortion, in order to test the relationship between the altered DNA structure and enzyme specificity. The Q138F HincII enzyme was found to possess an altered preference for the three cognate sequences, GTTAAC, GTCGAC, and GTTGAC (with complementary sequence GTCAAC), favoring that with TA at the center step 36 fold over that with CG¹⁵. The wild type enzyme also favors the TA sequence, but by only a factor of 6. Binding and kinetic studies showed that the effect occurred predominantly at the DNA cleavage, rather than at the DNA binding step¹⁵. In each Q138F HincII/DNA structure, differences in the positioning and conformation of the bound DNA in the two crystallographically independent active sites of the dimeric enzyme were found. Such asymmetry does not occur in any of the wild type HincII/DNA structures. Further, the structures of Q138F HincII bound to the two different DNA sequences differ. Ca²⁺ was soaked into the preformed Q138F HincII/DNA crystals to allow the determination of the Q138F/TA/Ca²⁺ (soak) and Q138F/CG/Ca²⁺ (soak) structures. The binding of Ca²⁺ clearly induces changes in the structures of Q138F bound to TA DNA and to CG DNA, and changes them differently. However, in both structures, the positioning and conformation of the DNA bound to the active site of one subunit was found to be very similar (light blue Fig. 2(a), pink, Fig. 2(d)). Comparison of the positioning and conformation of the DNA bound to the other active site was found to be better (closer to that bound to wild type HincII) when the center YR step is TA (pink, Fig. 2(e)) than CG (light blue, Fig. 2(b)). In addition, the overlap of the CSPS was found to show the opposite trend, being better with CG at the center YR step (light blue, Fig. 2(c)) than with TA (pink, Fig. 2(f)). As a result, a hypothesis was put forward that disruption of the CSPS leads to better active site alignment, and consequently faster cleavage. Since the stacking energy of adenine is weaker than that of guanine²⁹, the TA step unstacks more readily than CG, resulting in better active site alignment, and faster DNA cleavage.

New Structures

The distortions at the actual site of the mutation are very similar in all structures of Q138F HincII bound to DNA (Fig. 1(d), steps 1–5 shown in purple circles), however the response to misalignments at the protein-DNA interface, particularly the positions of the 3'P and SP, vary depending on DNA sequence and Ca²⁺ binding¹⁵ (a detailed description is described in the Supplementary Material). We have determined five new structures of Q138F HincII with three different DNA sequences (TA, CG, and GC DNA) and three different metal ions (Ca²⁺, Mg²⁺, and Mn²⁺). These structures have been prepared by cocrystallization with the metal ions. The previously analyzed structures¹⁵ were determined from crystals prepared by soaking preformed crystals of Q138F HincII/DNA with a solution containing 50 mM Ca²⁺, and are asymmetric (light blue, Fig. 2(a)–(b), pink, Fig. 2(d)–(e)), having the DNA bound differently in the two active sites of the enzyme dimer. This is unlike structures determined of wild type HincII, also prepared from soaking metal ions into preformed crystals of the enzyme and DNA,

which are identical within experimental error. The new structures presented in the current work were determined using crystals where all components, including the metal ions, were present prior to, as well as during the crystallization.

The structures of Q138F HincII bound to TA or CG DNA and cocrystallized with Ca^{2+} show significant differences from those prepared by soaking Ca^{2+} into preformed crystals (Fig. 2). Panel (a) and (d) of Fig. 2 show that the conformation and positioning of the DNA in the active sites of subunit A is very similar in the structures whether derived from soaking or cocrystallization. This common conformation has the 3'P making a close approach and ligating the Ca^{2+} ion (dashed lines, Fig. 2(a), (d)). Significantly, the oxygen atom of the 3'P that ligates the Ca^{2+} overlaps with the position of the putative nucleophile of the DNA cleavage reaction (pink sphere, Fig. 2(a), blue sphere, Fig. 2(d)). Such a conformation would be of low reactivity due to the interference with nucleophile binding and positioning.

The conformation and positioning of the DNA bound to the other subunit differs in the various structures with Ca^{2+} (Fig. 2(b), Fig. 2(e)). The position of the DNA in the active sites shown in Fig. 2(b) and (e) was used in the original analysis¹⁵, which concluded that the displacement is greater when the DNA contains CG (light blue, Fig. 2(b)) at the center step than TA (pink, Fig. 2(e)). The new structures show a decreased degree of mispositioning of the DNA in the active sites, however, the SP groups are still more mispositioned when the DNA contains CG at the center step (dark blue Fig. 2(a)–(b)) than TA (red, Fig. 2(d)–(e)). Similarly, though the CSPA is better formed in the new structures, the CSPA is still better formed with CG at the center step (Fig. 2(c), (f)). The asymmetry found in the structures derived from soaking Ca^{2+} into pre-formed crystals is absent in Q138F/TA/ Ca^{2+} (cocrystal) (red, Fig. 2(d)–(e)). Instead, the active sites of both subunits present the same conformation as that of the A subunits of the soak (pink, Fig. 2(d)). The structure of Q138F/CG/ Ca^{2+} (cocrystal) (dark blue, Fig. 2(a)–(b)) is less asymmetric than the corresponding soak (light blue, Fig. 2(a)–(b)). Importantly, the analysis of the structures described above shows a correlation between correct (i.e. like that in wild type HincII) positioning of the SP, and the rotation and positioning of the 3'P. In all structures with correct positioning of the SP, the 3'P is rotated, shifted toward the active site, and ligated to the Ca^{2+} . However, in none of the structures with a mispositioned SP does this interaction occur.

Q138F HincII was also cocrystallized with metal ions capable of supporting cleavage activity, namely Mg^{2+} and Mn^{2+} , and TA DNA. The DNA is cleaved in the Mn^{2+} bound structure, but uncleaved in that with Mg^{2+} . In the Q138F/TA/ Mg^{2+} structure, the SP of either DNA strand is mispositioned in both active sites, and the 3'P of either strand is rotated and shifted (orange, Fig. 3(a)–(b)). While two metal ions are expected to bind, in sites M1 and M2, only a single putative Mg^{2+} was located, in the M1 site (Q138F/TA/ Mg^{2+} , orange, Fig. 3(a)–(b)). The resolution of the diffraction data made unambiguous identification difficult, but the DNA is found to bind deeper in the active site than in the structure without any bound metal ions (brown, Fig. 3(a)–(b)). It is possible that only partial binding of Mg^{2+} occurs in the enzyme-DNA complex. In addition, though cocrystallized with the metal ion, the structure shows asymmetry, having different conformations in the two active sites. The oxygen of the 3'P is turned toward and approaches the nucleophile binding site but is more distant (1.0–1.3 Å) from the nucleophile than in the Q138F/TA/ Ca^{2+} structure (red, Fig. 2(d)–(e), 0.6 Å). The CSPA is disrupted (orange, Fig. 3(f)) to a much greater extent than in the Q138F/TA/ Ca^{2+} structure (red, Fig. 3(f)). The absence of DNA cleavage in the Q138F/TA/ Mg^{2+} structure suggests that crystallization conditions, including the solution conditions and the temperature (17°C), may have resulted in a greatly reduced DNA cleavage rate, perhaps due to weak binding of Mg^{2+} at the concentration of 2.5 mM.

The conformation and positioning of the DNA in the Q138F/TA/Mn²⁺ structure (purple, Fig. 4(a)–(f), 5(a)–(c)) is more symmetric, and the CSPA is found to be highly disrupted (purple, Fig. 5(c)). The positioning of the DNA in the active sites resembles that bound to the wild type enzyme, much more than any other structure of Q138F/TA (Fig. 4(a)–(b)). Having passed through the active conformation necessary to cleave DNA, the structure may more closely resemble the active conformation. Comparison with the structures of Q138F HincII bound to uncleaved DNA suggests two distinct features that appear to relate to the reduced DNA cleavage activity of the mutant enzyme, one is the mispositioning of the SP, relative to that found in the wild type HincII/DNA structure. The second is the orientation and position of the 3'P. This group is thought to play a catalytic role¹⁷, and hence positioning is likely to be important. Further, the position and orientation of this group in many of the Q138F HincII structures would be inhibitory for catalysis, as it overlaps with the expected position of the nucleophile (Fig. 1(d)). The many examples of this conformation found in Q138F/DNA structures indicate that the 3'P is rotated as a direct result of the Q138F mutation (Fig. 1(d)), and also show a correlation between correct positioning of the SP and the rotation/positioning of the 3'P; when the SP is correctly positioned, the 3'P is rotated and positioned such to overlap with the nucleophile binding site. Opening of the nucleophile binding site, while keeping the correct position of the SP, requires an increase in the separation of the 3'P and SP (Fig. 7(b)). This is only accomplished in the Q138F/TA/Mn²⁺ structure with the large disruption of the CSPA (purple, Fig. 4(c)).

The various structures of Q138F HincII bound to DNA and metal ions can be fit into the cartoon in Fig. 8. Distances between selected atoms in the structures are given in Table 2. The distance between the alpha carbon atoms of residue 138 of each subunit is used to describe the relative positioning of the two subunits in the dimer. The distance between the beta carbon atoms of residue 127 of each subunit measures the distance between the active sites in the dimer. The distance listed for CSPA is the smallest distance measured between atoms of the bases at nucleotide 8 of the bound DNA. A cartoon depicting the conformation in the wild type HincII/DNA structures (Fig. 8(a)) and relevant distances (Table 2) is given for reference. In Fig. 8(b), the symmetric conformation found in Q138F/TA/Ca²⁺ (cocrystal) is depicted, where the SP (red ball, Fig. 8(b)) of each DNA strand is correctly positioned in each active site (grey sphere, Fig. 8(b)), however the 3'P (yellow box, Fig. 8(b)) is also shifted into the active site. Some disruption of the CSPA may be necessary for the conformation of Fig. 8(b), as is found in Q138F/TA/Ca²⁺ (cocrystal) (red, Fig. 2(f)), and may be the reason Q138F/CG DNA structures do not exhibit this conformation, as the CSPA with two guanine bases requires more energy to disrupt. The mutation at residue 138 causes the Q138F HincII dimer to “open” to relieve the potential steric conflicts at the protein DNA interface. However, the subunits are oriented to position the active sites of the two subunits at the SP groups of the bound duplex DNA. Therefore, the distance measured between residues 138 of each subunit of Q138F/TA/Ca²⁺ (cocrystal) is longer than in wtHincII/CG/Ca²⁺ (28.4 Å vs. 26.7 Å). However, the distance between the active sites, measured from the alpha carbon of residue 127 of each subunit, is similar in Q138F/TA/Ca²⁺ (cocrystal) and wtHincII/CG/Ca²⁺ (26.8 Å vs. 26.6 Å). This conformation is inactive, according to the model, since the 3'P interferes with nucleophile binding. It may also only occur in structures with Ca²⁺, since the double ligation of the SP and 3'P is only possible with the larger Ca²⁺ ion.

The asymmetric conformation found in all of the other Q138F/DNA structures, excluding Q138F/TA/Ca²⁺ (cocrystal) and Q138F/TA/Mn²⁺, is illustrated by Fig. 8(c). The positioning and orientation of the DNA bound to the active site of subunit A is similar to that found in both subunits of Fig. 8(b). However, the SP of the other strand of the DNA (red circle, Fig. 8(c)) is not placed correctly in the active site (grey sphere, Fig. 8(c)) of the B subunit. The CSPA is relatively maintained. The dimer is opened up, with distances between residues 138 of 27.5–28.4 Å (wild type HincII/DNA = 26.7 Å) and active site residues 127 of 27.2–28.7 Å (wild

type HincII/DNA = 26.6 Å). This conformation is also inactive, according to the model, since the SP is mispositioned in one subunit, and the 3'P interferes with the nucleophile binding of the other.

Finally, the conformation shown in Fig. 8(d) is that found in Q138F/TA/Mn²⁺, the DNA cleavage competent conformation. The DNA is cleaved in Q138F/TA/Mn²⁺, indicating that the SP groups (red spheres, Fig. 8(d)) had once been positioned correctly in the active sites (grey spheres, Fig. 8(d)) of both subunits. The subunits are still separated to a greater extent than they are in the wild type structure (28.1 Å vs. 26.7 Å), as are the active sites (28.2 Å vs. 26.6 Å), however, the 3'P groups (yellow boxes) are shifted away from the respective active sites allowing binding of the nucleophile to the metal ion at M1 (Fig. 7(a)–(b)). The CSPA is largely disrupted (4.5 Å vs. 3.4 Å), allowing the separation of the 3'P from the SP (Fig. 7(b)). Finally, the conformation found in Q138F/TA/Mg²⁺ may better be represented as in-between the conformations shown in Fig. 8(c) and (d), as the CSPA is more disrupted than in the Ca²⁺ structures with the TA DNA (Fig. 3(f)), the SP groups are both slightly mispositioned (orange, Fig. 4(d)–(e)), and the 3'P groups of both strands take on the conformations seen in Q138F/TA/Mn²⁺ (Fig. 8(d), purple, Fig. 4(d)–(e)).

Since disruption of the CSPA appears to be linked to the formation of the active conformation (Fig. 8(d)), DNA containing a pyrimidine at nucleotide 8 should be more reactive with the Q138F HincII enzyme, as the CSPA cannot form (green, Fig. 5(f)). However, DNA cleavage experiments show that Q138F HincII does not cleave the noncognate GC DNA under any of the conditions tested. In addition, the structure of Q138F HincII bound to noncognate GC DNA and Ca²⁺ is best described by the cartoon shown in Fig. 7(b), an inactive conformation. The 3'P is rotated and positioned very close to the bound Ca²⁺, ligating the ion singly in subunit A, and doubly in subunit B (green, Fig. 5(d)–(e)). This position of the 3'P overlaps with the site of the putative nucleophile (blue sphere, Fig. 5(d)–(e)), and separation of the SP from the 3'P, to allow nucleophile binding (Fig. 7(b)) may not be possible within the complex of Q138F HincII bound to noncognate GC DNA and Ca²⁺. Binding measurements show that Q138F HincII binds only 300 fold more weakly to noncognate GC DNA than cognate in the presence of Ca²⁺, however no binding is observed in the absence of Ca²⁺, or in the presence of Mg²⁺ or Mn²⁺. The ligation of Ca²⁺ by the SP and 3'P may increase the binding affinity of Q138F HincII for the noncognate GC DNA, allowing trapping of the complex in the crystal. However, this complex may not form with Mg²⁺, as the smaller diameter of Mg²⁺ may not allow the double ligation of adjacent phosphate groups. The double ligation, involving both the SP and 3'P, is not observed in the Q138F/TA/Mg²⁺ structure (orange, Fig. 4(d)–(e)).

Therefore the lack of noncognate GC DNA cleavage by Q138F HincII may be due to the reduced binding affinity in the presence of Mg²⁺. The decreased affinity of Q138F HincII for the noncognate GC DNA, compared to CG or TA DNA, may lie in the loss of the hydrogen bond to the Pur 8 N7 of each strand, as well as differences in the energy of DNA distortion of the different sequences. Loss of the hydrogen bond to the Pur 8 N7, as well as to the Ade 9 N6, in the N141A mutant of HincII, reduces cognate DNA binding affinity at least 900 fold (Little, *et al.*, *submitted for publication*). The ability to ligate Ca²⁺ with both SP and 3'P in the Q138F HincII/DNA structures may compensate for, and hide the true impact of the Q138F mutation on the DNA binding affinity in the presence of Mg²⁺. Such a possibility is encompassed in the concept of structural adaptability, originally invoked to explain the apparent lack of effect on binding affinity by this mutation¹⁵. Although the K_D of wild type HincII with noncognate GC DNA is similar to that with Q138F HincII, cocrystallization proved unsuccessful, and the stability of the complex appears to be much weaker, as much more dissociation occurs during electrophoresis with wild type but not Q138F HincII (Fig. S1). The similar equilibrium constants may result from different on and off rates, and the complex of wild type HincII with noncognate GC DNA, also inactive, may have a fundamentally different

structure, possibly involved in DNA scanning and translocating (Little, *et al.*, *unpublished results*).

The smaller size of Mg^{2+} versus Ca^{2+} could be the source of the difference in observed ligation, and the Mg^{2+} , as well as the ligated water nucleophile, would be expected to occupy slightly different positions than found in structures with Ca^{2+} . The slight difference in positioning may be why the close approach of the 3'P to the putative nucleophile (predicted from wtHincII/CG/ Ca^{2+}) in Q138F/TA/ Mn^{2+} , although very close (1.3–1.4 Å), allows for the DNA cleavage activity, while that of Q138F/TA/ Ca^{2+} , with the distance of 0.6 Å, does not, according to the model.

The Q138F HincII/DNA complex is an ensemble sampling different conformational states. Unlike wild type HincII bound to cognate DNA, many of the conformations sampled by Q138F HincII appear to be inactive. The active conformation is achieved when the CSPS can be pulled apart, to allow separation of the SP from the 3'P and binding of the nucleophile (Fig. 7(b)). This happens more readily with adenine at Pur 8, as adenine has a weaker stacking energy than guanine³⁰. If the single turnover DNA cleavage rate constants represent the rates of this unstacking, then the difference between unstacking the CSPS with two guanines relative to that with two adenines is the difference in the free energy of activations derived from the rate constants¹⁵, or approximately 1.0 kcal/mol. The evolution of new sequence specificities in DNA binding proteins may occur through similar changes in the protein-DNA interface, allowing ample opportunity for DNA conformational energy to play a role in specificity.

Materials and Methods

Protein purification

Purification of wild type and mutant HincII was performed as described³¹. Purified enzyme was dialyzed exhaustively against buffer A (50 mM bis-trispropane, 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 prior to use. Purity has been estimated from overloaded SDS-PAGE as greater than 99.6% and 95.9% for Q138F and wild type HincII, respectively.

Preparation of DNA substrates

Both HEX (hexachlorofluorescein)-labeled, and non-HEX-labeled oligonucleotides were obtained from commercial synthetic sources and purified using C18 reverse phase HPLC³². The concentration of each single strand was measured spectrophotometrically, with extinction coefficients calculated from standard values for the nucleotides³³ and $32.2\text{ mM}^{-1}\text{ cm}^{-1}$ (at 260 nm) for the HEX molecule. Equimolar amounts of each strand (typically 100–400 mM each) were mixed and annealed by heating to 90°C for 10 minutes, followed by slow-cooling to 4°C over 4–5 hours in a thermocycler. Radiolabeled oligonucleotides were prepared using γ - ^{32}P -ATP (MP Biomedicals, Irvine, CA) and T4 polynucleotide kinase (Fermentas, Hercules, CA). Excess ^{32}P -ATP removed using Micro Bio-spin 6 Columns P-polyacrylamide P-6 gel (Biorad, Hercules, CA).

The DNA sequences used in the binding and single turnover assays are shown below, with the recognition sequence in bold, center step underlined, and potential HincII cleavage sites marked by | :

Noncognate GC:

5'-GGG CCG **GTG|CAC** CAA C-3'

3'-CCC GGCC**AC|GTG** GTT G-5'

Abasic site DNA:

5'-GGG CCG **GTC**XAC CAA C-3'

3'-CCC GGC **CAG**XTG GTT G-5'

Oligonucleotides used in crystallizations are shown below (all are self complementary):

TA DNA: Cognate DNA with TA at center step:

5'-G CCG **GTT**AAC CGG C-3'

CG DNA: Cognate DNA with CG at center step:

5'-G CCG **GTC**GAC CGG C-3'

GC DNA: Noncognate DNA with GC at center step:

5'-G CCG **GTG**CAC CGG C-3'

Binding affinity measurement using gel shift

The equilibrium dissociation constant K_D was measured with the 5' ^{32}P radiolabeled 16 bp DNA by a gel shift assay^{13; 34–37}. The binding buffer was composed of 50 mM bis-tris propane (pH 7.5 at 25°C), 140 mM NaCl, 5 mM CaCl_2 (or 20 mM MgCl_2), 1 mM DTT mixed with 50 mM bis-tris propane (pH 7.5 at 25°C), 5 mM CaCl_2 (or 20 mM MgCl_2), 1 mM DTT to the conductivity (19.00 mS/cm) of 50 mM bis-tris propane (pH 7.5 at 25°C), 140 mM NaCl, 1 mM DTT. Protein concentration (in terms of monomers) was measured spectrophotometrically using the extinction coefficient $\epsilon=36,840 \text{ cm}^{-1}\text{M}^{-1}$ at 280 nm. The calculated concentration was then adjusted for molarity of dimers. Protein and DNA were incubated at 37°C for 30 minutes, rapidly mixed with pre-warmed 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 subjected to electrophoresis at 190 V and 4°C with recirculation of the running buffer. Gels were autoradiographed using a phosphorimage plate (Molecular Dynamics). The data were fit by nonlinear regression to the equation (derived from the quadratic solution to the equilibrium association binding equation):

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

where A represents the intensity of the protein bound DNA band. Both $A_{\max}$ and $A_{\min}$ are fit in addition to K_D . The DNA concentration, O_T , was held constant, and P_T , the total protein concentration, varies with each measurement of A. Each reported K_D is the average of at least three independent measurements.

Binding affinity measurements by fluorescence polarization

Fluorescence polarization anisotropy measurements were made using a PC1 fluorimeter (ISS, Inc., Champaign, IL) equipped with a double grating excitation monochromator. Excitation was performed at 535 nm and emission recorded using a 50.8 mm diameter 570 nm cut-on filter with 580–2750 nm transmittance range (ThermoOriel Inc., no. 59510). Slit widths of 1 mm were used and HEX-labeled DNA at 1 nM in 2 mL binding buffer. Measurements were performed at room temperature after the addition of increasing amounts of enzyme. Sample stirring was performed during data measurement. Each measurement was recorded after averaging the signal over 5 minutes. Simple binding data were fit using the program Kaleidagraph (Macromedia, Inc.). Competition experiments were performed using 1 nM HEX-DNA, 221 nM Q138F HincII and varied concentrations of an unlabeled competing oligonucleotide, and the data fit using the program Scientist (MicroMath Scientific Software) as described³⁸.

Single turnover DNA cleavage reactions

Single turnover DNA cleavage measurements were attempted for wild type and Q138F HincII with the noncognate GC DNA. Assays were performed as described¹⁵ with 10 nM ³²P radiolabeled DNA incubated with 3 μM enzyme at 37°C for up to 24 hours. The reaction conditions included a buffer prepared by mixing 50 mM bis-tris propane (pH 7.5 at 25°C), 20 mM MgCl₂, 0.1 mg/ml BSA, 1 mM DTT with 50 mM bis-tris propane (pH 7.5 at 25°C), 20 mM MgCl₂, 0.1 mg/ml BSA, 1 mM DTT, 500 mM NaCl until the conductivity of the reference buffer 50 mM bis-tris propane (pH 7.5 at 25°C), 0.1 mg/ml BSA, 1 mM DTT, 140 mM NaCl was reached.

Crystallization, Diffraction Data Collection, and Structure Refinement

Crystals were prepared with Q138F HincII and TA, CG, and GC DNA, in the presence of 5 mM CaCl₂. In addition, Q138F HincII was cocrystallized with 2.5 mM MnCl₂ or MgCl₂ with the TA DNA. DNA was mixed at a molar ratio of 2:1 with 16 to 23 mg/ml of Q138F HincII. The crystallization conditions included 1.5 to 3.0 μl of the protein:DNA mixture with 1.0 to 1.5 μl of the precipitating solution: 25 to 21 % PEG 4K, 0.1 M imidazole buffer (pH 6.5) or 0.1 M HEPES buffer (pH 7.5), 0.15 M NaCl, per drop and placed over 1 ml of the precipitating solution. Crystals grow overnight to 1 week at 17°C. The crystals were then exchanged into a cryoprotection solution (25% PEG 4K, 0.1M buffer (pH 6.5 or 7.5 above), 0.3 M NaCl, 30% glycerol) and flash-frozen in liquid nitrogen. X-ray diffraction amplitudes were measured using synchrotron radiation at the Stanford Synchrotron Light Source (SSRL) BL9-1 (Q138F/TA/Mn²⁺) or BL9-2 (all others). Data collection was performed while maintaining the crystal at 100K. Image processing and data reduction were performed with MOSFLM³⁹ and SCALA⁴⁰.

All structures are isomorphous with the previously published structures of Q138F with cognate DNA in the P2₁ space group¹⁵, and were refined with the program CNS⁴¹ and PHENIX⁴². Manual model building was performed using XtalView⁴³.

Structural analyses

Superpositions were performed with LSQKAB⁴⁴ or PyMOL⁴⁵. Figures were prepared with PyMOL⁴⁵. Solvent accessible surface area (SASA) was calculated using the algorithm of Lee and Richards⁴⁶ with a probe radius of 1.4 Å as implemented in CNS⁴¹, and only the atoms of the two nucleotide 8 bases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

bp, base pairs; DTT, dithiothreitol; SDS-PAGE, sodium dodecyl sulfate polyacrylamide electrophoresis; BSA, bovine serum albumin.

Acknowledgments

This work was supported by the grant (to N.C.H.) NIH 5R01GM066805. 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 U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

References

1. Garvie CW, Wolberger C. Recognition of specific DNA sequences. *Mol Cell* 2001;8:937–946. [PubMed: 11741530]
2. Otwinowski Z, Schevitz RW, Zhang RG, Lawson CL, Joachimiak A, Marmorstein RQ, Luisi BF, Sigler PB. Crystal structure of trp repressor/operator complex at atomic resolution [published erratum appears in *Nature* 1988 Oct 27;335(6193):837]. *Nature* 1988;335:321–329. [PubMed: 3419502]
3. Segal DJ, Crotty JW, Bhakta MS, Barbas CF 3rd, Horton NC. Structure of Aart, a designed six-finger zinc finger peptide, bound to DNA. *J Mol Biol* 2006;363:405–421. [PubMed: 16963084]
4. Koudelka GB, Mauro SA, Ciubotaru M. Indirect readout of DNA sequence by proteins: the roles of DNA sequence-dependent intrinsic and extrinsic forces. *Prog Nucleic Acid Res Mol Biol* 2006;81:143–177. [PubMed: 16891171]
5. Blakaj DM, Kattamuri C, Khrapunov S, Hegde RS, Brenowitz M. Indirect readout of DNA sequence by papillomavirus E2 proteins depends upon net cation uptake. *J Mol Biol* 2006;358:224–240. [PubMed: 16513133]
6. Rice PA, Yang S, Mizuuchi K, Nash HA. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* 1996;87:1295–1306. [PubMed: 8980235]
7. Djuranovic D, Oguey C, Hartmann B. The role of DNA structure and dynamics in the recognition of bovine papillomavirus E2 protein target sequences. *J Mol Biol* 2004;339:785–796. [PubMed: 15165850]
8. Chen S, Gunasekera A, Zhang X, Kunkel TA, Ebright RH, Berman HM. Indirect readout of DNA sequence at the primary-kink site in the CAP- DNA complex: alteration of DNA binding specificity through alteration of DNA kinking. *J Mol Biol* 2001;314:75–82. [PubMed: 11724533]
9. Bareket-Samish A, Cohen I, Haran TE. Direct versus indirect readout in the interaction of the trp repressor with non-canonical binding sites. *Journal of Molecular Biology* 1998;277:1071–1080. [PubMed: 9571023]
10. Bareket-Samish A, Cohen I, Haran TE. Signals for TBP/TATA box recognition. *Journal of Molecular Biology* 2000;299:965–977. [PubMed: 10843851]
11. Garvie CW, Phillips SE. Direct and indirect readout in mutant Met repressor-operator complexes. *Structure with Folding & Design* 2000;8:905–914. [PubMed: 10986458]
12. Schumacher MA, Hurlburt BK, Brennan RG. Crystal structures of SarA, a pleiotropic regulator of virulence genes in *S. aureus*. *Nature* 2001;409:215–219. [PubMed: 11196648]
13. Martin AM, Sam MD, Reich NO, Perona JJ. Structural and energetic origins of indirect readout in site-specific DNA cleavage by a restriction endonuclease. *Nature Structural Biology* 1999;6:269–277.
14. Pingoud A, Jeltsch A. Structure and function of type II restriction endonucleases. *Nucleic Acids Res* 2001;29:3705–3727. [PubMed: 11557805]
15. Joshi HK, Etzkorn C, Chatwell L, Bitinaite J, Horton NC. Alteration of sequence specificity of the type II restriction endonuclease HincII through an indirect readout mechanism. *J Biol Chem* 2006;281:23852–23869. [PubMed: 16675462]
16. Horton NC, Dorner LF, Perona JJ. Sequence selectivity and degeneracy of a restriction endonuclease mediated by DNA intercalation. *Nat Struct Biol* 2002;9:42–47. [PubMed: 11742344]
17. Etzkorn C, Horton NC. Ca²⁺ binding in the active site of HincII: implications for the catalytic mechanism. *Biochemistry* 2004;43:13256–13270. [PubMed: 15491133]
18. Etzkorn C, Horton NC. Mechanistic insights from the structures of HincII bound to cognate DNA cleaved from addition of Mg²⁺ and Mn²⁺. *J Mol Biol* 2004;343:833–849. [PubMed: 15476804]
19. Galburt EA, Stoddard BL. Catalytic mechanisms of restriction and homing endonucleases. *Biochemistry* 2002;41:13851–13860. [PubMed: 12437341]
20. Horton NC, Perona JJ. Making the most of metal ions. *Nature Structural Biology* 2002;9:42–47. [comment].
21. Dupureur C. An Integrated Look at Metallonuclease Mechanism. *Current Chemical Biology* 2008;2:159–173.
22. Horton, NC. DNA Nucleases. In: Rice, PA.; Correll, CC., editors. *Protein-Nucleic Acid Interactions Structural Biology*. Cambridge, UK: The Royal Society of Chemistry; 2008. p. 333-363.

23. Pingoud A, Fuxreiter M, Pingoud V, Wende W. Type II restriction endonucleases: structure and function. *Cell. Mol. Life Sci* 2005;62:685–707. [PubMed: 15770420]
24. Horton NC, Otey C, Lusetti S, Sam MD, Kohn J, Martin AM, Ananthnarayan V, Perona JJ. Electrostatic contributions to site specific DNA cleavage by EcoRV endonuclease. *Biochemistry* 2002;41:10754–10763. [PubMed: 12196013]
25. Selent U, Ruter T, Kohler E, Liedtke M, Thielking V, Alves J, Oelgeschlager T, Wolfes H, Peters F, Pingoud A. A site-directed mutagenesis study to identify amino acid residues involved in the catalytic function of the restriction endonuclease EcoRV. *Biochemistry* 1992;31:4808–4815. [PubMed: 1591242]
26. Horton NC, Newberry KJ, Perona JJ. Metal ion-mediated substrate-assisted catalysis in type II restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* 1998;95:13489–13494. [PubMed: 9811827]
27. Horton NC, Perona JJ. DNA cleavage by EcoRV endonuclease: two metal ions in three metal ion binding sites. *Biochemistry* 2004;43:6841–6857. [PubMed: 15170321]
28. Horton NC, Connolly BA, Perona JJ. Inhibition of EcoRV endonuclease by deoxyribo-3'-S-phosphorothiolates: a high resolution x-ray crystallographic study. *J. Am. Chem. Soc* 2000;122:3314–3324.
29. Hobza P, Sponer J. Structure, energetics, and dynamics of the nucleic acid base pairs: Nonempirical ab initio calculations. *Chemical Reviews* 1999;99:3247–3276. [PubMed: 11749516]
30. Sponer J, Berger I, Spackova N, Leszczynski J, Hobza P. Aromatic base stacking in DNA: From ab initio calculations to molecular dynamics simulations. *Journal of Biomolecular Structure & Dynamics* 2000:383–407.
31. Horton NC, Dorner LF, Schildkraut I, Perona JJ. Crystallization and preliminary diffraction analysis of the HincII restriction endonuclease-DNA complex. *Acta Crystallographica. Section D: Biological Crystallography* 1999;55:1943–1945.
32. Aggarwal AK. Crystallization of DNA binding proteins with oligodeoxynucleotides. *Methods: A Companion to Methods in Enzymology* 1:83–90.
33. Fasman, GD. *CRC Handbook of Biochemistry and Molecular Biology*. 3rd edit. Cleveland, OH: CRC; 1975.
34. Fried M, Crothers DM. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res* 1981;9:6505–6525. [PubMed: 6275366]
35. Taylor JD, Badcoe IG, Clarke AR, Halford SE. EcoRV restriction endonuclease binds all DNA sequences with equal affinity. *Biochemistry* 1991;30:8743–8753. [PubMed: 1909572]
36. Thielking V, Selent U, Kohler E, Wolfes H, Pieper U, Geiger R, Urbanke C, Winkler FK, Pingoud A. Site-directed mutagenesis studies with EcoRV restriction endonuclease to identify regions involved in recognition and catalysis. *Biochemistry* 1991;30:6416–6422. [PubMed: 1647200]
37. Martin AM, Horton NC, Lusetti S, Reich NO, Perona JJ. Divalent metal dependence of site-specific DNA binding by EcoRV endonuclease. *Biochemistry* 1999;38:8430–8439. [PubMed: 10387089]
38. Reid SL, Parry D, Liu HH, Connolly BA. Binding and recognition of GATATC target sequences by the EcoRV restriction endonuclease: a study using fluorescent oligonucleotides and fluorescence polarization. *Biochemistry* 2001;40:2484–2494. [PubMed: 11327870]
39. Leslie AG. Integration of macromolecular diffraction data. *Acta Crystallographica. Section D: Biological Crystallography* 1999;55:1696–1702.
40. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 1994;50:760–763. [PubMed: 15299374]
41. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica. Section D: Biological Crystallography* 1998;54:905–921.
42. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 2002;58:1948–1954. [PubMed: 12393927]

43. McRee DE. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *Journal of Structural Biology* 1999;125:156–165. [PubMed: 10222271]
44. Kabsch W. A solution for the best rotation to relate two sets of vectors. *Acta Cryst* 1976;A32:922–923.
45. DeLano, WL. *The PyMOL User's Manual*. Palo Alto, CA: DeLano Scientific; 2002.
46. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* 1971;55:379–400. [PubMed: 5551392]

Figure 1.

(a). Ribbon diagram of HincII dimer (subunits shown in light and dark green) showing bound DNA (pale beige, red, and blue), site of Gln 138 (cyan) intercalation, position of Ca^{2+} (yellow) in the two active sites, and center step of the recognition sequence (red, blue). (b). The numbering scheme used throughout the paper for the HincII recognition sequence. Y indicates either C or T, R indicates either G or A. (c). Structure of the pyrimidine-purine step as found in the structure of wild type HincII bound to cognate DNA containing GTCGAC (left) and in B-form DNA (right). The Y 7 is unstacked from R 8 within the same strand in the HincII bound DNA, and the center step purines (R 8 and R 8') have increased stacking across the DNA duplex, forming a cross-strand purine stack (CSPS). (d). Overlay of portions of the structures

of wtHincII/CG/Ca²⁺ (white) and Q138F/TA/Ca²⁺ (cocrystal) (color), nuc, putative water nucleophile of the DNA cleavage reaction, SP, phosphate at the scissile phosphodiester bond. Arrows emphasize the largest structural differences: the side chain of residue 138 sits differently between the adjacent cytosine bases at the site of intercalation and the main chain around Phe 138 is shifted away from the DNA to accommodate this different position (1). As a result, the side chain of Ala 137 no longer makes a van der Waals contact to the DNA (2), which may be the reason for the altered pucker of the deoxyribose at Cyt 10, from C1'exo to C2'endo (3). The different pucker at Cyt 10 causes a twisting of the sugar-phosphate backbone of the DNA strand (4) propagating to the phosphate between the Pyr 7 and Pur 8, potentially resulting in the O5' of Ade 9 jutting into the space normally occupied by the side chain of Thr 130 (5). To avoid the steric conflict, the phosphate of Ade 9 is rotated blocking the position of the putative nucleophilic water (6).

Figure 2.

(a). Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/CG/Ca²⁺ (cocystal) (dark blue), and wtHincII/CG/Ca²⁺ (grey) and subunit B of Q138F/CG/Ca²⁺ (soak) (light blue). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of each subunit. Ca²⁺ are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the phosphorus of the phosphate at the scissile phosphodiester bond (SP) is shown as a pink sphere marked nuc, with the direction of attack as a pink arrow. Dashes indicate direct ligation to a Ca²⁺, or hydrogen bond from the phosphate 3' of the SP (3'P) to the nucleophilic water. (b). As in (a), but using the alpha carbon atoms of residues 114, 127, and 129 of subunit B of Q138F/CG/Ca²⁺ (cocystal) (dark blue) and wtHincII/CG/Ca²⁺ (grey), or subunit A of Q138F/CG/Ca²⁺ (soak) (light blue) in the superposition. (c). Stereo view of the two center step bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/CG/Ca²⁺. Colors as in (a). (d). Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/TA/Ca²⁺ (soak) (pink), Q138F/TA/Ca²⁺ (cocystal) (red), and wtHincII/CG/Ca²⁺ (grey). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of subunit A. Ca²⁺ are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the SP is shown

as a blue sphere marked nuc, with the direction of attack as a blue arrow. Dashes indicate direct ligation to a Ca^{2+} , or hydrogen bond from the 3'P to the nucleophilic water. **(e)**. As in (a), but subunit B. **(f)**. Stereo view of the two center step bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/CG/ Ca^{2+} . Colors as in (e).

Figure 3.

(a). Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/TA/Mg²⁺ (orange), Q138F/TA (brown), and wtHincII/CG/Ca²⁺ (grey). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of subunit A. Metal ions are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the phosphorus of the phosphate at the scissile phosphodiester bond (SP) is shown as a blue sphere marked nuc, with the direction of attack as a blue arrow. Dashes indicate direct ligation to a metal ion, M1, or hydrogen bond from the phosphate 3' from the SP (3'P) to the nucleophilic water. (b). As in (a), but subunit B. (c). Stereo view of the two center step bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/CG/Ca²⁺. Colors as in (a). (d). Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/TA/Mg²⁺ (orange), Q138F/TA/Ca²⁺ (cocystal) (red), and wtHincII/CG/Ca²⁺ (grey). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of subunit A. Metal ions are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the SP is shown as a blue sphere marked nuc, with the direction of attack as a blue arrow. Dashes indicate direct ligation to a metal ion, or hydrogen bond from the 3'P to the nucleophilic water. (e). As in (d), but subunit B. (f). Stereo view of the two center

step bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/
CG/Ca²⁺. Colors as in (d).

Figure 4.

(a). Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/TA/Mn²⁺ (purple), Q138F/CG/Mn²⁺ (black), and wtHincII/CG/Ca²⁺ (grey). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of subunit A. Metal ions are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the phosphorus of the phosphate at the scissile phosphodiester bond (SP) is shown as a blue sphere marked nuc, with the direction of attack as a blue arrow. Dashes indicate direct ligation to a metal ion, or hydrogen bond from the phosphate 3' from the SP (3'P) to the nucleophilic water. **(b).** As in (a), but subunit B. **(c).** Stereo view of the two center step bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/CG/Ca²⁺. Colors as in (a). **(d).** Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/TA/Mg²⁺ (orange), Q138F/TA/Mn²⁺ (purple), and wtHincII/CG/Ca²⁺ (grey). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of subunit A. Metal ions are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the SP is shown as a blue sphere marked nuc, with the direction of attack as a blue arrow. Dashes indicate direct ligation to a metal ion, or hydrogen bond from the 3'P to the nucleophilic water. **(e).** As in (d), but subunit B. **(f).** Stereo view of the two center step

bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/CG/
Ca²⁺. Colors as in (d).

Figure 5.

(a). Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/TA/Mn²⁺ (purple), Q138F/TA/Ca²⁺ (red), and wtHincII/CG/Ca²⁺ (grey). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of subunit A. Metal ions are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the phosphorus of the phosphate at the scissile phosphodiester bond (SP) is shown as a blue sphere marked nuc, with the direction of attack as a blue arrow. Dashes indicate direct ligation to a metal ion, or hydrogen bond from the phosphate 3' from the SP (3'P) to the nucleophilic water. (b). As in (a), but subunit B. (c). Stereo view of the two center step bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/CG/Ca²⁺. Colors as in (a). (d). Stereo view of the backbone of nucleotides 7–9 bound to the active site of subunit A of Q138F/GC/Ca²⁺ (green), Q138F/TA/Ca²⁺ (red), and wtHincII/CG/Ca²⁺ (grey). The superposition was done with the alpha carbon atoms of residues 114, 127, and 129 of subunit A. Metal ions are shown as spheres colored as in their parent structure. The position of the water molecule in wtHincII/CG/Ca²⁺ in-line for attack on the SP is shown as a blue sphere marked nuc, with the direction of attack as a blue arrow. Dashes indicate direct ligation to a metal ion, or hydrogen bond from the 3'P to the nucleophilic water. (e). As in (d), but subunit B. (f). Stereo view of the two center step

bases after superposition onto the base atoms of one of the nucleotide 8 bases of wtHincII/CG/
Ca²⁺. Colors as in (d).

Figure 6.

(a). Fit of the intensity of the unbound DNA from a gel shift experiment with 1 nM noncognate GC DNA and wild type HincII from 3.0×10^{-8} M to 7×10^{-6} M in the presence of 5 mM CaCl_2 . (b). Fluorescence polarization of a HEX-labeled abasic site DNA (1 nM) as a function of different concentrations of Q138F HincII. Data was fit 1:1 binding giving a K_D of 101 ± 4 nM. (c). Change in fluorescence polarization of HEX-labeled abasic site DNA in the presence of 221 nM Q138F HincII upon different amounts of unlabeled noncognate GC DNA. Data were fit to a competition equilibria giving a K_D for Q138F HincII and noncognate GC DNA of 321 nM.

Figure 7.

(a). Model of DNA cleavage mechanism of HincII. M1 and M2 mark the two metal ion binding sites. Dashes indicate direct ligation to a metal ion, or hydrogen bond from the 3'P to the nucleophilic water, shown as a pink sphere marked nuc. The scissile bond is colored red.

(b). Cartoon of the connection between the CSPS and the distance between the phosphorus of the phosphate at the scissile phosphodiester bond (SP) and the phosphate 3' from the SP (3'P). Dashes indicate direct ligation to the metal ion, M1, or hydrogen bond from the 3'P to the nucleophile (water or hydroxide molecule) shown as pink sphere. Left: inactive conformation with the 3'P blocking the nucleophile binding site on the M1 ion. Right: active conformation

where disruption of the CSPA separates SP from 3'P on each strand to allow the nucleophile to bind to the M1 ion. The scissile bond is shown in red.

Figure 8.

(a). Cartoon of wild type HincII bound to cognate DNA. cyan, residue 138, grey sphere, active site, red sphere, phosphate at the scissile phosphodiester bond (SP), yellow box, phosphate 3' from the SP (3'P), blue line, nucleotide 8 base. (b). Cartoon of Q138F/TA/Ca²⁺ (cocrystal) and Q138F/GC/Ca²⁺, where the SP of either DNA strand is bound correctly in each subunit active site, however, the 3'P blocks the nucleophile binding site. The CSPS is partly disrupted. (c). Cartoon of the conformation found in Q138F/CG/Ca²⁺ (soak), Q138F/TA/Mg²⁺, Q138F/CG/Ca²⁺ (cocrystal), Q138F/TA/Ca²⁺ (soak) where the disruption of the enzyme/DNA interface by the Q138F mutation has caused an increase in the separation between the subunits. Only one of the SP groups is correctly positioned in an active site (Subunit A), however, the 3'P

blocks the nucleophile binding site. The CSPA is maintained. **(d)**. Cartoon of the conformation found in Q138F/TA/Mn²⁺, where both of the SP groups are correctly positioned in each active site, and the 3'P does not block the nucleophile binding site. The CSPA is highly disrupted.

Table 1

Crystallographic Data Statistic

	Q138F/TA/Ca²⁺ cocrystal 1	Q138F/TA/Ca²⁺ cocrystal 2	Q138F/CG/Ca²⁺ cocrystal 1	Q138F/CG/Ca²⁺ cocrystal 2
Code	PDB ID 3E3Y RCSB048844	PDB ID 3E40 RCSB048846	PDB ID 3E41 RCSB048847	PDB ID 3E42 RCSB048848
Space Group	P2 ₁	P2 ₁	P2 ₁	P2 ₁
Cell	50.48 Å, 91.64 Å, 65.78 Å, β=104.91°	50.33 Å, 91.63 Å, 66.25 Å, β=104.77°	50.06 Å, 90.14 Å, 66.33 Å, β=104.6°	50.03 Å, 90.11 Å, 66.44 Å, β=104.4°
Resolution	2.13 Å	2.10 Å	2.73 Å	2.68 Å
Total	118,374	126,920	57,725	61,689
Observations	32,390	33,581	15,276	15,869
Unique				
Observations	99.9% (100%)	99.1% (98.7%)	100% (100%)	98.5% (97.9%)
% Complete				
I/sigma	5.6 (2.0)	5.6 (2.0)	11.2 (2.0)	12.4 (2.0)
Multiplicity	3.7	3.8	3.8	3.9
R _{merge} ¹	8.5% (35.6%)	8.1% (32.9%)	5.4% (37.5%)	4.8% (37.3%)
R _{cryst} ²	18.0%	16.9%	18.4%	18.8%
R _{free} ³	23.4%	23.2%	26.6%	28.3%
Overall B factor (Å ² , Wilson plot)	30.7	29.0	56.1	54.9
RMSD –bonds	0.005 Å	0.008 Å	0.009 Å	0.006 Å
RMSD angles	1.00°	1.20°	1.31°	1.14°
Number of water molecules	448	501	128	134
Number of metal ions	2 Ca ²⁺ , 3 Na ⁺	2 Ca ²⁺ , 3 Na ⁺	2 Ca ²⁺ , 1 Na ⁺	2 Ca ²⁺ , 2 Na ⁺
Protein residues	485	490	491	491
Nucleotides	28	28	28	28

	Q138F/TA/Mg²⁺ cocrystal	Q138F/TA/Mn²⁺ cocrystal	Q138F/CG/Ca²⁺ cocrystal
Code	PDB ID 3E43 RCSB048840	PDB ID 3E44 RCSB048850	PDB ID 3E45 RCSB048851
Space Group	P2 ₁	P2 ₁	P2 ₁
Cell	50.27 Å, 90.94 Å, 66.53 Å, β=105.1°	50.69 Å, 91.79 Å, 66.18 Å, β=105.3°	50.62 Å, 92.20 Å, 66.61 Å, β=105.19°
Resolution	2.73 Å	2.5 Å	2.78 Å
Total Observations	41,380	82,805	54,645
Unique Observations	14,388	22,903	14,947
% Complete	93.4% (94.7%)	99.4% (95.8%)	100% (100%)
I/sigma	9.7 (2.0)	9.8 (1.9)	5.1 (2.0)
Multiplicity	2.9	3.6	3.7
R _{merge} ¹	8.5% (49.0%)	5.5% (38.3%)	10.5% (35.3%)
R _{cryst} ²	19.2%	18.0%	17.9%
R _{free} ³	27.6%	25.7%	27.4%
Overall B factor (Å ² , Wilson plot)	55.5	39.7	37.3
RMSD –bonds	0.005 Å	0.008 Å	0.008 Å
RMSD – angles	0.94°	1.21°	1.18°
Number of water molecules	106	292	363
Number of metal ions	2 Mg ²⁺ (putative)	4 Mn ²⁺	2 Ca ²⁺ , 1 Na ⁺
Protein residues	488	488	497
Nucleotides	28	28	28

¹ $R_{\text{merge}} = \sum_{\text{hkl}} (| \langle I_{\text{hkl}} \rangle - I_{\text{hkl}} |) / (\sum_{\text{hkl}} I_{\text{hkl}})$ where $\langle I_{\text{hkl}} \rangle$ is the average intensity over symmetry related and equivalent reflections and I_{hkl} is the observed intensity for reflection hkl.

² $R_{\text{cyst}} = \sum_{\text{hkl}} (|F_{\text{obs}}| - |F_{\text{calc}}|) / (\sum_{\text{hkl}} |F_{\text{obs}}|)$ where $|F_{\text{obs}}|$ and $|F_{\text{calc}}|$ are the observed and calculated structure factor amplitude for reflection hkl. The sum is carried out over the 95% of the observed reflections which are used in refinement.

³ R_{free} refers to the R factor for the test reflection set (5% of the total observed) which was excluded from refinement

Table 2

Selected Structural Measurements

Structure	Distance between residue 138 C α of each subunit(Å)	Distance between residue 127 C β of each subunit(Å)	Distance between base atoms in CSPS (Å)	Solvent Accessible Surface Area buried in CSPS (Å ²) (relative SASA) ^I
wtHincII/CG/Ca ²⁺ (1TW8) ¹⁷	26.7	26.6	3.4	141 (100%)
wtHincII/TA (2GIE) ¹⁵	27.0	27.1	3.4	116 (100%)
Q138F/TA/Ca ²⁺ (cocrvstal)	28.4	26.8	4.0	99 (85%)
Q138F/GC/Ca ²⁺	28.5	26.9	5.3	29 (21–25%)
Q138F/CG/Ca ²⁺ (soak) (2GIH) ¹⁵	28.0	28.0	3.7	112 (79%)
Q138F/TA/Ca ²⁺ (soak) (2GIJ) ¹⁵	28.4	28.1	3.9	92 (79%)
Q138F/CG/Ca ²⁺ (cocrvstal)	27.8	27.2	3.6	128 (91%)
Q138F/TA/Mg ²⁺	28.0	28.7	3.9	84 (72%)
Q138F/TA/Mn ²⁺	28.1	28.2	4.5	75 (69%)

^I The relative buried surface area is the ratio with that of the wild type HincII structure bound to the same DNA sequence.