### Instructions for use

**Title**

Structures of restriction endonuclease HindIII in complex with its cognate DNA and divalent cations

**Author(s)**

Watanabe, Nobuhisa; Takasaki, Yozo; Sato, Chika; Ando, Shoji; Tanaka, Isao

**Citation**

Acta Crystallographica. Section D, Biological Crystallography, 65(12): 1326-1333

**Issue Date**

2009-12

**Doc URL**

http://hdl.handle.net/2115/39986

**Type**

article

**File Information**

watanabe_ActaCristD65.pdf

---

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP
Structures of restriction endonuclease *Hind*III in complex with its cognate DNA and divalent cations

Nobuhisa Watanabe, Yozo Takasaki, Chika Sato, Shoji Ando and Isao Tanaka


Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see [http://journals.iucr.org/services/authorrights.html](http://journals.iucr.org/services/authorrights.html)

*Acta Crystallographica Section D: Biological Crystallography* welcomes the submission of papers covering any aspect of structural biology, with a particular emphasis on the structures of biological macromolecules and the methods used to determine them. Reports on new protein structures are particularly encouraged, as are structure–function papers that could include crystallographic binding studies, or structural analysis of mutants or other modified forms of a known protein structure. The key criterion is that such papers should present new insights into biology, chemistry or structure. Papers on crystallographic methods should be oriented towards biological crystallography, and may include new approaches to any aspect of structure determination or analysis.

Crystallography Journals Online is available from [journals.iucr.org](http://journals.iucr.org)
Structures of restriction endonuclease *HindIII* in complex with its cognate DNA and divalent cations

The three-dimensional crystal structures of *HindIII* bound to its cognate DNA with and without divalent cations were solved at 2.17 and 2.00 Å resolution, respectively. *HindIII* forms a dimer. The structures showed that *HindIII* belongs to the *EcoRI*-like (α-class) subfamily of type II restriction endonucleases. The cognate DNA-complex structures revealed the specific DNA-recognition mechanism of *HindIII* by which it recognizes the palindromic sequence A/AGCTT. In the Mg$^{2+}$ ion-soaked structure the DNA was cleaved and two ions were bound at each active site, corresponding to the two-metal-ion mechanism.

1. Introduction

Restriction endonucleases are produced in bacteria, archaea and Chlorella viruses. These enzymes hydrolyze foreign DNA and thus allow microorganisms to avoid attack by phages etc. Based on subunit composition, cofactor requirements, DNA-recognition specificity and mode of action, restriction endonucleases have been classified into four types (Roberts *et al.*, 2003). Type II restriction endonucleases recognize and digest specific palindromic sequences of double-stranded DNA. About 4000 type II restriction endonucleases have been discovered to date. The specificity of the type II restriction endonucleases has made them indispensable tools in biotechnology. However, the tertiary structures of only about 30 type II restriction endonucleases are available at present in REBASE (Roberts *et al.*, 2007). Most of the type II restriction endonucleases have a common sequence motif PD-(D/E)xK (Stahl *et al.*, 1998). Pingoud *et al.* (2005) showed that they contain a core comprised of a five-stranded mixed β-sheet flanked by α-helices and that the second and third β-strands form a scaffold for the catalytic residues of PD-(D/E)xK. Typically, type II restriction enzymes utilize Mg$^{2+}$ ion as a divalent cation. Recent reviews have focused on their structure and function, the role of metal ions and their catalytic mechanisms (Pingoud & Jeltsch, 2001; Pingoud *et al.*, 2005).

The type II restriction endonuclease *HindIII* produced in *Haemophilus influenzae* Rd digests DNA at the palindromic sequence A/AGCTT. *HindIII* is comprised of 300 amino acids and has a molecular weight of 34 950 Da. The amino-acid sequences of this nuclease and the corresponding methyltransferase were deduced from their genes (Nwankwo *et al.*, 1994). We have previously reported protein-engineering studies on *HindIII* (Tang *et al.*, 2000). One mutant, D108L, showed a complete loss of enzyme activity, suggesting that...


**Table 1**

Crystallographic data-collection and refinement statistics.

<table>
<thead>
<tr>
<th></th>
<th>Iodine derivative</th>
<th>Native</th>
<th>Mg$^{2+}$-soaked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamline</td>
<td>BL-5A</td>
<td>NW-12A</td>
<td>BL-6A</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.5</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Space group</td>
<td>$P_2_1$</td>
<td>$P_2_1$</td>
<td>$P_2_1$</td>
</tr>
<tr>
<td>Unit-cell parameters (Å)</td>
<td>$a = 83.9$, $b = 130.9$, $c = 93.6$, $\beta = 110.2$</td>
<td>$a = 83.1$, $b = 131.5$, $c = 93.7$, $\beta = 111.2$</td>
<td>$a = 83.5$, $b = 132.2$, $c = 94.1$, $\beta = 111.0$</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>2.77 (2.87–2.77)</td>
<td>2.00 (2.07–2.00)</td>
<td>2.17 (2.25–2.17)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>48465</td>
<td>130036</td>
<td>100664</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.6</td>
<td>7.6</td>
<td>1.9</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.3 (67.4)</td>
<td>99.9 (99.5)</td>
<td>99.8 (98.0)</td>
</tr>
<tr>
<td>$R_{merge}$ (%)</td>
<td>9.2 (37.4)</td>
<td>5.6 (14.3)</td>
<td>5.4 (40.4)</td>
</tr>
<tr>
<td>$I/\sigma(I)$</td>
<td>7.6 (1.6)</td>
<td>23.2 (12.5)</td>
<td>24.7 (3.5)</td>
</tr>
<tr>
<td>Mosaicity (°)</td>
<td>1.00</td>
<td>0.25</td>
<td>0.49</td>
</tr>
<tr>
<td>Refinement resolution (Å)</td>
<td>2.00</td>
<td>2.00</td>
<td>2.17</td>
</tr>
<tr>
<td>$R_{merge}/R_{free}$ (5% of reflections)</td>
<td>17.7/21.7</td>
<td>17.5/22.1</td>
<td></td>
</tr>
<tr>
<td>No. of waters</td>
<td>968</td>
<td>998</td>
<td>598</td>
</tr>
<tr>
<td>R.m.s.d. bond lengths (Å)</td>
<td>0.015</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>R.m.s.d. bond angles (°)</td>
<td>1.50</td>
<td>1.84</td>
<td>1.84</td>
</tr>
<tr>
<td>Average B factor (Å$^2$)</td>
<td>23.8</td>
<td>36.1</td>
<td>36.1</td>
</tr>
</tbody>
</table>

**Figure 1**

Overall structure of the *Hin* III monomer. (a) *Hin* III monomer with labelled helices and $\beta$-strands. (b) Topology diagram of *Hin* III. $\beta$-Sheets are shown in red and helices are shown in cyan.

Asp108 is located in the active site of the enzyme. We also found that the mutant E86K showed higher activity than the wild-type enzyme (Tang et al., 1999). However, the tertiary structure of *Hin* III has not been determined. To facilitate further studies of substrate recognition and enzyme catalysis, we have determined the *Hin* III crystal structure in its cognate DNA complex with and without metal ions.

### 2. Materials and method

#### 2.1. Protein purification and crystallization

Wild-type *Hin* III was expressed in a His-tagged form (with ten His residues at the N-terminus). The *Hin* III protein was purified by a modification of the method reported previously (Tang et al., 2000). The *Hin* III-RM gene inserted into the pET16b vector was expressed in *Escherichia coli* BL21 (DE3) pLysS. The bacteria were grown in 3 l LB medium containing 50 μg ml$^{-1}$ ampicillin, harvested and sonicated to obtain crude extract. The extract in binding buffer (5 mM imidazole, 0.5 M NaCl and 20 mM Tris–HCl pH 8.0) was applied onto a column packed with His-Bind Resin (Novagen, San Diego, California, USA) in 50 mM NiSO$_4$. The column was washed with binding buffer and then with wash buffer (60 mM imidazole, 0.5 M NaCl and 20 mM Tris–HCl pH 8.0). The protein was eluted by adding 0.05 M EDTA, 0.1 M EDTA + 0.1 M imidazole, 0.1 M EDTA + 0.2 M imidazole, 0.1 M EDTA + 0.5 M imidazole and then 1.0 M imidazole to the column in a stepwise manner. *Hin* III protein eluted in the 0.1 M and 0.2 M imidazole fractions. The purified protein thus obtained was dialyzed against a buffer containing 20 mM Tris–HCl pH 8.0 and 0.2 M NaCl overnight. Protease Xa was added to the sample to remove the His tag by treatment at 285 K for 14 h. The sample was then added to the His-Bind resin column and the flowthrough fraction was collected and concentrated to more than 0.1% and therefore no trace was observed in the concentrated sample.

DNA oligonucleotides (5’-GCAAGCTTGGC-3’) were synthesized and purified by HPLC (Hokkaido System Science, Sapporo, Japan). The oligonucleotides for use in crystallization trials were dissolved in Milli-Q reverse-osmose purified water. The oligonucleotide solution was heated to 368 K for 10 min, followed by slow cooling to room temperature. Protein–DNA complexes were prepared by mixing annealed oligonucleotides with *Hin* III and incubating at 277 K overnight. The final protein–DNA complex solution contained *Hin* III protein at a concentration of 6.0 mg ml$^{-1}$; the DNA duplex was present in a 1.5-fold molar excess relative to the protein dimer.

Crystals of the *Hin* III–DNA complex were grown by the hanging-drop vapour-diffusion method over a reservoir of 12–16% (v/v) PEG 3350, 80–120 mM ammonium acetate. Drops were formed by mixing 1 μl *Hin* III–DNA complex solution...
and 1 μl reservoir solution and were equilibrated over the reservoir at 293 K.

To elucidate the metal-binding structure of the HindIII–DNA complex, Mg\(^{2+}\) was introduced into the crystals of the HindIII–DNA complex using microdialysis buttons. The buttons were left in dialysis solution containing 20 mM MgCl\(_2\), 20% glycerol and reservoir solution for one week.

### 2.2. Data collection

The initial phases were determined at 2.8 Å resolution by the SAD method using an iodinated derivative. Tyrosine residues in the HindIII–DNA complex crystals were iodinated using the vaporizing-iodine labelling method (Miyatake et al., 2006). The iodine SAD data were collected at 1.5 Å wavelength and all other data sets were collected at 1.0 Å wavelength using flash-frozen crystals at 95 K. Native and iodinated crystals were cryoprotected with 20%\((v/v)\) glycerol prior to cooling. All data sets were collected at the Photon Factory, Tsukuba, Japan. Native, iodine-derivative and Mg\(^{2+}\) data sets were collected on beamlines NW12A, BL5A and BL6A, respectively. Diffraction data were processed using HKL-2000 (Otwinowski & Minor, 1997).

### 2.3. Structure solution and refinement

Iodine positions were determined using SHELXC and SHELXD (Sheldrick, 2008) using the HKL2MAP interface (Pape & Schneider, 2004); the iodine sites were refined and the initial SAD phases were calculated using the program SOLVE (Terwilliger & Berendzen, 1999). Further phase improvement and auto model building were carried out using RESOLVE (Terwilliger & Berendzen, 1999). The structure of the HindIII–DNA complex was then refined with REFMACS (Murshudov et al., 1997) using 2.0 Å resolution native data and manually fitted with Coot.
Figure 5
Stereo figure of the flexible-loop region of monomer D of the metal-soaked structure. Grey helices are those of monomer C. The two metals Mg\(^{2+}\) and Mn\(^{2+}\) are shown as green and orange spheres, respectively. The side chain of Glu86 is shown as a stick model.

Figure 6
HindIII–DNA interactions. (a) Stereoview of the HindIII–DNA interaction. Only one monomer of HindIII is shown. Side chains with direct interactions to the bases are shown as stick models. (b) View of the base-pair recognition of HindIII, with hydrogen bonds indicated as dashed lines. The DNA is shown in yellow and magenta to indicate the two parts of the post-reaction product. (c) NUCPLOT sketch illustrating contacts between HindIII and the DNA bases. Contacts between HindIII AB dimer and the bases of the DNA strands E and G are shown. Asterisks indicate residues that appear more than once.
3. Results and discussion

3.1. Overall structure

The HindIII monomer (35 kDa) consists of 16 helices and two β-sheets: one small sheet consisting of two short strands and a larger sheet comprised of five β-strands (Fig. 1). Helices 7, 11 and 16 are 3_{10}-helices. Helix 10 is a proline-kinked helix with Pro195 located near its centre. A β-sheet flanked by α-helices is a core structure of restriction endonucleases (Venclavas et al., 1994). HindIII also has this core structure (α4, βb, βc, βd, α6, βl, α8, βg, α9). The three key residues of the canonical PD-(D/E)xK catalytic motif, Asp108, Ala109 and Lys110, are located in this larger β-sheet. As βg is parallel to βb, HindIII belongs to the EcoRI-like (α-class) subfamily (Venclavas et al., 1994; Pingoud & Jeltsch, 1997; Kovall & Matthews, 1999). Based on the classification of Niv and coworkers, HindIII belongs to class I, H1-S-H2-up, and H2 (α9 in HindIII) is almost perpendicular to the β-strands (Niv et al., 2007).

The HindIII homodimer measures approximately 70 × 80 × 50 Å (Fig. 2). The dimer formation of HindIII is interesting. The long arm consisting of the C-terminal four helices (α13, α14, α15 and α16) protrudes from the core and interacts with its counterpart in the dimer. In particular, the long helix α13 and three helices (α14, α15 and α16) of the adjacent monomer mainly make contact via hydrophobic interactions and form a hinge-like structure. Flexibility of the hinge can be observed when the structures of the four monomers in the asymmetric unit are superposed. The loop between α13 and α14 shows a displacement of more than 6.5 Å (Fig. 3).

In the crystal, there are four protein monomers and three DNA duplexes in an asymmetric unit. The DNA strands are uncleaved in the cocrystal structure without metals. Protein subunits A and B or C and D form dimers which encircle double-stranded DNA molecules. In the asymmetric unit, the HindIII dimers form a tetramer with 222 symmetry (Fig. 4). Apart from the tetramer of the asymmetric unit, PISA (Krissinel & Henrick, 2007) indicated that only the dimers are potentially biologically relevant, which is also consistent with the estimated molecular weight of the DNA complex on gel-filtration chromatography (data not shown). Interestingly, an extra DNA duplex, the occupancy of which was fixed to 0.7 in the final refinement cycles, exists like a ‘shore’ between the two parts of the post-reaction product. (b) Superposition of the DNA structures before and after the scission reaction. DNA before the reaction is shown in green and the cleaved DNA is shown in magenta.

Figure 7
DNA structure near the scissile bond. (a) View of the HindIII–DNA interactions, with hydrogen bonds indicated as dotted lines. Residues of HindIII are coloured green and the DNA is shown in yellow and magenta to indicate the two parts of the post-reaction product. (b) Superposition of the DNA structures before and after the scission reaction. DNA before the reaction is shown in green and the cleaved DNA is shown in magenta.
Fig. 6(c). The four residues that act at the major groove, Asp120, Lys122, Asp123 and Lys125, are on the short helix α5. Of the residues that act at the minor groove, Glu60 and Lys61 are located at the N-terminal end of helix α4, Ser56 is located in a loop between helices α3 and α4 and Arg88 is in the loop between the βb and βc strands. In the major groove, six direct hydrogen bonds and several water-mediated hydrogen bonds are formed between the protein and the bases of the specific sequence A/AGCTT of the DNA. Asn120 Oα forms a hydrogen bond to N6 of adenine 4A. Asn120 Nα also makes contact with O4 of thymine 8T. Lys122 Nε makes direct contact with O6 of guanine 6G and Asp123 Oδ makes direct contact with N6 of guanine 6G and Asp123 Oδ makes direct contact with N4 of cytosine 7C. Lys125 Nε forms a water-mediated hydrogen bond to N7 of adenine 5A. In the minor groove, two direct hydrogen bonds and three water-mediated hydrogen bonds are formed to the specific bases. Lys61 Nε makes direct contact with N2 of thymine 9T. Ser56 Oγ makes contact with O2 of thymine 8T. Contacts between Arg88 Nδ2 and Nε and N3 of adenine 4A, between Lys61 Nε and Glu60 Oδ and N3 of adenine 5A and between Ser56 Oδ and N3 of guanine 6G are water-mediated. Direct interactions with the bases by HindIII are only formed to part of the specific sequence, with most occurring in the major groove. No direct contacts were identified using *NUCPLOT* (Luscombe et al., 1997) to the bases of the nonspecific sequence at both ends of recognition sequence, although there are many interactions between the sugar-phosphate backbone and the protein in this region, as observed for other endonucleases (Martin et al., 1999).

The conformation of the bound DNA is distorted from its ideal B-DNA structure. The overall amount of DNA distortion was about 57° and that of shortening was about 8%, while the DNA distortion and shortening of the extra free DNA found as a ‘shore’ between the artificial tetramers were 12° and 0.6%, respectively, as determined using the program *CURVES* (Lavery & Sklenar, 1988). HindIII–DNA interactions and superposition of DNA structures around the scissile bond before and after the reaction are shown in Fig. 7. The overall DNA structures are almost the same, but the position of the ribose of adenosine 4A is significantly shifted when the phosphate backbone is broken, even though the purine position does not change. As shown in the figure, interactions between Asn120 and adenine 4A cause the purine to tilt about 25° from the pyrimidine of its counterpart thymine 9T, resulting in an accumulation of strain in the DNA structure. In addition, Thr117, which forms hydrogen bonds to the backbone phosphate of thymine 8T, is the only residue in the disallowed region of the Ramachandran plot.
Diagrammatic representations of the metal ion-coordination geometries. Mn$^{2+}$ A site (a) and Mg$^{2+}$ B site (b) are shown with the distances of each hydrogen bond.

Table 2
Refined temperature factors of the metal ions.

Refined B factors of the A site are shown using Mn$^{2+}$ and Mg$^{2+}$ as the metal ions. Those of the B-site Mg$^{2+}$ ions are also shown for comparison.

<table>
<thead>
<tr>
<th>A site</th>
<th>Using Mn$^{2+}$</th>
<th>Using Mg$^{2+}$</th>
<th>B site</th>
<th>Refined B factor (Å$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>27.05</td>
<td>7.34</td>
<td>S</td>
<td>30.86</td>
</tr>
<tr>
<td>P</td>
<td>32.61</td>
<td>12.28</td>
<td>T</td>
<td>34.80</td>
</tr>
<tr>
<td>Q</td>
<td>31.64</td>
<td>11.49</td>
<td>U</td>
<td>37.11</td>
</tr>
<tr>
<td>R</td>
<td>33.98</td>
<td>13.34</td>
<td>V</td>
<td>31.85</td>
</tr>
</tbody>
</table>

3.3. Active site

In the Mg$^{2+}$-soaked crystal two metal ions were found in the active site of each chain and the DNA was cleaved (Fig. 8). The presence of the two metal ions in the active centre of the Mg$^{2+}$-soaked structure near the cleaved phosphate allowed us to unambiguously determine the catalytic mechanism of HindIII as the two-metal-ion mechanism (Beese et al., 1993; Viadiu & Aggarwal, 1998). The distance between the two metal ions is 3.98 Å and is consistent with the transition-state model of the two-metal-ion mechanism. Interestingly, the two metals seem to be different although the crystal was soaked with Mg$^{2+}$. The metal ions located at the B site seemed to be Mg$^{2+}$, but those at the A site were likely to be Mn$^{2+}$. As we could not reproduce this high-resolution DNA-complex crystal, we were not able to confirm this using the anomalous signal of Mn$^{2+}$ with longer wavelength X-rays, but the temperature factors of the refined metal sites suggested that the sites probably contained Mn$^{2+}$ and not Mg$^{2+}$ (Table 2). The temperature factor of the A-site metal tends to be lower than that of B site. For example, in the case of NgoMIV the B factor of the A-site Mg$^{2+}$ is 11.0–15.0 Å$^2$, while that of the B site is 22.0–31.0 Å$^2$ (Deibert et al., 2000). For the HindIII structure, however, the temperature factors of the A-site metal ions are not consistent with those of the neighbouring atoms when using Mg$^{2+}$ as the metal ion. As MgCl$_2$ was used in the soaking experiment, the Mn$^{2+}$ ion at the A site is probably the preferred selection by HindIII from the trace elements present. It is not clear why Mn$^{2+}$ is preferred over Mg$^{2+}$, but the difference in reaction kinetics between Mn$^{2+}$ and Mg$^{2+}$ in EcoRV has been discussed based on the intrinsic affinities of these ions for phosphates (Sam & Perona, 1999) and seems to be consistent with the observation that the Mn$^{2+}$ concentration optimum is about ten times lower than that of Mg$^{2+}$ for many PD-(D/E)K-family endonucleases (Pingoud et al., 2009).

The geometry around the two metal ions is shown in Fig. 9. The A-site Mn$^{2+}$ ion is pentacoordinated and the B-site Mg$^{2+}$ ion is hexacoordinated. The coordination architecture of the two ions is as follows. Asp93 O$^\text{C14}$, 5'-O3P and 3'-O3 of cleaved DNA and three water molecules ligate the B-site Mg$^{2+}$ ion. Asp93 O$^\text{D2}$, Asp108 O$^\text{D2}$, the main-chain O atom of Ala109, 5'-O1P of cleaved DNA and one water molecule bind to the Mn$^{2+}$ ion at the A site. The coordination distances are reasonable compared with the distributions in the MESPEUS database (Hsin et al., 2008). Asp93 of HindIII corresponds to Asp94 of BamHI, which is important for two-metal binding, as shown in Fig. 8. Our previous studies showed that another aspartic acid, Asp108, is also important for HindIII activity (Tang et al., 1999, 2000). As can be seen in Figs. 8(a) and 8(b), the conformation of Asp108 is changed to coordinate to the two metal ions and Asp108 interacts directly with Mn$^{2+}$ at the A site and with Mg$^{2+}$ at the B site by water-mediated hydrogen bonding.

In summary, we have presented the three-dimensional crystal structure of HindIII bound to cognate DNA in metal-free and Mg$^{2+}$-soaked forms. The DNA was cleaved in the Mg$^{2+}$-soaked form, indicating that the solved structure is the active form. Further studies are required as the metal ion at the A site is responsible for activating the nucleophile. However, other crystallization conditions are required that do not lead to the free extra DNA 'shore' between the artificial tetramers, as the present crystallization conditions were poorly reproducible.

This work was supported in part by the Takeda Science Foundation. The authors would like to thank Yukari Minami and Kaori Onishi for help with the crystallization of HindIII.

References