The Recognition Domain of the BpuJI Restriction
Endonuclease in Complex with Cognate DNA at 1.3-Å
Resolution

Rasa Sukackaite¹, Saulius Grazulis¹, Matthias Bochtler²,³,⁴ and Virginijus Siksnys¹⁎

¹Institute of Biotechnology, Graiciuno 8, 02241 Vilnius, Lithuania
²International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland
³Max Planck Institute for Molecular Cell Biology and Genetics, Pfortenauerstr. 108, 01309 Dresden, Germany
⁴Schools of Chemistry and Biosciences, Cardiff University, Park Place, Cardiff CF10 3AT, Wales, UK

Received 8 February 2008; received in revised form 17 March 2008; accepted 19 March 2008
Available online 28 March 2008

Keywords: restriction endonuclease; BpuJI; DNA recognition; helix–turn–helix; crystal structure

Introduction

Type II restriction endonucleases recognize 4- to 8-bp DNA sequences and cleave both strands of the DNA at fixed positions.¹ They follow many different modes of action and are categorized into 11 overlapping subtypes.² Type IIS enzymes recognize asymmetric nucleotide sequences and cleave both DNA strands at fixed positions downstream of the recognition site. The archetypal Type IIS enzyme FokI is composed of a specific DNA-binding module fused to a cleavage domain that possesses a conserved restriction endonuclease catalytic core but cuts DNA in a non-specific manner.³ The FokI recognition domain is made of three smaller subdomains (D1, D2 and D3) that contain helix–turn–helix (HTH) motifs.⁴ The first two subdomains are involved in DNA recognition, whereas the third subdomain has been co-opted for protein–protein interactions. Modular architecture is also characteristic for another structurally characterized Type IIS enzyme, BfiI, which is composed of a DNA-binding domain fused to a catalytic module similar to the non-specific nuclease belonging to the phospholipase D family.⁵ The DNA recognition domain of BfiI exhibits a barrel-like structure that shows no similarity to FokI.⁶

Type IIS restriction endonucleases recognize asymmetric DNA sequences and cleave both DNA strands at fixed positions downstream of the recognition site. The restriction endonuclease BpuJI recognizes the asymmetric sequence 5′-CCCGT; however, it cuts at multiple sites in the vicinity of the target sequence. BpuJI consists of two physically separate domains, with catalytic and dimerization functions in the C-terminal domain and DNA recognition functions in the N-terminal domain. Here we report the crystal structure of the BpuJI recognition domain bound to cognate DNA at 1.3-Å resolution. This region folds into two winged-helix subdomains, D1 and D2, interspaced by the DL subdomain. The D1 and D2 subdomains of BpuJI share structural similarity with the similar subdomains of the FokI DNA-binding domain; however, their orientations in protein–DNA complexes are different. Recognition of the 5′-CCCGT target sequence is achieved by BpuJI through the major groove contacts of amino acid residues located on both the helix–turn–helix motifs and the N-terminal arm. The role of these interactions in DNA recognition is also corroborated by mutational analysis.

© 2008 Elsevier Ltd. All rights reserved.

*Corresponding author. E-mail address: siksnys@ibt.lt.
Abbreviations used: HTH, helix–turn–helix; wt, wild type; wHTH, winged helix–turn–helix; PEG, polyethylene glycol.
Restriction endonuclease BpuJI, like most Type IIS enzymes, recognizes the asymmetric sequence 5\'-CCCGT; however, it cuts at multiple sites in the vicinity of the target sequence. Like FokI and other Type IIS enzymes, BpuJI consists of two physically separate domains: an N-terminal domain that binds to the recognition sequence as a monomer and a C-terminal domain that forms a dimer with non-specific nuclease activity. In contrast to FokI, which is a monomer in solution and dimerizes in the presence of target DNA, full-length BpuJI is a dimer in the DNA-free form. The dimerization interface is formed by the C-terminal domains; therefore, wild-type (wt) BpuJI has two surfaces for specific DNA binding provided by the N-terminal domains. As a consequence, BpuJI binds two copies of cognate DNA, making a synaptic complex that promotes cleavage. The isolated catalytic domain of BpuJI is an active stand-alone nuclease that possesses a non-specific end-directed nuclease activity and shows preference for blunt-ended DNA. In full-length BpuJI, the nuclease activity is repressed and becomes activated only upon specific DNA binding, resulting in a complicated pattern of specific DNA cleavage close to the target site. It is likely that interdomain interactions or the linker connecting catalytic and DNA-binding domains may control the BpuJI cleavage activity, similarly to FokI and BfiI. Fold recognition revealed that the BpuJI C-terminal domain is structurally related to the archaeal Holliday junction resolvases; however, it failed to predict the fold of the N-terminal DNA recognition domain.

Here we report the crystal structure of the BpuJI recognition domain (residues 1–285) bound to a cognate 12-bp oligoduplex at 1.3-Å resolution. The structure reveals two winged-helix subdomains and provides a detailed view of 5\'-CCCGT sequence recognition by the restriction endonuclease BpuJI.

Results and Discussion

Overall structure of the protein–DNA complex

Using limited proteolysis and mass spectrometry, we identified a stable N-terminal fragment (residues 1–285) constituting the cognate DNA-binding domain of the BpuJI restriction endonuclease. The BpuJI N-terminal domain–DNA complex obtained by limited proteolysis crystallized in space group P2\_12\_12. The structure of the BpuJI recognition domain in complex with a 12-bp oligoduplex was determined by the single-wavelength anomalous dispersion method using a mercury derivative and refined to 1.3-Å resolution with a free R-factor of 16.6% and a working R-factor of 13.8%. Electron density shows all the DNA bases and the amino acid residues except the five C-terminal residues 281–285. The asymmetric unit contains one protein monomer bound to the DNA duplex, in accordance with the biochemical data.

The BpuJI N-terminal domain (Fig. 1a) contains two winged-helix subdomains. D1 and D2, that are interspaced by the DL subdomain. The D1 subdomain (residues 1–112) contains an N-terminal arm, five helices (H1–H5) and three strands (B1–B3) (Fig. 1b). Helices H2 and H3 form an HTH structure, which is followed by a signature wing region (B2 and B3) and helices H4 and H5. The loop between H1 and H2 adopts an extended configuration and is incorporated as a third strand (B1) in the sheet. The D2 subdomain (residues 137–245) is composed of six helices (H7–H12) and two strands (B6 and B7) (Fig. 1b). D2 also bears an HTH structure, formed by helices H8 and H11. The HTH motif is followed by a \(\beta\)-hairpin wing (B6 and B7) and helix H12. In contrast to canonical transcription factors, an expected “turn” in the HTH motif of D2 is replaced by a 35-aa insertion that contains helices H9 and H10. A 24-aa linker region with the chain topology...
B4–H6–B5 connects subdomains D1 and D2 (Fig. 1b). Helix H13 forms a C-terminal extension that stands apart from the D2 subdomain and is close to the linker region. The solvent-inaccessible surface area of 1663 Å² buried at the interface between the linker and helix H13 is larger in comparison with the surfaces buried between the linker–H13 fragment and the D1 and D2 subdomains (1192 and 1140 Å², respectively). Thus, the internal contacts between the linker and the H13 residues are more extensive than the external contacts to the D1 and D2 subdomains. Based on this criterion, the linker and the C-terminal extension, including helix H13, are classified as an interrupted subdomain DL.

The DNA in complex with the BpuJI recognition domain maintains a canonical B-DNA conformation with no major bend or kink (Fig. 1a). The recognition domain maintains a canonical B-DNA conformation and the C-terminal extension, including helix H13, are classified as an interrupted subdomain DL.

The DNA in complex with the BpuJI recognition domain harbors a winged-helix motif, some of them, such as reverse gyrase11 and Werner syndrome protein,12 have a nuclease function. However, in all cases, the region of similarity is limited to one winged HTH (wHTH) subdomain of BpuJI, and no overall match is observed.

With the use of the structure of the D1 subdomain as a DALI search target, the closest similarity was encountered with the restriction endonuclease FokI (Table 1), which was not found using the entire structure. According to a structural alignment of the TOP3D program,13 the D1 subdomains of FokI and BpuJI superimpose with an r.m.s. distance of 2.1 Å over 63 Cα atoms. The region of topological similarity spans over the winged-helix motif and the two following helices (Fig. 2a). The winged-helix motif of the BpuJI D1 subdomain is more compact than that of FokI and resembles the domains of the MarR family.10 The winged-helix motifs of the BpuJI and FokI D2 subdomains also superimpose with an r.m.s. distance of 1.6 Å over 45 Cα atoms (Fig. 2a). They are both extensively modified, with large insertions upstream of the recognition helices, although the insertions show no topological similarity. However, the similarity of the BpuJI D2 subdomain to the FokI D2 subdomain was not revealed by a DALI search using the C-terminal part of the protein as a search target (Table 1).

Comparison with other DNA-binding proteins

A DALI search against the Protein Data Bank (PDB) database, using the entire structure of the BpuJI recognition domain as a target, revealed significant similarity to other DNA-binding proteins (Table 1). Most of the proteins showing similarity to the BpuJI recognition domain harbor a winged-helix motif, some of them, such as reverse gyrase11 and Werner syndrome protein,12 have a nuclease function. However, in all cases, the region of similarity is limited to one winged HTH (wHTH) subdomain of BpuJI, and no overall match is observed.

Table 1. Proteins showing similarity to the BpuJI recognition domain

<table>
<thead>
<tr>
<th>Search target</th>
<th>PDB accession code</th>
<th>Protein</th>
<th>Z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>BpuJI-Nt</td>
<td>198</td>
<td>Chromosome partition protein</td>
<td>5.3</td>
</tr>
<tr>
<td>(1–280 aa)</td>
<td></td>
<td>MukF fragment</td>
<td></td>
</tr>
<tr>
<td>lgs</td>
<td>4.8</td>
<td>Multiple antibiotic resistance protein</td>
<td>5.0</td>
</tr>
<tr>
<td>1gku</td>
<td>4.8</td>
<td>Reverse gyrase</td>
<td></td>
</tr>
<tr>
<td>1ly7</td>
<td>4.8</td>
<td>Esal histone acetyl-transferase fragment</td>
<td></td>
</tr>
<tr>
<td>1axl</td>
<td>4.7</td>
<td>Werner syndrome protein</td>
<td></td>
</tr>
<tr>
<td>BpuJI-D1</td>
<td>198</td>
<td>FokI restriction endonuclease</td>
<td>5.5</td>
</tr>
<tr>
<td>(20–131 aa)</td>
<td></td>
<td>MukF fragment</td>
<td>5.4</td>
</tr>
<tr>
<td>1yg2</td>
<td>5.2</td>
<td>Gene activator ApmA</td>
<td></td>
</tr>
<tr>
<td>1ec1</td>
<td>5.0</td>
<td>Escherichia coli DNA topoisomerase</td>
<td></td>
</tr>
<tr>
<td>lgs</td>
<td>4.8</td>
<td>Multiple antibiotic resistance protein</td>
<td></td>
</tr>
<tr>
<td>BpuJI-D2-DL</td>
<td>1gn</td>
<td>Aminoacyl-tRNA synthase</td>
<td>3.6</td>
</tr>
<tr>
<td>(91–280 aa)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1kyq</td>
<td>3.5</td>
<td>Oxidoreductase, lyase</td>
<td></td>
</tr>
<tr>
<td>2ga1</td>
<td>3.2</td>
<td>crpFR family transcription regulatory protein</td>
<td></td>
</tr>
<tr>
<td>1gku</td>
<td>3.2</td>
<td>Reverse gyrase</td>
<td></td>
</tr>
<tr>
<td>1w5s</td>
<td>3.0</td>
<td>DNA replication initiation protein</td>
<td></td>
</tr>
</tbody>
</table>

The similarities were revealed by a DALI database search using the BpuJI recognition domain (residues 1–280), the D1 subdomain (residues 20–131) or the C-terminal part of the BpuJI recognition domain (91–280 aa) as a search target. Only five proteins with the highest Z scores are shown.

The DNA-binding domain in BpuJI contains two wHTH units, while the FokI recognition domain is composed of three wHTH motifs (Fig. 2b). The BpuJI subdomains show similarity to the respective FokI subdomains D1 and D2 that interact with DNA. The third wHTH subdomain of FokI, which is missing in BpuJI, mediates protein–protein rather than protein–DNA interactions.4 Thus, both BpuJI and FokI employ two wHTH subdomains to interact with their target sequences.

Although the BpuJI and FokI subdomains D1 and D2 share structural similarity, their orientations in protein–DNA complexes differ significantly (Fig. 2b). In BpuJI, both the recognition helices of the HTH motifs pack against the major groove with their helical axes perpendicular to the long DNA axis and almost parallel with the DNA base pair edges, while in the FokI–DNA complex, only the D1 recognition helix is perpendicular to the DNA axis (Fig. 2b). The recognition helix of the FokI D2 subdomain juts away from the DNA, and its helical axis is tilted by ~35° with respect to the plane of the base pairs.4

Duplications of HTH units are found in a large number of protein families, such as sigma factors,14 POU family domains,15 the paired domains of the Pax and Prd transcription factors and the recombinase DNA-binding domains.16 The HTH units of bipartite domains usually bind to different regions of an extended binding site. In certain cases, the unstructured linker region fills in the minor groove.
between domains, as shown in Fig. 2b for the DNA-binding domain of the Tc3 transposase Tc3A bound to 26-bp DNA. The wHTH units of the bipartite recognition domains of restriction endonucleases BpuJl and FokI bind to the 5-bp target sequences. The wHTH motif was also identified in the restriction endonuclease SdaI, which recognizes the continuous palindromic 8-bp sequence. However, the wHTH unit of SdaI forms a single globular domain and has to dimerize to achieve the recognition of the 8-bp palindromic target site.

**DNA recognition by the BpuJl N-terminal domain**

Interactions between the BpuJl recognition domain and the target sequence occur along the two HTH motifs and the N-terminal arm, as in the FokI–
DNA complex, although the BpuJI target sequence 5′-CCCGT shows no similarity to the FokI recognition sequence 5′-GGATG. The BpuJI N-terminal domain recognizes the entire target sequence by direct readout in the major groove. The interactions between the protein and the target bases are mediated by residues located solely in the recognition helices of the HTH motifs (H3 and H11) and the N-terminal arm (Fig. 3).

The recognition helix of the BpuJI D1 subdomain (H3) recognizes the two CG base pairs at the 5′-end of the recognition sequence (CCCGT). The contacts with the DNA bases are exclusively mediated by Lys63, Asn67 and Glu71 residues (Fig. 3): Lys63 donates a hydrogen bond to O6 of the guanine G5 (bases are numbered as in Fig. 3); Asn67 forms hydrogen bonds with N4 of the cytosine C1 and O6 of the guanine G4; and Glu71 accepts a hydrogen bond from N4 of the cytosine C2. The inner CG base pairs (CCCGT) are recognized by Arg15 and Lys17 residues, located in the N-terminal arm (Fig. 3). Lys17 donates a hydrogen bond to O6 of the guanine G3 and interacts via water with N7 of the same guanine, while Arg15 donates a bidentate hydrogen bond to the guanine G4. The D2 subdomain recognizes only one TA base pair at the 3′-end of the recognition sequence (CCCGT), and contacts with DNA bases are mediated by the N-terminal portion of the recognition helix H11: Gln208 forms a bidentate hydrogen bond with the adenine and Ser204 donates a hydrogen bond to O4 of the thymine (Fig. 3).

The backbone phosphate groups form the extensive set of hydrogen bonds and salt bridges with the residues located in the recognition helices (H3 and H11) and in the loops of the HTH motifs: there are four direct and three water-mediated contacts with the HTH motif of the D1 subdomain, and there are four direct and four water-mediated contacts with the HTH motif of the D2 subdomain (Fig. 4). The positive helix dipoles at the N-terminus of the recognition helices H3 and H11 could be in favor of interaction with negatively charged phosphate groups of the DNA backbone. There are also two salt bridges, nine water-mediated hydrogen bonds and two van der Waals contacts between the backbone phosphates and residues located in the N-terminal arm (Fig. 4).

The DL subdomain interacts with the DNA upstream of the target sequence. The residues of the linker segment interact with DNA in the minor groove, making one direct and two water-mediated contacts with the backbone phosphate groups (Fig. 4). Lys121 forms a hydrogen bond and a water-mediated contact to the DNA bases upstream of the recognition site. Lys266 located in the C-terminal helix H13 forms a water-mediated contact to the phosphate located 3 bp upstream of the target site (Fig. 4), while Glu263, Lys266 and Lys267 form salt bridges with the phosphates of a symmetry-related DNA molecule (not shown).

**Mutational analysis of DNA-binding interface**

The protein–DNA contacts observed in the crystal structure were tested by site-directed mutagenesis. The residues Lys63, Asn67, Glu71, Arg15, Lys17, Ser204 and Gln208, which form hydrogen bonds with the target bases in the major groove, as well as Lys121, interacting with the flanking DNA bases in the minor groove, were separately mutated to alanines to evaluate these residues for their individual contributions to DNA binding. Both the wt and mutant proteins were expressed and purified as the N-terminal domain (residues 1–285) fusions with a C-terminal hexahistidine tag, and their DNA-
binding properties were studied by gel shift analysis. The 33P-labelled 16/16(specific) and 16/16 oligoduplexes (Table 2) were used in DNA-binding experiments as cognate DNA and non-cognate DNA, respectively.

Gel shift analysis demonstrates that the BpuJI recognition domain binds to the cognate DNA with high affinity ($K_d = 7.3\pm0.3$ nM), while non-cognate DNA binding remains very low (Fig. 5). The ability to bind cognate DNA is almost completely abolished by single N67A, E71A, R15A or Q208A mutation and highly compromised by K63A ($K_d = 48\pm7.4$ nM) or K17A ($K_d = 67\pm14$ nM) mutation. On the other hand, S204A mutation has a minor effect on DNA binding ($K_d = 9.6\pm1.0$ nM). Taken together, mutational analysis data are consistent with the major groove interactions observed in the crystal structure. Surprisingly, alanine replacement of Lys121, which interacts with the TA dinucleotide upstream of the BpuJI target site 5′-CCCGT (Fig. 5), increased the binding affinity of the BpuJI N-terminal domain to cognate DNA nearly 10

Fig. 4. Schematic diagram summarizing DNA contacts by the BpuJI recognition domain. Hydrogen bonds and salt bridges (<3 Å) are indicated by blue lines, and van der Waals contacts (<3.35 Å) are indicated by red lines. Non-bridging water molecules are not shown. P, phosphate group; w, water molecule; *, residue or water on plot more than once.
times ($K_d = 0.83 \pm 0.03 \text{nM}$). However, it still should be tested if the Lys121 mutant shows the same binding phenotype in the context of the full-length BpuJI.

**Conclusions**

The restriction endonuclease BpuJI cuts DNA at multiple sites in the vicinity of the asymmetric target sequence. It is composed of two separate domains that perform different functions. The C-terminal catalytic domain cuts DNA non-specifically and is structurally related to the archaeal Holliday junction resolvases, while the N-terminal domain binds to the 5'CCCGT target site.7 Protein sequence analysis and fold recognition failed to predict the structure of the N-terminal domain of BpuJI.

In this study, we determined the crystal structure of the BpuJI recognition domain bound to cognate DNA. The structure reveals two winged-helix subdomains, D1 and D2, interspaced by the DL subdomain. The BpuJI D1 and D2 subdomains share structural similarity with the DNA-binding subdomains of the archetypal Type IIS restriction endonuclease FokI.

Both FokI and BpuJI contact the respective target sequences 5'-GGATG and 5'-CCCGT by residues located in the two wHTH units and the N-terminal arm. However, the orientations of the wHTH subdomains in FokI and BpuJI bound to cognate DNA are different (Fig. 2b). Thus, two topologically similar wHTH motifs and an extended N-terminal arm are adapted to recognize different asymmetric pentanucleotide sequences in FokI and BpuJI. The presence of the bipartite wHTH motifs in the two Type IIS enzymes recognizing different sequences demonstrates the versatility of HTH units in protein–DNA interactions. Indeed, while the HTH units of bipartite domains, such as the recognition domain of Tc3 transposase, usually bind to different regions of an extended binding site (Fig. 2b), both the wHTH units of the restriction endonucleases BpuJI and FokI bind to the pentanucleotide target sequences. In addition, the extended N-terminal arm also plays an important role in the sequence-specific DNA binding by BpuJI and FokI. The use of two HTH motifs and an N-terminal arm allows to form a highly cooperative hydrogen bond network that is a characteristic feature of the specific protein–DNA complexes of Type II restriction endonucleases.1 One may expect that similar structural motifs could be present in other Type IIS enzymes that exhibit modular structures and cleave DNA outside of their target sites.

**Materials and Methods**

**Preparation of the N-terminal domain–DNA complex**

The restriction endonuclease BpuJI was purified as described previously.7 The BpuJI–DNA complex was prepared in buffer A (10 mM Tris–HCl, pH 7.5 at 25 °C, 100 mM KCl and 2 mM CaCl2) by diluting BpuJI and a cognate oligoduplex (Table 2) to a final protein concentration of 0.3 mg/ml and 1.1-fold molar ratio of DNA duplex to protein monomer. The complex was digested with thermolysin (Sigma) at a 1:10 mass ratio of the protease to BpuJI for 1 h at 25 °C. The reaction was quenched by adding ethylenediaminetetraacetic acid (EDTA) to a final concentration of 10 mM. The resulting proteolysis mixture was loaded onto a heparin–Sepharose column (GE Healthcare), pre-equilibrated with buffer B (50 mM Tris–HCl, pH 7.5 at 25 °C, 100 mM KCl, and 1 mM CaCl2) diluting BpuJI and a cognate oligoduplex (Table 2) to a final protein concentration of 0.3 mg/ml and 1.1-fold molar ratio of DNA duplex to protein monomer. The complex was digested with thermolysin (Sigma) at a 1:10 mass ratio of the protease to BpuJI for 1 h at 25 °C. The reaction was quenched by adding ethylenediaminetetraacetic acid (EDTA) to a final concentration of 10 mM. The resulting proteolysis mixture was loaded onto a Superdex 200 HR column (GE Healthcare). The flow-through was concentrated by centrifugal filter units with a molecular-weight cutoff of 5000 (Millipore) and loaded onto a Superdex 200 HR column (GE Healthcare). Pooled fractions containing the N-terminal domain–DNA complex were concentrated to 6 mg/ml of protein, changing buffer B into buffer C (10 mM Tris–HCl, pH 7.5 at 25 °C, 0.1 mM EDTA and 0.02% NaN3).

**Table 2. Oligoduplexes used for crystallization and in DNA-binding experiments**

<table>
<thead>
<tr>
<th>Oligoduplexes</th>
<th>5'-GGATCCCGTGGA</th>
<th>5'-CCATGGGCACCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/12(SP)</td>
<td>5'-GGATCCCGTGGA</td>
<td>5'-CCATGGGCACCT</td>
</tr>
<tr>
<td>16/16(SP)</td>
<td>5'-TCGGTACCCCGTGATC</td>
<td>3'-AGCCATGGGCACCTAG</td>
</tr>
<tr>
<td>16/16</td>
<td>5'-GGGATCCCGTGGA</td>
<td>5'-CCATGGGCACCT</td>
</tr>
</tbody>
</table>

The BpuJI recognition sequence is underlined. The oligoduplexes were prepared by annealing HPLC-purified oligonucleotides from Metabion.
Crystallographic and data collection

Crystallography was carried out using the sitting-drop method with screens from Hampton Research. The first crystals of the BpuJ N-terminal domain were obtained with a 16-bp cognate oligoduplex 16/16(Sp) (Table 2); however, all the crystals tested were polycrystalline. The problem was solved using the N-terminal domain complex with a 12-bp DNA duplex 12/12(Sp), which gave orthorhombic crystals diffracting to 1.3 Å. The best crystals were grown in drops consisting of 1 μL of the N-terminal domain complex with the 12/12(Sp) in buffer C and 1 μL of a reservoir solution [0.2 M ammonium tartrate, 20% PEG 4000 and 25% PEG 400] at room temperature. The crystals were soaked in 2 mM HgCl₂ for about 1 week to obtain a Hg derivative. For data collection, the crystals were transferred into a cryo-protecting buffer (0.2 M potassium sodium tartrate, 0.2 M NH₄Cl, 20% PEG 4000 and 25% PEG 400) for 30 min prior to flash cryo-cooling. Complete diffraction data sets for the native N-terminal domain complex and the Hg derivative (Table 3) were collected at the EMBL X12 beamline at the DORIS storage ring (DESY, Hamburg, Germany). Data were processed using the MOSFLM, SCALA and TRUNCATE programs.

Structure determination and analysis

Hg positions were identified on anomalous Patterson map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data). Mercury peaks were brought to the map Harker sections using the HARA program (S.G., unpublished data).

Table 3. Data collection statistics

<table>
<thead>
<tr>
<th>Data set</th>
<th>Native</th>
<th>Hg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature (K)</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1,000</td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2</td>
<td></td>
</tr>
<tr>
<td>Cell constants (Å)</td>
<td>56.2</td>
<td>59.0</td>
</tr>
<tr>
<td>A</td>
<td>167.2</td>
<td>168.5</td>
</tr>
<tr>
<td>B</td>
<td>43.8</td>
<td>43.8</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>83.6–1.30</td>
<td>84.2–1.80</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.2 (86.1)</td>
<td>99.9 (99.9)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>10.8 (8.0)</td>
<td>15.5 (15.3)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>4.7 (2.4)</td>
<td>3.7 (2.6)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.082 (0.31)</td>
<td>0.11 (0.24)</td>
</tr>
</tbody>
</table>

* Values in parentheses refer to data in the highest-resolution shell.

Cloning and mutagenesis

The pAL-BpuJ1 plasmid-containing bpuJIR gene was used as a PCR template to clone the N-terminal domain (residues 1–285) into pET21b (Novagen) between the Ndel and Xhol sites. The megaprimer PCR method was used to introduce a silent mutation into the Ndel site present in the BpuJ1 gene. The resulting construct pET-BpuJ1NH overexpresses the BpuJ1 N-terminal domain with a hexahistidine tag at the C-terminus. Site-directed mutagenesis was carried out by the one-step method; the gene regions were sequenced to confirm that only designed mutations had been introduced. Oligonucleotide primers used in the cloning and mutagenesis experiments are presented in Supplementary Table 1.

Gel mobility shift assay for DNA binding

The cognate 16/16(Sp) and non-cognate 16/16 oligoduplexes (Table 2) were 5'-labelled with [γ-³²P]ATP and T4 polynucleotide kinase (Fermentas). The ³²P-labelled DNA was incubated with the wt or mutant N-terminal domain in a binding buffer [30 mM Mes, pH 6.5, 30 mM histidine, 10% (v/v) glycerol and 0.2 mg/mL of bovine serum albumin] for 10 min at room temperature. Binding mixtures were annealed by mixing together the protein and DNA for 1 min and the temperature was increased to 37 °C for 1 min. The complexes were run on a 5% polyacrylamide gel at 1300 V. A typical binding assay is shown in Figure 1.
where $y$ is the concentration of protein–DNA complex (in terms of nanomolar) at the total protein concentration $x$, $s_0$ is the total DNA concentration in the binding mixture and $K_d$ is the dissociation constant. Data were analyzed using the KyPlot program.37

**PDB accession code**

Coordinates and structure factors are deposited in the PDB under accession code 2VLA.

---

**Acknowledgements**

Experiments at the Vilnius Institute of Biotechnol-
y for discussions and comments. Grazvydas Lukinavicius for mass spectrometry and

**Supplementary Data**

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2008.03.041](https://doi.org/10.1016/j.jmb.2008.03.041)

---

**References**


---

from the fusion of a DNA recognition domain to a non-specific nuclease from the phospholipase D superfamily. *J. Mol. Biol.* **336**, 81–92.


*Note added in the proof:* While the article was being prepared and submitted, the coordinates of the site-specific nicking endonuclease N.BspD6I (PDB:2EWF) in the DNA free form were released. The N-terminal part of N.BspD6l shows significant similarity to the BpuJI N-terminal domain (DALLI Z-score=13.4). Further modelling studies should reveal if BpuJI and N.BspD6l share similar mode of target recognition.