

# Conformational Changes and Cleavage by the Homing Endonuclease I-PpoI: A Critical Role for a Leucine Residue in the Active Site

Eric A. Galburt<sup>1</sup>, Meggen S. Chadsey<sup>2</sup>, Melissa S. Jurica<sup>1</sup>  
Brett S. Chevalier<sup>1</sup>, David Erho<sup>1</sup>, Weiliang Tang<sup>2</sup>  
Raymond J. Monnat Jr<sup>2</sup> and Barry L. Stoddard<sup>1\*</sup>

<sup>1</sup>Fred Hutchinson Cancer  
Research Center and the  
Graduate Programs in  
Molecular and Cell Biology and  
Biomolecular Structure and  
Design, 1100 Fairview Ave. N.  
A3-023, Seattle  
WA 98109, USA

<sup>2</sup>University of Washington  
Department of Pathology  
Box 357705, Seattle  
WA 98195, USA

The homing endonuclease I-PpoI severely bends its DNA target, resulting in significant deformations of the minor and major groove near the scissile phosphate groups. To study the role of conformational changes within the protein catalyst and the DNA substrate, we have determined the structure of the enzyme in the absence of bound DNA, performed gel retardation analyses of DNA binding and bending, and have mutagenized a leucine residue that contacts an adenine nucleotide at the site of cleavage. The structure of the L116A/DNA complex has been determined and the effects of the mutation on affinity and catalysis have been measured. The wild-type protein displays a rigid-body rotation of its individual subunits upon DNA binding. Homing site DNA is not detectably bent in the absence of protein, but is sharply bent in both the wild-type and L116A complexes. These results indicate that binding involves a large distortion of the DNA and a smaller change in protein conformation. Leucine 116 is critical for binding and catalysis: it appears to be important for forming a well-ordered protein-DNA complex at the cleavage site, for maximal deformation of the DNA, and for desolvation of the nucleotide bases that are partially unstacked in the enzyme complex.

© 2000 Academic Press

**Keywords:** intron homing; endonuclease; X-ray crystallography; DNA-binding protein; double-strand break

\*Corresponding author

## Introduction

Many interactions between DNA-binding proteins and their targets involve a conformational change in one or both molecules. For non-enzymatic proteins such as transcription factors or histones, changes in protein and DNA structure facilitate high-affinity binding, alter DNA accessibility, bring bound proteins into close proximity, or allow the assembly of multi-protein complexes at specific DNA target sites. Within these complexes, the perturbation of protein or nucleic acid may range from subtle to extreme. The DNA may remain essentially unperturbed while the protein undergoes a conformational change upon binding. Examples include bHLH transcription factors (Ferre-D'Amare *et al.*, 1993, 1994), arc repressor (Bowie & Sauer, 1989; Raumann *et al.*, 1994), and  $\lambda$

repressor (Beamer & Pabo, 1992; Clarke *et al.*, 1991). For other systems, protein binding results in significant distortions of the DNA. Proteins that bind primarily within the DNA major groove can induce DNA bend angles that range from 40° ( $\lambda$  Cro) (Brennan *et al.*, 1990) to 90° (CAP) (Schultz *et al.*, 1991). TATA box-binding protein (TBP) binds to the minor groove and induces an 80°–100° bend in its substrate by inserting phenylalanine residues between nucleotide bases (J.L. Kim *et al.*, 1993; Y. Kim *et al.*, 1993). Integration host factor disrupts base-stacking *via* intercalating proline residues and generates a 160° DNA bend (Rice *et al.*, 1996). Chromatin DNA is wrapped around the nucleosomal core particle in a left-handed superhelix, with half of the binding interactions formed between protein main-chain and DNA backbone atoms (Luger *et al.*, 1997).

Many enzymes that act on nucleic acid substrates also induce or undergo conformational changes. These perturbations contribute to catalysis

E-mail address of the corresponding author:  
bstoddard@fred.hcrc.org

by aligning reactive functional groups in the active site, by increasing binding energy in the enzyme-substrate complex, or by inducing localized strain in the substrate. Cre recombinase (Guo *et al.*, 1997) catalyzes homologous recombination by binding to loxP sites and bending the DNA by approximately 100°. Methyltransferases such as *Hae*III (Reinisch *et al.*, 1995) and *Hha*I (Klimasauskas *et al.*, 1994) induce a conformational change in their substrates by flipping a cytosine nucleotide out of the DNA helix. Similarly, uracil-DNA glycosylase is a base excision enzyme that flips out and removes uracil residues (Slupphaug *et al.*, 1996). Topoisomerases induce changes in DNA structure in order to resolve topological strain in the double helix. The type II family of topoisomerases pass an intact double helix through a double-stranded break in another strand (Berger *et al.*, 1996), while members of the type I family make a single-stranded break, allow the DNA to undergo a rotation along the helical axis, and then re-ligate the broken strand (Redinbo *et al.*, 1998; Stewart *et al.*, 1998).

Restriction endonucleases catalyze double-stranded breaks in DNA and are believed to use lateral diffusion along the double helix to search out their target sites efficiently. Type II restriction endonucleases and their bound DNA exhibit a wide range of conformational changes when the correct cognate sequence is encountered. *Eco*RV binds non-cognate DNA sequences in an unperturbed B-form conformation, but induces a 50° bend in its specific restriction sequence (Winkler *et al.*, 1993). The structure of the *Eco*RI-DNA complex reveals a kink of 25° in the DNA substrate (McClarín *et al.*, 1986), whereas smooth bending is observed in DNA bound to *Bgl*III (23°) (Lukacs *et al.*, 2000) and *Bgl*II (20°) (Newman *et al.*, 1998). In contrast, cognate DNA sites are bound in a B-form conformation by both *Bam*HI (Newman *et al.*, 1995) and *Pvu*II (Athanasiadis *et al.*, 1994; Cheng *et al.*, 1994). *Bgl*III and *Bam*HI share similar recognition sequences, yet bind their respective substrates in different conformations (Lukacs *et al.*, 2000). A range of protein conformational changes has been observed upon restriction endonuclease binding. *Bam*HI and *Eco*RV show a wide range of changes upon binding, including a rotation of subunits, the folding of disordered regions, and the unfolding of C-terminal  $\alpha$ -helices (Newman *et al.*, 1995; Winkler *et al.*, 1993). *Pvu*II displays a conformational change that results in the formation of a histidine-histidine hydrogen bond by completely encircling the DNA (Cheng *et al.*, 1994).

Homing endonucleases are encoded by open reading frames nested within mobile introns or inteins, and recognize lengthy DNA target sites (14 to 40 bp). Cleavage of these sites initiates "homing", a high-frequency, site-specific gene conversion event that transfers a copy of the intervening sequence to a specific insertion point at or near the cleavage site (Belfort & Perlman, 1995; Belfort *et al.*, 1995; Belfort & Roberts, 1997; Dujon, 1989; Jurica & Stoddard, 1999; Lambowitz & Belfort, 1993;

Lambowitz *et al.*, 1998; Mueller *et al.*, 1993). Homing of group I introns requires homology between the donor and recipient alleles. Their homing endonucleases have been grouped into four distinct families based on conserved sequence motifs; LAGLIDADG, His-Cys Box, GIY-YIG, and HNH. I-PpoI was the first identified member of the His-Cys box homing endonuclease family and is encoded by an open reading frame in the third intron of the nuclear, extrachromosomal 26 S rRNA gene of the slime mold *Physarum polycephalum* (Muscarella *et al.*, 1990b; Muscarella & Vogt, 1989). The His-Cys box sequence motifs are responsible for forming two structural zinc-binding sites per monomer and the enzyme is dependent on divalent cations (usually magnesium) for activity. I-PpoI is a homodimer that cleaves a partially symmetric 15 bp homing site generating four base, 3' cohesive overhangs of sequence 5'-TTAA-3' (Ellison & Vogt, 1993; Lowery *et al.*, 1992; Wittmayer & Raines, 1996).

Unlike restriction sites, homing sites may vary in sequence at many of their individual nucleotide positions while still being recognized and cleaved. I-PpoI tolerates base-pair substitutions at several positions within the homing site, but exhibits a strong preference for AT base-pairs in the central four positions of the homing site (Argast *et al.*, 1998). The structures of I-PpoI in complex with its DNA substrate and product have been determined at high resolution (Flick *et al.*, 1998; Galburt *et al.*, 1999). The homing site is severely bent in the protein-DNA complex, resulting in a sharp bend at the cleavage site. In the complex, leucine residues 116 and 116' are observed to make edge-on contacts to the adenine nucleotides directly at the primary bend sites. The conformational changes observed in this enzyme-DNA complex appear to be important for proper alignment of scissile phosphate groups with catalytic groups in the enzyme active site. These conformational changes may also promote binding and catalysis by maximizing surface complementarity, binding energy, and sequence-specific contacts between the enzyme and DNA. Here, we report the results of structural and functional analyses of I-PpoI, designed to determine the degree to which the bound complex represents an induced fit of enzyme and DNA substrate, and to determine the role of Leu116 in binding and catalysis.

## Results

### Protein conformational changes

I-PpoI displays an unusual dimeric structure, with the final 18 residues of each monomer (146 to 163) forming a domain-swapped C-terminal tail that wraps around the surface of the opposite subunit. The structure of individual I-PpoI monomers, excluding these domain-swapped tails, is unchanged when comparing the unbound and DNA-bound proteins (r.m.s.d 0.7 Å). The overall

fold of the subunits is identical and all of the secondary structural elements are intact. The DNA-binding  $\beta$ -sheets are well formed in the absence of DNA, and can be independently superimposed from the same two structures with an r.m.s.d. of 0.8 Å. However, a rigid-body rotation of the monomers relative to one another is observed when the bound and unbound protein structures are superimposed (Figure 1). The DNA-binding sheets are approximately 5 Å closer together in the bound protein-DNA complex than in the unbound structure. This is a result of a hinge motion of 5° that changes the relative orientation of the two monomers. This motion involves a smooth, incremental series of protein backbone rotations between residues 144 and 148, near the dimer interface and directly preceding the domain-swapped tails of each subunit. When single monomers from each structure, excluding their C-terminal tails, are superimposed as shown in Figure 1, the r.m.s.d between the unrestrained subunits is 4.4 Å, with deviations increasing from 1 Å close to the dimer interface to greater than 7 Å at the distal end of the DNA binding  $\beta$ -sheet.

The two structural zinc atoms are present in each monomer, but the catalytic magnesium atoms are not seen in either active site. This was expected, as two DNA oxygen atoms and only one uncharged protein ligand (Asn119) directly contribute to the coordination of the active-site metal in the complex. A bound magnesium-water cluster in a similar active site is observed in an apo structure of *Serratia* nuclease when the crystal is soaked in high concentrations of magnesium chloride (Miller *et al.*, 1999). The active-site residues (His98, Asn119, Cys105, Arg61) overlay very closely in the aligned monomer (r.m.s.d. 0.9 Å); however, the positions of the two active sites relative to one another are significantly different as a result of the 5° hinge motion between monomers.

### DNA conformational changes

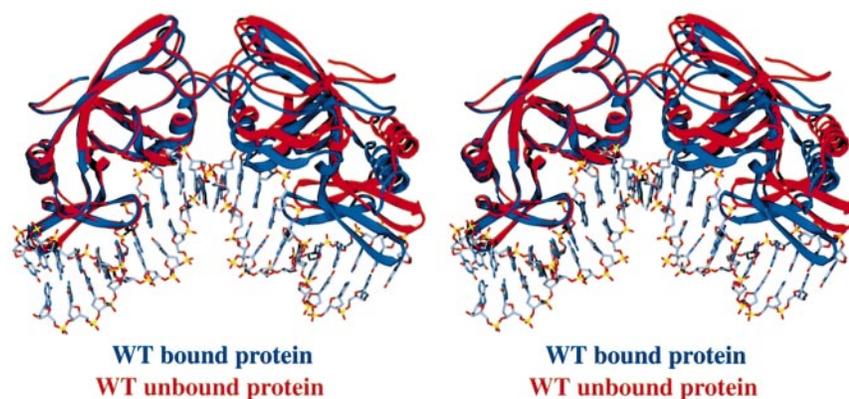
DNA bending induced by binding of wild-type I-PpoI to its homing site was measured using a circular permutation assay (data not shown). The

bend angle calculated from subsequent analyses (see Materials and Methods) is 81.7(±2.4)°. No homing site DNA bending was detected in the absence of I-PpoI, indicating that the bent conformation of the DNA is induced upon protein binding.

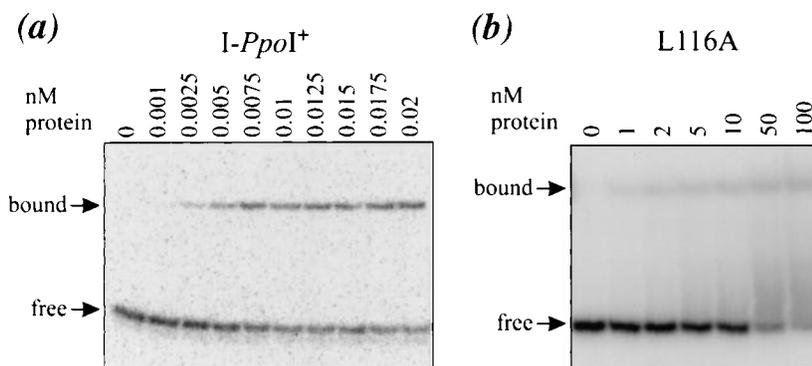
### Effects of the L116A mutation

The positioning of Leu116 near A+2 in the protein-DNA crystal structure suggested a role for this residue during DNA binding and/or cleavage. Substitution of alanine at this position had a dramatic effect on both of these activities. Gel mobility shift assays (Figure 2) revealed that the binding constant ( $K_d$ ) of the L116A mutant is increased at least 5000-fold above that of wild-type I-PpoI. It was not possible to assign an absolute  $K_d$  value to L116A by this method, as protein concentrations greater than 50 nM caused complexes to migrate aberrantly. The poor DNA-binding activity of the mutant precluded use of the circular permutation assay to estimate DNA bending. The ability of L116A to cleave a supercoiled plasmid substrate is also compromised; the mutant protein is 72,000-fold less active than wild-type I-PpoI at pH 10.0 (Figure 3).

To study the structural effects of Leu116, the structure of an L116A/DNA complex was determined. The overall structure of the L116A/DNA complex is very similar to the wild-type complex, and individual contacts between the protein, nucleotides in the major groove, and the phosphate backbone are conserved. However, the overall protein conformation aligns better with the unbound wild-type structure (r.m.s.d. 0.7 Å) than with the bound wild-type structure (r.m.s.d. 1.1 Å) (Figure 4). The difference in conformation between bound L116A and bound wild-type I-PpoI is the same as the difference described above for the wild-type bound and unbound I-PpoI structures. Because of the relative change in orientation between monomers, the DNA-binding  $\beta$ -sheets of the subunits are further away from each other by approximately 5 Å. This less compact protein conformation is matched by a slightly less bent DNA



**Figure 1.** An alignment of the structure of the wild-type protein-DNA complex (1A73) with the structure of the wild-type protein alone (1EVX). The structures were aligned using only the DNA-binding  $\beta$ -sheet on the left. The bound protein is shown in blue and the unbound protein is shown in red. The bound DNA homing site is colored according to atom type.



**Figure 2.** Gel mobility-shift assay of specific DNA-binding activity by (a) wild-type I-PpoI and (b) L116A. Increasing levels of dimeric I-PpoI (nM concentrations shown above lanes) were incubated with  $^{32}\text{P}$ -labeled dsPpo38 oligonucleotide (2 pM in (a); 2 nM in (b)) in 25 mM Caps/Ches (pH 10.0), 50 mM NaCl, 2 mM DTT, 10 mM  $\text{MgCl}_2$ , 0.1 mg/ml bovine serum albumin for 30 minutes at 25 °C. Apparent  $K_d$  values of wild-type I-PpoI and L116A are 10 pM and >50 nM, respectively.

duplex. The conformations of protein and DNA in the wild-type and L116A complexes are mutually exclusive. For example, in the wild-type complex, there are 2.8 Å contacts between both Gln63-A+6 and Arg74-G+7. When the DNA structure from the wild-type complex is docked to the protein from the mutant L116A complex, these contacts extend to 6.1 Å and 6.2 Å, respectively. However, these contact distances are still 2.8 Å in the mutant protein-DNA complex, indicating that enzyme-substrate complementarity is maintained across the interface of the L116A mutant despite the described conformational changes.

In the wild-type complex, L116 makes an edge-on contact to A+2. The contact is such that these groups are not in close van der Waals contact, but their calculated solvent-accessible surfaces intersect (Figure 5). The closest approach between the two groups is 3.7 Å between  $\text{C}^{\delta 2}$  of leucine and N-3 of the adenine nucleotide. In the mutant complex, the average distance from  $\text{C}^{\beta}$  of Ala116 and N-3 of adenine is 6.5 Å. The L116A substitution creates a hole in the protein-DNA interface by removing three carbon atoms. This results in an increase in the solvent-accessible surface area and a decrease in the buried surface area of the interface at this position.

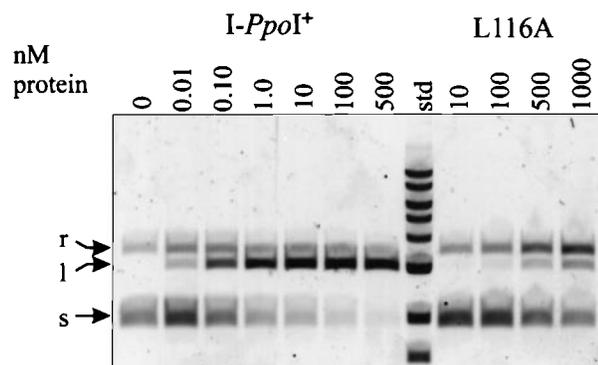
I-PpoI cleaves two phosphodiester bonds, leaving four base, 3' overhangs. In the wild-type product complex, the four base overhangs are well ordered ( $B_{\text{avg}} = 13.5 \text{ \AA}^2$ ) compared to the rest of the DNA atoms ( $B_{\text{avg}} = 19.3 \text{ \AA}^2$ ). In the L116A complex, the central four bases are extremely disordered. The poor visual quality of the map in this region was corroborated by the high group  $B$ -values for these atoms ( $B_{\text{avg}} = 86.9 \text{ \AA}^2$ ) as compared to the rest of the DNA atoms ( $B_{\text{avg}} = 21.5 \text{ \AA}^2$ ). More specifically, the  $B$ -values increase dramatically at the adenine nucleotide that leucine 116 contacts in the wild-type structure. While the average  $B$ -value for the preceding guanine nucleotide is 7.8 Å<sup>2</sup>, it jumps

to 90.4 Å<sup>2</sup> for the adenine nucleotide lacking the leucine contact (Figure 6(a)).

## Discussion

In the wild-type complex, approximately 12,500 Å<sup>2</sup> of surface area is buried between I-PpoI and DNA. As in many biomolecular complexes, it is likely that the burial of surface area drives the association through the free energy gain of releasing solvent molecules. For I-PpoI, the formation and desolvation of the protein-DNA interface requires a severe distortion of  $B$ -form DNA to a geometry with lower twist angles and moderately high roll angles (Figure 6(b)). This distortion results in a compaction of the major groove and an expansion of the minor groove such that at the center of the homing site, the minor groove is 5 Å wider than the major groove. The nucleotide step at which L116 makes contact shows the largest deviation in base-pair slip and slide, which results in the unstacking of an AT-CG step (Figure 6(c) and (d)). The 4 bp stretch of DNA that lies between L116 and L116' in the complex exhibits the most distorted major and minor groove widths. As mentioned in Results, this region is well ordered in the wild-type-product complex and is disordered in the L116A-product crystal structure (Figure 6(a)). Based on these results, L116 appears to be important for maintaining a well-ordered DNA complex at the cleavage site, in which the DNA substrate is maximally deformed for desolvating nucleotide bases that are partially unstacked in the DNA complex.

I-PpoI's use of leucine residues to stabilize a deformed DNA conformation appears fundamentally different from other proteins that have been visualized structurally and that use hydrophobic residues in protein-DNA interfaces. We can contrast the leucine-DNA contact that I-PpoI makes to that made by the purine repressor (purR) bound to its operator sequence. The structure of the purR-

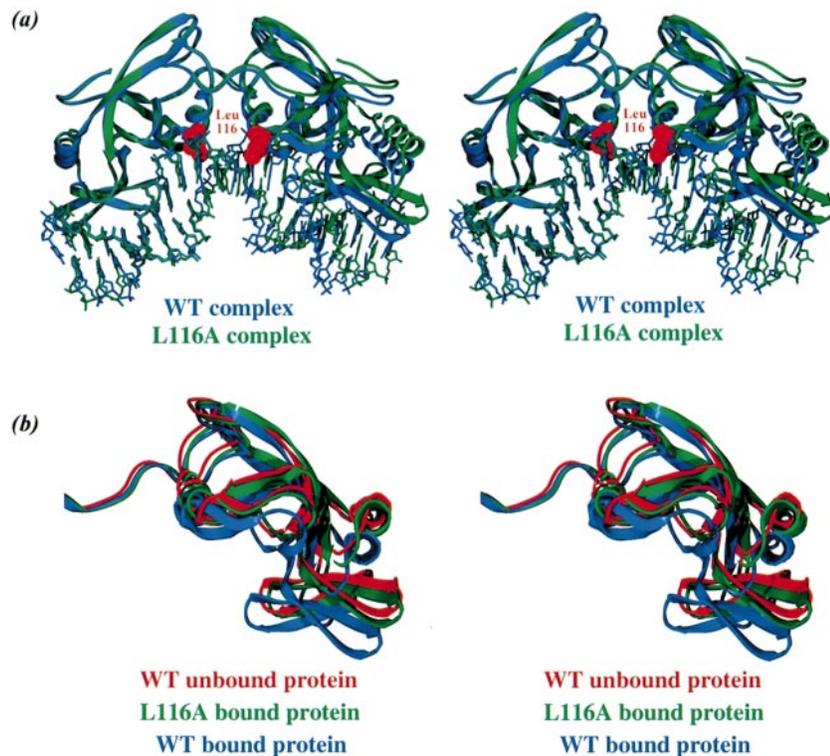


**Figure 3.** Agarose gels showing the relative cleavage activities of wild-type I-PpoI and L116A. Cleavage reactions (10  $\mu$ l) contained the concentrations of protein indicated above the lanes (nM), 10 nM p42 plasmid substrate, 25 mM Caps/Ches (pH 10.0), 50 mM NaCl, 2 mM DTT, 10 mM MgCl<sub>2</sub>, 0.1 mg/ml bovine serum albumin. Reactions were incubated at 37°C for 60 minutes before being quenched and electrophoresed. Arrows labeled r, l, and s indicate relaxed, linear, and supercoiled forms of p42, respectively.

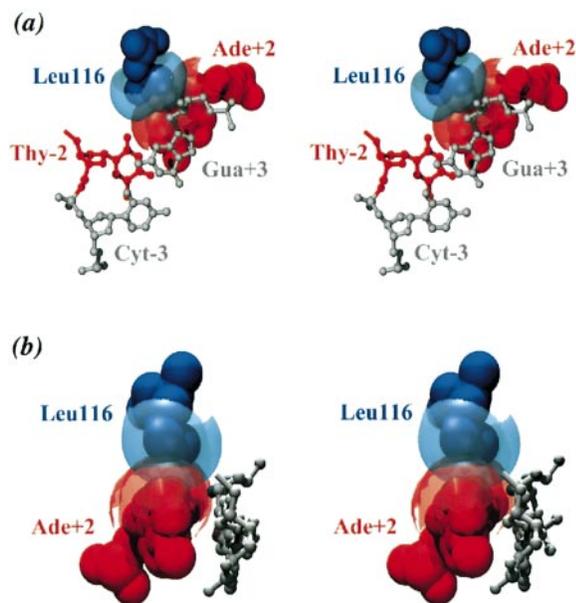
purF operator complex reveals a bound DNA conformation that is somewhat similar to that of I-PpoI bound DNA (Schumacher *et al.*, 1994). An alignment of the two DNA structures shows that both have centrally compacted major grooves and expanded minor grooves, and that both helix axes are bent by similar amounts. However, a closer analysis yields insight into the different ways that

proteins use hydrophobic residues to affect DNA conformation. PurR intercalates a leucine residue, L54, between the central CG-GC base-step nucleotides to induce a large roll angle of almost 50° (Figure 6(b)) and a drastically reduced twist angle. In addition, the inserted leucine residue causes the individual base-pairs to buckle such that they are aplanar by  $\sim$ 20°. This use of a protein residue as a wedge is similar to TBP's use of phenylalanine residues to unstack the base-pairs TA and CG (J.L. Kim *et al.*, 1993; Y. Kim *et al.*, 1993). Bound TATA box DNA also displays large roll and buckle angles where the phenylalanine residues are positioned. In contrast, I-PpoI uses leucine residues to make edge-on contacts with adenine nucleotides such that there is only a slight deviation in roll and twist angles from the B-form. Instead, the majority of the distortion is caused by the sliding of the bases past one another.

The L116 residue of I-PpoI is too far away and not properly positioned to act as a wedge, but it is close enough to locally desolvate the minor groove at the site where bases in the substrate are most dramatically unstacked in the protein complex. This local desolvation appears to stabilize the position of the adenine base in the bound complex and compensates for the energetic penalty of unstacking base-pairs. The distorted DNA conformation has almost 10% more accessible surface area than B-form DNA of the same length and sequence; this increase is localized almost entirely to the central six base-pairs of the homing site. To stabilize this DNA conformation and bind tightly, I-PpoI must desolvate this additional surface area.



**Figure 4.** Alignments of the structures of the wild-type protein-DNA complex (1A73) and the L116A protein-DNA complex (1EVW). The alignments were performed in the same way as in Figure 1 by constraining only one  $\beta$ -sheet. (a) The wild-type complex is shown in blue with the mutated leucine residue in red van der Waals spheres and the L116A complex in green. (b) The unconstrained monomers of the unbound wild-type protein (red), the bound L116A protein from panel a (green), and the bound wild-type protein (blue) are rotated to show the extent of the conformational changes and the similarity between the bound L116A protein and the unbound wild-type protein.



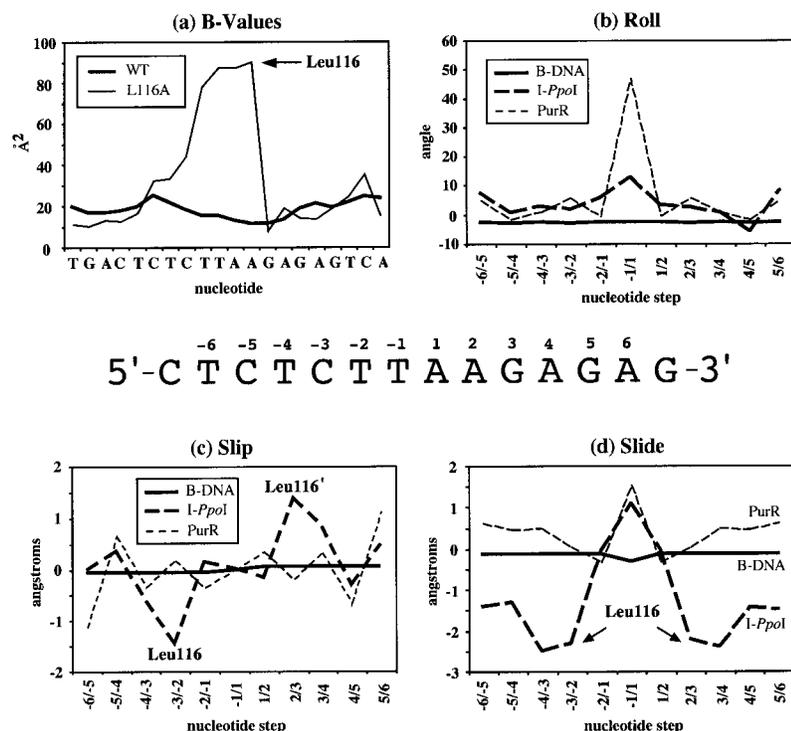
**Figure 5.** Protein-DNA contacts between L116 and A+2. Leu116 is in blue, A+2 and T-2 are in red, and the neighboring DNA nucleotides are in gray. The leucine residue and the adenine nucleotide are shown in van der Waals spheres with their corresponding solvent-accessible surface areas. The van der Waals spheres are not in contact, but the solvent-accessible surface areas intersect. (a) The large slip and slide between A+2 and G+3 is shown. (b) The structure has been rotated 90° and T-2 has been removed to show an edge-on view of the leucine-adenine interaction.

Without L116, additional water molecules can enter the pocket between protein and DNA, and solvate the edge of the adenine and any unstacked base-pairs. The solvation of the exposed hydrophobic surface area would cause an increase in the free energy of the productive protein-DNA complex. This hydrophobic interaction can explain the experimentally observed reduction in binding affinity for L116A. This hypothesis is further supported by the observation that the central four base-pairs are well ordered in the wild-type complex and completely disordered in the mutant complex. In summary, while both purR and TBP use hydrophobic residues as pry bars or wedges to lever apart stacked bases, I-PpoI appears to use L116 to locally desolvate the minor groove, thereby stabilizing a distorted DNA conformation and reducing the energetic cost of base-pair unstacking.

In general, I-PpoI can recognize target sites with one or two base substitutions. However, randomized substrate screens reveal that the central sequence TTAA is strongly preferred by the enzyme (Argast *et al.*, 1998). Most characterized restriction endonucleases specify their target site sequences by making all possible hydrogen bonds to the edges of the nucleotide bases that are accessible in the major and minor grooves. Interestingly,

I-PpoI makes no base-specific contact to any of the four preferred central base-pairs in the target sequence. The only interactions at these positions are between K120 and O-2 of T-1 and the L116-A-2 interaction discussed above. There is no steric clash if other base-pairs are modeled into these positions, and the hydrogen bond made by the lysine residue would almost certainly be conserved with a CG in place of the TA or with a switch of the base-pair from TA to AT. Thus, the observed protein-DNA interactions are unable to explain the sequence conservation. The conservation of these base-pairs might be due to the sequence-dependent conformational preferences of DNA or to the sequence-dependent flexibility of the DNA polymer. This dependence of binding affinity on non-contacted base-pairs has been observed before for various repressor-operator complexes. The crystal structure of the 434 repressor-operator complex reveals that no protein-DNA contact is made to the inner 4 bp of the operator. Despite this, mutations in this region of the operator sequence can significantly alter repressor-operator affinities (Koudelka *et al.*, 1987). More specifically, operators with an AT or TA base-pair bind repressor with higher affinity than those with CG or GC. Further analysis revealed that higher affinity of the AT operators could be explained by a matching of the intrinsic twist of the DNA to the twist observed in the protein-DNA complex (Koudelka & Carlson, 1992). The affinity of the P22 operator varies similarly with the sequence of the central non-contacted bases. The central bases were observed to affect the structure (groove geometry) of both the unbound operator and the operator-repressor complex (Wu & Koudelka, 1993; Wu *et al.*, 1992). The catabolite gene activator protein (CAP) and nucleosomes both induce severe DNA bending without base-specific protein contacts. For both proteins, sequence-dependent analyses revealed that AT content is preferred where the minor groove faces the protein and GC content is preferred where the major groove faces the protein (Drew & Travers, 1985; Gartenberg & Crothers, 1988; Travers, 1989). Based on the lack of specific protein-DNA contacts to the central TTAA sequence in the I-PpoI target site, it is likely that the preferred DNA sequence affects the physical characteristics of the DNA molecule such that it is able to more easily achieve the bent conformation required for high-affinity binding and cleavage.

DNA bending is commonly observed in protein-DNA complexes. In some cases, the bend is important expressly for altering DNA conformation either for compaction (nucleosome core particle) (Luger *et al.*, 1997) or for subsequent molecular recognition (CAP) (Schultz *et al.*, 1991). In other cases, bent DNA is necessary to maximize the shape complementarity of surfaces between a globular protein and an initially linear DNA molecule. Lastly, DNA-bending enzymes may generate a DNA conformation that results in a productive catalytic alignment. I-PpoI appears to bind and



**Figure 6.** A compilation of DNA parameters calculated from the crystal structures using the program FREEHELIX (Dickerson, 1998). (a) A plot of DNA  $B$ -values for the wild-type and L116A complexes. (b), (c) and (d) Plots of DNA roll angles, slip distances, and slide distances, for  $B$ -form DNA, wild-type *I-PpoI* bound DNA, and PurR bound DNA, respectively. Roll refers to the total angle between subsequent base-pair normal vectors, slip refers to the displacement of base-pairs along their short axes, and slide is the analogous metric in the direction of the long axes. The homing site of *I-PpoI* is showed in the middle of the Figure with the numbering scheme used in the graphs.

bend its DNA homing site for a combination of these reasons. The bent conformation increases the buried surface between protein and DNA, thus stabilizing the complex. At the same time, the bend results in the proper positioning of the substrate phosphodiester bonds relative to the two active sites and allows *I-PpoI* to cleave across the minor groove. In  $B$ -form DNA, the scissile phosphate groups leading to four base, 3' overhangs are close together, making it difficult to surround them with two independent endonuclease active sites. The LAGLIDADG homing endonuclease family addresses this problem by tightly packing its active sites together, such that their catalytic aspartate residues and bound divalent cations are only 10 Å apart in the bound DNA complex. In contrast, *I-PpoI* distorts the DNA substrate to widen the minor groove and facilitate the proper positioning of two sets of active-site residues.

LAGLIDADG and His-Cys box family enzymes both cleave across the DNA minor groove (generating four base, 3' overhangs) and form sequence-specific interfaces in the major grooves of the homing site. LAGLIDADG enzymes accomplish this by adopting a structure with tightly packed active sites. In contrast, *I-PpoI* and its target site exhibit conformational motions that allow maximum complementarity across the protein-DNA interface (tight binding) and the alignment of scissile phosphates with more spatially separated active sites (efficient cleavage). It is interesting to note that a subfamily of His-Cys box endonucleases, typified by *I-NjaI*, have been shown to generate five base, 3' overhangs (Elde *et al.*, 1999). In these cases, the homing site scissile phosphate groups would be

further apart in unperturbed  $B$ -form DNA (15 Å instead of 10 Å), and their corresponding homing endonucleases do not contain residues that correspond to Leu116 in *I-PpoI*. This increased required spacing of active sites and the corresponding lack of a leucine residue makes it likely that the DNA-binding mode of these enzymes is different from the mode observed for *I-PpoI*. In conclusion, even homing endonucleases within the same family appear to have evolved a variety of DNA-binding modes and active-site chemistries that accomplish the same biological function.

## Materials and Methods

### Oligonucleotides and plasmids

DNA oligonucleotides used for the biochemical analyses (Table 1) were purchased from Operon Technologies (Alameda, CA). Single-stranded (ss) oligonucleotides were purified from denaturing (7 M urea) 15% polyacrylamide gels; annealed double-stranded (ds) oligonucleotides were purified from non-denaturing 6% polyacrylamide gels. Bands were excised, crushed, and soaked overnight in 0.3 M sodium acetate buffer (pH 5.2), 1 mM EDTA to elute DNA. Supernatant containing DNA was ethanol-precipitated, resuspended in STE (100 mM NaCl, 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA), and quantified by UV spectroscopy. To prepare substrate for the binding assays, a gel-purified 38 bp oligonucleotide, Ppo38<sup>-</sup>, was radiolabeled on the 5'-OH group with [ $\gamma$ -<sup>32</sup>P]ATP and phage T4 polynucleotide kinase (NEB), and then annealed with a 30-fold excess of the complementary oligonucleotide, Ppo38<sup>+</sup>, by heating to 90 °C in 10 mM Tris (pH 8.0), 200 mM NaCl followed by slow cooling to room temperature. Unlabeled non-specific competitor dsDNA, PpoComp

(chemically identical with Ppo38, but with a scrambled I-PpoI homing site), was annealed prior to gel purification. pBENDPpo was constructed by cloning the DNA duplex oligomer PpoHSup/PpoHSlow into the *Xba*I site of the DNA circular permutation assay vector pBEND3 (a gift from Dr Sanker Adhya) (Zwieb & Adhya, 1994). The I-PpoI cleavage assay target plasmid, p42, was a gift from Dr Volker Vogt (Muscarella *et al.*, 1990a).

### Protein purification

The overexpression and purification of wild-type I-PpoI and the crystallization of the wild-type enzyme in complex with DNA have been described (Flick *et al.*, 1997). An I-PpoI L116A mutant was created using CLONTECH's Transformer Site-Directed Mutagenesis Kit. It was expressed and purified exactly as was done for the wild-type.

### Crystallography

Unbound wild-type I-PpoI was crystallized in hanging drops (2  $\mu$ l well with 2  $\mu$ l of protein solution at 6 mg/ml) with a well solution containing 200 mM ammonium sulfate, 18% polyethylene glycol (PEG) 8000, 100 mM sodium cacodylate (pH 6.5). The protein crystallized in the space group  $P4_12_12$  with unit cell dimensions  $a = b = 52.8$  Å,  $c = 278.8$  Å. The crystal was flash frozen after soaking in a cryoprotectant (30% (w/v) glucose, 10% (v/v) glycerol). The asymmetric unit contains a single copy of a protein homodimer. The protein structure from the DNA complex was used as a search model for molecular replacement (EPMR) (Kissinger & Gehlhaar, 1997, Agouron Pharmaceuticals, La Jolla, CA). During the initial rigid-body refinement, each half of the homodimer was allowed to move independently, as were the DNA-binding  $\beta$ -sheets. L166A I-PpoI (4.5 mg/ml) was mixed with a DNA duplex oligomer (XtalOligo2 and its complement) as described (Flick *et al.*, 1997) and the L116A/DNA complex was crystallized in hanging drops (2  $\mu$ l well with 2  $\mu$ l of a 1:1 molar ratio protein/DNA mixture) with a well solution containing 35% (v/v) methyl-2,4-pentanediol, 100 mM Mes (pH 6.5), 20 mM NaCl, 10 mM MgCl<sub>2</sub>. The protein crystallized in space group C2 with unit cell dimensions  $a = 182.0$  Å,  $b = 73.1$  Å,  $c = 92.7$  Å,  $\beta = 95.4^\circ$ . The crystal was flash-frozen directly out of the crystallization medium. A protein monomer combined with the extended C-terminal arm from its dimer partner was used as a search model for molecular replacement (AMoRe CCP4, 1994). After finding four monomer solutions that corresponded to two dimers in the asymmetric unit, maps were calculated and a DNA model could be easily built into density

for each enzyme dimer. Both data sets were collected on an in-house RAXIS IV imaging plate area detector. The DENZO/SCALEPACK program suite was used to process all data sets (Otwinowski & Minor, 1997). X-PLOR was used to perform the refinements of all structures (Brunger, 1992). To calculate an  $R_{\text{free}}$ , 6% of reflections were set aside at the beginning of refinement for both structures (Brunger, 1993). Data and refinement statistics are shown for all data sets in Table 2.

### Protein Data Bank accession numbers

Both structures have been deposited into the RCSB PDB (apo enzyme: 1EVX; L116A complex: 1EVW). The wild-type product complex structure has been solved and reported previously (Brookhaven PDB 1A73) (Flick *et al.*, 1998).

### Gel mobility-shift assay for DNA binding

Gel mobility-shift assays were based on retardation of the electrophoretic mobility of a <sup>32</sup>P-labeled DNA molecule upon the binding of I-PpoI (Fried & Crothers, 1981; Garner & Revzin, 1981). Either 0.02 fmol (I-PpoI<sup>+</sup> assay) or 0.02 pmol (L116A assay) of a <sup>32</sup>P-labeled homing site oligonucleotide, dsPpo38, were incubated for 30 minutes at 25 °C with increasing amounts of I-PpoI in 10  $\mu$ l of binding buffer (25 mM Caps/Ches (pH 10.0), 2 mM EDTA, 50 mM NaCl, 10% glycerol, 50  $\mu$ g/ml bovine serum albumin) lacking Mg<sup>2+</sup> to inhibit cleavage. Non-specific competitor oligonucleotide, dsPpoComp, was present at a 50-fold molar excess. Samples were loaded onto a 0.7 mm  $\times$  16 cm  $\times$  16 cm 10% (75:1 (w/w) acrylamide to bisacrylamide) polyacrylamide gel containing 2% glycerol, and run at 175 V in 0.25  $\times$  TBE for three hours at 4 °C. Dried gels were imaged on a Storm Phosphorimager 840 (Molecular Dynamics, Sunnyvale, CA) and the intensity of the free and bound DNA bands quantified using ImageQuant Software (Molecular Dynamics). The apparent  $K_d$  of the I-PpoI target site complex was defined as the concentration of I-PpoI at which 50% of the DNA was shifted into a complex having retarded electrophoretic mobility (Ausubel *et al.*, 1989). The reported  $K_d$  values represent the average of three independent assays.

### Circular permutation assay for DNA bending

The degree to which I-PpoI bends homing site DNA upon binding was determined by circular permutation analysis, a variation of the gel mobility-shift assay described above. This assay is based on the empirical observation that the mobility of a bent DNA fragment is

**Table 1.** Oligonucleotide sequences

Oligonucleotide	Sequence (5' $\rightarrow$ 3')
Ppo38 <sup>-</sup>	GGCATTGGCTAC <b>CTTAAGAGAG</b> TCATAGTTACTAATT
Ppo38 <sup>+</sup>	AATTAGTAACTATGACT <b>CTCTCTTAAGGTAGC</b> CAAAATGCC
PpoComp <sup>-</sup>	GGCATTGGCGATATACTGAGCTTTACAATAGTCTAATT
PpoComp <sup>+</sup>	AATTAGACTATTGTAAAGCTCAGTATATCGCCAATGCC
PpoHSup	CTAG <b>CTCTCTTAAGGTAGC</b>
PpoHSlow	CTAG <b>GCTACCTTAAGAGAG</b>
XtalOligo1	TGACT <b>CTCTCTTAAGAGAG</b> TCA
XtalOligo2	TGACT <b>CTCTCTTAAGGTAGC</b> CA

I-PpoI homing sites are shown in bold.

**Table 2.** X-ray data collection and refinement statistics

	Wild-type apo protein	L116A product complex	Wild-type product complex <sup>a</sup>
<b>A. Data</b>			
Source	5.0.2 beamline; ALS	Rotating anode	5.0.2 beamline; ALS
Space group	<i>P4</i> <sub>1</sub> <i>2</i> <sub>1</sub> <i>2</i>	<i>C2</i>	<i>P3</i> <sub>1</sub> <i>21</i>
Resolution limit (Å)	2.0	3.1	1.8
Wavelength (Å)	0.98	1.54	0.98
Unit cell <i>a</i> , <i>b</i> , <i>c</i> (Å)	52.8, 52.8, 278.8	182.0, 73.1, 92.7, β 95.4°	114.0, 114.0, 89.0
Unique reflections	25,509	19,707	71,892
Average <i>I</i> / $\sigma$ <i>I</i>	12.2 (7.0)	7.4 (3.8)	19.1
<i>R</i> -merge (%)	5.2 (20.4)	11.0 (23.8)	4.9 (26.7)
Completeness (%)	90.8 (74.8)	88.5 (63.4)	99.8 (98.1)
<b>B. Refinement</b>			
<i>R</i> -factor (%)	19.4	27.3	22.6
<i>R</i> -free <sup>b</sup> (%)	22.9	32.0	24.7
Resolution (Å)	50-2.0	50-3.1	50-1.8
Total number of atoms	2827	6612	3751
Number of water molecules	352	0	403
r.m.s.d bond length (Å)	0.006	0.012	0.011
r.m.s.d bond angles (deg.)	1.30	1.45	1.70
r.m.s.d impropers (deg.)	0.67	0.86	1.04
r.m.s.d dihedrals (deg.)	26.9	28.2	27.6
Mean <i>B</i> value, overall (Å <sup>2</sup> )	17.4	22.3	19.8
Mean <i>B</i> value, protein (Å <sup>2</sup> )	16.2	20.0	15.5
Mean <i>B</i> value, DNA (Å <sup>2</sup> )	NA	19.9	17.3
Mean <i>B</i> value, solvent (Å <sup>2</sup> )	25.0	NA	20.7

The numbers in parentheses are last resolution shell statistics. ALS, advanced light source; NA, not applicable.

<sup>a</sup> Flick *et al.* (1998).

<sup>b</sup> *R*<sub>free</sub> was calculated with 6% of the data.

related to the position of the bend relative to the ends of the molecule. Five different 140 bp fragments were generated from pBENDPpo by *Bam*HI, *Kpn*I, *Eco*RV, *Nhe*I, and *Bgl*II digestions for the circular permutation assay (Zwieb & Adhya, 1994). The center of the I-PpoI target site is positioned 17, 32, 70, 108 and 114 bp from one end of the molecule, respectively. These DNA fragments were gel-purified and labeled at their 5' ends by treatment with [ $\gamma$ -<sup>32</sup>P]ATP and T4 polynucleotide kinase. DNA binding, gel electrophoresis, and visualization of the shifted complexes was carried out as described above, except that the concentration of DNA was increased to 0.2 nM, I-PpoI was present at 0.5 nM, and non-specific competitor was not included. The deviation from linearity,  $\alpha$ , was calculated using the formula  $\mu_M/\mu_E = \cos(\alpha/2)$ , where  $\mu_M$  is the migration distance of a fragment bent in the center, and  $\mu_E$  that of a fragment bent at one end (Thompson & Landy, 1988).  $\mu_M$  and  $\mu_E$  were determined by measuring the distance of the protein-bound *Eco*RV and *Bam*HI fragments from their loading origin. The reported bend angle is an average of five independent assays. Bend angles of less than 10° were beyond the limit of detection of this method.

### Catalytic activity assays

The relative catalytic activity of wild-type I-PpoI and the L116A variant were measured as described (Mannino *et al.*, 1999). Briefly, serial dilutions of enzyme were incubated with the substrate plasmid p42, which contains a single I-PpoI target site, for 60 minutes at 37°C. The concentration of p42 in 10  $\mu$ l reactions was 10 nM; the range of enzyme concentrations tested was 10 pM to 1  $\mu$ M. Cleavage reactions contained 25 mM Caps/Ches (pH 10.0), 50 mM NaCl, 2 mM DTT, 10 mM MgCl<sub>2</sub>, 0.1 mg/ml bovine serum albumin. The assay was performed at pH 10.0, where specific activity is opti-

mal (Mannino *et al.*, 1999). Reactions were quenched by the addition of 2 $\times$  Stop buffer (2% (w/v) SDS, 100 mM EDTA, 20% glycerol, 0.2% (w/v) bromophenol blue). Cleavage products were resolved from uncut DNA by agarose gel electrophoresis. The intensity of the ethidium bromide-stained DNA bands was measured on a Fluorimager SI (Molecular Dynamics). The relative activities of L116A and wild-type I-PpoI were determined by comparing the percentage of linear plasmid generated at each enzyme concentration using ImageQuant software (Molecular Dynamics).

### References

- Argast, G. M., Stephens, K. M., Emond, M. J. & Monnat, R. J. J. (1998). I-PpoI and I-CreI homing site sequence degeneracy determined by random mutagenesis and sequential *in vitro* enrichment. *J. Mol. Biol.* **280**, 345-353.
- Athanasiadis, A., Vlassi, M., Kotsifaki, D., Tucker, P. A., Wilson, K. S. & Kokkinidis, K. (1994). Crystal structure of *Pvu*II endonuclease reveals extensive structural homologies to *Eco*RV. *Nature Struct. Biol.* **1**, 469-475.
- Ausubel, F. M., Thomas, P. J., Mills, A., Moss, D. S. & Palmer, R. A. (1989). *Current Protocols in Molecular Biology*, Wiley, New York.
- Beamer, L. J. & Pabo, C. O. (1992). Refined 1.8 Å crystal structure of the lambda repressor-operator complex. *J. Mol. Biol.* **227**, 177-196.
- Belfort, M. & Perlman, P. S. (1995). Mechanisms of intron mobility. *J. Biol. Chem.* **270**, 30237-30240.
- Belfort, M. & Roberts, R. J. (1997). Homing endonucleases - keeping the house in order. *Nucl. Acids Res.* **25**, 3379-3388.

- Belfort, M., Reaban, M. E., Coetzee, T. & Dalgaard, J. Z. (1995). Prokaryotic introns and inteins: a panoply of form and function. *J. Bacteriol.* **177**, 3897-3903.
- Berger, J. M., Gamblin, S. J., Harrison, S. C. & Wang, J. C. (1996). Structure and mechanism of DNA topoisomerase II. *Nature*, **379**, 225-232.
- Bowie, J. U. & Sauer, R. T. (1989). Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry*, **28**, 7139-7143.
- Brennan, R. G., Roderick, S. L., Takeda, Y. & Matthews, B. W. (1990). Protein-DNA conformational changes in the crystal structure of a  $\lambda$  Cro-operator complex. *Proc. Natl Acad. Sci. USA*, **87**, 8165-8169.
- Brunger, A. (1992). *XPLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
- Brunger, A. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallog. sect. D*, **49**, 24-36.
- CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760-763.
- Cheng, X., Balendiran, K., Schildkraut, I. & Anderson, J. E. (1994). Structure of *PvuII* endonuclease with cognate DNA. *EMBO J.* **13**, 3927-3935.
- Clarke, N. D., Beamer, L. J., Goldberg, H. R., Berkower, C. & Pabo, C. O. (1991). The DNA binding arm of lambda repressor: critical contacts from a flexible region. *Science*, **254**, 267-270.
- Dickerson, R. E. (1998). DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucl. Acids Res.* **26**, 1906-1926.
- Drew, H. R. & Travers, A. A. (1985). DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* **186**, 773-790.
- Dujon, B. (1989). Group I introns as mobile genetic elements: facts and mechanistic speculations - a review. *Gene*, **82**, 91-114.
- Elde, M., Haugen, P., Willassen, N. P. & Johansen, S. (1999). I-NjaI, a nuclear intron-encoded homing endonuclease from *Naegleria*, generates a pentanucleotide 3' cleavage-overhang within a 19 base-pair partially symmetric DNA recognition site. *Eur. J. Biochem.* **259**, 281-288.
- Ellison, E. L. & Vogt, V. M. (1993). Interaction of the intron-encoded mobility endonuclease I-PpoI with its target site. *Mol. Cell. Biol.* **13**, 7531-7539.
- Ferre-D'Amare, A. R., Prendergast, G. C., Ziff, E. B. & Burley, S. K. (1993). Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature*, **363**, 38-45.
- Ferre-D'Amare, A. R., Pognonec, P., Roeder, R. G. & Burley, S. K. (1994). Structure and function of the b/HLH/Z domain of USF. *EMBO J.* **13**, 180-189.
- Flick, K. E., McHugh, D., Heath, J. D., Stephens, K. M. R. J. M., Jr & Stoddard, B. L. (1997). Crystallization and preliminary X-ray studies of I-PpoI: a nuclear, intron-encoded homing endonuclease from *Phy-sarum polycephalum*. *Protein Sci.* **6**, 2677-2680.
- Flick, K. E., Jurica, M. S., Monnat, R. J. & Stoddard, B. L. (1998). DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-PpoI. *Nature*, **394**, 96-101.
- Fried, M. & Crothers, D. M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucl. Acids Res.* **9**, 6505-6525.
- Galburt, E. A., Chevalier, B., Tang, W., Jurica, M. S., Flick, K. E., Monnat, R. J. & Stoddard, B. L. (1999). A novel endonuclease mechanism directly visualized for I-PpoI. *Nature Struct. Biol.* **6**, 1096-1099.
- Garner, M. M. & Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of *Escherichia coli* lactose regulatory system. *Nucl. Acids Res.* **9**, 3047-3060.
- Gartenberg, M. R. & Crothers, D. M. (1988). DNA sequence determinants of CAP-induced bending and protein binding affinity. *Nature*, **333**, 824-829.
- Guo, F., Gopaul, D. N. & Duyne, G. D. V. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, **389**, 40-46.
- Jurica, M. S. & Stoddard, B. L. (1999). Homing endonucleases: structure, function, and evolution. *Cell. Mol. Life Sci.* **55**, 1304-1326.
- Kim, J. L., Nikolov, D. B. & Burley, S. K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **265**, 520-527.
- Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. (1993). Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512-520.
- Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. (1994). *HhaI* methyltransferase flips its target base out of the DNA. *Cell*, **76**, 357-369.
- Koudelka, G. B. & Carlson, P. (1992). DNA twisting and the effects of non-contacted bases on affinity of 434 operator for 434 repressor. *Nature*, **355**, 89-91.
- Koudelka, G. B., Harrison, S. C. & Ptashne, M. (1987). Effect of non-contacted bases on the affinity of 434 operator for 434 repressor and Cro. *Nature*, **326**, 886-888.
- Lambowitz, A. M. & Belfort, M. (1993). Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62**, 587-622.
- Lambowitz, A. M., Caprara, M. G., Zimmerly, S. & Perlman, P. S. (1998). Group I and group II ribozymes as RNPs: clues to the past and guides to the future. In *RNA World II*, Cold Spring Harbor Press, Cold Spring Harbor, NY.
- Lowery, R., Hung, L., Knoche, K. & Bandziulis, R. (1992). Properties of I-PpoI: a rare-cutting intron-encoded endonuclease. *Promega Notes*, **38**, 8-12.
- Luger, K., Mader, A. W., Richmond, R. K. & Sargent, D. F. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251-260.
- Lukacs, C. M., Kucera, R., Schildkraut, I. & Aggarwal, A. K. (2000). Understanding the immutability of restriction enzymes: crystal structure of *BglIII* and its DNA substrate at 1.5 Å resolution. *Nature Struct. Biol.* **2**, 134-140.
- Mannino, S. J., Jenkins, C. L. & Raines, R. T. (1999). Chemical mechanism of DNA cleavage by the homing endonuclease I-PpoI. *Biochemistry*, **38**, 16178-16186.
- McClarín, J. A., Frederick, C. A., Wang, B., Greene, P., Boyer, H. W., Grable, J. & Rosenberg, J. M. (1986). Structure of the DNA-*EcoRI* endonuclease recognition complex at 3 Å resolution. *Science*, **234**, 1526-1541.
- Miller, M. D., Cai, J. & Krause, K. L. (1999). The active site of *serratia* endonuclease contains a conserved magnesium-water cluster. *J. Mol. Biol.* **288**, 975-987.
- Mueller, J. E., Brysk, M., Loizos, N. & Belfort, M. (1993). Homing endonucleases. In *Nucleases* (Linn, S. M., Lloyd, R. S. & Roberts, R. J., eds), vol. 2, pp. 111-

- 143, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Muscarella, D. E. & Vogt, V. M. (1989). A mobile group I intron in the nuclear rDNA of *Physarum polycephalum*. *Cell*, **56**, 443-454.
- Muscarella, D., Ellison, E., Ruoff, B. & Vogt, V. (1990a). Characterization of I-Ppo, an intron-encoded endonuclease that mediates homing of a group I intron in the ribosomal DNA of *Physarum polycephalum*. *Mol. Cell. Biol.* **10**, 3386-3396.
- Muscarella, D. E., Ellison, E. L., Ruoff, B. M. & Vogt, V. M. (1990b). Characterization of I-Ppo, an intron-encoded endonuclease that mediates homing of a group I intron in the ribosomal DNA of *Physarum polycephalum*. *Mol. Cell. Biol.* **10**, 3386-3396.
- Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I. & Aggarwal, A. K. (1995). Structure of BamHI endonuclease bound to DNA: partial folding and unfolding on DNA binding. *Science*, **269**, 656-663.
- Newman, M., Lunnen, K., Wilson, G., Greci, J., Schildkraut, I. & Philips, S. E. V. (1998). Crystal structure of restriction endonuclease BglI bound to its interrupted DNA recognition sequence. *EMBO J.* **17**, 5466-5476.
- Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
- Raumann, B. E., Rould, M. A., Pabo, C. O. & Sauer, R. T. (1994). DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature*, **367**, 754-757.
- Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J. & Hol, W. G. J. (1998). Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science*, **279**, 1504-1513.
- Reinisch, K. M., Chen, L., Verdine, G. L. & Lipscomb, W. N. (1995). The crystal structure of HaeIII methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. *Cell*, **82**, 143-153.
- Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. (1996). Crystal structure of an IHD-DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295-1306.
- Schultz, S. C., Shields, G. C. & Steitz, T. A. (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science*, **253**, 1001-1007.
- Schumacher, M. A., Choi, K. Y., Zalkin, H. & Brennan, R. G. (1994). Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science*, **266**, 763-770.
- Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E. & Tainer, J. A. (1996). A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87-92.
- Stewart, L., Redinbo, M. R., Qiu, X. & Hol, W. G. J. (1998). A model for the mechanism of human topoisomerase I. *Science*, **279**, 1534-1541.
- Thompson, J. E. & Landy, A. (1988). Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. *Nucl. Acids Res.* **20**, 9687-9705.
- Travers, A. A. (1989). DNA conformation and protein binding. In *Annual Review of Biochemistry* (Richardson, C. C., ed.), vol. 58, pp. 427-452, Annual Reviews, Inc., Palo Alto.
- Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K. & Wilson, K. S. (1993). The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.* **12**, 1781-1795.
- Wittmayer, P. K. & Raines, R. T. (1996). Substrate binding and turnover by the highly specific I-PpoI endonuclease. *Biochemistry*, **35**, 1076-1083.
- Wu, L. & Koudelka, G. B. (1993). Sequence-dependent differences in DNA structure influence the affinity of P22 operator for P22 repressor. *J. Biol. Chem.* **268**, 18975-18981.
- Wu, L., Vertino, A. & Koudelka, G. B. (1992). Non-contacted bases affect the affinity of synthetic P22 operators for P22 repressor. *J. Biol. Chem.* **267**, 9134-9139.
- Zwieb, C. & Adhya, S. (1994). Improved plasmid vectors for the analysis of protein-induced DNA bending. *Methods Mol. Biol.* **30**, 281-294.

Edited by I. A. Wilson

(Received 23 March 2000; received in revised form 28 April 2000; accepted 6 May 2000)