letters

A novel endonuclease mechanism directly visualized for I-*Ppo*I

Eric A. Galburt[1], Brett Chevalier[1], Weiliang Tang[2], Melissa S. Jurica[1], Karen E. Flick[1], Raymond J. Monnat, Jr.[2] and Barry L. Stoddard[1]

[1]Fred Hutchinson Cancer Research Center and the Graduate Programs in Molecular and Cell Biology and Biomolecular Structure and Design, University of Washington, 1100 Fairview Ave. N. A3-023, Seattle, Washington 98109, USA. [2]Department of Pathology, University of Washington, Box 357705, Seattle, Washington 98195, USA.

A novel mechanism of DNA endonucleolytic cleavage has been visualized for the homing endonuclease I-*Ppo*I by trapping the uncleaved enzyme–substrate complex and comparing it to the previously visualized product complex. This enzyme employs a unique single metal mechanism. A magnesium ion is coordinated by an asparagine residue and two DNA oxygen atoms and stabilizes the phosphoanion transition state and the 3'oxygen leaving group. A hydrolytic water molecule is activated by a histidine residue for an in-line attack on the scissile phosphate. A strained enzyme–substrate–metal complex is formed before cleavage, then relaxed during the reaction.

Any phosphodiester cleavage event requires several important mechanistic elements. These include activation of a nucleophile by a general base, stabilization of the phosphoanion transition state and protonation of the leaving group[1]. A common feature of most nucleases is the use of a bound metal ion as a Lewis acid to lower the pK_a of a water molecule, which may then act as a nucleophile or as a base. However, recent studies of catalysts such as homing endonucleases have indicated that the active sites and structural mechanisms of nucleases are more diverse than originally thought[2,3]. Homing endonucleases are encoded by open reading frames within mobile introns[4,5] and are also found as independently folded domains in self-splicing inteins[6,7]. Enzymes in the largest family of these enzymes contain conserved LAGLIDADG or dodecapeptide protein motifs[8]. This family appears to have loosely reproduced the active site architecture of restriction endonucleases[9–12]. In contrast, I-*Ppo*I, a His-Cys box homing endonuclease, utilizes a different mechanism.

I-*Ppo*I is encoded within the third intron of the nuclear 26S rRNA gene of *Physarum polycephalum*[13,14]. It cleaves a 16 bp homing site to generate four-base, 3' overhangs with a k_{cat}/K_m of 10^8 $M^{-1}s^{-1}$ and is activated by many divalent metal ions[15–17]. The structure of I-*Ppo*I in complex with its DNA homing site has been determined at high resolution, both as an uncleaved complex in the absence of bound metal ions (trapped by sulfur substitution of a metal-binding DNA oxygen atom), and as a cleaved product complex that contains a single bound divalent cation per subunit[18]. The metal is coordinated by a conserved asparagine residue, the 3'-OH of the cleaved DNA and four bound water molecules. It is positioned to interact with the scissile phosphate, and cannot coordinate a water molecule for an in-line nucleophilic attack. Instead, a strictly conserved histidine residue (His 98) is properly positioned to activate a water molecule.
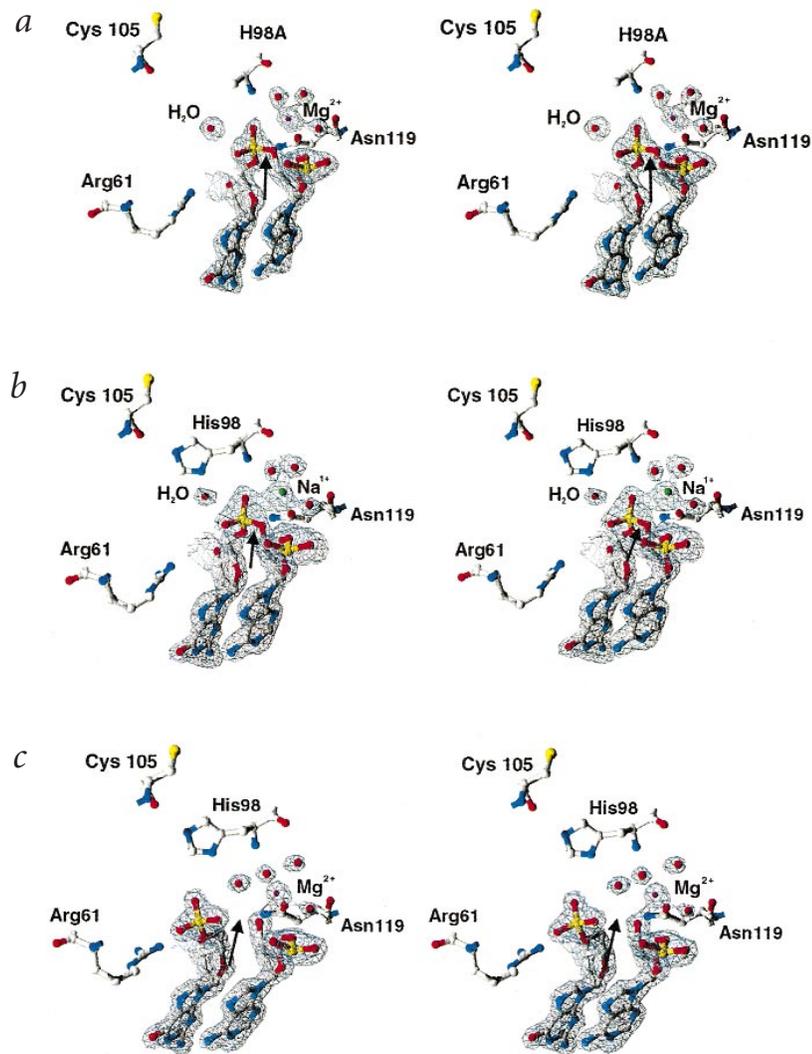
Fig. 1 Electron density difference maps for three catalytic complexes formed by I-*Ppo*I. *a*, Complex of H98A I-*Ppo*I, DNA and magnesium. *b*, Complex of wild type I-*Ppo*I, DNA and monovalent sodium ion. *c*, Complex of wild type I-*Ppo*I, DNA and magnesium. In all three maps, F_o - F_c density corresponding to the atoms omitted from the refinement model is displayed at 3σ contour levels. The maps were calculated after a full round of positional and simulated annealing refinement. Substitution of alanine for histidine at position 98 (*a*) or substitution of an octahedrally coordinated monovalent Na^+ for a similarly coordinated Mg^{2+} (*b*) both fully inhibit cleavage in the crystal. In both trapped E-S complexes a bound hydrolytic water is clearly visible, as is a bound cation associated with the scissile phosphate. In the complex of wild type enzyme in the presence of magnesium (*c*), a similar difference map clearly demonstrates complete cleavage and formation of a product complex. The free 5'-phosphate group has undergone a significant conformational motion and clearly exhibits tetrahedral coordination, consistent with nucleophilic attack and formation of a PO_3^{2-} end group. The side chain of Arg 61 moves slightly to facilitate an electrostatic contact with the released phosphate in the product complex. With motion of the 5'-phosphate, the metal acquires an additional water ligand to complete a sixfold octahedral coordination shell.

Fig. 2 Anomalous difference Fourier map of I-*Ppo*I and DNA with bound manganese ion, shown in the same orientation as Fig. 1. The sole feature of the difference maps when contoured at 3σ is a peak corresponding to the position of the modeled metal cation from the experimental electron density map in Fig. 1.

A nonspecific nuclease from *Serratia marcescens*, which displays a different fold from I-*Ppo*I, also contains conserved, essential asparagine and histidine residues in its active site[19,20]. Superposition of the active site regions of I-*Ppo*I and *Serratia* nuclease indicate that they are very similar and exhibit a similar metal coordination scheme[21,22]. It has been proposed that a similar, novel mechanism of DNA cleavage is used by both enzymes[22,23]. In this mechanism a histidine residue activates a water molecule for in-line attack, and a single metal ion stabilizes the phosphoanion transition state and the 3' leaving group. An asparagine residue is the only metal ligand donated by the protein.

To directly visualize and characterize the mechanism of bond cleavage exhibited by I-*Ppo*I, its structure has been determined as three distinct species on the reaction pathway, using intermediate trapping to capture early states preceding bond cleavage. The enzyme–substrate (E-S) complexes were trapped by substitution of a monovalent cation for an activating divalent cation in the active site, or by substitution of alanine for the His 98 general base. Both of these substitutions inhibit DNA cleavage.

Catalytic metal coordination, substrate destabilization

A single monovalent (Na$^+$) or divalent (Mg^{2+}) cation is bound to the protein–DNA complex in each active site of the protein dimer; this binding site is visible and occupied in all three structures (Fig. 1). The binding of cations at this site has been verified in separate experiments by soaking crystals with electron-rich manganese (Mn^{2+}) ions and examination of isomorphous and anomalous difference Fourier maps (Fig. 2). These experimental results indicate the likelihood that this site is occupied by bound metal cations rather than by a water molecule, in agreement with the recent structural analysis of the *Serratia* nuclease[22]. In the structure of the H98A enzyme–DNA–Mg^{2+} complex, the bound Mg^{2+} ion is coordinated in a six-fold geometry by the side chain oxygen of Asn 119, the bridging 3' oxygen and a single nonbridging oxygen atom of

the scissile phosphate, and three well-ordered water molecules (Figs 1*a*, 3*a*). The average metal-oxygen ligand bond distance is 2.48 Å (σ = 0.09 Å). This distance is slightly longer (by ~0.2 Å) than typical magnesium bond distances observed in a more highly charged metal binding site, and is indicative of the relatively neutral character of the metal ligands in I-*Ppo*I. In an ideal octahedral coordination geometry, all bond angles between nonopposing metal ligands would be near 90°. This geometry is observed in the cleaved product complex. In contrast, both structures of the trapped substrate complex demonstrate that the bond angle from the nonbridging phosphate oxygen, through the bound cation to the 3' bridging oxygen, is significantly strained with a value of 58°. The complementary angle from the same nonbridging phosphate oxygen to 'water 1' is 120°, while all other orthogonal angles through the bound metal are near 90° (σ = 5.6°; Fig. 3*a*).

The structure of the complex of wild type enzyme, DNA and Na$^+$ (Fig. 1*b*), trapped by the substitution of monovalent sodium cations for activating divalent magnesium cations, is in agreement with these observations. Electron density in omit maps supports the presence of an uncleaved phosphodiester bond and a bound cation in a strained six-fold octahedral coordination, similar to that described for the substrate complex trapped by the H98A mutation in the presence of Mg^{2+}. This was predicted, as sodium resembles magnesium in its ionic radii and its number of electrons and can display octahedral coordination both in solution and in crystal lattices[24]. Substitution of sodium for magnesium in the active site appears to inhibit

Fig. 3 Interactions of DNA, solvent and bound cation in the region of the scissile phosphate. *a*, The uncleaved substrate complex. Distances and angles shown are an average for the two trapped E-S complexes as described in the text. In each structure, the distances and interactions are identical for the two active sites in the crystallographic asymmetric unit. The angle from the refined water position, through the phosphate atom to the 3' oxygen leaving group (H$_2$O-P-3' O) is 168°, which is reasonable for a direct S$_N$2 displacement leading to inversion of configuration of the phosphate group. The octahedral coordination of the bound metal is clearly strained, due to the interaction of the metal with adjacent bridging and nonbridging oxygen atoms from the scissile phosphate. *b*, The product complex. Movement of the 5'-phosphate away from the bound cation has allowed the binding of an additional inner shell water ligand and optimization of the six-fold coordination around the metal.

letters

Fig. 4 Proposed catalytic mechanism for the I-*PpoI*/*Serratia* endonucleases. The single bound metal participates solely in transition state stabilization and protonation of the 3' oxygen leaving group; it does not activate the hydrolytic water molecule. The general base for the reaction is a histidine imidazole ring, activated by a hydrogen bond between its ε-nitrogen and a backbone carbonyl oxygen. Residue numberings are from I-*Ppo*I.

cleavage by reducing the charge density of the bound cation, leading to inadequate stabilization of the phosphoanion transition state.

The single bound metal ion in I-*PpoI* appears to serve three distinct roles in catalysis (Fig. 4). Direct interaction of the bound metal ion with the scissile phosphate indicates that Mg^{2+} stabilizes the phosphoanion intermediate and the 3' hydroxylate leaving group. Second, a water molecule in the inner coordination sphere of the metal is appropriately positioned to donate a proton to the 3' hydroxylate leaving group. The metal ion decreases the pK_a of this water molecule and accelerates proton transfer. Finally, the bound metal forms a geometrically strained octahedral complex with surrounding protein, DNA and solvent atoms that is relaxed after DNA bond cleavage. A growing number of studies have demonstrated that substrate destabilization is an important mechanism for rate enhancement for many biological catalysts.

It is interesting to note that the I-*PpoI* endonuclease displays full catalytic function with a wide variety of divalent cations, while most restriction enzymes and the LAGLIDADG homing endonucleases are far more restricted in their use of metals. Promiscuous use of metal ions by an endonuclease might indicate that the bound cation is used predominantly for charge stabilization of the phosphoanion transition state. In contrast, greater sensitivity of cleavage rates with respect to metal species might imply a metal ion is also used to precisely position and activate a nucleophile. Recent studies on the *Eco*RV endonuclease, measuring rates of E-S association, hydrolysis and dissociation with different metal ion species, appear to support this conclusion[25].

Alignment and activation of the hydrolytic water
In both E-S complexes we observed an electron density peak corresponding to an ordered water molecule positioned for in-line attack on the scissile phosphate (Figs. 1*a,b*, 3*a*). In these structures the modeled water molecule is 3.9 Å from the phosphorus atom, and the angle from the water molecule through the phosphorus atom to the 3' oxygen leaving group is 168°. The appearance and position of the bound solvent molecule is not simply a result of the H98A mutation, but is a consistent feature of the uncleaved enzyme–DNA complex. The δN of His 98 appears to be directly hydrogen-bonded to the water molecule, with a distance between the imidazole nitrogen and water oxygen atom of 2.6 Å. This histidine appears to act as a general base by activating the water molecule (Fig. 4) and may also participate in stabilization of the phosphoanion transition state.

Because the observed nucleophilic water molecule is not associated with a bound cation or any other electrophilic group, its pK_a is likely to be higher than the metal-bound water nucleophiles that are postulated for enzymes such as *Bam*HI or *Eco*RV. Because the pK_a of an uncharged histidine residue is only about 6, it would seem likely that such a side chain must be rendered a stronger base through an interaction with a hydrogen bond acceptor in order to effectively deprotonate this water molecule. In both substrate complexes the backbone carboxylate oxygen of Cys 105 is 2.8 Å from the His 98 εN, and is positioned to form a linear hydrogen bond (Fig. 3*a*). In the structure of the *Serratia* nuclease[19], the putative general base (His 89) displays a similar interaction between its εN and Asn 106. However, there are currently no reported pH dependence studies of the chemical step of the I-*PpoI* reaction, nor has the importance of this interaction been experimentally tested by mutation of Asn 106 in *Serratia* nuclease or by measurement of the pK_a of His 98 in I-*PpoI* .

Conformational changes and transition state stabilization
A series of conformational changes are observed in the active site as a result of DNA bond cleavage (Figs 1*c*, 3*b*). The free 5'-phosphate moves by over 2.5 Å from its position in the substrate complex and forms a 2.8 Å electrostatic bond with a guanidonitrogen of Arg 61, which moves by ~0.5 Å. The movement of the 5'-phosphate disrupts the interaction between its non-bridging oxygen and the bound metal ion. A fourth well-resolved water molecule is added to the inner metal coordination sphere, which assumes a more ideal octahedral geometry. The previously strained metal bond angles relax by ~20°, and the average orthogonal coordination angle for the ligands is 89.6° ($\sigma = 5.3$). The metal ion does not move significantly upon cleavage, and maintains interactions with Asn 119 and the 3' oxygen leaving group of the cleaved phosphodiester bond.

These structures indicate that the phosphoanion transition state is stabilized through contacts with the bound metal ion and the imidazole ring of His 98 (Fig. 3). This contact exists in the E-S complex as a polar interaction with the hydrolytic water molecule and is maintained in the free 5'-PO_3 group of the E-P complex. Arg 61 does not appear to play a role in transition state stabilization, because the distance from this side chain to the scissile phosphate before bond cleavage is too long, at 5.5 Å. Arg 61 does, however, appear to stabilize the final product complex, and thus may help to drive the equilibrium of the reaction forward by inhibiting re-ligation.

Methods

The overexpression and purification of wild type I-*PpoI* have been described[26]. The H98A was expressed and purified identically to wild type. CHELEX 100 Resin (Bio-Rad Laboratories) was used to remove divalent cations in the metal-free crystallization. For this latter experiment, the protein and DNA were dialyzed extensively against buffer containing solid CHELEX resin, as were all crystallization buffers. The relative cleavage activity of the wild type enzyme, the H98A mutant and the wild type enzyme in the absence of divalent cations was determined by digestion of circular plasmid and linearized substrates (unpublished results). Both the H98A substitution and the removal of divalent cation inhibit cleavage.

The crystallization of the enzyme in complex with the cleaved product DNA and divalent cations has been described[26]. Under similar conditions cocrystallization of the wild type protein and DNA in the absence of divalent cations, or of the H98A mutant and DNA in the presence of divalent cations yield uncleaved substrate complexes. All complexes crystallized in $P3_121$ with unit cell dimensions a = b = 113.9 Å, c = 89.0 Å; the asymmetric unit is an enzyme dimer and its bound DNA palindrome. Both the data for H98A and for the metal-free complex were collected from cryocooled crystals, on an in-house RAXIS IV imaging plate area detector and at beamline 5.0.2 at the ALS (λ = 0.9792 Å), respectively. Data for Mn^{2+}-substituted structures were collected by soaking wild type crystals in 100 mM manganese for 1 h and then collecting a data set at beamline 5.0.2 at ALS, using an incident wavelength of 1.89 Å (6.54 keV), corresponding to the anomalous edge of the manganese ion. For this data set, a full quadrant of reciprocal space (90°) was collected to maximize redundancy; and the crystal was oriented with the three-fold crystallographic symmetry axis offset by ~20° from the X-ray axis to increase the percentage of Friedel mates measured on the same exposure. All data sets were processed using the DENZO/SCALEPACK program suite[27]. For the manganese-soaked data set, isomorphous and anomalous difference maps were calculated after an initial round of refinement with all metal and solvent coordinates removed from the model. The refinements of all structures were performed using the X-PLOR package[28]. Deoxyribonucleotide and phosphate atoms that span the cleavage site and all water and cation atoms were omitted from the model before initial refinement cycles and map calculations. To calculate an R_{free} during refinement, 6% of reflections were set aside before any refinement cycles[29]. Data and refinement statistics are shown for all data sets in Table 1.

Coordinates. Coordinates of the uncleaved substrate complexes and the product complex have been deposited in the Protein Data Base (PDB accession numbers 1CZ0, 1CYQ and 1a73).

Correspondence should be addressed to B.L.S. *email:bstoddar@fhcrc.org*

Table 1 X-ray data collection and refinement statistics[1]

Data set	H98A with Mg^{2+}	Wild type with Na^+	Wild type with Mn^{2+}
Source	Rotating anode	5.0.2 beamline (ALS)	5.0.2 beamline (ALS)
Resolution limit (Å)	1.93	2.10	2.20
Wavelength (Å)	1.54	0.98	1.89
Unit cell (Å)	113.9, 113.9, 89.0	113.9, 113.9, 89.0	113.9, 113.9, 89.0
Measured reflections	171,353	206,866	387,325
Unique reflections	49,539	38,078	34,139
R_{merge} (%)	4.0 (30.2)	6.7 (17.8)	5.8 (22.9)
Completeness (%)	98.3 (82.2)	96.9 (97.8)	99.6 (95.3)
Refinement			
R-factor (%)	20.4	19.1	19.7
R_{free} (%)[2]	23.5	21.6	24.7
Resolution (Å)	50–1.93	50–2.1	50–2.2
Total number of atoms	3,702	3,744	3,717
Number of water molecules	362	396	358
R.m.s. deviations			
Bond length (Å)	0.005	0.006	0.007
Bond angles (°)	1.15	1.18	1.21
Impropers (°)	1.03	1.04	1.07
Dihedrals (°)	27.0	27.3	26.9
Mean B value			
Overall (Å^2)	18.6	21.9	20.3
Protein (Å^2)	17.0	20.6	20.7
DNA (Å^2)	20.6	22.8	21.4
Solvent (Å^2)	25.5	28.1	27.2
Cations (Å^2)	Mg^{2+}, 17.9	Na^+, 21.8	Mn^{2+}, 18.9

[1]The numbers in parentheses are statistics from the highest resolution shell as reported by X-PLOR.
[2]R_{free} was calculated with 6% of the data.

1. Pingoud, A. & Jeltsch, A. *Eur. J. Biochem.* **246**, 1–22 (1997).
2. Jurica, M.S. & Stoddard, B.L. *Cellular and Molecular Life Sciences* **55**, 1304–1326 (1999).
3. Aggarwal, A.K. & Wah, D.A. *Curr. Opin. Struct. Biol.* **8**, 19–25 (1998).
4. Belfort, M. & Roberts, R.J. *Nucleic Acids Res.* **25**, 3379–3388 (1997).
5. Belfort, M. & Perlman, P.S. *J. Biol. Chem.* **270**, 30237–30240 (1995).
6. Belfort, M., Reaban, M.E., Coetzee, T. & Dalgaard, J.Z. *J. Bacteriol.* **177**, 3897–3903 (1995).
7. Pietrokovski, S. *Protein Sci.* **63**, 64–71 (1998).
8. Dalgaard, J.Z. *et al. Nucleic Acids Res.* **25**, 4626–4638 (1997).
9. Silva, G.H., Dalgaard, J.Z., Belfort, M. & Roey, P.V. *J. Mol. Biol.* **286**, 1123–1136 (1999).
10. Heath, P.J., Stephens, K.M., Monnat, R.J. & Stoddard, B.L. *Nature Struct. Biol.* **4**, 468–476 (1997).
11. Duan, X., Gimble, F.S. & Quiocho, F.A. *Cell* **89**, 555–564 (1997).
12. Jurica, M.S., Monnat, R.J. & Stoddard, B.L. *Mol. Cell* **2**, 469–476 (1998).
13. Muscarella, D.E. & Vogt, V.M. *Cell* **56**, 443–454 (1989).
14. Muscarella, D.E., Ellison, E.L., Ruoff, B.M. & Vogt, V.M. *Mol. Cell. Biol.* **10**, 3386–3396 (1990).
15. Lowery, R., Hung, L., Knoche, K. & Bandziulis, R. *Promega Notes* **38**, 8–12 (1992).
16. Whittmayer, P.K. & Raines, R.T. *Biochemistry* **35**, 1076–1083 (1996).
17. Ellison, E.L. & Vogt, V.M. *Mol. Cell. Biol.* **13**, 7531–7539 (1993).
18. Flick, K.E., Jurica, M.S., Monnat, R.J. & Stoddard, B.L. *Nature* **394**, 96–101 (1998).
19. Miller, M.D., Tanner, J., Alpaugh, M., Benedik, M.J. & Krause, K.L. *Nature Struct. Biol.* **1**, 461–468 (1994).
20. Friedhoff, P. *et al. Nucleic Acids Res.* **24**, 2632–2639 (1996).
21. Friedhoff, P. *et al. Nature Struct. Biol.* **6**, 112 (1999).
22. Miller, M.D., Cai, J. & Krause, K.L. *J. Mol. Biol.* **288**, 975–987 (1999).
23. Friedhoff, P., Franke, I., Krause, K.L. & Pingoud, A. *FEBS Lett.* **443**, 209–214 (1999).
24. Cotton, F.A. & Wilkinson, G. In *Advanced inorganic chemistry* 253–270 (John Wiley and Sons, New York; 1980).
25. Baldwin, G.S., Sessions, R.B., Erskine, S.G. & Halford, S.E. *J. Mol. Biol.* **288**, 87–103 (1999).
26. Flick, K.E. *et al. Protein Sci.* **6**, 2677–2680 (1997).
27. Otwinowski, Z. & Minor, W. *Methods Enzymol.* **276**, 307–326 (1997).
28. Brünger, A. *X-PLOR: a system for X-ray crystallography and NMR.* (Yale University Press, New Haven, Connecticut; 1992).
29. Brünger, A. *Acta Crystallogr. D* **49**, 24–36 (1993).