doi:10.1016/S0022-2836(02)01092-6 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2002) 324, 227–236

JMB

The Crystal Structure of the Nuclease Domain of Colicin E7 Suggests a Mechanism for Binding to Double-stranded DNA by the H–N–H Endonucleases

Yi-Sheng Cheng[1]†, Kuo-Chiang Hsia[1]†, Lyudmila G. Doudeva[1] Kin-Fu Chak[2] and Hanna S. Yuan[1]*

[1]*Institute of Molecular Biology Academia Sinica, Taipei 11529 Taiwan, ROC*

[2]*Institute of Biochemistry National Yang Ming University, Taipei 11221 Taiwan, ROC*

**Corresponding author*

The bacterial toxin ColE7 contains an H–N–H endonuclease domain (nuclease ColE7) that digests cellular DNA or RNA non-specifically in target cells, leading to cell death. In the host cell, protein Im7 forms a complex with ColE7 to inhibit its nuclease activity. Here, we present the crystal structure of the unbound nuclease ColE7 at a resolution of 2.1 Å. Structural comparison between the unbound and bound nuclease ColE7 in complex with Im7, suggests that Im7 is not an allosteric inhibitor that induces backbone conformational changes in nuclease ColE7, but rather one that inhibits by blocking the substrate-binding site. There were two nuclease ColE7 molecules in the P1 unit cell in crystals and they appeared as a dimer related to each other by a non-crystallographic dyad symmetry. Gel-filtration and cross-linking experiments confirmed that nuclease ColE7 indeed formed dimers in solution and that the dimeric conformation was more favored in the presence of double-stranded DNA. Structural comparison of nuclease ColE7 with the His-Cys box homing endonuclease I-*Ppo* I further demonstrated that H–N–H motifs in dimeric nuclease ColE7 were oriented in a manner very similar to that of the ββα-fold of the active sites found in dimeric I-*Ppo* I. A mechanism for the binding of double-stranded DNA by dimeric H–N–H nuclease ColE7 is suggested.

Keywords: Endonuclease structure; H–N–H motif; colicin; DNA binding; Zn enzyme

Introduction

Many endonucleases are homodimeric enzymes that recognize and cleave palindromic or non-palindromic sequences of double-stranded DNA, examples include type II restriction endonucleases,[1–3] homing endonucleases,[4] topoisomerases[5,6] and recombinases.[7–9] The advantages of dimeric enzymes in DNA recognition and cleavage are as follows: palindromic DNA has a 2-fold symmetry, which can be "read" proficiently by a homodimeric protein with an identical 2-fold symmetry, as demonstrated by the early research in the helix-turn-helix (H-T-H) family of DNA-binding proteins.[10–12] Furthermore, each of the monomeric subunits of an endonuclease contains a single active site for hydrolysis, therefore dimeric endonucleases can make two nicks and cleave double-stranded DNA efficiently, resulting in blunt ends or overhangs of DNA products, depending on the relative orientation and distance of the two active sites.[13]

Here, we show that protein colicin E7 (ColE7) can dimerize and it may be another example of a dimeric endonuclease that cleaves double-stranded DNA. ColE7 is a non-specific endonuclease classified as one of the *Escherichia coli* secretary toxins that are released under stress conditions to kill other related bacteria.[14–17] After ColE7 enters target cells with the help of its receptor-binding domain and membrane translocation domain, it hydrolyzes cellular DNA and RNA non-specifically, using its nuclease domain, leading to cell death. In order to protect the host cell from the cytotoxic nuclease activity of ColE7, an immunity protein, Im7, is co-expressed with ColE7 to inhibit the nuclease activity of ColE7 in host cells. It has been shown recently, that ColE7 and Im7 are secreted from the host cell as a heterodimeric complex; however, ColE7 is processed in the periplasm with only the

† These two authors contributed equally to this work.
Abbreviations used: ColE7, colicin E7.
E-mail address of the corresponding author: hanna@sinica.edu.tw; http://hyuan.imb.sinica.edu.tw

nuclease domain of ColE7 (referred to as nuclease ColE7) traversing the inner membrane and reaching the cytoplasm of target bacterial cells.[18,19]

The C-terminal domain of nuclease ColE7 contains a 30 amino acid residue H–N–H motif, which has been identified in numerous endonucleases, including homing endonucleases: I-*Hmu* I,[20–22] I-*Hmu* II,[21,22] I-*Hmu* III,[22] I-*Tev* III,[23] I-*Cmoe* I,[24] RF253,[25] *yos*Q,[26] Avi,[27] Cpc,[27] PetD,[28] I-*Sec*V,[29] I-*Sec*VI,[30] I-*Lla* I,[5] I-*Two*,[31] bacteria toxins: ColE2,[14] ColE7,[15] ColE8,[16] ColE9,[17] S1, S2,[32,33] and repair or restriction enzymes: McrA[34] and MutS.[35] Therefore, it is possible that the DNA recognition and hydrolysis mechanism mediated by nuclease ColE7 may apply to other H–N–H family proteins. Crystal structures of nuclease ColE7/Im7[36] and the highly homologous nuclease ColE9/Im9[37] showed that the H–N–H motif has a topology similar to that of a zinc finger motif[38] or treble clef finger motif,[39] containing two β-strands and one α-helix linked together by a divalent metal ion. Enzymological characterization of ColE9 has revealed that the cleavage of DNA produced 3′-hydroxyl and 5′-phosphate termini, with the enzyme preferring to make nicks at thymine bases in double-stranded DNA.[40] Extensive studies on metal-dependence activities[24,38,40–43] for H–N–H proteins have been done; however, the roles of transition metal or alkaline earth metal ions in DNA hydrolysis are still matters for discussion.

The crystal structures of Im7[44] and the nuclease domain of ColE7 (or ColE9) in complex with Im7 (or Im9)[36,37] provide a clear picture of how an inhibitor protein binds to a colicin. Since the nuclease active site in the H–N–H motif of ColE7 is not covered by the inhibitor directly, it was proposed that either the inhibitor blocks the DNA-binding site or inhibitor binding induces conformational changes in colicins.[36,45] In order to better understand the recognition process between H–N–H proteins and DNA, as well as to determine if there are structural changes in nuclease ColE7 after inhibitor binding, we analyzed the crystal structure of the unbound nuclease ColE7 at a resolution of 2.1 Å. The crystal structure of nuclease ColE7 reveals that there is no allosteric effect induced by Im7, in that the enzyme retains its conformation before and after Im7 binding. Biochemical data show that nuclease ColE7 can form dimers in solution, and the presence of double-stranded DNA promotes dimer formation. A mechanism for the recognition of double-stranded DNA by the H–N–H nuclease ColE7 is proposed based on the structural comparison between dimeric nuclease ColE7 and the His–Cys box homing endonuclease I-*Ppo* I.

Results

Structure determination and overall structure

The crystal structure of the unbound nuclease ColE7 was solved by molecular replacement using the structure of the bound nuclease ColE7 (bound with Im7; PDB entry: 7cei) as the searching model. The final structural model contains two nuclease ColE7 (residues 446–576), two zinc ions in the center of each H–N–H motif, and two phosphate ions that bind directly to the zinc ion. The two nuclease ColE7 molecules (labeled 1 and 2) in the asymmetric unit of the P1 unit cell are related to each other by a non-crystallographic 2-fold symmetry and are displayed in blue with only the H–N–H motifs in red (Figure 1). Molecule 2 makes different contacts with molecule 3 (in green in Figure 1), which is symmetrically identical with molecule 1 (moved one unit in the *z*-direction). Molecules 2 and 3 are related to each other by a pseudo 2-fold symmetry.

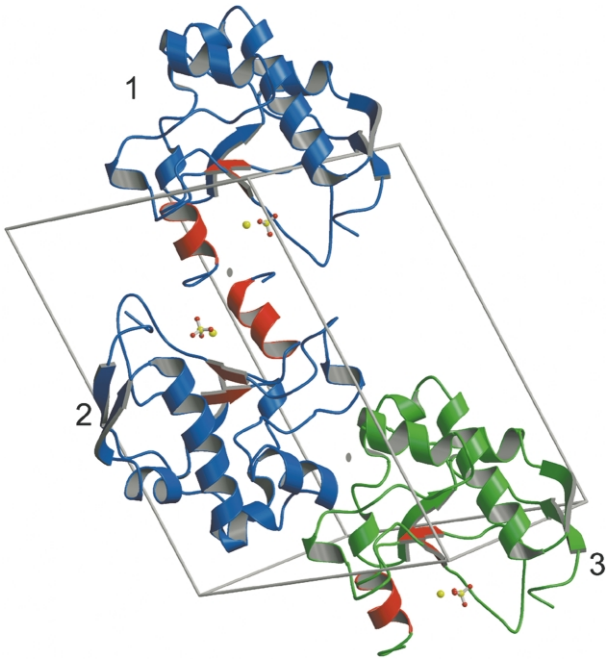

Figure 1. Three molecules of nuclease domain of ColE7 packed in the *P*1 unit cell. The two nuclease ColE7 molecules (labeled 1 and 2) in one asymmetric unit are displayed in blue with only the H–N–H motif in red. Zinc ions (yellow sphere) and the bound phosphate ions in the center of the H–N–H motif are represented by ball-and-stick models. Molecule 3, displayed in green, is symmetrically identical with molecule 1; however, it makes different contacts with molecule 2. Molecules 1 and 2 are related by a pseudo 2-fold symmetry (position roughly labeled as an ellipsoid).

Comparison between the structures of the bound and the unbound nuclease ColE7

In order to find out if the structure of nuclease ColE7 remains the same before and after inhibitor Im7 binding, we compared the structures of the bound (PDB entry: 7cei) and the unbound nuclease ColE7. Before the comparison, the C^{α} backbones of the two molecules (1 and 2) in the P1 unit cell were superimposed, giving an rms difference of only 0.22 Å, indicating that the two structures

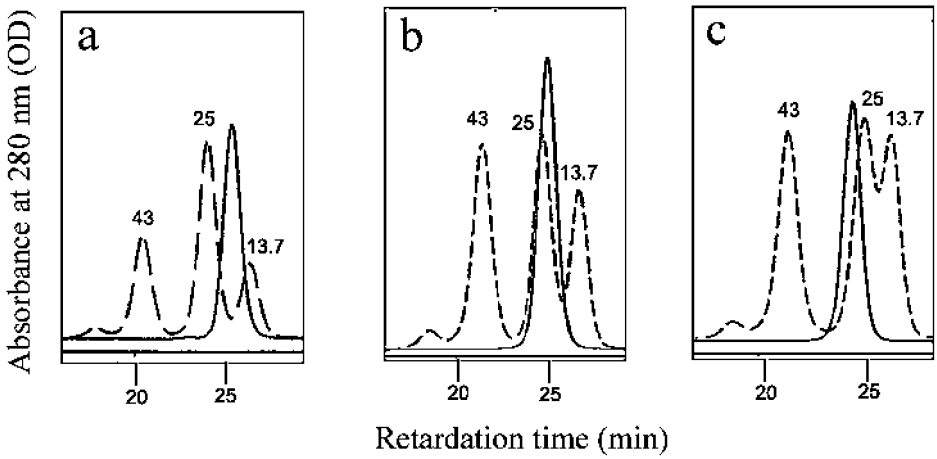

Figure 2. Analysis of the oligomeric state of nuclease ColE7 in different solutions by gel-filtration chromatography. Broken lines represent the elution profiles for molecular mass markers ovalbumin (43 kDa), chymotrypsinogen (25 kDa) and ribonuclease A (13.7 kDa). The molecular mass of the monomeric nuclease ColE7 is 15,731 Da and that of the dimeric nuclease ColE7 is 31,462 Da. Continuous lines represent the elution profiles of nuclease ColE7 in (a) 50 mM sodium phosphate (pH 7.0) and 150 mM NaCl; (b) 10 mM $ZnCl_2$ and 0.7 M CH_3COONH_4; and (c) 10 mM $ZnCl_2$ and 1.5 M CH_3COONH_4.

were almost identical. Therefore, only one of the unbound nuclease ColE7 structures (molecule 1) was used in the following comparison with the structure of the bound enzyme. The main difference between the bound and unbound nuclease ColE7 is that the temperature factors for the unbound enzyme were much lower than those of bound enzyme, an average of 27.0 Å^2 for unbound and 41.2 Å^2 for bound nuclease-ColE7. Particularly, the long loop in the H–N–H motif of bound nuclease ColE7 had higher temperature factors and became less flexible in the free form. This loop was in contact with the neighboring molecule (2) in the unbound structure. Residues Pro548, Ser550 and Gln551 in the loop were buried in the protein–protein interface. Thus, the lower flexibility of the H–N–H motif in the unbound nuclease ColE7 was likely the result of protein–protein contacts.

The superimposition of the C^{α} atoms of the bound and the unbound nuclease ColE7 gave an average rms difference of 0.62 Å. The core regions of the two structures were almost identical. The higher rms differences of 2.9–10.9 Å are positioned in a loop region from residues 466 to 472; however, this loop is located distantly from the Im7-recognition surface in ColE7. In the unbound structure, this loop makes contacts with the neighboring molecule (3) and, therefore, the larger shifts possibly resulted from crystal packing. So, in summary, the structures of the bound and unbound nuclease ColE7 are almost identical, and the binding of Im7 to ColE7 does not induce obvious backbone conformational changes in colicin.

Protein–protein interfaces

The protein interfaces between molecules 1 and 2, and between molecules 2 and 3 were examined closely to determine if they resulted from crystal packing. We found that the accessible surface area buried in the interface between molecules 1 and 2 was 558.5 Å^2 and between molecules 2 and 3 it was 595.9 Å^2 for each molecule. These values are in the lower range as compared to the values of 368–4746 Å^2 of buried interface in 32 homodimers.[46] The interface between molecules 1 and 2 contained 52% polar atoms and 48% non-polar atoms, so the interface had a similar chemical character as compared to the accessible surface of an average protein.[47] There were four hydrogen bonds in the interface (molecules 1 to 2): Asp571 (OD1) to Gln551 (NE2), ILE572 (O) to Gln574 (NH1), His573 (ND1) to Arg574 (O), and His573 (CE1) to Arg574 (O). Several hydrophobic residues, including Pro548, Ile570 and Ile572, were involved in the interactions. Two H–N–H motifs were located close to the interface and the three histidine residues directly bound to the zinc ion, His544, His569 and His573, were buried in the interface. On the basis of these data, it was difficult to make a conclusive statement concerning whether protein–protein interactions between molecules 1 and 2 is genuine or the result of crystal packing.

The interface between molecules 2 and 3 was filled with more charged and polar residues. There are seven hydrogen bonds in the interface (molecule 2 to molecule 3): Asn462 (OD1) to Gln532 (NE2), Asn462 (ND2) to Gln532 (OE1), Lys511 (O) to Gln532 (N), Phe513 (O) to Thr531 (N), Thr531 (N) to Phe513 (O), Gln532 (N) to Lys511 (O) and Gln532 (OE1) to Lys511 (NZ). Only a single hydrophobic residue, Phe513, was identified among the 15 polar or charged residues in the interface. Therefore, the interactions between molecules 2 and 3 were likely generated from crystal packing.

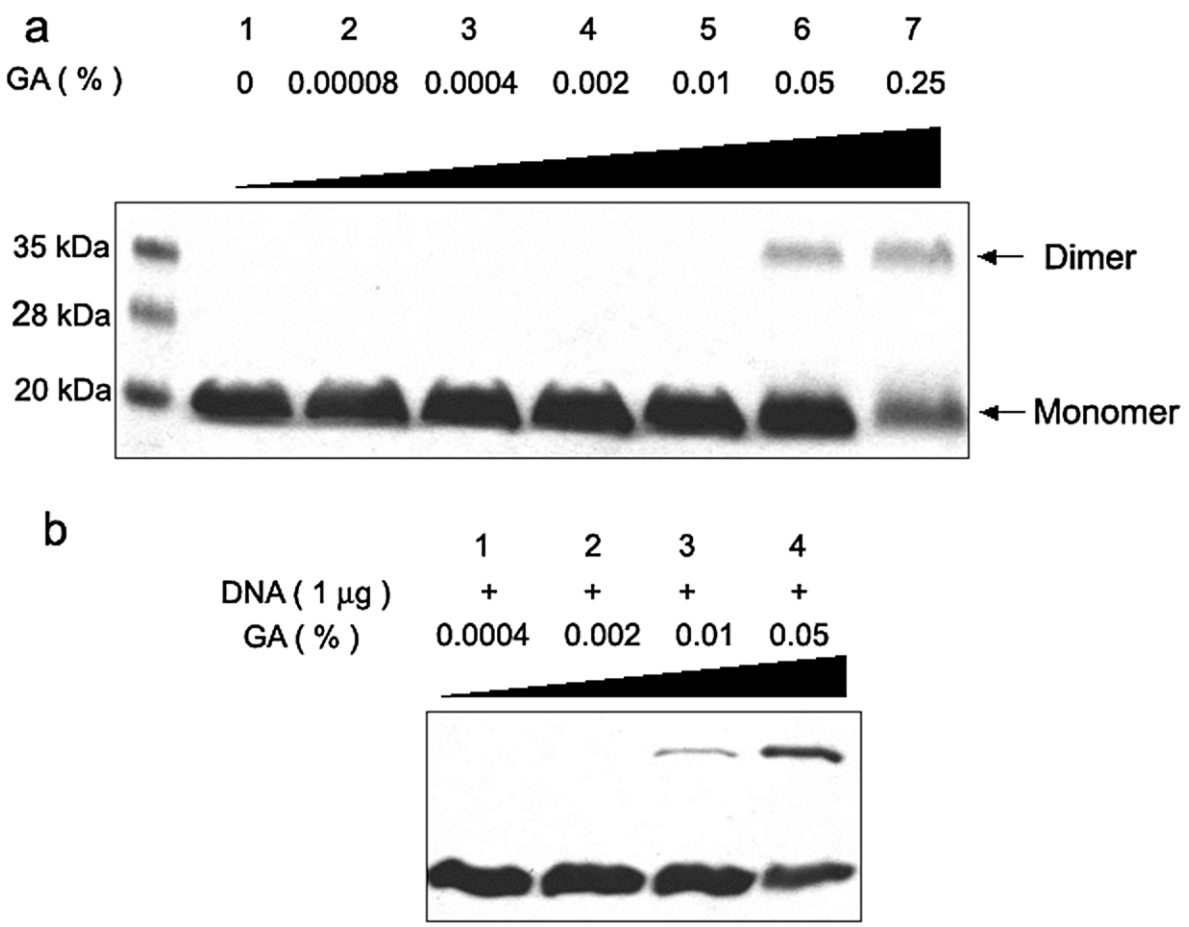

Figure 3. Glutaraldehyde (GA) cross-linking analysis of nuclease ColE7. (a) Nuclease ColE7 (1 μg) was cross-linked by >0.05% glutaraldehyde. (b) Nuclease ColE7 (1 μg) was cross-linked by less glutaraldehyde (>0.01%) in the presence of 27 bp double-stranded DNA (1 μg).

Biochemical characterization of monomer and dimer

It was unexpected that nuclease ColE7 could form a dimer, because a previous study using analytical ultracentrifugation had shown that nuclease-ColE9 exists only as monomers.[48] In order to clarify the oligomeric state of nuclease ColE7, we further analyzed nuclease ColE7 by gel-filtration chromatography and biochemical cross-linking methods. From gel-filtration experiments, we found that nuclease ColE7 was indeed a monomer in a low-salt buffer (150 mM NaCl) in the absence of metal ions (Figure 2(a)); however, nuclease ColE7 shifted to a dimer in buffers with higher concentrations of salt in the presence of 10 mM $ZnCl_2$ (see Figures 2(b) and 3(c)). Nuclease ColE7 samples in similar conditions were assayed by analytical ultracentrifugation (data not shown); however, only monomeric protein was detected.

Cross-linking experiments further identified the dimeric conformation of nuclease ColE7 in a low-salt buffer. Figure 3 shows that nuclease ColE7 was cross-linked by >0.05% (v/v) glutaraldehyde. In the presence of a 27 bp double-stranded DNA, nuclease ColE7 was cross-linked by less glutaraldehyde (>0.01%). This result indicated that nuclease ColE7 formed dimers in solution, and the presence of double-stranded DNA promoted dimer formation.

Discussion

Hydrolysis of a phosphodiester bond requires three general function entities in the active site: a general base to activate a nucleophile; a Lewis acid to stabilize the phosphoanion transition state; and a general acid to protonate the leaving group.[3] Therefore, some endonucleases share no sequence homology or overall structural resemblance, but have a similar active site. One example is the similarity observed in a small region between the homing endonuclease I-*Ppo* I, Serratia nuclease, phage T4 endonuclease VII (Endo VII) and nuclease ColE7/ColE9.[49–51] I-*Ppo* I is a His–Cys box homing endonuclease that site-specifically cleaves a 14 bp pseudo-palindromic homing site to yield 3′ overhangs of four nucleotides.[52] Serratia nuclease is similar to nuclease ColE7, in that it is a non-specific endonuclease that cleaves both single-stranded and double-stranded DNA and RNA without apparent base preference.[53] Endo VII is more different in

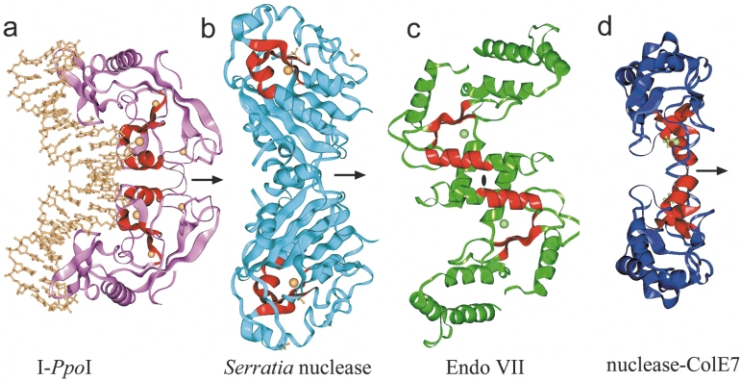

Figure 4. Structural comparison of dimeric nuclease ColE7 with other dimeric nucleases. Ribbon models for the crystal structures of (a) I-*Ppo* I in complex with DNA;[57] (b) Serratia nuclease;[55] (c) phage T4 Endo VII;[60] and (d) nuclease ColE7. The common ββα-fold of the active sites are displayed in red. The H–N–H motifs in nuclease ColE7 are arranged similarly as compared to the ββα-fold of active sites in I-*Ppo* I.

that it recognizes a variety of branched and single-base mismatched DNA and cleaves Holliday junctions, cruciform DNA, Y-junctions, single-stranded overhangs, and abasic sites.[54] Although these nucleases are different in many aspects, their active-site regions all consist of two β-strands and one α-helix linked by a divalent metal ion. The overall crystal structures of these nucleases are displayed in Figure 4 with the common ββα-metal fold colored in red.

As well as the structural similarity in the active-site region, I-*Ppo* I, Serratia nuclease and Endo VII are all dimeric enzymes. The crystal structure of I-*Ppo* I in complex with DNA demonstrates at the atomic level how this enzyme works.[55–57] I-*Ppo* I has an elongated fold with a concave surface to bind and cleave DNA at the minor groove of its homing site. The two active sites in the dimeric I-*Ppo* I are about 20 Å apart, therefore in order to place the two farthest away scissile phosphate groups near the active sites, the DNA is bent by approximately 50°. The two active sites in dimeric Endo VII are separated by ~25 Å and arranged in a different relative orientation, as compared to those in I-*Ppo* I. This difference is likely due to the fact that Endo VII recognizes branched but not typical double-stranded DNA substrates, as demonstrated in the structural model constructed for the Endo VII in complex with a Holiday junction.[58] The two active sites in Serratia nuclease, however, are distant from the dimeric interface, and it has been shown that a dimeric state is not essential for the catalytic function of Serratia nuclease.[59]

The distance between the two H–N–H motifs in the dimeric nuclease ColE7 (molecules 1 and 2) is 14.7 Å (between two zinc ions), which is close to the distance of 18.7 Å (between two magnesium ions) in I-*Ppo* I. More worthy of note is the fact that the relative orientation of the two H–N–H motifs in the dimeric nuclease ColE7 are arranged in a similar manner as compared to the two ββα-folds of active sites in I-*Ppo* I (see Figure 4). The dimeric nuclease ColE7 has an elongated crescent shape, with the two H–N–H motifs located close to the interface. Superimposing the two H–N–H motifs of dimeric nuclease ColE7 onto the active sites of I-*Ppo* I/DNA (see Figure 5(a)) shows that the ββα folds in the two enzymes match well with an rms deviation of 3.8 Å (for the C^{α} atoms in the ββα motif). Without any further adjustment, after the superimposition at the active-site regions, the nuclease ColE7 fits to the DNA snugly with the α-helices (α2) binding at the two major grooves and the zinc ions located close to the phosphate backbone at the central minor groove. Although the structures of the two enzymes are different, nuclease ColE7 has an overall concave shape similar to that of I-*Ppo* I. An electrostatic surface plot of dimeric nuclease ColE7 further shows that the concave surface is more basic, especially at the region close to the active sites, indicating that this surface is appropriate for DNA binding (see Figure 6). Gel-filtration and cross-linking experiments confirm that nuclease ColE7 forms dimers in solution, and that dimeric conformation increases in the presence of double-stranded DNA. Moreover, a molecular ratio of two nuclease ColE9 molecules bound to one 12 bp double-stranded DNA molecule has been characterized earlier in a fluorescence study.[40] Therefore, all these results support the possibility that nuclease ColE7 may form a dimer upon binding double-stranded DNA.

However, the dimeric interface of 558 Å is in the lower range for stable dimer formation.[48] This could explain the result that the nuclease ColE7 is not a permanent dimer, but exists as a dimer only in some conditions. In fact, H–N–H endonucleases show great versatility in substrate specificity. Some H–N–H nucleases, such as colicins, are non-specific nucleases and some are site-specific endonucleases but cleave DNA differently, such as I-*Hmu* I,[20–22] I-*Hmu* II[21,22] and I-*Two* I,[31] which cleave only one strand of their double-stranded DNA targets; but I-*Cmoe* I[24] and I-*Tev* III[23] make two nicks in DNA. Therefore the H–N–H family enzymes likely cleave nucleic acids either as monomers or dimers depending on substrate requirement. It is possible that some H–N–H family proteins may bind and cleave double-stranded DNA as homodimers, especially in cases where they need to make two nicks. However, the possibility is not excluded that some of the H–N–H proteins, including ColE7, may bind and cleave nucleic acids, especially single-stranded DNA or RNA, as monomers.

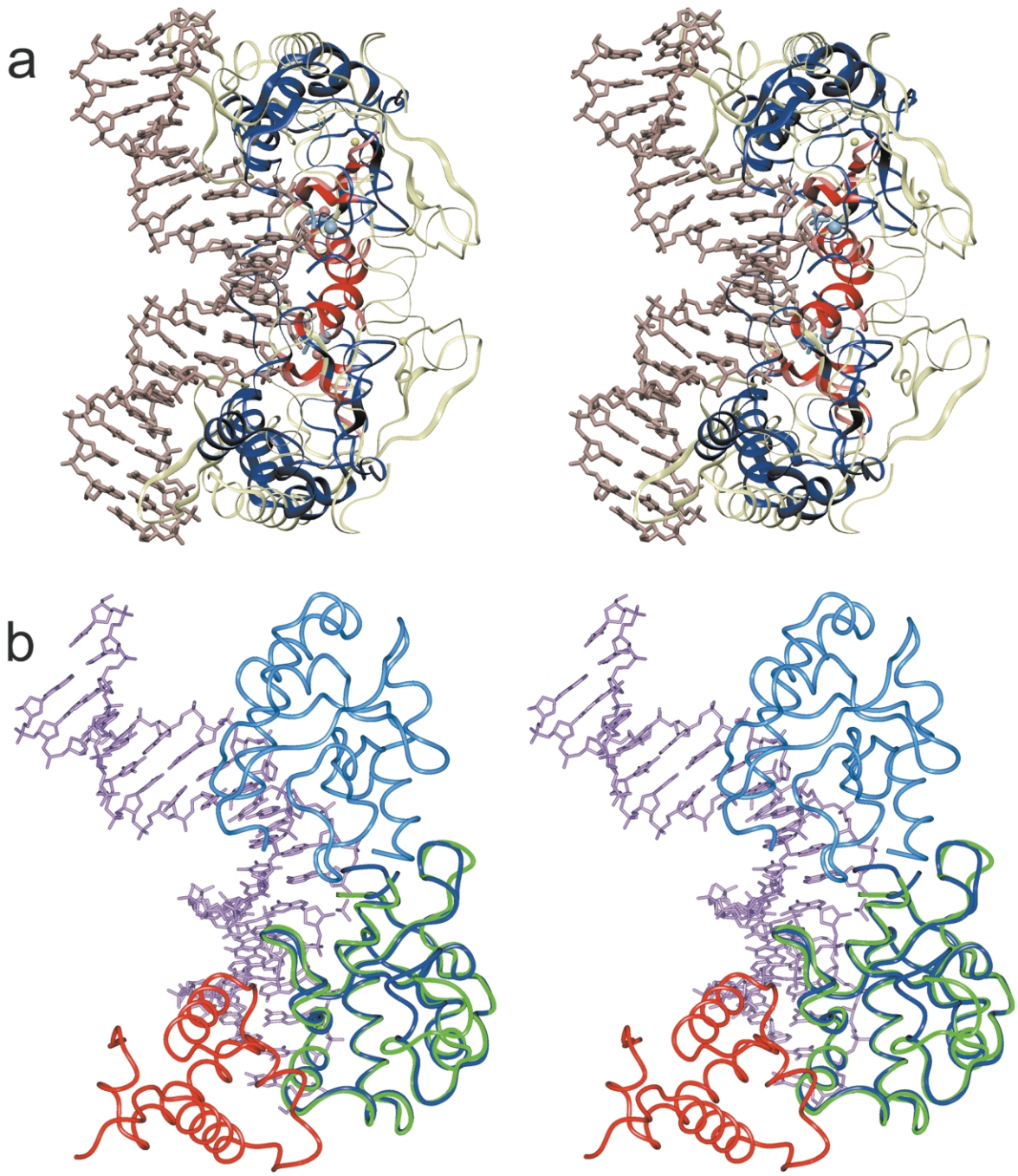

Figure 5. (a) Stereo views of the superimposition of the crystal structures of I-*Ppo* I/DNA and nuclease ColE7. Only the active sites (the ββα folds) were used for the fitting. The nuclease ColE7 is colored in blue, with the H–N–H motifs colored in red and the zinc ions colored in blue. I-*Ppo* I is colored in light yellow, with the ββα-Mg^{2+} active site colored in pink. The dimeric nuclease ColE7 shows a similar concave structure with two α-helices binding at the major grooves and two zinc ions located close to the phosphate backbonds in the minor groove. (b) Stereo views of the superimposition of the structural model of nuclease ColE7/DNA and the crystal structure of nuclease ColE7/Im7 (PDB entry 7cei). The DNA is tilted as compared to the view in (a). Im7 overlaps with DNA, therefore it is likely that Im7 blocks the DNA substrate-binding site thus inhibiting the nuclease activity of ColE7.

An intriguing yet unanswered question was how Im7 inhibits the nuclease activity of ColE7. In the crystal structure of the ColE7/Im7 complex, the putative nuclease active site in H–N–H motif is not blocked by Im7. It had been suggested that either Im7 induces conformational change or Im7 blocks the DNA-binding site so that ColE7 loses its activity upon binding to Im7.[36,45] Our analysis shows that the crystal structure of the unbound nuclease ColE7 is almost identical with the bound one, indicating that Im7 does not induce conformational changes in ColE7. By superimposing the nuclease-ColE7/DNA structural model onto the crystal structure of nuclease-ColE7/Im7 (Figure

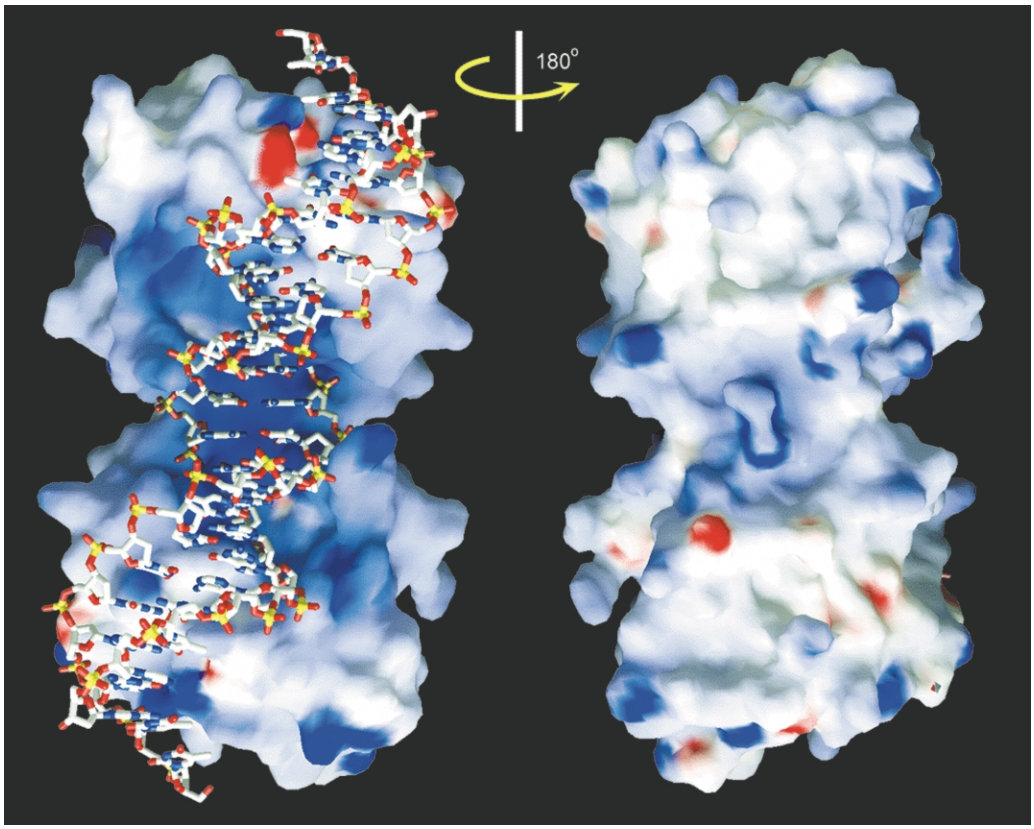

Figure 6. The electrostatic molecular surface of dimeric nuclease ColE7 produced with GRASP.[62] Blue and red correspond to positively and negatively charged areas, respectively. The concave surface of the dimeric nuclease ColE7 facing DNA (left panel) is more basic than the opposite side of the molecule.

5(b)), it can be seen that Im7 overlaps with DNA. Our study supports the theory that Im7 blocks the DNA-binding site in ColE7, resulting in the inhibition of its nuclease activity. The crystal structure of nuclease ColE7 in complex with DNA substrates is eagerly awaited to further confirm the hypothesis proposed here.

Experimental Procedures

Protein over expression and purification

The expression vector pQE70 (Qiagen) containing a C-terminal His_6 affinity tag was used for the over-expression of nuclease ColE7 and Im7 proteins. *E. coli* M15 (*nals*, *strs*, *rif s*, *lac*2, *ara*2, *gal*2, *mtl*2, *recA*1, *uvr*1) was used as the bacterial host for the expression vector. PCR amplification and cloning of the nuclease domain *ceiE7* and *ceiE7* gene fragments into the pQE expression system are described elsewhere.[38] Cells were cultured at 37 °C until the $A_{600\ nm}$ reached 0.6 in LB medium. IPTG was then added to a final concentration of 1 mM, to induce the expression of the nuclease ColE7/Im7 complex, for four hours. Crude cell extracts were loaded onto a Ni-NTA resin affinity column (Qiagen, Germany). The bound protein was then eluted by an imidazole gradient solution from 75 mM to 500 mM. The eluent was dialyzed in 20 mM glycine–HCl buffer (pH 3.0) overnight to denature the protein complex. The resulting protein solution was loaded onto a Sepharose-SP column (HiTrap SP, Pharmacia) equilibrated in 20 mM glycine–HCl buffer (pH 3.0). The nuclease ColE7 was eluted by a NaCl gradient (0–2.0 M). Im7 was further eluted by 20 mM sodium phosphate buffer at pH 7.0. The eluent containing nuclease ColE7 was dialyzed in 20 mM sodium phosphate buffer (pH 7.0) and then applied to a heparin column (HiTrap SP, Pharmacia). The free nuclease ColE7 (containing residues 444–576 without a His-tag) was then eluted by a NaCl gradient (0–1.0 M) and concentrated to 10 mg/ml in 50 mM Tris–HCl (pH 7.5).

Gel-filtration and glutaraldehyde cross-linking

The nuclease ColE7 protein samples were dialyzed against different buffers, including (a) 50 mM sodium phosphate (pH 7.0) and 150 mM NaCl; (b) 10 mM $ZnCl_2$ and 0.7 M CH_3COONH_4 (close to the crystallization condition); (c) 10 mM $ZnCl_2$ and 1.5 M CH_3COONH_4. Aliquots (100–200 μg) of pure nuclease ColE7 were then applied to a Superdex-75 gel-filtration column (HR, Pharmacia). Protein markers of ovalbumin (43 kDa), chymotrypsinogen (25 kDa) and ribonuclease A (13.7 kDa) were applied to the columns in the same buffers.

Purified nuclease ColE7 (1 μg) was cross-linked at room temperature by 0–0.25% glutaraldehyde, in the absence or presence of 27 bp double-stranded DNA (1 μg), in a total of 5 μl of 50 mM Tris–HCl (pH 8.0) for ten minutes at 37 °C. The reaction solution was quenched with 1 μl of stop solution (10% (v/v) β-mercaptoethanol, 50 mM Tris–HCl (pH 8.0), 10 mM EDTA). The products

Table 1. X-ray diffraction statistics for the nuclease ColE7

A. *Data collection and processing*	
Space group	*P*1
Cell dimensions	
a, *b*, *c* (Å)	37.7, 39.2, 53.2
α, β, γ (°)	96.8, 103.9, 91.4
Resolution (Å)	2.1
Observed reflections	53,594
Unique reflections	16,335
Completeness, all data (%)	95.0
Completeness, last shell (%)	92.7
(Resolution range, Å)	(2.18–2.10)
R_{merge} all data (%)[a]	4.4
R_{merge} last shell (%)	17.3
B. *Refinement*	
Resolution range (Å)	50.0–2.1
Reflections	16,016
Non-hydrogen atoms	
Protein	2116
Solvent molecules	209
R-factor (%)[b]	18.5
R_{free} (%)[c]	24.3
$I/\sigma(I)$, all data (50.0–2.1 Å)	22.5
$I/\sigma(I)$, last shell (2.18–2.10 Å)	7.0
C. *Model quality*	
rms deviations in	
Bond lengths (Å)	0.017
Bond angles (°)	1.655
Average *B*-factor (Å^2)	
Protein atoms	27.0
Solvent atoms	34.3
Ramachandran plot (%)	
Most favored	88.3
Additionally allowed	10.8
Generously allowed	0.9
Disallowed	0

[a]
$$R_{\text{merge}} = \frac{\sum_h \sum_i |I_h, i - \langle Ih \rangle|}{\sum_h \sum_i I_h, i},$$

where $\langle I_h \rangle$ is the mean intensity of the i observations for a given reflection h.

[b]
$$R\text{-factor} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|},$$

where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor amplitudes, respectively.

[c] R_{free} was calculated using a random set of 10% of observations that were omitted during refinement.

were resolved by SDS-PAGE (20% (w/v) polyacrylamide) and visualized by staining with Coomassie brilliant blue.

Crystallization and X-ray data collection

Crystals of nuclease ColE7 were obtained by the hanging-drop, vapor-diffusion method. Drops of a solution containing 10 mg/ml of nuclease ColE7, 5 mM $ZnCl_2$, 0.35 M ammonium acetate, 6.35% (w/v) PEG4000, 0.025 M Tris–HCl (pH 7.5) were set up against a reservoir of 10 mM $ZnCl_2$, 0.7 M ammonium acetate and 12.7% PEG4000. Plates of crystal appeared within a week at room temperature. X-ray diffraction intensities were collected at −150 °C from an R-AXIS-II imaging plate equipped with a double-mirror focusing system mounted on a rotating anode X-ray generator (MSC Co., USA). The nuclease ColE7 crystallized in a *P*1 unit cell and diffracted at least to 2.1 Å resolution. Diffraction data were processed with the HKL package[60] and the diffraction statistics are listed in Table 1.

Structure determination and refinement

The crystal structure of nuclease ColE7 was determined by molecular replacement using the program CNS.[61] Two nuclease ColE7 molecules were expected in one asymmetry unit of the *P*1 cell that gave a Matthew's coefficient of 2.6 Å^3/Da with the corresponding solvent content of 53.1% (v/v). The bound structure of the nuclease ColE7 without Im7 (PDB accession number 7cei) was used as the searching model. All reflection data above 2σ in the resolution range of 8–4 Å were subjected for rotational search using a direct rotation search function. The two highest peaks in the rotational search had correlation coefficients of 26.4 and 25.5, respectively. Subsequently, the two rotation solutions were used for a fast translation search with the data in the same resolution range, and the top two molecules gave a correlation coefficient of 54.4. Fast rigid-body refinement using data above 2σ between 50 Å and 2.5 Å gave an *R*-factor of 40.4%.

Structure refinement of nuclease ColE7 was carried out by the program CNS[61] with 10% of data selected randomly and set aside for the calculation of *R*-free values. Manual rebuilding of the model was performed using TURBO-FRODO,† alternating with torsional angle-simulated annealing, standard positional refinement, individual isotropic *B*-value refinement, calculation of electron density maps and automatic solvent molecules placement in CNS. The two nuclease ColE7 molecules in the asymmetry unit were numbered as A446-A576 and B446-B576, respectively. The final refinement statistics are listed in Table 1.

Protein Data Bank accession code

The atomic coordinates for the two nuclease ColE7 molecules have been deposited at the RCSB Protein Data Bank with accession code: 1M08.

Acknowledgements

This work was supported by research grants from Academia Sinica and the National Science Council of the Republic of China to H. S. Yuan (NSC90-2320-B001-024).

References

1. Kovall, R. A. & Matthews, B. W. (1999). Type II restriction endonucleases: structural, functional and evolutionary relationships. *Curr. Opin. Chem. Biol.* **3**, 578–583.

† http://afmb.cnrs-mrs.fr/TURBO_FRODO/turbo.html

2. Aggarwal, A. (1995). Structure and function of restriction endonucleases. *Curr. Opin. Struct. Biol.* **1995**, 11–19.
3. Pingoud, A. & Jeltsch, A. (1997). Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur. J. Biochem.* **246**, 1–22.
4. Chevalier, B. S. & Stoddard, B. L. (2001). Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucl. Acids Res.* **29**, 3757–3774.
5. Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J. & Mohr, G. (1997). A bacteria group II intron encoding reverse transcriptase, maturase and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new generic information within the intron. *Genes Dev.* **11**, 2910–2924.
6. Berger, J. M., Gamblin, S. J., Harrison, S. C. & Wang, J. C. (1996). Structure and mechanism of DNA topoisomerase II. *Nature*, **379**, 225–232.
7. Yang, W. & Steitz, T. A. (1995). Crystal structure of the site-specific recombinase γδ resolvase complexed with a 34 bp cleavage site. *Cell*, **82**, 193–207.
8. Guo, F., Gopaul, D. N. & Van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, **389**, 40–46.
9. Subramanya, H. S., Arciszewska, L. K., Baker, R. A., Bird, L. E., Sherratt, D. J. & Wigley, D. B. (1997). Crystal structure of the site-specific recombinase, XerD. *EMBO J.* **16**, 5178–5187.
10. Brennan, R. G. & Matthew, B. W. (1989). The helix-turn-helix DNA binding motif. *J. Biol. Chem.* **264**, 1903–1906.
11. Freemont, P. S., Lane, A. N. & Sanderson, M. R. (1991). Structural aspects of protein–DNA recognition. *Biochem. J.* **278**, 1–23.
12. Steitz, T. A. (1990). Structural studies of protein–nucleic acid interaction: the sources of sequence-specific binding. *Quart. Rev. Biophys.* **23**, 205–279.
13. Anderson, J. E. (1993). Restriction endonucleases and modification methylases. *Curr. Opin. Struct. Biol.* **3**, 24–30.
14. Schaller, K. & Nomura, M. (1976). Colicin E2 is a DNA endonuclease. *Proc. Natl Acad. Sci. USA*, **68**, 3989–3993.
15. Chak, K.-F., Kuo, W.-S., Lu, F.-M. & James, R. (1991). Cloning and characterization of the ColE7 plasmid. *J. Gen. Microbiol.* **137**, 91–100.
16. Toba, M., Masaki, H. & Ohta, T. (1988). Colicin E8, a DNase which indicates an evolutionary relationship between colicins E2 and E3. *J. Bacteriol.* **170**, 3237–3242.
17. Wallis, R., Moore, G. R., Kleanthous, C. & James, R. (1992). Molecular analysis of the protein–protein interaction between the E9 immunity protein and colicin E9. *Eur. J. Biochem.* **210**, 923–930.
18. Chak, K.-F., Hsieh, S.-Y., Liao, C.-C. & Kan, L.-S. (1998). Change of thermal stability of colicin E7 triggerred by acidic pH suggests the existence of unfolded intermediate during the membrane-translocation phase. *Proteins: Struct. Funct. Genet.* **32**, 17–25.
19. Liao, C.-C., Hsia, K.-C., Liu, Y.-W., Liang, P. H., Yuan, H. S. & Chak, K.-F. (2001). Processing of DNase domain during translocation of Colicin E7 across the membrane of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **284**, 556–562.
20. Goodrich-Blair, H., Scarlato, V., Gott, J. M., Xu, M.-Q. & Shub, D. A. (1990). A self-splicing group I intro in the DNA polymerase gene of *Bacillus subtilis* bacteriophage SPO1. *Cell*, **63**, 417–424.
21. Goodrich-Blair, H. & Shub, D. A. (1996). Beyond homing: competition between intron endonucleases confers a selective advantage on flanking genetic markers. *Cell*, **84**, 211–221.
22. Goodrich-Blair, H. (1994). The DNA polymerase genes of several HMU-bacteriophages have similar group I introns with highly divergent open reading frames. *Nucl. Acids Res.* **22**, 3715–3721.
23. Eddy, S. R. & Gold, L. (1991). The phage T4 *nrd* B intron: a deletion mutant of a version found in the wild. *Genes Dev.* **5**, 1032–1041.
24. Drouin, M., Lucas, P., Otis, C., Lemieux, C. & Turmel, M. (2000). Biochemical characterization of I-C*moe* I reveals that this H–N–H homing endonuclease shares functional similarities with H–N–H colicins. *Nucl. Acids Res.* **28**, 4566–4572.
25. Foley, S., Bruttin, A. & Brussow, H. (2000). Widespread distribution of a group I intron and its three deletion derivatives in the lysin gene of *Streptococcus thermophilius* bacteriophages. *J. Virol.* **74**, 611–618.
26. Lazarevic, V., Soldo, B., Dusterhoft, A., Hilbert, H., Mauel, C. & Karamata, D. (1998). Introns and intein coding sequence in the ribonucleotide reductase genes of *Bacillus substitls* temperate bacteriophage SPβ. *Proc. Natl Acad. Sci. USA*, **95**, 1692–1697.
27. Ferat, J. L. & Michel, F. (1993). Group II self-splicing introns in bacteria. *Nature*, **364**, 358–361.
28. Kuck, U. (1989). The intron of a plasmid gene from a green alga contains an open reading frame for a reverse transcriptase-like enzyme. *Mol. Gen. Genet.* **218**, 257–265.
29. Zimmerly, S., Guo, H., Perlman, P. S. & Lambowitz, A. M. (1995). Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell*, **82**, 545–554.
30. Zimmerly, S., Guo, H., Eskes, R., Yang, J., Perlman, P. S. & Lambowitz, A. M. (1995). A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell*, **83**, 529–538.
31. Landthaler, M., Begley, U., Lau, N. C. & Shub, D. A. (2002). Two self-splicing group I introns in the ribonucleotide reductase large gene of *Staphylococcus aureus* phage Twort. *Nucl. Acids Res.* **30**, 1935–1943.
32. Sano, Y. & Kageyama, M. (1993). A novel transposon-like structure carries the genes for pyocin AP41, a *Pseudomonas aeruginosa* bacteriocin with a DNase domain homology to E2 group colicins. *Mol. Gen. Genet.* **237**, 161–170.
33. Sano, Y., Matsui, H., Kobayashi, M. & Kageyama, M. (1993). Molecular structures and functions of Pyocins S1 and S2 in *Pseudomonas aeruginosa*. *J. Bacteriol.* **175**, 2907–2916.
34. Hiom, K. & Sedgwick, S. G. (1991). Cloning and structural characterization of the mcrA locus of *Escherichia coli*. *J. Bacteriol.* **173**, 7368–7373.
35. Eisen, J. A. (1998). A phylogenomic study of the MutS family of proteins. *Nucl. Acids Res.* **26**, 4291–4300.
36. Ko, T.-P., Liao, C.-C., Ku, W.-Y., Chak, K.-F. & Yuan, H. S. (1999). The crystal structure of the DNase domain of colicin E7 in complex with its inhibitor Im7 protein. *Structure*, **7**, 91–102.
37. Kleanthous, C., Kuhlmann, U. C., Pommer, A. J., Ferguson, N., Radford, S. E., Moore, G. R. *et al.*

(1999). Structural and mechanistic basis of immunity toward endonuclease colicins. *Nature Struct. Biol.* **6**, 243–252.
38. Sui, M.-J., Tsai, L.-C., Hsia, K.-C., Doudeva, L.-G., Chak, K.-F. & Yuan, H. S. (2002). Metal ions and phosphate binding in the HNH motif: crystal structures of the nuclease domain of ColE7/Im7 in complex with a phosphate ion and different divalent metal ions. *Protein Sci.* In press.
39. Grishin, N. V. (2001). Treble clef finger—a functionally diverse zinc-binding structural motif. *Nucl. Acids Res.* **29**, 1703–1714.
40. Pommer, A. J., Cal, S., Keeble, A. H., Walker, D., Evans, S. J., Kuhlmann, U. C. *et al.* (2001). Mechanism and cleavage specificity of the H–N–H endonuclease colicin E9. *J. Mol. Biol.* **314**, 735–749.
41. Ku, W.-Y., Liu, Y.-W., Hsu, Y.-C., Liao, C.-C., Liang, P.-H., Yuan, H. S. & Chak, K.-F. (2002). The zinc ion in the HNH motif of the endonuclease domain of colicin E7 is not required for DNA binding but is essential for DNA hydrolysis. *Nucl. Acids Res.* **30**, 1670–1678.
42. Keeble, A. H., Hemmings, A. M., James, R., Moore, G. R. & Kleanthous, C. (2002). Multistep binding of transition metals to the H–N–H endonuclease toxin colicin E9. *Biochemistry*, **41**, 10234–10244.
43. Pommer, A. J., Kuhlmann, U. C., Cooper, A., Hemmings, A. M., Moore, G. R., James, R. & Kleanthous, C. (1999). Homing in on the role of transition metals in the HNH motif of colicin endonuclease. *J. Biol. Chem.* **274**, 27153–27160.
44. Chak, K.-F., Safo, M. K., Ku, W.-Y., Hsieh, S.-Y. & Yuan, H. S. (1996). The crystal structure of the ImmE7 protein suggests a possible colicin-interacting surface. *Proc. Natl Acad. Sci. USA*, **93**, 6437–6442.
45. Kleanthous, C. & Walker, D. (2001). Immunity proteins: enzyme inhibitors that avoid the active site. *Trends Biochem. Sci.* **26**, 624–631.
46. Jones, S. & Thornton, J. M. (1996). Principles of protein–protein interactions. *Proc. Natl Acad. Sci. USA*, **93**, 13–20.
47. Janin, J. & Chothia, C. (1990). The structure of protein–protein recognition sites. *J. Biol. Chem.* **265**, 16027–16030.
48. Pommer, A. J., Wallis, R., Moore, G. R., James, R. & Kleanthous, C. (1998). Enzymatic characterization of the nuclease domain from the bacterial toxin colicin E9 from *Escherichia coli*. *Biochem. J.* **334**, 387–392.
49. Friedhoff, P., Franke, I., Meiss, G., Wende, W., Krause, K. L. & Pingoud, A. (1999). A similar active site for non-specific and specific endonucleases. *Nature Struct. Biol.* **6**, 112–113.
50. 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.
51. Kuhlmann, U. C., Moore, G. R., James, R., Kleanthous, C. & Hemmings, A. M. (1999). Structural parsimony in endonuclease active site: should the number of homing endonuclease families be redefined? *FEBS Letters*, **463**, 1–2.
52. Ellison, E. L. (1993). Interaction of the intro-encoded mobility endonuclease I-P*po*I with its target site. *Mol. Cell. Biol.* **13**, 7351–7539.
53. Miller, M. D., Tanner, J., Alpaugh, M., Benedik, M. J. & Krause, K. L. (1994). 2.1 Angstrom structure of *Serratia* endonuclease suggests a mechanism for binding to double-stranded DNA. *Nature Struct. Biol.* **1**, 461–468.
54. Kemper, B. (1997). Branched DNA resolving enzymes. In DNA Damage and Repair (Nickoloff, J. A. & Hoekstra, M. F., eds), vol. 1, pp. 179–204, Humana Press Inc, Totowa, NJ.
55. 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-P*po*I. *Nature Struct. Biol.* **6**, 1096–1099.
56. 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-P*po*I. *Nature*, **394**, 96–101.
57. Galburt, E. A., Chadsey, M. S., Jurica, M. S., Chevalier, B. S., Erho, D., Tang, W. *et al.* (2000). Conformational changes and cleavage by the homing endonuclease I-P*po*I: acritical role for a leucine residue in the active site. *J. Mol. Biol.* **300**, 877–887.
58. Raaijmakers, H., Vix, O., Toro, I., Golz, S., Kemper, B. & Suck, D. (1999). X-ray structure of T4 endonuclease VII: a DNA junction resolvase with a novel fold and unusual domain-swapped dimer architecture. *EMBO J.* **18**, 1447–1458.
59. Franke, I., Meiss, G., Blecher, D., Gimadutdinow, O., Urbanke, C. & Pingoud, A. (1998). Genetic engineering, production and characterisation of monomeric variants of the dimeric *Serratia marcescens* endonuclease. *FEBS Letters*, **425**, 517–522.
60. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Enzymology (Carter, C. W. & Sweet, R. M., eds), vol. 276, pp. 307–326, Academic Press, New York.
61. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
62. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.

Edited by K. Morikawa

(Received 16 July 2002; received in revised form 23 September 2002; accepted 24 September 2002)