JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Conserved Asparagine in the HNH Motif Serves an Important Structural Role in Metal Finger Endonucleases

Hsinchin Huang and Hanna S. Yuan*

Institute of Molecular Biology Academia Sinica, Taipei Taiwan, ROC

The HNH motif is a small nucleic acid binding and cleavage module, widespread in metal finger endonucleases in all life kingdoms. Here we studied a non-specific endonuclease, the nuclease domain of ColE7 (N-ColE7), to decipher the role of the conserved asparagine and histidine residues in the HNH motif. We found, using fluorescence resonance energy transfer (FRET) assays, that the DNA hydrolysis activity of H545 N-ColE7 mutants was completely abolished while activities of N560 and H573 mutants varied from 6.9% to 83.2% of the wild-type activity. The crystal structures of three N-ColE7 mutants in complex with the inhibitor Im7, N560A–Im7, N560D–Im7 and H573A–Im7, were determined at a resolution of 1.9 Å to 2.2 Å. H573 is responsible for metal ion binding in the wild-type protein, as the zinc ion is still partially associated in the structure of H573A, suggesting that H573 plays a supportive role in metal binding. Both N560A and N560D contain a disordered loop in the HNH motif due to the disruption of the hydrogen bond network surrounding the side-chain of residue 560, and as a result, the imidazole ring of the general base residue H545 is tilted slightly and the scissile phosphate is shifted, leading to the large reductions in hydrolysis activities. These results suggest that the highly conserved asparagine in the HNH motif, in general, plays a structural role in constraining the loop in the metal finger structure and keeping the general base histidine and scissile phosphate in the correct position for DNA hydrolysis.

© 2007 Published by Elsevier Ltd.

Keywords: ββα-metal motif; metal finger motif; DNase; DNA hydrolysis; crystal structure

*Corresponding author

Introduction

The HNH motif is a small nucleic acid binding and cleavage module, containing about 30 to 40 amino acid residues and bound with a single divalent metal ion. This motif was first identified based on the sequence similarity among several intron-encoded homing endonucleases and bacterial toxins.[1,2] HNH stands for the three most conserved histidine and asparagine residues in the degenerate motif with less than a dozen of consensus amino acids. To date, more than 1000 proteins containing an HNH motif have been identified in all life kingdoms of the prokaryotes, archaea and eukaryotes (see the HNH endonuclease domain in Pfam Protein Families database[3]).

The biggest subgroup of HNH proteins with known functions are the site-specific group I and group II homing endonucleases, including PetD,[4] Avi,[5] Cpc,[5] I-TevIII,[6] I-HmuI,[7,8] yosQ,[9] ORF253,[10] I-CmoeI,[11] I-TwoI[12] and I-BasI.[13] A group of bacterial toxins with non-specific endonuclease activity represent another prototype sub-group of HNH proteins, including the colicins, ColE7[14] and ColE9,[15] and the pyocins, S1 and S2.[16,17] The HNH motif has also been identified in restriction or repair enzymes, such as McrA,[18] mtMSH,[19] KpnI,[20] MnlI,[21] Hin4II[22] and HphI.[23] In eukaryotes, most HNH proteins have unknown functions, but the HNH motifs are often found in proteins bearing helicase or reverse transcriptase domains.[24–26] Figure 1 shows the sequence alignment of a number of representa-
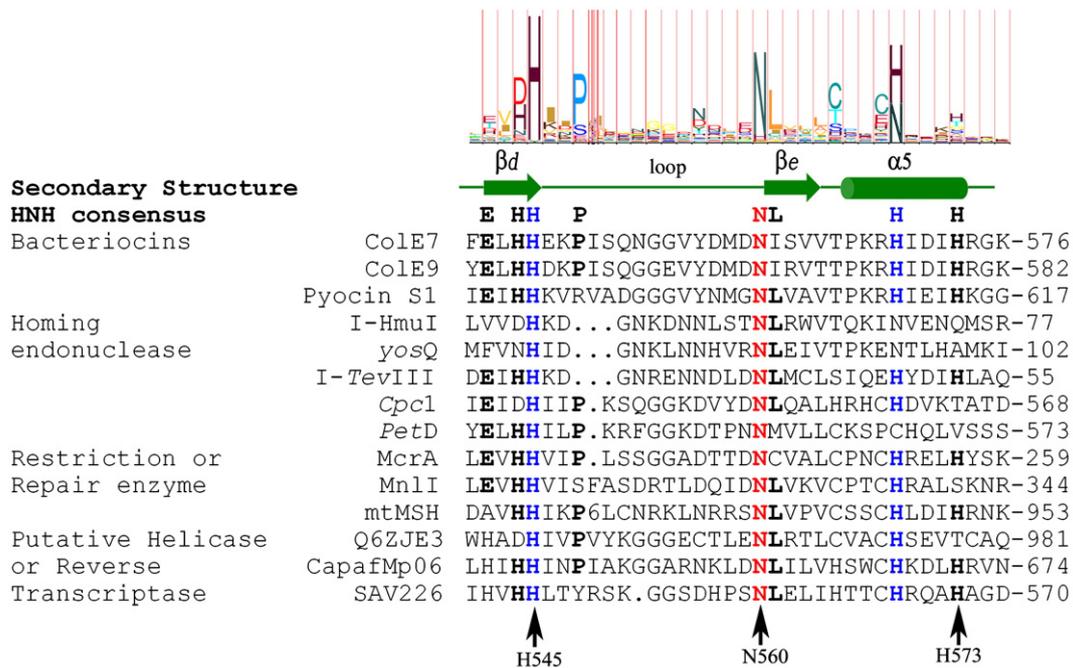
Figure 1. Sequence alignment of HNH proteins. The amino acid sequences of the HNH motifs in several representative proteins are aligned and these proteins are classified into four sub-groups: bacteriacin, homing endonuclease, restriction or repair enzyme and putative helicase or reverse transcriptase. Relative amino acid probabilities at each position are listed on the top of the aligned sequences, based on the analysis of 922 HNH protein sequences in the Pfam database.[48] The secondary structure derived from the crystal structure of ColE7[28] and the consensus sequence of the HNH motif are also listed on the top of the aligned sequence. The three most conserved histidine (blue) and asparagine (red) residues standing for the HNH are highlighted in color. The mutation points in this study, at H545, N560 and H573 in ColE7, are marked at the bottom line of the aligned sequences.

tive HNH proteins, classified into four sub-groups of bacterial toxin, homing endonuclease, restriction or repair enzyme and putative helicase or reverse transcriptase.

The three-dimensional structure of an HNH motif was first revealed in the crystal structure analysis of the nuclease domain of ColE7 (referred to as N-ColE7 hereinafter) in complex with its inhibitor Im7.[27] Subsequently, the crystal structures of several HNH proteins in complex with DNA duplexes have been determined, including N-ColE7 in complex with an 8 bp,[28] a 12 bp[29] and an 18 bp DNA,[30] the nuclease domain of ColE9 in complex with an 8 bp DNA[31] and the homing endonuclease I-HmuI in complex with a cleaved 25 bp DNA.[32] The HNH motifs in these structures all fold into a similar conformation, containing two antiparallel β-strands, one α-helix and a divalent metal ion bound in the center (see Figure 2). This type of "ββα-metal fold" has been observed in a number of endonucleases, including the homing endonuclease I-PpoI,[33] *Serratia* nuclease,[34] phage T4 Endo VII,[35] Vvn[36] and CAD[37], and is a general structure of the endonuclease active site[38–40].

Different metal ions are associated with the HNH motif as the cofactor for DNA hydrolysis. Those proteins containing an H-N-*N* rather than the HNH consensus sequence are often activated by Mg^{2+}, such as I-HmuI. The Mg^{2+} in the HNH/N motif of I-HmuI is bound to six coordinated groups in an octahedral geometry: D74, N96, the cleaved DNA 5′-phosphate group, the 3′-hydroxyl leaving group, and two water molecules (PDB entry 1U3E). On the contrary, some HNH proteins are bound to a Zn^{2+}, such as ColE7 in which the zinc ion is bound to three histidine residues (H544, H569 and H573) and the scissile phosphate in a tetrahedral geometry[41] (see Figure 2). Since metal ions are always bound to the scissile phosphate, it has been concluded that the divalent metal ion in the HNH motif serves three roles during hydrolysis: polarization of the P–O bond for nucleophilic attack, stabilization of the phospho anion transition state and stabilization of the cleaved product.[29] H544 and H569 in ColE7 geometrically match well with D74 and N96 in I-HmuI, all responsible for metal ion binding.

However, the most conserved residues in the HNH motif of ColE7 are H545 and N560, but not the metal ion binding residues of H544 and H569. H545 is strictly conserved in the HNH motif because it functions as the general base, which activates a water molecule for DNA hydrolysis. N560 is the second most conserved residue in the HNH motif, located in the loop between two β-strands and far away from the endonuclease active site and the protein–DNA interface. It was therefore intriguing why N560 is so conserved in all of the HNH proteins, since it can be neither a catalytic residue, nor responsible for DNA interactions. The side-chain of N560 forms a hydrogen bond network with the main-chain atoms in the loop region in the HNH motif. The side-chain of the general base H545 makes
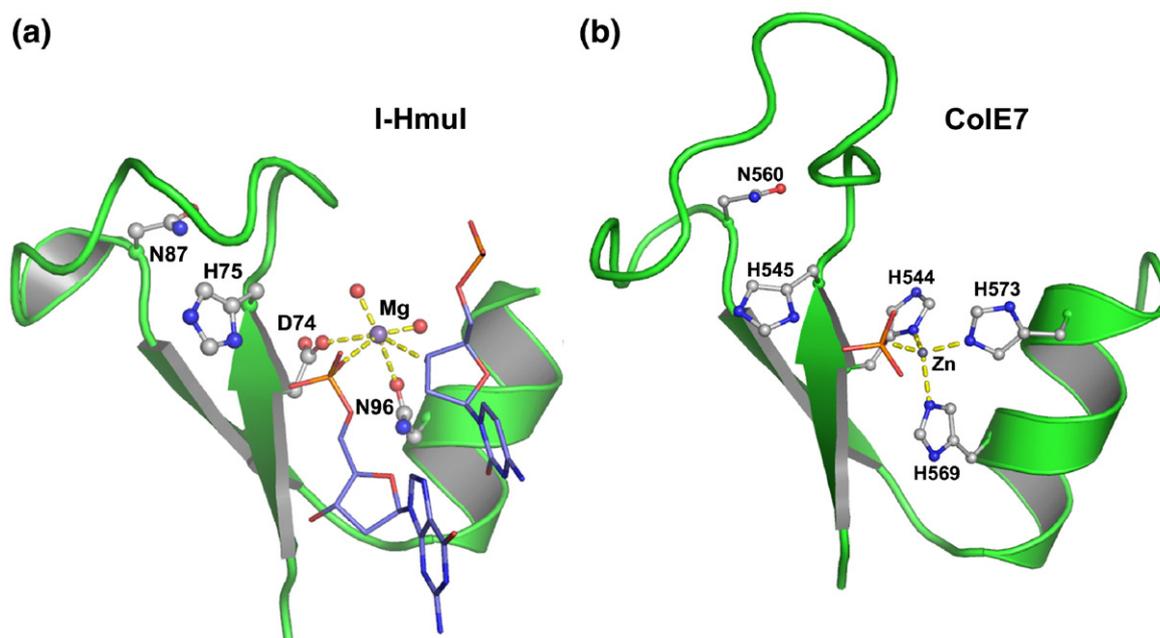
Figure 2. Structures of HNH motif in I-HmuI and ColE7. (a) The HNH motif in I-HmuI contains two β-strands and one α-helix folded in a ββα-metal structure with a Mg^{2+} bound in an octahedral geometry in the center. The most conserved histidine (H75) and asparagine (N87) are located at the loop region between the two β-strands. The magnesium ion is bound to D74, N96, the cleaved DNA 5′-phosphate group, the 3′-hydroxyl leaving group, and two water molecules (PDB entry 1U3E). Two nucleotide fragments of the 25 bp DNA are shown as a stick model. (b) The HNH motif in ColE7 folds into a ββα-metal structure with a Zn^{2+} bound in a tetrahedral geometry in the active site. The zinc ion is bound to three histidine side-chains, H544, H569 and H573, and a phosphate ion (PDB entry 1MZ8). The most conserved H545 and N560 are located at the loop region.

a hydrogen bond to the main-chain carbonyl atom of V555 in the loop. Therefore it was suspected that the N560 is involved in constraining the loop structure, and in turn holding the general base residue in an appropriate position for hydrolysis.[41,42]

Here we have used the ColE7 as the model system to decipher the role of conserved residues in the HNH motif in DNA binding and hydrolysis. We mutated several conserved residues, including H545 (general base), H573 (metal ion binding), and N560 (function unclear), and measured the endonuclease activity of these mutants by the fluorescence resonance energy transfer (FRET) method. Moreover, the crystal structures of three mutants in complex with inhibitor Im7, H573A–Im7, N560A–Im7 and N560D–Im7, were determined at a resolution of 1.9 Å to 2.2 Å. Upon consideration of both the biochemical and structural results, we suggest that the conserved asparagine plays important structural roles in the HNH motif not only in restraining the loop structure but also in maintaining the precise location of the general base histidine.

Results

Mutations at H545, N560 and H573 interrupt ColE7 endonuclease activity

We had constructed five N-ColE7 mutants, H545A, H545E, H545Q, H573E and H573A, and found that the mutation of the general base residue H545 or the metal-ion-binding residue H573 abolished or reduced the endonuclease activity of N-ColE7 by plasmid digestion assays.[29] Besides these five mutants, here we constructed a further five more N560 and H573 mutants: N560A, N560D, N560H, H573N and H573Q. Altogether, the ten single-point mutants and the wild-type N-ColE7 were expressed and purified, respectively. The overall structure of each mutant was then assayed by CD, which gave a similar spectrum as compared to that of the wild-type N-ColE7, indicating that the overall fold of each mutant was unchanged (data not shown). The melting points for thermal denaturation of these proteins were further analyzed by CD (see Figure 3). The wild-type N-ColE7, and the H545 mutants and N560 mutants, had a melting point of ~81 °C, suggesting that the mutation did not change the overall structure nor the thermal stability of the protein. However, all of the H573 mutants had lower melting points in the range of 73 °C–77 °C. Since H573 is responsible for metal ion binding, the lower melting points are likely resulted from the reduced metal binding ability (see Discussion).

To obtain a quantitative estimate, the endonuclease activities of each N-ColE7 mutant was measured by the FRET method using a 16 nucleotide single-stranded DNA, bound to a fluorophore at the 5′ end and a quencher at the 3′ end, as the substrate. The cleavage of the DNA by N-ColE7 abolishes the
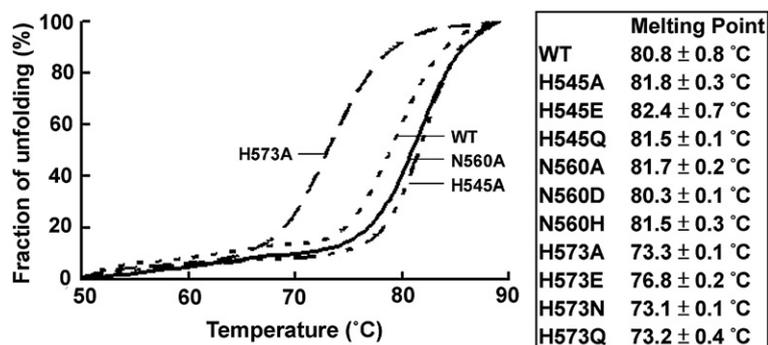
Figure 3. The melting points of the wild-type and mutated N-ColE7 analyzed by circular dichroism. Thermal denaturation experiments were performed at a wavelength of 222 nm with each protein sample first dialyzed with $ZnCl_2$ and then concentrated to 0.1 mg/ml in 10 mM Tris–HCl (pH 8.0). The melting points of wild-type and mutated N-ColE7 derived from three measurements are listed in the right panel.

fluorescence quenching and results in the increased fluorescence emission intensity so that the kinetics data of K_m and k_{cat} can be derived. As expected, H545A and H545E had no measurable endonuclease activity whereas H545Q had low (11.38%) activity as compared to the wild-type enzyme (see Table 1). H573A and H573E showed reduced cleavage efficiencies of 30.92% and 83.24% of the wild-type enzyme, respectively. The K_m values of these two H573 mutants were close to that of the wild-type enzyme, indicating that the different catalytic activities did not result from reduced binding affinity to the DNA substrate. The three N560 mutants also cleaved DNA with reduced activities, 23.76% for N560A, 6.93% for N560D and 27.08% for N560H, although they still bound DNA with a similar affinity to the wild-type protein. In summary, the mutations at H545 and N560 did not change the protein's overall structure or thermal stability but did abolish or greatly reduce the endonuclease activity. The mutations at H573 lowered the thermal stability of the protein and reduced the endonuclease activity of N-ColE7.

Crystal structure of ColE7 mutants- N560A and N560D

To determine the structural basis of the reduced enzyme activity, we attempted to crystallize the three N560 mutants in complex with Im7, N560A–Im7, N560D–Im7 and N560H–Im7, for detailed structural determination. Two of the mutant complexes, N560A–Im7 and N560D–Im7, were successfully crystallized in unit cells isomorphous to the wild-type phosphate-bound N-ColE7–Im7 complex.[41] The crystal structures of N560A–Im7 and N560D–Im7 were refined to a resolution of 2.2 Å and 2.0 Å, respectively, with reasonable refinement and geometry statistics (see Table 2). The omit difference maps displayed in Figure 4 show clearly the side-chain density of A560 and D560.

The electron density for the loop between the two β-strands in the HNH motif was poorly defined in the crystal structures of N560A–Im7 and N560D–Im7. Therefore the loop between residues 547 to 554 was not modeled in the structures. As a result, a

Table 1. The endonuclease activities of the wild-type and mutated nuclease domain of ColE7 measured by FRET methods

	K_m (nM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1}s^{-1})
WT	74.9±3.2	0.0045±0.0010	60,600 (100%)
H545A	–	–	≈ 0 (0%)
H545E	–	–	≈ 0 (0%)
H545Q	53.3±2.2	0.0003±0.0001	6200 (11.38%)
N560A	52.3±6.0	0.0008±0.0002	14,400 (23.76%)
N560D	95.3±4.8	0.0004±0.0002	4200 (6.93%)
N560H	20.7±8.2	0.0003±0.0000	16,400 (27.08 %)
H573A	62.4±4.4	0.0012±0.0001	18,700 (30.92%)
H573E	24.8±5.0	0.0013±0.0001	50,400 (83.24%)

Table 2. Diffraction and refinement statistics for N-ColE7 mutants in complex with Im7

	N560A-Im7	N560D-Im7	H573A-Im7
Data collection and procession			
Space group	$P2_12_12$	$P2_12_12$	$I222$
Cell dimensions (Å)	$a=119.2$	$a=120.0$	$a=63.1$
	$b=62.7$	$b=62.9$	$b=74.4$
	$c=74.8$	$c=74.2$	$c=119.6$
Resolution (Å)	2.2	2.0	1.9
Observed reflections	194,023	871,615	138,124
Unique reflections	29,109	35,931	21,837
Completeness- all data (%)	99.9 (40.0–2.2 Å)	92.5 (50.0–2.0 Å)	97.4 (50.0–1.9 Å)
Completeness- last shell (%)	100.0 (2.28–2.20 Å)	89.6 (2.10–2.03 Å)	86.1 (1.98–1.91 Å)
R_{sym}- all data (%)[a]	7.1	5.1	3.9
R_{sym}- last shell (%)[a]	38.5	35.4	26.9
$I/\sigma(l)$, all data	27.1	21.4	39.2
$I/\sigma(l)$, last shell	5.4	3.0	4.8
Refinement			
Resolution range (Å)	40.0–2.2	50.0–2.0	50.0–1.9
Reflections	28,439	34,191	22,607
Non-hydrogen atoms			
Protein	3363	3327	1659
Solvent molecules	335	322	195
R-factor (%)[b]	20.1	20.6	21.1
R_{free} (%)	25.0	25.0	25.2
Model quality			
RMS deviations in			
Bond lengths (Å)	0.007	0.007	0.006
Bond angles (deg.)	1.12	1.22	1.08

[a] $R_{sym}=\sum_{hkl}\sum_i |I_i(hkl)-\langle I(hkl)\rangle| / I_i(hkl)$.
[b] R-factor$=\sum_{hkl} ||F_o(hkl)|-|F_c(hkl)|| / \sum_{hkl} F_o(hkl)$.

246 number of hydrogen bonds in the loop region were
247 lost in the mutants as compared to the wild-type N-
248 ColE7 (see Figure 5). In the wild-type enzyme, the
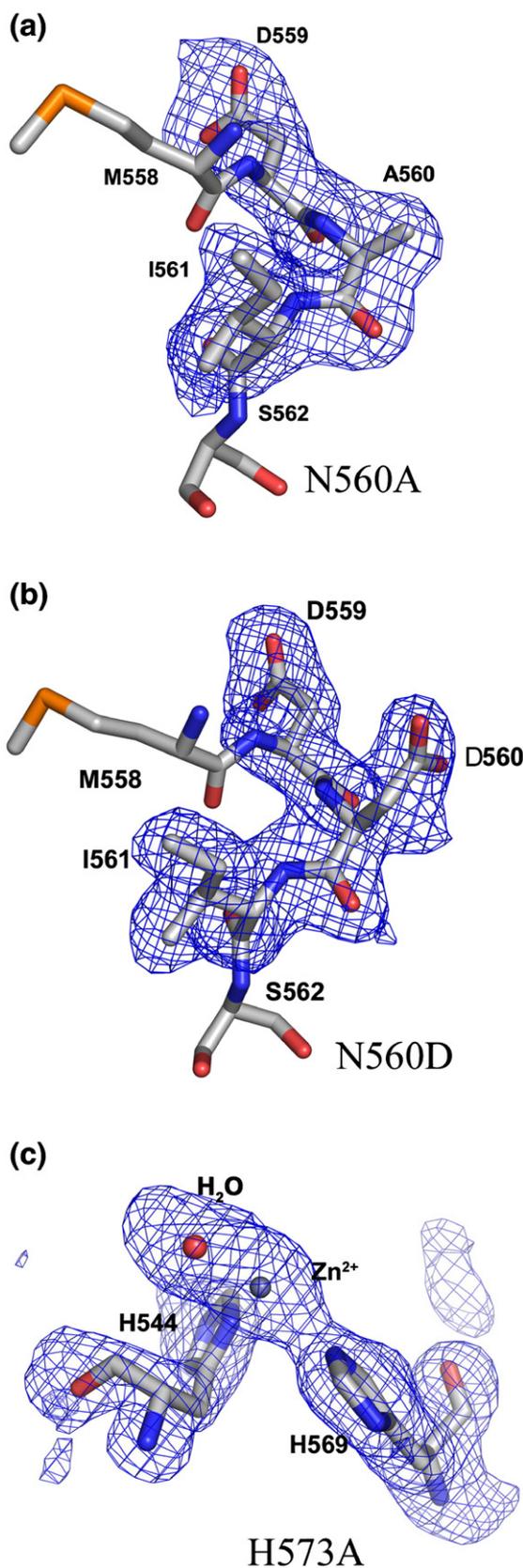
side-chain of N560 makes a hydrogen bond network 249
to the peptide backbone atoms in the loop, the $N^{\delta 2}$ 250
atom making two hydrogen bonds with the carbo- 251
nyl groups of G554 and K547, and the $O^{\delta 1}$ atom 252
making two hydrogen bonds with the amino groups 253
of K547 and E546. In the N560A–Im7 structure, the 254
side-chain of A560 cannot make any hydrogen bond 255
to stabilize the loop structure. In the N560D–Im7 256
structure, the side-chain of D560 swung out of the 257
loop region and did not make any hydrogen bonds 258
with the loop backbone atoms either. Because of the 259
disruption of the hydrogen bond network surround- 260
ing the side-chain of residue 560, half of the loop was 261
disordered and not seen in the N560A and N560D 262
mutants. 263

The general base residue H545 makes a hydrogen 264
bond to the carbonyl group of the loop residue 265
V555 in the wild-type N-ColE7.[41] It was expected 266
that this hydrogen bond would be disrupted by the 267
mutation at N560 due to the disordered loop 268
structure. However, to our surprise, this hydrogen 269
bond was maintained in the structures of N560A– 270
Im7 and N560D–Im7. Therefore, the imidazole ring 271
of H545 was still polarized by the carbonyl group 272
of Val555, and was probably still capable of 273
functioning as a general base to activate a water 274
molecule for the nucleophilic attack on the scissile 275
phosphate. 276

Crystal structure of ColE7 mutant H573A   277

To examine the structural basis for the reduced 278
endonuclease activity of H573 mutants we also 279
purified H573 mutants in complex with Im7 for 280
structural studies. All of the purified protein 281
complexes were first dialyzed with $ZnCl_2$ before 282
protein concentration and crystallization. Only 283
H573A–Im7 was successfully crystallized in the 284
*I*222 unit cell, isomorphous to that of the wild- 285
type zinc-bound N-ColE7.[27] The structure of 286
H573A–Im7 was refined to a resolution of 1.9 Å 287
with reasonable statistics as listed in Table 2. The 288
final model of H573A contained residues 449 to 289
572, without visible electron density after residue 290
A573. The omit difference map surrounding the 291
metal ion binding site is shown in Figure 4(c). A 292
zinc ion was bound to three ligands, H544 and 293
H569 and a water molecule, whereas the fourth 294
ligand was not present in the map. Compared to 295
the peak height of 9σ in the N560A–Im7 structure 296
and 12σ in the N560D–Im7 structure, the zinc ion 297
in the H573A–Im7 had a peak height of only 5σ 298

Figure 4. The omit difference (F_o-F_c) maps of N560A–
Im7, N560D–Im7 and H573A–Im7 at the mutation site. (a)
The N560A map contoured at 3σ shows the short side-
chain density of A560. (b) In N560D, the side-chain of
D560 swings out of the loop and points outwards (map
contoured at 3σ). (c) In H573A, the electron density for the
mutated residue A573 is not seen in the difference map
(contoured at 3σ). A zinc ion is located in the active site
and binds to H544, H569 and a water molecule.

(a) Wild type N560

(b) N560A

(c) N560D
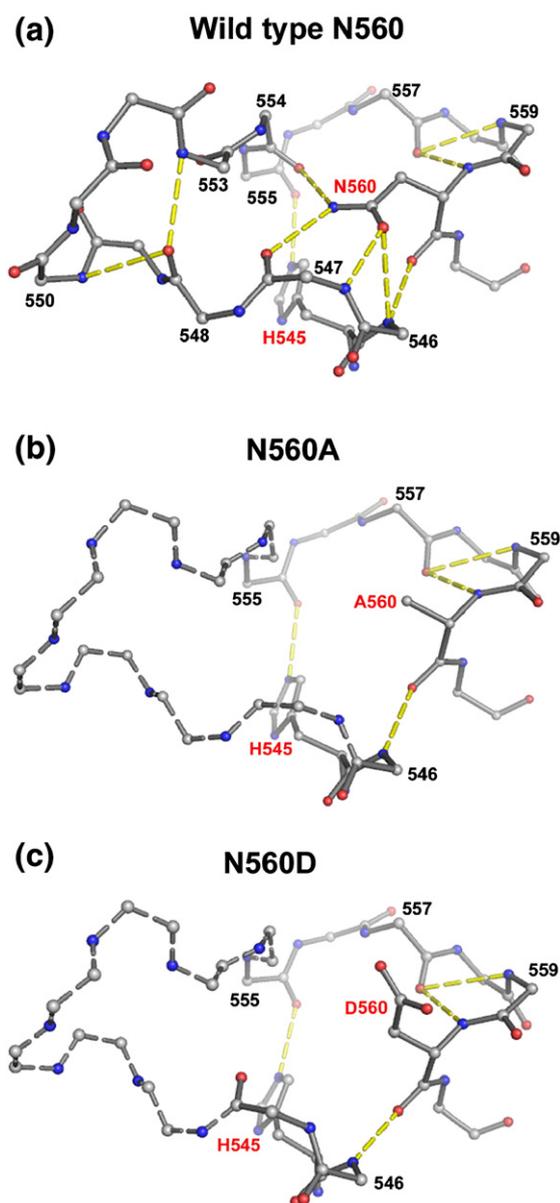
Figure 5. The hydrogen bond networks in the loop region of the HNH motif in N-ColE7. (a) In the wild-type N-ColE7–Im7, the side-chain of N560 makes four hydrogen bonds to the backbone atoms to constrain a well ordered loop structure. For clarity, only the side-chain atoms of H545 and N560 are shown with the loop backbone atoms. Hydrogen bonds are shown in yellow dotted lines. (b) In the N560A–Im7, the hydrogen bond network surrounding residue 560 is disrupted due to the replacement of N560 to Ala. As a result, half of the loop is disordered (as shown in broken grey sticks). (c) In the N560D–Im7, the side-chain of D560 swings out of the loop region and does not make hydrogen bonds to the backbone atoms in the loop. Half of the loop is disordered without a defined structure.

299 in the 2Fo–Fc Fourier map, indicating that with-
300 out the metal-binding residue of H573, the zinc
301 ion did still bind to the mutant but with a lower
302 occupancy.

Discussion   303

H573 is responsible but not crucial for metal ion binding in ColE7   304 305

H573, together with H544 and H569, is responsible for Zn^{2+} binding in the HNH motif of N-ColE7. The H573 mutants, H573A and H573E, have decreased endonuclease activities and melting points of 73.3($\pm$0.1) °C for H573A and 76.8($\pm$0.2) °C for H573E. The zinc-bound wild-type N-ColE7 has a melting point of 80.8($\pm$0.8) °C and the apo-N-ColE7 has a melting point of 74.3($\pm$0.7) °C.[29] Therefore, we had suspected that the zinc ion was not bound in H573A and H573E. We therefore further mutated H573 to Asn or Gln, which contains an amide group supposedly capable of metal ion binding. However, the melting points for H573N and H573Q were still close to that of the apo-enzyme, indicating that neither N573 nor Q573 bind zinc ions.   306–320

To find out if the zinc was still able to bind to the H573 mutants, we determined the crystal structure of H573A–Im7 complex. We found that the zinc ion was still bound in the HNH motif of the N560A mutant but the peak height was reduced from ~12σ (N560D) and ~9σ (N560A) to ~5σ (H573A), indicating that the zinc could bind but with lower occupancy. The zinc-binding capability of the mutant seems to be correlated to the enzyme activity. H573E has the highest melting point and also the highest enzyme activity among the H573 mutants. Therefore, we conclude that the divalent metal ion is required for the endonuclease activity, and that H573 plays a supportive role for metal ion binding in the HNH motif of N-ColE7. With fewer metal ions bound at the active site, H573 mutants are less active than the wild-type enzyme.   321–337

N560 plays a structural role in the HNH motif   338

In contrast to the H573 mutants, all of the N560 mutants, including N560A, N560D and N560H, had a melting point of ~81 °C, close to that of the wild-type N-ColE7, indicating that all of these mutants fold as well as the wild-type protein. However, all of the N560 mutants had reduced endonuclease activities, ranging from 6.93% to 27.08%, indicating that the conserved asparagine in the HNH motif plays an important role in DNA hydrolysis. Since they all fold well and bind to DNA with high affinity, it was intriguing why these mutants have low endonuclease activity. The crystal structures of N560A–Im7 and N560D–Im7 reveal a disordered loop in the HNH motif due to the disruption of the hydrogen-bond network surrounding the N560 side-chain. Therefore, N560 plays an important role in retention the loop structure in the HNH motif.   339–356

The loop is located remotely from the endonuclease active site or protein–DNA interface, hence it was difficult to explain how a disordered loop led to the decreased endonuclease activity. A previous   357–360

steady-state and pre-steady-state cleavage experiments demonstrated that the DNA hydrolysis reaction catalyzed by the non-specific *Serratia* nuclease was not limited by the product dissociation step but by the association step or the chemical cleavage of the phosphodiester bnd.[43] *Serratia* nuclease, similar to ColE7, is a member of ββα-metal nucleases as a result, ColE7 likely cleaves DNA in a way similar to that of *Serratia* nuclease. We therefore looked into the residues that are involved in the possible rate-limiting step of phosphate binding and cleavage in ColE7. We found that the general base residue H545 makes a hydrogen bond with the carbonyl group of V555 in the loop region in the wild-type N-ColE7, and this hydrogen bond is still retained in the N560A and N560D mutants. But a closer examination by superimposition of the HNH motif in the wild-type and N560 mutants shows that the imidazole ring of H545 is tilted, and as a result, the phosphate, which mimics the scissile phosphate in DNA, is shifted (see Figure 6). The $N^{\delta 1}$ atom in the imidazole ring is shifted by 0.63 Å in N560A and 0.75 Å in N560D mutant compared to that of wild-type protein. The average root-mean-square differences between wild-type N-ColE7 and N560A/N560D mutants are only 0.131/0.330 Å for main-chain atoms and 0.207/0.347 Å for side-chain atoms. Therefore, the displacement of the general base residue of ~0.6 Å to 0.7 Å is significant. Moreover, the distance between H545 $N^{\delta 1}$ atom to the phosphate oxygen atom is 2.41 Å in the wild-type enzyme, 2.90 Å in N560A mutant and 3.72 Å in the N560D mutant (see Figure 6). Therefore, the disruption of the endonuclease activities observed for N560A and N560D is possibly due to the misalignment of the scissile phosphate of the DNA substrate. The longer the $N^{\delta 1}$–O distances resulted in greater reduction of endonuclease activity: 100% for wild-type, 23.76% for N560A and only 6.93% for N560D of endonuclease activity.

It has been shown that the scissile phosphate of a specific DNA site has to be bound at a precise location close to the active site of restriction enzymes for efficient cleavage.[44] In those cases, where a non-specific DNA site is bound to a restriction enzyme, the DNA is not cleaved because the scissile phosphate is located too far away from the active site.[45] We also showed that N-ColE7 cleaves DNA with a preference for making nicks after thymine because the scissile phosphate moves closer to the endonuclease active site.[30] Another example from the study with the Klenow fragment of *Escherichia coli* DNA polymerase I have shown that the movement of the scissile bond by only 0.6 Å from the native position substantially reduces the exonuclease activity.[46] Here we show one more example of surrounding residues fine-tuning the geometry of a set of catalytic residues for precise substrate binding and hydrolysis. We conclude that the N560 plays a structural role to restrain the loop structure in the HNH motif and in turn it restrains the precise location and orientation of the general base H545 for DNA binding and efficient DNA hydrolysis.
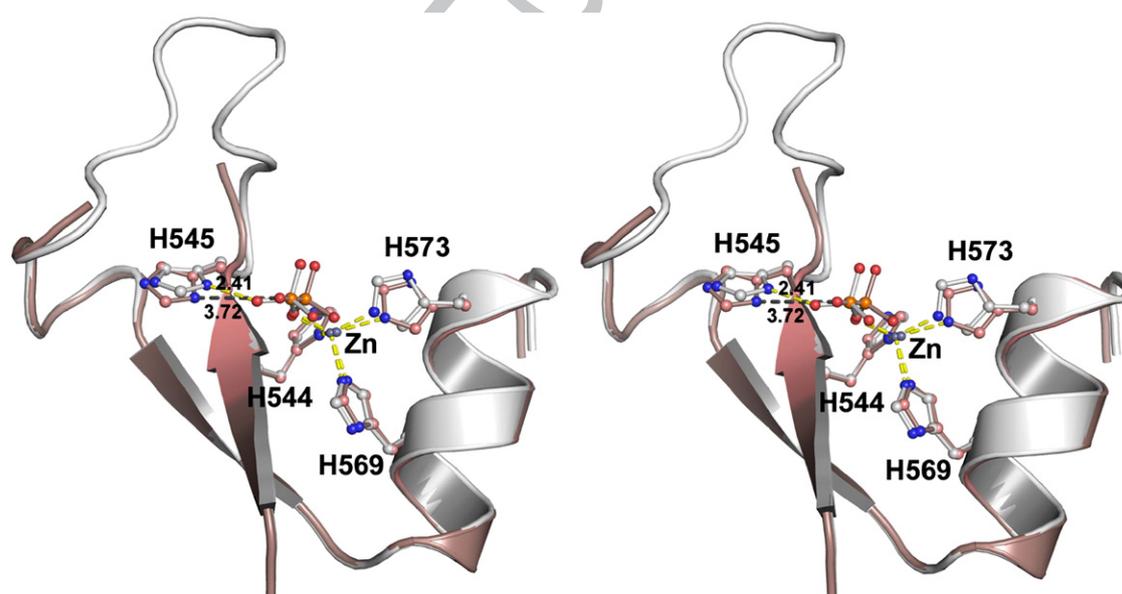
Figure 6. Stereo views of the superposition of the HNH motif in the wild-type N-ColE7 (in grey) and N560D mutant (in pink). Only the backbone atoms in the HNH motif were used for the least-squares fitting. Half of the loop is disordered in the N560D mutant, and as a result, the side-chain of H545 is tilted and the phosphate ion bound in the active site is shifted. The distance between H545 $N^{\delta 1}$ atom to the phosphate oxygen atom changes from 2.41 Å in the wild-type enzyme to 3.72 Å in the N560D mutant. This result indicates that the replacement of N560 in the HNH motif disturbs the loop structure and in turn changes the orientation of the imidazole ring of the general base histidine (H545), which is responsible for the reduced endonuclease activity.

Methods

Cloning, protein expression and purification

The expression vectors pQE-70 (QIAGEN), containing a 6-histidine affinity tag at the C terminus of the cloning site, were used to express wild-type N-ColE7 and Im7 heterodimer in *E. coli* M15 cells. The N-ColE7 mutants, H545A, H545E, H545Q, N560A, N560D, N560H, H573A, H573E, H573N and H573Q were constructed using the Quick-Change™ Site-Directed Mutagenesis Kit (Stratagene), and were expressed with Im7 and purified in the same manner as that of wide-type N-ColE7. The free-form of the wild-type and mutated N-ColE7 were then separated from Im7 and purified by methods described.[47] The molecular mass of each mutant was measured by mass spectroscopy and the measured molecular weight matched well with the calculated weights (data not shown).

CD measurements

Before melting point measurements, all the proteins were treated with Chelex 100 resin (Bio-Rad) and EDTA to remove any associated metal ions using the method described.[29] The metal-free N-ColE7 was then dialyzed against 1 mM $ZnCl_2$ (99.999%) in 10 mM Tris–HCl (pH 8.0) at 4 °C overnight. The metal-treated N-ColE7 was then dialyzed against 10 mM Tris–HCl buffer (pH 8.0) with four separate changes of buffer independently. The melting point of each zinc-bound protein sample was measured three times by circular dichroism. Thermal denaturation experiments were performed in 10 mm cuvettes at a wavelength of 222 nm on a Jasco J720 spectropolarimeter. The protein concentrations used for all measurements were 0.1 mg/ml in 10 mM Tris–HCl (pH8.0). The temperature was increased from 25 °C to 95 °C at a rate of 50 °C/h. The transition curve (T_m) was fitted to a two-state unfolding model.

Endonuclease activity measured by the FRET method

The 16 nucleotide single-stranded DNA substrate was labeled with fluorogenic material FAM (6-carboxyl-fluorescein) at the 5′ end and TAMRA (6-carboxyl-tetramethyl-rhodamine) at the 3′ end (Biotech, Taiwan): FAM-5′-CTGTCGCTACCTGTGG-3′-TAMRA. The wild-type or mutated zinc-bound N-ColE7 was then mixed with the annealed fluorogenic DNA in buffers of 10 mM Tris–HCl (pH8.0) and 1 mM $MgCl_2$ at 25 °C. With an excitation frequency of 515 nm, the increased fluorescence emission intensity at 486 nm, resulting from DNA cleavage, was measured on a fluorescence plate reader (PerkinElmer 1420 Victor2 multi-label counter) over a period of 200 s at 25 °C. The protein concentration used in the measurement was fixed at 10 nM with varied DNA concentrations of 7, 14, 35, 70, 140, 280 and 700 nM. The FRET measurement of each protein was repeated for at least three times and the resultant kinetic data of K_m and k_{cat} are summarized in Table 1.

Crystallization and data collection

Before crystallization, the purified N-ColE7 mutant complexes of N560A–Im7, N560D–Im7 and H573A–Im7 were dialyzed against 1 mM $ZnCl_2$ to make sure that the active site contained a Zn ion. Crystals of N560A–Im7 were grown by hanging drop vapor-diffusion method at 4 °C against a reservoir containing 20% (w/v) polyethylene glycol monomethyl ether 2000, 0.2 M ammonium sulfate, and 0.1 M sodium acetate trihydrate (pH 4.6). Crystals of N560D–Im7 and H573A–Im7 complexes were also grown by hanging drop vapor-diffusion method at room temperature with a reservoir of 20% (w/v) PEG3350 and 0.1 M di-ammonium hydrogen citrate for N560D–Im7; and 20% PEG3350 and 0.2 M di-ammonium hydrogen citrate for H573A–Im7.

All crystals were flash cooled in liquid nitrogen before data collection. The diffraction data of N560A–Im7 were collected by a R-AXIS-IV imaging plate using Micro-Max007 X-ray generator. Diffraction data of N560D–Im7 and H573A–Im7 were collected using synchrotron X-ray radiation by the ADSC Quantum-315 CCD detector at SPXF beamline BL13B1 at NSRRC (Taiwan, ROC), and by the Quantum 315 CCD detector at the BL12B2 in SPring-8, Japan, respectively. These mutant complex crystals diffracted X-ray to a resolution of 1.9 Å–2.2 Å. All the diffraction statistics are listed in Table 2.

Structural determination and refinement

N560A–Im7 and N560D–Im7 were crystallized in an isomorphous $P2_12_12$ unit cell of the phosphate-bound N-ColE7–Im7.[41] H573A–Im7 was crystallized in an isomorphous $I222$ unit cell of the wild-type zinc-bound N-ColE7–Im7.[27] Therefore these two structures (PDB entry of 1MZ8 and 7CEI) were used as the starting models for structural refinement by the program CNS. Metal ions and phosphate ions were added into the model before the last cycle of refinement. The final refinement statistics are listed in Table 2.

Protein Data Bank accession codes

Structural coordinates and diffraction structure factors have been deposited in the RCSB Protein Data Bank with ID codes of 2JBG for N560A–Im7, 2JAZ for N560D–Im7 and 2JB0 for H573A–Im7.

References

1. Shub, D. A., Goodrich-Blair, H. & Eddy, S. R. (1994). Amino acid sequence motif of group I intron endonucleases is conserved in open reading frames of group II introns. *Trends Biochem. Sci.* **19**, 402–404.
2. Gorbalenya, A. E. (1994). Self-splicing group I and group II introns encode homologous (putative)

DNA endonucleases of a new family. *Protein Sci.* **3**, 1117–1120.

3. Bateman, A., Barney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R. *et al.* (2002). The Pfam protein families database. *Nucl. Acids Res.* **30**, 276–280.

4. 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.

5. Ferat, J. L. & Michel, F. (1993). Group II self-splicing introns in bacteria. *Nature*, **364**, 358–361.

6. Eddy, S. R. & Gold, L. (1991). The phage T4 nrdB intron: a deletion mutant of a version found in the wild. *Genes Develop.* **5**, 1032–1041.

7. 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.

8. 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.

9. 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.

10. 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.

11. Drouin, M., Lucas, P., Otis, C., Lemieux, C. & Turmel, M. (2000). Biochemical characterization of I-*Cmoe*I reveals that this H-N-H homing endonuclease shares functional similarities with H-N-H colicins. *Nucl. Acids Res.* **28**, 4566–4572.

12. 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.

13. Landthaler, M. & Shub, D. A. (2003). The nicking homing endonuclease I-BasI is encoded by a group I intron in the DNA polymerase gene of the *Bacillus thuringiensis* phage Bastille. *Nucl. Acids Res.* **31**, 3071–3077.

14. 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.

15. 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.

16. 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.

17. 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.

18. Hiom, K. & Sedgwick, S. G. (1991). Cloning and structural characterization of the *mcr*A locus of *Escherichia coli*. *J. Bacteriol.* **173**, 7368–7373.

19. Pont-Kingdon, G. A., Okada, N. A., Macfarlane, J. L., Beagley, C. T., Wolstenholme, D. R., Cavalier-Smith, T. & Clark-Walker, G. D. (1995). A coral mitochondrial *mutS* gene. *Nature*, **375**, 109–111.

20. Saravanan, M., Bujnicki, J. M., Cymerman, I. A., Rao, D. N. & Nagaraja, V. (2004). Type II restriction endonuclease R.KpnI is a member of the HNH nuclease superfamily. *Nucl. Acids Res.* **32**, 6129–6135.

21. Kriukiene, E., Lubiene, J., Lagunavicius, A. & Lubys, A. (2005). MnlI- The member of H-N-H subtype of type IIS restriction endonucleases. *Biochem. Biophys. Acta*, **1751**, 194–204.

22. Azarinskas, A., Maneliene, Z. & Jakubauskas, A. (2006). Hin4II, a new prototype restriction endonuclease from *Haemophilus influenzae* RFL4: discovery, cloning and expression in *Escherichia coli*. *J. Biotechnol.* **123**, 288–296.

23. Cymerman, I. A., Obarska, A., Skowronek, K. J., Lubys, A. & Bujnicki, J. M. (2006). Identification of a new subfamily of HNH nucleases and experimental characterization of a representative member. HphI restriction endonuclease. *Proteins: Struct. Func. Bioinform.* **65**, 867–876.

24. Matsumoto, T., W. J. *et al.* (2005). The map-based sequence of the rice genome. *Nature*, **436**, 793–800.

25. Nosek, J., Novotna, M., Hlavatovicova, Z., Ussery, D. W., Fajkus, J. & Tomaska, L. (2004). Complete DNA sequence of the linear mitochondrial genome of the pathogenic yeast *Candida parapsilosis*. *Mol. Genet. Genom.* **272**, 173–180.

26. Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M. *et al.* (2001). Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl Acad. Sci. USA*, **98**, 12215–12220.

27. 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.

28. Hsia, K.-C., Chak, K.-F., Liang, P.-H., Cheng, Y.-S., Ku, W.-Y. & Yuan, H. S. (2004). DNA binding and degradation by the HNH endonuclease ColE7. *Structure*, **12**, 205–214.

29. Doudeva, L. G., Huang, H., Hsia, K.-C., Shi, Z., Li, C.-L., Shen, Y. & Yuan, H. S. (2006). Crystal structural analysis and metal-dependent stability and activity studies of the ColE7 endonuclease domain in complex with DNA/Zn^{2+} or inhibitor/Ni^{2+}. *Protein Sci.* **15**, 269–280.

30. Wang, Y.-T., Yang, W.-J., Li, C.-L., Doudeva, L. G. & Yuan, H. S. (2007). Structural basis for sequence-dependent DNA cleavage by nonspecific endonucleases. *Nucl. Acids Res.* In the press.

31. Mate, M. J. & Kleanthous, C. (2004). Structure-based analysis of the metal-dependence mechanism of H-N-H endonucleases. *J. Biol. Chem.* **279**, 34763–34769.

32. Shen, B. W., Landthaler, M., Shub, D. A. & Stoddard, B. L. (2004). DNA binding and cleavage by the HNH homing endonuclease I-HmuI. *J. Mol. Biol.* **342**, 43–56.

33. 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-*Ppo*I. *Nature*, **394**, 96–101.

34. 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.

35. 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.

36. Li, C.-L., Hor, L.-I., Chang, Z.-F., Tsai, L.-C., Yang, W.-Z. & Yuan, H. S. (2003). DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. *EMBO J.* **22**, 4014–4025.

37. Woo, E.-J., Kim, Y.-G., Kim, M.-S., Han, W.-D., Shin, S., Robinson, H. *et al.* (2004). Structural mechanism for inactivation and activation of CAD/DFF40 in the apoptotic pathway. *Mol. Cell*, **14**, 531–539.

38. 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.

39. 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.

40. Grishin, N. V. (2001). Treble clef finger- a functionally diverse zinc-binding structural motif. *Nucl. Acids Res.* **29**, 1703–1714.

41. 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 H-N-H motif: crystal structures of the nuclease domain of ColE7/Im7 in complex with a phosphate ion and different divalent metal ions. *Protein Sci.* **11**, 2947–2957.

42. Scholz, S. R., Korn, C., Bujnicki, J. M., Gimadutdinow, O., Pingoud, A. & Meiss, G. (2003). Experimental evidence for a ββα-metal-finger nuclease motif to represent the active site of the capase-activated DNase. *Biochemistry*, **42**, 9288–9294.

43. Friedhoff, P., Meiss, G., Kolmes, B., Pieper, U., Gimadutdinow, O., Urbanke, C. & Pingoud, A. (1996). Kinetic analysis of the cleavage of natural and synthetic substrates by the *Serratia nuclease*. *Eur. J. Biochem.* **241**, 572–580.

44. Pingoud, A. & Jeltsch, A. (2001). Structure and function of type II restriction endonucleases. *Nucl. Acids Res.* **29**, 3705–3727.

45. Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P. *et al.* (1993). The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.* **12**, 1781–1795.

46. Brautigam, C. & Steitz, T. A. (1998). Structural principles for the inhibition of the 3′-5′ exonuclease activity of *Escherichia coli* DNA polymerase I by phosphorothioates. *J. Mol. Biol.* **277**, 363–377.

47. Cheng, Y.-S., Hsia, K.-C., Doudeva, L. G., Chak, K.-F. & Yuan, H. S. (2002). The crystal structure of the nuclease domain of ColE7 suggests a mechanism for binding to double-stranded DNA by the H-N-H endonucleases. *J. Mol. Biol.* **324**, 227–236.

48. Schuster-Boeckler, B., Schultz, J. & Rahmann, S. (2004). HMM Logos for visualization of protein families. *BMC Bioinformat.* **5**, 7.

Edited by K. Morikawa

UNCORRECTED PROOF