5194–5207 *Nucleic Acids Research, 2015, Vol. 43, No. 10* *Published online 27 April 2015*
doi: 10.1093/nar/gkv348

Structure-based functional identification of *Helicobacter pylori* HP0268 as a nuclease with both DNA nicking and RNase activities

Ki-Young Lee[1], Kyu-Yeon Lee[1], Ji-Hun Kim[1], In-Gyun Lee[1], Sung-Hee Lee[2], Dae-Won Sim[2], Hyung-Sik Won[2] and Bong-Jin Lee[1,*]

[1]Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Seoul 151-742, Korea and [2]Department of Biotechnology, College of Biomedical and Health Science, Konkuk University, Chungju, Chungbuk 380-701, Korea

Received December 09, 2014; Revised March 18, 2015; Accepted April 03, 2015

ABSTRACT

HP0268 is a conserved, uncharacterized protein from *Helicobacter pylori*. Here, we determined the solution structure of HP0268 using three-dimensional nuclear magnetic resonance (NMR) spectroscopy, revealing that this protein is structurally most similar to a small MutS-related (SMR) domain that exhibits nicking endonuclease activity. We also demonstrated for the first time that HP0268 is a nicking endonuclease and a purine-specific ribonuclease through gel electrophoresis and fluorescence spectroscopy. The nuclease activities for DNA and RNA were maximally increased by Mn^{2+} and Mg^{2+} ions, respectively, and decreased by Cu^{2+} ions. Using NMR chemical shift perturbations, the metal and nucleotide binding sites of HP0268 were determined to be spatially divided but close to each other. The lysine residues (Lys7, Lys11 and Lys43) are clustered and form the nucleotide binding site. Moreover, site-directed mutagenesis was used to define the catalytic active site of HP0268, revealing that this site contains two acidic residues, Asp50 and Glu54, in the metal binding site. The nucleotide binding and active sites are not conserved in the structural homologues of HP0268. This study will contribute to improving our understanding of the structure and functionality of a wide spectrum of nucleases.

INTRODUCTION

Although the past decade has been marked by outstanding advances in understanding the genomic information throughout all organisms, many genes have yet to be biologically characterized. Therefore, functional studies on gene products, especially unknown or hypothetical proteins, are necessary to obtain a comprehensive view of gene function in the biological world and to discover novel proteins that may be targeted by drugs. To accomplish this task, it would be helpful to determine the three-dimensional structures of proteins and subsequently obtain clues about their biological function by comparison with structural homologues with known function.

Helicobacter pylori (*H. pylori*) is a human pathogenic bacterium that survives in the extreme acidic environment of the stomach. This bacterium is present in almost half of the world's population and is associated with the development of diverse duodenogastric diseases, such as chronic gastritis, peptic ulcers, duodenal ulcers and stomach cancer (1–3). In an effort to treat *H. pylori* infections, multiple antibiotic therapies have been commonly used for many years. However, this treatment history has been accompanied by the global prevalence of *H. pylori* resistance to various antibiotics, which has emerged as a major issue in the 21st century (4,5). Therefore, much effort has been focused on discovering new antibiotic targets among *H. pylori* proteins. The database of essential *H. pylori* genes that are required for cell growth and maintenance has thus been established using a combination of transposon-mediated mutagenesis and high-throughput sequencing (6). This database suggests that the product of the *hp0268* gene (HP0268) might be one of the antibiotic targets. HP0268 is a small, uncharacterized protein that is conserved in *H. pylori* strains and consists of 80 amino acid residues with a molecular weight of approximately 9.5 kDa.

In this study, we determined the solution structure of HP0268 using nuclear magnetic resonance (NMR) spectroscopy and revealed for the first time, that this protein has both DNA nicking and RNase activities through gel electrophoresis and fluorescence spectroscopy. Further-

*To whom correspondence should be addressed. Tel: +82 2 880 7869; Fax: +82 2 872 3632; Email: lbj@nmr.snu.ac.kr
Present address: Ji-Hun Kim, Center for Structural Biology and Departments of Biochemistry and Pharmacology, Vanderbilt University School of Medicine, Nashville, TN, 37232-8725, USA.

more, we investigated the relationship between the structure, nucleotide/metal binding and catalysis of HP0268 using mutagenesis and NMR titration experiments.

MATERIALS AND METHODS

Protein expression, purification and site-directed mutagenesis

The *hp0268* gene was acquired from *H. pylori* 26695 genomic DNA using PCR amplification and subcloned into the plasmid pET28a (Novagen, Inc.). The recombinant plasmid was transformed into *Escherichia coli* BL21 (DE3) host cells. This construct was expressed as a fusion protein with 20 non-native residues containing a hexahistidine-tag and thrombin cleavage site at the N-terminus. Overexpression was induced with 0.5 m*M* isopropyl ß-D-thiogalactopyranoside (IPTG) in *E. coli* BL21 (DE3) for 3 h at 37°C after the cells had grown to an OD_{600} of 0.6. These cells were harvested, resuspended and sonicated in lysis buffer (pH 8.0) containing 50 mM Tris-HCl and 500 mM NaCl. The disrupted cells were then centrifuged at 16 000 g for 1 h. After centrifugation, the supernatant was loaded onto a His_6-tag affinity column (Chelating Sepharose Fast Flow resin; GE Healthcare, Inc.), and the purified protein was eluted with an imidazole gradient from 100 to 300 mM imidazole. The eluted protein was concentrated and thrombin-cleaved to produce HP0268 as a native protein. This protein was finally purified using gel filtration chromatography (Superdex^{TM} 75 10/300; GE Healthcare, Inc.) with the AKTA prime^{TM} system (GE Healthcare, Inc.). Uniformly ^{15}N-, ^{13}C-labeled protein was prepared by growing the cells in M9 medium, which includes 99% $^{15}NH_4Cl$ and 99% [^{13}C]-D-glucose (Cambridge Isotope Laboratories, Inc.). The NMR sample was finally obtained at a concentration of approximately 1 mM in 90% H_2O/10% D_2O containing 20 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 6.0) and 50 mM NaCl. Site-directed mutagenesis of HP0268 was performed using an EZchange^{TM} site-directed mutagenesis kit (Enzynomics, Inc.). To exclude the possibility that the observed catalysis was caused by impurities, control samples for the nicking endonuclease and RNase assays were purified from a bacterial culture containing an empty vector, using the same purification scheme as for HP0268 and its mutants.

NMR measurements and structure determination

The NMR experiments were performed at 298 K on a Bruker AVANCE DRX 600 spectrometer that was equipped with a cryogenic probe, and a Bruker AMX 500 spectrometer. Backbone and side-chain peak assignments were performed using a series of triple resonance spectra [3D HNCO, HN(CA)CO, HNCACB, CBCA(CO)NH, HBHA(CO)NH, HCCH-TOCSY and C(CO)NH]. Aromatic ring resonances were assigned using 2D NOESY, 3D ^{15}N NOESY-HSQC and 3D ^{13}C NOESY-HSQC. Chemical shifts were referenced to DSS externally. NMR data were processed using the program NMRPipe (7) and analyzed using the program NMRView (8). The distance restraints for the structure calculation were collected from the 3D ^{15}N NOESY-HSQC and ^{13}C NOESY-HSQC data by manual and automatic NOE assignments, for which the program CYANA 2.1 was used (9). Dihedral angle restraints were calculated from the chemical shifts using the program TALOS (10), and the overall secondary structure was predicted from the chemical shift index (CSI) (11) and NOE pattern. Hydrogen-bond restraints were obtained using a H–D exchange experiment and an observation of regular secondary elements from the CSI search and NOE patterns. A total of 200 structures of HP0268 were initially generated with the program CYANA 2.1 through the standard simulated annealing protocol and incorporation of the NOE distance, dihedral angle and hydrogen bond restraints. The structure calculations were guided by the macro file, in which several variables were set: 10 000 torsion angle dynamics steps and the chemical shift tolerances of proton (0.05) and carbon/nitrogen (0.5). A total of 100 structures with the lowest target function values were selected and further refined by the standard simulated annealing and torsion angle dynamics using the program CNS 1.3 (12). The structure calculation protocol included three stages: (i) 1000 steps of high-temperature molecular dynamics in torsion angle space at 50 000 K; (ii) 3000 steps of a slow-cooling annealing stage in torsion angle space from 50 000 K to 250 K and (iii) 200 steps and 10 cycles for final conjugate gradient minimization. The 20 lowest-energy conformers were finally refined in explicit water using the RECOORD script (13). The programs MOLMOL and PyMOL (DeLano Scientific LLC) were used to visualize the result of the 20 energy-minimized conformers. The quality of the final structure was analyzed using PROCHECK-NMR and Aqua (14).

Nicking endonuclease assay

Relaxed circular plasmid DNAs [pET-15b(+), pET-21a(+) and pET-28a(+)] were prepared using a Wizard Plus SV Miniprep kit (Promega). Recombinant HP0268 proteins (1, 2, 4 and 8 μM) were incubated at 37°C for 30 min with each circular DNA (10 ng/μl) in a final volume of 30 μl. The reaction was conducted in 20 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 7.0) containing 150 mM NaCl. The reaction conditions included variable pH levels (pH 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0 and 9.5) and the presence of 1 mM metal ions (Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}). The commercial enzymes Nt.BsmAI and NdeI (NEB, Inc.) were used as references to validate the DNA nicking and double-cutting activity of HP0268, respectively. Ten units of the two enzymes were used in a final volume of 30 μl. Reactions were stopped by the addition of 6 μl of a 6×loading buffer (100 mM EDTA, 0.7% SDS and 70% glycerol), and the reaction solutions were loaded onto a 1.0% agarose gel containing 0.5 mM TBE buffer (90 mM Tris borate and 2 mM EDTA) and electrophoresed in the same buffer. DNA was visualized under UV light at 254 nm using ethidium bromide staining. At least three independent experiments were performed for each protein and yielded similar results. The digested, nicked and relaxed circular plasmids were quantified using the Gel Doc^{TM} EQ system (Bio-Rad, Inc.).

Fluorescence nuclease assay

The RNase activities of HP0268 and its single-site mutants were examined using fluorescence spectroscopy com-

bined with an RNaseAlert kit (IDT, Inc.). The fluorescence that was emitted from the cleavage of the substrate was measured at 520 nm after excitation at 490 nm. Basically, the reaction solution contained 20 mM Tris (pH 8.0) and 150 mM NaCl. Ribonuclease A (0.2 μM) and the purified HP0268 protein (1, 2, 4 and 6 μM) were incubated at 37°C for 150 min with the fluorophore-labeled RNA substrate (1 μM, calculated from 50 pmol in a total volume of 50 μl). The reactions were conducted in the presence and absence of various divalent metal ions (1 mM), such as Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}. Eight base-specific single-stranded RNAs and DNAs, termed poly(A)$_{10}$, poly(U)$_{10}$, poly(G)$_{10}$, poly(C)$_{10}$, poly(dA)$_{10}$, poly(dT)$_{10}$, poly(dG)$_{10}$ and poly(dC)$_{10}$, were synthesized by Bioneer (Daejeon, Republic of Korea). The fluorescein fluorophore and lowa Black FQ quencher were appended to the 5' and 3' ends of the oligonucleotide, respectively. These substrates (100 pmol) were incubated at 37°C for 150 min with the HP0268 protein (2 μM). In all of the nuclease experiments, at least three independent experiments were performed for each protein and yielded similar results.

NMR chemical shift experiments

An NMR chemical shift perturbation analysis of ^{15}N-labeled HP0268 was performed on a Bruker AMX 500 spectrometer in the presence of the unlabeled nucleoside monophosphates (AMP, GMP, UMP, CMP, dAMP, dGMP, dTMP and dCMP) and the diamagnetic Mg^{2+} ion. The chemical shift changes were recorded by the difference in the chemical shift values of each peak before the titration and at ligand saturation. The average chemical shift perturbation values ($\Delta\delta_{avg}$) for ^{15}N and ^{1}H nuclei were derived as follows from Equation 1:

$$\Delta\delta_{avg} = [(0.2 \times \Delta\delta_N^2 + \Delta\delta_H^2)/2]^{1/2} \quad (1)$$

where $\Delta\delta_N$ and $\Delta\delta_H$ represent the chemical shift perturbation values of the amide nitrogen and proton, respectively. The resonance broadenings were quantified with regard to the HP0268 residues using the ratio of the peak intensities in the presence and absence of the paramagnetic Mn^{2+} ion. Throughout the NMR titration experiments, the protein concentration was maintained at 0.4 mM, while the concentrations of the nucleoside monophosphates and diamagnetic metals ranged from 0 to 2 mM. The paramagnetic metals were used at a relatively low concentration of 0.2 mM to avoid resonance broadening induced by the non-specific binding of the metal to the protein. The NMR samples were dissolved in 20 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 6.0) containing 50 mM NaCl at 298 K.

RESULTS

Solution structure of HP0268

To determine the solution structure of HP0268 by NMR spectroscopy, 2D-NMR (^{1}H-^{15}N HSQC) spectra were first obtained to examine whether HP0268 is well-structured and to optimize the solvent conditions for additional 3D NMR measurements. The ^{1}H-^{15}N HSQC spectrum of ^{15}N-enriched HP0268 exhibited well-dispersed and homogeneous cross-peaks, indicating that the overall structure of HP0268 is well-defined. Conventional 3D NMR data were collected and used to assign most ^{1}H, ^{13}C and ^{15}N resonances of the backbone and side-chain atoms of HP0268. Based on the resonance assignments of HP0268, the final solution structure of HP0268 was determined using 1534 distance constraints that were derived from Nuclear Overhauser Enhancement (NOE), 98 TALOS-derived dihedral angles, and 24 hydrogen bond constraints that were derived from a hydrogen-deuterium exchange experiment. During the simulated annealing protocol, the 20 lowest-energy structures out of 200 were chosen and subjected to water refinement. None of the refined conformers had NOE violations greater than 0.5 Å and dihedral angle violations greater than 5°. The solution structure of HP0268 is described as the backbone superposition of the ensemble of the final 20 structures (Figure 1A). The structural statistics are summarized in Table 1 and show a high quality NMR structure of HP0268. The NMR structures of residues 17–78 are well-converged with an average root-mean-square deviation (RMSD) of 0.402 Å for backbone atoms and 0.921 Å for all of the heavy atoms. The average RMSD values for backbone and heavy atoms of all the residues are 0.545 and 1.092 Å, respectively. The structural quality was also assessed using the Ramachandran analysis of the ensemble. The residues of HP0268 are categorized into the most favored (86.2%), additionally allowed (12.5%), generously allowed (0.8%) and disallowed (0.5%) regions. This protein folds into a classical two-layered α/β-sandwich structure with a topology described as β1α1β2α2β3α3β4, comprising a mixed four-stranded β-sheet that is stacked against two α-helices, both of which are nearly parallel to the strands of the β-sheet (Figure 1B). The β-strands correspond to residues 3–5 (β1), 31–36 (β2), 59–63 (β3) and 73–78 (β4), while the α-helices correspond to residues 17–28 (α1), 43–54 (α2) and 66–70 (α3).

We searched for structural homologues of HP0268 within the Protein Data Bank (15) by submitting the lowest energy-minimized NMR conformer of HP0268 to the fold recognition program DALI (16). The structural homologues with a DALI Z-score higher than 6.5 are the small mutS-related (SMR) domain of MutS2 from *Deinococcus radiodurans* (687–766 aa, PDB ID = 4OCH, Z-score = 9.8 and RMSD = 3.0), the uncharacterized protein YDAL from *E. coli*. (85–175 aa, PDB ID = 3QD7, Z-score = 7.3 and RMSD = 2.7), and the SMR domain of Bcl3-binding protein (B3BP) from *Homo sapiens* (1688–1770 aa, PDB ID = 2VKC, Z-score = 6.9 and RMSD = 2.5). The structures of the proteins are shown in Figure 1C. The proteins also have a secondary structure topology that is similar to that of HP0268. Notably, HP0268 is structurally most similar to the SMR domain, although they lack significant sequence homology. The SMR domain is a nicking endonuclease that is involved in DNA mismatch repair (MMR) or gene recombination (17). This fact prompted us to test whether HP0268 has nuclease activity similar to that of the SMR domain.

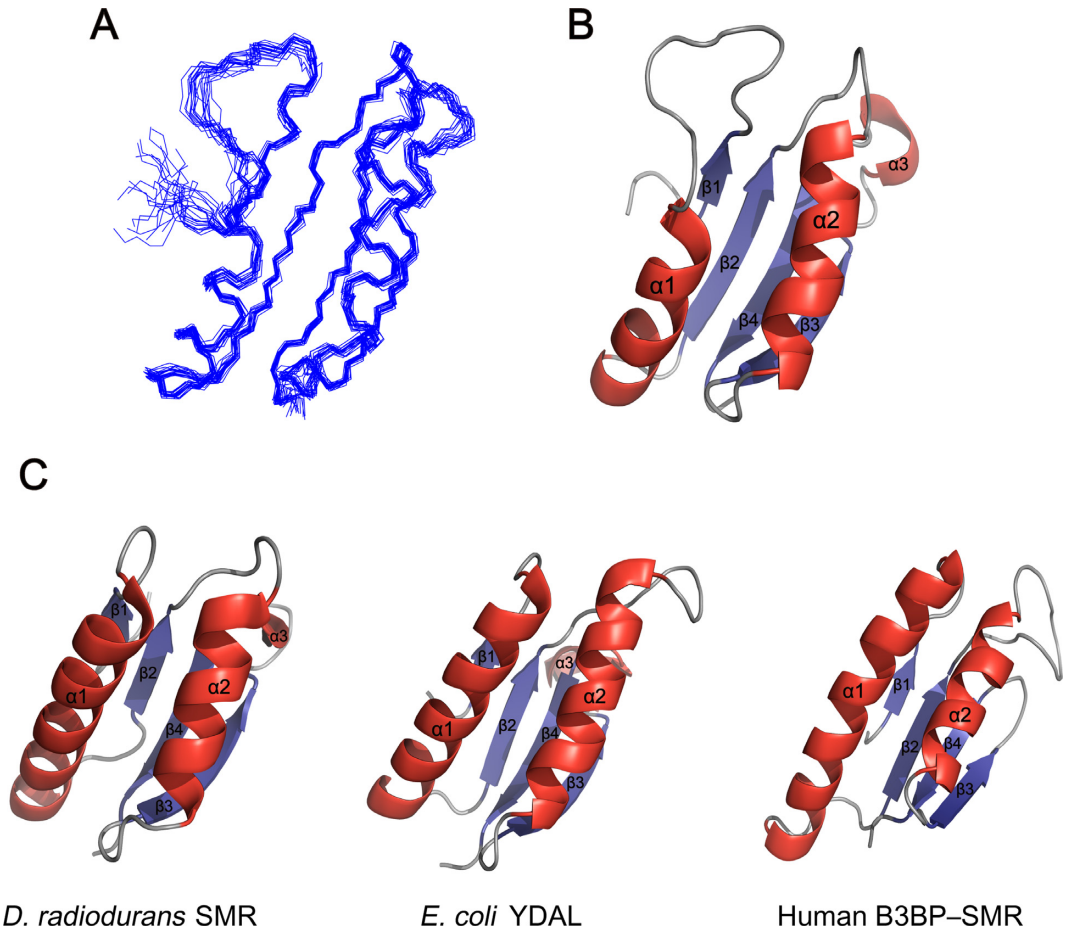

Figure 1. Solution structure of HP0268 and its structural homologues. **(A)** NMR structure of HP0268. The structure is described as the backbone superposition of the 20 best energy-refined conformers. **(B)** Ribbon representation of the structure of HP0268. The protein adopts a two-layered α/β-sandwich structure with a β1α1β2α2β3α3β4 topology. **(C)** Ribbon representation of structural homologues of HP0268. From left to right, three structures of the SMR domain from *D. radiodurans* (687–766 aa, Z-score = 9.8), the YDAL protein from *E. coli* (85–175 aa, Z-score = 7.3) and the B3BP-SMR domain from *Homo sapiens* (1688–1770 aa, Z-score = 6.9) are shown. The secondary structures of (B) and (C) are displayed as follows: α-helices in red, β-strands in deep blue and loops in gray.

Nicking endonuclease activity of HP0268

The DNA nicking activity of HP0268 was investigated by conversion of relaxed circular DNA to nicked open-circular DNA when one strand was cleaved. The DNA substrate of the protein was the plasmid pET-21a(+), which had been obtained in a relaxed circular form rather than a covalently closed-circular form. In general, the conformations of plasmid DNAs have different mobilities in a gel during electrophoresis. Nicked open-circular DNA with one strand cut, relaxed circular DNA with both strands uncut and linear DNA with free ends are the slowest to the fastest in order of electrophoretic speed. The DNA nicking activity was first measured by incubating the plasmid DNA with increasing amounts of HP0268. We found that the relaxed circular DNA changed into nicked open circular DNA, and this conversion linearly depended on the protein concentration up to 8 μM (Figure 2A). The nicked plasmid DNA was located above the undigested plasmid DNA on the agarose gel. Some cleavage of double-stranded DNA also occurred with a high concentration of HP0268, producing a linear DNA that was located below the nicked and undigested plasmid DNAs on the agarose gel. The appearance of linear DNA is caused by single-stranded cuts at close sites on opposite strands. Each band was identified as one of the DNA conformations by comparing the reference bands that were generated by the commercial enzymes Nt.BsmAI (a nicking enzyme) and NdeI (a restriction enzyme). The other plasmids, pET-15b(+) and pET-28a(+), were also used as substrate DNA for HP0268, leading to similar results (Supplementary Figure S1). Linear DNAs were commonly produced at a relatively high concentration (20 μM) of HP0268. Moreover, the nicking endonuclease activity was surveyed over a broad pH range of 5.5–9.5, indicating that HP0268 displays maximal activity at a slightly higher pH (pH 8.0–8.5) (Figure 2B). The optimal pH range permitted the subsequent production of some digested DNA. There was no significant difference between the activities at a pH range of 5.5–7.5. Additionally, to evaluate the metal dependence of the nuclease activity, similar experiments were conducted at pH 7.5 in the presence and absence of diverse metal ions, such as Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}. These eight metals are cofactors of many functional proteins in the cell. Compared

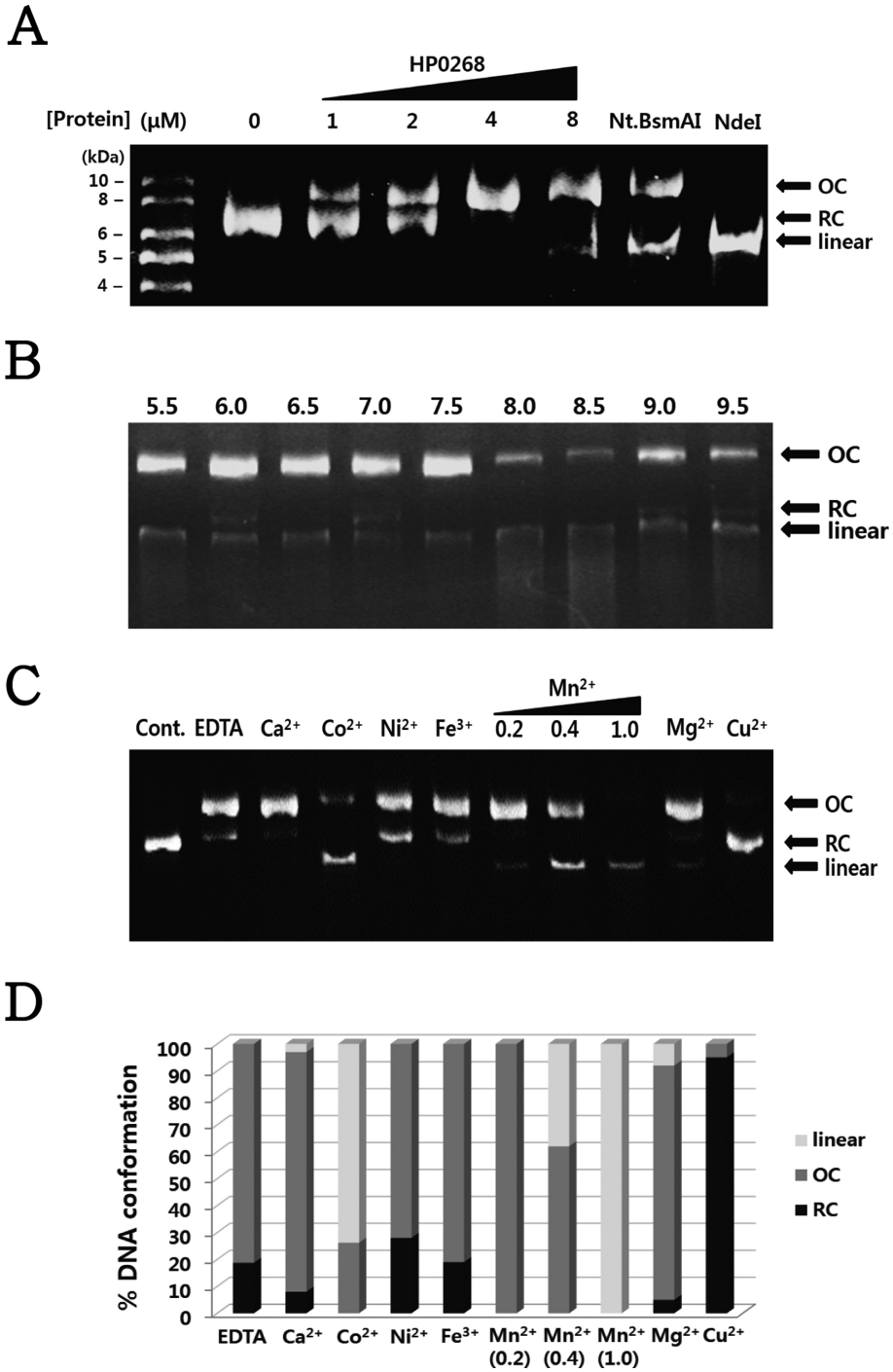

Figure 2. Nicking endonuclease activity of HP0268. **(A)** The nicking endonuclease activity at various protein concentrations (1, 2, 4 and 8 μM) after incubation at 37°C for 30 min. OC, RC and linear are abbreviations for the nicked open-circular, relaxed circular and linear DNA, respectively. **(B)** The pH dependence of the DNA nicking activity. The substrate plasmid DNA was incubated with 4 μM HP0268 at 37°C for 30 min under various pH conditions (pH 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0 and 9.5). **(C)** Effect of metal ions on the DNA nicking activity of HP0268. The substrate plasmid DNA was incubated with 4 μM HP0268 at 37°C for 30 min in the presence and absence of 1 mM metal ion (Ca^{2+}, Co^{2+}, Ni^{2+}, Fe^{3+}, Mn^{2+}, Mg^{2+} and Cu^{2+}). Increasing concentrations (0.2, 0.4 and 1 μM) of Mn^{2+} ion were used. Excess EDTA was used to remove the residual metal ions during the protein preparation. **(D)** The percentages of the resulting DNA conformations were plotted with regard to metal ion used. Cont. represents the substrate plasmid pET-21a(+) without HP0268, and Nt.BsmAI and NdeI represent the positive controls for the nicked and linear DNA, respectively. Commonly, 10 units of control enzyme were used in a final volume of 30 μl.

Table 1. Structural statistics for the final 20 energy-minimized conformers of HP0268

NOE upper distance constraints	
All	1534
Intraresidue ($\mid i-j \mid = 0$)	576
Sequential ($\mid i-j \mid = 1$)	453
Medium range ($2 \leq \mid i-j \mid \leq 4$)	198
Long range ($\mid i-j \mid \geq 5$)	307
Dihedral angle constraints	
All	98
Φ	49
Ψ	49
Hydrogen-bond constraints	24
Residual NOE and angle violations	
Number of NOE violations > 0.5 Å	0
Number of torsion angle violations > 5°	0
RMSD from the average coordinate (Å)	
Residue 17–78	
Backbone atoms (N, Cα, CO) 0.402 ± 0.081	
All of the heavy atoms 0.921 ± 0.126	
Residue 1–80	
Backbone atoms (N, Cα, CO)	0.545 ± 0.084
All of the heavy atoms	1.092 ± 0.132
Deviation from idealized geometry	
Bonds (Å)	0.000968 ± 0.000087
Angles (°)	0.294 ± 0.0032
RECOORD water refinement energies (kcal/mol)[a]	
$E_{overall}$	– 2894.88 ± 36
E_{bond}	45.68 ± 0.78
E_{angles}	431.62 ± 12
$E_{improper}$	44.28 ± 0.65
E_{vdw}	– 338.75 ± 11
E_{elec}	– 3082.15 ± 11
E_{noe}	0.32 ± 0.07
E_{cdih}	4.12 ± 1.1
Ramachandran plot (%)[b]	
Most favored region	86.2
Additionally allowed region	12.5
Generously allowed region	0.8
Disallowed region	0.5

[a]The default parameters and force constants of protein-allhydg5-4.param and re.h2o.inp in RECOORD were used for water refinement.
[b]The program PROCHECK-NMR was used for the calculation.

to the absence of the metal ions, the DNA nicking activity was most elevated with Mn^{2+} and Co^{2+} ions, and the nicked DNA was subsequently converted into a linear DNA (Figure 2C). The linear DNA was further fragmented, particularly in the presence of the Mn^{2+} ion. The Ca^{2+} and Mg^{2+} ions slightly increased the DNA nicking activity. In contrast, the Cu^{2+} ion most suppressed the nicking activity, leaving a large amount of the relaxed circular DNA fully intact. In particular, aggregation of HP0268 was promoted during incubation, even at a low concentration of Zn^{2+}, suggesting that the Zn^{2+} ion damages the structural integrity of the protein with a loss of function. Similar patterns of metal dependence were observed when the assay was performed with variable metal concentrations ranging from 0.5 to 5 mM (data not shown). To calculate the degree of the nuclease activity of HP0268 in the presence of these metal ions, different ratios of nicked and linear DNA to unreacted DNA were plotted against the metal ions as shown in Figures 2D. The total amount of converted plasmid DNA was indirectly estimated by quantifying the residual uncut DNA. The values were determined by averaging three replicate experiments. We concluded that the nicking endonuclease activity of HP0268 was maximal at pH 8.5 and in the presence of the Mn^{2+} ion, and was most suppressed by the Cu^{2+} ion.

Purine-specific ribonuclease activity of HP0268

The RNase activities were examined using the RNaseAlert kit combined with fluorescence spectroscopy, which is a highly sensitive assay for the quantification and real-time kinetics of RNase activity. In principle, this method uses a single-stranded RNA substrate that contains a quencher on one end, and if the RNA is cleaved by an RNase, then a fluorescent molecule on the other end emits green fluorescence. The RNA substrate in the kit is composed of four typical ribonucleotides that are randomly distributed with many possible combinations and thus has commonly been used as an alternative substrate for characterizing an ill-studied or putative RNase without knowledge of the sequence specificity. The fluorescence emission, which reflects the RNase activity, increased at elevated concentrations of HP0268 (1, 2, 4 and 6 μM) (Figure 3A), showing that the protein has RNase activity. The kinetic pattern of the enzymatic activity at higher protein concentrations was consistent with a first-order reaction, where the velocity at any time is proportional to the existing substrate concentration, similar to

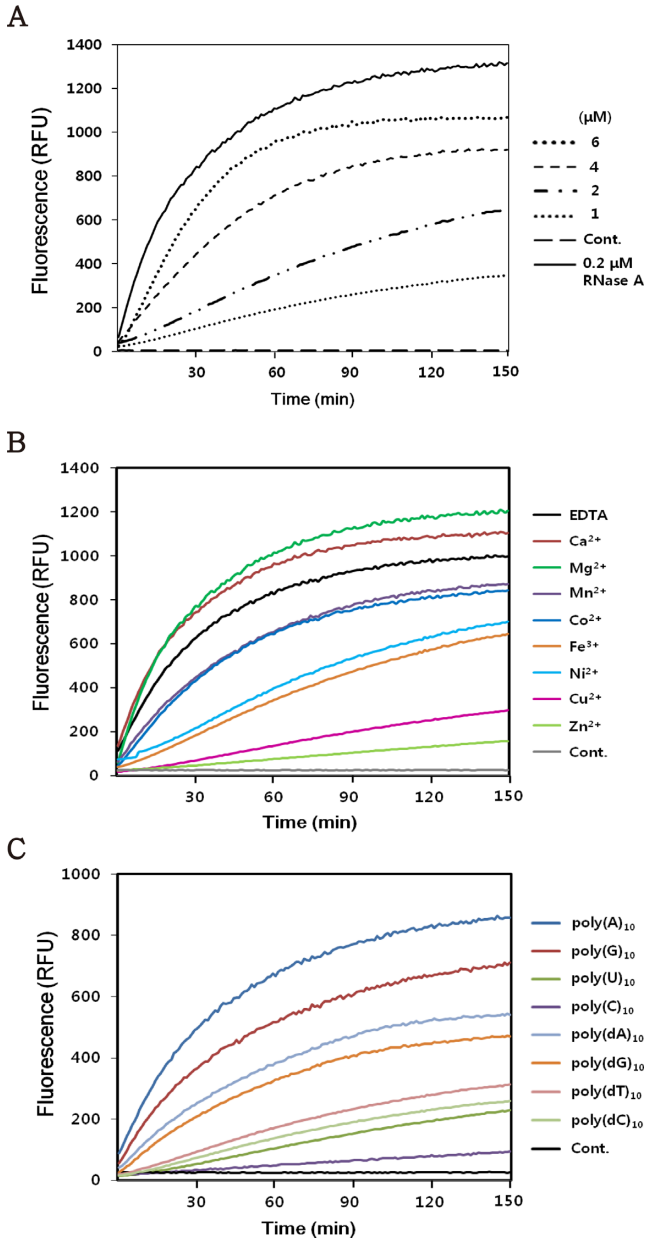

Figure 3. Fluorometric assay for the detection of the ribonuclease activities of HP0268. **(A)** Ribonuclease activities depending on the protein concentration. Fluorescence spectra were recorded in the presence of increasing concentrations of HP0268 (0, 1, 2, 4 and 6 μM). As a positive control for ribonuclease activity, 0.2 μM ribonuclease A was used. **(B)** Metal dependence of the ribonuclease activity. The fluorescence spectra were recorded in the presence and absence of 1 mM metal ions including Ca^{2+}, Mg^{2+}, Mn^{2+}, Co^{2+}, Fe^{3+}, Ni^{2+}, Cu^{2+} and Zn^{2+}. The fluorescence spectra are shown in a color-coded mode. The protein concentrations were maintained at 5 μM. EDTA represents the reference spectrum that was acquired by the sample that was purified in an excess of EDTA, which eliminates almost all of the metal ions in solution. **(C)** Base specificity of the nucleolytic activity. The substrates of HP0268 are eight base-specific oligonucleotides [$poly(A)_{10}$, $poly(U)_{10}$, $poly(G)_{10}$, $poly(C)_{10}$, $poly(dA)_{10}$, $poly(dT)_{10}$, $poly(dG)_{10}$ and $poly(dC)_{10}$], which are composed of 10-mer adenine, uracil, guanine, cytosine, deoxyadenine, deoxythymidine, deoxyguanine and deoxycytosine, respectively. For comparison, each spectrum is color-coded according to the type of the nucleotides. The protein concentrations were maintained at 2 μM. Cont. in (B) and (C) represents the reference spectra of the buffer solution containing 20 mM Tris (pH 8.0) and 150 mM NaCl, which are included in all of the reaction solutions.

that of many nucleases. As a positive control for this RNase assay, 0.2 μM ribonuclease A was used according to the recommendation in the RNaseAlert kit protocol. An approximately forty-times greater concentration of HP0268 than ribonuclease A is required for similar RNase activity, suggesting that the RNase activity of HP0268 is relatively low. The tested RNA sequence has been carefully optimized to detect several RNases, including RNase A, RNase T1, RNase I, micrococcal nuclease, S1 nuclease, mung bean nuclease and Benzonase®. In a previous report, a relatively high concentration of HP0315 (10 μM) was used to detect its RNase activity using the same RNaseAlert kit (18).

To further characterize the metal dependence of the RNase activity, a similar fluorescence assay was carried out using diverse metal ions, as in the nicking endonuclease study. The results showed that RNA degradation depends, in order of influence, on the presence of Mg^{2+}, Ca^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, Fe^{3+}, Cu^{2+} and Zn^{2+} ions (Figure 3B). Among these ions, the two metal ions Mg^{2+} and Ca^{2+} slightly stimulated the RNase activity compared to no metal. Interestingly, the Mn^{2+} and Co^{2+} ions decreased the RNase activity to some degree, which is in contrast to the previous result from the nicking endonuclease assay. The Cu^{2+} ion exhibits the most inhibitory effects on both the DNA nicking and RNase activities. The effect of the Zn^{2+} ion was excluded from the analysis of the metal dependence because Zn^{2+} promoted protein precipitation during the fluorescence measurement

To assess which type of nucleotide is more amenable to cleavage by HP0268, eight 10-mer single-stranded oligonucleotides that were labeled with fluorogenic and quencher moieties on opposite ends, termed $poly(A)_{10}$, $poly(U)_{10}$, $poly(G)_{10}$, $poly(C)_{10}$, $poly(dA)_{10}$, $poly(dG)_{10}$, $poly(dT)_{10}$ and $poly(dC)_{10}$, were synthesized *in vitro* and used as the RNA and DNA substrates for HP0268 in fluorescence experiments. The results showed that two substrates, $poly(A)_{10}$ and $poly(G)_{10}$, emitted fluorescence that was much greater than that of the other substrates $poly(U)_{10}$ and $poly(C)_{10}$ when they were reacted with the protein (Figure 3C). These fluorescence data indicate that the adenine and guanine ribonucleotides can be readily degraded by HP0268. Notably, the $poly(C)_{10}$ substrate remained nearly uncleaved. In the case of the poly-deoxyribonucleotides $poly(dA)_{10}$, $poly(dG)_{10}$, $poly(dT)_{10}$ and $poly(dC)_{10}$, the fluorescence emissions of these oligonucleotides showed no significant difference, indicating a more relaxed base specificity of HP0268 for DNA (Figure 3C). The substrate preference could be ranked in the order of $poly(dA)_{10} \geq poly(dG)_{10} > poly(dT)_{10} \geq poly(dC)_{10}$. The $poly(dA)_{10}$ and $poly(dG)_{10}$ substrates emitted relatively low fluorescence in comparison with that of the two RNA substrates $poly(A)_{10}$ and $poly(G)_{10}$. The other substrates, $poly(dT)_{10}$ and $poly(dC)_{10}$, showed similar fluorescence profiles to those of $poly(U)_{10}$ and $poly(C)_{10}$, respectively. We concluded that HP0268 is a purine-specific RNase and a nicking endonuclease with relaxed base specificity. However, it remains unclear whether the nuclease activities requires exquisite or relaxed nucleotide sequence specificity.

Nucleotide binding site of HP0268

A transient interaction of nucleotides with HP0268 is inevitably necessary for this protein to exert its nuclease activity. To identify the nucleotide binding site with residue-specific information, we measured how much mono-nucleotide can cause chemical shift perturbations for the backbone amide protons and nitrogens of the protein. Mono-nucleotide substrates have the intrinsic advantage of being resistant to cleavage during NMR measurements; thus, eight mono-nucleotides, AMP, GMP, CMP, UMP, dAMP, dGMP, dCMP and dTMP, were used as alternative substrates for HP0268. This procedure may be appropriate to study aspects of substrate recognition, as demonstrated by a previous report that proposed a mechanism for RNA cleavage by RNase Sa2 using a complex structure with mono-nucleotides (19). Unfortunately, we did not find oligonucleotides or nucleotide mimics that were not completely degraded by HP0268. This result possibly occurred because the NMR titration method requires high concentrations of HP0268 and its nucleotide substrate, making the enzyme becomes more reactive. A heterogeneous mixture of degraded RNAs would produce complicated and difficult-to-interpret chemical shift perturbation data, preventing a reliable determination of the nucleotide binding site. Similar results were also obtained in the presence of two catalytically attenuated mutants, D50A or E54A. To tackle this problem, the less active double mutant D54A/E54A was cloned but not expressed.

We compared ^{1}H-^{15}N HSQC spectra that were acquired from the ^{15}N-labeled HP0268 (the observed partner) alone and in the presence of increasing amounts of the unlabeled mono-ribonucleotides (the unobserved partner). The peaks showing the chemical shift perturbations during the NMR titration correspond to amino acid residues that experience changes in the local chemical environments due to the protein–RNA interaction. As a result, the largest chemical shift changes were observed for Lys7, Lys11 and Lys43 of HP0268 during the course of the titrations with AMP, GMP and UMP, respectively (Figure 4A), indicating that these basic residues are important for nucleotide binding. AMP and GMP induced larger chemical shift changes than did UMP, and their saturable concentrations were relatively low, possibly due to a higher binding affinity of the purine nucleotide to HP0268. Lower chemical shift perturbations upon binding to UMP could also be explained by the binding being achieved in a dynamic complex that involves a broad coverage of the binding interface around the core nucleotide binding site. However, no significant chemical shift changes were observed in the presence of CMP. The chemical shift data agreed well with our previous finding that HP0268 is a purine-specific ribonuclease. In all of the titration experiments, the perturbed residues were almost identical and their NMR peaks were shifted in similar patterns. This observation indicates that the mono-ribonucleotides bind at the same site and to nearly identical conformations of the protein. The peak movements without resonance broadening in the spectra indicated that the protein-nucleotide interaction is achieved in a fast exchange on the NMR time scale, suggesting an intermolecular interaction with a weak affinity. Chemical shift perturbations that were elicited by titration with AMP could serve as an indicator for the nucleotide binding site of HP0268. These values were calculated from the standard chemical shift perturbation equation (Equation 1, Materials and Methods) and were plotted with respect to the HP0268 residues (Figure 4B). For this plot, the NMR titration was performed at a 5-fold molar excess of AMP over HP0268 to reach binding saturation. These values were color-mapped onto the surface representation of HP0268 to spatially locate the area of large chemical shift changes (Figure 4C). Similar chemical shift experiments with four mono-deoxyribonucleotides (dAMP, dGMP, dTMP and dCMP) were performed to compare the region and base preference for DNA binding with those for RNA binding. Notably, the chemical shift perturbation patterns and the affected residues with the mono-deoxyribonucleotides were almost similar to those with the mono-ribonucleotides. However, the degrees of the overall chemical shift perturbations with the mono-deoxyribonucleotides tended to be lower than those with the mono-ribonucleotides. The chemical shift perturbations upon binding to the mono-nucleotides were quantified with regard to the amino acid residues (Supplementary Figure S2). The results are well consistent with the substrate specificity that was determined by our fluorescence experiments using oligonucleotides. Comparative analyses suggested that (i) the residues that are involved in the DNA and RNA binding are nearly identical and that (ii) the RNA binding is slightly more favorable than the DNA binding. The positively charged residues Lys7, Lys11 and Lys15 in loop 1 and Lys43 in the boundary of loop 2 and the α2 region are the main constituents of the common nucleotide recognition site. It is conceivable that the flexibility of the loop is important for efficient binding to the RNA/DNA substrates.

Metal binding site of HP0268

The metal binding site of HP0268 was investigated using a similar NMR titration method as that presented in the previous section. The divalent metal ions Mg^{2+} and Mn^{2+} were chosen as ligands for HP0268 because they are the cofactors that maximally increase the RNase and DNA nicking activities, respectively. Chemical shift perturbations could detect the amino acid residues that were involved in metal binding. We compared the 2D ^{1}H–^{15}N HSQC spectra that were acquired from the ^{15}N-labeled HP0268 alone and with increasing concentrations of Mg^{2+} or Mn^{2+} ions. The results showed that the Mg^{2+} ion caused significant chemical shift changes, especially for the two residues Asp50 and Glu54 (Figure 5A), indicating that these acidic residues mainly participate in the metal binding. Little resonance broadening of the peaks indicated that the interaction of the Mg^{2+} ion with HP0268 is dynamic in a fast-exchange regime. To quantify the chemical shift changes at a binding saturation between the Mg^{2+} ion and HP0268, we calculated the degrees of the chemical shift perturbation at a 5-fold molar excess of the Mg^{2+} ion over the protein. The values are plotted against the amino acid residues of HP0268 as shown in Figure 5B. Based on the perturbation data, the magnesium binding site can be color-mapped onto the surface representation of HP0268 (Figure 5C). The residues with a relatively

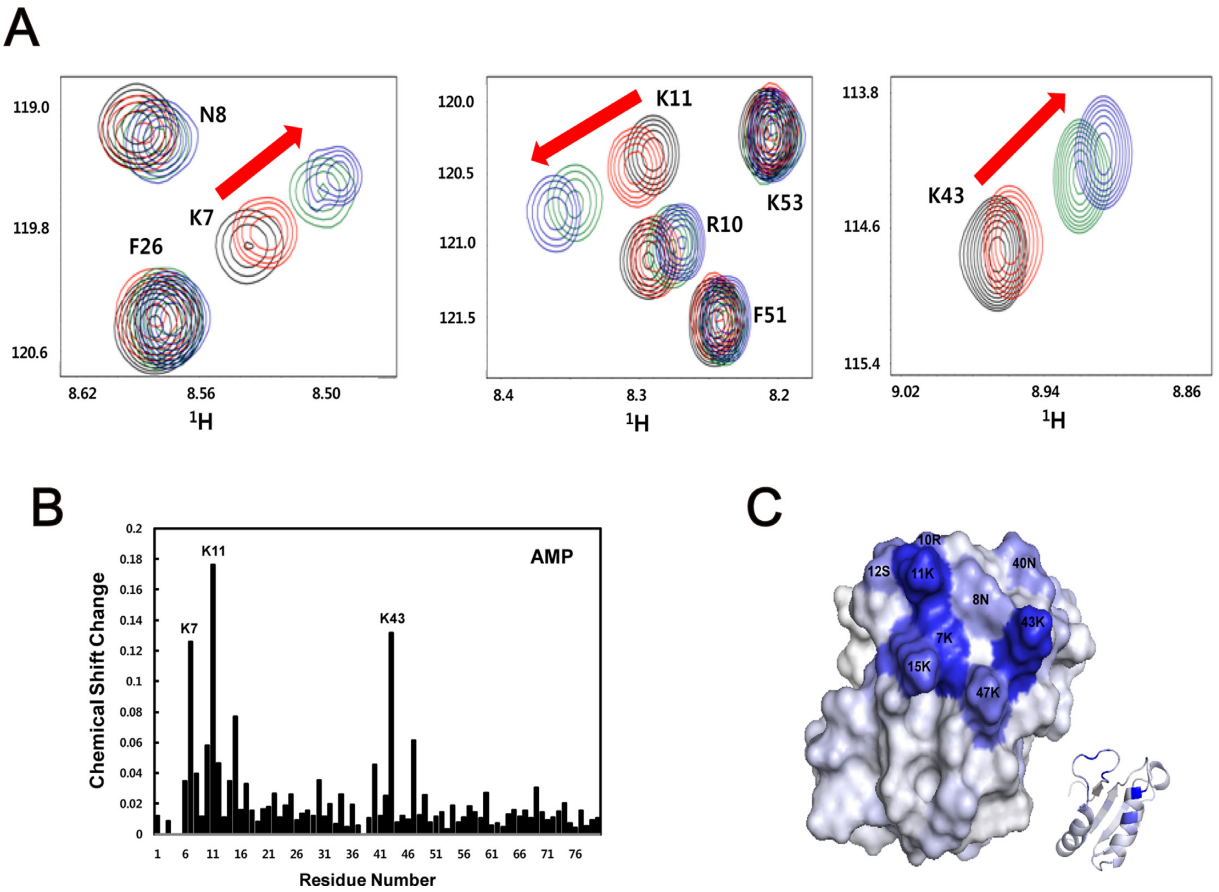

Figure 4. Chemical shift perturbation data of ^{15}N-labeled HP0268 with mono-nucleotides. **(A)** Expanded views of the ^{15}N-^{1}H HSQC spectra. Significant chemical shift changes for the residues, especially K7, K11 and K43, are shown in the presence and absence of the mono-nucleotides AMP, GMP and UMP, respectively. The arrow indicates the directions of the NMR peak shifts, and the peaks of red, green and blue represent the chemical shifts as changed by titration with 2 mM UMP, GMP and AMP, respectively. The protein concentrations were maintained at 0.4 mM. **(B)** Graph of the amide chemical shift changes of HP0268 upon AMP binding. The values of the chemical shift changes upon AMP binding were plotted against the HP0268 residues. AMP was used as a representative of the mono-nucleotide substrates for HP0268. The residues with chemical shift change values that were greater than 0.12 ppm are indicated. **(C)** Chemical shift mapping onto a surface representation of HP0268. The residues that were affected by AMP binding are colored according to a gradient that was derived from their chemical shift perturbation values; darker blue represents a larger perturbation. On the lower right-hand side, the corresponding residues are also colored onto a ribbon representation of HP0268.

large value are mostly distributed in a part of the α2 region of the protein. Of these residues, Asp50 and Glu54 showed the largest values at greater than 0.14 ppm and were thought to be essential for magnesium binding.

In the case of the NMR titration with the Mn^{2+} ion, its intrinsic property of paramagnetic relaxation enhancement (PRE) was utilized to identify the residues for manganese binding. In contrast to the diamagnetic Mg^{2+} ion, manganese binding was manifested as reductions in the peak intensities for the residues in the vicinity of the Mn^{2+} ion. In theory, the level of the reductions is proportional to the inverse sixth power of the distance between the residue and the Mn^{2+} ion. However, decreases in the overall peak intensities were observed in the titration with an excess amount of the Mn^{2+} ion, presumably due to its non-specific binding to the solvent-exposed regions of the protein. To accurately determine the metal binding site, the Mn^{2+} ion was used at half of the protein concentration, although this concentration does not completely saturate the binding site with the Mn^{2+} ion. To quantify the PRE effect, we calculated the ratios of the peak intensities that were obtained in the presence and absence of the Mn^{2+} ion. These values are plotted with respect to the residues of the protein (Figure 5B), and the lower values correspond to the residues involved in manganese binding. The most strongly affected peaks corresponded to the residues Leu46, Val49, Asp50, Lys53 and Glu54. The regions that were susceptible to the PRE effect are color-mapped onto the surface representation of HP0268 (Figure 5C). The binding site of the Mn^{2+} ion is primarily confined to a part of the α2 region and is highly similar to that of the Mg^{2+} ion. Considering all of these data, we speculated that the closely located acidic residues Asp50 and Glu54 of the α2 region are important for the interaction with positively charged metal ions.

Catalytically active site of HP0268

A site-directed mutagenesis study was performed to determine the HP0268 residues that are engaged in the DNA nicking and RNase activities. Given that nuclease activation is commonly stimulated by the histidine or acidic residues

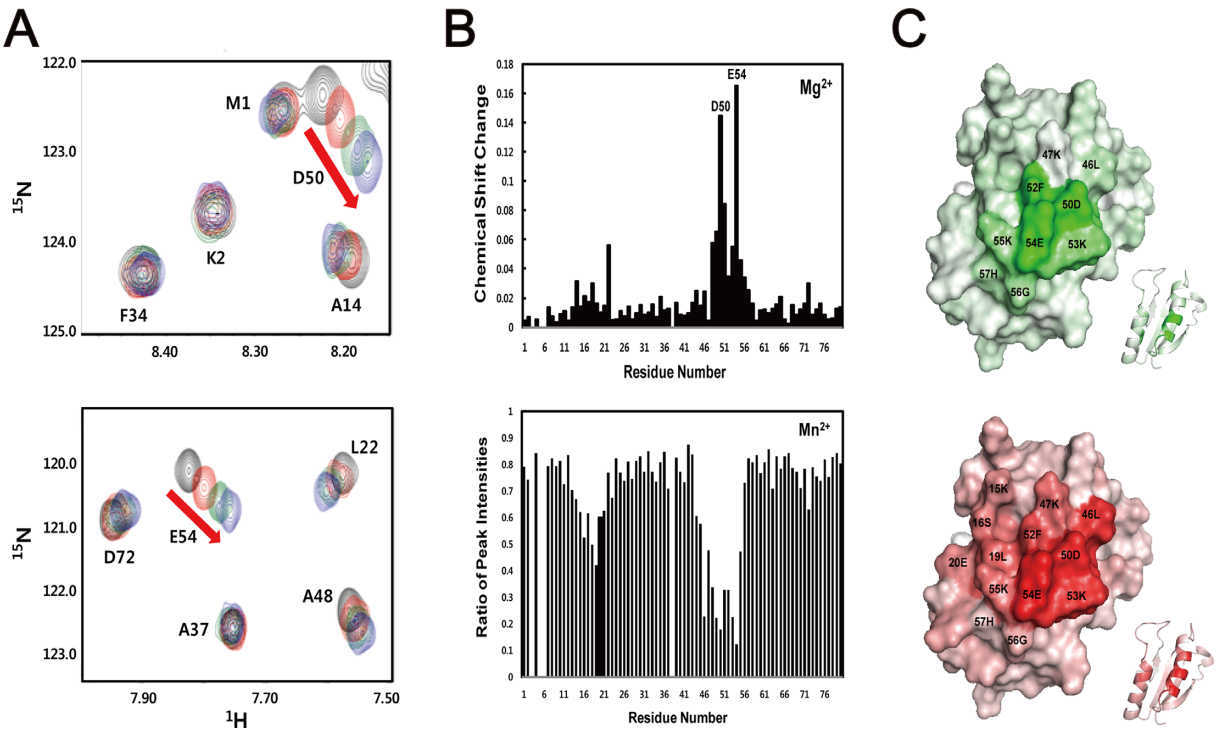

Figure 5. The metal binding site of HP0268. **(A)** Expanded views of the ^{15}N-^{1}H HSQC spectra. Significant chemical shift changes for the residues D50 and E54 are shown in the presence of the Mg^{2+} ion. The arrow refers to the directions of the NMR peak shifts in the course of the titration with increasing concentrations of Mg^{2+} ion. The peaks of red, green and blue correspond to the chemical shifts in the metal concentrations of 0.4, 1.0 and 2.0 mM, respectively. **(B)** Graph of the resonance changes of HP0268 upon binding to the Mg^{2+} and Mn^{2+} ions. The upper figure shows the values of the chemical shift changes upon magnesium binding plotted against the HP0268 residues. The residues with chemical shift change values that were greater than 0.14 ppm are indicated. The lower figure shows the intensity ratios of the individual peaks for the HP0268 residues. The comparative peaks were obtained in the presence and absence of 0.2 mM Mn^{2+} ion. Lower values represent more reductions in the peak intensities due to the paramagnetic property of the Mn^{2+} ion. **(C)** Mapping of the metal binding site onto a surface representation of HP0268. The residues that were affected by binding to the Mg^{2+} and Mn^{2+} ions are gradient-colored according to the values of chemical shift perturbations and peak intensity reductions, respectively. Darker colors of green and red represent greater changes in the values, showing the binding site of the Mg^{2+} and Mn^{2+} ions, respectively. On the lower right-hand side, the corresponding residues are also colored onto a ribbon representation of HP0268.

of the catalytic site, we selected Asp18, His24, Asp50, Glu54 and His57 as candidate residues to be essential for catalysis. It is possible that the metal-independent nuclease activity of HP0268 involves Asp18 and the two histidines that are located apart from the metal binding site. The other possibility is that Asp50 and Glu54, which are capable of metal binding, are involved in the metal-mediated nuclease activity. In addition, the three residues Lys7, Lys11 and Lys43 of the nucleotide binding site were selected for the mutational analysis to examine how strongly the nuclease activation is related to nucleotide binding. It also remains possible that nucleotide binding and catalysis are coincident events at the same region. The selected eight residues were mutated to alanine, and the individual single mutants (named K7A, K11A, E18A, H24A, K43A, D50A, E54A and H57A) were tested for both DNA nicking and RNase activities. Far-UV circular dichroism (CD) analysis showed that these single mutations do not perturb the secondary structure of the protein, causing no significant differences between the CD spectra of the wild-type and mutants (Supplementary Figure S3). The electrophoretic results of the nicking endonuclease activity were obtained with respect to the wild-type and individual mutants (Figure 6A), and the amounts of the resulting nicked-DNA were calculated and normalized by the wild-type activity (Figure 6B). The two mutants D50A and E54A showed the greatest reductions in the nicking activities among the native and mutants, suggesting that the two acidic residues play a crucial role in catalysis. The four mutants K7A, K11A, E18A and K43A showed relatively moderate reductions in DNA nicking activity. The nucleotide binding event involving the lysine residues seems to be important in the nuclease activation process. The other two mutants H24A and H57A have almost the same activities as the wild-type protein.

The RNase activities of the mutants were assessed by the previous method using the RNaseAlert kit (IDT Inc.). The mutants could be categorized into three groups according to the level of the fluorescence emission as shown in Figure 6C. Group 1 (D50A and E54A) showed a significantly decreased fluorescence emission due to a great loss of RNase activity compared to that of the wild-type. Group 2 (K7A, K11A, E18A and K43A) and group 3 (H24A and H57A) showed a moderately and slightly decreased fluorescence emission, respectively. The fluorescence data were consistent with the mutational effects on the nicking endonuclease activity. We concluded that the residues Asp50 and Glu54 form the potential catalytic site for both the nicking and RNase activities.

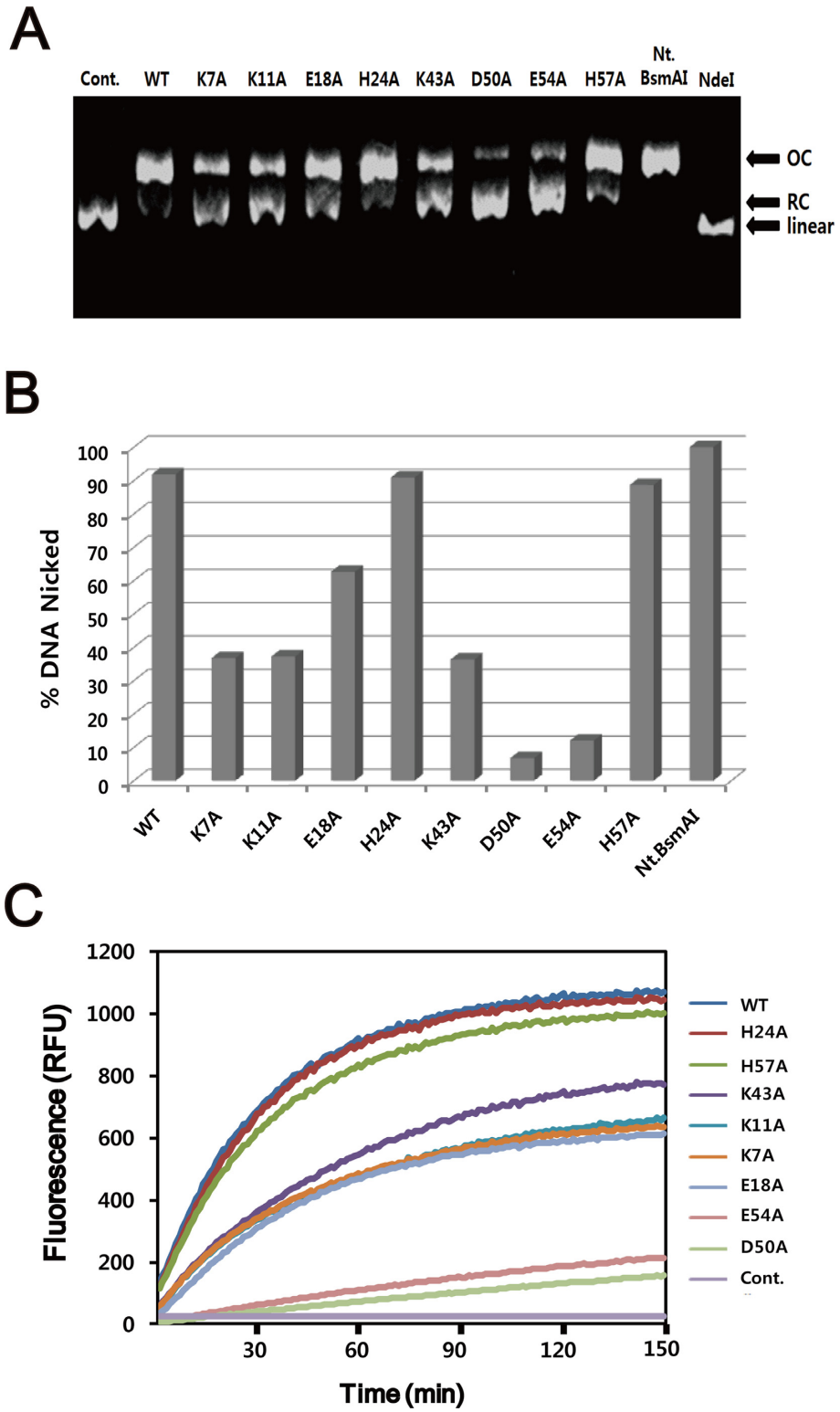

Figure 6. Nuclease activity of HP0268 mutants. **(A)** Nicking endonuclease assay of wild-type and mutant HP0268 using gel electrophoresis. The substrate plasmid DNA was incubated with the wild-type and mutants (2 μM) at 37°C for 30 min. Nt.BsmAI and NdeI represent the positive controls for the nicked and linear DNA, respectively. **(B)** Graph of the nicking endonuclease activities of wild-type and mutant HP0268. The DNA nicking activities of the mutants are normalized by that of the wild-type. **(C)** Fluorometric ribonuclease activities of wild-type and mutant HP0268. The protein concentrations were maintained at 6 μM. The fluorescence spectra are shown in a color-coded mode. In every figure, Cont. and WT indicate the reference condition of having only a buffer and the wild-type protein, respectively. The reaction buffer consisted of 20 mM Tris (pH 8.0) and 150 mM NaCl.

DISCUSSION

We determined the solution structure of HP0268 and discovered that this protein has nicking endonuclease and RNase activities, both of which are specific for a single-strand of nucleotides. To date, bifunctional nucleases (BFN) with both RNase and DNase activities are not well documented. The BFN domain (IPR003729) that is categorized by the InterPro database exists as a homodimer, with each monomer composed of approximately 140 amino acids. Another functionally similar domain is annotated as a DNA/RNA non-specific endonuclease domain (IPR001604) that is mainly found in extracellular endonucleases. This domain contains a common 'ββα-metal' motif that primarily binds to the phosphate backbone of nucleic acids. However, significant differences between the two domains and HP0268 are found with regard to sequence and structure. The two domains can perform a double-strand cut of DNA but do not exhibit nicking endonuclease activity similar to that of HP0268. Based on the biological role of some sugar-nonspecific nucleases in prokaryotes (20), it could be inferred that HP0268 degrades foreign DNA and RNA molecules in a sequence-independent manner for host defense. In general, single-strand specific nucleases are widely distributed and implicated in recombination, repair and replication in the cell (21). These biological roles are consistent with the previous finding that the *hp0268 gene* is essential for cell survival and maintenance (6). Some small proteins that are related to cell toxicity are known to have purine-specific RNase activity, which is similar to HP0268. In the InterPro database, these proteins are classified into the toxin YoeB (IPR009614) and toxin RelE/ParE (IPR007712) families. The two toxins have been proposed to be an mRNA interferase as a toxic component of a toxin-antitoxin (TA) module, inhibiting protein translation and thus cell growth. However, their sequences and structures are not analogous to those of HP0268. Further detailed studies should be performed to determine the correlation between the structure and the substrate preference or specificity of HP0268.

The elution profiles of size exclusion chromatography showed that HP0268 exists as a monomer even in an excess of diverse metals or nucleotides (data not shown). Based on previous reports that a dimeric restriction enzyme engineered into a monomer form changes into a DNA nicking endonuclease (22,23), HP0268 is thought to have one catalytic site for the cleavage of single-stranded RNA or DNA. A combined analysis using NMR titration and mutagenesis experiments suggested that HP0268 has nucleotide binding and catalytic sites that are spatially close to each other as shown in Figures 7A and B. This orientation could provide a rational model in which an initial substrate binding to the nucleotide binding site increases the accessibility of the substrate to the catalytic site, facilitating its efficient digestion. HP0268 is structurally similar to SMR domains that can be categorized roughly into three subfamilies according to their arrangement in the domain architecture (24). The SMR domains that are a part of *D. radiodurans* MutS2 and human B3BP fall into subfamilies 1 and 2, respectively, and *E. coli* YDAL and HP0268 fall into subfamily 3 that is found as stand-alone type proteins. However, the biological function of subfamily 3 has not been well established, leaving the possibility of functional or physical associations with other nuclease-related proteins. Among the SMR domains that have been functionally well characterized, the *Thermus thermophilus* SMR domain, which is composed of 76 amino acids, exhibits nicking endonuclease activity for supercoiled plasmid DNA and cleavage activity for linear double-stranded DNA, showing that a small protein by itself can have significant DNase activity (17). It has also been reported that the human B3BP-SMR domain and *E. coli* YDAL have nicking endonuclease activities (25,26). Despite the use of different DNA substrates and reaction conditions, the nicking endonuclease activity of HP0268 is thought to be approximately one order of magnitude weaker than those of the *T. thermophilus* SMR domain and *E. coli* YDAL, and to be slightly stronger than those of the human B3BP-Smr domain.

HP0268 has characteristic lysine residues that are the primary constituents of a positively charged nucleotide binding site. It could be hypothesized that this highly basic surface binds to the negatively charged phosphate backbone of nucleic acids without discrimination between RNA and DNA. In addition, compared with large nucleases, this small protein might have greater difficulty to specifically recognize the sugar parts of RNA or DNA with an exquisite binding mechanism. This possibility could be supported by our chemical shift perturbation and fluorescence data showing that HP0268 is a sugar-nonspecific nuclease with a shared binding site for DNA and RNA. In the strict sense, HP0268 has slight ribonuclease activity rather than nicking endonuclease activity. It is possible that 2-OH of the RNA sugar ring facilitates slightly greater electrostatic interactions with the basic nucleotide binding site than does the 2-H of the DNA sugar ring. Mutagenesis experiments suggest that the two acidic residues Asp50 and Glu54 are involved in metal binding and catalysis. It is known that DNA/RNA non-specific nucleases have a common catalytic site for the hydrolysis of RNA and DNA (27). It seems likely that the two residues regulate nucleotide cleavage through their binding to specific metal ions. We revealed the different effects of diverse metals on the nuclease activities of HP0268; among the eight metals that were used in this study, the most preferred metals for the DNA nicking and RNase activities are Mn^{2+} and Mg^{2+} ions, respectively, whereas Cu^{2+} ion has the greatest inhibition of these activities. The exact mechanism behind the nuclease activity and its metal dependence remains to be elucidated. Many divalent metals in living cells play critical roles as cofactors that stabilize the structure of various enzymes and help to catalyze a variety of chemical reactions. The Mg^{2+} ion is frequently related to many nucleases because of its abundance, solubility, redox stability, small size and stringent coordination geometry compared with those of other divalent cations (28,29). The Mn^{2+} ion tends to broaden the substrate specificity of nucleases and metalloenzymes, perhaps because the Mn^{2+} ion requires less rigid coordination than does the Mg^{2+} ion (30). This fact could explain how HP0268 readily recognizes and digests plasmid DNA with relaxed sequence specificity due to the Mn^{2+} ion. The metal dependence that was studied here could provide a clue to link the regulation of the nuclease activities to the environmental metal stress of *H. py-*

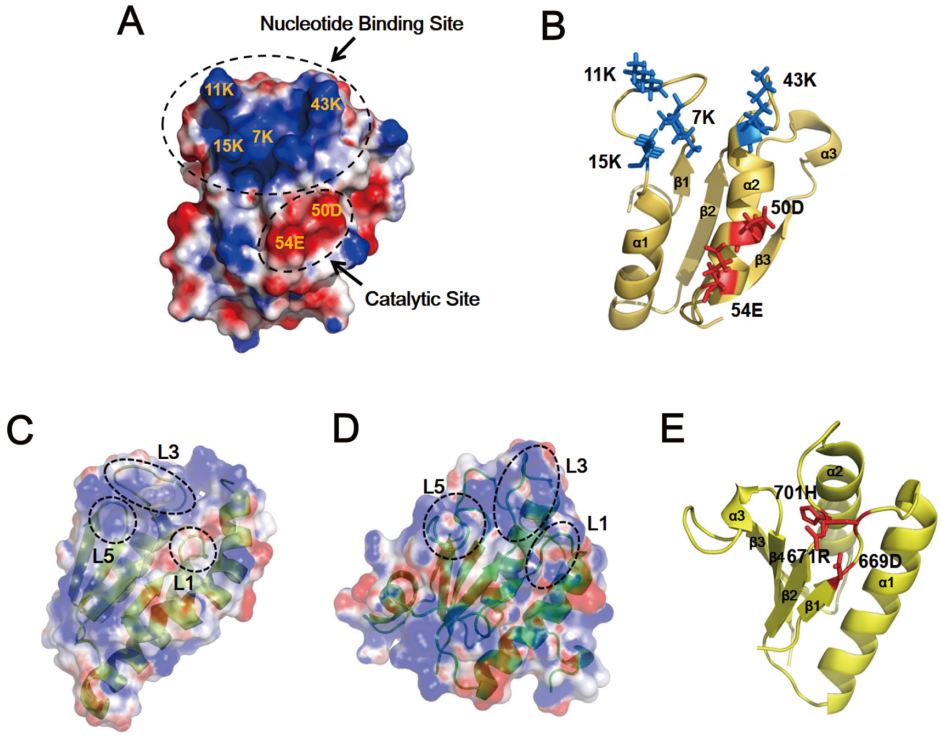

Figure 7. Comparison of the nucleotide binding and catalytic sites between HP0268 and its structural homologues. **(A)** Electrostatic potential surface of HP0268 with the location of the residues that are involved in nucleotide binding and catalysis. Dotted circles represent the nucleotide binding and catalytic sites, respectively. **(B)** Ribbon representation of (A). Blue and red colors represent the side-chains of the residues that are involved in nucleotide binding and catalysis, respectively. **(C, D)** The nucleotide binding sites of human B3BP-SMR domain and *E. coli* YDAL. These structures are shown as an overlay of the electrostatic potential surface and ribbon models. Dotted circles show loops 1, 2 and 3 (named L1, L3 and L5), which form the primary DNA binding region. **(E)** The catalytic site of the SMR domain from *T. thermophilus.* The catalytic site is predicted to encompass residues 669D, 671R, 701H whose side-chains are shown in red.

lori from ingested food sources in the stomach. The nuclease activities of HP0268 are not only affected by diverse metal ions but also stimulated in the absence of metal ion. It is possible that a small number of residues, including Asp50 and Glu54, form a less coordinated network with metal ions. This coordination might entail low affinity between HP0268 and the metal ion, giving rise to a low metal dependence of the nuclease activities.

We compared the nucleotide binding and catalytic sites of HP0268 with those of its structural homologues. Previous studies of human B3BP-SMR domain using mutagenesis and chemical shift perturbation suggested that loops L1, L5 and, in particular, L3, form the primary DNA binding region of the SMR domain of human B3BP (Figure 7C) (31). Based on this site, *E. coli* YDAL is assumed to have similar loops L1, L3 and L5 for DNA binding (Figure 7D) (26). These nucleotide binding sites commonly have the positive electrostatic potential that is required to bind the negatively charged phosphate backbone of nucleic acids. The catalytic residues of the B3BP-SMR domain and YDAL protein have not yet been identified. In the case of the SMR domain from *T. thermophilus*, the residues Asp669, Arg671 and His701 were proposed to be involved in nicking endonuclease activity on the basis of a structural comparison with *E. coli* RNase E and previous mutagenesis data (Figure 7E) (24,32).

The comparative analysis revealed that the distribution of the nucleotide binding and catalytic sites of HP0268 is different from the distribution in the structural homologues although the folds of the proteins are similar. Therefore, the molecular mechanism of nuclease activation might also be different. HP0268 is considered a hypothetical protein primarily because overall sequence conservation between HP0268 and functionally known proteins has not been found. Therefore, more careful and comprehensive analyses using bioinformatics were performed in a search for sequence conservation between the nucleotide binding, metal binding, and catalytic sites of HP0268 and many nucleases. However, well-defined sequence motifs for these sites are not conserved in HP0268, suggesting that this protein belongs to a new family of nucleases. The only finding was that the overall electrostatic properties are conserved in the basic nucleotide binding and acidic metal binding sites. We propose that HP0268 has become an evolutionary intermediate between RNases and nicking endonucleases during *H. pylori* adaptation to the extremely acidic environment of the stomach. This state might be accompanied by the unusual nature of nucleotide binding and active sites, although this protein shares an evolutionarily conserved fold with the SMR domain. Each activity might have been compromised by an incomplete bifunctional evolution with a size limitation. The findings of the present study will contribute to

our knowledge of the structure and functionality of sugar-nonspecific and single strand-specific nucleases.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the National Center for Inter-University Research Facilities (NCIRF) for the use of their NMR machines.

FUNDING

National Research Foundation of Korea (NRF) (MEST) [2012R1A2A1A01003569, 20110001207]; Korea Healthcare Technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea [A092006]; 2013 BK21 project for Medicine, Dentistry, and Pharmacy. Funding for open access charge: National Research Foundation of Korea (NRF); Korea Healthcare technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea.
Conflict of interest statement. None declared.

REFERENCES

1. Cover,T.L. and Blaser,M.J. (1996) Helicobacter pylori infection, a paradigm for chronic mucosal inflammation: pathogenesis and implications for eradication and prevention. *Adv. Intern. Med.*, **41**, 85–117.
2. Forman,D., Newell,D.G., Fullerton,F., Yarnell,J.W., Stacey,A.R., Wald,N. and Sitas,F. (1991) Association between infection with Helicobacter pylori and risk of gastric cancer: evidence from a prospective investigation. *BMJ*, **302**, 1302–1305.
3. Peek,R.M. Jr and Blaser,M.J. (2002) Helicobacter pylori and gastrointestinal tract adenocarcinomas. *Nat. Rev. Cancer*, **2**, 28–37.
4. Selgrad,M., Kandulski,A. and Malfertheiner,P. (2009) Helicobacter pylori: diagnosis and treatment. *Curr. Opin. Gastroenterol.*, **25**, 549–556.
5. Lochmannova,J. (2010) Current perspective of the resistance of Helicobacter pylori strains to antimicrobial drugs. *Klin. Mikrobiol. Infekc. Lek.*, **16**, 199–202.
6. Luo,H., Lin,Y., Gao,F., Zhang,C.T. and Zhang,R. (2014) DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res.*, **42**, D574–D580.
7. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
8. Johnson,B.A. and Blevins,R.A. (1994) NMR View: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
9. Guntert,P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.*, **278**, 353–378.
10. Shen,Y., Delaglio,F., Cornilescu,G. and Bax,A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*, **44**, 213–223.
11. Wishart,D.S. and Sykes,B.D. (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR*, **4**, 171–180.
12. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
13. Nederveen,A.J., Doreleijers,J.F., Vranken,W., Miller,Z., Spronk,C.A., Nabuurs,S.B., Guntert,P., Livny,M., Markley,J.L., Nilges,M. *et al.* (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins*, **59**, 662–672.
14. Laskowski,R.A., Rullmannn,J.A., MacArthur,M.W., Kaptein,R. and Thornton,J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.
15. Berman,H.M., Battistuz,T., Bhat,T.N., Bluhm,W.F., Bourne,P.E., Burkhardt,K., Feng,Z., Gilliland,G.L., Iype,L., Jain,S. *et al.* (2002) The Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 899–907.
16. Dietmann,S., Park,J., Notredame,C., Heger,A., Lappe,M. and Holm,L. (2001) A fully automatic evolutionary classification of protein folds: Dali Domain Dictionary version 3. *Nucleic Acids Res.*, **29**, 55–57.
17. Fukui,K., Kosaka,H., Kuramitsu,S. and Masui,R. (2007) Nuclease activity of the MutS homologue MutS2 from Thermus thermophilus is confined to the Smr domain. *Nucleic Acids Res.*, **35**, 850–860.
18. Kwon,A.R., Kim,J.H., Park,S.J., Lee,K.Y., Min,Y.H., Im,H., Lee,I. and Lee,B.J. (2012) Structural and biochemical characterization of HP0315 from Helicobacter pylori as a VapD protein with an endoribonuclease activity. *Nucleic Acids Res.*, **40**, 4216–4228.
19. Bauerova-Hlinkova,V., Dvorsky,R., Perecko,D., Povazanec,F. and Sevcik,J. (2009) Structure of RNase Sa2 complexes with mononucleotides–new aspects of catalytic reaction and substrate recognition. *FEBS J.*, **276**, 4156–4168.
20. Hsia,K.C., Li,C.L. and Yuan,H.S. (2005) Structural and functional insight into sugar-nonspecific nucleases in host defense. *Curr. Opin. Struct. Biol.*, **15**, 126–134.
21. Desai,N.A. and Shankar,V. (2003) Single-strand-specific nucleases. *FEMS Microbiol. Rev.*, **26**, 457–491.
22. Besnier,C.E. and Kong,H. (2001) Converting MlyI endonuclease into a nicking enzyme by changing its oligomerization state. *EMBO Rep.*, **2**, 782–786.
23. McConnell Smith,A., Takeuchi,R., Pellenz,S., Davis,L., Maizels,N., Monnat,R.J. Jr and Stoddard,B.L. (2009) Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 5099–5104.
24. Fukui,K. and Kuramitsu,S. (2011) Structure and Function of the Small MutS-Related Domain. *Mol. Biol. Int.*, **2011**, 691735.
25. Diercks,T., Ab,E., Daniels,M.A., de Jong,R.N., Besseling,R., Kaptein,R. and Folkers,G.E. (2008) Solution structure and characterization of the DNA-binding activity of the B3BP-Smr domain. *J. Mol. Biol.*, **383**, 1156–1170.
26. Gui,W.J., Qu,Q.H., Chen,Y.Y., Wang,M., Zhang,X.E., Bi,L.J. and Jiang,T. (2011) Crystal structure of YdaL, a stand-alone small MutS-related protein from Escherichia coli. *J. Struct. Biol.*, **174**, 282–289.
27. Rangarajan,E.S. and Shankar,V. (2001) Sugar non-specific endonucleases. *FEMS Microbiol. Rev.*, **25**, 583–613.
28. Cowan,J.A. (2002) Structural and catalytic chemistry of magnesium-dependent enzymes. *Biometals*, **15**, 225–235.
29. Maguire,M.E. and Cowan,J.A. (2002) Magnesium chemistry and biochemistry. *Biometals*, **15**, 203–210.
30. Yang,W., Lee,J.Y. and Nowotny,M. (2006) Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. *Mol. Cell*, **22**, 5–13.
31. Diercks,T., Ab,E., Daniels,M.A., de Jong,R.N., Besseling,R., Kaptein,R. and Folkers,G.E. (2008) Solution Structure and Characterization of the DNA-Binding Activity of the B3BP–Smr Domain. *J. Mol. Biol.*, **383**, 1156–1170.
32. Fukui,K., Nakagawa,N., Kitamura,Y., Nishida,Y., Masui,R. and Kuramitsu,S. (2008) Crystal Structure of MutS2 Endonuclease Domain and the Mechanism of Homologous Recombination Suppression. *J. Biol. Chem.*, **283**, 33417–33427.