Activity, stability, and structure of metagenome-derived LC11-RNase H1, a homolog of *Sulfolobus tokodaii* RNase H1

Tri-Nhan Nguyen,¹ Clement Angkawidjaja,^{1,2} Eiko Kanaya,¹ Yuichi Koga,¹ Kazufumi Takano,¹ and Shigenori Kanaya^{1*}

¹Department of Material and Life Science, Graduate School of Engineering, Osaka University, Suita, Osaka 565-0871, Japan

²International College, Osaka University, Toyonaka, Osaka 560-0043, Japan

Received 28 November 2011; Revised 3 February 2012; Accepted 6 February 2012

DOI: 10.1002/pro.2043

Published online 9 February 2012 proteinscience.org

Abstract: Metagenome-derived LC11-RNase H1 is a homolog of *Sulfolobus tokodaii* RNase H1 (Sto-RNase H1). It lacks a C-terminal tail, which is responsible for hyperstabilization of Sto-RNase H1. Sto-RNase H1 is characterized by its ability to cleave not only an RNA/DNA hybrid but also a double-stranded RNA (dsRNA). To examine whether LC11-RNase H1 also exhibits both RNase H and dsRNase activities, LC11-RNase H1 was overproduced in *Escherichia coli*, purified, and characterized. LC11-RNase H1 exhibited RNase H activity with similar metal ion preference, optimum pH, and cleavage mode of substrate with those of Sto-RNase H1. However, LC11-RNase H1 did not exhibit dsRNase activity at any condition examined. LC11-RNase H1 was less stable than Sto-RNases H1 and its derivative lacking the C-terminal tail (Sto-RNase H1 Δ C6) by 37 and 13°C in T_m , respectively. To understand the structural bases for these differences, the crystal structure of LC11-RNase H1 was determined at 1.4 Å resolution. The LC11-RNase H1 structure is highly similar to the Sto-RNase H1 structure. However, LC11-RNase H1 has two grooves on protein surface, one containing the active site and the other containing DNA-phosphate binding pocket, while Sto-RNase H1 has one groove containing the active site. In addition, LC11-RNase H1 contains more cavities and buried charged residues than Sto-RNase H1. We propose that LC11-RNase H1 does not exhibit dsRNase activity because dsRNA cannot fit to the two grooves on protein surface and that LC11-RNase H1 is less stable than Sto-RNase H1 Δ C6 because of the increase in cavity volume and number of buried charged residues.

Keywords: RNase H1; metagenome; crystal structure; stability; dsRNase activity

Additional Supporting Information may be found in the online version of this article.

Kazufumi Takano's current address is Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, 1-5 Hangi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan.

Grant sponsor: Ministry of Education, Culture, Sports, Science, and Technology of Japan; Grant number: 21380065; Grant sponsor: New Energy and Industrial Technology Development Organization (NEDO), Japan

*Correspondence to: Shigenori Kanaya, Department of Material and Life Science, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan. E-mail: kanaya@mls.eng.osaka-u.ac.jp

Introduction

RNase H is an endoribonuclease that specifically cleaves the RNA strand of RNA/DNA hybrids.¹ It hydrolyzes phosphodiester bonds of RNA to yield 3'-hydroxyl and 5'-phosphate groups with a two-metal-ion catalysis mechanism.² RNases H are classified into two major families, type 1 (RNase H1) and type 2 (RNases H2 and H3) RNases H, based on the difference in the amino acid sequences.^{3,4} These RNases H are functional in a monomeric form,⁴ except for eukaryotic RNases H2, which are functional only in a heterotrimeric form.⁵⁻⁷ RNases H are involved in DNA replication, repair and transcription.^{4,8-10} RNase H is also present in retroviruses, including human

immunodeficiency virus type-1 (HIV-1), as a C-terminal domain of reverse transcriptase (RT).³ Retroviral RNase H activity is required for viral proliferation and is therefore regarded as one of the targets for AIDS chemotherapy.¹¹ RNase H also plays a crucial role in antisense therapy.¹²

LC11-RNase H1 is one of the 12 RNases H1 enzymes (LC1 ~ LC12-RNases H1) isolated from leaf-branch compost by a metagenomic approach.¹³ It is a *Sulfolobus tokodaii* RNase H1 (Sto-RNase H1) homolog, which is phylogenetically different from bacterial RNases H1, eukaryotic RNases H1, and retroviral RNases H.¹⁴ LC11-RNase H1 is the smallest RNase H enzyme with 140 amino acid residues (molecular mass of 15.6 kDa) so far examined and shows the amino acid sequence identity of 34% to Sto-RNase H1. It lacks a C-terminal tail, which is not important for enzymatic function but is responsible for hyperstabilization of Sto-RNase H1.¹⁵ Sto-RNase H1 is unique in that it can cleave not only the RNA strand of an RNA/DNA hybrid but also the RNA strand of a double-stranded RNA (dsRNA).¹⁴ None of other RNases H, except for some retroviral RNases H, cleave dsRNA. Therefore, it would be informative to examine whether LC11-RNase H1 exhibits dsRNA-dependent RNase (dsRNase) activity.

In this report, we overproduced LC11-RNase H1 in *Escherichia coli* (*E. coli*), purified, and compared its activity, stability, and structure with those of Sto-RNase H1. Unlike Sto-RNase H1, LC11-RNase H1 did not exhibit dsRNase activity. LC11-RNase H1 was less stable than Sto-RNase H1 by 37°C in T_m . We discuss the structural features for these differences.

Results and Discussion

Overproduction and purification

We used the *rnhA/rnhB* double mutant strain *E. coli* MIC2067(DE3) as a host strain for overproduction of LC11-RNase H1 to avoid contamination of host-derived RNases H (RNases H1 and H2). Upon induction for overproduction, the recombinant protein accumulated in the cells in a soluble form. It was purified to give a single band on SDS-PAGE (data not shown). The amount of the protein purified from 1 L of culture was ~10 mg. The molecular mass of LC11-RNase H1 was estimated to be 16 kDa by SDS-PAGE and 18 kDa by gel filtration column chromatography. Both values are comparable to that calculated from the amino acid sequence (15.6 kDa), indicating that LC11-RNase H1 exists in a monomeric form, like other RNases H1.

Enzymatic activity

The dependencies of the LC11-RNase H1 activity on pH, salt, and metal ion were analyzed at 37°C by changing one of the conditions used for assay [10 mM Tris-HCl (pH 8.5), 10 mM NaCl, and 10 mM

MgCl₂]. The M13 DNA/RNA was used as a substrate. The enzymatic activity was determined at this temperature (37°C), which would be lower than the optimum one, because no substrates used in this study are fully stable at ≥50°C.

When the enzymatic activity was determined over the pH range of 4–11, LC11-RNase H1 exhibited the highest activity at pH 10 (Fig. S1 in the Supporting Information). However, the RNase H activity was determined at pH 8.5, at which LC11-RNase H1 exhibits ~80% of the maximal enzymatic activity, because the stability of the substrate and the solubility of the metal cofactor decrease as the pH increases beyond 9.0.

When the enzymatic activity was determined at various concentrations of NaCl, LC11-RNase H1 exhibited the highest activity in the presence of 10 mM NaCl (data not shown). Its activity gradually decreased as the salt concentration increased. As a result, LC11-RNase H1 exhibited 50 and 10% of the maximal activity in the presence of 0.4 and 1M NaCl, respectively. A similar result was obtained when KCl was used as a salt (data not shown).

When the enzymatic activity was determined in the presence of various concentrations of MgCl₂, LC11-RNase H1 exhibited the highest activity at 10 mM MgCl₂ (Fig. S2 in the Supporting Information). It also exhibited activity in the presence of MnCl₂, CoCl₂, NiCl₂, ZnCl₂, and CuCl₂ with the maximum at 10, 0.1, 10, 1, and 1 mM, respectively, but did not exhibit activity in the presence of CaCl₂ (Fig. S2 in the Supporting Information). The maximal Mn²⁺-dependent activity was slightly lower than but comparable to the maximal Mg²⁺-dependent activity, while the maximal Co²⁺-, Ni²⁺-, Zn²⁺-, and Cu²⁺-dependent activities were lower than the maximal Mg²⁺-dependent activity by 8–15 fold. The specific activities of LC11-RNase H1 and Sto-RNase H1 were determined to be 3.5 and 2.2 units mg⁻¹, respectively, in the presence of 10 mM MgCl₂, and 3.4 and 0.8 units mg⁻¹, respectively, in the presence of 10 mM MnCl₂. It has been reported that Sto-RNase H1 exhibits activity in the presence of Co²⁺ and Ni²⁺ as well, but does not exhibit activity in the presence of Cu²⁺, Zn²⁺, and Ca²⁺.¹⁴ The maximal Co²⁺- and Ni²⁺-dependent activities of Sto-RNase H1 are lower than its maximal Mg²⁺-dependent activity by 10–20 fold. Thus, LC11-RNase H1 shows a similar metal ion preference for activity to that of Sto-RNase H1, although its preference to Mg²⁺ is weaker than that of Sto-RNase H1.

Substrate and cleavage-site specificities

The substrate and cleavage-site specificities of LC11-RNase H1 were analyzed by using 12 bp RNA/DNA hybrid (R12/D12), 12 bp RNA/RNA duplex (R12/R12), 29 b RNA, 29 b DNA, 29 bp DNA₁₅-RNA₁-DNA₁₃/DNA duplex (D15-R1-D13/D29), and 18 bp RNA₉-DNA₉/DNA duplex (R9-D9/D18) as a

Figure 1. Cleavage of the 12 bp RNA/DNA substrate with LC11-RNase H1 and Sto-RNase H1. A: Separation of the hydrolysates by urea gel. The 5'-end labeled 12 bp RNA/DNA substrate was hydrolyzed by the enzyme at 37°C for 15 min and the hydrolysates were separated on a 20% polyacrylamide gel containing 7M urea. The concentration of the substrate was 1.0 μ M. The concentration of the enzyme in the reaction mixture is indicated above each lane. The sequence of the RNA strand of the substrate is indicated along the gel. B: Schematic representation of the sites and extents of cleavage by LC11-RNase H1 and Sto-RNase H1. Cleavage sites are shown by arrows. The differences in the lengths of the arrows reflect relative cleavage intensities at the position indicated. These lengths do not necessarily reflect the amount of the products accumulated upon complete hydrolysis of the substrate.

substrate. For comparative purposes, these substrates were cleaved by Sto-RNase H1 as well. D15-R1-D13 is the chimeric oligonucleotide in which single ribonucleotide is flanked by 15 and 13 b DNAs at the 5' and 3' sides, respectively. R9-D9/D18 is an Okazaki fragment-like substrate, in which the 18 b RNA₉-DNA₉ chimeric oligonucleotide is hybridized to the 18 b complementary DNA.

Cleavage of the R12/D12 substrate with LC11-RNase H1 and Sto-RNase H1 is summarized in Figure 1. Both proteins cleaved this substrate at multiple sites, but with different cleavage-site specificities. LC11-RNase H1 cleaved this substrate most preferably at a6–u7 and a9–c10 in the presence of 10 mM MgCl₂ and nearly equally at all sites between g5 and g12 in the presence of 10 mM MnCl₂, whereas Sto-RNase H1 cleaved this substrate at all sites between a4 and g12, most preferably at g5–a6, in the presence of 10 mM MgCl₂ and most preferably at a4–g5 and less preferably at g3–a4, u7–g8, c10–g11, and g11–g12 in the presence of 10 mM MnCl₂.

Cleavage of the R9-D9/D18 substrate with LC11-RNase H1 and Sto-RNase H1 in the presence of 10 mM MgCl₂ is summarized in Figure 2. Both proteins cleaved this substrate at multiple sites, but with different cleavage-site specificities. LC11-RNase H1 preferably cleaved this substrate at three sites (a5–u6, u6–g7, and g7–c8) with similar efficiencies, whereas Sto-RNase H1 preferably cleaved this substrate at five sites with different efficiencies (most preferably at a5–u6 and u6–g7 and less preferably at g7–c8, c8–c9, and c9–T10). This result is consistent with the previous one that Sto-RNase H1 can

cleave the (5')RNA-DNA(3') junction of the Okazaki fragment-like substrate.¹⁴ The cleavage-site specificities of both proteins for this substrate were not seriously changed in the presence of 10 mM MnCl₂ (data not shown). These results indicate that LC11-

Figure 2. Cleavage of the 18 bp RNA-DNA/DNA substrate with LC11-RNase H1 and Sto-RNase H1. The 3'-end labeled 18 bp RNA₉-DNA₉/DNA substrate was hydrolyzed by the enzyme in the presence of 10 mM MgCl₂ and the hydrolysates were separated by urea gel as in Figure 1(A). The partial sequence of the RNA₉-DNA₉ strand of the substrate is indicated along the gel, in which deoxyribonucleotides are indicated by capital letters and ribonucleotides are indicated by lowercase letters. The substrate and the products containing number one and two ribonucleotides at the 5'-termini are schematically shown on the right side of the gel. The asterisk indicates the fluorescence-labeled site.

Figure 3. Cleavage of the 12 bp RNA/RNA substrate with LC11-RNase H1 and Sto-RNase H1. A: The 5'-end labeled 12 bp RNA/RNA substrate was hydrolyzed by the enzyme in the presence of 10 mM MnCl_2 and the hydrolysates were separated by urea gel as in Figure 1(A). B: The sites and extents of cleavage by LC11-RNase H1 and Sto-RNase H1 are schematically shown as in Figure 1(B).

RNase H1 cannot cleave the (5')RNA-DNA(3') junction. It is noted that Sto-RNase H1 can cleave the (5')RNA-DNA(3') junction, but very poorly even after extensive hydrolysis, suggesting that its RNase H activity is not sufficient for complete removal of RNA from Okazaki fragment *in vivo*.

Cleavage of the R12/R12 substrate with LC11-RNase H1 and Sto-RNase H1 in the presence of 10 mM MnCl_2 is summarized in Figure 3. Sto-RNase H1 cleaved this substrate at all sites between g5 and g12, most preferably at a9–c10, as reported previously,¹⁴ whereas LC11-RNase H1 did not cleave this substrate. Neither LC11-RNase H1 nor Sto-RNase H1 cleaved this substrate in the presence of 10 mM MgCl_2 (data not shown). These results indicate that, unlike Sto-RNase H1, LC11-RNase H1 does not have an ability to cleave double-stranded (ds) RNA.

Both LC11-RNase H1 and Sto-RNase H1 did not cleave D15-R1-D13/D29, 29 b RNA, and 29 b DNA either in the presence of 10 mM MgCl_2 or MnCl_2 , indicating that both proteins do not have an ability to cleave dsDNA, single ribonucleotide-embedded dsDNA, single-stranded (ss) RNA, and ssDNA.

Thermal stability

Heat-induced unfolding of LC11-RNase H1, Sto-RNase H1, and the Sto-RNase H1 derivative lacking the C-terminal six-residue tail (Sto-RNase H1ΔC6) was analyzed by differential scanning calorimetry (DSC) at pH 3.0. All DSC curves were reproduced by repeating thermal scans, indicating that thermal unfolding of these proteins is reversible. The denaturation curves of all proteins showed single transitions (Fig. 4). The peak of this curve represents the melting temperature (T_m) of the protein, which was 64°C for LC11 RNase H1, 101°C for Sto-RNase H1, and 77°C for Sto-RNase H1ΔC6. The T_m values of Sto-

RNase H1 and Sto-RNase H1ΔC6 are identical with those previously reported.¹⁵ Thus, LC11-RNase H1 is less stable than Sto-RNase H1 and Sto-RNase H1ΔC6 by 37°C and 13°C in T_m , respectively. These results indicate that destabilization by removal of the C-terminal tail accounts for two thirds of the difference in T_m between LC11-RNase H1 and Sto-RNase H1.

Crystal structure

The crystal structure of LC11-RNase H1 was determined at a resolution of 1.4 Å. The asymmetric unit of the crystal structure consists of four-protein molecule (A–D). The structures of these four protein molecules are virtually identical with one another with root-mean-square deviation (RMSD) values of 1.00 Å between molecules A and D, 1.04 Å between molecules B and D, and 0.50 Å between molecules C and D for 137 C α atoms. We used the structure of molecule D in this study.

The overall structure of LC11-RNase H1 consists of a five-stranded mixed β -sheet with three antiparallel (β 1– β 3) and two parallel (β 4 and β 5) strands and four α -helices (α A, α B, α D, and α E) [Fig. 5(A,B)]. This structure is highly similar to that of Sto-RNase H1¹⁵ with the RMSD value of 1.49 Å for 126 C α atoms [Fig. 5(A,B)]. However, this structure lacks a C-terminal tail, which is responsible for hyperstabilization of Sto-RNase H1.¹⁵ In addition, several loops, especially that between β 3-strand and α A-helix vary in size for these proteins.

The amino acid sequence of LC11-RNase H1 is compared with those of Sto-RNase H1, Bh-RNase H1, HIV-1 RNase H, Hs-RNase H1, and Ec-RNase H1 on the bases of their crystal structures in Figure 6. The sequence of LC11-RNase H1 is 34% identical to Sto-RNase H1, 16% identical to *Bacillus halodurans* RNase H1 (Bh-RNase H1), 22% identical to HIV-1 RNase H, 24% identical to human (*Homo sapiens*)

Figure 4. DSC curves. The DSC curves of LC11-RNase H1 (thick solid line), Sto-RNase H1ΔC6 (dashed line), and Sto-RNase H1 (thin solid line) were measured at a scan rate of 1°C min⁻¹. These proteins were dissolved in 20 mM Gly-HCl (pH 3.0) at ~0.5 mg mL⁻¹.

Figure 5. Crystal structure of LC11-RNase H1. A and B: Stereoview of the structure of LC11-RNase H1 superimposed on that of Sto-RNase H1 (PDB code 2EHG). The LC11-RNase H1 and Sto-RNase H1 structures are colored green and salmon, respectively. Five active-site residues of LC11-RNase H1 (Asp9, Glu49, Asp77, Arg121, and Asp128) and the corresponding residues of Sto-RNase H1 (Asp7, Glu52, Asp76, Arg118, and Asp125) are shown by stick models. The C-terminal tail of Sto-RNase H1 is colored red. NT and CT represent N- and C-termini. The structures shown in (A) and (B) differ in the view direction. C: Stereoview of the active-site structures of RNases H. The side chains of the active-site residues of LC11-RNase H1 (green) and Sto-RNase H1 (salmon) are superimposed onto those in the cocrystal structure of Bh-RNase H with the substrate and metal ions (orange, PDB code 1ZBI). The positions of the RNA strand of the substrate with the scissile phosphate group between R(−1) and R(0) and two metal ions A and B are also shown. D: A model for the complex between LC11-RNase H1 and RNA/DNA hybrid. Negative and positive surface potentials of LC11-RNase H1 are shown in red and blue, respectively. The electrostatic potential value ranges from -15 to $+15$ kT e^{-1} . The RNA and DNA strands of RNA/DNA hybrid are shown in magenta and cyan, respectively. The DNA phosphate group located two-base pairs away from the scissile bond is shown in yellow. The highly negatively charged region represents the active site. E: The structures around the DNA binding grooves containing the phosphate binding pockets. A model of the LC11-RNase H1-substrate complex, the structure of the Bh-RNase H1-substrate complex, and a model of the Sto-RNase H1-substrate complex are shown, where LC11-RNase H1, Bh-RNase H1, and Sto-RNase H1 are colored gray, orange, and salmon, respectively. F: Stereoview of the core structure of LC11-RNase H1 superimposed on that of Sto-RNase H1ΔC6 (PDB code 3ALY). The amino acid residues located at the hydrophobic cores of LC11-RNase H1 (green) and Sto-RNase H1ΔC6 (salmon) are shown by stick models. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 6. Alignment of the amino acid sequences. The amino acid sequences of LC11-RNase H1 (LC11), Sto-RNase H1 (Sto), Bh-RNase H1 (Bh), HIV-1 RT RNase H (HIV-1), Hs-RNase H1 (Hs), and Ec-RNase H1 (Ec) are compared with one another on the basis of their crystal structures. The ranges of the secondary structures of LC11-RNase H1 and Ec-RNase H1 are shown above and below the sequences, respectively. The amino acid residues that are not located within the α -helices or β -strands are shown in italic. The amino acid residues that are conserved in at least four different proteins are highlighted in black. The positions of the five active-site residues are indicated by arrows. The amino acid residues that contact with the RNA and DNA strands of the substrate in the cocrystal structure of Bh-RNase H1 with the substrate are indicated by the open and closed circles above the sequences, respectively. Gaps are denoted by dashes. The numbers represent the positions of the amino acid residues relative to the initiator methionine for each protein.

RNase H1 (Hs-RNase H), and 26% identical to *E. coli* RNase H1 (Ec-RNase H1).

Catalytic center

The four acidic residues, Asp9, Glu49, Asp77, and Asp128, form the active site of LC11-RNase H1. The corresponding residues are Asp7, Glu52, Asp76, and Asp125 for Sto-RNase H1, and Asp71, Glu109, Asp132, and Asp192 for Bh-RNase H1. Superimposition of the active-site structures of LC11-RNase H1 and Sto-RNase H1 on the cocrystal structure of Bh-RNase H1 with the substrate and metal cofactor¹⁶ indicates that the steric configuration of the active-site residues of LC11-RNase H1 is very similar to those of Sto-RNase H1 and Bh-RNase H1 [Fig. 5(C)]. Therefore, these four acidic residues of LC11-RNase H1 probably form the metal binding sites A and B, as seen in the cocrystal structure of Bh-RNase H1. In addition, Arg121 of LC11-RNase H1 is located at the similar position where Arg118 of Sto-RNase H1 and Glu188 of Bh-RNase H1 are located [Fig. 5(C)]. This residue probably promotes a product release by perturbing the coordination of the metal ion A, as proposed for Glu188 of Bh-RNase H1.¹⁶

Substrate binding grooves on protein surface

According to the cocrystal structures of Bh-RNase H1¹⁶ and Hs-RNase H1¹⁷ with the RNA/DNA sub-

strate, these enzymes bind to the minor groove of the substrate by two grooves on protein surface, such that one groove containing the active site interacts with the RNA backbone and the other groove containing the phosphate-binding pocket interacts with the DNA backbone. Hs-RNase H1 has a DNA binding channel in the basic protrusion in addition to these grooves. dsRNA may not fit to these grooves and channel on protein surface because of its wide-groove A-form conformation, and therefore may not be recognized as a substrate by Bh-RNase H1 and Hs-RNase H1. This may be the reason why these enzymes cannot cleave dsRNA.

To examine whether LC11-RNase H1 also has these grooves on protein surface, a model for the complex between LC11-RNase H1 and the substrate (12 bp RNA/DNA hybrid) was constructed based on the cocrystal structure of Bh-RNase H1 with the substrate [Fig. 5(D,E)]. According to this model, RNA and DNA binding grooves are present on the surface of LC11-RNase H1. Lys92 and Glu93 apparently form humps, which separate inside and outside of the DNA-binding groove. These humps may interfere with the binding of LC11-RNase H1 to dsRNA. A similar DNA binding groove is present on the surface of Bh-RNase H1 [Fig. 5(E)]. Lys146 and Thr148 form humps as Lys92 and Glu93 of LC11-RNase H1 do.

However, according to a model for the Sto-RNase H1-substrate complex, a DNA-binding groove is not clearly detected on protein surface [Fig. 5(E)]. Lys92 of LC11-RNase H1 and Lys146 of Bh-RNase H1 are conserved as Lys91 in Sto-RNase H1. However, this residue does not form a clear hump to separate inside and outside of the DNA-binding groove. HIV-1 RNase H also does not have this DNA-binding groove (data not shown). These results suggest that Sto-RNase H1 and HIV-1 RNase H can cleave dsRNA because they do not have a sharply-defined DNA-binding groove and therefore can accommodate dsRNA. However, it has been reported that the mutation of Lys91 of Sto-RNase H1 to Ala significantly reduces the dsRNase activity without seriously affecting the RNase H activity.¹⁵ This result suggests that electrostatic interactions are responsible for efficient binding of Sto-RNase H1 to dsRNA. We propose that only the RNase H or RNase H-fold containing proteins, which do not have a clear DNA binding groove but have positively charged residues near the DNA binding site, can cleave dsRNA.

It is noted that the structure of LC11-RNase H1 is highly similar to that of a PIWI domain of Argonaute, a key catalytic component of the RNA-induced silencing complex.^{18,19} According to the crystal structure of the complex between *Thermus thermophilus* Argonaute and a substrate,¹⁸ and a model for it between *Neurospora crassa* QDE-2 MID-PIWI lobe and a substrate,¹⁹ the PIWI domains of these Argonautes also do not have a clear binding groove for the guide strand and have positively charged residues at the guide strand binding site.

Structural features for the difference in stability between LC11-RNase H1 and Sto-RNase H1

Comparison of the crystal structure of LC11-RNase H1 with that of Sto-RNase H1ΔC6²⁰ indicates that the number of buried nonpolar residues and the number of ion pairs of LC11-RNase H1 are comparable to those of Sto-RNase H1ΔC6, while the number of buried charged residues and the total volume of cavities of LC11-RNase H1 are higher than those of Sto-RNase H1ΔC6 (Table I). These results suggest that LC11-RNase H1 is destabilized as compared to Sto-RNase H1 not only by removal of the C-terminal anchoring tail but also by increasing the cavity volume and the number of buried charged residues. It has been reported that increased cavity volume^{21,22} and burial of non-hydrogen-bonded polar or charged groups²³ contribute to protein instability.

The difference in the cavity volume at the hydrophobic core between LC11-RNase H1 and Sto-RNase H1ΔC6 is shown in Figure 5(F) as a typical example. Calculation of the cavity volume using the program CASTp²⁴ indicates that two cavities exist around Thr8 and Val74 of LC11-RNase H1. The total volume of these cavities is 158 Å³. These cavities do not exist in the hydrophobic core of Sto-RNase

Table 1. Structural Features of LC11-RNase H1 and Sto-RNase H1ΔC6

	LC11-RNase H1	Sto-RNase H1ΔC6
No. of amino acids	140	141
<i>T_m</i> (°C)	64	77
Content of residues (%)		
Buried charged	18	10
Buried nonpolar	48	49
No. of ion pairs (< 5.0 Å)	15	12
Total volume of cavities (Å ³)	223	20

H1ΔC6, because these cavities are filled by the amino acid substitutions from Thr8 to Phe6, Ile81 to Val80, Leu57 to Met60, and Val74 to Ile73.

Materials and Methods

Cells and plasmids

E. coli MIC2067 [F[−], λ[−], IN(*rnhD*−*rnhE*)1, *rnhA*399:cat, *rnhB*716:kam] was kindly donated by M. Itaya.¹⁴ The *E. coli* strain MIC2067(DE3),¹⁵ plasmid pBR322-LC11 containing the gene encoding LC11-RNase H1,¹³ and plasmids pET-Sto²⁵ and pET-StoΔC6¹⁵ for overproduction of Sto-RNase H1 and Sto-RNase H1ΔC6, respectively, were previously constructed in our laboratory. Plasmid pET25b was obtained from Novagen.

Plasmid construction, overproduction, and purification

Plasmid pET-LC11 for overproduction of LC11-RNase H1 was constructed by amplifying the gene encoding LC11-RNase H1 by PCR and ligating it into the *Nde*I-*Bam*HI sites of pET25b. Plasmid pBR322-LC11 was used as a template. All DNA oligomers for PCR were synthesized by Hokkaido System Science. PCR was performed with a GeneAmp PCR system 2400 (Applied Biosystems). The DNA sequences were confirmed by a Prism 310 DNA sequencer (Applied Biosystems).

Overproduction of LC11-RNase H1 using *E. coli* MIC2067(DE3) transformants with pET-LC11, disruption of the cells by sonication lysis, and centrifugation were performed as described previously.²⁶ The supernatant was collected, dialyzed against 10 mM sodium phosphate (pH 7.0), and applied to a column (5 mL) of DE52 (GE Healthcare) equilibrated with the same buffer. The protein was eluted from the column with a linear gradient of NaCl from 0 to 1M. The fractions containing the protein were collected, dialyzed against 20 mM Tris-HCl (pH 9.0), and loaded onto a HiTrap Q column (GE Healthcare) equilibrated with the same buffer. The protein was eluted from the column with a linear gradient of NaCl from 0 to 1M. The fractions containing the protein were collected, dialyzed against 20 mM Tris-HCl (pH 7.5) containing 50 mM NaCl, and loaded onto the Superdex 200 column (GE Healthcare)

equilibrated with the same buffer. The fractions containing the protein were collected and dialyzed against 10 mM Tris-HCl (pH 7.5). Overproduction of Sto-RNase H1 and Sto-RNase H1ΔC6 using *E. coli* MIC2067(DE3) transformants with pET-Sto and its derivatives and purification of these proteins were performed as described previously.²⁵

The purity of the protein was analyzed by SDS/PAGE on a 15% polyacrylamide gel,²⁷ followed by staining with Coomassie Brilliant Blue. The protein concentration was determined from UV absorption using an A₂₈₀ value of 0.30 for LC11-RNase H1, 0.97 for Sto-RNase H1, and 1.0 for Sto-RNase H1ΔC6 for a 0.1% (0.1 mg mL⁻¹) solution. These values were calculated by using ϵ of 1576 M⁻¹ cm⁻¹ for Tyr and 5225 M⁻¹ cm⁻¹ for Trp at 280 nm.²⁸

Enzymatic activity

The RNase H activity was determined by measuring the amount of the acid-soluble digestion product from the substrate, ³H-labeled M13 DNA/RNA hybrid, accumulated upon incubation at 37°C for 15 min, as described previously.²⁶ The buffer used for assay was 10 mM Tris-HCl (pH 8.5) containing 10 mM MgCl₂, 10 mM NaCl, 1 mM 2-mercaptoethanol, and 50 μg mL⁻¹ bovine serum albumin. One unit was defined as the amount of enzyme producing 1 μmol of acid-soluble material per min at 37°C. The specific activity was defined as the enzymatic activity per milligram of protein.

For cleavage of the oligomeric substrates, 5'-fluorescein-labeled 12 bp RNA/DNA hybrid (R12/D12), 12 bp RNA/RNA duplex (R12/R12), 29 bp DNA₁₅-RNA₁-DNA₁₃/DNA duplex (D15-R1-D13/D29), 29 b RNA and 29 b DNA, and 3'-fluorescein-labeled 18 bp RNA₉-DNA₉/DNA duplex (R9-D9/D18) were prepared as described previously.²⁶ Hydrolysis of the substrate at 37°C for 15 min and separation of the products on a 20% polyacrylamide gel containing 7M urea were carried out as described previously.²⁶ The products were detected by Typhoon 9240 Imager (GE Healthcare). The reaction buffers were the same as those for the hydrolysis of the M13 DNA/RNA hybrid. The products were identified by comparing their migration on the gel.

Differential scanning calorimetry (DSC)

The DSC measurement was performed on a high sensitivity VP-DSC and controlled by the VPVIEWERTM software package (Microcal) at scan rate of 1°C min⁻¹. The proteins were dissolved in 20 mM Gly-HCl (pH 3.0) at a concentration of ~0.5 mg mL⁻¹. Before performing the measurement, the proteins were degassed in a vacuum for 5 min.

Crystallization

The crystallization conditions were initially screened using crystallization kits from Hampton Research (Crystal Screens I and II) and Emerald Biostruc-

Table 2. Data Collection and Refinement Statistics

Data collection	
Wavelength (Å)	0.9
Space group	P21
Unit cell a, b, c (Å)	54.3, 104.9, 57.2
α, β, γ (°)	90.0, 110.3, 90.0
Resolution (Å)	50.00 - 1.40 (1.42 - 1.40)
Observations	1,394,496
Unique reflections	115,103
Completeness (%)	97.4 (85.8) ^a
R _{merge} (%) ^b	6.6 (45.1) ^a
Average I/s (I)	31.7 (2.8) ^a
Refinement statistics	
Resolution limits (Å)	47.76 - 1.40
No. of molecules	4
No. of protein atoms	4434
No. of water atoms	904
R _{work} (%)	17.2
R _{free} (%) ^c	19.2
Deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.061
Mean B factors (Å ²)	
Protein atoms	13.2
Water atoms	24.6

^a Values in parentheses are for the highest resolution shell.

^b $R_{\text{merge}} = \sum |I_{hkl} - \langle I_{hkl} \rangle| / \sum I_{hkl}$, where I_{hkl} is the intensity measurement for reflection with indices hkl and $\langle I_{hkl} \rangle$ is the mean intensity for multiply recorded reflections.

^c R_{free} was calculated using 5% of the total reflections chosen randomly and omitted from refinement.

tures (Wizard I and II). The conditions were surveyed using sitting-drop vapor-diffusion method at 4°C. Drops were prepared by mixing 1 μL each of protein solution and reservoir solution and were vapor-equilibrated against 100 μL reservoir solution. Native LC11-RNase H1 crystals appeared after 3 weeks using Wizard I solution No. 20 (0.1M imidazole pH 8, 0.2M NaCl, 0.4M NaH₂PO₄, and 1.6M K₂HPO₄). The crystallization conditions were further optimized and a few crystals appeared when the drop was prepared by mixing 1.5 μL protein solution at the concentration of 10 mg mL⁻¹ and 1.5 μL reservoir solution at 4°C. Despite the same crystallization conditions, only three crystals were obtained from 48 crystallization wells.

X-ray diffraction data collection and structure determination

X-ray diffraction data set of LC11-RNase H1 was collected at -173°C using synchrotron radiation on the BL44XU station at SPring-8 (Hyogo, Japan). All data sets were indexed, integrated, and scaled using the HKL2000 program suite.²⁹ The structure was solved by the molecular replacement method using MOLREP³⁰ in the CCP4 program suite. The crystal structure of RNase H1 from *Chlorobium tepidum* (Protein Data Bank entry 3H08) was used as a starting

model. The statistics for data collection and refinement are summarized in Table II. The figures were prepared by PyMol (<http://www.pymol.org/>).

Accession number

The coordinates and structure factors for LC11-RNase H1 have been deposited in the Protein Data Bank under accession codes 3U3G.

Acknowledgments

The synchrotron radiation experiments were performed at beamlines BL38B1 and BL44XU in the SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI; Proposal numbers 2009A1357 and 2010A6915).

References

- Crouch RJ, Dirksen M-L, Ribonucleases H. In: Linn SM, Roberts RJ, Eds. (1982) Nuclease. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp 211–241.
- Yang W, Lee JY, Nowotny M (2006) Making and breaking nucleic acids: two-Mg²⁺-ion catalysis and substrate specificity. *Mol Cell* 22:5–13.
- Ohtani N, Haruki M, Morikawa M, Kanaya S (1999) Molecular diversities of RNases H. *J Biosci Bioeng* 88: 12–19.
- Tadokoro T, Kanaya S (2009) Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes. *FEBS J* 276: 1482–1493.
- Figiel M, Chon H, Cerritelli SM, Cybulska M, Crouch RJ, Nowotny M (2011) The structural and biochemical characterization of human RNase H2 complex reveals the molecular basis for substrate recognition and Aicardi-Goutieres syndrome defects. *J Biol Chem* 286:10540–10550.
- Shaban NM, Harvey S, Perrino FW, Hollis T (2010) The structure of the mammalian RNase H2 complex provides insight into RNA/DNA hybrid processing to prevent immune dysfunction. *J Biol Chem* 285:3617–3624.
- Reijns MA, Bubeck D, Gibson LC, Graham SC, Baillie GS, Jones EY, Jackson AP (2011) The structure of the human RNase H2 complex defines key interaction interfaces relevant to enzyme function and human disease. *J Biol Chem* 286:10530–10539.
- Cerritelli SM, Frolova EG, Feng C, Grinberg A, Love PE, Crouch RJ (2003) Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell* 11:807–815.
- Nick McElhinny SA, Kumar D, Clark AB, Watt DL, Watts BE, Lundström EB, Johansson E, Chabes A, Kunkel TA (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol* 6:774–781.
- Kim N, Huang SN, Williams JS, Li YC, Clark AB, Cho JE, Kunkel TA, Pommier Y, Jinks-Robertson S (2011) Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science* 332:1561–1564.
- Tramontano E, Di Santo R (2010) HIV-1 RT-associated RNase H function inhibitors: recent advances in drug development. *Curr Med Chem* 17:2837–2853.
- Dean NM, Bennett CF (2003) Antisense oligonucleotide-based therapeutics for cancer. *Oncogene* 22:9087–9096.
- Kanaya E, Sakabe T, Nguyen NT, Koikeda S, Koga Y, Takano K & Kanaya S (2010) Cloning of the RNase H genes from a metagenomic DNA library: identification of a new type 1 RNase H without a typical active-site motif. *J Appl Microbiol* 109:974–983.
- Ohtani N, Yanagawa H, Tomita M, Itaya M (2004) Cleavage of double-stranded RNA by RNase HI from a thermoacidophilic archaeon, *Sulfolobus tokodaii* 7. *Nucleic Acids Res* 32:5809–5819.
- You DJ, Chon H, Koga Y, Takano K, Kanaya S (2007) Crystal structure of type 1 ribonuclease H from hyperthermophilic archaeon *Sulfolobus tokodaii*: role of arginine 118 and C-terminal anchoring. *Biochemistry* 46:11494–11503.
- Nowotny M, Gaidamakov SA, Crouch RJ, Yang W (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* 121:1005–1016.
- Nowotny M, Gaidamakov SA, Ghirlando R, Cerritelli SM, Crouch RJ, Yang W (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol Cell* 28:513–513.
- Wang Y, Juranek S, Li H, Sheng G, Wardle GS, Tuschl T, Patel DJ (2009) Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* 461:754–761.
- Boland A, Huntzinger E, Schmidt S, Izaurralde E, Weichenrieder O (2011) Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. *Proc Natl Acad Sci U S A* 108:10466–10471.
- Takano K, Okamoto T, Okada J, Tanaka S, Angkawidjaja C, Koga Y, Kanaya S (2011) Stabilization by fusion to the C-terminus of hyperthermophile *Sulfolobus tokodaii* RNase HI: a possibility of protein stabilization tag. *PLoS One* 6:e16226.
- Ishikawa K, Nakamura H, Morikawa K, Kanaya S (1993) Stabilization of *Escherichia coli* ribonuclease HI by cavity-filling mutations within a hydrophobic core. *Biochemistry* 32:6171–6178.
- Kadonosono T, Chatani E, Hayashi R, Moriyama H, Ueki T (2003) Minimization of cavity size ensures protein stability and folding: structures of Phe46-replaced bovine pancreatic RNase A. *Biochemistry* 42:10651–10658.
- Pace CN, Fu HL, Fryar KL, Landua J, Trevino SR, Shirley BA, Hendricks MM, Iimura S, Gajiwala K, Scholtz JM, Grimsley GR (2011) Contribution of hydrophobic interactions to protein stability. *J Mol Biol* 408:514–528.
- Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res* 34:W116–118 (Web Server issue).
- You DJ, Chon H, Koga Y, Takano K, Kanaya S (2006) Crystallization and preliminary crystallographic analysis of type 1 RNase H from the hyperthermophilic archaeon *Sulfolobus tokodaii* 7. *Acta Crystallogr Sect F* 62:781–784.
- Jongruja N, You D-J, Kanaya E, Koga Y, Takano K, Kanaya S (2010) The N-terminal hybrid binding domain of RNase HI from *Thermotoga maritima* is important for substrate binding and Mg²⁺-dependent activity. *FEBS J* 277:4474–4489.
- Laemmli UK (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nature* 227:680–685.
- Goodwin TW, Morton RA (1946) The spectrophotometric determination of tyrosine and tryptophan in proteins. *Biochem J* 40:628–32.
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromol Crystallogr Pt A* 276:307–326.
- Vagin A, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* 30:1022–1025.