Importance of an N-Terminal Extension in Ribonuclease HII from *Bacillus stearothermophilus* for Substrate Binding

AYUMU MUROYA,1 RIKITA NAKANO,1 NAOTO OHTANI,1 MITSURU HARUKI,1 MASAAKI MORIKAWA,1* AND SHIGENORI KANAYA1

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

Received 29 October 2001/Accepted 20 November 2001

The gene encoding ribonuclease HII from *Bacillus stearothermophilus* was cloned and expressed in *Escherichia coli*. The overproduced protein, Bst-RNase HII, was purified and biochemically characterized. Bst-RNase HII, which consists of 259 amino acid residues, showed the highest amino acid sequence identity (50.2%) to *Bacillus subtilis* RNase HIII. Like *B. subtilis* RNase HII, it exhibited Mn2+-dependent RNase H activity. It was, however, more thermostable than *B. subtilis* RNase HII. When the Bst-RNase HII amino acid sequence is compared with that of *Thermococcus kodakaraensis* RNase HII, to which it shows 29.8% identity, 30 residues are observed to be truncated from the C-terminus and there is an extension of 71 residues at the N-terminus. The C-terminal truncation results in the loss of the α9 helix, which is rich in basic amino acid residues and is therefore important for substrate binding. A truncated protein, Δ59-Bst-RNase HII, in which most of the N-terminal extension was removed, completely lost its RNase H activity. Surface plasmon resonance analysis indicated that this truncated protein did not bind to the substrate. These results suggest that the N-terminal extension of Bst-RNase HII is important for substrate binding. Because *B. subtilis* RNase HII has an N-terminal extension of the same length and these extensions contain a region in which basic amino acid residues are clustered, the *Bacillus* enzymes may represent a novel type of RNase H which possesses a substrate-binding domain at the N-terminus.

[Key words: RNase H, *Bacillus stearothermophilus*, gene cloning, N-terminal truncation, substrate binding]

Ribonuclease H (RNase H) cleaves the P-O3' bond of the RNA portion of DNA/RNA hybrids (1). The enzyme, which is universally present in diverse organisms, is thought to be involved in the removal of RNA primers from Okazaki fragments and of R-loops associated with transcription, but its physiological functions are not yet fully understood (2). Based on differences in their amino acid sequences, RNases H are classified into two major families, Type 1 and Type 2 (3, 4). The Type 1 enzymes can be further divided into bacterial RNases HI, eukaryotic RNases HI, and the RNase H domains of reverse transcriptases, among which *Escherichia coli* RNase H1 (5, 6) and the RNase H domain of HIV-1 reverse transcriptase (7) have been most extensively studied in terms of their structures and functions. The Type 2 enzymes can be further divided into bacterial RNases HII and HIII, eukaryotic RNases H2, and archaeal RNases HIII, among which the crystal structures of some archaeal RNases HII have thus far been determined (8-10). Because these structures highly resemble one another, we selected and studied the RNase HII from *Thermococcus kodakaraensis* as a representative of the archaeal RNases HII. Despite having poor amino acid sequence similarity, *T. kodakaraensis* RNase HII (8) and *E. coli* RNase HII (11, 12) share a main chain fold consisting of a five-stranded β-sheet and two α-helices. Also, the geometrical arrangement of the four acidic active-site residues (Asp7, Glu8, Asp105, and Asp135 for *T. kodakaraensis* RNase HII; Asp10, Glu48, Asp70 and Asp134 for *E. coli* RNase HII) is similar in these two proteins. These findings strongly suggest that the Type 1 and Type 2 enzymes share a common catalytic mechanism. According to a general acid-base mechanism proposed for *E. coli* RNase HII (3), one divalent metal cation is required for activity and the hydroxyl ion, which attacks the phosphate group for the RNA cleavage, is activated by an amino acid residue.

Because of similarities in their amino acid sequences and biochemical properties, bacterial and archaeal RNases HII are envisaged to share a common three-dimensional structure (13). However, RNases HII from mesophilic bacteria have relatively long N-terminal extensions as compared to their archaeal counterparts (4). For example, the N-terminal of *Bacillus subtilis* RNase HII, which has the longest known extension, is extended by 71 residues as compared to *T. kodakaraensis* RNase HII. In contrast, RNases HII from thermophilic bacteria, such as *Aquifex aeolicus* and *Thermotoga maritima*, do not have such long extensions. These findings may indicate that the absence of a large N-terminal extension is related to the functional adaptation of RNase HII enzymes to a thermophilic environment. Hence, it will...
be informative to examine whether RNases HII from thermophilic *Bacillus* strains have an N-terminal extension similar in length to that of *B. subtilis* RNase HII.

*Bacillus stearothermophilus* is a thermophilic *Bacillus* strain that grows optimally at 60°C. In the work reported here, we cloned the gene from this strain that encodes RNase HII, expressed it in *E. coli*, and purified and biochemically characterized the overproduced protein (*Bst-RNase HII*). We found that *Bst*-RNase HII contains an N-terminal extension similar to that of *B. subtilis* RNase HII. Construction and examination of a truncated protein with most of this N-terminal extension removed, followed by biochemical characterization, suggest that the N-terminal extension is important for substrate binding.

**MATERIALS AND METHODS**

**Cells and plasmids**  
*B. stearothermophilus* CU21 was previously isolated (14). The *rnhH* mutant strain *E. coli* MC3009 [F-, supE44, supF58, lacY1 or ΔlacZΔYΔ6, thrR53, gatK2, gatT2, metaB1, hsdR1Δ(γ, m), rnhH3390::carF] (15) was kindly donated by M. Itaya. Competent cells of *E. coli* HB101 [F, hsdS20(γ, m)], recA13, ara-13, proA2, lacY1, gatK2, rpsL20(Sm), xyl-5, mil-1, supE44, λ] and the plasmid pUC18 were obtained from Takara Shuzo (Kyoto). The plasmid pJLA503 was constructed by Schauer et al. (16). *E. coli* transformants were grown in Luria-Bertani medium containing 0.1 g/l ampicillin.

**Gene cloning**  
The genomic DNA of *B. stearothermophilus* CU21 was prepared as previously described (17) and used as a template to amplify a part of the gene (*Bst-rnhH*) encoding *Bst*-RNase HII by PCR. The sequences of the PCR primers were 5'-GACGAGGTCGGCCGGGGGCC-3' for the 5'-primer and 5'-GCGGTCGACCU2-3' for the 3'-primer. PCR was performed twice using a 5'-primer (5'-AGGGAGAGACATATGAAGGAGTA-CACG-3'), 3'-primer (5'-GCGGTCGACAGCGTG-3': the underlined bases show the position of the AideI site) and the 3'-primer mentioned above. The DNA sequence was determined by a Prism 310 DNA sequencer (Perkin-Elmer, Tokyo) as recommended by the supplier. The amplified DNA fragment was used as a probe for Southern blotting and colony hybridization to clone the entire *Bst*-rnhH fragment. These procedures were carried out using the AlkPhos Direct system (Amersham Pharmacia Biotech, Piscataway, NJ, USA) equilibrated with the same buffer. The pass-through fraction was applied to a column of Hitrap heparin (Amersham Pharmacia Biotech) equilibrated with the same buffer. The enzyme was eluted from the column at an NaCl concentration of 0.3 M by linearly increasing the NaCl concentration from 0 to 0.5 M. The fractions containing the enzyme were combined and used for further analyses. The purity of the enzyme was analyzed by SDS-PAGE on a 12% polyacrylamide gel (20), followed by staining with Coomassie brilliant blue R250.

**Biochemical characterization**  
The molecular mass of the protein was estimated by gel filtration chromatography using a column (1.6 x 60 cm) of Superdex 200 (Amersham Pharmacia Biotech) equilibrated with 10 mM Tris-HCl (pH 7.5) containing 150 mM NaCl. Bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and RNase A (13.7 kDa) were used as standard proteins.

The far-UV CD spectra were measured at 30°C in 50 mM Tris-HCl (pH 8.0) on an automatic spectropolarimeter (J725; Japan Spectroscopic, Tokyo). The protein concentration was ~0.1 mg/ml and a cell with an optical path of 2 mm was used. The mean residue ellipticity, θ (deg cm2 dmoll'), was calculated using an average amino acid molecular mass of 110 Da.

**Enzymatic activity**  
The RNase H activity was determined at 30°C for 15 min in 10 mM Tris-HCl (pH 8.0) containing 10 mM MnCl2, 50 mM NaCl, 1 mM 2-mercaptoethanol, and 10 μg/ml bovine serum albumin, using (7) by measuring the acid-soluble digestion product from the substrate, a 3H-labeled M13 DNA/RNA hybrid, as previously described (21). One unit was defined as the amount of enzyme producing 1 nmol of acid-soluble material per min at 30°C. The specific activity was defined as the enzymatic activity per mg of protein. The protein concentration was determined from the UV absorption with A280 values of 0.91 for *Bst*-RNase HII and 0.66 for *Asp*-RNase HII. These values were calculated using ε values of 15.6 M⁻¹ cm⁻¹ for 1yr and 5225 M⁻¹ cm⁻¹ for Trp at 280 nm (22).

**Binding analysis**  
Interaction between the protein and substrate was analyzed with a BIACore instrument (Biacore, Tokyo) using a sensor chip on which a 36-bp DNA/RNA hybrid was immobilized, as described previously (23). The sensorgrams were analyzed using BIAlveluation software (Biacore) to estimate the association constant K.

**3D modeling**  
The amino acid sequences of *Bst*-RNase HII and *T. kodakaraensis* RNase HII were aligned based on the secondary structures of *Bst*-RNase HII predicted by the EMPIR PredictProtein server. This alignment and the coordinates from the
RESULTS AND DISCUSSION

Cloning of the Bst-rnhB gene When the amino acid sequences of various bacterial RNases HI are compared, the sequences DEVGRGP and AKVTRDR, which respectively correspond to Asp^74-Pro^84 and Ala^209-Arg^209 of B. subtilis RNase HI, are observed to be highly conserved (4). On the assumption that the nucleotide sequence of the B. subtilis RNase HIgene is conserved in the Bst-rnhB gene in the regions where the amino acid sequences are conserved, DNA oligomers with the sequences encoding Asp^74-Pro^84 and Ala^209-Arg^209 of B. subtilis RNase HI were synthesized and used to amplify a part of the Bst-rnhB gene. PCR using the genomic DNA of B. steaethermophilus CU121 as a template produced only a 386-bp DNA fragment encoding a part of the Bst-RNase HI sequence. Southern blotting and colony hybridization using this DNA fragment as a probe indicated that a 2.0-kb BamHI fragment of the CU21 genomic contains the entire Bst-rnhB gene (data not shown).

Determination of the nucleotide sequence of the Bst-rnhB gene revealed that Bst-RNase HI is composed of 259 amino acid residues with a calculated molecular mass of 28,892 Da and an isoelectric point of 8.6. A potential Shine-Dalgarno (SD) sequence (5'-GGAG-3'), which is complementary to the 3'-terminal sequence of 16S rRNA from CU21, is located six nucleotides upstream of the initiation codon for translation. The nucleotide sequence of the Bst-rnhB gene has been deposited in DDBJ with accession number AB073670.

Comparison of amino acid sequences The amino acid sequence of Bst-RNase HI deduced from the nucleotide sequence showed the highest identity (50.2%) to B. subtilis RNase HI. In Fig. 1, these sequences, as well as the amino acid sequences of RNases HI from Methanococcus jannaschii and T. kodakaraensis, for which the crystallographic structures are available, are aligned based on the elements of their secondary structures. Both Bst-RNase HI and B. subtilis RNase HI have a 71-residue extension at the N-terminus, as compared to the sequences of M. jannaschii and T. kodakaraensis RNases HI. Without this N-terminal extension, Bst-RNase HI would have amino acid sequence identities of 26.0% to M. jannaschii RNase HI and 29.5% to T. kodakaraensis RNase HI. The four acidic amino acid residues involved in divalent cation binding and catalytic function, which are fully conserved in the various RNase HI sequences, are also conserved in the Bst-RNase HI sequence. These residues are Asp^78, Glu^79, Asp^170, and Asp^187. In addition, several conserved sequence motifs, such as ORGP and DSK, are conserved in the Bst-RNase HI sequence. These features suggest that Bst-RNase HI resembles other RNases HI both structurally and functionally.

Biochemical properties of recombinant Bst-RNase HI

Upon induction, Bst-RNase HI accumulated in E. coli cells in a soluble form. Its production level was roughly 5 mg/l culture. The protein was purified to give a single band on SDS-PAGE with a yield of ~20% (data not shown). The molecular mass of the protein was estimated to be ~31 kDa by both SDS-PAGE and gel filtration column chromatography, which is comparable to the calculated value. These results strongly suggest that like B. subtilis (3) and T. kodakaraensis (31) RNases HI, Bst-RNase HI exists in a monomeric form.

Bst-RNase HI requires a divalent metal cation for activity. Like B. subtilis (3) and E. coli (32) RNases HI, it prefers Mn^2+ to Mg^2+, and shows the highest RNase H activity at pH 8 in the presence of 10 mM MnCl_2 and 50 mM NaCl. The temperature dependence of the Bst-RNase HI activity was not analyzed because the substrate is unstable at high temperatures. The specific activity of Bst-RNase HI was determined to be 20±2 units/mg at 30°C (average of values obtained from two independent experiments), which is 25- and 15-fold lower than those of B. subtilis and E. coli RNases HI, respectively.

When Bst-RNase HI and B. subtilis RNase HI (0.1 mg/ml) were incubated at various temperatures for 10 min in 10 mM Tris-HCl (pH 7.5) containing 1 mM EDTA, 0.1 M NaCl, and 10% glycerol and the residual activities were determined at 30°C, Bst-RNase HI and B. subtilis RNase HI lost half their activity at incubation temperatures of ~60°C and ~40°C, respectively. These results suggest that the thermostability of Bst-RNase HI is higher than that of B. subtilis RNase HI. Thus, an absence of long N-terminal extensions in the amino acid sequences of RNases HI from thermophilic sources seems to be unrelated to the adaptation of these enzymes to high temperatures. The factors that make Bst-RNase HI more stable than B. subtilis RNase HI remain to be determined.

Properties of Δ59-Bst-RNase HI

By analyzing the biochemical properties of truncated proteins of T. kodakaraensis RNase HI in which 15, 21, 25, and 30 residues were removed from the C-terminus, it has previously been shown that the α9 helix of this enzyme is important for substrate binding (8). Because this helix is rich in basic amino acid residues, it may bind to the negatively charged DNA/RNA substrate through electrostatic interaction. However, Bst-RNase HI lacks most of the amino acid residues that form this helix. The question thus arises as to whether Bst-RNase HI possesses an alternative substrate binding domain. Because basic amino acid residues are clustered at positions 41–55 within the N-terminal extension of Bst-RNase HI, it was thought that this extension may function as a substrate binding domain in this protein. To verify whether this is the case, the truncated protein Δ59-Bst-RNase HI, in which 59 N-terminal residues were removed, was constructed and its biochemical properties were analyzed. This truncated protein was overproduced in E. coli cells and purified to give a single band on SDS-PAGE, as was the intact protein. The truncated and intact proteins gave almost identical far-UV CD spectra (data not shown), suggesting that the protein conformation was not seriously changed upon N-terminal truncation. Nevertheless, Δ59-Bst-RNase HI completely lost...
The amino acid sequence of Bst-RNase HII (Bst) is compared with those of RNases HII from B. subtilis (Bsu) (GenBank accession Bsu: rnh), M. jannaschii (Mja) (GenBank, U76470), and T. kodakaraensis KODI (Tko) (DDBJ, AB012613). Gaps are denoted by dashes. Amino acid residues conserved in at least three different enzymes are shown by white letters on a black background. The four acidic catalytic residues are marked by black circles. Numbers indicate the positions of the amino acid residues, which start from the initiator methionine for each enzyme. The ranges of the nine a-helices and five b-strands of T. kodakaraensis (8) and M. jannaschii (9) RNases HII are shown below the sequences. The position at which the N-terminal extension is truncated to create A59-Bst-RNase HII is indicated by a black inverted triangle.

FIG. 1. Alignment of amino acid sequences of RNases HII. The amino acid sequence of Bst-RNase HII (Bst) is compared with those of RNases HII from B. subtilis (Bsu) (GenBank accession Bsu: rnh), M. jannaschii (Mja) (GenBank, U76470), and T. kodakaraensis KODI (Tko) (DDBJ, AB012613). Gaps are denoted by dashes. Amino acid residues conserved in at least three different enzymes are shown by white letters on a black background. The four acidic catalytic residues are marked by black circles. Numbers indicate the positions of the amino acid residues, which start from the initiator methionine for each enzyme. The ranges of the nine a-helices and five b-strands of T. kodakaraensis (8) and M. jannaschii (9) RNases HII are shown below the sequences. The position at which the N-terminal extension is truncated to create A59-Bst-RNase HII is indicated by a black inverted triangle.

its RNase H activity.
To examine whether N-terminal truncation seriously affects substrate-binding affinity, interaction between the protein and substrate was analyzed using surface plasmon resonance. The sensorgrams for the interaction between the intact protein and the substrate are shown in Fig. 2A. The responses increased with increasing protein concentration. From the plot of RU/C as a function of RU, shown in Fig. 2B, the association constant \(K_a\) was estimated to be \((2.15 \pm 0.43) \times 10^7\) (mean \(\pm S.E., n = 7\). In contrast, the sensorgrams for the interaction between the truncated protein and the substrate did not give such positive signals. The response did not increase beyond the background level at any protein concentration examined, indicating that A59-Bst-RNase HII does not bind to the substrate. These results suggest that the N-terminal extension of Bst-RNase HII is important for substrate binding.

Substrate binding domain. T. kodakaraensis RNase HII and E. coli RNase H share a main chain fold termed the “RNase H fold” (8). This fold is conserved in all members of nucleotidyltransferase superfamilies (33). However, these two enzymes differ in the location of the domain involved in substrate binding. In T. kodakaraensis RNase HII, it occurs as an extra C-terminal domain (8) whereas in E. coli RNase H it is as an internal domain termed the “basic protrusion” (34). Our findings indicate that Bst-RNase HII may represent a novel type of RNase H that differs from T. kodakaraensis RNase HII and E. coli RNase H in its substrate binding domain location. Because the N-terminal extension of bacterial RNases HII differ greatly in length, it remains to be determined whether other bacterial RNases HII, including that of E. coli, also have a substrate binding
domain at their N-termini. In contrast, none of the amino acid sequences of archaeal RNases HII so far available contain an N-terminal extension, suggesting that archaical RNases HII possess a substrate binding domain at their C-termini. To understand the role of the N-terminal extensions of Bacillus RNases HII, it will be necessary to determine their crystal structures.

3D modeling The fact that the amino acid sequence of Bst-RNase HIII at the core region does not contain large insertions when compared to that of T. kodakaraensis RNase HIII (Fig. 1) allowed us to build a model of the three-dimensional structure of the Bst-RNase HIII core region. Comparison of this 3D-model with the crystal structure of T. kodakaraensis RNase HIII suggests that the Bst-RNase HIII structure is similar to that of T. kodakaraensis RNase HIII (Fig. 3). The steric configurations of the four acidic catalytic residues are well conserved in the two structures. However, the former lacks the α5 helix and most of the α9 helix. This difference reflects the large internal deletion and C-terminal truncation in the amino acid sequence of Bst-RNase HIII (Fig. 1).

According to the crystal structure of T. kodakaraensis RNase HIII, the α5 helix is located in the vicinity of a catalytic site and forms a hydrophobic core with a central β-sheet (Fig. 3). Because bacterial RNases H prefer Mn2+ to Mg2+ for activity (13), whereas T. kodakaraensis RNase HIII exhibits its activity almost equally in the presence of Mn2+ and Mg2+, this helix may be related to the metal ion preference of the enzyme.

REFERENCES


