Nucleic Acids Research, Vol. 19, No. 16 4443–4449

Molecular cloning of a ribonuclease H (RNase HI) gene from an extreme thermophile *Thermus thermophilus* HB8: a thermostable RNase H can functionally replace the *Escherichia coli* enzyme *in vivo*

Mitsuhiro Itaya and Kanae Kondo
Laboratory of Microbiology, Department of Molecular Biology, Mitsubishi Kasei Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan

Received May 20, 1991; Revised and Accepted July 26, 1991 EMBL accession no. X60507

ABSTRACT

A DNA fragment encoding Ribonuclease H (EC 3. 1. 26. 4) was isolated from an extreme thermophilic bacterium, *Thermus thermophilus* HB8, by its ability to complement the temperature-sensitive growth of an *Escherichia coli rnhA* deficient mutant. The primary amino acid sequence showed 56% similarity to that of *E. coli* RNase HI but little or no homology to *E. coli* RNase HII. Enzymes derived from thermophilic organisms tend to have fewer cysteines than their bacterial counterparts. However, *T. thermophilus* RNase H has one more cysteine than its *E. coli* homologue. Stability of the RNase H in extracts of *T. thermophilus* to elevated temperatures was the same for the protein expressed in *E. coli*. *T. thermophilus* RNase H should, therefore, be a useful tool for editing RNA-DNA hybrid molecules at higher temperatures and may also be stable enough to be used in a cyclical process. It was suggested that regulation of expression of the RNase H may be different from that of *E. coli*. RNase HI.

INTRODUCTION

Ribonuclease H (RNase H, EC 3.1.26.4) is an endonuclease which degrade specifically the RNA moiety of RNA-DNA hybrid molecules. The wide distribution of RNases H in various organisms, from bacteria to human cell lines, suggests important physiological functions for the enzyme (1).

Extensive studies on *Escherichia coli* RNase H (RNase HI, encoded by the *rnhA* gene) have implicated the enzyme in such processes as DNA replication (2–4), or DNA repair (5–7). However, the recent finding of the second RNase H (RNase HII, encoded by the *rnhB* gene) might require re-evaluation of the largely studied *E. coli* RNase H mutants (*rnhA*$^-$ *rnhB*$^+$) (8). The structural and functional relationship of the *E. coli* RNase HI has been under investigation using kinetic and physicochemical analyses (9, 10) and by X-ray crystallographic analysis (11, 12). Furthermore, functional and structural similarities between the C-terminus of retroviral Reverse Transcriptases (RTs) and *E. coli* RNase HI have been indicated by computer analyses (13) as well as experimental observations (14–18). Many genes encoding Reverse Transcriptases have been cloned (13). In contrast, only four clones of cellular RNase H genes [*E. coli* RNase HI (19, 20),*E. coli* RNase HII (8), *Salmonella typhimurium* RNase HI (21) *Saccharomyces cerevisiae* RNase HI (21)] and Bacteriophage T4 RNase H (22) have been described to date. However, there are multiple reports of purified cellular RNases H from various species (see for example 1, 23–26). We though it would be useful to obtain an RNase H from *Thermus thermophilus*, a thermophilic eubacteria which grows at 85°C (27). Obtaining such a clone would: (i) demonstrate the presence of an RNase H gene in the species,(ii) provide information on the requirements of an RNase H protein to functionally substitute in vivo for the *E. coli* cell, (iii) permit study of the thermostability of the RNase H, and (iv) possibly provide an enzyme useful for RNA editing (28) at elevated temperatures.

In this paper, we report the molecular cloning of a *T. thermophilus* RNase H gene by the use of an *E. coli* RNase H deficient mutant (8, 21). We also provide evidence that the RNase H similar in primary amino acid sequence to *E. coli* RNase HI yet possesses significantly increased thermostability.

MATERIALS AND METHODS

Bacterial Strains and plasmids

T. thermophilus HB8 (ATCC 27634) was from T. Oshima (27), and highly purified chromosomal DNA was from T. Tanaka (29). LE392 [F$^-$ *supE44 supF58 lacY or (lacIZY)6 trpR55 galK2 galT22 metB1 hsdR14* (r_k^- m_k^+) *rnhA*$^+$ *rnhB*$^+$] and MIC3037 [*rnhA339::cat rnhB*$^+$ *recC271(Ts)*, derived from LE392] have been described (8, 21). JA221 (F$^-$ *hsdR hsdM*$^+$ *trpE5 leuB6 lacY recA1 rnhA*$^+$ *rnhB*$^+$) as a general cloning host has been described previously (8). Cells were grown in LB medium, *T. thermophilus* HB8 at 80°C and *E. coli* at 37°C. pBR322 was used as vector for construction of the *T. thermophilus* chromosomal DNA library and also for subclonings. pGEM4 (Promega Biotec) was used for exonuclease III-mediated sequential deletion analysis (30).

Enzymes and DNA manipulations

Exonuclease III and all the type II restriction endonucleases were from Toyobo (Tokyo, Japan). T4-DNA ligase and Klenow DNA polymerase were from Boehringer (USA). Other chemicals were reagent grade. DNA cloning techniques were according to either Maniatis *et al.* (31) or suppliers' manuals.

Construction and screening of a *T. thermophilus* DNA library

DNA fragments (>4 kb) of a *Sau*3A partial digest were isolated from low melting point agarose gel(Sigma, TypeVII) and were ligated to alkaline phosphatase treated *Bam*HI site of pBR322 and transformed into the *E. coli* strain MIC3037. Competent MIC3037 cells were prepared according to $CaCl_2$ treatment method (32). Colonies growing on LB plate supplemented with ampicillin (100μg/ml) at 30°C were tested on fresh LB plate containing tetracycline (20μg/ml). A total of 753 ampicillin-resistant, tetracycline-sensitive transformants were pooled. After appropriate dilution, they were plated on LB plate with ampicillin and incubated at 42°C for 2 days. Plasmid DNAs were isolated from colonies formed at 42°C and used to transform MIC3037. Cloned DNAs which repeatedly gave MIC3037 transformants at 42°C were considered positive. This process is termed a temperature-sensitive complementation assay.

Assay of RNase H

Specific substrate for RNase H, poly([^{32}P]rA). poly(dT) was prepared according to Keller & Crouch (33). [^{32}P]ATP was from Amersham (USA). Enzymatic activity was determined by the method of either renaturation gel assay, or acid-soluble radioactivity measurement (8). Renaturation gel assay was done as described in ref. 8. Coomassie brilliant blue staining of the gel was conducted before exposure to an X-ray film.

Preparation of mini-scale crude lysate for RNase H activity measurement was according to Ogawa & Okazaki (34). Briefly, in 1.5-ml microcentrifuge tube, 10μl of a stationary culture of *T. thermophilus* HB8 or *E. coli* was mixed with 10μl of 40mM Tris-HCl, pH7.8, 10mM EDTA, Lysozyme (800μg/ml). The mixture was rapidly frozen in dry ice-ethanol and thawed in a water-bath at 37° for 1 min. This freezing and thawing process was repeated three times. RNase H assay buffer (200μl of 40mM Tris-HCl, pH7.6, 4mM $MgCl_2$, 1mM dithiothreitol) was added and followed by brief mixing. The solution was immediately heated to various temperatures. Usually 5μl was used per RNase H assay (50μl of 40mM Tris-HCl, pH7.6, 4mM $MgCl_2$, 1mM dithiothreitol, 30μg of bovine serum albumin at 37°C for 15 min) for measuring the residual RNase H activity.

Subcloning and DNA Sequence Analysis

A 1.3 kbp *Xho*I−*Bam*HI fragment from pRET4 (Fig. 2) was made blunt with Klenow DNA polymerase and inserted at the *Sma*I site of pGEM4. The resulting plasmid, pRET44, was used for exonuclease III-mediated deletion analysis (30). The derivatives were used for DNA sequence analysis and temperature-sensitive complementation assay. Nucleotide sequence was determined using dideoxynucleotides and double-stranded plasmid DNAs were used as templates. Toyobo M13 sequencing kit (Toyobo, Tokyo Japan) with a labeling nucleotide [^{35}S] dCTP (New England Nuclear, USA) was used. SP6 (17 mer) and T7 primers (16 mer) for DNA sequencing were from Promega (USA).

RESULTS

Isolation of the *rnh* gene from *T. thermophilus* HB8

To isolate the DNA fragment encoding *T. thermophilus* RNase H we adopted the method used for cloning of the *rnhB* gene of *E. coli* (8). *E. coli* strain MIC3037 shows a temperature-sensitive growth phenotype which can be alleviated by supplying a functional RNase H gene such as the *rnhB* of *E. coli* (8), *rnhA* of *S. typhimurium* (21), or *RNH1* of *S. cerevisiae* (21). If RNase H of *T. thermophilus* could be expressed and function properly in *E. coli*, MIC3037 was expected to form colonies at 42°C, the non-permissive temperature (temperature-sensitive complementation assay). By screening a *T. thermophilus* DNA library, as described in Materials and Methods, a total of 3 distinctive clones were obtained. Only one clone (pRET4) produced RNase H activity, as shown in lane B of Fig. 1.

Deletion analysis of the fragment in pRET4

A partial restriction map of a 5 kbp insert of pRET4 plasmid was shown in Fig. 2. Subsequent deletion analyses, monitored by both temperature-sensitive complementation assays and

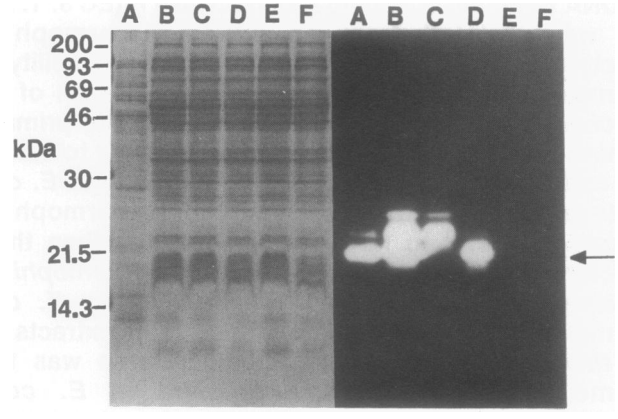

Fig. 1. Measurement of *T. thermophilus* RNase H activity by renaturation gel assay. The left panel shows a Coomassie brilliant blue stained renaturation gel. In the right panel RNase H activities are visible as white bands. Size markers are indicated at the left side (kilodaltons). Lane A: lysate of *T. thermophilus* HB8. Lanes B−F: lysate of MIC3037 transformant harboring; (B), pRET4;(C), pRET45A; (D), pRET45D;(E), pRET45E; (F), pBR322. The position of the *T. thermophilus* RNase H activity is indicated by an arrow. The higher molecular weight bands in lanes B and C are described in Discussion section.

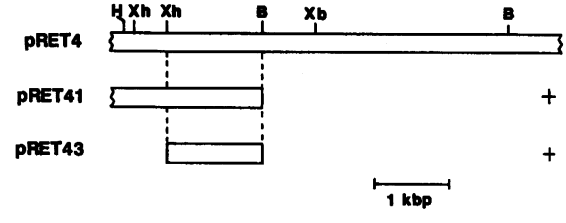

Fig. 2. A Restriction map of the 5 kbp insert fragments of pRET4. Restriction sites are indicated with B, *Bam*HI; H, *Hin*dIII; Xb, *Xba*I; Xh, *Xho*I. The orientation of the insert from left to right is the same as the tetracycline-resistance gene of pBR322. The *Bam*HI site at the left *Sau*3A/*Bam*HI junction was recreated. Construction of pRET41, and 43 are described in the Results section. Size is indicated by a bar and positive temperature-sensitive complementation of MIC3037 are indicated by +.

renaturation gel assays, located the *T. thermophilus rnh* gene on the 1.3 kbp *XhoI*−*BamHI* fragment. Excision of the *Bam*HI fragment from pRET4 generated pRET41. pRET43 was obtained from pRET41 by excision of the segment from *Xho*I site to the *Cla*I site of the pBR322 vector. The 1.3 kbp *Xho*I−*Bam*HI fragment was moved from pRET4 to a *Sma*I site within the multiple cloning site of pGEM4 for further deletion analysis, resulting in pRET44(shown in Fig. 3). The insert occurred to generate the order SP6 promoter-*Xho*I−*Bam*HI-T7 promoter. When pRET44 was used as a probe for Southern hybridization analysis of *T. thermophilus* DNA, the expected patterns deduced from the restriction map of pRET4 were obtained (data not shown). To determine the precise position of the RNase H gene and its nucleotide sequence, derivatives of pRET44 were obtained by exonuclease III-mediated deletion. Plasmids pRET44A−H (Fig. 3) were obtained through unidirectional deletion (30) from the *Bam*HI side using *Bam*HI and *Pst*I sites. An *Acc*I site had been found within the *T. thermophilus* DNA as indicated in Fig. 4. From the *Xho*I side, pRET44A was digested with the *Acc*I and *Kpn*I treated with exonuclease III, yielding plasmids pRET45A−H (shown in Fig. 3). The minimal DNA fragments which gave positive in the renaturation gel assay as well as the temperature-sensitive complementation assay were on plasmid pRET44B and pRET45D. This encompasses a total of 553 base pairs (nucleotide 385−937).

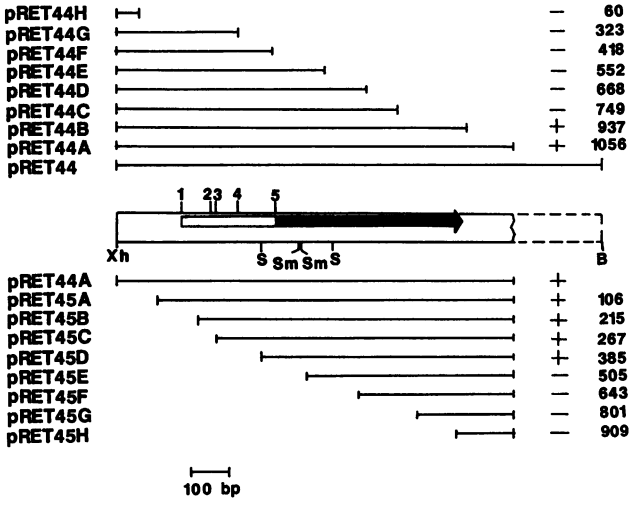

Fig. 3. Derivatives of pRET44 carrying an RNase H gene of *T. thermophilus*. The procedures for generating pRET44 derivatives are described in the Results section. The arrow indicates the orientation and location of an ORF predicted from the nucleotide sequence. S and Sm stand for *Sac*I and *Sma*I. Numerals at the right of each deletion derivative indicate the end of nucleotide sequence shown in Fig. 4. GTG(423) is the start site for translation of the *T. thermophilus* RNase H gene (the closed portion of the arrow). Size is represented by a bar. Positive or negative temperature-sensitive complementation is indicated by + or −.

DNA sequence analysis

T7-promoter primer was used for determination of the nucleotide sequence for plasmids pRET44A−H. For plasmids pRET45A−H, SP6-primer was used. An open reading frame starting at nucleotide position 174(GTG) to 915(TAA) was found. However, the plasmid pRET45D started at nucleotide 385 and remained functional. At nucleotide position 174 there was a GTG sequence which has been found to be an initiation codon for some proteins

```
        10         20         30         40         50         60         70         80         90
TCGAGGCCATCCCCGTGGCCGACCTCATCGCCCGGCTCAAGTAGTGGAGGTTCTGGAGACTTCCGTCTACGCCGAGGACCTGGAGAAGGC

       100        110        120        130        140        150        160        170   1    180
CCGGGCCTCTACGAAGGGGTCTTGGGCCTTCCCTCCTTTCAGTTCAAACCGCCCCGGCACGCCTTCTTCCGCGCCGGGCGGGGGTGTTCC

       190        200        210        220        230        240        250   2    260   3    270
TGGTCTTCAACCCAAAGCTCACGGCAAAGGAAACCCACCTCCCTCCCCACGGGGCCAAAGGCAGCGTCCACGTGGCCTTCCGAGTGGAGG

       280        290        300        310        320   4    330        340        350        360
AGGAGGAGCTTCCCGCCTGGGAAGCGCGCCTCAAGGCGGCGGGCTACCCGGTGTGGTGGGCCGAGTGGCAAGGGAAAGAGCCTCTACACC

       370        380        390        400        410        420   5    430        440        450
CAGGACCCCGCCGGCAACCTGGTGGAGCTCGCCCCGGCGCGGATCTGGGGCCTGGAGGAGGCGTGAACCCCTCCCCCAGGAAACGCGTGG
                                                                    V  N  P  S  P  R  K  R  V

       460        470        480        490        500        510        520        530        540
CCCTCTTCACCGACGGGGCCTGCCTGGGAAACCCCGGGCCCGGGGGGTGGGCGGCCCTCCTCCGCTTCCACGCCCACGAGAAGCTCCTCT
A  L  F  T  D  G  A  C  L  G  N  P  G  P  G  G  W  A  A  L  L  R  F  H  A  H  E  K  L  L

       550        560        570        580        590        600        610        620        630
CCGGGGGAGAGGCCTGCACCACCAACAACCGCATGGAGCTCAAGGCGGCCATAGAGGGCCTAAGGCTCAAGGAGCCTTGCGAGGTGGACC
S  G  G  E  A  C  T  T  N  N  R  M  E  L  K  A  A  I  E  G  L  R  L  K  E  P  C  E  V  D

       640        650        660        670        680        690        700        710        720
TCTACACCGACAGCCACTACCTCAAGAAGGCCTTCACCGAGGGCTGGCTGGAAGGCTGGCGGAAAAGGGGCTGGCGGACGGCGGAGGGCA
L  Y  T  D  S  H  Y  L  K  K  A  F  T  E  G  W  L  E  G  W  R  K  R  G  W  R  T  A  E  G

       730        740        750        760        770        780        790        800        810
AGCCCGTGAAAAACCGCGACCTCTGGGAGGCCCTCCTCCTCGCCATGGCCCCCCACGGTGCGCTTCACTTCGTGAAGGGGCACACGGGCC
K  P  V  K  N  R  D  L  W  E  A  L  L  L  A  M  A  P  H  G  A  L  H  F  V  K  G  H  T  G

       820        830        840        850        860        870        880        890        900
ACCCGGAGAACGAACGGGTGGACCGGGAGGCGAGGCGCCAGGCCCAGTCCCAGGCCAAAACGCCCTGCCCGCCCCGGGCCCCCACGCTTT
H  P  E  N  E  R  V  D  R  E  A  R  R  Q  A  Q  S  Q  A  K  T  P  C  P  P  R  A  P  T  L

       910        920        930        940        950        960        970        980        990
TTCACGAAGAGGCATAAAAGCGTTATAATCGGGCCATGCTGGAAACCGCCCTGCGCCCGGAGGACCGGGAAAGCCCCACCTTTAAGGCCA
F  H  E  E  A

      1000       1010       1020       1030       1040       1050
ACAAGGACGCCTGGGTGGCCCTGGTGCGGGACTTCAGGGAAAGCCTGGAAAGGGTGCGCCAGGGAG
```

Fig. 4. Nucleotide sequence of the *T. thermophilus rnhA* gene and a deduced amino acid sequence. Numbering of the nucleotides starts from the *Xho*I site of pRET44A. Wavy underline is the *Acc*I site which was used in exonuclease III-mediated deletion experiment (see Deletion analysis of the fragment in pRET4). Initiation codons (GTG) found upstream of position 423 are boxed with numerals (see Fig. 3 and text). A potential Shine-Dalgarno sequence is underlined (36). Junctions between the insert and vector sequences of pRET45D (position 385) and pRET45E (position 505) are indicated by arrows.

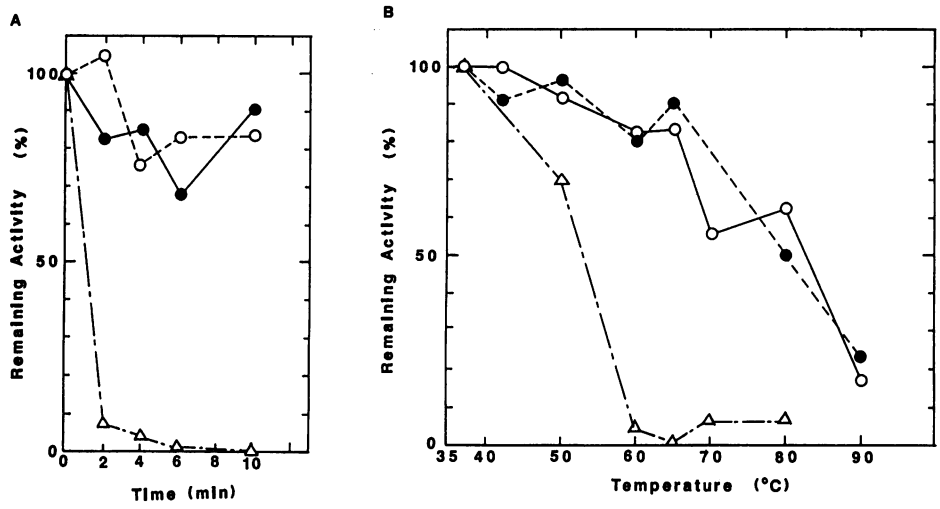

Fig. 5. Rate of heat inactivation of *T. thermophilus* RNase H and *E. coli* RNase HI. (A) A crude extract from each cell was incubated at 65°C and at each time interval acid soluble radionucleotide was measured at 37°C using poly([^{32}P]rA).poly(dT). The reaction conditions are in Materials and Methods. Lysates are from: (●), *T. thermophilus* HB8; (○), MIC3037 harboring pRET45D; (△), LE392. (B) Crude extracts are incubated for 10 min at the designated temperatures. Measurement of the remaining RNase H activity and symbols are the same as in (A).

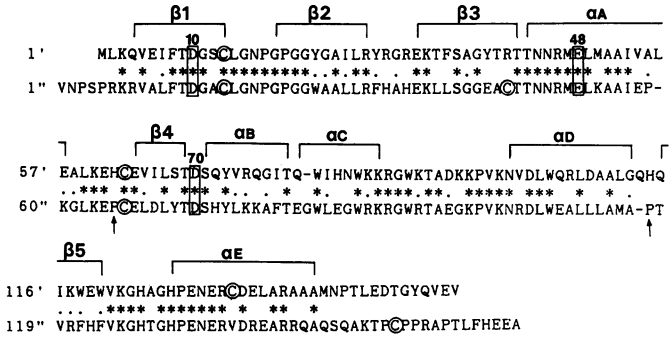

Fig. 6. Alignment of amino acid sequences of *E. coli* RNase HI and *T. thermophilus* RNase H. The 155 amino acid sequence of *E. coli* RNase HI (top lane, refs. 19, 20) and the 164 amino acid sequence of *T. thermophilus* RNase H (bottom lane, Fig. 4) are aligned. Numerals with ' or " indicate the position of the first amino acid of each lane. Identical amino acids in both sequences are indicated by * and similar amino acids by dots. Regions of helices (αA−αE) and strands (β1−β5) on the *E. coli* RNase HI sequence are drawn according to Yang. et al.(11) and Katayanagi et al.,(12). Amino acid residues involved in the active site of *E. coli* RNase HI (9) are numbered and boxed. Cysteine residues are circled. Two proline residues of *T. thermophilus* RNase H discussed in the text are indicated by arrows.

and may be used here since there was no ATG found. The protein product expected from initiation at GTG(174) would be 242 amino acids in length (Fig. 3, 4) whereas initiation at GTG(423) would yield a protein of 164 amino acids.

Detection of *T. Thermophilus* RNase H

When we measured the RNase H activities from *T. thermophilus* HB8 by renaturation gel assay, a major RNase H activity (about 20 kilodaltons) similar to that of *E. coli* (17,596 daltons) was observed as shown in lane A of Fig. 1. Crude lysates were prepared from stationary cultures of *T. thermophilus* HB8 (*T. thermophilus* RNase H proficient), MIC3037 harboring pRET45D(*E. coil* RNase HI deficient, but *T. thermophilus* RNase H expressed), and LE392 (*E. coli* RNase HI proficient) as described in Materials and Methods. From the result of renaturation gel assay shown in Fig. 1, major RNase H activities (lanes A and D) were attributed to *T. thermophilus* RNase H. The contribution of the second RNase H of *E. coli* (RNase HII, Mr=23,225 daltons) in crude *E. coli* lysate was negligible, since RNase HII was not seen in lanes E, F of Fig. 1. Activity of *T. thermophilus* RNase H in the lysate from MIC3037 harboring pRET45D was estimated about 8% of that from LE392, the wild type *E. coli* RNase HI, based on the assumption that the lysates were prepared from the same number of cells. We failed to detect an increase of *T. thermophilus* RNase H protein band on CBB-stained SDS-PAGE (Fig. 1, left panel). Liquid RNase H assays of crude lysates of *T. thermophilus* gave activities about 20% of that found for LE392.

Thermostability of *T. thermophilus* RNase H

Enzymes from *T. thermophilus* are generally expected to exhibit increased thermostability compared to those from other mesothermal bacteria (35). For thermostability of *T. thermophilus* RNase H, residual activity of crude lysate containing the enzyme was measured after heating at various temperatures and for variable length of time. This approach was taken since the RNA-DNA hybrid substrate is altered at the elevated temperatures.

T. thermophilus RNase H retained about 90% of its activity after a 10 minute incubation at 65°C while the activity of *E. coli* RNase HI dropped sharply, losing 95% of its activity within the first 2 minutes (Fig. 5A). The result suggested that *T. thermophilus* RNase H was either stable at 65°C or could renature after heat treatment, whereas the *E. coli* enzyme was inactivated. The thermostability was also observed when residual activities were measured at various temperatures as shown in Fig. 5B. The *T. thermophilus* RNase HI retained about 50% of its activity even when exposed to 80°C for 10 minutes. Because heat inactivation profiles are indistinguishable between *T. thermophilus* RNase H derived from *T. thermophilus* cells and that expressed in *E. coli*, we conclude that we measured a property intrinsic to the *T. thermophilus* RNase H.

DISCUSSION

The presence of RNase H enzyme has been reported in various species (1) and multiple RNase H 'genes' has been suggested to be common in prokaryotes(8). The existence of at least one RNase H gene in the extreme thermophilic bacteria, *T. thermophilus* HB8, was indicated by the following criteria: i) The gene product expressed in *E. coli* showed RNase H activity (Fig. 1) and ii) The gene product alleviated the temperature-sensitive growth of an *E. coli* RNase HI deficient mutant (Figs. 2 and 3)

Structure of the *T. thermophilus* RNase H gene

We concluded that the GTG(423) was the initiation site for translation of *T. thermophilus* RNase H (Figs. 3 and 4) based on the following observations: i) The calculated molecular weight of GTG (423) product (164 amino acids, 18,335 dalton) agreed with the size determined by the renaturation gel assay (about 20 kDa—Fig. 1) and was similar to RNase HI of *E. coli* (155 amino acids, 17,569 daltons; refs. 19, 20) and ii) Two initiation codons, GTG(423) and GTG(447), were found between 384 and 505 (pRET45D and pRET45E) but only GTG(423) had a Shine−Dalgarno sequence(36). The larger bands observed in lanes B and C of Fig. 1 (which varied in amount from experiment to experiment—data not shown) may be products expressed from one of the four upstream GTGs indicated in Figs. 3 and 4. These lager bands have never been detected in the lysate from *E. coli* bearing the plasmid pRET45D (lane D of Fig. 1).

Structure of *T. thermophilus* RNase H protein

Homology between RNases H from *T. thermophilus* and *E. coli* was detected using the GENETYX computer program (shown in Fig. 6). Fifty six percent similarity of primary amino acid sequence to RNase HI (19, 20) was observed but less than 30% similarity to RNase HII (8) was found. Kanaya et al. (9) identified three amino acids (Asp^{10}, Glu^{48}, and Asp^{70}) to be at the catalytic site of *E. coli* RNase HI. From X-ray crystallographic analysis, the active site of *E. coli* RNase HI is composed of the essential carboxylic acid triad (Asp^{10}, Glu^{48}, and Asp^{70}) spatially surrounded by Ser^{71}, His^{124}, Asn^{130}, and Asp^{134} (11, 12). All of these amino acids were found (Fig. 6) in *T. thermophilus* RNase H. *T. thermophilus* RNase H probably is similar in three dimensional structure and mechanism of enzymatic cleavage to that suggested for *E. coli* RNase HI (9, 11, 12).

The deduced amino acid composition is listed in Table I. Prominent differences between the amino acid composition of *E. coli* RNase HI and *T. thermophilus* RNase H were observed [Ile (7 to 1), Pro (5 to 12), Gln (8 to 3) and Tyr (5 to 12)]. It has been suggested that a decrease in cysteine residues contributes to the thermostability of enzymes from thermophilic bacteria (38). However, *T. thermophilus* RNase H had one more cysteine than did *E. coli* RNase HI. Two of the 4 cysteine residues corresponded to Cys^{13} and Cys^{63} of *E. coli* (see Fig. 6). Kanaya et al. reported that none of the cysteine residues was required for enzymatic activity of *E. coli* RNase HI (10). The fourth cysteine residue, at position 152 of *T. thermophilus* RNase H, may not be important for RNase H activity, because in RNase HI of *S. typhimurium* this residue also differed from *E. coli* RNase HI (21). We think *T. thermophilus* RNase H might be an exception in the general tendency of thermostable enzymes to have fewer cysteines.

Table I. Amino acid composition of RNases HI from *E. coli* and *T.thermophilus*.

Amino acid	*E. coli*	*T. thermophilus*
Gly	14	15
Ala	14	19
Val	9	6
Leu	12	18
Ile	7	1
Met	4	2
Phe	2	5
Trp	6	5
Pro	5	12
Ser	4	4
Thr	10	9
Asn	7	6
Gln	8	3
Cys	3	4
Asp	7	5
Glu	12	15
Lys	11	11
His	5	8
Arg	10	14
Tyr	5	2
total	155	164

Data for *E. coli* RNase HI are from Kanaya and Crouch(19). *T. thermophilus* RNase HI is based on the deduced amino acid sequence shown in Fig. 4.

There are a myriad plausible explanations to account for the thermostability of *T. thermophilus* RNase H since about half of the amino acids differ from the *E. coli* enzyme. Among the changes, however, we think the positioning of two proline residues is intriguing. The tertiary structure of *E. coli* RNase H is composed of five β strands ($\beta1-\beta5$) and five α-helices (A−E) (11, 12). In *T. thermophilus* RNase H, one proline residue (Pro^{62}) is located at the start site of $\beta4$ and the other (Pro^{114}) is at the junction between helix D and $\beta5$ (indicated by arrows in Fig. 6). His^{62} and His^{144} are conserved in *S. typhimurium* RNase HI (21) but converted to Ile and Phe, respectively in an alignment with *S. cerevisiae* RNase H sequence (21). One could imagine that the presence of these two proline residues increase the rigidity of the tertiary structure of *T. thermophilus* enzyme, thereby contributing to an increased thermostability.

Expression and physiological roles of *T. thermophilus* RNase H

The promoters for the *E. coli rnhA* gene and the *E. coli dnaQ* gene (the gene encoding epsilon subunit of DNA polymerase III) are overlapping and transcription proceeds in opposite directions on the chromosome(20). Such a *dnaQ-rnhA* relationship is conserved in *S. typhimurium* (21). However, no indication of the presence of a *dnaQ*-like gene around the *T. thermophilus rnh* gene was found. The lack of association with a gene similar to *dnaQ* may imply a different mechanism of expression of *T. thermophilus* from that suggested for *E. coli* (39, 40). The choice of the relatively rare initiation codon (GTG) also implies that the level of transcription of *T. thermophilus* RNase H might be different in *T. thermophilus* than in *E. coli* (41). Since no obvious promoter for *E. coli* RNA polymerase was found upstream of the *T. thermophilus rnh* gene, we suspect that transcription of the *T. thermophilus* RNase H in MIC3037 could be from vector sequences. This is particularly the case for pRET45D, the smallest clone which was positive for both temperature-sensitive complementation (Fig. 3) and RNase H activity (Fig. 1).

Table II. Biased third letter guanine or cytosine usage of the *rnh* gene of *T. thermophilus*.

		E. coli	T. thermophilus			E. coli	T. thermophilus
UU	U	1	1	UA	U	2	0
	C	1	4		C	3	2
	A	2	0		A	*	*
	G	3	0		G	*	*
CU	U	2	2	CA	U	4	0
	C	1	13		C	1	8
	A	0	1		A	4	0
	G	4	2		G	4	3
AU	U	4	0	AA	U	3	0
	C	3	0		C	4	6
	A	0	1		A	11	4
	G	4	2		G	0	7
GU	U	3	0	GA	U	5	0
	C	4	0		C	2	5
	A	2	0		A	9	3
	G	0	6		G	3	12
UC	U	0	0	UG	U	2	0
	C	0	3		C	1	4
	A	0	0		A	*	*
	G	1	0		G	6	5
CC	U	1	1	CG	U	4	0
	C	1	9		C	6	5
	A	2	0		A	0	0
	G	1	2		G	0	5
AC	U	0	0	AG	U	1	0
	C	8	5		C	2	1
	A	2	0		A	0	0
	G	0	4		G	0	4
GC	U	5	0	GG	U	3	1
	C	3	13		C	6	6
	A	2	1		A	3	2
	G	4	5		G	2	6

*stand for ocher (UAA), amber(UAG), and opal (UGA). Numbers of codon usage were derived from the data of Kanaya and Crouch(19) for *E. coli* RNase HI and data shown in Fig. 4 for *T. thermophilus*.

The high guanine and cytosine content of the *rnh* gene (69.7%) and the biased usage of G or C at the third position (Table II) are also as expected in for this species (35, 37). As a result of the high G+C content,termination codons are rare. The fact that there is a significant open reading frame upstream of the region necessary for temperature-sensitive complementation and RNase H activity may or may not be important. The RNase H portion of Reverse Transcriptases (RTs) and *S. cerevisiae* RNase H are both located on the carboxyl-portion of the proteins(21). It may be that *T. thermophilus* RNase H product possesses yet unidentified physiological roles in *T. thermophilus*.

The level of RNase H activity required for temperature-sensitive complementation of MIC3037 was estimated less than 1% of wild type *E. coli* RNase HI (8, 21, 42). The overall activity of *T. thermophilus* RNase H expressed from pRET45D in MIC3037 was about 8% of that of wild type *E. coli* RNase HI (LE392). It is, therefore, readily conceivable that the amount of *T. thermophilus* RNase H activity was sufficient for temperature-sensitive complementation of MIC3037.

The construction of systems to overexpress *T. thermophilus* RNase H in *E. coli* are now in progress. The thermostable enzyme, when purified, will serve for the study of thermostability as well as in vitro RNA editing (28) at higher temperatures. We have had no indication of the presence of RNAse HII counterpart in *T. thermophilus*. Because we fully expect to find a second RNase H (RNase HII) in *T. thermophilus*, we propose that the enzyme described in this report be called RNase HI of *T. thermophilus*(encoded by the *rnhA* gene).

ACKNOWLEDGMENTS

We wish to thank Drs. T.Oshima and T.Tanaka respectively for the *T. thermophilus* HB8 strain and its chromosomal DNA sample and also Drs. S.Kanaya and R.J.Crouch for helpful discussions. We sincerely thank Dr. R.J.Crouch particularly for his careful reading of the manuscript and many suggestions. We thank Miss Y.Yamada for her help to prepare this manuscript.

REFERENCES

1. Crouch, R. J., and Dirksen, M.-L. (1982) in Linn, S. M., and Roberts, R. J., (eds) *Nuclease*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 211–241.
2. Ogawa, T., Pickett, G. G., Kogoma, T., and Kornberg, A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1040–1044.
3. Horiuchi, T., Maki, H., and Sekiguchi, M. (1984) *Mol. Gen. Genet.* **195**, 17–22.
4. Dasgupta, S., Masukata, H., and Tomizawa, J. (1987) *Cell* **24**, 1113–1122.
5. Bockrath, R., Wolff, L., Farr, A., and Crouch, R. J. (1987) *Genetics* **115**, 33–40.
6. Foster, P.L., Sullivan, A. D., and Franklin, S. B. (1989) *J. Bacteriol.* **171**, 3144–3151.
7. Itaya, M. and Crouch, R. J (1991) *Mol. Gen. Genet.* in press.
8. Itaya, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8587–8591.
9. Kanaya, S., Kohara, A., Miura, Y., Sekiguchi, A., Iwai, S., Inoue, H., Otsuka, E., and Ikehara, M. (1990) *J. Biol. Chem.* **265**, 4615–4621.
10. Kanaya, S., Kimura, S., Katsuda, C., and Ikehara, M. (1990) *Biochem. J.* **271**, 59–66.
11. Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzaki, T., and Morikawa, K. (1990) *Nature* **347**, 306–309.
12. Yang, W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. (1990) *Science* **249**, 1398–1405.
13. Doolittle, R., Feng, D. F., Johnson, M. S., and McClure, M. A. (1989) *Q. Rev. Biol.* **64**, 1–30.
14. Tanese, N., and Goff, S. P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1777–1781.
15. Hansen, J., Schulze, T., Mellert, W., and Moelling, K. (1988) *EMBO J.* **7**, 239–243.
16. Levin, J. G., Crouch, R. J., Post, K., Hu, S. T., Mckelrin, D., Zweig, M., Court, D. L., and Gerwin, B. I. (1988) *J. Virol.* **62**, 4376–4380.
17. Krug, M. S., and Berger, S. L. (1989) *Proc. Natl. Acad. Sic. USA* **86**, 3539–3543.
18. Oyama, F., Kikuchi, R., Crouch, R. J., and Uchida, T. (1989) *J. Biol. Chem.* **265**, 18808–18817.
19. Kanaya, S., and Crouch, R. J. (1983) *J. Biol, Chem.* **258**, 1276–1281.
20. Maki, H., Horiuchi, T., and Sekiguchi, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 7137–7141.
21. Itaya, M., McKelvin D., Chatterjie, S. K., and Crouch, R. J. (1991) *Mol. Gen. Genet. in press.*.
22. Hollingsworth, H. C. and Nossal, N. G. (1991) *J. Biol. Chem.* **266**, 1888–1897.
23. Wintersberger, U., Kuhne, C., and Karwan,. R. (1988) *Biochim. Biophys. Acta* **951**, 322–329.
24. Kane, C. M. (1988) *Biochemistry* **27**, 3187–3196.
25. Rong, Y. W. and Carl, P. L. (1990) *Biochemistry*, **29** 383–389.
26. Vonwirth H., Frank, P., Kedinger, C., and Busen, W. (1990) *Biochim. Biophys. acta*, **1087**, 31–38.
27. Oshima, T., and Imahori, K. (1974) *J. Biochem (Tokyo)*, **75**, 179–183.
28. Hayase, Y., Inoue, H., and Ohtsuka, E. (1990) *Biochemistry* **29**, 8793–8797.
29. Tanaka, T., Kawano, N., and Oshima, T. (1981) *J. Biochem. (Tokyo)* **89**, 677–682.

30. Henikoff, S,(1984) *Gene* **28** 351–359.
31. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular* Cloning: *A Laboratory Manual*, Cold Spring Harbor University Press, Cold Spring Harbor.
32. Mandel, M. and Higa A. (1970) *J. Mol. Biol.* **53**, 159–162.
33. Keller, W., and Crouch, R. J. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3360–3364 .
34. Ogawa, T., and Okazaki, T. (1984) *Mol. Gen. Genet.* **193** 231–237.
35. Oshima,T. (1986) in Brock, T. D. (ed. *Thermophiles: General, Molecular, and Applied Microbiology*, John Wiley & Sons, Inc. USA pp. 137–157.
36. Shine J., and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
37. Muto, A. and Osawa, S. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 166–169.
38. Koyama, Y. and Furukawa, K. (1990) *J. Bacteriol.* **172**, 3490–3495.
39. Cox, E. C. and Hormer, D. L. (1986) *J. Mol. Biol.* **190**, 113–117.
40. Quinones, A., Kucherer, C., Piechocki, R., and Messer, W (1987) *Mol. Gen. Genet.* **206**, 95–100.
41. Gualerzi, C. O. and Pon, C. L. (1990) *Biochemisty*, **29**, 5881–5889.
42. Itaya, M. and Crouch, R. J (1991) *Mol. Gen. Genet.* in press.