PROBING SUBSTRATE SPECIFICITY OF RNA REPAIR PROTEIN PNKP/HEN1 FROM *ANABAENA VARIABILIS*

BY

CAN ZHANG

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Biochemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Advisor:

Associate Professor Raven Huang

Abstract

This is an *in vitro* study of the substrate specificity of a RNA-repair protein Pnkp/Hen1 from bacteria *Anabaena Variabilis.* Following the previous discovery of a RNA repair protein—Pnkp/Hen1, which can repair ribotoxin cleaved tRNA in vitro and can add a methyl group at the cleaved site to prevent further cleavage, this study investigated the RNA secondary structure that this protein complex recognizes and repairs. First, three different truncated versions of aspartic acid and arginine tRNAs are produced by deleting the D loop (tRNA-ΔD), TΨC loop (tRNA-ΔT), or both loops (tRNA-ΔDT) from the tRNA secondary structure. Then these three versions of RNA as well as the original tRNA are cleaved at the anticodon loops *in vitro* with Colicin. Then protein Pnkp/Hen1 is used to repair each pair of the cleaved RNAs, and it is found that truncated versions of RNAs could be repaired as efficiently as the original tRNA, suggesting that Pnkp/Hen1 is not tRNA specific. Next tRNAs with only stem-loop structures are produced and tested by Pnkp/Hen1, the repair result is negative in comparison to the repair of tRNA-ΔDT as position control. Then nicked & bulged RNA secondary structures are tested and it turns out either nicked or bulged RNA structures is repaired. However the negative controls of the same experiment shows are ligated by Pnkp/Hen1. Following this discovery, pairs of single strand RNAs with similar or distinct sequences are mixed and tested, and we found that RNAs with a particular structural pattern can be ligated to the target RNA; while others can't be ligated. The study shows that Pnkp/Hen1 repairs a relatively broader spectrum of RNA substrates than we expected but also has secondary structure specificity. Based on the experimental data, a model of preferred RNA substrates of the Pnkp/Hen1 repair system is proposed.

ACKNOWLEDGMENT

I would like to thank my advisor Dr. Raven Huang for supervising my Master thesis, especially I want like to thank him for providing insight and guides for this study and technical support during my research, as well as his generous research assistant funding for the first half year of my Master study. Then I would like to thank Chio Miu Chan (Ella), the PhD graduate in Dr. Huang's laboratory for her work to establish the foundation of this research and for her warm help to set my footing in the laboratory. Also thank Pei Wang and Kiruthika, the two other PhD students in Dr. Huang's laboratory for support and help during my Master thesis.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION TO RNA REPAIR

RNA repair is a relatively rare phenomenon compared to DNA repair. Since DNA is the storage of heritage information which needed to be maintained with high fidelity, DNA repair is highly evolutionarily advantageous to achieve this goal. Normally cells don't need an RNA repair system, because there are multiple copies of the same RNA in cell to cover its functional role in case one of them is damaged and because RNA has a high turnover rate and can be readily transcribed from DNA when damaged. However on occasion of massive RNA destruction—for example when RNAs are cleaved by ribotoxin from a foreign species (Ribotoxin is ribonuclease that cleaves the RNAs involved in protein synthesis, thus killing the cell)—emergent RNA repair system is evolutionary advantageous because it helps bacteria to survive such attack. RNA repair could also be utilized by viruses to ensure the success attack of a cell. For example, when infected by virus, some bacteria produce tRNA endonuclease to cleave their own tRNA, thus shutting down the virus's protein synthesis to prevent the virus from replicate. To counter this bacterial defense system, viruses have RNA repair protein to restore the functional tRNA in the cell to ensure their success of the attack.

One well studied RNA repair system –T4 RNA repair—is an example of the later occasion. Some strains of *Escherichia Coli* cleave their own $tRNA^{Lys}$ at the wobble position in the anticodon loop by turning on the tRNA endonuclease PrrC, thereby shutting off protein synthesis before the phage can replicate, the virus evades the *E. Coli's* host response by

encoding a two-component RNA repair system, consisting of T4 polynucleotide kinase-phosphatase (Pnkp) and T4 RNA ligase 1 (Rnl1). (Amitsur et al. 1987); T4 Pnkp "heals" the broken 2',3' cyclic phosphate and 5'-OH tRNA ends in the cleaved tRNA site by hydrolysis of the 2',3' cyclic phosphate to form a 3'-OH and phosphorylation of the 5'-OH to form a 5'-PO_4, Then T4 Rnl1 ligates the two processed ends to restore the pool of functional tRNA (Amitsur et al. 1987) (Figure 1). Thus, the RNA repair carried out by bacteriophage T4 system ensures the survival of a bacteriophage in a host cell.

A new RNA repair protein Pnkp/Hen1 is discovered recently from our laboratory. First, aspartic acid tRNA ($tRNA^{asp}$) and arginine tRNA ($tRNA^{arg}$) are cleaved *in vitro* with Colicin E and Colicin D respectively at the anticodon loop. Colicin E5(Ogawa et al. 1999) and colin D (Tomita et al. 2000) are a type of ribotoxin produced by some strains of *E.Coli* and has DNase activity or RNase activity. They are tRNA-specific ribotoxins, which make a single cut in the anticodon loop of a specific tRNA, leaving a 2', 3' cyclic phosphate group and 5'-OH group at the cleaved site. Pnkp and Hen1 from the cyanobacteria *Anabaena Variabilis*, physically interacting with each other and forming a heterotetramer, can repair ribotoxin-cleaved tRNAs *in vitro* (Chan et al. 2009b). There are two functional domains in the protein: polynucleotide kinase-phosphatase (Pnkp) and Hen1. Pnkp has 5' kinase, 2', 3' phosphatase, and adenylyltrans-ferase activities (Martins and Shuman 2005), but Pnkp alone could not repair a cleaved RNA (Keppetipola et al. 2007). Hen1 has a methyl transferase domain and a ligase domain and Hen1's the eukaryotic homologue serves as the methyl transferase in RNA interference. During the repair, the kinase of Pnkp adds a phosphate group to the 5'-OH end of the cleaved site, while phosphatase of Pnkp removes the 2', 3' cyclic phosphate group at the cleaved site, exposing 2'-OH and 3'-OH

groups. This way the colicin-cleaved tRNA is “healed”. Then Hen1 can add one methyl group to 2’-OH group of the tRNA while the rest of the protein complex function together as ligase to ligate the 3’-OH group with 5’ phosphate group. After the repair of tRNA by Pnkp/Hen1, one extra methyl group is attached to the previously cleaved site, the adding of methyl group can prevent the same site from being cleaved by Colin again, thus the repaired RNA is “better than new”. (Chan et al, 2009a) (Figure 2)

In contrast to the extensively characterized T4 system, Much less is known about the more recently discovered bacterial Pnkp/Hen1 RNA repair system. Prior studies from the Shuman laboratory and our laboratory have focused on *in vitro* characterization, either of the Pnkp/Hen1 complex (Chan et al. 2009b) or of components of the complex (Martins and Shuman 2005; Keppetipola et al. 2007; Chan et al. 2009a; Jain and Shuman 2010, 2011). The next big question for Pnkp/Hen1 would be what *in vivo* RNA substrate it repairs. If RNA damage is caused internally in bacterial species possessing the Pnkp/Hen1 RNA repair system, it is possible that an *in vivo* study on one of these bacterial species would provide insight into RNA damage and repair. On the other hand, it is possible that RNA damage in these bacteria could be caused by external agents such as ribotoxins released from other surrounding organisms. If that were the case, then RNA repair would only occur when these bacteria live under conditions that induce such damage, which might not be readily reproduced in a laboratory setting.

This thesis study focus on answering this question by finding out the repair efficiency of different RNA substrates by Pnkp/Hen1 *in vitro*. Here I am presenting a comprehensive *in vitro* RNA repair study using Pnkp/Hen1 from *Anabaena variabilis* (*Ava*Pnkp/Hen1) and a variety of

RNA substrates, with the aim of providing insight into the likely *in vivo* RNA substrates of the Pnkp/Hen1 RNA repair system. The first highly likely substrate is tRNA because of its importance in protein synthesis and because tRNA has been used as substrate during the previous study. In order to determine tRNA secondary structure specificity of Pnkp/Hen1, we created three different truncated version of Aspartic acid and Arginine tRNA ($tRNA^{asp}$ and $tRNA^{arg}$) by deleting either the D loop (tRNA-ΔD), TΨC loop (tRNA-ΔT), or both loops (tRNA-ΔDT) from the tRNA secondary structure, then Colicin E or Colicin D are used to cleave either $tRNA^{asp}$ or $tRNA^{arg}$ as well as their truncated versions at the anti-codon sites. A time course repair by Pnkp/Hen1 is carried out on each pair of cleaved RNA. We also tested RNAs with stem loop, dumbbell, nick and bulge secondary structure, as well as single strand RNAs of different sequences. Eventually we found that PnkP/Hen1 may exhibit broader substrate specificity than T4 PnkP/Rnl1. Based on the studies presented here, the likely *in vivo* RNA substrates of Pnkp/Hen1 and the possible biological roles of Pnkp/Hen1 are discussed.

CHAPTER 2

METHODS AND MATERIALS

2.1 PREPARATION OF RECOMBINANT PROTEINS

*Ava*Pnkp/Hen1 and *Ava*Pnkp-N were overexpressed in *E. coli* BL21(DE3) and purified from soluble bacterial lysates asj described previously (Chan et al. 2009b).

2.2 PREPARATION AND PURIFICATION OF TRNAS AND THEIR TRUNCATED VERSIONS

Design and purification of DNA template for *in vitro* transcription

Two different tRNAs are used in this study: $tRNA^{asp}$ and $tRNA^{arg}$. The truncated versions of them are created by deleting either the D loop (tRNA-ΔD), TΨC loop (tRNA-ΔT), or both loops (tRNA-ΔDT). Each RNA is produced by *in vitro* transcription. The DNA primers for each RNA (with added ACGCACGCTGTAATACGACTCACTATA consensus promoter sequence before the 5' +1 Guanine in the RNA sequence) are purchased from integreated DNA technology. **Table 1** shows the sequence of all RNA and corresponding DNA primers used in tRNA truncation study. The DNA primers are purified with 8% denaturing polyacrylamide gel electrophoresis (DPAGE). The gel that contains desired DNA band (located under UV light) is cut out and crushed, then 10 ml elution buffer (4% NH_4OAC, 2.5μM EDTA in DEPC water) is added to the crushed gel and the gel is eluted in a rotator for half an hour. Then the crushed gel and elution buffer mixture is centrifuged for 10 minutes at 3,000 rpm, and then the elution buffer supernatant which contains the DNA is collected. The total supernatant is concentrated to 100μl~400μl depending

on the original volume, then -20 °C 100% ethanol is added to the concentrated buffer to get a 70% final ethanol concentration to precipitate DNA. The ethanol precipitation tube is incubated in -20 °C fridge overnight, and then the solution is centrifuged at 13,000rpm for 30 minutes, and then supernatant is removed leaving only the precipitated DNA pellet to air dry for 10 minutes. Before use the dried purified DNA is dissolved in DEPC H_2O and to get a 2 mg/ml concentration. The DNA absorbance is measured at 280nm using a spectrophotometer in comparison to DEPC water as blank, the extinction coefficients of DNAs are calculated from the DNA sequence, then the DNA concentration is calculated using beer's law. Two DNA templates were mixed in a 1:1 ratio to make double strand DNA concentration 1mg/ml, and the templates are annealed by heating up to 95 °C for 3 min and then slowly cooling to 25 °C in half an hour.

***In vitro* transcription and purification of RNA**

The *in vitro* transcription reaction contains 1.25 mM ATP, CTP, GTP, and UTP, 40 mM Tris-HCl (pH 7.9), 6 mM $MgCl_2$, 50 mM dithiothreitol (DTT), 2 mM spermidine, 1 μM DNA template and 1 mg purified T7 polymerase with DEPC water adding the total volume to 1ml. The reaction tube is incubated at 37 °C for 5-10 hours then stopped by adding 1ml of DPAGE loading buffer (10 mM of EDTA, pH 8.0, in 100% formamide). The RNA is loaded to 10%-12% DPAGE and purified the same way as described in the purification of DNA template. The RNA pellets are saved as dry pellets until right before usage to avoid degradation in water solution. The preparation of ^{33}p internally radiolabeled RNA follows the same protocol as above, only in a smaller scale (40ul), with ^{33}P-α-GTP added to the reaction.

2.3 CLEAVING OF TRNA SUBSTRATE WITH COLICIN AND PURIFICATION OF THE CLEAVED PRODUCT

Mixing "cold" and "hot" RNA before cleavage

The "cold" RNA (un-radiolabeled RNA) dry pellets from previous *in vitro* transcription are dissolved in TE buffer to get a 50μM concentration. And then the 50mM RNA solution is used to dissolve the same "hot" RNA (internally ^{33}P radiolabeled) dry pellets. For example, cold 50μM $tRNA^{asp}$-ΔD solution is used to dissolve hot $tRNA^{asp}$-ΔD, while cold 50μM $tRNA^{arg}$-F solution is used to dissolve hot $tRNA^{arg}$-F. Because the radiolabel RNA concentration is much lower than the unlabeled RNA, it's impossible to visualize the radiolabeled RNA in DPAGE gel under UV light, so the addition of 50μM cold RNA is to help locate the right band in the gel during purification of cleaved product.

Cleavage reaction of RNAs by colicins.

The cleavage reaction of RNA by Colicins contains the following: For $tRNA^{asp}$ and its three truncated versions: 20μl of RNA (~50μM), 4ul 10X cleavage buffer, 100mM MgCl2, 0.5μl of RNase inhibitor, Colicin E, and with DEPC H_2O to bring the final volume up to 40μl. During the addition of each contents, the RNA, DEPC H_2O, cleavage buffer as well as $MgCl_2$ are mixed and fast annealed (heat up to 95°C for 3 minutes then immediately put on ice for 5 minutes), and then RNase inhibitor and Colicin are added. The reaction mixture is incubated at room temperature for 10 minutes, 2 hours, 15minutes, 2.5 hours respectively for $tRNA^{asp}$-F, $tRNA^{asp}$-ΔD, $tRNA^{asp}$-ΔT, $tRNA^{asp}$-ΔDT. The reaction is stopped by mixing 40μl DPAGE loading buffer and heating to 95°C for 3 minutes. For $tRNA^{arg}$ and its three truncated versions, the differences are

that Colin D is used instead of Colicin E, and the reaction times are 1 hour, 3 hours, 1.5 hours and 2 hours repectively for $tRNA^{arg}$-F, $tRNA^{arg}$-ΔD, $tRNA^{arg}$-ΔT, $tRNA^{arg}$-ΔDT.

The cleaved products

Both $tRNA^{asp}$ and $tRNA^{arg}$ and their truncated versions (all 8 different RNAs) are purified by 20% DPAGE gel. The desired bands of cleaved product are located under UV light and the gel containing the desired bands are cut out, crushed, eluted and ethanol precipitated following the same protocol as in DNA purification. 15 cleaved products are obtained, they are: $tRNA^{asp}$-F-5'(35 nucleotides, abbreviated as Asp-F35), $tRNA^{asp}$-F-5'(Asp-F42), Asp-D42, Asp-D20, Asp-T35, Asp-T27, Asp-DT27, Asp-DT20, Arg-F38, Arg-F37, Arg-D37, Arg-D24, Arg-T38, Arg-T22, Arg-DT24/22 (because $tRNA^{asp}$-ΔDT-5' and $tRNA^{asp}$-ΔDT-3' can't be separated by DPAGE gel, so both strands are purified as a mixture). **Table 2** shows the sequence and sizes of all these 15 products. All of them are saved as dried RNA pellets in the tube until use to avoid degradation from water.

2.4 TIME COURSE RNA REPAIR ASSAY ON TRNAS AND THEIR TRUNCATED VERSIONS

Verification and adjusting of Radioactivity of RNAs before repair.

Each RNA product is dissolved in 10 μl H_2O, and 1 μl is taken from this and mixed with 6 μl DPAGE loading buffer, the mixture is heated up to 95 °C for 3 minutes, and then loaded to a 15% DPAGE thin gel, the electrophoresis is done at 900 volts for 40minutes, then the DPAGE gel is transferred to a filter paper and covered by syring wrap. The gel is then dried by using the Sep-pak system. Next the dried gel is exposed to film for overnight (about 14 hours) before the

image is scanned and quantified by using a PhosphorImager system (Molecular Dynamics). Each band showed up in this prep DPAGE is indicating the concentration and radioactivity of RNAs in each tube. Tubes are diluted or not to obtain a final homogeneous radioactivity for all the RNA substrates.

Time course repair of $tRNA^{arg}$ and $tRNA^{asp}$

The tRNA time course repair reaction is 20 μl scale containing: 25 mM Tris-HCl (pH 8.0), 50 mM KCl, 2.5 mM MgCl2, 0.05 mM EDTA, 2.5% glycerol, 5 mM DTT, 0.2 mM ATP, 0.5 mM MnCl2, 0.05mM S-AdoMet (Methyl group donor), internally ^{33}P-radiolabeled cleaved tRNA, RNasin (RNase inhibitor) and Pnkp/Hen1. Firstly RNA substrate pairs that reconstitute desired un-cleaved products (e.g. Asp-F35 and Asp-F42 are mixed to re-constitute $tRNA^{asp}$-F) are mixed together with DEPC H_2O and buffer, a slow annealing is performed to maximize the amount of repairable secondary RNA structure by heating the RNA solution to 95 °C and let it slowly cool down to 25 °C during a 30 minutes time period. Then the rest of the reaction mixture is added into the reaction including protein Pnkp/Hen1 and RNase inhibitor. The 20 μl reaction mixture was incubated at 37°C. At each time point 0 min, 5 minutes, 15 minutes, 30 minutes and 60 minutes from adding of the Pnkp/Hen1, 3 μl of the reaction mixture is taken out and quenched by adding 4 μl of DPAGE loading buffer and heating up to 95 °C for 3 minutes. All the samples are collected and analyzed by 15% DPAGE thin gel. The gel is dried and exposed as well as the image is obtained and quantified as described in the 1.3.1 section of methods.

2.5 PREPARATION AND REPAIR OF STEM LOOP RNAS

Preparation of stem loop RNAs

Four strands of RNA that re-constitute the anticodon loop part of the $tRNA^{arg}$ and $tRNA^{asp}$ (with a break at the Colicin cleavage site) are purchased from IDT, named as $tRNA^{Arg}$-AC-SL-5', $tRNA^{Arg}$-AC-SL-3', $tRNA^{Asp}$- AC-SL-5', $tRNA^{Asp}$-AC-SL-3'. Four strands of RNA that reconstitute $tRNA^{arg}$-ΔDT and $tRNA^{asp}$-ΔDT but without the residual bulge of TΨC loop and D loop are also purchased from the same source, and they are named as $tRNA^{Arg}$-ΔDT-ΔB-5', $tRNA^{Arg}$-ΔDT-ΔB-3', $tRNA^{Asp}$-ΔDT-ΔB-5' and $tRNA^{Asp}$-ΔDT-ΔB-3'. **Table 3** showed the sequence of these 8 synthetic RNAs. They are purified by 20% DPAGE gel followed by elution and ethanol precipitation. The final RNA dry pellets are saved until use. $tRNA^{arg}$ -ΔDT and $tRNA^{asp}$ -ΔDT produced from the tRNA truncation study are used as positive control.

Radiolabeling stem loop RNA with ^{33}p-γ-ATP

Each synthetic stem loop RNA is radiolabeled by ^{33}p-γ-ATP using Pnkp-N, the N terminal domain of Pnkp, which has strong kinase activity (unpublished data). The reaction is 10ul scale containing 1 μM stem loop RNA substrate, ^{33}p-γ-ATP, 25 mM Tris-HCl (pH 8.0), 50 mM KCl, 2.5 mM MgCl2, 0.05 mM EDTA, 2.5% glycerol, 5 mM DTT, 0.2 mM ATP, 0.5 mM MnCl2, 0.05mM S-AdoMet, RNasin and Pnkp-N protein. This phosphorylation reaction is incubated at 37 °C for 40 minutes.

Repair Assay of stem loop RNA

Then 0.5 μM radiolabeled RNA pairs (either internally radiolabeled $tRNA^{arg}$-DT, $tRNA^{asp}$-DT as positive control or externally radiolabeled stem loop RNA as described above) are mixed to reconstitute the desired stem loop structure (e.g. radiolabeled Asp-DT-stem loop-5' and Asp-DT-stem loop-3' are mixed together to reconstitute the anti-codon stem loop of $tRNA^{asp}$), after

slow anneal, RNasin and Pnkp/Hen1 protein is added to the reaction. The repair assay is incubated at 37 °C for 40 minutes before quenched by DPAGE loading buffer. 6 μl of RNA and DPAGE loading buffer mixture is loaded to a 20% DPAGE thin gel along with Radiolabeled RNA molecular marker. The DPAGE gel is later dried, exposed and scanned the same way as previously.

2.6 PREPARATION AND REPAIR OF NICKED AND BULGED DOUBLE STRAND RNAS

A 27-nucleotide RNA (lead RNA) is taken from storage of previous research in the same lab. A ^{33}p internally labeled 30-nucleotide RNA (tail RNA) is produced and purified by *in vitro* transcription. Four Bridge RNA strands are produced by *in vitro* transcription as well. **Table 4** shows the sequence of Lead RNA, Tail RNA and the 4 different bridge RNAs as well as the DNA primers for Tail and Bridge RNAs.

In the repair reaction of nicked and bulged RNA, only the tail RNA is internally radiolabeled with ^{33}p, in order to tell if any repair/ligation reaction happens, one needs to compare the migration distance of tail RNA on the RNA molecular standard to find the size of repaired/ligated product (which is also internally labeled because of tail RNA). If the size of the migrated tail RNA is the same size as predicted by ligated product, it means the tail RNA is ligated to other non-radiolabeled RNAs.

The repair reaction contains 1μM of radiolabeled tail RNA and 1 μM of each other un-radiolabeled RNA. The rest of the reaction recipe is the same as in the time course tRNA repair and externally labeled stem loop repair. Reaction is incubated at 37 °C for 40 minutes, and then only 1 μl of the reaction mix is taken out to be mixed with 1 μl of very high concentration

(50 μM) un-radiolabeled tail RNA and 5 μl DPAGE loading buffer. The purpose of this step is to replace the binding of radiolabeled tail RNA & repair product to bridge RNA with un-radiolabeled tail RNA, so that the replaced "free" radiolabeled RNA can easily migrate in the DPAGE gel according to its right size. The rest of process is the same as tRNA time course assay and stem loop repair till eventually get the scanned gel of radioactivity.

2.7 Preparation and repair of random single strand RNAs

7 different short RNAs (S1-S7) as well as lead RNA are mixed with tail RNA to test the repair. All these 7 RNAs are obtained from the previous research in the laboratory. **Table 5** shows the sequences of all the RNA used in this random repair study (Including tail and lead RNA). Each of the lead RNA and S1-S7 RNA is mixed with tail RNA and a 10 μl repair reaction containing 1 μM each RNA is performed. The rest the procedure is the same as described in nick & bulge repair.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 TIME COURSE REPAIR OF TRUNCATED VERSIONS OF TRNA ASPARTIC ACID AND ARGININE

Previously, several RNA ligase enzymes with RNA editing or repair abilities have been biochemically characterized (Amitsur et al. 1987; Ho and Shuman 2002; Simpson et al. 2003; Martins and Shuman 2004a,b). Among them, the T4Pnkp/Rnl1 is the most extensively characterized and the *in vivo* RNA repair target of the T4 system was specifically $tRNA^{Lys}$ cleaved by the endogenous ribonuclease PrrC. Furthermore, the T4Pnkp/Rnl1 repair system appears to have evolved to optimize RNA repair of the cleaved tRNA; this has been suggested by respective T4Pnkp/tRNA and T4Rnl1/tRNA docking experiments (Galburt et al. 2002; El Omari et al. 2006), as well as a subsequent study by Shuman and coworkers (Nandakumar et al. 2008). Using both the cleaved full-length and tRNA deletion mutants, Shuman and coworkers were able to demonstrate that the efficiency of repair for the cleaved full-length tRNA is significantly higher than the ones for the cleaved tRNA deletion mutants (Nandakumar et al. 2008).

Speculating that the target of bacterial Pnkp/Hen1 RNA repair system is also the cleaved tRNAs, we performed similar experiments with cleaved tRNAAsp and its deletion derivatives. Unlike the *E. coli* $tRNA^{Lys}$ cleaved by PrrC in the T4 RNA repair system; our tRNAs are cleaved by colicin E5 at a position near the wobble nucleotide (Figure 3).

For $tRNA^{asp}$ and the truncated versions, in full length $tRNA^{asp}$ ($tRNA^{asp}$-F), 65% of all the RNA is repaired after 1 hour, which is very efficient. In comparison, the repair percentage of $tRNA^{asp}$-

ΔD, which is the tRNA with deleted D stem-loop, is 15%, much lower than the repair $tRNA^{asp}$-F. The repair percentage of $tRNA^{asp}$-ΔT, which is the tRNA with deleted TψC stem-loop, is not as bad, with a percentage of 38%. Deletion of both D stem-loop and TψC stem-loop results in a very bad repair percentage – 13%. (Figure 4A,C)

This result seems to support our hypothesis that tRNA is the *in vivo* substrate of protein Pnkp/Hen1, by deleting the TΨC loop or D loop of the full length $tRNA^{asp}$ the three dimensional structure change could render the truncated tRNA less affinity to Pnkp/Hen1, thus lower repair efficiency.

Next, we carried out RNA repair of the cleaved $tRNA^{Arg}$ and its deletion derivatives. The substrates $tRNA^{Arg}$ and its deletion derivatives were cleaved by colicin D at a position that is significantly different from the one by colicin E5 (Figure 3). Surprisingly, this time the $tRNA^{arg}$-F has only 29% of repair yield, compared to the 72%, 73% and 64% of $tRNA^{arg}$-ΔD, $tRNA^{arg}$-ΔT and $tRNA^{arg}$-ΔDT. Compared with the repair of the cleaved $tRNA^{Asp}$, cleaved full-length $tRNA^{Arg}$ was repaired less efficiently. Unlike $tRNA^{Asp}$ deletion derivatives, deletion of the D stem–loop, TΨC stem–loop, or both in $tRNA^{Arg}$ resulted in RNA substrates that were much more efficiently repaired (Figure 4B,D). The efficiency of repair of these cleaved $tRNA^{Arg}$ deletion derivatives by *Ava*Pnkp/Hen1 is quite remarkable. For example, after only 15 min of repair reaction, the repair yields of the cleaved $tRNA^{Arg}$-ΔD, $tRNA^{Arg}$-ΔT, and $tRNA^{Arg}$- ΔDT were 54%, 67%, and 53%, respectively (Figure 4D). At that time point, the repair yields of the cleaved $tRNA^{Asp}$ and $tRNA^{Arg}$ were 22% and 12%, respectively (Figure 4C).

Additionally in the fourth panel of the figure, above where the ligated tRNA-ΔDT product is, there are extra ligation products of higher molecular weight (Figure 4A, marked with *). The explanation for these bands is that repaired product undergoes further rounds of ligation by Pnkp/Hen1 thus generating various secondary structured RNA of larger size. The reason why these extra bands are only observed in ΔDT sample but not the others could be that tRNA-F, tRNA-ΔD, tRNA-ΔT have more base pairs that that of tRNA-ΔDT, thus didn't provide alternative base pair for alternative ligation to happen. The fact is that the deleted TΨC loop and D loop in tRNA-ΔDT form two large bulges, which are thermodynamically less stable than the TΨC loop or D loop, so the secondary structure of tRNA-ΔDT could easily rearrange in given buffer conditions, giving rise to possibility of various different ligation combination. The existence of these extra bands also support that Pnkp/Hen1 is more general RNA repair protein than tRNA specific.

These results overthrow the previous hypothesis about the tRNA specificity of Pnkp/Hen1. It leads us to explore the possible secondary structure that Pnkp/Hen1 could recognize and repairs in a much broader spectrum, the result of which will be explained in the following experiment.

3.2 REPAIR OF STEM-LOOP RNA

RNA stem loop structure is one of the most common secondary structures in different RNAs, including the anti-codon loop of tRNA. Having demonstrated that the drastically deleted tRNAs (e.g., tRNA-ΔDT) could still be effectively repaired by *Ava*Pnkp/Hen1, we next tested if further truncated RNA substrates would be repaired, aiming to define a minimum RNA repair

substrate—step-loop structure. A cleavage in the stem loop part of various RNA should also be very common in cells because of stem loop's single strand is naturally less stable than double strand.

A total of four additional RNA substrates were prepared for the RNA repair assays. The first two are further deletion of tRNA-ΔDT, with the bulge region between the acceptor stem and anticodon stem deleted (named tRNA-ΔDT-ΔB). The deletion of the bulge region produces a stem–loop RNA with a long and continuous stem (Figure 5A, middle). The other two RNAs correspond to the anticodon stem–loops of $tRNA^{Asp}$ and $tRNA^{Arg}$ (named tRNA-AC-SL). To increase the stability of these two RNAs and to make the in vitro repair assay feasible, two additional base pairs were added on the stem (Figure 5A, left).

Because colicin D did not cleave $tRNA^{Arg}$-ΔDT-ΔB and $tRNA^{Arg}$-AC-SL, we had to reconstitute these four RNA substrates differently. Specifically, we first prepared the 5' half and the 3' half of each RNA substrate individually via chemical synthesis (tRNA-AC-SL) or in vitro transcription (tRNA-ΔDT-ΔB), followed by mixing and annealing the two halves. Thus, these RNA substrates do not have the 2', 3'-cyclic phosphate at the 3' terminus of the 5' half. To ensure that the missing 2', 3'-cyclic phosphate does not significantly affect the outcome of RNA repair, we used the same approach to prepare tRNA-ΔDT substrates as positive controls (Figure 5A, right).

Each of the six RNA substrates was incubated with either Pnkp-N (containing only the kinase and phosphatase domains of Pnkp) or Pnkp/Hen1 in the presence of ^{33}P-γ-ATP, and the resulting products were analyzed by a denaturing gel (Figure 5B). The broken $tRNA^{Asp}$-ΔDT and $tRNA^{Arg}$-ΔDT were efficiently repaired (Figure 5B, lanes 9-12), and their repair yields were

comparable to the ones observed previously (Figure 4A, B, far right panel), indicating that the lack of the 2', 3'-cyclic phosphate does not significantly affect the outcome of RNA repair.

Deleting the bulge in tRNA-ΔDT has a significant effect on RNA repair. Although a repaired product was produced with the broken $tRNA^{Arg}$-ΔDT-ΔB as the substrate (Figure 5B, lane 8 marked with an *), its size (z29 nt) did not match the expected size of 35 nt that would have resulted from the ligation of the 5' half to the 3' half. Therefore, the identity of the repaired product is unknown. It is possible that, as in the case of $tRNA^{Asp}$-ΔDT (Figure 4A, far right panel), continuous 4Gs in $tRNA^{Asp}$-ΔDT-ΔB-5' (Table 3) could be the culprit of this repaired product. This uncertainty, however, does not affect our general conclusion that the deletion of the bulge region results in a RNA substrate with significantly reduced repair efficiency by *Ava*Pnkp/Hen1 (Figure 5B, lanes 8 and 12).

Furthermore, deletion of the bulge region in $tRNA^{Asp}$-ΔDT resulted in a RNA substrate that was not repaired at all by *Ava*Pnkp/Hen1 (Figure 5B, lane 6). Further shortening the stem of $tRNA^{Arg}$-ΔDT-ΔB produced an unrepaired RNA substrate (Figure 5B, lane 4), and like $tRNA^{Asp}$-ΔDT-ΔB, $tRNA^{Asp}$-AC-SL was not a substrate of *Ava*Pnkp/Hen1 (Figure 5B, lane 2).

But it is important to mention that the structure of RNA in test tubes are also affected by the buffer content, temperature and various other factors, so our designed secondary structure according to base pairs might not be there in actual test tube.

3.3 REPAIR OF NICKED AND BULGED RNA

In addition to pure ssRNAs and loops of the stem–loop RNAs, bulges that interrupt continuation of double-stranded RNAs are also frequently present as ssRNAs in living cells. Therefore, we decided to carry out an investigation to see if *Ava*Pnkp/Hen1 was able to repair RNAs damaged in the bulge region.

Two ssRNAs available in our laboratory (named Lead and Tail RNAs, respectively) are used as the targets of repair. The tail and lead RNAs were brought together by base-pairing with a complementary strand (named Bridge). By varying the nucleotide sequence of Bridge RNA, four RNA repair substrates differing in their secondary structures were produced (Figure 6A-D):

Bridge 1 connects the leader and tail RNA in perfect base pair, leaving only a Nick at the 3' end of leader RNA and 5' end of Tail RNA (Figure 6A); Bridge 2 base pairs with leader and tail RNA in such way that four bases from the 3' end of leader RNA and four bases from the 5' end of tail RNA protrude out of the RNA double strand, if the 4 nucleotide overhang from each Leader and Tail RNA got ligated, there will be a 8 nucleotide bulge structure protruding out of the double stranded RNA (Figure 6B); Bridge 3 is similar to Bridge 2, only that it also forms another 4 nucleotide bulge at the side of Bridge RNA (Figure 6C); Bridge 4 forms a 4 nucleotide larger bulge than Bridge 3 (Figure 6D).

To simplify interpretation of repair results, only Tail RNA was radiolabeled with ^{33}P (Figure 6, marked with a *). The Tail RNA is mixed with Leader RNA and Bridge RNA, and the repair assay is done with Pnkp/Hen1. The size of the ligated product (which contains 33p thus visible) minus the size of Tail RNA (30nt) will show the size of the strand that's ligated to the Tail RNA. Since the size of the Leader RNA is 27 nucleotide, and we are expecting Pnkp/Hen1 to ligate the 3'

end of Leader RNA and the 5’ end fof tail RNA (if it does repair nick or bulge structure), then the size of the product should be 57. For negative control, we did the same repair assay on Tail RNA alone, Tail RNA with Leader RNA, and Tail RNA with four bridge RNAs.

First, only a small amount of the repaired products was produced with the nicked and bulged RNA substrates (Figure 6E, lanes 8–11), indicating that, unlike the cleaved stem–loop RNAs, RNAs damaged within the double-stranded region (nicked RNA) (Figure 6A) or in the bulge region of the double-stranded RNA cannot be effectively repaired by *Ava*Pnkp/Hen1.

Surprisingly, efficient RNA repair was observed with some RNA substrates originally designed as negative controls. Specifically, in the absence of Lead RNA, Tail RNA was efficiently ligated to Bridge3 and Bridge4 RNAs (Figure 6E, lanes 6, 7), but not Bridge1 and Bridge2 RNAs (Figure 6E, lanes 4, 5). The result suggest that the Leader RNA and the Tail RNA can form a structure pattern that gives very high Pnkp/Hen1 repair affinity, because can be seen from the figure the most of the tail RNA is converted to ligated 57mer (Leader + Tail).

Without Lead RNA, Tail RNA and Bridge3 or Bridge4 RNAs form a double-stranded RNA with a 4-nt single-stranded 59 overhang of Tail RNA, and a 27-nt or 31-nt single-stranded 3’ overhang of Bridge3 or Bridge4 (Figure 7C, D). Efficient RNA repair of Tail/Bridge3 and Tail/Bridge4 combinations indicates ligation of the 5’ overhang of Tail RNA to the 3’ overhang of the Bridge3 or Bridge4 RNAs to form a 31-nt or 35-nt size loop. The inability of Tail RNA to be ligated to Bridge1 (Figure 6E, lane 4) can be explained by the lack of the flexible single-stranded 5’ overhang of Tail RNA in the Tail/Bridge1 combination (Figure 7A). On the other hand, the negative result of RNA repair in the Tail/Bridge2 combination was initially puzzling, as it should

have produced a result similar to the Tail/Bridge3 combination because they appeared to have similar single-stranded 5’ and 3’ overhangs (Figure 6B,C). Further inspection of the nucleotide sequences of these two Bridge RNAs explains why. In the absence of Lead RNA, three of the four nucleotides in Bridge2 originally designed to base-pair with the 3’ end of Lead RNA are now able to base-pair with the single-stranded 5’ overhang of Tail RNA (Figure 7B). In essence, the Tail/Bridge2 combination has a secondary structure similar to the one of the Tail/Bridge1 combination (Figure 7A, B). It is therefore only logical that RNA repair of these two combinations by *Ava*Pnkp/Hen1 produced similar results (Figure 6E, lanes 4, 5).

Another observation was the efficient ligation of Tail RNA to Lead RNA in the absence of any Bridge RNAs (Figure 6E, lane 3). Again, inspection of the nucleotide sequences of these two RNAs provides rationale as to why these two RNAs were efficiently ligated by *Ava*Pnkp/Hen1. Within a region of 11 nt near the 3’ end of Lead RNA and the 5’end of Tail RNA, these two RNAs are able to form a duplex of a total of 10 bp, and eight of them are continuous (Figure 8A). Base-pairing between Lead RNA and Tail RNA produces an RNA substrate with a 12-nt single-stranded 5’ overhang of Lead RNA and a 19-nt single-stranded 3’ overhang of Tail RNA. Presumably, these two ends were ligated by *Ava*Pnkp/Hen1 during RNA repair, producing a stem–loop RNA with a 31-nt size loop.

In addition to the efficient RNA repair of the Tail/Lead, Tail/Bridge3, and Tail/Bridg4 combinations, a small amount of self-ligated product of Tail RNA was also observed in the absence of other RNAs (Figure 6E, lane 2). A possible explanation for this inefficient self-ligation is the potential formation of four GC pairs via 5’ ends of two Tail RNAs.

3.4 REPAIR OF RANDOM SELECTED RNA

Based on the inspiration from the nicked and bulged RNA repair experiments, which showed a very unpredictable way of RNA repair by Pnkp/Hen1, we decided to randomly select different RNA from the pool of previous research, and mix this RNA with tail RNA to see the repair result, at the same time the Tail RNA plus Lead RNA combination is used at positive control.

We have chosen seven ssRNAs (Table 5) for the study. Based on their nucleotide sequences, these seven ssRNAs can be divided into three groups. S1 and S2 RNAs belong to the first group, both of which are related to Lead RNA due to their high nucleotide sequence identities (Figure 8B). Indeed, like Lead RNA, S1 and S2 RNAs were efficiently ligated to Tail RNA by *Ava*Pnkp/Hen1 (Figure 9, lanes3, 4, 5). The positions of the ligation products in the gel are consistent with the expected ligation sizes of Lead (27 nt), S1 (21 nt), and S2 (24 nt) RNAs to Tail RNA (30 nt). S3 is the only RNA that belongs to the second group. S3 is also related to Lead RNA. But instead of sharing the same nucleotide sequences with Lead RNA as S1 and S2, S3 is complementary to Lead RNA (Figure 8C). RNA ligation of S3 to Tail RNA was not observed (Figure 9, lane 6). Therefore, the positive results from S1 and S2, and the negative result from S3, provide further support for the explanation that efficient ligation of Tail RNA to Lead RNA is due to formation of a secondary structure as depicted in Figure 8A. RNAs in the third group are S4, S5, S6, and S7 (ranging in size from 17–35 nt) (Table 5). Based on their nucleotide sequences, these RNAs are not related to Lead RNA or to Tail RNA. None of them were able to ligate to Tail RNA (Figure 9, lanes 7–10). Therefore, we conclude that if two ssRNAs do not physically interact

with each other via base-pairing to bring them together, their ligation by *Ava*Pnkp/Hen1 will not occur.

It is interesting to notice that whereas a small amount of self-ligated product was produced with Tail RNA alone (Figure 9, lane 2), such a ligation product was not observed with the addition of S3–S7 (Figure 9, lanes 6–10). Presumably, the addition of excess ssRNAs S3–S7 sequesters a significant amount of enzyme, making inefficient self-ligation of Tail RNA unlikely to occur.

CHAPTER 4

CONCLUSION

Based on results of RNA repair assays presented in this study, an effective RNA substrate for repair by *Ava*Pnkp/Hen1 is summarized in Figure 10.

To be an effective RNA repair substrate of *Ava*Pnkp/Hen1, two elements are essential. The first element is the two single-stranded ends from the original RNA cleavage, which are the repair targets and are depicted here as 5’ and 3’ overhangs (Figure 10). The second element is a means to bring these two ends together to facilitate repair by *Ava*Pnkp/Hen1. Our data indicate that the lengths of both 5’ and 3’ overhangs are quite flexible, and a stable stem is the most effective method to bring these two overhangs together for efficient repair.

For the 5’ overhang, we have observed efficient repair of RNA substrates with its length from 4–12 nt (Figure 6E, lanes 3, 6, 7, Figure 7C, D, Figure 8A). It is likely that RNA substrates with the 5’ overhang >12 nt can also be efficiently repaired by *Ava*Pnkp/Hen1, but it requires future experiments for confirmation. The shortest length of 5’ overhang allowed for effective repair is in fact zero, demonstrated by the effective repair of the cleaved $tRNA^{Arg}$ and its truncated derivatives (Figures 3, 4). However, $tRNA^{Arg}$ and its deletion derivatives are likely to be an exception as similar configurations in the Tail/Bridge1 and Tail/Bridge2 combinations did not work well for repair (Figure 6E, lanes 4, 5; Figure 7A, B). It is possible that, after cleavage by colicin D, the 5’-terminal nucleotide (nucleotide 39) in $tRNA^{Arg}$ and its deletion derivatives might no longer base-pair with its partner (nucleotide 31) due to limited stability of the anticodon

stem (5 bp). This would provide the 5' end some flexibility for the repair reactions (phosphorylation and ligation) to be carried out. This hypothesis is consistent with the observations that the cleaved $tRNA^{Arg}$ deletion mutant is repaired more efficiently than the full-length $tRNA^{Arg}$ (Figure 2B, D).

Deleting the D or/and TψC stem–loop in $tRNA^{Arg}$ would likely produce less stable RNAs, which might be better substrates of *Ava*Pnkp/Hen1 due to more potential flexibility of the 5'-terminal nucleotide. Conversely, because both Tail/Bridge1 and Tail/Bridge2 have long stems (Figure 7A, B), the flexibility of the 5'-terminal nucleotide seen in $tRNA^{Arg}$ and its deletion derivatives might not be present in the Tail/Bridge1 and Tail/Bridge2 combinations, resulting in inefficient repair of Tail/Bridge1 and Tail/Bridge2 (Figure 6E, lanes 4, 5). For the 3' overhang, RNA substrates with its length ranging from 3 nt ($tRNA^{Asp}$ and its derivative) (Figure 4A, C) to 31 nt (Tail/Bridge4 combination) (Figure 6E, lane 7) were seen to be efficiently repaired by *Ava*Pnkp/Hen1. It is possible that the 3' overhang could even be <3 nt, but we have not tested this possibility. We propose that a stable stem, as depicted Stem 1 in Figure 6, is the best means to bring two overhangs together for efficient repair by *Ava*Pnkp/Hen1 based on the following observations. First, *Ava*Pnkp/Hen1 is unable to repair an RNA substrate that does not form a secondary structure to bring the two ends together. This is supported by the negative results of ligating Tail RNA to other unrelated ssRNAs (Fig 9, lanes 6–10).

The general conclusions based on our *in vitro* RNA repair study of Pnkp/Hen1 from *A. variabilis* could be applicable for Pnkp/Hen1 from other bacterial species. However, Pnkp/Hen1 in different bacterial might have evolved different RNA substrate specificity due to the

requirement of repairing specific RNA damages caused by RNA-damaging agents present in their local living environments. Future studies similar to the one presented here, but using Pnkp/Hen1 from other organisms, should address this issue.

In addition, recently it is found a second subfamily of the bacterial Pnkp/Hen1 RNA repair system that is significantly different from the first subfamily in the amino acid sequences of both Pnkp and Hen1 (data not shown). Therefore, it is possible that Pnkp/Hen1 from the second subfamily might have a different RNA substrate specificity from those in the first subfamily, of which *Ava*Pnkp/Hen1 is a member.

FIGURES

Figure 1

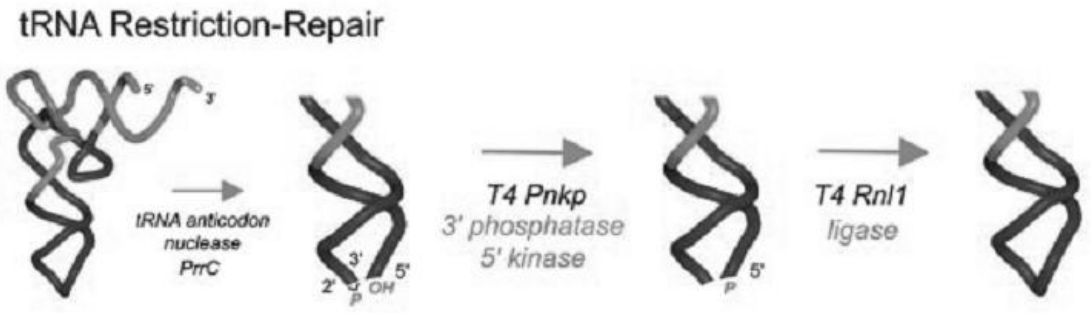

Figure 1: Under bacteriophage T4's infection, some strains of *Escherichia Coli* will produce the enzyme PrrC to cleave its tRNA in the anticodon loop, leaving 2', 3'-cyclic phosphate and 5'-OH ends. To counter this reaction, T4's repair system activates T4 Pnkp, which removes the phosphate at the 3' end and phosphorylates the 5' terminus. Then T4 Rnl1 seals the 3'-OH and 5'-PO_4 termini to form a standard 3'-5'phosphodiester linkage and thus restore tRNA function in protein synthesis. (Schwer et al., 2004)

Figure 2

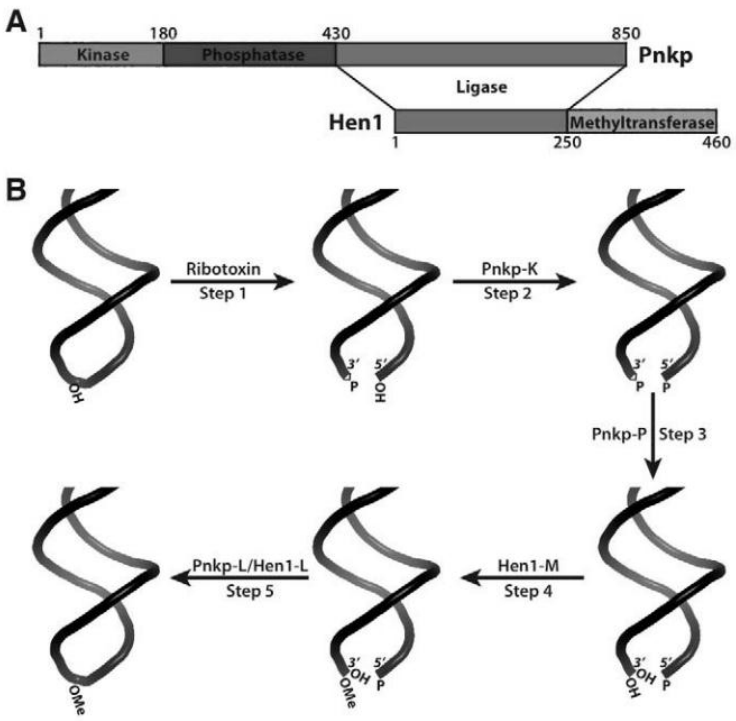

Figure 2A: RNA repair protein Pnkp/Hen1 from bacteria *Anabaena Variabilis* have two functional domains: polynucleotide kinase-phosphatase (Pnkp) and Hen1. Pnkp has kinase, phosphatase, and adenylyltrans-ferase activities. Hen1 has a methyl transferase domain and a ligase domain. **2B:** Step1, Ribotoxin cleaves RNA leaving a 2', 3' cyclic phosphate group and a 5' OH group. Step 2: During the repair, the kinase of Pnkp adds a phosphate group to the 5'-OH end of the cleaved site. Step 3: phosphatase of Pnkp removes the 2', 3' cyclic phosphate group at the cleaved site, exposing 2'-OH and 3'-OH groups. Step 4: Hen1 adds one methyl group to 2'-OH group of the tRNA Step 5: The ligase domain of both Hen1 and Pnkp ligates the 3'-OH group with 5' phosphate group. After the repair of tRNA by Pnkp/Hen1, one extra methyl group is attached to the previously cleaved site.

Figure 3

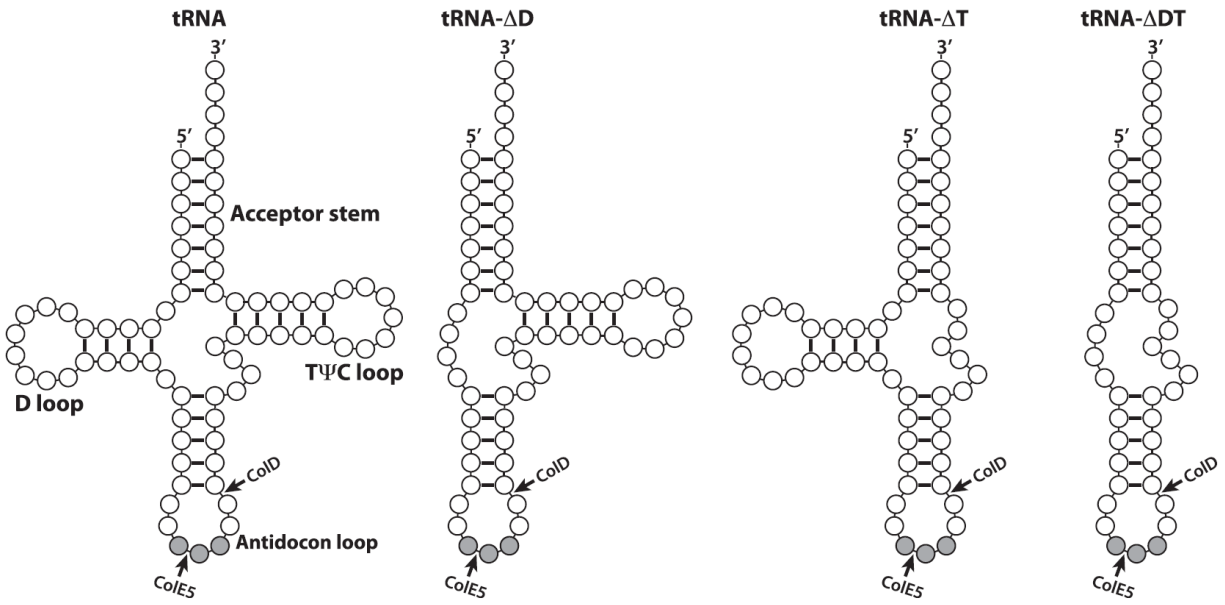

Figure 3: Cloverleaf view of full-length and tRNA deletion mutants employed in this study. Nucleotides are depicted as cycles, and they are linked with each other by short thin lines. Base pairs are represented with thick lines. The cleavage sites of $tRNA^{Asp}$ by colicin E5 (ColE5) and $tRNA^{Arg}$ by colicin D (ColD) are marked with arrows. Three nucleotides constituting the anticodon of tRNA are highlighted in gray.

Figure 4

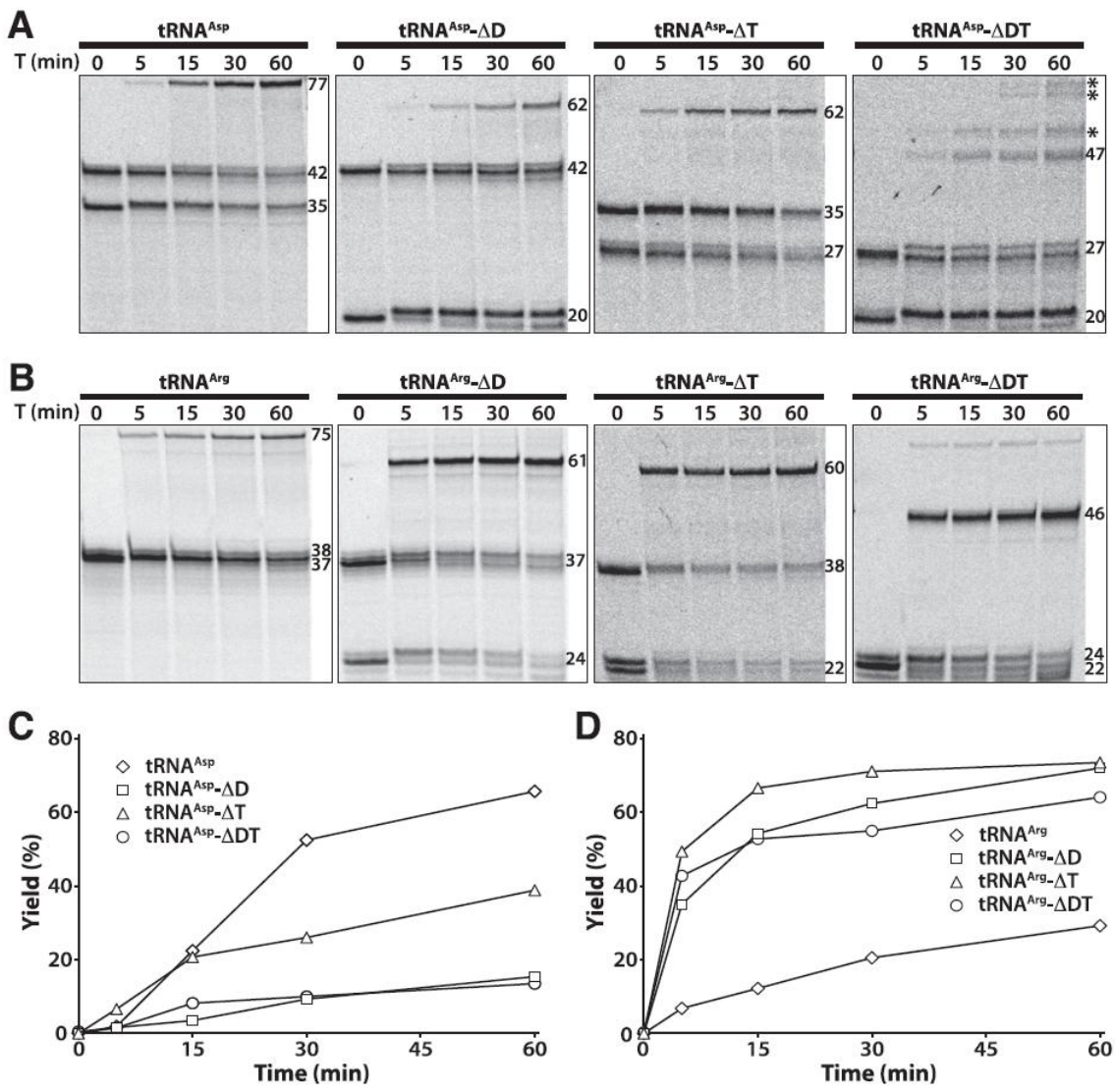

Figure 4: Repair of cleaved full-length and tRNA deletion mutants depicted in Figure 3 by *Ava*Pnkp/Hen1. **A,C**: Time course of RNA repair of $tRNA^{Asp}$ and its deletion derivatives. **B,D:** Time course of RNA repair of $tRNA^{Arg}$ and its deletion derivatives. The positions of RNA substrates and repair products are marked with numbers representing their sizes in nucleotides. T indicates time in minutes.

Figure 5

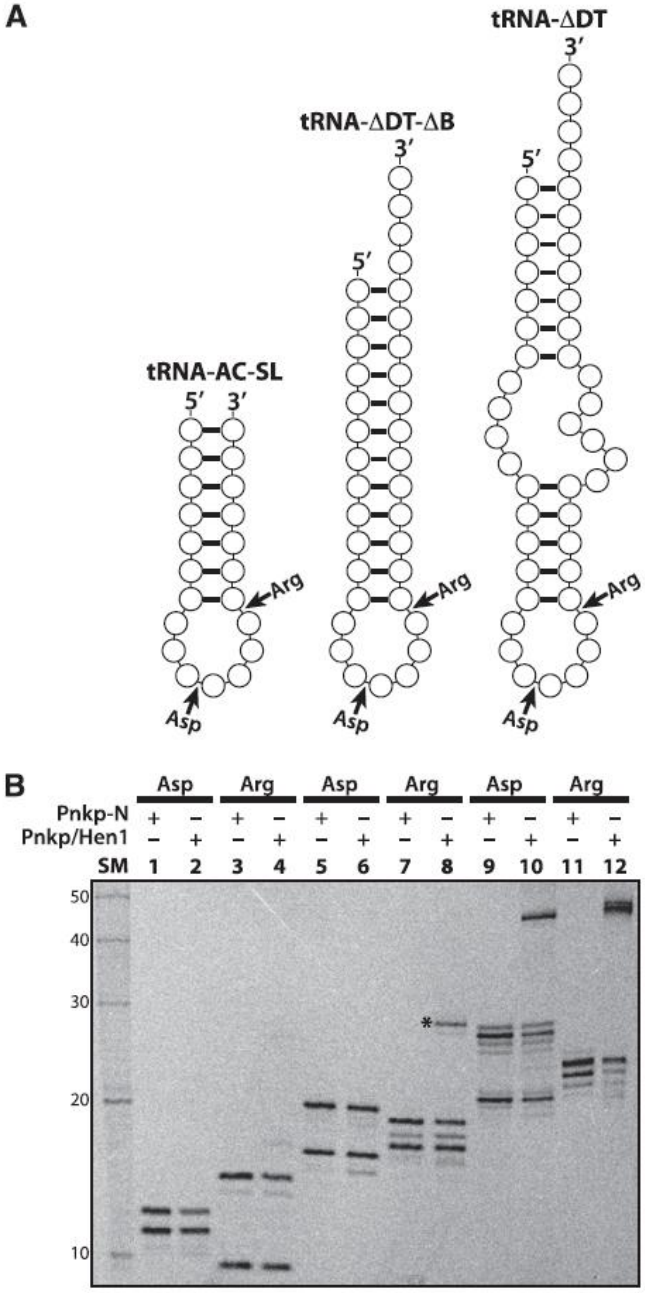

Figure 5 continue

Figure 5: Repair of broken stem–loop RNAs by *Ava*Pnkp/Hen1. **A**: Schematic view of three sets of broken stem–loop RNAs used for the assay. Since each substrate is reconstituted by mixing and annealing separately prepared 5'-half and 3'-half RNAs, its breaking point is marked with an arrow and labeled as Asp (for nucleotide sequence matching $tRNA^{Asp}$) or Arg (for nucleotide sequence matching $tRNA^{Arg}$). **B:** Analysis of phosphorylation and repair reactions by DPAGE. Pnkp-N indicates the N-terminal half of Pnkp consisting of the kinase and phosphatase domains. SM indicates RNA size marker.

Figure 6

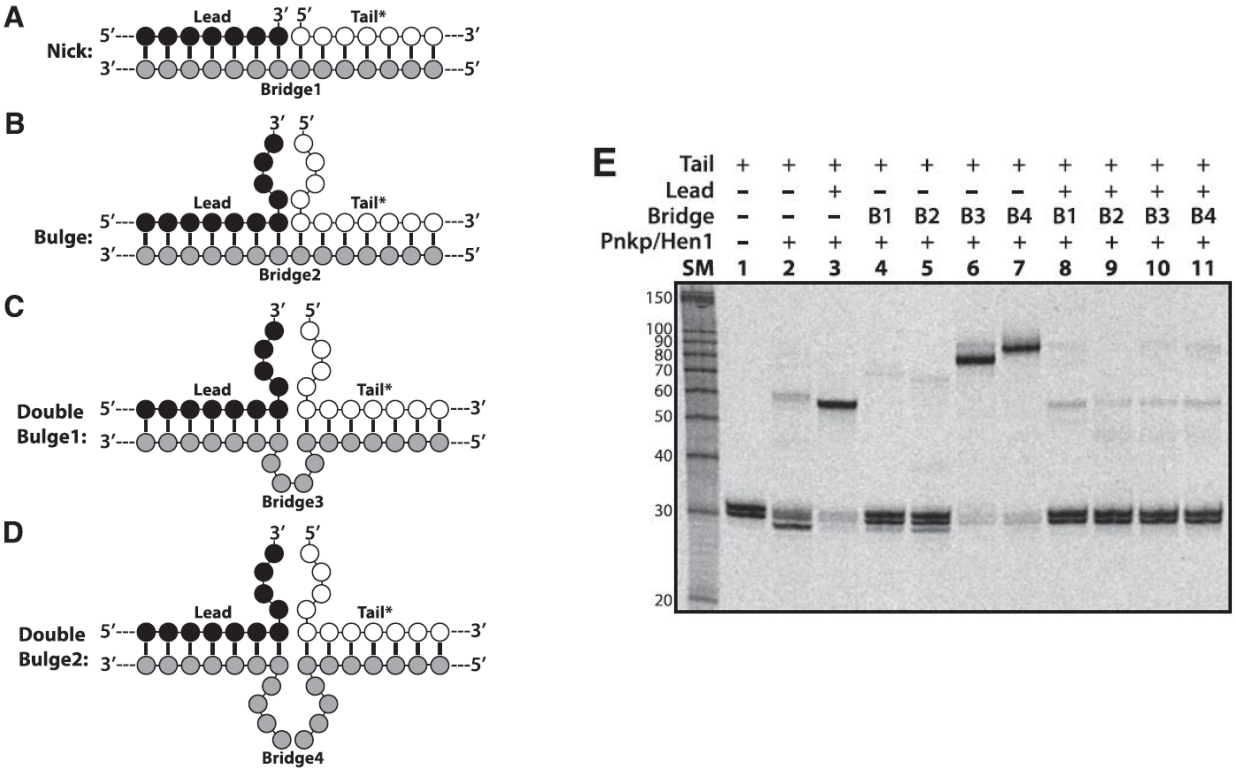

Figure 6: Repair of broken double-stranded RNAs by AvaPnkp/Hen1. **A–D**: Schematic depiction of different secondary structures of RNA repair substrates with single RNA break in the regions of double-stranded RNA (Nick) and bulges (Bulge, Double Bulge1, and Double Bulge2). Only Tail RNA, which is depicted with white cycles, is radiolabeled (with *). Lead RNA, the potential repair partner of Tail, is colored black, and the complementary strand bridging Tail and Lead RNAs (Bridge1 to Bridge4) is in gray. **E**: Analysis of RNA repair products by DPAGE. B1 to B4 indicate Bridge1 to Bridge4.

Figure 7

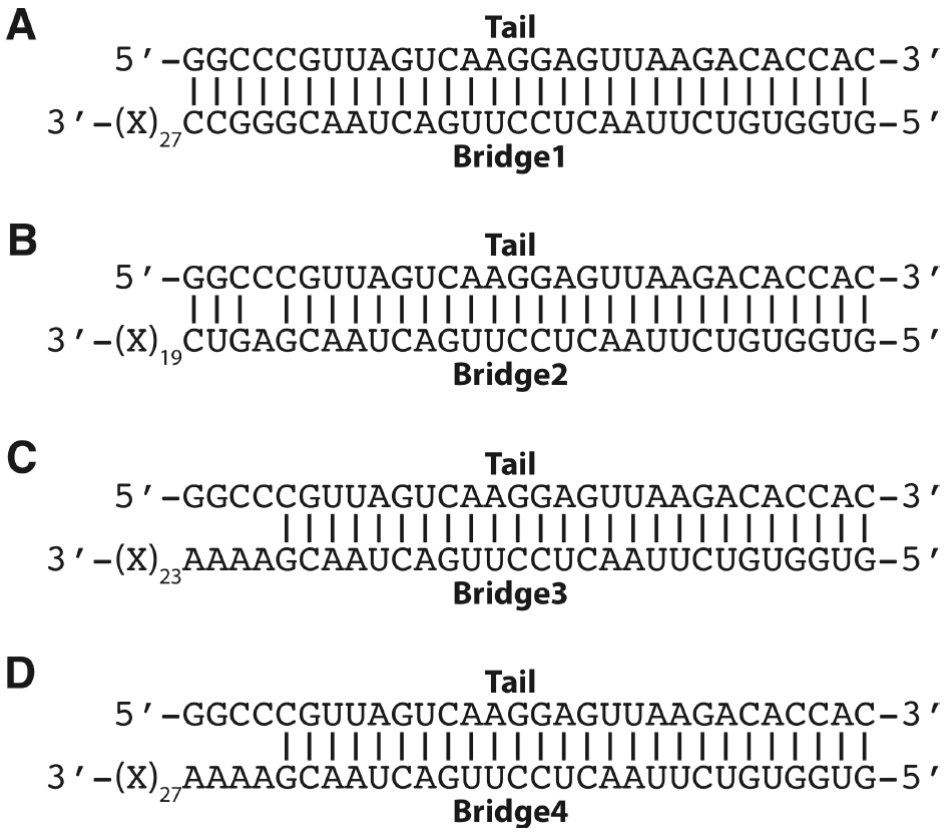

Figure 7: Proposed formation of secondary structures of **A:** Tail/Bridge1, **B:** Tail/Bridge2, **C:** Tail/Bridge3, and **D:** Tail/Bridge4 combinations.

Figure 8

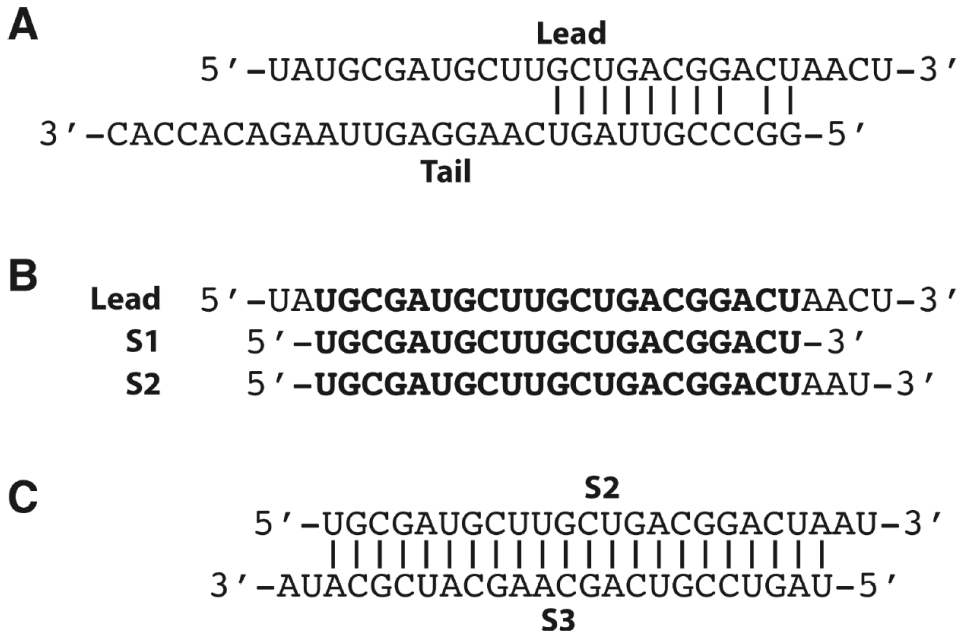

Figure 8: Comparison of Nucleotide sequences of some single-stranded RNAs used in this study their potential to form RNA duplexes. **A**: Proposed formation of base-pairs between Lead and Tail RNAs. **B**: Nucleotide sequence similarities among Lead, S1 and S2 single-stranded RNAs. The sequence that is identical among these three RNAs is highlighted in bold. **C**: Complementary of nucleotide sequence of S3 to the ones of Lead RNA.

Figure 9

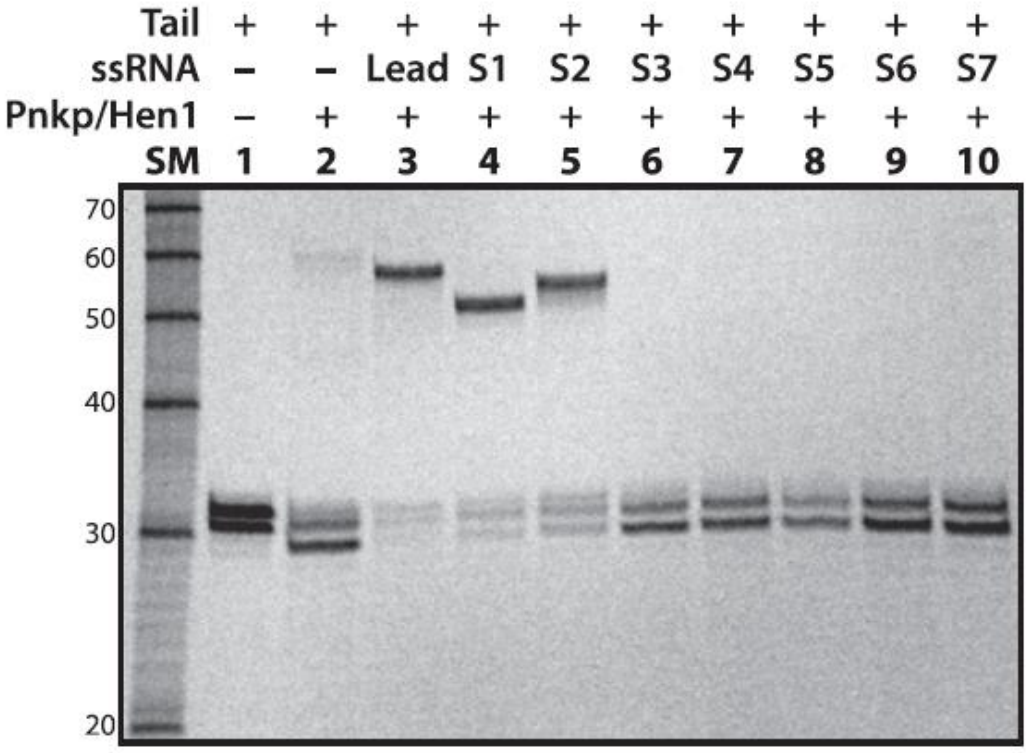

Figure 9: Repair of RNA substrates consisting of two single stranded RNAs (ssRNAs) by *Ava*Pnkp/Hen1. As in the case of Figure 6, only Tail RNA is radiolabeled.

Figure 10

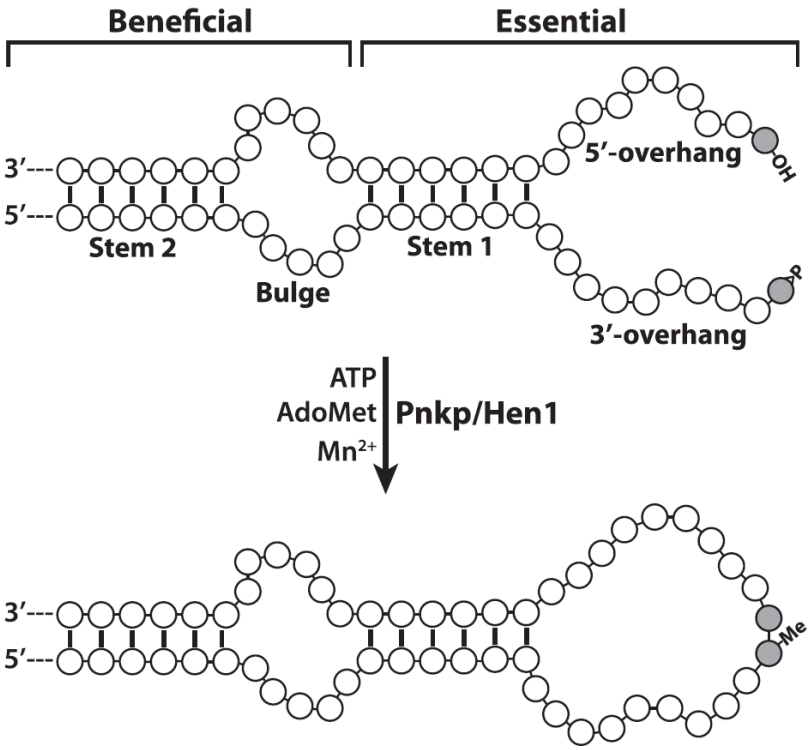

Figure 10: Schematic depiction of an effective RNA substrate of AvaPnkp/Hen1 and its repaired product based on experimental data presented in this study. The two terminal nucleotides to be ligated by *Ava*Pnkp/Hen1 are highlighted in gray, with the depicted 5'-OH and 2', 3'-cyclic phosphate groups that are directly involved in repair chemistry. ATP is required for phosphorylation of the 5'-OH group, and AdoMet is the methyl donor for 2'-O-methylation on the 3'-terminal nucleotide during repair, which requires a manganese ion (Mn2+) to carry out the reaction. The numbers of nucleotides for the different regions of RNA (5' overhang, 3' overhang, Stem 1, Bulge, and Stem 2) were chosen arbitrarily. Secondary structures of the Bulge, the 5' overhang, the 3' overhang, and the resulting repaired loop were also arbitrary depicted.

TABLES

Table 1. tRNA and corresponding DNA primer sequence used in tRNA truncation study

RNAs	Sequence (5'-3')	Size
$tRNA^{Asp}$	GGA GCG GUA GUU CAG UCG GUU AGA AUA CCU GCC UGU CAC GCA GGG GGU CGC GGG UUC GAG UCC CGU CCG UUC CGC CA	77
$tRNA^{Asp}$-ΔD	GGA GCG GUA GUA CCU GCC UGU CAC GCA GGG GGU CGC GGG UUC GAG UCC CGU CCG UUC CGC CA	62
$tRNA^{Asp}$-ΔT	GGA GCG GUA GUU CAG UCG GUU AGA AUA CCU GCC UGU CAC GCA GGG GGU CGU CCG UUC CGC CA	62
$tRNA^{Asp}$-ΔDT	GGA GCG GUA GUA CCU GCC UGU CAC GCA GGG GGU CGU CCG UUC CGC CA	47
$tRNA^{Arg}$	GUC CUC UUA GUU AAA UGG AUA UAA CGA GCC CCU CCU AAG GGC UAA UUG CAG GUU CGA UUC CUG CAG GGG ACA CCA	75
$tRNA^{Arg}$-ΔD	GUC CUC UUA GCG AGC CCC UCC UAA GGG CUA AUU GCA GGU UCG AUU CCU GCA GGG GAC ACC A	61
$tRNA^{Arg}$-ΔT	GUC CUC UUA GUU AAA UGG AUA UAA CGA GCC CCU CCU AAG GGC UAA UUG CAG GGG ACA CCA	60
$tRNA^{Arg}$-ΔDT	GUC CUC UUA GCG AGC CCC UCC UAA GGG CUA AUU GCA GGG GAC ACC A	46
DNA primers	**Sequence (5'-3')**	**Size**
$tRNA^{Asp}$ primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GGA GCG GTA GTT CAG TCG GTT AGA ATA CCT GCC TGT CAC GCA GGG GGT CGC GGG TTC GAG TCC CGT CCG TTC CGC CA	104
$tRNA^{Asp}$ primer 2	TGG CGG TTC GGT CGG GTC TCG TTC CCG CGT CCC CCT GCG TGT CTG GCT GGT TTT CTT TCC GTC TGT TCT TCC GCT CCT ATA GTG AGT CGT ATT ACA GCG TGC GT	104
$tRNA^{Asp}$-ΔD primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GGA GCG GTA GTA CCT GCC TGT CAC GCA GGG GGT CGC GGG TTC GAG TCC CGT CCG TTC CGC CA	89

Table 1 continue

DNA primers	Sequence (5'-3')	Size
$tRNA^{Asp}$-ΔD primer 2	TGG CGG AAC GGA CGG GAC TCG AAC CCG CGA CCC CCT GCG TGA CAG GCA GGT ACT ACC GCT CCT ATA GTG AGT CGT ATT ACA GCG TGC GT	89
$tRNA^{Asp}$-ΔT primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GGA GCG GTA GTT CAG TCG GTT AGA ATA CCT GCC TGT CAC GCA GGG GGT CGT CCG TTC CGC CA	89
$tRNA^{Asp}$-ΔT primer 2	TGG CGG AAC GGA CGA CCC CCT GCG TGA CAG GCA GGT ATT CTA ACC GAC TGA ACT ACC GCT CCT ATA GTG AGT CGT ATT ACA GCG TGC GT	89
$tRNA^{Asp}$-ΔDT primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GGA GCG GTA GTA CCT GCC TGT CAC GCA GGG GGT CGT CCG TTC CGC CA	74
$tRNA^{Asp}$-ΔDT primer 2	TGG CGG AAC GGA CGA CCC CCT GCG TGA CAG GCA GGT ACT ACC GCT CCT ATA GTG AGT CGT ATT ACA GCG TGC GT	74
$tRNA^{Arg}$ primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTC CTC TTA GTT AAA TGG ATA TAA CGA GCC CCT CCT AAG GGC TAA TTG CAG GTT CGA TTC CTG CAG GGG ACA CCA	102
$tRNA^{Arg}$ primer 2	TGG TGT CCC CTG CAG GAA TCG AAC CTG CAA TTA GCC CTT AGG AGG GGC TCG TTA TAT CCA TTT AAC TAA GAG GAC TAT AGT GAG TCG TAT TAC AGC GTG CGT	102
$tRNA^{Arg}$-ΔD primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTC CTC TTA GCG AGC CCC TCC TAA GGG CTA ATT GCA GGT TCG ATT CCT GCA GGG GAC ACC A	88
$tRNA^{Arg}$-ΔD primer 2	TGG TGT CCC CTG CAG GAA TCG AAC CTG CAA TTA GCC CTT AGG AGG GGC TCG CTA AGA GGA CTA TAG TGA GTC GTA TTA CAG CGT GCG T	88
$tRNA^{Arg}$-ΔT primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTC CTC TTA GTT AAA TGG ATA TAA CGA GCC CCT CCT AAG GGC TAA TTG CAG GGG ACA CCA	87
$tRNA^{Arg}$-ΔT primer 2	TGG TGT CCC CTG CAA TTA GCC CTT AGG AGG GGC TCG TTA TAT CCA TTT AAC TAA GAG GAC TAT AGT GAG TCG TAT TAC AGC GTG CGT	87

Table 1 continue

DNA primers	Sequence (5'-3')	Size
$tRNA^{Arg}$-ΔDT primer 2	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTC CTC TTA GCG AGC CCC TCC TAA GGG CTA ATT GCA GGG GAC ACC A	73
$tRNA^{Arg}$-ΔDT primer 2	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTC CTC TTA GCG AGC CCC TCC TAA GGG CTA ATT GCA GGG GAC ACC A	73

Table 2: sequences of all cleaved products from tRNA truncation study

Cleaved RNAs	Sequence (5'-3')	Size
Asp-F35	GGA GCG GUA GUU CAG UCG GUU AGA AUA CCU GCC UG	35
Asp-T35		
Asp-F42	UCA CGC AGG GGG UCG CGG GUU CGA GUC CCG UCC GUU CCG CCA	42
Asp-D42		
Asp-D20	GGA GCG GUA GUA CCU GCC UG	20
Asp-DT20		
Asp-T27	UCA CGC AGG GGG UCG UCC GUU CCG CCA	27
Asp-DT27		
Arg-F38	GUC CUC UUA GUU AAA UGG AUA UAA CGA GCC CCU CCU AA	38
Arg-T38		
Arg-F37	GGG CUA AUU GCA GGU UCG AUU CCU GCA GGG GAC ACC A	37
Arg-D37		
Arg-D24	GUC CUC UUA GCG AGC CCC UCC UAA	24
Arg-DT24		
Arg-T22	GGG CUA AUU GCA GGG GAC ACC A	22
Arg-DT22		

Table 3: sequences of 8 synthetic stem loop RNA

Cleaved RNAs	Sequence (5'-3')	Size
$tRNA^{Asp}$-AC-SL-5'	CGC CUG CCU G	10
$tRNA^{Asp}$-AC-SL-3'	GCA GCC CCU CCU AA	11
$tRNA^{Arg}$-AC-SL-5'	UCA CGC AGG CG	14
$tRNA^{Arg}$-AC-SL-3'	CGC CUG CCU G	7
$tRNA^{Asp}$-ΔDT-ΔB-5'	GGG CUA GGG GAC ACC A	15
$tRNA^{Asp}$-ΔDT-ΔB-3'	GUC CUC UAG CCC CUC CUA A	20
$tRNA^{Arg}$-ΔDT-ΔB-5'	UCA CGC AGG CCG UUC CGC CA	19
$tRNA^{Arg}$-ΔDT-ΔB-3'	GGA GCG GCC UGC CUG	16

Table 4: tRNA and corresponding DNA primer sequence used in tRNA truncation study

RNAs	Sequence (5'-3')	Size
Tail RNA	GGC CCG UUA GUC AAG GAG UUA AGA CAC CAC	30
Lead RNA	UAU GCG AUG CUU GCU GAC GGA CUA ACU	27
Bridge 1 (BR1)	GUG GUG UCU UAA CUG CUU GAC UAA CGG GCC AGU UAG UCC GUC AGC AAG CAU CGC AUA	57
Bridge 2 (BR2)	GUG GUG UCU UAA CUG CUU GAC UAA CGA GUC CGU CAG CAA GCA UCG CAU A	49
Bridge 3 (BR3)	GUG GUG UCU UAA CUG CUU GAC UAA CGA AAA AGU CCG UCA GCA AGC AUC GCA UA	53
Bridge 4 (BR4)	GUG GUG UCU UAA CUG CUU GAC UAA CGA AAA UAC CAG UCC GUC AGC AAG CAU CGC AUA	57
DNA primers	**Sequence (5'-3')**	**Size**
Tail RNA primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GGC CCG TTA GTC AAG GAG TTA AGA CAC CAC	57
Tail RNA primer 2	GTG GTG TCT TAA CTC CTT GAC TAA CGG GCC TAT AGT GAG TCG TAT TAC AGC GTG CGT	57
BR1 primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTG GTG TCT TAA CTG CTT GAC TAA CGG GCC AGT TAG TCC GTC AGC AAG CAT CGC ATA	84
BR1 primer 2	TAT GCG ATG CTT GCT GAC GGA CTA ACT GGC CCG TTA GTC AAG CAG TTA AGA CAC CAC TAT AGT GAG TCG TAT TAC AGC GTG CGT	84
BR2 primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTG GTG TCT TAA CTG CTT GAC TAA CGA GTC CGT CAG CAA GCA TCG CAT A	76
BR2 primer 2	TAT GCG ATG CTT GCT GAC GGA CTC GTT AGT CAA GCA GTT AAG ACA CCA CTA TAG TGA GTC GTA TTA CAG CGT GCG T	76
BR3 primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTG GTG TCT TAA CTG CTT GAC TAA CGA AAA AGT CCG TCA GCA AGC ATC GCA TA	80
BR3 primer 2	TAT GCG ATG CTT GCT GAC GGA CTT TTT CGT TAG TCA AGC AGT TAA GAC ACC ACT ATA GTG AGT CGT ATT ACA GCG TGC GT	80

Table 4 continue

DNA primers	Sequence (5’-3’)	Size
BR4 primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GTG GTG TCT TAA CTG CTT GAC TAA CGA AAA TAC CAG TCC GTC AGC AAG CAT CGC ATA	84
BR4 primer 2	TAT GCG ATG CTT GCT GAC GGA CTG GTA TTT TCG TTA GTC AAG CAG TTA AGA CAC CAC TAT AGT GAG TCG TAT TAC AGC GTG CGT	84

Table 5: sequences of RNA used in random repair study

RNAs	Sequence (5'-3')	Size
Tail RNA	GGC CCG UUA GUC AAG GAG UUA AGA CAC CAC	30
Lead RNA	UAU GCG AUG CUU GCU GAC GGA CUA ACU	27
S1	UGC GAU GCU UGC UGA CGG ACU	21
S2	UGC GAU GCU UGC UGA CGG ACU AAU	24
S3	UAG UCC GUC AGC AAG CAU CGC AUA	24
S4	GGC ACG GCU GUA AAC CGU GC	20
S5	GCC GGC UGU AAA CCG GC	17
S6	GGG UGC UCA GUA CGA GAG GAA CCG CAC CC	29
S7	GGA GCG GUA GUU CAG UCG GUU AGA AUA CCU GCC UG	35

Table 6: Sequences of dumbbell RNA and their DNA primers

RNAs	Sequence (5'-3')	Size
$tRNA^{asp}$-ΔDT-DB	UCA CGC AGG GGG UCG UCC GUU CCU UUU GGA GCG GUA GUA CCU GCC UG	47
$tRNA^{arg}$-ΔDT-DB	GGG CUA AUU GCA GGG GAC UUU UGU CCU CUU AGC GAG CCC CUC CUA A	46
Asp-stem loop-DB	UCA CGC AGG CGU UUU CGC CUG CCU G	25
Arg-stem loop-DB	GGG CUG CUU UUG CAG CCC CUC CUA A	2
DNA primers	**Sequence (5'-3')**	**Size**
$tRNA^{asp}$-ΔDT-DB primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA TCA CGC AGG GGG TCG TCC GTT CCT TTT GGA GCG GTA GTA CCT GCC TG	74
$tRNA^{asp}$-ΔDT-DB primer 2	CAG GCA GGT ACT ACC GCT CCA AAA GGA ACG GAC GAC CCC CTG CGT GAT ATA GTG AGT CGT ATT ACA GCG TGC GT	74
$tRNA^{arg}$-ΔDT-DB primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GGG CTA ATT GCA GGG GAC TTT TGT CCT CTT AGC GAG CCC CTC CTA A	73
$tRNA^{arg}$-ΔDT-DB primer 2	TTA GGA GGG GCT CGC TAA GAG GAC AAA AGT CCC CTG CAA TTA GCC CTA TAG TGA GTC GTA TTA CAG CGT GCG T	73
Asp-stem loop-DB primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA TCA CGC AGG CGT TTT CGC CTG CCT G	52
Asp-stem loop-DB primer2	CAG GCA GGC GAA AAC GCC TGC GTG ATA TAG TGA GTC GTA TTA CAG CGT GCG T	52
Arg-stem loop-DB primer 1	ACG CAC GCT GTA ATA CGA CTC ACT ATA GGG CTG CTT TTG CAG CCC CTC CTA A	52
Arg-stem loop-DB primer 2	TTA GGA GGG GCT GCA AAA GCA GCC CTA TAG TGA GTC GTA TTA CAG CGT GCG T	52

REFERENCE

Amitsur M, Levitz R, Kaufmann G. 1987. Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA. *EMBO J* 6: 2499–2503.

Chan C, Zhou C, Brunzelle JS, Huang RH. 2009a. Structural and biochemical insights into 29-O-methylation at the 3'-terminal nucleotide of RNA by Hen1. *Proc Natl Acad Sci* 106: 17699-17704.

Chan CM, Zhou C, Huang RH. 2009b. Reconstituting bacterial RNA repair and modification in vitro. *Science* 326: 247.

El Omari K, Ren J, Bird LE, Bona MK, Klarmann G, LeGrice SF, Stammers DK. 2006. Molecular architecture and ligand recognition determinants for T4 RNA ligase. *J Biol Chem* 281: 1573–1579.

Galburt EA, Pelletier J, Wilson G, Stoddard BL. 2002. Structure of a tRNA repair enzyme and molecular biology workhorse: T4 polynucleotide kinase. *Structure* 10: 1249–1260.

Ho CK, Shuman S. 2002. Bacteriophage T4 RNA ligase 2 (gp24.1) exemplifies a family of RNA ligases found in all phylogenetic domains. *Proc Natl Acad Sci* 99: 12709–12714.

Jain R, Shuman S. 2010. Bacterial Hen1 is a 3' terminal RNA ribose 29-O-methyltransferase component of a bacterial RNA repair cassette. *RNA* 16: 316–323.

Jain R, Shuman S. 2011. Active site mapping and substrate specificity of bacterial Hen1, a manganese-dependent 3' terminal RNA ribose 29O-methyltransferase. *RNA* 17: 429–438.

Keppetipola N, Nandakumar J, Shuman S. 2007. Reprogramming the tRNA-splicing activity of a bacterial RNA repair enzyme. *Nucleic Acids Res* 35: 3624–3630.

Martins A, Shuman S. 2004a. An RNA ligase from Deinococcus radiodurans. *J Biol Chem* 279: 50654–50661.

Martins A, Shuman S. 2004b. Characterization of a baculovirus enzyme with RNA ligase, polynucleotide 59-kinase, and polynucleotide 3'-phosphatase activities. *J Biol Chem* 279: 18220–18231.

Martins A, Shuman S. 2005. An end-healing enzyme from Clostridium thermocellum with 5' kinase, 2', 3' phosphatase, and adenylyltransferase activities. *RNA* 11: 1271–1280.

Nandakumar J, Schwer B, Schaffrath R, Shuman S. 2008. RNA repair: an antidote to cytotoxic eukaryal RNA damage. *Mol Cell* 31: 278–286.

Ogawa T, Tomita K, Ueda T, Watanabe K, Uozumi T, Masaki H. 1999. A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science* 283: 2097–2100.

Schwer B, Sawaya R, Ho CK, Shuman S. 2004. Portability and fidelity of RNA-repair systems. *Proc Natl Acad Sci* USA 101: 2788-2793.

Simpson L, Sbicego S, Aphasizhev R. 2003. Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* 9: 265–276.

Tomita K, Ogawa T, Uozumi T, Watanabe K, Masaki H. 2000. A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. *Proc Natl Acad Sci* 97: 8278–8283.

Zhang C, Chan CM, Wang P, Huang R. 2012. Probing the substrate specificity of the bacterial Pnkp/Hen1 RNA repair system using synthetic RNAs. *RNA* 18: 335-44.