Expression of RNA Nanoparticles Based on Bacteriophage Phi29 pRNA in Escherichia coli and Bacillus subtilis

Le Zhang

University of Kentucky, le.zhang@uky.edu

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation
Zhang, Le, "Expression of RNA Nanoparticles Based on Bacteriophage Phi29 pRNA in Escherichia coli and Bacillus subtilis" (2013). Theses and Dissertations--Pharmacy. 36.
https://uknowledge.uky.edu/pharmacy_etds/36

This Master's Thesis is brought to you for free and open access by the College of Pharmacy at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Pharmacy by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student’s advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student’s dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Le Zhang, Student

Dr. Peixuan Guo, Major Professor

Dr. Jim Pauly, Director of Graduate Studies
EXPRESSION OF RNA NANOPARTICLES BASED ON BACTERIOPHAGE PHI29 PRNA IN ESCHERICHIA COLI AND BACILLUS SUBTILIS

THESIS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Pharmaceutical Sciences at the University of Kentucky

By

Le Zhang

Lexington, Kentucky

Director: Peixuan Guo, Professor of Pharmaceutical Sciences

Lexington, Kentucky

2013

Copyright© Le Zhang 2013
ABSTRACT OF THESIS

EXPRESSION OF RNA NANOPARTICLES BASED ON BACTERIOPHAGE PHI29 PRNA IN ESCHERICHIA COLI AND BACILLUS SUBTILIS

Currently, most of the RNAs used in lab research are prepared by *in vitro* transcription or chemical synthesis, which can be costly. *In vivo* expression in bacterial cells is another approach to RNA preparation that allows large scale production at a lower cost. However, there are some obstacles in bacterial expression, including RNA degradation in host cell, as well as RNA extraction and purification. tRNA and 5S RNA have been reported as scaffolds to circumvent the degradation problem. These scaffolds can not only make the RNA product survive in the cell but also increase the stability after extraction.

The packaging RNA (pRNA) of bacteriophage phi29 is a small non-coding RNA with a compact structure. The three-way junction (3WJ) region from pRNA is a thermodynamically stable RNA motif good for constructing therapeutic RNA nanoparticles. The 3WJ can not only integrate multiple RNA modules, but also stabilize them.

Here I report a series of approaches made to express recombinant RNAs based on pRNA or 3WJ in bacteria, including 1) Investigating the mechanism of RNA folding *in vitro* and *in vivo* using 3WJ. 3WJ-based RNAs were expressed in *E. coli* using pET system. The results show that the folding of RNA is affected by both overall and regional energy landscape. 2) Expression of an RNA nanoparticle harboring multiple functional modules, a model of therapeutic RNA, in *E. coli* using a combination of tRNA scaffold and pRNA-3WJ. The expression was successful and all of the RNA modules were functional. 3) Expression of pRNA-based recombinant RNAs in *B. subtilis*. This is a novel system of expressing recombinant RNAs in Gram-positive bacteria.

KEYWORDS: pRNA, bacteriophage phi29, Bacillus subtilis, RNA expression, three-way junction
Expression of RNA Nanoparticles Based on Bacteriophage Phi29 pRNA in
*Escherichia coli* and *Bacillus subtilis*

By

Le Zhang

Dr. Peixuan Guo

(Director of Thesis)

Dr. Jim Pauly

Director of Graduate Studies

2013
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Peixuan Guo, for his inspiring guidance and great patience throughout all my research. I would also like to give my thanks to my committee members, Dr. Kimberly Nixon and Dr. Steven Van Lanen, for their continuous encouragement, valuable advice, and selfless help for this thesis.

I would like to acknowledge all members of Dr. Guo’s laboratory. Firstly, I want to highlight Dr. Dan Shu for her great support in all the RNA expression projects. I would like to thank Dr. Emil Khisamutdinov for our collaboration in the RNA folding project. Also, I want to express my appreciation to Dr. Yi Shu and Dr. Jia Geng who helped me to get familiar with the laboratory and life in America. I am also thankful to the other members in the laboratory, past and present, Dr. Hui Zhang, Dr. Farzin Haque, Dr. Mathieu Cinier, Dr. Randall Reif, Dr. Peng Jing, Dr. Gian Marco De Donatis, Dr. Zhanxi Hao, Dr. Huaming Fang, Dr. Chad Schwartz, Dr. Anne Vonderheide, Wei Li, Daniel Binzel, Fengmei Pi, Zhengyi Zhao, Hui Li, Danny Jasinski, Shaoying Wang, Zheng Cui, Eva Beabout, Jeannie Haak, and Nayeem Hossain, for all their collaborations and friendship.

Finally, I want to express my love and gratitude to my parents, my father Ailin Zhang and my mother Ling Wang, for their everlasting support and encouragement in my life.
## Table of Contents

Acknowledgements…………………………………………………………………………………………………iii

Table of Contents………………………………………………………………………………………………iv

Table of Figures…………………………………………………………………………………………………vi

Chapter 1 Introduction ............................................................................... 7

1.1 *In vivo* expression of non-coding RNAs........................................... 7

1.1.1 Background of recombinant RNA expression ........................... 7

1.1.2 Expression System ................................................................... 2

1.1.3 RNA scaffolds ........................................................................... 3

1.1.4 Elements That Have Been Successfully Expressed *in vivo* ...... 5

1.2 The bacteriophage phi29 and its packaging motor .......................... 7

1.2.1 Introduction of bacteriophage phi29 .......................................... 7

1.2.2 The packaging motor of phi29 ................................................... 7

1.2.3 Phi29 packaging RNA ............................................................... 8

1.2.4 The three-way junction motif in phi29 pRNA ........................... 10

Chapter 2 Study of RNA Folding Using the Malachite Green Aptamer .... 13

2.1 Background ................................................................................... 13

2.2 MATERIALS AND METHODS ....................................................... 14

2.2.1 Design of RNA constructs ....................................................... 14

2.2.2 *In vitro* synthesis and purification of RNAs ............................. 15

2.2.3 Malachite Green (MG) aptamer fluorescence assay ............... 16

2.2.4 Vector construction for *in vivo* expression ............................... 16

2.2.5 RNA *in vivo* expression and purification ............................... 16

2.2.6 Gel electrophoresis of *in vivo* prepared RNAs ...................... 17

2.3 Results and Discussion ................................................................. 17

2.3.1 In vitro comparison of 3WJ-pRNA folding after transcription and after denaturation/annealing ............................................. 17

2.3.2 In vivo assessment of the RNA folding after transcription ...... 18

2.4 Conclusion.......................................................................................... 20

Chapter 3 Expression of RNA nanoparticle harboring multiple functional elements in *Escherichia coli* using tRNA scaffold .......... 22
3.1 Background ................................................................................... 22
3.2 MATERIALS AND METHODS....................................................... 24
  3.2.1 Design of RNA construct......................................................... 24
  3.2.2 Construction of expression vector........................................... 24
  3.2.3 RNA expression and extraction............................................... 24
  3.2.4 Purification of tRNA-MG-HBV-STV from agarose gel .......... 25
  3.2.5 Streptavidin binding test.......................................................... 25
  3.2.6 HBV ribozyme cleavage assay ............................................... 26
3.3 Result and Discussion ................................................................... 26
  3.3.1 Verifying the function of MG aptamer, streptavidin aptamer and
       HBV ribozyme.......................................................................... 26
  3.3.2 The yield of RNA production ................................................... 27
3.4 Conclusion ..................................................................................... 27
Chapter 4 RNA Expression in Bacillus subtilis Using Phi29 pRNA as
Scaffold 31
4.1 Background ................................................................................... 31
4.2 MATERIALS AND METHODS....................................................... 33
  4.2.1 Design of RNA constructs....................................................... 33
  4.2.2 Construction of expression vector........................................... 33
  4.2.3 Bacillus subtilis transformation................................................ 33
  4.2.4 Fluorescence assay of living B. subtilis cells........................... 34
  4.2.5 RNA extraction from B. subtilis .............................................. 35
  4.2.6 Fluorescence assay of extracted RNA.................................... 36
4.3 Result and Discussion ................................................................... 36
  4.3.1 Verifying the RNA products..................................................... 36
  4.3.2 The method of RNA extraction and purification............... 37
  4.3.3 Current limitations of pRNA scaffold ....................................... 38
4.4 Conclusion........................................................................................ 39
4.5 Work in progress and failed attempts ............................................ 40
References.......................................................................................... 45
Vita....................................................................................................... 51
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Structure of phi29 pRNA.</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Example of RNA expression using scaffold</td>
<td>12</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Folding assay of <em>in vivo</em> expressed pRNA-3WJ MG-apt</td>
<td>21</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Design of tRNA-MG-HBV-STV</td>
<td>28</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Activity of the aptamers and ribozyme in tRNA-MG-HBV-STV</td>
<td>29</td>
</tr>
<tr>
<td>Figure 6</td>
<td>AFM image of tRNA-MG-HBV-STV</td>
<td>30</td>
</tr>
<tr>
<td>Figure 7</td>
<td>pRNA-based design constructs to be expressed in <em>B. subtilis</em></td>
<td>41</td>
</tr>
<tr>
<td>Figure 8</td>
<td>MG and spinach fluorescence spectrum of pRNA-based structures</td>
<td>42</td>
</tr>
<tr>
<td>Figure 9</td>
<td>RNAs extracted from <em>B. subtilis</em> 12A</td>
<td>43</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Image of living <em>B. subtilis</em> cells under confocal microscope</td>
<td>44</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction

1.1 In vivo expression of non-coding RNAs

1.1.1 Background of recombinant RNA expression

Novel functions and applications of RNAs have been discovered in recent decades, including siRNAs, aptamers, and ribozymes. While the interest in the potential usage of RNAs in biomedical and molecular studies is increasing, the fragile nature of RNAs is an obstacle in RNA studies. The recombinant protein expression system has contributed a lot to the studies of protein structure and function that it makes it possible to prepare proteins in large scale. [1] Similarly, finding a way to prepare RNA for high yield and with cost-efficiency is important in RNA studies. Currently most RNAs used in research are prepared by in vitro transcription or chemical synthesis, which are costly and difficult to operate on a large scale, or limit the length of product. [2]

Compared to these in vitro approaches, in vivo expression is cheaper and much more practical in large scale. Nowadays recombinant plasmids are widely used in protein expression. Although these in vivo protein expression systems will make an mRNA before the protein is translated, most them are not feasible for making RNAs because the RNA products are often degraded quickly. The system needs to be redesigned to be available for RNA expression. The first three successful approaches of large scale
RNA preparation in *E.coli* were reported in 1988. [1, 3-5] Two of them were tRNAs and another was 5S ribosomal RNA. All these RNAs naturally exist in the host strain and have compacted structures, this may contribute to their stability in the cell. After the initial success, kinds of wild type and modified RNAs have been expressed in *E.coli*. In most recent approaches, different RNA structures were inserted into the anticodon loop, protecting them from being degraded by nucleases. [2] This offers a method of preparing desired RNA structures *in vivo*. Similar approaches were also made with 5s RNA. [6]

### 1.1.2 Expression System

#### Host Strain

*Escherichia coli* (*E. coli*) is one of the most widely used bacterium in biological studies. The advantage of *E. coli* is that it grows fast and can be cultured large scale at low cost. Moreover, it produces RNAs with a very high efficiency. [2]

#### Plasmid Vector for RNA Expression

Just like the expression of recombinant proteins, a plasmid that is constructed for RNA expression generally contains elements including replication origin, promoter, multiple cloning site, and selection tags.

One vector designed for RNA expression is constructed by inserting an *lpp* promoter and a downstream multiple cloning site into pBS plasmid. (Agilent Technologies). [59] It is derived from a suppressor tRNA encoding plasmid and it has been found that recombinant RNAs based on tRNA scaffolds can also be expressed. [2] Other research using 5S RNA scaffold has been based on the T7 expression system. [6]

#### Promoters
T7 expression system and \textit{lpp} promoter are widely used in the expression of recombinant RNAs. Each of them has advantages and drawbacks. The \textit{lpp} promoter is one of the strongest known natural bacterial promoters and is available in a lot of strains. [60] The expression is constant and the efficiency is high. However, it cannot be regulated. The lac oparon-T7 system can be induced and also has a very high efficiency. This is a great advantage when the RNA product is toxic to the host, and it prevents the host from growing. But, its termination efficiency is poor; therefore it often makes undesirably long products. Additionally, it can only be used in some strains specifically designed for T7 expression, like BL21 (DE3). [2]

1.1.3 RNA scaffolds
Different kinds of structured RNAs has been tested to transcribe directly from a vector that works with tRNA, and the result has been heterogeneous. The product failed to accumulate to the desirable amount. Therefore, it has been found that the structure, rather than the vector, contributes to the stability of RNA products.

The degradation of RNAs in bacteria is a complex process involving several kinds of RNase. Many of them will attack the terminal single strand, either the 5'-end or the 3'-end. A compacted double strand in the end is helpful to the survival of the RNA product. Unfortunately, \textit{in vivo} transcription often makes a terminal single strand, especially at the 3'-end. Unlike protein that has a stop codon to terminate translation at a specific site, the terminators of transcription are based on a stem-loop structure. The site of termination in transcription is often unspecific and the efficiency is rather low. [7] Thus, post transcriptional processing is important to make a compact
RNA structure that resists degradation. Flanking the desired RNA structure with scaffolds based on stable and abundant natural RNAs is an effective way to protect the product. The scaffold not only allows the product to be processed to give a compact end, but also mimics the product in the structure of a natural RNA, therefore greatly increases the stability of the product. [2]

tRNA is one of the earliest available scaffolds found for recombinant RNA expression. The advantage of tRNA scaffold is the small size of the scaffold and high natural. There are up to 130,000 tRNA molecules per genome in the condition of exponential growth. For comparison, this is almost twice as the most abundant cytoplasmic protein, the elongation factor EF-Tu, which is approximately 75,000. [25] In previous applications the desired RNA is inserted into the anticodon loop. ΨC and D stem-loops need to be kept to maintain the L-shape secondary structure as acceptor. Insertions up to approximately 300 nucleotides have been successfully expressed in tRNA scaffold. [2]

Recent studies have also proven 5S RNA to be an effective scaffold for recombinant RNAs. There are around 12,500 ribosomes per cell and thus, the same number of 5S RNAs. Ribosomal RNAs are resistant to degradation and have a half life of several days. [12] The 5S RNA scaffold has some advantages that complement the tRNA, including different cloning strategy that allows overcoming the RNA misfolding problems in the tRNA scaffold designs. [6] In many case the scaffold will not affect the activity of the inserted elements. However, it is beneficial to remove the scaffold in many situations. The strategy of scaffold removal includes ribozymes, DNAzyme,
and RNase H. RNase H has been proven to work well in releasing the insertion from tRNA scaffold. [2] RNase H cleaves DNA-RNA hybrid double strands, therefore a complementary single strand DNA is needed for the cleavage. DNAzymes are catalytic DNA molecules, usually possessing specific cleavage activity like ribozymes.

1.1.4 Elements That Have Been Successfully Expressed \textit{in vivo}
Currently there is a wide range of functional RNA elements that are available for \textit{in vivo} expression.

**Malachite Green Aptamer (Fluorescence tag)**
Malachite Green (MG) is a triphenyl methane dye that has low fluorescence by itself, but the intensity of the fluorescence increases greatly when it is binding to an RNA aptamer. [48, 51] The MG aptamer is a compact double strand structure that has been expressed using either 5S RNA scaffold or tRNA scaffold. [2, 6] The MG fluorescence can be used in RNA tracking and quantification.

**Spinach Aptamer (Fluorescence tag)**
The fluorophore in Green fluorescent protein (GFP) is formed by intramolecular cyclization of three residues, Ser65-Tyr66-Gly67. The resulting fluorophore, 4-hydroxybenzlidene imidazolinone (HBI), is nonfluorescent by itself. It needs to be enclosed into a correctly folded GFP molecule to start the emission of fluorescence. [20]

An RNA aptamer (spinach) has been discovered by SELEX that is able to bind to 3,5-dimethoxy-4-hydroxybenzylidene imidazolinone (DMHBI) and activate its fluorescence. This fluorogenic RNA aptamer processes a similar pattern of spectrum as EGFP, and the brightness of the fluorescence is
satisfying. This aptamer has been used as a fluorescent tag to observe the
motion of 5S rRNA in living cells by fusing it to the 3’ end of 5S rRNA. [19]

**Sephadex Aptamer (Affinity tag)**
The sephadex aptamer is acquired by *in vitro* selection. The specific binding
provides an effective method of RNA purification using sephadex column.[8]
High concentration of UREA is used in elution to denature the RNA and
disturb the secondary structure of the aptamer, thus disable the binding.

**Streptavidin Aptamer (Affinity tag)**
Streptavidin is a protein acquired from *Streptomyces avidinii* that has very
high affinity to biotin. Commercial streptavidin resin is now available. The
streptavidin-binding RNA aptamer is developed by *in vitro* selection and
offers an alternative way of RNA purification. [9] The capacity of streptavidin
column is not as high as sephadex, but the RNA is eluted by biotin without
denaturation. This is valuable when the folding of RNA product need to be
preserved.

**Hammerhead ribozyme (Therapeutic application)**
A hammerhead ribozyme is found to be able to recognize and cleave the
polyA signal in hepatitis B virus, thus disabling the virus. Applying this
ribozyme to infested cells will inhibit the replication of hepatitis B virus. [10] *In vivo* expression makes large scale preparation of this kind of therapeutic
RNAs more practicable.

**HIV-DIS (Self assembly)**
The HIV-1 type dimerization initiation signal (DIS) loop enables *in vivo* dimer
formation. HIV-DIS is a stem-loop structure with complementary nucleotides
in the loop. Dimer forms by forming basepairs through a loop-loop interaction.
This allows assemblage of two RNA molecules harboring different functional groups [11].

1.2 The bacteriophage phi29 and its packaging motor

1.2.1 Introduction of bacteriophage phi29
Bacteriophage phi29 is a tailed dsDNA phage that infects *Bacillus subtilis*. Phi29 is one of the smallest and simplest known dsDNA phage. Like other viruses, the replication of phi29 contains two steps. First, the viral proteins, RNAs, and genome DNA are synthesized in the host. Then the components assemble to form an infectious unit. Prohead is the first particle assembled during the construction of virion, containing scaffolding protein (gp7), major capsid protein (gp8), head fibers (gp8.5). A dodecameric connector (gp10) is incorporated into the bottom of the prohead, which acts as a path for DNA packaging, as well as connecting the tail to the prohead. [21, 22] A 174-base RNA (pRNA) is attached to the connector. [24, 25] A virus-encoded ATPase (gp16) drives the 19.8 kbp phage genome into the prohead by viral ATPase (gp16). The process of packaging will trigger the release of gp7. [26] gp16 and pRNA detach from the connector after the packaging is completed. After that, the lower collar, appendages, and tail knob are attached to the capsid to form a mature viral unit.

1.2.2 The packaging motor of phi29
During the process of phage assembly, the viral genome DNA is transported into the procapsid. This packaging process is entropically unfavorable and is driven by the hydrolysis of ATP. The packaging motor of phi29 is constructed
by three coaxial rings: the connector, the ATPase hexamer, and the pRNA hexamer. The connector is a dodecameric ring incorporated into the prohead. It is constructed by 12 copies of protein gp10 and serves as the pathway for the DNA. The hexameric DNase gp16 generates the energy needed for DNA translocation. The hexametric packaging RNA (pRNA) connects the ATPase to the connector. [27, 28]

The packaging motor of bacterial phage phi29 has been studied for a long time. The one-way traffic of dsDNA through the channel of the connector has been verified by different methods, including voltage ramping, electrode polarity switching, and sedimentation force assessment. [29] However, the detailed mechanism of the one-way transportation is still unclear. Although a rotation model has been proposed, a later study revealed that none of the three layers of the motor rotate during the packaging process. [27, 28, 30] Recently, our laboratory has discovered a novel revolving mechanism. During packaging, dsDNA revolves along the channel wall of the connector. The affinity between the ATPase subunit and dsDNA is increased while the ATPase subunit is binding to an ATP molecule. The hydrolysis of ATP results in a conformational change and lower affinity for dsDNA, thus transferring the dsDNA to the next adjunct ATPase subunit. [27, 31]

1.2.3 Phi29 packaging RNA

The packaging RNA (pRNA) is essential for packaging the DNA into the prohead. [25] pRNA is an 174-nucleotide (nt) RNA encoded by the viral genome. The first 117 nt in pRNA is the functional domain; the rest is the 3′-domain, the terminator, that is not required for the packaging function. The
117 nt pRNA contains two domains, the gp16 binding domain (domain/motif 1 in Fig.1) and the prohead binding domain (domain/motif 2-5 in Fig.1). The prohead binding domain is located at the central region of the pRNA. Two interlocking loops (right-hand loop, motif 3 and left hand loop, motif 5, Fig.1) containing complementary nucleotide sequence are located in this domain, allowing hexamer forming in the packaging motor. [16, 25, 32, 33] The sequence in the loops can also be engineered to form stable dimers, trimers, and other pRNA oligomers.

The properties of pRNA make it a good vector for constructing RNA therapeutic RNA nanoparticles. First, the diameter of hexametric pRNA is around 30 nanometers (nm), which is suitable for endocytosis. Particles larger than 100 nm are difficult to be engulfed by the cells, while particles smaller than 20 nm can be cleaned out of the serum in a short time. [34] Secondly, the prohead binding domain and the gp16 binding domain fold independently. Therefore, the gp16 binding domain can be replaced by another sequence (e.g. siRNA) and the hexamer forming activity (that is induced by the left- and right-hand loop in prohead binding domain) will not be affected. [10] Moreover, while single strand RNAs can be vulnerable to RNase, small and compact RNAs constructed by double-strand motifs are often more stable. [2]

Our laboratory has proven that pRNA can be used as a scaffold to escort siRNA or hepatitis B virus ribozyme (HBV ribozyme). While the HBV ribozyme is escorted by the pRNA its activity is even higher than the ribozyme itself. [10] It is possible that the stable double strand pRNA can
help the folding of ribozyme, making a larger portion of the ribozyme folded into the correct conformation.

1.2.4 The three-way junction motif in phi29 pRNA

The three-way junction (3WJ) motif is the central region of pRNA. It can be constructed by three synthesized single strand RNA oligos, denoted as $a_{3WJ}$, $b_{3WJ}$, and $c_{3WJ}$. When the three RNA oligos are mixed in distilled water at room temperature in 1:1:1 ratio, they form the 3WJ motif automatically. Formed 3WJs remain stable in distilled water at room temperature for up to several weeks without dissociation. Moreover, the 3WJ does not dissociate in the presence of 8M UREA, which is common in denaturing buffers. The $T_m$ of pRNA-3WJ is 58°C, which is significantly higher than most of similar RNA 3WJ motifs. This evidence shows that the pRNA-3WJ is thermodynamically stable. [35]

One of the problems in using therapeutic RNA nanoparticle, though, is that RNAs need to fold correctly to exhibit functions. Unlike proteins whose structure can be stabilized by disulfide bond, the folding of RNAs is maintained by hydrogen bonds and other weak interactions. Therefore, the folding of RNAs is relatively unstable because of the lack of covalent bonds and crosslinking. After systemic injection, the therapeutic RNA nanoparticle can be diluted into extremely low concentration. Further, although not always necessary, tens of millimoles of magnesium are required for the correct folding of many RNA motifs. However, under physiological conditions, the concentration of magnesium is usually lower than 1 mM. Therefore, misfolding or dissociation can occur to therapeutic RNA nanoparticles because of a low concentration of magnesium or a low concentration of RNA
itself. The pRNA-3WJ, however, is able to assemble automatically in distilled water without any magnesium added. In addition, it has been diluted to a concentration as low as 160 pM and there was no detectable dissociation. The stable nature of pRNA-3WJ makes it a potential scaffold to escort functional RNA modules, preventing them from dissociation or misfolding. [35, 36]

Figure 1 Structure of phi29 pRNA. Structures of the components of phi29 DNA-packaging RNA (pRNA), showing its sequence in domains. The 3WJ domain sequence is highlighted in red.
Figure 2 Example of RNA expression using scaffold (A) *E. coli* tRNA\textsubscript{Lys3}. (B) *E. coli* tRNA\textsubscript{Lys3} with anticodon loop replaced by malachite green (MG) aptamer. (C) Example of a plasmid engineered for RNA expression in bacteria [1, 2].
Chapter 2  Study of RNA Folding Using the Malachite Green Aptamer

2.1 Background

Non-coding RNAs need to fold to appropriate tertiary structures in order to perform their intended functions. Until now, the folding of RNA has been studied for decades, and a lot of principles of RNA folding have been elucidated. However, most of them are based on \textit{in vitro} studies. [37-39] The difference of RNA folding between \textit{in vitro} and \textit{in vivo} conditions is still unclear. The common methods of denaturing/reannealing (denaturing by high salt concentration or high temperature) RNAs \textit{in vitro} are impractical \textit{in vivo}. It is very difficult to change salt concentration \textit{in vivo}. Heating would also kill the cells and possibly degrade the \textit{in vivo} expressed RNAs.

Several methods have been reported to be effective in elucidating RNA structure \textit{in vivo}. Dimethyl sulfide (DMS) is sensitive to RNA structure and has been successful at probing RNA structure in a variety of organisms from bacteria to eukaryotes. [40-41] Lead-(II)-acetate has also been used to probe RNA structures in bacteria for its ability of inducing specific cleavages at the position of tight metal ion binding. [42]

Here I report a different approach to studying RNA folding \textit{in vivo} by the fluorogenic malachite green (MG) aptamer. The MG aptamer needs to fold into a double-strand conformation in order to bind to MG dye and fluoresce. The fluorescence will disappear when the aptamer is degraded or misfolded. Therefore, it can be used as a reporter of RNA folding both \textit{in vitro} and \textit{in vivo}. 
vivo. I hypothesize that the presence of 3WJ will help the MG aptamer keep its correct conformation.

2.2 MATERIALS AND METHODS

2.2.1 Design of RNA constructs

To investigate the mechanism of RNA both in vitro and in vivo, we designed an RNA nanostructure unit based on a thermodynamically stable pRNA-3WJ recently discovered by our laboratory. [35] The pRNA-3WJ acts as a scaffold to direct the folding and an MG-aptamer inserted into the 3WJ to serve as a folding reporter. The RNA constructs were named after the length and position of their interfering sequences: for example, “5’+12” means that there were 12 nt added to the 5’-end complementary to the MG-aptamer sequence, and “3’+15” means that there were 15 overhanging nt added to the 3’-end complementary to the MG-aptamer sequence. Similarly, “3’NM+15” means that 15 nt non-complementary (NM = no match) to the MG-aptamer region were added to the 5’-end of the pRNA 3WJ.

During the transcription, the sequence at the 5’-end was synthesized earlier than the sequence at the 3’-end. Single stranded nt of different lengths (6, 12, and 15) were added to the 5’-end complementary to the MG aptamer sequence in order to interfere with the folding that was originally driven by the pRNA-3WJ (5’+15, 5’+12, and 5’+6). (All RNAs constructed for in vitro assay were designed by Dan Shu.)

Several controls were designed with overhanging nt of different lengths that did not match the MG aptamer sequence (5’NM+12, 5’NM+9, and 5’NM+6). Another control contained 15 nt that matched the MG aptamer
inserted into the 3'-end. Since the 3'-end would not come out until the entire MG aptamer had been synthesized during transcription, the 3'-end overhanging sequence was not expected to disturb the folding.

Unlike protein translation, where expression starts after a ribosome receives a signal from the ATG start codon of mRNA and ceases at the corresponding stop codon(s), effective mechanisms of termination in RNA transcription are lacking. [43-45] Furthermore, the efficiency of the T7-ТΦ RNA transcription terminator has shown to be about 66% by in vitro assays. [46] If the terminator failed to block the RNA polymerase, long RNA products would be transcribed; often, these RNA products have been subjected to rapid degradation. In addition, the terminator sequences undergo transcription and produce a stem-loop at the 3'-end of RNA transcripts. This additional sequence would confront the experimental design. Therefore, we introduced a cis-acting ribozyme sequence placed at the 3'-end to avoid an unwanted sequence at the 3'-end. [10, 47]

### 2.2.2 In vitro synthesis and purification of RNAs

3WJ RNAs harboring the MG aptamer and different insertions, used in MG aptamer fluorescence assay, were prepared by in vitro transcription using T7 RNA polymerase. The DNA templates and primers were synthesized chemically by IDT (Iowa). DNA templates of in vivo transcription were amplified by PCR. RNAs were prepared by in vitro T7 transcription and then purified by 8 M urea 8% PAGE. The corresponding bands were excised under UV shadow and eluted from the gel for 4 h at 37°C in the elution buffer (0.5 M NH4OAc, 0.1 mM EDTA, 0.1% SDS, and 0.5 mM MgCl2), followed by ethanol precipitation overnight at -20°C (2.5x volume of 100% ethanol and
1/10 volume of 3M NaOAc). The precipitate was pelleted by centrifugation (16500 x g, 30 min), washed with 70% ethanol, and dried by speed vacuum. Finally, the RNA dried pellet was rehydrated in 0.05% DEPC treated water and stored at -20°C.

2.2.3 Malachite Green (MG) aptamer fluorescence assay

Gel-purified RNAs were mixed with MG (2 µM) in binding buffer containing 100 mM KCl, 5 mM MgCl2, and 10 mM HEPES (pH 7.4) and incubated at room temperature for 30 min. The refolded RNAs (after heating) were treated by heating them to 95°C for 5 min before staining. The fluorescence was measured using a fluorospectrometer (Horiba Jobin Yvon; SPEX Fluolog-3), excited at 615 nm, and scanned from 625 to 800 nm for emission. [48, 49, 51] (The in vitro fluorescence assay was completed by Dan Shu.)

2.2.4 Vector construction for in vivo expression

A sequence of cis-ribozyme (Rz) was fused onto the 3'-end of DNA templates of corresponding RNA for terminal processing. [47] Reference, 5'+6, 5'+12, and 5'+15 RNAs were inserted between BglII/Ndel sites of expression vector pET-3b. The BglII cleavage would remove the original T7 promoter in the vector to prevent any undesired sequence in the 5'-end of the RNA product. The insertion fragment contained the T7 promoter in its 5'-end. The cloning was completed in E. coli strain DH5α. The recombinant plasmids were verified by sequencing (GENEWIZ).

2.2.5 RNA in vivo expression and purification

E. coli strain BL21 (DE3) (Invitrogen) was transformed by the recombinant plasmids. The colony was inoculated by 5 ml LB medium containing 100
µg/ml ampicillin grown at 37°C and shaken at 250 rpm until A600nm reached 0.5. IPTG solution of 50 µl (1 M) was added to 5 ml cell culture, cells were allowed to continue to grow for 1.5 h, and then were pelleted and resuspended in 250 µl of 10 mM magnesium acetate, 1 mM Tris-HCl, pH 7.4 (buffer L; Ponchon et al. [50]). The total soluble RNAs were extracted using 500 µl of water saturated phenol (pH 4.5) (Fisher). The aqueous phase was ethanol precipitated and then dissolved in 50 µl of 0.05% DEPC treated water. [1, 50]

2.2.6 Gel electrophoresis of in vivo prepared RNAs

In vivo expressed 5’+0 (reference), 5’+6, 5’+12, and 5’+15 RNAs were analyzed by PAGE gel by loading 5 µl of each sample. The denaturing gel was 8% PAGE gel containing 8 M of urea run in 1x TBE (89 mM Tris base, 89 mM Boric acid, 2 mM EDTA) at 120 V at room temperature for 1 h. After electrophoresis, the gels were stained by 10 µM MG in binding buffer (100 mM KCl, 5 mM MgCl2, and 10 mM HEPES, pH 7.4) for 15 min at room temperature. The MG fluorescence image was acquired by the Cy5 channel (635 nm excitation/670 nm observation) using a Typhoon scanner. [48, 51]

The gels were then stained by Ethidium Bromide (EB) and scanned in an EB channel (532 nm excitation/580 nm observation).

2.3 Results and Discussion

2.3.1 In vitro comparison of 3WJ-pRNA folding after transcription and after denaturation/reannealing

To reveal the folding differences of the pRNA-3WJ prior to and after denaturation/annealing, the spectrum of MG fluorescence was recorded

17
firstly when the transcription was completed. After that, the RNAs were heated to 95° C, and then cooled to room temperature. The fluorescence spectrum was measured again.

RNA design 5'+0, 5'+6, 3'+15 and those with non-match insertions in 5'-end exhibited similar fluorescence intensity before and after heating. However, for design 5'+12 and 5'+15, the intensity of fluorescence increased significantly after heating. Therefore, the folding of RNA was impacted by the length of 5'-interference sequence. In the situation of the designs with long interference sequence in 5'-end, such as 5'+12 and 5'+15, the 5'-sequence would pair with the first several nucleotides in the MG aptamer, preventing the rest of the aptamer to fold. After heating and reannealing, the low free energy promoted the 3WJ to fold first, thus recovering the folding of MG aptamer. [35] The 3'-interference sequence had no effect on the folding, since the 3WJ-MG structure had already folded when the 3'-sequence was transcribed. The 5'-non-match sequences had no effect either, for they could not pair with the 3WJ-MG.

2.3.2 In vivo assessment of the RNA folding after transcription
To investigate the RNA folding in vivo, designed 5'+0, 5'+6, 5'+12 and 5'+15 RNAs were expressed in E.coli BL21(DE3) strain. The sequences of in vivo designs were the same as those in vitro, but with an additional cis-ribozyme inserted to the 3'-end to process the RNA product. Unlike protein expression that starts at the ATG codon and stops at the stop codon, effective mechanisms of termination are lacking in RNA transcription. The efficiency of T7-ТΦ terminator has been found to be ~66% at in vitro assays. [46] Long RNA products would be transcribed if the terminator failed to block the RNA
polymerase. These RNAs are often degraded rapidly. Additionally, the terminator itself would be transcribed as a stem-loop in the 3’-end of RNA product, which is undesired in this experiment. Therefore, we used a cis-ribozyme to remove the unwanted sequences in 3’-end after transcription, thus reducing the additional sequence to only 7 nt. [10, 47]

The RNA product of in vivo expression varies while the length of interference sequence changes. Designed 5'+0 and 5'+6 RNA products with MG binding activity were detected in MG-stained PAGE gels. However, when the interference sequences were longer in design, 5'+12 and 5'+15, there were hardly any products showing MG binding activity. Instead, there was a very small RNA fragment that could not be stained by MG. These results indicated that the folding of the designs with short insertions and those with longer insertions may have been different. In 5'+6 the 5’ interference sequence was not long enough to disturb the folding of 3WJ-MG, thus the 3WJ and MG aptamer folded normally and a complete RNA product was transcribed.

In 5'+15 and 5'+12, the interference sequence would bind to the first strand of MG aptamer before the last part of the RNA, including the other strand of MG aptamer and C strand of the 3WJ, was transcribed. When the first strand of MG aptamer is transcribed, the RNA may fold as a long, GC-rich hairpin that resembles the structure of many terminators. This hairpin may stall the RNA polymerase and cause premature termination. On the other hand, if the transcription continues, the newly synthesized RNA will be left unpaired because their complementary sequences were already paired with the interference sequence. Single strand RNAs are much more
vulnerable to RNases than those with stable double strand. As a result, the 3’ region of the RNA can be degraded even though transcription continues. In both situations the product will be incomplete with only a part of the MG aptamer. Thus, the products of 5’+15 and 5’+12 were much smaller than expected and had no MG binding activity.

The results of in vivo expression corresponded with the in vitro transcription before heating, but not after heating. In 5’+0 and 5’+12 the RNA folded normally and the product was complete with MG binding activity. In 5’+12 and 5’+15, while the interference sequence was longer, the folding of 3WJ and MG aptamer was disturbed. For in vitro transcription, this results in a misfolded MG aptamer whose activity can be recovered by denaturing and annealing. In the case of in vivo expression, the misfolding can cause degradation of unpaired sequences, resulting in an incomplete product.

2.4 Conclusion

The folding of RNA is affected by both overall and regional energy landscape. Sometimes motifs may form in 5’ region of RNA product during transcription, preventing the remaining part of RNA to fold correctly. RNA constructs 5’+12 and 5’+15 have a long interfering sequence in 5’-end that was more likely to bind to the MG aptamer sequence, and thus prevent the folding of MG aptamer. After denaturation/annealing, these RNAs folded into the minimum energy structure. While expressed in living cells, the loose structure of misfolded RNAs can cause degradation. For in vitro transcription, the product can be denatured and refolded by changing temperature or contents.
in the solution. For *in vivo* expression, however, the temperature is often stable, and denaturation/reannealing is not practical. Misfolded RNA products can be degraded rapidly. Predicted conformation of RNA during transcription needs to be taken into consideration for *in vivo* RNA expression.

Acknowledgement: This RNA folding project was initially carried out by Dan Shu, who finished the *in vitro* study and concluded that long interference sequence in 5'-end would disturb the folding of MG aptamer. I would like to thank Dr. Shu for allowing me to join this project and I am glad to contribute to this project with my skills in bacterial expression. Unlike the *in vitro* study, we were not able to get quantitative data for *in vivo* assessment because denaturation/reannealing is impractical in living cells. However, the *in vivo* results show some synergy with the *in vitro* data. The *in vitro* conclusion that long interference sequence in 5'-end will disturb RNA folding was supported by the fact that these RNA designs were degraded *in vivo*.

![Figure 3](image)

**Figure 3** Folding assay of *in vivo* expressed pRNA-3WJ MG-apt.
Chapter 3 Expression of RNA nanoparticle harboring multiple functional elements in *E. coli* using tRNA scaffold

3.1 Background

So far, methods of massive production and purification in bacteria are only available for proteins. Most of the RNAs used in research are prepared by *in vitro* transcription or chemical synthesis, which can be costly in large scale or when producing long RNAs. Although *in vivo* expression of RNA has been considered as an alternate method, there are some obstacles. For example, RNA without a certain structure is vulnerable to RNase and can be degraded rapidly in living cell. The low efficiency of termination can often result in heterogeneous products. [52] Also, large scale purification of certain RNA
products can be difficult. Therefore, scaffolds based on native RNAs have been developed to protect RNA products from degradation, as well as allowing them to be processed by cellular enzymes. tRNA and 5S rRNA have been proven effective scaffolds for recombinant RNA expression. [2, 6] The stability and high copy number of tRNA make it a promising scaffold for large scale expression. In previous applications the exogenous RNA sequence is inserted into the anticodon loop. Insertions up to approximately 300 nucleotides have been successfully expressed using the tRNA scaffold in *E. coli*. [2]

Our lab has developed a tetravalent X-shaped RNA motif based on the pRNA-3WJ, which is able to carry up to four different RNA modules. For therapeutic RNA nanoparticles, multiple functional modules are often needed to be integrated into one molecule. [36] For example, one affinity tag is required to purify the RNA after production (e.g. streptavidin aptamer or sephadex aptamer), similar to his-tag or strep-tag in proteins. One fluorogenic tag is required to detect the transportation, folding and degradation of RNA after delivery (e.g. MG aptamer or spinach aptamer). Finally, one or more therapeutically active elements (e.g. siRNA, ribozymes, or riboswitches) are the most important part of therapeutic RNA nanoparticle. The 3WJ or X-shaped motif can not only integrate multiple modules into one RNA particle, but can also stabilize the folding of these modules.

Here we report a method of combining both tRNA scaffold and 3WJ to make an RNA nanoparticle harboring multiple functional modules that can be expressed in *E. coli* in large scale. Although the pRNA can not be expressed in *E.coli*, the attempt of protecting the pRNA by tRNA scaffold is successful.
Therefore, we hypothesize the 3WJ motif is also capable for *E.coli* expression with tRNA scaffold.

### 3.2 MATERIALS AND METHODS

#### 3.2.1 Design of RNA construct

An X-shaped motif was constructed by combining one 3WJ and one reversed 3WJ. One of the four arms of the X-shaped motif was connected to the anti-codon loop of *E. coli* tRNALys3. The MG aptamer, streptavidin aptamer, and HBV ribozyme were connected to the other three arms. (Fig. 4B) This RNA will be described hereafter as tRNA-MG-HBV-STV.

#### 3.2.2 Construction of expression vector

The tRNA expression vector was constructed using the method developed by L. Ponchon and F. Dardel. [2] The *lpp* promoter, *rrnC* terminator, and several restriction sites were inserted between SacI and XhoI sites of pBluescript II SK (+/-) (pBS). (Fig. 4A) (The construction was done by Keyclone Technologies). This plasmid will be described hereafter as pBS-lpp.

Template DNA sequence of tRNA-MG-HBV-STV was acquired by overlap PCR and cloned between EcoRI/HindIII sites in plasmid pBS-lpp. The cloning was completed in *E. coli* strain DH5α. The recombinant plasmids were verified by sequencing (GENEWIZ).

#### 3.2.3 RNA expression and extraction

*E. coli* strain DH5α containing pBS-lpp with tRNA-MG-HBV-STV sequence inserted was inoculated into 1 liter of LB medium containing 100 µg/ml
ampicillin and was grown at 37°C shaker at 250 rpm for 16 hr. Cells were pelleted and resuspended in 10 ml of 10 mM magnesium acetate, 1 mM Tris-HCl, pH 7.4 (buffer L; Ponchon et al. [50]). The total soluble RNAs were extracted using 12 ml of water saturated phenol (pH 4.5) (Fisher). The aqueous phase was ethanol precipitated then dissolved in 1.5 ml of 0.05% DEPC treated water.

3.2.4 Purification of tRNA-MG-HBV-STV from agarose gel
The extracted RNA was loaded into 1% agarose gel and separated by electrophoresis at 121 V for 30 min in 1x TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0). The band of tRNA-MG-HBV-STV was cut down and sealed in a dialysis bag (3000 MWCO) with 1 ml of TAE. The bag was put back to the electrophoresis chamber and the electrophoresis was continued at 150 V for 15 min. In the last 30 s of the electrophoresis, the electrodes were inverted to release the RNAs attached to the dialysis bag. After electrophoresis, the solution inside the bag was collected. The purified RNAs were concentrated by ethanol precipitation.

3.2.5 Streptavidin binding test
The tRNA-MG-HBV-STV sample was loaded to streptavidin resin in 1x PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 1.8 mM KH2PO4) [13] containing 10 mM of additional MgCl2. The resin was washed by the same buffer. Then, the RNA was eluted by 5 mM of biotin. Samples were collected during washing and elution and verified in 8% 8 M UREA PAGE in 1x TBE by MG and EB staining, as described in 2.2.6. [9]
3.2.6 HBV ribozyme cleavage assay
The DNA substrate of HBV ribozyme was tagged by Cy3. The cleavage reaction was performed at 37°C in 20 mM Tris (pH 7.5) and 20 mM MgCl₂ for 60 min. An in vitro transcribed, pRNA escorted HBV ribozyme was used as positive control of cleavage. The result of cleavage was confirmed in 8% 8 M UREA PAGE in 1x TBE. [10]

3.3 Result and Discussion
3.3.1 Verifying the function of MG aptamer, streptavidin aptamer, and HBV ribozyme
In vivo product of pRNA-MG-HVB-STV was able to bind streptavidin resin. Bound RNA can be eluted competitively by 5 mM biotin. (Fig. 5B) The cleavage of HBV ribozyme had been confirmed (Fig. 5A), however, the efficiency of cleavage was lower than the positive control (pRNA escorted HBV ribozyme, lane 2). This could be caused by the structure of RNA rather than in vivo expression, since both in vivo expressed and in vitro transcribed pRNA-MG-HVB-STV presented similar cleavage efficiency. It was possible that different modules in the RNA molecule could interfere with each other, since the result of HBV ribozyme was positive. The MG fluorescence was detected even in the denaturing gel (Fig. 5A). Connecting to the 3WJ would stabilize the MG aptamer, protecting its conformation under the presence of 8 M UREA.

The design of pRNA-MG-HVB-STV is an example of hypothesized therapeutic RNA nanoparticle. After production, it can be purified using streptavidin column. The HBV ribozyme is a therapeutic module which can
destroy viral genome. After delivery, the presence of the RNA can be observed by MG fluorescence. Our laboratory has created a lot of RNA nanoparticles based on pRNA 3WJ, harboring multiple functional modules. [36, 53] The result of this experiment reveals the possibility of massive production of these RNA nanoparticles in bacteria.

3.3.2 The yield of RNA production

Roughly 5 mg of purified pRNA-MG-HVB-STV was acquired from 1 L of *E. coli* culture. This is less than the reported yield (10-50 mg) of other recombinant RNAs based on tRNA scaffold. [2] One possible reason of the reduction in yield is the HBV ribozyme, which has a single strand region in its hammerhead. Single strand regions are vulnerable to RNases even when they are in the middle of the RNA. Another possible reason is that our RNA was purified from agarose gel, in which more samples could be lost when comparing with the chromatography used by L. Ponchon and F. Dardel. [1,2]

3.4 Conclusion

The pRNA 3WJ is a potential scaffold for therapeutic RNA nanoparticles, it is able to harbor multiple elements and stabilize them. Until now, these pRNA-based nanoparticles were made mainly by *in vitro* transcription. In this project we incorporated the 3WJ and the tRNA scaffold to express 3WJ-based RNA nanoparticles in *E. coli*. The *in vivo* product of pRNA-MG-HBV-STV has been acquired successfully, with all three modules proving functional, thus demonstrating the viability of massive production of 3WJ-based RNA nanoparticle *in vivo*. The current method is not perfect,
yet. Further study is required to improve the yield of expression, as well as the compatibility of different modules.

**Figure 4** Design of tRNA-MG-HBV-STV
(A) The expression region of tRNA expression vector (refered as pBS-lpp). The lpp promoter, rmC terminator and several restriction sites are inserted between Sacl and XhoI sites of pBluescript II SK (+/-) (pBS). (Keyclone Technologies). (B) Construct of tRNA-MG-HBV-STV. The MG aptamer, streptavidin aptamer, and HBV ribozyme were incorporated into the anti-codon loop of *E. coli* tRNALys3 using an X-shaped motif.
Figure 5 Activity of the aptamers and ribozyme in tRNA-MG-HBV-STV

(A) HBV ribozyme cleavage of DNA substrate labeled by Cy3. Cy3 channel shows the DNA substrate, MG channel shows tRNA-MG-HBV-STV, EB channel shows all DNAs and RNAs. Lane 1: substrate only. Lane 2: substrate and HBV ribozyme escorted by pRNA as positive control of cleavage, cleaved substrate appears as a lower band in Cy3 channel. Lane 3: substrate and cell-expressed tRNA-MG-HBV-STV, cleavage detected in Cy3 channel, the MG fluorescence of tRNA-MG-HBV-STV detected at Cy5 channel. Lane 4-5: substrate and in vitro transcribed tRNA-MG-HBV-STV, cleavage detected in Cy3 channel, the MG fluorescence of tRNA-MG-HBV-STV detected at Cy5 channel. Lane 6: tRNA-MG-HBV-STV, purified from cell. Lane 7: tRNA-MG-HBV-STV, made by in vitro transcription.

(B) MG fluorescence and streptavidin binding. The tRNA-MG-HBV-STV sample was extracted from E. coli and purified from agarose gel. 1: tRNA-MG-HBV-STV before loading to streptavidin column, 2: unbound RNA, 3: wash 4-5: elution.

Acknowledgement: Thanks to Dan Shu for the gel images.
Figure 6 AFM image of tRNA-MG-HBV-STV
tRNA-MG-HBV-STV expressed in *E. coli* and purified from gel. An overall “X” shape is apparent. (scale: nm)
Chapter 4  RNA Expression in *Bacillus subtilis* Using Phi29 pRNA as Scaffold

4.1 Background

An efficient method of recombinant RNA expression has been developed in *E. coli* using tRNA and 5S rRNA as a scaffold. [2, 6] However, until now, not a lot of attempts have been made to express recombinant RNAs in gram-positive bacteria.

Currently most of the RNA extraction kits and reagents are primarily designed for animal cells. When applied to bacteria, cell lysis can become a problem. Gram-negative bacteria like *E. coli* are often lysed under harsh conditions (e.g. SDS/NaOH, phenol). [50, 54] These reagents make the kits unavailable. Additionally, phenol or detergents need to be removed after RNA extraction, or they may denature the RNA product. Most RNA extraction kits suggest crushing the bacteria cells mechanically using a homogenizer or a French Press. However, RNA degradation should be taken into consideration since cellular nuclease will not be deactivated in this way. On the other hand, the cell wall of gram-positive bacteria can be removed by lysozyme, leaving the protoplasts. [55, 56] Protoplasts can be simply lysed using the same method as animal cells, and the reagents are often offered in the kits.

Our lab has developed many RNA nanoparticles based on phi29 pRNA or 3WJ. [10, 36, 53] Most of them are prepared by *in vitro* transcription. In one of our previous studies, pRNA was designed as a scaffold to escort the HBV ribozyme. Animal cells were transfected by DNA sequence of pRNA
with HBV ribozyme insertion. The cleavage activity of HBV ribozyme was confirmed in the cells. [10] However, no efforts had been made to observe or extract the RNA product, and the amount of RNA product was not determined. To study the function of pRNA, wild-type pRNA has been expressed in \textit{B. subtilis}, as well as pRNAs with several modified bases. [24] But, pRNA’s capacity of being a scaffold to escort functional elements in large scale expression has not been verified.

The strategy of expressing recombinant RNAs in \textit{E. coli} uses a natural RNA (tRNA or 5S rRNA) as scaffold to deceive the cell, rendering the exogenous RNAs as harmless natural products. The structure of the scaffold plays an important role of the stability of RNA product. [2, 6] Unfortunately, phi29 pRNA is recognized as alien by \textit{E. coli}. Although we have been able to express the pRNA-3WJ in \textit{E. coli} using T7 expression system, applying the same method to full-length pRNA results in degradation of RNA product. On the other hand, \textit{B. subtilis} is the natural host of phi29, and pRNA is rather stable in \textit{B. subtilis}. The full gene of wild-type pRNA, including the promoter and terminator, has been cloned into a plasmid for \textit{B. subtilis} expression. The amount of pRNA product has been satisfying. [24] Here we report the result of the RNA expression in \textit{B. subtilis}, using phi29 pRNA as scaffold to escort exogenous RNA modules. We hypothesize that some regions of the pRNA can be replaced by functional RNA elements while the RNA is still stable in \textit{B. subtilis}. 
4.2 MATERIALS AND METHODS

4.2.1 Design of RNA constructs
We used fluorogenic RNA aptamers because they are easily detected. In our first attempt, MG aptamer was inserted into the left-hand loop of pRNA (pRNA-MG). (Fig.7A) In another design, the head-loop of pRNA was replaced by spinach aptamer (pRNA-spi). (Fig.7B) To test pRNA’s capacity to harboring more than one functional module, both MG aptamer and spinach aptamer were integrated into the left-hand loop of pRNA using a reversed 3WJ (pRNA-MG-spi). (Fig.7C) Template DNA sequences of pRNA-MG, pRNA-spi, and pRNA-MG-spi were acquired by overlap PCR.

4.2.2 Construction of expression vector
Shuttle vector pHT315 has been created by O. Arantes and D. Lereclus that can replicate in either *E. coli* or *Bacillus*. pHT315 was constructed by inserting *Bacillus* replication origin, erythromycin resistant gene, and multiple cloning site into *E. coli* vector pUC19. [58]

In our experiment, template DNA sequence of pRNA-MG/ pRNA-spi/ pRNA-MG-spi was inserted between EcoRI/XbaI sites in pHT315. The cloning was completed in *E. coli* strain DH5α. The recombinant plasmids were verified by sequencing (GENEWIZ).

4.2.3 *Bacillus subtilis* transformation
*B. subtilis* strain 12A was transformed following the method by S. Chang and S. N. Cohen, 1979. [23] SMM buffer (0.5 M sucrose, 0.02 M Maleate and 0.02 M MgCl₂, pH 6.5. Wyrick and Rogers, 1973 [57]) was used to maintain the osmotic pressure and prevent the protoplast from lysis. SMMP medium
was prepared by mixing equal volumes of 4 X Penassay broth (Difco Antibiotic Metium 3) and 2 X SMM. *B. subtilis* 12A was grown in 37°C shaker until mid-log phase. Cells were harvested and resuspended in 1/10 volume of SMMP containing 2 mg/ml lysozyme. The resuspended cells were incubated at 37°C with gentle shaking for 2 hr to form protoplasts. After that, the protoplasts were pelleted and washed by SMMP to remove lysozyme, then resuspended in the same volume of SMMP.

For each transformation, 500 μl of protoplast was mixed with 1.5 ml of 30% PEG 8000 in 1X SMM (w/v), as well as 500 ng of plasmid. Two minutes later, 5 ml of SMMP was added to the mixture. The protoplasts were pelleted again and resuspended in 1 ml of SMMP, then incubated at 30°C for 1.5 hr with gentle shaking at 100 rpm. 200 μl of the protoplasts were plated on DM3 regeneration medium (for one liter, 200 ml 4% agar, 500 ml 1 M sodium succinate (pH 7.3), 100 ml 5% Casamino acids, 50 ml 10% yeast extract, 100 ml 3.5% K₂HPO₄ and 1.5% KH₂PO₄, 25 ml 20% glucose, 20 ml 1 M MgCl₂, and 5 ml 2% bovine serum albumin (sterilized by filtering, added to the mixture when the temperature is about 55°C after autoclave)). The recovered colonies were transferred to LB plate containing 25 μg/ml erythromycin for selection. [23]

**4.2.4 Fluorescence assay of living B.subtilis cells**

Overnight culture of *B. subtilis* expressing pRNA-MG/ pRNA-spi/ pRNA-MG-spi was pelleted and resuspended in 2 volumes of binding buffer containing 100 mM KCl, 5 mM MgCl₂, and 10 mM HEPES (pH 7.4). To detect MG fluorescence, 2 μM MG dye was added to the solution. To detect spinach fluorescence, 2 μM DFHBI was added to the solution. [19] The cell
suspension was incubated at room temperature for 30 min. The fluorescence was measured using a fluorospectrometer (Horiba Jobin Yvon; SPEX Fluolog-3). The MG fluorescence was excited at 570 nm, and scanned from 600 to 800 nm to detect emission. [48, 49] The spinach fluorescence was excited at 450 nm, and scanned from 570 to 700 nm to detect emission. [19] To acquire the microscope image of living *B. subtilis* cells, overnight culture of expressing pRNA-MG-spi was pelleted and resuspended in 1/2 volume of binding buffer containing 5 μM MG dye (for MG fluorescence) or DFHBI (for spinach fluorescence). The cell suspension was incubated at room temperature for 30 min. 3 μl of cells was added to a slide and examined under Olympus FLUOVIEW FV1000 confocal laser scanning microscope using a 60X oil lens. MG fluorescence was acquired by Cy5 channel, and spinach fluorescence was acquired by Cy2 channel.

### 4.2.5 RNA extraction from *B. subtilis*

*B. subtilis* expressing pRNA-MG/ pRNA-spi/ pRNA-MG-spi was inoculated into 10 ml of LB broth containing 25 μg/ml erythromycin and grown in 37°C shaker at 250 rpm for 14 h. Cells were pelleted at 4,000 rpm for 10 min at room temperature, and then resuspended in 1 ml of 1X SMM containing 5 mg of lysozyme and 2 u RNase-free DNase I. [23, 57] Protoplasts were formed by incubating at 37°C for 30 min. After that, RNA was extracted from the protoplasts using RNeasy Mini Kit (QIAGEN), following the instructions from the official manual.
4.2.6 Fluorescence assay of extracted RNA

The extracted RNAs were analyzed in 8% PAGE gel in 1x TBE (89 mM Tris base, 89 mM Boric acid, 2 mM EDTA). After electrophoresis, the gels were stained by 5 µM DFHBI in binding buffer (100 mM KCl, 5 mM MgCl2, and 10 mM HEPES, pH 7.4) and scanned using the Cy2 channel (473 nm excitation/520 nm observation) of Typhoon scanner to acquire the spinach fluorescence signal. After that, the gel was stained by 5 µM MG dye in binding buffer and scanned using the Cy5 channel (635 nm excitation/670 nm observation) to acquire the MG fluorescence signal. Finally, the gel was stained by Ethidium Bromide (E. B.) and then scanned in the E. B. channel (532 nm excitation/580 nm observation) to display the total RNA.

The MG and spinach fluorescence spectrums of extracted RNAs were also acquired using the same method described in 4.2.4.

4.3 Result and Discussion

4.3.1 Verifying the RNA products

The fluorescence activity of both MG aptamer and spinach aptamer has been confirmed by spectrum. For cells expressing pRNA-MG, the emission peak was discovered at 650 nm. [48, 51] For cells expressing pRNA-spi, the emission peak was discovered at 500 nm. [20] For cells expressing pRNA-MG-spi, both peaks were detected, but not as high as the constructs with only one aptamer inserted. Fluorescence assay of extracted RNAs had a similar result. The peaks were more obvious since there was not a cell background.
For cells expressing pRNA-MG-spi, both MG and spinach fluorescence have been observed under confocal microscope. (Fig. 10) The intensity of fluorescence differs from cell to cell. This is possibly caused by different levels of expression. It is important to use freshly transformed cells for expression, since constant expression of RNAs is a burden on growth. Cells that produce more RNA product may grow slower, thus finally eliminated by selection, leaving only those have a low level of expression. [1]

The fluorescence was also visible in PAGE gel of extracted RNAs. (Fig. 9) The size difference was caused by different length of insertions.

4.3.2 The method of RNA extraction and purification

Total soluble RNAs can also be extracted from the protoplasts using phenol extraction, as described in 3.2.3. [1] However, the purity of RNA product is not as good as using the RNeasy kit. The MG fluorescence can be detected immediately after phenol extraction and ethanol precipitation. The spinach fluorescence, however, is missing. The RNA needs to be cleaned by passing through a NucAway Spin column (Invitrogen) to recover its spinach fluorescence. The structure of spinach aptamer is more flexible than the double-strand MG aptamer. Therefore the spinach aptamer is possibly denatured by the phenol left after ethanol precipitation. While using the RNeasy kits, the RNA product is ready to use immediately after purification. For *E. coli*, it is difficult to remove the cell wall completely. Additional steps of cell lysis (e.g. homogenizer) are required before the kit can be used. This is why phenol extraction is preferred in *E. coli*, since the cell lysis and RNA extraction can be completed in one step. For *B. subtilis*, the cell wall can be
easily removed by lysozyme digestion. Thus, the kit is preferred for a better quality product.

4.3.3 Current limitations of pRNA scaffold

Although we were able to acquire the expected product of pRNA-MG-spi, its amount was significantly less than pRNA-MG or pRNA-spi. It seems the long insertion had an impact on the stability of RNA. To protect the RNA from being degraded it is important to maintain a compact structure. While using tRNA as scaffold, only the anti-codon loop can be replaced by the insertion. [2] There are three loops in phi29 pRNA that can be replaced by insertions without changing the 3WJ structure. We found each of the three loops can be replaced by the MG aptamer, and that the amount of product is satisfying. However, when all three loops were replaced by insertions (one by MG aptamer, one by streptavidin aptamer, one by HBV ribozyme), there was no detectable product. In addition we were not able to alter the sequence at either the 5' or 3' end of pRNA.

Currently the amount of RNA product with large insertions is not as good as desired. One possible solution to this problem is the modification of the host strain. In our experiment expressing 3WJ in E. coli (Chapter 2) we tried both regular BL21 (DE3) and BL21 Star (DE3) (Invitrogen) whose RNase E is disabled by mutation. The results withBL21 Star (DE3) were better, with more RNA product. Therefore, selecting a suitable host strain can be important for RNA in vivo expression. Bacterial expression of proteins is common; and there are a lot of commercially viable strains that are designed for protein expression. On the other hand, very few strains are engineered specifically for RNA expression, if any. Therefore, we may need to create a
modified (e.g. deletion of RNases) strain for the best result of RNA expression.

4.4 Conclusion

Full-length phi29 pRNA is unstable in *E. coli*. However, it can be produced in a satisfying amount in *B. subtilis*, which is the natural host of phi29. Therefore, pRNA is a potential scaffold for large scale RNA expression in *B. subtilis*. The attempt to insert one aptamer into pRNA is successful. Currently we are able to replace one of the three loops of pRNA with an aptamer. However, while both MG and spinach aptamers are integrated into one loop, the amount of RNA product is significantly reduced. It is possible that both the pRNA scaffold and the host strain need to be modified to express large RNA insertions. Although further improvements are needed, the pRNA scaffold in *B. subtilis* is a potential system for large scale expression of recombinant RNA. The Gram-positive nature of *B. subtilis* will bring some advantages over *E. coli* in cell lysis and RNA extraction. For example, the cell wall can be easily removed by lysozyme, allowing the commercially viable RNA purification kits to be used easily.

Acknowledgement: Thanks to Dan Shu for designing the pRNA-MG, our first recombinant RNA that is available for *B. subtilis expression*. My design of pRNA-MG-spi was also inspired by previous RNA designs made by Shu for *in vitro* study.
4.5 Work in progress and failed attempts

Besides the designs described in this chapter, we also tested a series of different RNA constructions for bacterial expression. Some of them work was completed as desired, but the rest of them failed to make expected products. The result shows that, for an RNA molecule, its viability of in vivo expression heavily depends on the structure.

Besides the pRNA-MG design that integrated the MG aptamer into the left-hand loop, we also tried replacing the right-hand loop, or the head loop, by MG aptamer and the result was similar. However, we encountered problems replacing more than one of the loops in pRNA. One of the designs had the head loop replaced by spinach aptamer and the left-hand loop replaced by MG aptamer. Both MG and spinach fluorescence are detected in living cell. However, the extracted RNA showed several unexpected bands in UREA PAGE gel, suggesting degradation of product. This is why we integrated both MG and spinach aptamers into one loop in design pRNA-MG-spi. The attempt to replace all three loops (one by MG aptamer, one by streptavidin aptamer, one by HBV ribozyme) resulted in no detectable product. It was important to maintain the structure of the scaffold for in vivo expression. [2]

Another attempt was to replace the gp16 binding domain of pRNA by siRNA sequence, in order to produce siRNAs in B. subtilis. This modification would not change the overall structure of the pRNA much. However, this design failed to make any product in B. subtilis. Interestingly, although the full-length pRNA could not be produced in E. coli, the expression worked well
when the sequence of pRNA was inserted into the anti-codon loop of tRNA scaffold. [2] Therefore, the terminal sequence may be critical for in vivo expression. Sequence in the 5'- and 3'-ends can participate in the initiation and termination of transcription, as well as post-transcriptional processing. We are planning to replace either the 5'- or 3'-end of pRNA by random sequence to investigate which parts of the RNA are required for expression. Unlike the sophisticated protein expression, the study of RNA expression in bacteria is still in its initial stage. Modifications may be needed to be made to the host strain, the plasmid vector, as well as the RNA scaffold itself, to finally develop a high-efficiency expression system.

Figure 7 pRNA-based design constructs to be expressed in B. subtilis. All designs are based on wild-type pRNA (with terminator) as scaffold. (A) pRNA-MG, the left-hand loop replaced by MG aptamer. (B) pRNA-spi, the head loop replaced by spinach aptamer. (C) pRNA-MG-spi, both MG aptamer and spinach aptamer integrated into the left-hand loop by a reversed 3WJ.
A  Fluorescence spectrum of living *B. subtilis* cells

- **Cell Control**
- **pRNA-MG**
- **pRNA-spi**
- **pRNA-MG-spi**

**Spinach emission**

**MG emission**

B  Fluorescence emission of extracted RNA

**pRNA-MG**

**pRNA-spi**

**pRNA-MG-spi**

**Spinach emission**

**MG emission**

**Figure 8** MG and spinach fluorescence spectrum of pRNA-based structures (A) Fluorescence spectrum of living cell. The MG fluorescence was excited at 570 nm, and scanned from 600 to 800 nm to detect emission. The spinach fluorescence was excited at 450 nm, and scanned from 570 to 700 nm to detect emission. *B. subtilis* 12A without any exogenous plasmid was taken as negative control. (B) Fluorescence spectrum of extracted RNAs, measured under the same condition.
Figure 9 RNAs extracted from *B. subtilis* 12A 8% TBE PAGE with EB, MG, and DFHBI staining.
Figure 10 image of living *B. subtilis* cells under confocal microscope

The fluorescence from MG aptamer was acquired in Cy5 channel. The fluorescence from MG aptamer was acquired in Cy2 channel. *B. subtilis* 12A without any exogenous plasmids was taken as negative control.
References


13. http://cshprotocols.cshlp.org/content/2006/1/pdb.rec8247.full


Vita
Le Zhang

Date and Place of Birth:
March 17, 1987
Daqing, Heilongjiang Province, China

Education:
2005-2009 B.S. in Biological Sciences, Peking University

Professional Publications:

Presentations:
Dan Shu, Le Zhang, Emil Khisamutdinov, and Peixuan Guo. Differences Between *in vivo* and *in vitro* Folding of RNA. Markey Cancer Center Research Day, Apr 2013