Selective binding to mRNA duplex regions by chemically modified PNAs stimulates ribosomal frameshifting

Puah, Ru Ying

2018

Puah, R. Y. (2018). Selective binding to mRNA duplex regions by chemically modified PNAs stimulates ribosomal frameshifting. Master's thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/73165>

<https://doi.org/10.32657/10356/73165>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**Selective Binding to mRNA Duplex Regions by Chemically Modified
PNAs Stimulates Ribosomal Frameshifting**

PUAH RU YING

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2017

**Selective Binding to mRNA Duplex Regions by Chemically Modified
PNAs Stimulates Ribosomal Frameshifting**

PUAH RU YING

School of Physical and Mathematical Sciences

A thesis submitted to the Nanyang Technological University
in the partial fulfilment of the requirement for the degree of
Master of Science

2017

Acknowledgements

I would like to take this opportunity to express my sincere gratitude to my supervisor, Assistant Professor Chen Gang, for his kind guidance, support and advice throughout my course of research. I would also like to thank him for the opportunity to work on this project, enabling me to explore and learn valuable skills and knowledge.

I would also like to express my deepest appreciation and gratitude to Dr. Shu Zhiyu, Ms. Desiree-Faye Toh Kaixin and Ms. Jia Huan for their kind advice and guidance during the course of my research. I am especially grateful to Dr. Shu Zhiyu and Ms. Jia Huan for their great patience and for imparting their invaluable knowledge, skills and experiences to me during my initial learning phase.

My appreciation is extended to my other group members, Dr. Kiran Patil, Dr. Tan Jiazi, Mr. Manchugondanahalli S. Krishna, Mr. Alan Ong Ann Lerk, Ms. Yang Lixia and undergraduate students, Ms. Zoey Chua Jiayu and Mr. Wong Jun Yan for their encouragement, help and support. I am thankful for the enduring friendship and memories that I have attained during my time in Nanyang Technological University.

I would like to thank Nanyang Assistant Professor Shao Fangwei and Associate Professor Li Tianhu for kindly allowing access to their laboratory equipment. I would also like to thank the administrative and store staff in the Division of Chemistry and Biological Chemistry for their assistance whenever required.

Finally, I would like to dedicate this thesis to my family and loved ones. I am indebted for their unwavering love and care, and their continuous encouragement and support, helping me to pull through difficult periods during my candidature.

Table of Contents

Acknowledgements.....	i
Abstract	iii
List of Figures	iv
List of Tables	v
List of Abbreviations	vi
1. Introduction	1
1.1 Nucleic acids and their properties	1
1.2 Translation and minus-one ribosomal frameshifting	2
1.3 Peptide Nucleic Acids (PNAs) and their applications	4
2. Rational and Objectives	7
3. Materials and Methods	8
3.1 Construction of frameshift reporter plasmid	8
3.2 Synthesis of PNAs.....	10
3.3 In vitro transcription and RNA purification	10
3.4 Binding of PNAs to model mRNAs and in vitro translation.....	10
3.5 Frameshift assay and data analysis.....	11
4. Results and Discussion	13
4.1 Analysis of DNA sequencing results and RNA integrity.....	13
4.2 Frameshifting assay of model mRNA by PNA	15
4.3 Analysis of potential binding sites of PNA P1	20
4.4 Different conditions of in vitro translation assay	22
5. Conclusion	24
6. References	25
7. Appendices	32

Abstract

Minus-one programmed ribosomal frameshifting (-1 PRF) is a prominent mechanism that exists in ribonucleic acid (RNA) viruses. It allows precise maintenance of ratio between viral structural and enzymatic proteins, and is involved in the regulation of the half-life of cellular messenger ribonucleic acid (mRNAs). Variations made to the narrow window of protein ratio would disrupt the propagation of viral replication. The efficiency of -1 PRF is affected by the stability of the mRNA secondary structure, together with the presence of a slippery site which is located upstream of the secondary element embedded in the mRNA sequence. Peptide nucleic acids (PNAs) were thus, utilized as a ligand to regulate the stability of model mRNA secondary structures, altering the ratio of protein production. PNA strands containing modified nucleobases were designed to target the stem-loop region, forming major-groove parallel PNA·RNA-RNA triplex. They show selective binding to double-stranded RNAs (dsRNAs) over single-stranded RNAs (ssRNAs). In summary, this study suggests that the dsRNA-specific chemically modified PNA stimulate -1 RF in a model mRNA hairpin (from 2% to 32%), demonstrated by using cell-free *in vitro* translation assays. An unmodified control PNA, however, shows nonspecific inhibition of translation.

List of Figures

Figure 1	Diagram of ribosomal translation process	4
Figure 2	Chemical structures of base triples and base pairs	6
Figure 3	Scheme of p2luc dual-luciferase system and PNAs for <i>in vitro</i> translation assays	9
Figure 4	DNA sequencing results of DNA plasmids by BLAST	13
Figure 5	Agarose gel electrophoresis of model mRNAs with and without incubation with PNA	14
Figure 6	Frameshifting efficiency of model mRNAs by PNA P3 and P5	16
Figure 7	Effects of PNAs on the expression of FLuc and RLuc	17
Figure 8	Example potential off-target binding sites of unmodified PNA P1	21
Figure 9	Frameshifting efficiency of model mRNAs without pre-incubation with PNA P3 and P5	22
Figure 10	Frameshifting efficiency of model mRNAs upon post-translational addition of PNAs	23
Figure 11	Effects of PNAs on the expression of FLuc and RLuc by post-translational addition of PNAs to model mRNAs	23

List of Tables

Table 1	Raw luminescence values of FLuc and RLuc in ssRNA with the addition of PNA	18
Table 2	Raw luminescence values of FLuc and RLuc in rHP1 with the addition of PNA	19
Table 3	Raw luminescence values of FLuc and RLuc in rHP2 with the addition of PNA	19

List of Abbreviations

A	Adenine
bp	Base pairs
C	Cytosine
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dsRNA	Double-stranded ribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FLuc	Firefly luciferase
G	Guanine
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
kb	Kilobase
K _D	Dissociation constant
L	Thio-pseudoisocytosine
mRNA	Messenger ribonucleic acid
NaCl	Sodium chloride
nt	Nucleotides
PCR	Polymerase chain reaction
PNA	Peptide nucleic acid
PRF	Programmed ribosomal frameshifting
Q	<i>N</i> ⁴ -(2-guanidoethyl)-5-methylcytosine
rHP	RNA hairpin
rRNA	Ribosomal ribonucleic acid
RRL	Rabbit reticulocyte lysate

RLuc	Renilla luciferase
RNA	Ribonucleic acid
rNTP	Ribonucleoside triphosphate
ssRNA	Single-stranded ribonucleic acid
T	Thymine
TAE	Tris-acetate-EDTA
U	Uracil
w/v	Weight per volume
-1 PRF	Minus-one programmed ribosomal frameshifting
-1 RF	Minus-one ribosomal frameshifting

1. Introduction

1.1 Nucleic acids and their properties

Nucleic acids are involved in many important processes in biological systems. Nucleic acid strands are made up of nucleotides which are connected by phosphodiester bonds, where a single nucleotide consists of a ribose sugar, a nitrogenous base and phosphate groups. The two main types of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is the main repository of genetic information in cells, whereas RNA, on the other hand, is involved in the expression of those information stored in DNA through the processes of transcription and translation, and many regulatory and catalytic functions. However, an exception to this rule is that some viruses such as retroviruses, store their genetic information as RNA.

Different conformations of DNA or RNA structures exist and act as functional genomic elements. Unlike DNA, RNA adopts single-stranded conformations instead of the common double-stranded, B-form helical structure. Intramolecular interactions often take place in RNA, allowing them to fold and form purine-pyrimidine Watson-Crick base pairs within its strand¹. RNA also exists in double-stranded A-form helices and higher-order tertiary structures¹. The higher-order RNA conformations can be formed by the existence of mismatches, internal loops and bulges stabilized by interactions such as hydrogen bonds and stacking interactions²⁻³. These variations lead to functional diversity which are important in biological systems^{1,4}. Moreover, the presence of defects in a stem-loop structure could widen the major groove of the stem⁵, increasing the chance of its interactions with other molecules such as drugs and proteins⁶⁻⁹. Over the years, extensive research has demonstrated that RNA-RNA interactions are essential for performing catalytic and regulatory functions¹⁰.

1.2 Translation and minus-one ribosomal frameshifting

Translation is a core metabolic process for the synthesis of protein from messenger RNA (mRNA), the template of genetic code. Transfer RNA (tRNA) delivers amino acids to the translation machinery whereas ribosomal RNA (rRNA), a major component of ribosome is responsible for the catalysis of peptide bonds between amino acids. All organisms are dependent on translation to carry out routine specialized metabolic functions in order to execute cellular activities and to sustain life. In normal circumstances, the translating ribosome decodes the mRNA sequentially by encoding every three nucleotides (a codon) to a specific amino acid in an open reading frame. Progressing from the start codon (AUG), the ribosome will continue to build on the polypeptide chain until it encounters a stop codon (UAG, UAA, UGA) after which the full-length polypeptide is then detached from the ribosome for post-translational processes. The reading frame of an mRNA is precisely maintained to ensure correct protein synthesis¹¹. Translation is a very delicate and exquisite process which adhere very strictly to the specific sequences of the mRNA and any disruption to this high-fidelity process would result in incorrect incorporation of amino acids or even premature termination¹². The production of an erroneous protein may cause the protein to malfunction, and activate cellular clean-up and quality-control machineries^{10, 12}.

Minus-one programmed ribosomal frameshifting (−1 PRF) is a translation recoding event that is employed by many organisms, ranging from prokaryotes to higher eukaryotes^{10, 13-14}. Minus-one ribosomal frameshifting (−1 RF) is activated by a heptameric slippery sequence (X XXY YYZ) and a mRNA secondary structure that is positioned 2-8 nucleotides (nt) downstream from the slippery site. In viruses and eukaryotes, X, Y and Z are usually any bases, A or U, and non-G bases, respectively. When the ribosomal reading frame is shifted, upon −1 RF, from a normal zero frame to

a -1 frame with the two codons at the slippery site decoded as XXX and YYY, respectively, instead of XXY and YYZ (Figure 1). Previous research suggests that -1 RF at a slippery site may be stimulated by a downstream hairpin (Figure 1), pseudoknot, G-quadruplex, or a duplex formed between an antisense oligonucleotide and mRNA¹⁵⁻²¹. It is reported that higher thermodynamic stability of RNA secondary structures corresponds to the increase of frameshifting efficiency. For example, the thermodynamic stability of the first three to four base pairs in a RNA stem-loop structure is shown to be in strong correlation with the frameshifting efficiency¹⁵. Similarly, other research groups employed the usage of antisense oligonucleotides to induce -1 RF by stabilizing RNA secondary structures²⁰⁻²¹. It is also revealed that the mechanical stability of RNA secondary structure is dependent on triplex structures formed by stem-loop interactions¹⁶⁻¹⁷. Furthermore, it is shown that the higher stability of G-quadruplexes also causes -1 RF to increase¹⁹, indicating the RNA secondary structure could be a potential site for the design of therapeutic ligands in targeting viral RNA.

-1 PRF is known to take place predominantly during translocation²². A downstream mRNA secondary structure causes blockage at the entrance of the mRNA tunnel, hampering the rotation of the ribosomal subunits with respect to each other and therefore, promotes the backward movement of the ribosome by one nucleotide at the slippery site^{10, 14, 22-28}. The role of the slippery sequence is crucial as it enhances the progression of the ribosomes, allowing translocation to be completed at a faster rate. Minus-one PRF is often highly regulated with a narrow window of recoding efficiency, which allows the precise maintenance of the ratio between viral structural and enzymatic proteins^{10, 14, 29}. In addition to being essential for the replication of many viruses, -1 RF is also involved in the regulation of the half-life of cellular mRNAs^{14, 30}.

Figure 1. Diagram of ribosome during translational process with the presence of the two stimulatory elements (slippery site and mRNA secondary structure) of -1 PRF embedded in the mRNA sequence.

The *cis*-acting -1 RF-stimulatory mRNA hairpin or pseudoknot structures are often highly conserved, and thus, are potential target sites for the regulation of -1 RF³¹⁻³⁶. The structural stability of the mRNA secondary structures may be critical for stimulating -1 RF^{15, 37-40}. Thus, RNA structure-binding ligands may be developed to modulate the structural stability and -1 RF efficiency. In addition, development of site-specific RNA-binding ligands may facilitate a deeper understanding of the complex -1 RF processes.

1.3 Peptide Nucleic Acids (PNAs) and their applications

In this study, Peptide nucleic acid (PNA) strands were designed to target and stabilize the mRNA secondary structure via Hoogsteen hydrogen bonding, inflicting its unfolding rate. PNA is a type of artificial nucleic acid, which consists of an uncharged 2-aminoethyl glycine unit instead of the phosphodiester backbone, and a methylene carbonyl linker instead of a cyclic sugar moiety (Figure 2)⁴¹⁻⁴⁸. PNA's non-ionic flexible backbone allows it to form Watson-Crick duplexes with a complementary DNA or RNA in both parallel and antiparallel manners, with significantly enhanced stability compared to DNA/RNA duplexes. Furthermore, PNA has a lower sensitivity towards ionic

strength changes, and is resistant to nucleases and proteases⁴⁷. Thus, there is a great potential in developing PNAs as chemical probes and therapeutic agents.

Many chemical modifications of PNAs have been developed to enhance the binding selectivity and specificity towards targeted DNAs and RNAs⁴⁹⁻⁵². Our research group has successfully developed chemically-modified PNA L (thio-pseudoisocytosine) and Q (*N*⁴-(2-guanidoethyl)-5-methylcytosine) monomers to recognize G-C and C-G Watson–Crick base pairs, respectively, through major-groove parallel PNA·RNA-RNA triplex formation⁵²⁻⁵⁴. Incorporation of L base into PNAs facilitates the recognition of a G-C pair at physiological pH through Hoogsteen L-G pair formation, but disfavors the formation of Watson–Crick-like L-G pair due to a steric clash between the sulfur atom and amine group (Figure 2c, e). Q base has a methyl group attached to C5 of cytosine ring, resulting in the guanidine group to face towards the Watson–Crick edge, which facilitates the recognition of the Hoogsteen edge of G in a C-G pair, but generates steric clash with G in a Watson–Crick-like Q-G pair (Figure 2d, f). Thus, L- and Q-incorporated PNAs show selective binding to double-stranded RNAs (dsRNAs) over single-stranded RNAs (ssRNAs) (Figure 2c-f)⁵⁴. The sequence specificity and structural selectivity of L- and Q- modified PNAs may allow the precise targeting of desired viral and cellular RNA structures, and thus may serve as valuable biological tools for mechanistic studies and potential therapeutics for fighting diseases.

Figure 2. Chemical structures of stable PNA-RNA-RNA base triples and unstable PNA-RNA base pairs. The letter "R" indicates the sugar-phosphate backbone of RNA. (a-d) Chemical structures of triples C⁺•G-C, T•A-U, L•G-C, and Q•C-G. (e, f) Unstable Watson-Crick-like base pairs of L-G and Q-G⁵³⁻⁵⁴.

2. Rational and Objectives

RNA viruses utilize -1 PRF to control the synthesis of viral RNA replicase polyproteins required for genome replication and alterations made to the expression of polyproteins would disrupt the viral replication cycle. The *cis*-acting -1 RF-stimulatory mRNA secondary structure is a potential target region for the regulation of -1 PRF as it is highly conserved. This region often exists in double-stranded conformations, allowing it to be a suitable binding site for triplex-forming PNAs.

There have been few reports regarding the selectivity of chemically-modified PNAs towards RNA duplex over single-stranded conformation, and the specificity of modified PNA towards distinctive Watson-Crick base pairs to minimize nonspecific interactions in biological systems. Additionally, Watson-Crick inversions often exist in functional RNA, hence, it is useful to incorporate complementary chemically-modified monomer within a PNA strand, further improving its specificity.

In this work, cell-free *in vitro* translation assays using rabbit reticulocyte lysate was employed to study the modulation of -1 RF of model mRNAs upon the addition of chemically-modified PNAs. Cell-free *in vitro* translation assays were previously employed by many research groups, with the use of plasmids that were constructed based on p2luc vector, consisting of two luciferase reporter genes located between the test sequence^{15, 32, 55-56}. This method was utilized for mRNA structural studies whereby its intramolecular interactions affect the -1 RF^{18, 57}. Furthermore, the effects of -1 RF in RNA viruses were also studied based on cell-free *in vitro* translation assays before proceeding to cell culture studies in -1 RF⁵⁸⁻⁵⁹.

3. Materials and Methods

3.1 Construction of frameshift reporter plasmid

The -1 PRF reporter model were constructed based on the p2luc vector^{32, 56} by inserting the sequences of interest at its poly-linker region using BamHI and SacI (NEB) restriction enzyme sites, which are located between the Renilla luciferase (RLuc) and firefly luciferase (FLuc) reporter genes. DNA oligonucleotides (IDT), ssRNA-E (5'-GAT CCT TTT TTA GGG TAG AGA GAG AAA GTG AGC T-3'), ssRNA-C (5'-GAT CCC TTC TTA GGG ATA GAG AGA GAA AGT GAG CT-3'), rHP1-E (5'-GAT CCT TTT TTA GGG TAG AGA GAG AAA GTT TCG ACT TTC TCT CTC TAG AGC T-3') and rHP1-C (5'-GAT CCA CTT CTT AGG GTA GAG AGA GAA AGT TTC GAC TTT CTC TCT CTA GAG CT-3') and their complementary strands were phosphorylated, annealed, and ligated into p2luc vector to produce both experimental (denoted with the letter "E") and positive control (denoted with the letter "C") constructs. Site-directed mutagenesis by Polymerase Chain Reaction (PCR) was performed on rHP1-E/C plasmids to obtain rHP2-E/C. rHP2-E was constructed by 2-step point mutations with two sets of primers G1663C F (5'-GTC GAC GGA TCC TTT TTT AGG GTA GAG AGA CAA AGT TTC-3') and G1663C R (5'-GAA ACT TTG TCT CTC TAC CCT AAA AAA GGA TCC GTC GAC-3'); C1678G F (5'-TTT CGA CTT TGT CTC TCT AGA G-3') and C1678G R (5'-CTT TGT CTC TCT ACC CTA AAA AAG-3'). As for the construction of rHP2-C, single step mutagenesis was performed instead, modifying two nucleotides concurrently with one set of primers rHP2C-F (5'-CGA CTT TGT CTC TCT AGA GCT CGA AG-3') and rHP2C-R (5'-AAA CTT TGT CTC TCT ACC CTA AGA AGT G-3'). The resultant products were treated with DpnI (NEB) before transforming into *Escherichia coli* competent cells (DH5 α). Plasmid DNA was purified from bacterial cell cultures (Bio Basic Canada Inc)

and the sequences of all the constructs were verified by Applied Biosystems highest capacity-based genetic analyser platforms (1ST BASE).

(a) Experimental constructs (rHP1-E, rHP2-E, and ssRNA-E)

(b) Control constructs

Figure 3. Scheme of p2luc dual-luciferase system and PNAs for *in vitro* translation assays. (a) Experimental mRNA constructs with the insertion between the RLuc and FLuc genes of three model sequences: RNA hairpin 1 (rHP1), ssRNA and rHP2. The slippery sequence is underlined. The 0 frame stop codon (UAG) is shown in cyan. Minus-one RF results in the production of both RLuc (0 frame) and FLuc (-1 frame) proteins. The sequences of the PNAs aligned with the targeted regions of the mRNAs are shown. (b) mRNA constructs serving as positive controls for 100% -1 RF efficiency. The slippery sequences were mutated (indicated by cyan color) to obtain the non-slippy sequences and to have FLuc translated in 0 frame.

3.2 Synthesis of PNAs

The sequences of PNA containing control monomer (C) or modified monomer (L and Q) were designed to target the duplex RNA binding site of 'AGAGAAAG'. The PNA T and C monomers were purchased from ASM (Germany). The modified PNA L and Q monomers were synthesized and purified according to the reported method⁵³⁻⁵⁴. The PNA oligomers were synthesized as previously reported⁵³⁻⁵⁴.

3.3 *In vitro* transcription and RNA purification

The purified DNA plasmids were linearized with PmlI restriction enzyme (NEB) prior to *in vitro* transcription. Large scale synthesis of mRNA was carried out using Riboprobe® *in vitro* Transcription Systems (Promega). 100 µL reaction containing 1× transcription optimized buffer, 10 mM DTT, 100 units Recombinant RNasin® ribonuclease inhibitor, 2.5 mM rNTPs, 5 µg linearized template DNA and 40 units of T7 RNA Polymerase was incubated at 37 °C for 2 hours. mRNA (about 2.6 kilobase (kb)) purification and DNase on-column treatment were both performed by using PureLink RNA Mini Kit. The concentrations of the mRNA products were measured by Nanodrop and the RNA integrity and purity were analysed by 1% (w/v) agarose gel electrophoresis. The mRNAs were re-suspended in DEPC-treated water and were separated into aliquots which were stored at -80 °C.

3.4 Binding of PNAs to model mRNAs and *in vitro* translation

The model mRNAs (experimental and positive control) at 0.1 µM were mixed with different concentrations (0.02 µM, 0.2 µM, 2 µM, 5 µM and 10 µM) of PNA in DEPC-

treated water. The mRNA/PNA mixtures were incubated at room temperature for 20 minutes before adding 70 % of the total sample volume (12.5 µL) of nuclease treated rabbit reticulocyte lysate (RRL) mixture (Promega). *In vitro* translation was conducted at 30 °C for 90 minutes, followed by the quenching of reaction on ice for 20 minutes. The mRNA integrity for the mRNA/PNA (0.14 µM of RNA/12 µM PNA) mixture after the 20-minutes incubation (room temperature) was checked by 1 % (w/v) agarose gel electrophoresis at 100 V for 1 hour at room temperature. The gels were stained by ethidium bromide and imaged by using Syngene G:Box.

3.5 Frameshift assay and data analysis

The luminescence was quantified by using the dual-luciferase reporter assay (Promega). Tecan infinite M200 microplate reader was used for reading the luciferase expression of Renilla and firefly luciferase on Nunclon™ flat-bottom 96-well black microplate. The luciferase readings for both FLuc and RLuc were taken at room temperature immediately after the addition of 50 µL of the respective substrate with an integration time of 3 seconds. After the measurement of one translated sample in a well, the product mixture was removed completely prior to the measurement of the next translated sample in the next well. This is to prevent the interference of the luminescence readings. For each set of mRNA/PNA mixture, two rounds of measurements were carried out, while the *in vitro* translation reactions were conducted three times. Therefore, a total of nine data sets were obtained for each mRNA/PNA mixture. The efficiency of -1 RF was calculated using the equation, $-1 \text{ RF (\%)} = \frac{\text{FLuc (E)}/\text{RLuc (E)}}{\text{FLuc (C)}/\text{RLuc (C)}} \times 100\%^{56}$. The activity ratio of the positive control (C) in the equation was used to normalize the activity ratio of the test sequence – experimental (E) construct⁵⁶. It is assumed that ribosomes which

drop off during the synthesis of FLuc is proportional to the completion of the synthesis of this reporter⁵⁶. To prevent the underestimation of the proportion of ribosomes, the correction factor is provided by a control in which all ribosomes that complete the synthesis of RLuc would continue the translation of second reporter, FLuc⁵⁶. The average of the frameshifting efficiencies and the standard deviations of the six data sets were obtained.

4. Results and Discussion

4.1 Analysis of DNA sequencing results and RNA integrity

DNA sequencing was performed to verify that ligation and site-directed mutagenesis were successfully carried out. By aligning the DNA interest with the full sequence of its corresponding DNA plasmid on Basic Local Alignment Search Tool (BLAST), matching sequences were observed, showing that the desired plasmids and modifications were obtained (Figure 4).

Figure 4. Alignment of DNA sequences (ssRNA, rHP1 and rHP2) between full sequence of DNA plasmid (Query) and DNA interest (Sbjct) by BLAST.

To rule out the possibility that mRNAs degrade upon PNA addition, agarose gel electrophoresis of the model mRNAs in the absence and presence of PNAs was carried out. The gel results reveal no mRNA degradation upon the incubation with the PNAs (P3, P5, and P1) in DEPC-treated water (Figure 5b-d) and incubation buffer (Figure 5e).

Figure 5. Agarose gel electrophoresis (1% (w/v)) of model mRNAs with (b-d) and without the incubation with PNA (a). Incubation of mRNA/PNA (in DEPC-treated water) was conducted for 20 minutes at room temperature and gel electrophoresis was carried out at 100 V for 1 hour. (e) Incubation of rHP1 and PNA (P3, P5 and P1) was performed at 20 minutes at room temperature in incubation buffer (200 mM NaCl, 0.5 mM EDTA, 20 mM HEPES, pH 7.5) and the samples were electrophorized at 100 V for 90 minutes. All the gels were stained with ethidium bromide and imaged by using Syngene G:Box.

4.2 Frameshifting assay of model mRNA by PNA

The -1 RF efficiency of model mRNAs upon the addition of PNA was studied by frameshifting assay. Previous studies have revealed that chemically-modified PNAs P3 and P5 bind specifically to 32-mer hairpins rHP2 ($K_D = 4.7 \pm 0.6 \mu\text{M}$) and rHP1 ($K_D = 0.2 \pm 0.1 \mu\text{M}$), respectively⁵⁴. rHP2 differs from rHP1 by having an inverted C-G pair as the ninth base pair from the bottom, and ssRNA only contains the 5'-half of the hairpin rHP1 (Figure 3). We tested the effects of the formation of P3·rHP2 and P5·rHP1 complexes on -1 RF efficiency of the mRNAs containing rHP2 and rHP1, respectively (Figure 6). The PNAs were incubated with the mRNAs for 20 mins before the cell-free *in vitro* translation reaction. The *in vitro* assay reveals that the frameshifting efficiency of rHP2 increases with increasing concentration of P3, reaching 13% at 10 μM P3 (Figure 6a). Consistent with the fact that PNA P3 does not bind to ssRNA or rHP1⁵⁴, the addition of up to 10 μM of P3 essentially did not change -1 RF efficiency for ssRNA and rHP1, which remained at 1% and 2%, respectively, at all concentrations of P3.

In correlation with the fact that P5·rHP1 triplex is more stable than P3·rHP2 triplex, we observed that P5 had a relatively higher activity in stimulating -1 RF of rHP1 (Figure 6b). The -1 RF efficiency of rHP1 increases from 2% to 32% by the addition of 10 μM of P5 (Figure 6b). P5 also shows exquisite selectivity, with no significant effects on the -1 RF of both ssRNA and rHP2, in agreement with the weak binding of P5 for ssRNA and rHP2⁵³⁻⁵⁴.

Figure 6. PNAs P3 and P5 stimulate -1 RF through binding to dsRNAs. The modified PNAs tested are PNA P3 (NH₂-Lys-TLTQTTTL-CONH₂) (panel a) and PNA P5 (NH₂-Lys-TLTLTTTL-CONH₂) (panel b). With no addition of PNAs, the -1 RF efficiencies of mRNAs ssRNA, rHP1, and rHP2 are 1%, 2%, and 2%, respectively.

The high selectivity is also manifested by the fact that the expression of RLuc (in the 0 frame) is not significantly affected by the modified PNAs P3 and P5 (Figure 7a-b). Thus, the -1 RF assay results suggest that our chemically-modified PNAs P3 and P5 show structural selectivity and sequence specificity in stimulating the recoding of ribosomal reading frame. Presumably, PNAs P3 and P5 stabilize the structural stabilities of rHP2 and rHP1, respectively, and disfavor the formation of alternative mRNA structures⁶⁰⁻⁶². Our results support the hypothesis that PNAs P3 and P5 stimulate -1 RF through binding to and stabilizing rHP2 and rHP1 respectively. It is also possible that the stimulation of -1 RF is caused by the interactions of the PNA·RNA-RNA triplex with the ribosome. The orthogonal *trans*-activation of -1 RF by modified PNAs (P3 for rHP2 and P5 for rHP1) indicates that the ribosomal recoding process is mRNA template sequence and PNA sequence specific, and is not facilitated by direct interaction of PNAs with ribosome.

Unmodified PNAs are known to be able to bind to complementary ssRNAs and partially structured RNAs through strand invasion⁶³. As compared to our chemically-modified PNAs, unmodified PNAs may show less selectivity and specificity in binding to structured RNAs. For example, unmodified PNA P1 can bind to both ssRNA and rHP1⁵³⁻⁵⁴. The effects of PNA P1 on –1 RF was hence tested. As illustrated in Figure 7c, with increasing concentration of unmodified PNA P1 added, the expression levels of both FLuc and RLuc luciferase drop significantly to baseline level for all mRNAs (ssRNA, rHP1, and rHP2). Thus, it is possible that PNA P1 inhibits translation by binding to rRNAs and/or coding regions of Renilla/firefly luciferase genes.

Figure 7. Effects of PNAs on the expression of FLuc and RLuc. The expression levels are normalized. (a) PNA P3 selectively enhances the expression of FLuc of the experimental mRNA rHP2 construct. (b) PNA P5 selectively enhances the expression of FLuc of the experimental mRNA rHP1 construct. (c) Unmodified PNA P1 shows non-selective inhibition of the expression of both FLuc and RLuc.

Table 1-3 display the raw values of quantified FLuc and RLuc in ssRNA, rHP1 and rHP2 respectively upon the addition of PNA. The numbers shown in the tables are one of replicates and are before normalization. A general increasing trend is observed for both the FLuc of rHP1 and rHP2 with the addition of P5 and P3 respectively, while having a minimal decrease in their RLuc values. The value of FLuc for rHP1 increases from 1344 to 9954 by the addition of 2 μ M P5. However, the FLuc value starts to decrease when higher concentration of P5 was introduced. This could be due to nonspecific interaction of translation in rHP1 when higher concentration of PNA was added. The FLuc value of P3 increases from 1427 to 6933 by the addition of 10 μ M of P3. The addition of unmodified PNA P1, on the other hand, causes both the FLuc and RLuc values to decrease drastically for ssRNA, rHP1 and rHP2 with increasing concentration of P1.

Table 1. Raw luminescence values of FLuc and RLuc in ssRNA with the addition of PNA

Concentration of PNA (μ M)	Construct	ssRNA P3		ssRNA P5		ssRNA P1	
		FLuc	RLuc	FLuc	RLuc	FLuc	RLuc
0	Experimental	742	279260	757	321866	748	290055
	Control	28549	143243	22990	130947	24994	123539
0.02	Experimental	900	361251	744	297612	769	278905
	Control	27222	144917	24192	135198	22583	119363
0.2	Experimental	903	355749	1077	353502	1026	237846
	Control	27604	150223	25039	152334	19029	107550
2	Experimental	784	312955	1631	360995	261	36893
	Control	25764	134371	23874	150970	1642	23670
5	Experimental	710	276226	1915	312087	20	14292
	Control	22795	133373	23414	137586	197	11044
10	Experiment	712	276618	1911	233949	2	3435
	Control	15165	101131	18162	112360	23	2596

Table 2. Raw luminescence values of FLuc and RLuc in rHP1 with the addition of PNA

Concentration of PNA (μ M)	Construct	rHP1 P3		rHP1 P5		rHP1 P1	
		FLuc	RLuc	FLuc	RLuc	FLuc	RLuc
0	Experimental	1441	283055	1344	322944	1386	302756
	Control	30301	141672	27602	157649	26974	133892
0.02	Experimental	1405	299678	1545	297975	1410	314010
	Control	29050	131304	28923	152414	26285	138331
0.2	Experimental	1579	320957	2854	307437	1079	229756
	Control	28872	139399	27686	157083	23135	117360
2	Experimental	1378	303803	9954	245120	272	44128
	Control	27475	138370	24271	163809	2290	22067
5	Experimental	1415	306951	9641	210398	63	19467
	Control	20580	113364	19146	134315	302	9323
10	Experiment	1252	280770	8222	172748	4	4902
	Control	17716	112247	16353	120484	39	2563

Table 3. Raw luminescence values of FLuc and RLuc in rHP2 with the addition of PNA

Concentration of PNA (μ M)	Construct	rHP2 P3		rHP2 P5		rHP2 P1	
		FLuc	RLuc	FLuc	RLuc	FLuc	RLuc
0	Experimental	1427	333059	1108	279460	1348	318817
	Control	30503	159960	21586	127976	27087	147225
0.02	Experimental	1423	319355	1144	281703	1363	308042
	Control	30497	163174	20862	127346	25610	144302
0.2	Experimental	1853	341929	1104	296366	1042	237690
	Control	33875	173688	22601	128018	22177	114228
2	Experimental	3787	325357	1150	291630	128	51789
	Control	31046	165449	21098	129626	2434	23271
5	Experimental	5736	285438	1258	271730	14	26124
	Control	27368	146700	19005	117789	401	12393
10	Experiment	6933	256792	1340	225423	1	6630
	Control	23424	123420	14949	93201	66	3690

4.3 Analysis of potential binding sites of PNA P1

By examining the sequences of Homo sapiens ribosomal RNAs (rRNAs) (similar to rabbit rRNAs)⁶⁰ and p2luc vector, several possible off-target sites containing sequences similar to 5'-AGAGAAAG-3' and 3'-AGAGAAAG-5' were identified (Figure 8). We used the secondary structures of human rRNAs because (i) human rRNAs are very similar to rabbit rRNAs, and (ii) the secondary structure of the large subunit rabbit rRNA is not available⁶⁴. PNA P1 can form parallel and antiparallel duplexes with 5'-AGAGAAAG-3' and 3'-AGAGAAAG-5', respectively. It is highly probable that the nonspecific inhibition of translation by P1 is due to duplex formation between P1 and the identified rRNA regions (Figure 8a), resulting in the obstruction of ribosomal functions⁶⁵⁻⁶⁶. In addition, binding of P1 to the coding regions of RLuc gene (Figure 8b) may affect the ribosomal translocation on mRNA.

The identified potential binding sites in rRNAs for the unmodified PNA P1 are essentially single-stranded, and may not be recognized by chemically-modified PNAs P3 and P5 because P3 and P5 show selective binding towards dsRNAs over ssRNAs⁵⁴. Consistently, a less prominent reduction was observed in the RLuc readings upon the addition of P3 and P5 (Figure 7a-b). Our data indicate that chemically-modified PNAs P3 and P5 site-specifically bind to rHP2 and rHP1, respectively, with no significant binding to ribosome or other regions of mRNA.

Figure 8. Examples of potential off-target binding sites of unmodified PNA P1. The identified sequences are similar to RNA (5'-AGAGAAAG-3') or RNA 3'-AGAGAAAG-5', and may form parallel or antiparallel PNA-RNA duplexes, respectively, with unmodified PNA P1. (a) Examples of potential binding regions found in human small subunit rRNA (HS1 and HS2) and large subunit rRNA (HL2, HL3, and HL4). The red and blue sequences indicate the identified partially base paired regions may form parallel and antiparallel duplexes with PNA P1, respectively. (b) DNA sequences of part of rHP1-E construct based on p2luc vector (see Appendix for the complete sequence). PNA P1 may bind to two potential sites (blue underlined) to form RNA-PNA duplexes. The cyan region represents the DNA sequence inserted for the construction of rHP1-E construct.

4.4 Different conditions of *in vitro* translation assay

We tested if pre-incubation of PNA with mRNA would affect the efficacy of -1 RF stimulation. We observed that even without the 20-min pre-incubation, P3 and P5 are still able to induce -1 RF of rHP2 and rHP1, respectively (Figure 9). The results are consistent with the fact that PNA·RNA-RNA triplex formation is relatively fast⁶⁷⁻⁶⁸ and the two modified PNAs (P3 and P5) activate -1 RF during the multi-turnover of translation elongation. We also tested if the PNAs affect the activities of the luciferases by adding the PNA after the translation reactions were completed. For all the mRNAs, upon the post-translational addition of the PNAs (modified PNAs P3 and P5 and unmodified PNA P1), we observed no significant changes in the -1 RF efficiency (Figure 10) and expression levels of RLuc and FLuc (Figure 11). Thus, the unmodified and modified PNAs show no direct effect on the luciferase activities. Taken together, our results verify that the changes in -1 RF efficiency (Figure 6) result from the specific interactions between the mRNA hairpins and the dsRNA-binding modified PNAs.

Figure 9. PNAs P3 and P5 selectively stimulate -1 RF efficiencies of model mRNAs rHP2 and rHP1, respectively, without mRNA-PNA pre-incubation. The modified PNAs P3 (NH₂-Lys-TLTQTTTL-CONH₂) (panel a) and PNA P5 (NH₂-Lys-TLTLTTTL-CONH₂) (panel b) were added to the mRNA (0.1 μM) and mixed concurrently with the nuclease treated Rabbit Reticulocyte Lysate (RRL). *In vitro* translation was conducted once with two rounds of measurements. Thus, the error bars are smaller than the symbols.

Figure 10. Minus-one RF efficiency is not affected with PNAs added after the *in vitro* translation reactions. 0.25 μ L PNAs at different concentrations (Stock: 0, 0.2, 2, 20, 50 and 100 μ M) were incubated with 2.25 μ L translated product on ice for 20 minutes. (a) PNA P3 (NH₂-Lys-TLTQTTL-CONH₂) and mRNA rHP2. (b) PNA P5 (NH₂-Lys-TLTLTTTL-CONH₂) and mRNA rHP1. (c) PNA P1 (NH₂-Lys-TCTCTTTC-CONH₂) and mRNA rHP1. The mRNA concentration used was 0.1 μ M.

Figure 11. The activities of Renilla and firefly luciferases are not directly affected by the PNAs added after the *in vitro* translation reactions of the mRNAs (0.1 μ M). *In vitro* translation was conducted twice with one round of measurement each. 0.25 μ L PNAs at different concentrations (Stock: 0, 0.2, 2, 20, 50 and 100 μ M) were incubated with 2.25 μ L translated product on ice for 20 minutes. (a) PNA P3 (NH₂-Lys-TLTQTTL-CONH₂) and mRNA rHP2. (b) PNA P5 (NH₂-Lys-TLTLTTTL-CONH₂) and mRNA rHP1. (c) PNA P1 (NH₂-Lys-TCTCTTTC-CONH₂) and mRNA rHP1.

5. Conclusion

In summary, this study suggests that the modified dsRNA-binding PNAs are advantageous compared to unmodified PNAs for targeting structured RNAs. It also reveals that relatively short chemically-modified dsRNA-binding PNAs may be used to probe the complex ribosomal translation processes. Formation of a major-groove parallel PNA·RNA-RNA triplex may stimulate -1 RF due to the stabilization of mRNA duplex structures, and/or the interference of the direct interactions of mRNA duplex at the entrance tunnel with the ribosome at the slippery site. The chemically-modified PNAs platform may be applied in the future to the targeting of viral RNA structures to uncover the roles cis-acting mRNA structures in -1 PRF. To allow the production of a protein in the downstream -1 frame, the PNA·RNA-RNA triplex needs to be unwound for continued translation in -1 frame. It is known that an overly stable mRNA structure may block translation instead of stimulating -1 RF⁶⁹. Thus, it is expected that the stimulation of -1 RF and other translational regulations will be highly dependent on the sequence, length, and binding position of the dsRNA-binding PNAs. Future structural and mechanistic studies will help develop dsRNA-binding PNAs with further improved efficacy and selectivity in stimulating -1 RF. This study also provides insights into future development of highly selective and potent PNAs targeting rRNA structures. It is also important to consider the effects of PNAs on other types of RNA secondary conformations such as multiple RNA stem-loop and pseudoknot structures that are present in various RNA viruses. With the advances in PNA delivery technologies^{54, 70-74}, this study on the model mRNAs may also serve as a basis for the future development of antiviral drugs targeting viral mRNA structures that stimulate -1 PRF, and therapeutics targeting many other biomedically-important RNAs.

6. References

1. Bevilacqua, P. C.; Blose, J. M., Structures, kinetics, thermodynamics, and biological functions of RNA hairpins. *Annu Rev Phys Chem* **2008**, *59*, 79-103.
2. Davis, A. R.; Kirkpatrick, C. C.; Znosko, B. M., Structural characterization of naturally occurring RNA single mismatches. *Nucleic Acids Res* **2011**, *39* (3), 1081-1094.
3. Morse, S. E.; Draper, D. E., Purine Purine Mismatches in Rna Helices - Evidence for Protonated G-Center-Dot-a Pairs and Next-Nearest-Neighbor Effects. *Nucleic Acids Res* **1995**, *23* (2), 302-306.
4. Svoboda, P.; Di Cara, A., Hairpin RNA: a secondary structure of primary importance. *Cell Mol Life Sci* **2006**, *63* (7-8), 901-918.
5. Weeks, K. M.; Crothers, D. M., Major Groove Accessibility of Rna. *Science* **1993**, *261* (5128), 1574-1577.
6. Tok, J. B.; Bi, L.; Saenz, M., Specific recognition of naphthyridine-based ligands toward guanine-containing bulges in RNA duplexes and RNA-DNA heteroduplexes. *Bioorg Med Chem Lett* **2005**, *15* (3), 827-31.
7. Bevilacqua, P. C., Battle for the bulge: directing small molecules to DNA and RNA defects. *Chem Biol* **2002**, *9* (8), 854-5.
8. Hall, K. B., RNA-protein interactions. *Curr Opin Struc Biol* **2002**, *12* (3), 283-288.
9. Beuth, B.; Garcia-Mayoral, M. F.; Taylor, I. A.; Ramos, A., Scaffold-independent analysis of RNA-Protein interactions: The nova-1 KH3-RNA complex. *J. Am. Chem. Soc.* **2007**, *129* (33), 10205-10210.
10. Caliskan, N.; Peske, F.; Rodnina, M. V., Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. *Trends Biochem Sci* **2015**, *40* (5), 265-74.
11. Kurland, C. G., Translational accuracy and the fitness of bacteria. *Annu Rev Genet* **1992**, *26*, 29-50.
12. Drummond, D. A.; Wilke, C. O., The evolutionary consequences of erroneous protein synthesis. *Nat Rev Genet* **2009**, *10* (10), 715-24.
13. Farabaugh, P. J., Programmed translational frameshifting. *Microbiol Rev* **1996**, *60* (1), 103-&.

14. Dinman, J. D., Mechanisms and implications of programmed translational frameshifting. *Wires Rna* **2012**, *3* (5), 661-673.
15. Mouzakis, K. D.; Lang, A. L.; Vander Meulen, K. A.; Easterday, P. D.; Butcher, S. E., HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. *Nucleic Acids Res* **2013**, *41* (3), 1901-1913.
16. Chen, G.; Chang, K. Y.; Chou, M. Y.; Bustamante, C.; Tinoco, I., Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *P Natl Acad Sci USA* **2009**, *106* (31), 12706-12711.
17. Zhong, Z. S.; Yang, L. X.; Zhang, H. P.; Shi, J. H.; Vandana, J. J.; Lam, D. T. U. H.; Olsthoorn, R. C. L.; Lu, L. Y.; Chen, G., Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. *Sci Rep-Uk* **2016**, *6*.
18. Chen, Y. T.; Chang, K. C.; Hu, H. T.; Chen, Y. L.; Lin, Y. H.; Hsu, C. F.; Chang, C. F.; Chang, K. Y.; Wen, J. D., Coordination among tertiary base pairs results in an efficient frameshift-stimulating RNA pseudoknot. *Nucleic Acids Res* **2017**, *45* (10), 6011-6022.
19. Yu, C. H.; Teulade-Fichou, M. P.; Olsthoorn, R. C., Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res* **2014**, *42* (3), 1887-92.
20. Olsthoorn, R. C. L.; Laurs, M.; Sohet, F.; Hilbers, C. W.; Heus, H. A.; Pleij, C. W. A., Novel application of sRNA: Stimulation of ribosomal frameshifting. *Rna* **2004**, *10* (11), 1702-1703.
21. Yu, C. H.; Noteborn, M. H.; Olsthoorn, R. C., Stimulation of ribosomal frameshifting by antisense LNA. *Nucleic Acids Res* **2010**, *38* (22), 8277-83.
22. Caliskan, N.; Katunin, V. I.; Belardinelli, R.; Peske, F.; Rodnina, M. V., Programmed -1 frameshifting by kinetic partitioning during impeded translocation. *Cell* **2014**, *157* (7), 1619-31.
23. Chen, J.; Petrov, A.; Johansson, M.; Tsai, A.; O'Leary, S. E.; Puglisi, J. D., Dynamic pathways of -1 translational frameshifting. *Nature* **2014**, *512* (7514), 328-32.
24. Frank, J.; Gao, H.; Sengupta, J.; Gao, N.; Taylor, D. J., The process of mRNA-tRNA translocation. *Proc Natl Acad Sci U S A* **2007**, *104* (50), 19671-8.

25. Chen, C.; Zhang, H.; Broitman, S. L.; Reiche, M.; Farrell, I.; Cooperman, B. S.; Goldman, Y. E., Dynamics of translation by single ribosomes through mRNA secondary structures. *Nat Struct Mol Biol* **2013**, *20* (5), 582-8.
26. Kim, H. K.; Liu, F.; Fei, J.; Bustamante, C.; Gonzalez, R. L., Jr.; Tinoco, I., Jr., A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation. *Proc Natl Acad Sci U S A* **2014**, *111* (15), 5538-43.
27. Yan, S.; Wen, J. D.; Bustamante, C.; Tinoco, I., Jr., Ribosome excursions during mRNA translocation mediate broad branching of frameshift pathways. *Cell* **2015**, *160* (5), 870-81.
28. Qin, P. W.; Yu, D. M.; Zuo, X. B.; Cornish, P. V., Structured mRNA induces the ribosome into a hyper-rotated state. *Embo Rep* **2014**, *15* (2), 185-190.
29. Jacks, T.; Power, M. D.; Masiarz, F. R.; Luciw, P. A.; Barr, P. J.; Varmus, H. E., Characterization of Ribosomal Frameshifting in Hiv-1 Gag-Pol Expression. *Nature* **1988**, *331* (6153), 280-283.
30. Belew, A. T.; Meskauskas, A.; Musalgaonkar, S.; Advani, V. M.; Sulima, S. O.; Kasprzak, W. K.; Shapiro, B. A.; Dinman, J. D., Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. *Nature* **2014**, *512* (7514), 265-+.
31. Dinman, J. D.; Ruiz-Echevarria, M. J.; Peltz, S. W., Translating old drugs into new treatments: ribosomal frameshifting as a target for antiviral agents. *Trends Biotechnol* **1998**, *16* (4), 190-196.
32. Marcheschi, R. J.; Mouzakis, K. D.; Butcher, S. E., Selection and Characterization of Small Molecules That Bind the HIV-1 Frameshift Site RNA. *Acs Chem Biol* **2009**, *4* (10), 844-854.
33. Gareiss, P. C.; Miller, B. L., Ribosomal frameshifting: An emerging drug target for HIV. *Curr Opin Invest Dr* **2009**, *10* (2), 121-128.
34. Giedroc, D. P.; Cornish, P. V., Frameshifting RNA pseudoknots: Structure and mechanism. *Virus Res* **2009**, *139* (2), 193-208.
35. Brakier-Gingras, L.; Charbonneau, J.; Butcher, S. E., Targeting frameshifting in the human immunodeficiency virus. *Expert Opin Ther Tar* **2012**, *16* (3), 249-258.
36. Hilimire, T. A.; Bennett, R. P.; Stewart, R. A.; Garcia-Miranda, P.; Blume, A.; Becker, J.; Sherer, N.; Helms, E. D.; Butcher, S. E.; Smith, H. C.; Miller, B. L., N-Methylation as a Strategy for Enhancing the Affinity and Selectivity of RNA-

- binding Peptides: Application to the HIV-1 Frameshift-Stimulating RNA. *Acs Chem Biol* **2016**, *11* (1), 88-94.
37. Liphardt, J.; Naphine, S.; Kontos, H.; Brierley, I., Evidence for an RNA pseudoknot loop-helix interaction essential for efficient-1 ribosomal frameshifting. *J Mol Biol* **1999**, *288* (3), 321-335.
 38. Namy, O.; Moran, S. J.; Stuart, D. I.; Gilbert, R. J. C.; Brierley, I., A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* **2006**, *441* (7090), 244-247.
 39. Garcia-Miranda, P.; Becker, J. T.; Benner, B. E.; Blume, A.; Sherer, N. M.; Butcher, S. E., Stability of HIV Frameshift Site RNA Correlates with Frameshift Efficiency and Decreased Virus Infectivity. *J Virol* **2016**, *90* (15), 6906-17.
 40. Zhang, X. Y.; Xu, X. J.; Yang, Z. Y.; Burcke, A. J.; Gates, K. S.; Chen, S. J.; Gu, L. Q., Mimicking Ribosomal Vectorial Unfolding of RNA Pseudoknot in a Protein Channel. *Biophys J* **2016**, *110* (3), 408a-408a.
 41. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E., PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **1993**, *365* (6446), 566-8.
 42. Nielsen, P. E.; Egholm, M.; Buchardt, O., Peptide nucleic acid (PNA). A DNA mimic with a peptide backbone. *Bioconjug Chem* **1994**, *5* (1), 3-7.
 43. Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Norden, B., DNA-Like Double Helix Formed by Peptide Nucleic-Acid. *Nature* **1994**, *368* (6471), 561-563.
 44. Hyrup, B.; Nielsen, P. E., Peptide nucleic acids (PNA): Synthesis, properties and potential applications. *Bioorgan Med Chem* **1996**, *4* (1), 5-23.
 45. Eriksson, M.; Nielsen, P. E., PNA-nucleic acid complexes. Structure, stability and dynamics. *Q Rev Biophys* **1996**, *29* (4), 369-94.
 46. Nielsen, P. E.; Egholm, M., An introduction to peptide nucleic acid. *Curr Issues Mol Biol* **1999**, *1* (1-2), 89-104.
 47. Ray, A.; Norden, B., Peptide nucleic acid (PNA): its medical and biotechnical applications and promise for the future. *Faseb Journal* **2000**, *14* (9), 1041-1060.
 48. Jarver, P.; O'Donovan, L.; Gait, M. J., A Chemical View of Oligonucleotides for Exon Skipping and Related Drug Applications. *Nucleic Acid Ther* **2014**, *24* (1), 37-47.

49. Zhou, P.; Wang, M. M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H., Novel binding and efficient cellular uptake of guanidine-based peptide nucleic acids (GPNA). *J. Am. Chem. Soc.* **2003**, *125* (23), 6878-6879.
50. Gupta, P.; Muse, O.; Rozners, E., Recognition of Double-Stranded RNA by Guanidine-Modified Peptide Nucleic Acids. *Biochemistry-Us* **2012**, *51* (1), 63-73.
51. Moccia, M.; Adamo, M. F.; Saviano, M., Insights on chiral, backbone modified peptide nucleic acids: Properties and biological activity. *Artif DNA PNA XNA* **2014**, *5* (3), e1107176.
52. Devi, G.; Zhou, Y.; Zhong, Z. S.; Toh, D. F. K.; Chen, G., RNA triplexes: from structural principles to biological and biotech applications. *Wires Rna* **2015**, *6* (1), 111-128.
53. Devi, G.; Yuan, Z.; Lu, Y. P.; Zhao, Y. L.; Chen, G., Incorporation of thio-pseudoisocytosine into triplex-forming peptide nucleic acids for enhanced recognition of RNA duplexes. *Nucleic Acids Res* **2014**, *42* (6), 4008-4018.
54. Toh, D. K.; Devi, G.; Patil, K. M.; Qu, Q.; Maraswami, M.; Xiao, Y.; Loh, T. P.; Zhao, Y.; Chen, G., Incorporating a guanidine-modified cytosine base into triplex-forming PNAs for the recognition of a C-G pyrimidine-purine inversion site of an RNA duplex. *Nucleic Acids Res* **2016**, *44* (19), 9071-9082.
55. Marcheschi, R. J.; Staple, D. W.; Butcher, S. E., Programmed ribosomal frameshifting is SIV is induced by a highly structured RNA stem-loop. *J Mol Biol* **2007**, *373* (3), 652-663.
56. Grentzmann, G.; Ingram, J. A.; Kelly, P. J.; Gesteland, R. F.; Atkins, J. F., A dual-luciferase reporter system for studying recoding signals. *Rna* **1998**, *4* (4), 479-486.
57. Barry, J. K.; Miller, W. A., A-1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. *P Natl Acad Sci USA* **2002**, *99* (17), 11133-11138.
58. Nikolic, E. I. C.; King, L. M.; Vidakovic, M.; Irigoyen, N.; Brierley, I., Modulation of Ribosomal Frameshifting Frequency and Its Effect on the Replication of Rous Sarcoma Virus. *Journal of Virology* **2012**, *86* (21), 11581-11594.

59. Henderson, C. M.; Anderson, C. B.; Howard, M. T., Antisense-induced ribosomal frameshifting. *Nucleic Acids Res* **2006**, *34* (15), 4302-4310.
60. Ritchie, D. B.; Foster, D. A.; Woodside, M. T., Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc Natl Acad Sci U S A* **2012**, *109* (40), 16167-72.
61. Ritchie, D. B.; Soong, J.; Sikkema, W. K. A.; Woodside, M. T., Anti-frameshifting Ligand Reduces the Conformational Plasticity of the SARS Virus Pseudoknot. *J. Am. Chem. Soc.* **2014**, *136* (6), 2196-2199.
62. Ritchie, D. B.; Cappellano, T. R.; Tittle, C.; Rezajooei, N.; Rouleau, L.; Sikkema, W. K.; Woodside, M. T., Conformational dynamics of the frameshift stimulatory structure in HIV-1. *Rna* **2017**.
63. Armitage, B. A., The impact of nucleic acid secondary structure on PNA hybridization. *Drug Discov Today* **2003**, *8* (5), 222-228.
64. Cannone, J. J.; Subramanian, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y. S.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; Pande, N.; Shang, Z. D.; Yu, N.; Gutell, R. R., The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs: Correction (vol 3, pg 2, 2002). *Bmc Bioinformatics* **2002**, *3*.
65. Auerbach-Nevo, T.; Baram, D.; Bashan, A.; Belousoff, M.; Breiner, E.; Davidovich, C.; Camicata, G.; Eyal, Z.; Halfon, Y.; Krupkin, M.; Matzov, D.; Metz, M.; Rufayda, M.; Peretz, M.; Pick, O.; Pyetan, E.; Rozenberg, H.; Shalev-Benami, M.; Wekselman, I.; Zarivach, R.; Zimmerman, E.; Assis, N.; Bloch, J.; Israeli, H.; Kalaora, R.; Lim, L.; Sade-Falk, O.; Shapira, T.; Taha-Salaime, L.; Tang, H.; Yonath, A., Ribosomal Antibiotics: Contemporary Challenges. *Antibiotics-Basel* **2016**, *5* (3).
66. Kulik, M.; Markowska-Zagrajek, A.; Wojciechowska, M.; Grzela, R.; Witula, T.; Trylska, J., Helix 69 of Escherichia coli 23S ribosomal RNA as a peptide nucleic acid target. *Biochimie* **2017**, *138*, 32-42.
67. Sato, T.; Sato, Y.; Nishizawa, S., Triplex-Forming Peptide Nucleic Acid Probe Having Thiazole Orange as a Base Surrogate for Fluorescence Sensing of Double-stranded RNA. *J. Am. Chem. Soc.* **2016**, *138* (30), 9397-9400.

68. Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N., Nucleobase-Modified PNA Suppresses Translation by Forming a Triple Helix with a Hairpin Structure in mRNA InVitro and in Cells. *Angew Chem Int Edit* **2016**, 55 (3), 899-903.
69. Tholstrup, J.; Oddershede, L. B.; Sorensen, M. A., mRNA pseudoknot structures can act as ribosomal roadblocks. *Nucleic Acids Res* **2012**, 40 (1), 303-313.
70. Ma, X.; Devi, G.; Qu, Q. Y.; Toh, D. F. K.; Chen, G.; Zhao, Y. L., Intracellular Delivery of Antisense Peptide Nucleic Acid by Fluorescent Mesoporous Silica Nanoparticles. *Bioconjugate Chem* **2014**, 25 (8), 1412-1420.
71. Avitabile, C.; Accardo, A.; Ringhieri, P.; Morelli, G.; Saviano, M.; Montagner, G.; Fabbri, E.; Gallerani, E.; Gambari, R.; Romanelli, A., Incorporation of Naked Peptide Nucleic Acids into Liposomes Leads to Fast and Efficient Delivery. *Bioconjugate Chem* **2015**, 26 (8), 1533-1541.
72. Gupta, A.; Bahal, R.; Gupta, M.; Glazer, P. M.; Saltzman, W. M., Nanotechnology for delivery of peptide nucleic acids (PNAs). *J Control Release* **2016**, 240, 302-311.
73. Ellipilli, S.; Palvai, S.; Ganesh, K. N., Fluorinated Peptide Nucleic Acids with Fluoroacetyl Side Chain Bearing 5-(F/CF₃)-Uracil: Synthesis and Cell Uptake Studies. *J Org Chem* **2016**, 81 (15), 6364-6373.
74. Hnedzko, D.; Mcgee, D. W.; Karamitas, Y. A.; Rozners, E., Sequence-selective recognition of double-stranded RNA and enhanced cellular uptake of cationic nucleobase and backbone-modified peptide nucleic acids. *Rna* **2017**, 23 (1), 58-69.

7. Appendices

Sequences of plasmids and secondary structures of rRNAs

ssRNA-E/p2Luc

5417 bp ds-DNA

1	AGATCGATCT	GCGCAGCACC	ATGGCCTGAA	ATAACCTCTG	AAAGAGGAAC	TTGGTTAGGT
61	ACCTTCTGAG	GCGGAAAGAA	CCAGCTGTGG	AATGTGTGTC	AGTTAGGGTG	TGGAAAGTCC
121	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC	TCAATTAGTC	AGCAACCAGG
181	TGTGGAAAGT	CCCCAGGCTC	CCCAGCAGGC	AGAAGTATGC	AAAGCATGCA	TCTCAATTAG
241	TCAGCAACCA	TAGTCCCGCC	CCTAACTCCG	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC
301	GCCCATTCTC	CGCCCCATGG	CTGACTAATT	TTTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC
361	TCGGCCTCTG	AGCTATTCCA	GAAGTAGTGA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC
421	AAAAAGCTTG	ATTCTTCTGA	CACAACAGTC	TCGAACCTAA	GCTGCAGAAG	TTGGTCGTGA
481	GGCACTGGGC	AGGTAAGTAT	CAAGGTTACA	AGACAGGTTT	AAGGAGACCA	ATAGAAACTG
541	GGCTTGTCGA	GACAGAGAAG	ACTCTTGCGT	TTCTGATAGG	CACCTATTGG	TCTTACTGAC
601	ATCCACTTTG	CCTTCTCTCT	CACAGGTGTC	CACTCCCAGT	TCAATTACAG	CTCTTAAGGC
661	TAGAGTACTT	AATACGACTC	ACTATAGGCT	AGCCACCATG	ACTTCGAAAG	TTTATGATCC
721	AGAACAAGG	AAACGGATGA	TAAGTGGTCC	GCAGTGGTGG	GCCAGATGTA	AACAAATGAA
781	TGTTCTTGAT	TCATTTATTA	ATTATTATGA	TTCAGAAAAA	CATGCAGAAA	ATGCTGTTAT
841	TTTTTTACAT	GGTAACGCGG	CCTCTTCTTA	TTTATGGCGA	CATGTTGTGC	CACATATTGA
901	GCCAGTAGCG	CGGTGTATTA	TACCAGACCT	TATTGGTATG	GGCAAATCAG	GCAAATCTGG
961	TAATGGTTCT	TATAGGTTAC	TTGATCATT	CAAATATCTT	ACTGCATGGT	TTGAACCTCT
1021	TAATTTACCA	AAGAAGATCA	TTTTTGTCTG	CCATGATTGG	GGTGCTTGTT	TGGCATTTC
1081	TTATAGCTAT	GAGCATCAAG	ATAAGATCAA	AGCAATAGTT	CACGCTGAAA	GTGTAGTAGA
1141	TGTGATTGAA	TCATGGGATG	AATGGCCTGA	TATTGAAGAA	GATATTGCGT	TGATCAAATC
1201	TGAAGAAGGA	GAAAAAATGG	TTTTGGAGAA	TAAGTCTTTC	GTGGAAACCA	TGTTGCCATC
1261	AAAAATCATG	AGAAAGTTAG	AACCAAGAGA	ATTTGCAGCA	TATCTTGAAC	CATTCAAAGA
1321	GAAAGGTGAA	GTTTCGTCTG	CAACATTATC	ATGGCCTCGT	GAAATCCCGT	TAGTAAAAGG
1381	TGGTAAACCT	GACGTTGTAC	AAATTGTTAG	GAATTATAAT	GCTTATCTAC	GTGCAAGTGA
1441	TGATTTACCA	AAAATGTTTA	TTGAATCGGA	CCCAGGATTC	TTTTCCAATG	CTATTGTTGA
1501	AGGTGCCAAG	AAGTTTCCTA	ATACTGAATT	TGTCAAAGTA	AAAGGTCTTC	ATTTTTTCGCA
1561	AGAAGATGCA	CCTGATGAAA	TGGGAAAATA	TATCAAATCG	TTGTTTGAGC	GAGTTCTCAA
1621	AAATGAACAA	ATGTCGACGG	ATCCTTTTTT	AGGGTAGAGA	GAGAAAAGTGA	GCTCGAAGAC
1681	GCCAAAAACA	TAAAGAAAGG	CCCGGCGCCA	TTCTATCCTC	TAGAGGATGG	AACCGCTGGA
1741	GAGCAACTGC	ATAAGGCTAT	GAAGAGATAC	GCCCTGGTTC	CTGGAACAA	TGCTTTTACA
1801	GATGCACATA	TCGAGGTGAA	CATCAGGTAC	GCGGAATACT	TCGAAATGTC	CGTTCCGTTG
1861	GCAGAAGCTA	TGAAACGATA	TGGGCTGAAT	ACAAATCACA	GAATCGTCGT	ATGCAGTGAA
1921	AAGTCTCTTC	AATTCTTTAT	GCCGGTGTTC	GGCGCGTTAT	TTATCGGAGT	TGCAGTTGCG
1981	CCCGCGAACG	ACATTTTATA	TGAACGTGAA	TTGCTCAACA	GTATGAACAT	TTGCGAGCCT
2041	ACCGTAGTGT	TTGTTTCCAA	AAAGGGGTTC	CAAAAAATTT	TGAACGTGCA	AAAAAATTTA
2101	CCAATAATCC	AGAAAATTAT	TATCATGGAT	TCTAAAACGG	ATTACCAGGG	ATTTTCAGTCG
2161	ATGTACACGT	TCGTACACATC	TCATCTACCT	CCCGGTTTTT	ATGAATACGA	TTTTGTACCA
2221	GAGTCCTTTG	ATCGTGACAA	AACAATTGCA	CTGATAATGA	ATTCTCTGCG	ATCTACTGGG
2281	TTACCTAAGG	GTGTGGCCCT	TCCGCATAGA	ACTGCCTGCG	TCAGATTCTC	GCATGCCAGA
2341	GATCCTATTT	TTGGCAATCA	AATCATTCCG	GATACTGCGA	TTTTAAGTGT	TGTTCCATTC
2401	CATCACGGTT	TTGGAATGTT	TACTACACTC	GGATATTTGA	TATGTGGATT	TCGAGTCGTC
2461	TTAATGTATA	GATTTGAAGA	AGAGCTGTTT	TTACGATCCC	TTACGATTTA	CAAAATTCAA
2521	AGTGCCTTGC	TAGTACCAAC	CCTATTTTCA	TTCTTCGCCA	AAAGCACTCT	GATTGACAAA
2581	TACGATTTAT	CTAATTTTACA	CGAAATTGCT	TCTGGGGGCG	CACCTCTTTC	GAAAGAAGTC
2641	GGGGAAGCGG	TTGCAAAACG	CTTCCATCTT	CCAGGGATAC	GACAAGGATA	TGGGCTCACT
2701	GAGACTACAT	CAGCTATTCT	GATTACACCC	GAGGGGGATG	ATAAACCGGG	CGCGGTCCGT
2761	AAAGTTGTTT	CATTTTTTGA	AGCGAAGGTT	GTGGATCTGG	ATACCGGGAA	AACGCTGGGC
2821	GTTAATCAGA	GAGGCGAATT	ATGTGTCAGA	GGACCTATGA	TTATGTCCGG	TTATGTAAAC
2881	AATCCGGAAG	CGACCAACGC	CTTGATTGAC	AAGGATGGAT	GGCTACATTC	TGGAGACATA
2941	GCTTACTGGG	ACGAAGACGA	ACACTTCTTC	ATAGTTGACC	GCTTGAAGTC	TTTAATTAAA
3001	TACAAAGGAT	ATCAGGTGGC	CCCCGTGAAA	TTGGAATCGA	TATTTGTACA	ACACCCCAAC
3061	ATCTTCGACG	CGGGCGTGCG	AGGTCTTCCC	GACGATGACG	CCGGTGAACT	TCCCGCCGCC
3121	GTTGTTGTTT	TGGAGCACGG	AAAGACGATG	ACGGAAGGAG	AGATCGTGGA	TTACGTCGCC
3181	AGTCAAGTAA	CAACCGCGAA	AAAGTTGCGC	GGAGGAGTTG	TGTTTGTGGA	CGAAGTACCG

3241	AAAGGTCTTA	CCGGAAAAC	CGACGCAAGA	AAAATCAGAG	AGATCCTCAT	AAAGGCCAAG
3301	AAGGGCGGAA	AGTCCAAATT	GTAAACACGTG	TAATTCTAGA	GCGGCCGCTT	CGAGCAGACA
3361	TGATAAGATA	CATTGATGAG	TTTGGACAAA	CCACAAC TAG	AATGCAGTGA	AAAAAATGCT
3421	TTATTTGTGA	AATTTGTGAT	GCTATTGCTT	TATTTGTAAC	CATTATAAGC	TGCAATAAAC
3481	AAGTTAACAA	CAACAATTGC	ATTCATTTTA	TGTTTCAGGT	TCAGGGGGAG	GTGTGGGAGG
3541	TTTTTTAAAG	CAAGTAAAAC	CTCTACAAAT	GTGGTAAAAT	CGATAAGTTT	AAACAGATCC
3601	AGGTGGCACT	TTTCGGGGAA	ATGTGCGCGG	AACCCCTATT	TGTTTATTTT	TCTAAATACA
3661	TTCAAATATG	TATCCGCTCA	TGAGACAATA	ACCCTGATAA	ATGCTTCAAT	AATATTGAAA
3721	AAGGAAGAGT	ATGAGTATTC	AACATTTCCG	TGTCGCCCTT	ATTCCCTTTT	TTGCGGCATT
3781	TTGCCTTCCT	GTTTTTGCTC	ACCCAGAAAC	GCTGGTGAAA	GTAAGAGATG	CTGAAGATCA
3841	GTTGGGTGCA	CGAGTGGGTT	ACATCGAACT	GGATCTCAAC	AGCGGTAAGA	TCCTTGAGAG
3901	TTTTCGCCCC	GAAGAACGTT	TTCCAATGAT	GAGCACTTTT	AAAGTTCTGC	TATGTGGCGC
3961	GGTATTATCC	CGTATTGACG	CCGGGCAAGA	GCAACTCGGT	CGCCGCATAC	ACTATTCTCA
4021	GAATGACTTG	GTTGAGTACT	CACCAGTCAC	AGAAAAGCAT	CTTACGGATG	GCATGACAGT
4081	AAGAGAATTA	TGCAGTGCTG	CCATAACCAT	GAGTGATAAC	ACTGCGGCCA	ACTTACTTCT
4141	GACAACATC	GGAGGACCGA	AGGAGCTAAC	CGCTTTTTTG	CACAACATGG	GGGATCATGT
4201	AACTCGCCTT	GATCGTTGGG	AACCGGAGCT	GAATGAAGCC	ATACCAAACG	ACGAGCGTGA
4261	CACCACGATG	CCTGTAGCAA	TGGCAACAAC	GTTGCGCAAA	CTATTAAGTG	GCGAACTACT
4321	TACTCTAGCT	TCCCGGCAAC	AATTAATAGA	CTGGATGGAG	GCGGATAAAG	TTGCAGGACC
4381	ACTTCTGCGC	TCGGCCCTTC	CGGCTGGCTG	GTTTATTGCT	GATAAATCTG	GAGCCGGTGA
4441	GCGTGGGTCT	CGCGGTATCA	TTGCAGCACT	GGGGCCAGAT	GGTAAGCCCT	CCCGTATCGT
4501	AGTTATCTAC	ACGACGGGGA	GTCAGGCAAC	TATGGATGAA	CGAAATAGAC	AGATCGCTGA
4561	GATAGGTGCC	TCACTGATTA	AGCATTGGTA	ACTGTCAGAC	CAAGTTTACT	CATATATACT
4621	TTAGATTGAT	TTAAAACCTC	ATTTTAAATT	TAAAAGGATC	TAGGTGAAGA	TCCTTTTTTGA
4681	TAATCTCATG	ACCAAAATCC	CTTAACGTGA	GTTTTCGTTC	CACTGAGCGT	CAGACCCCGT
4741	AGAAAAGATC	AAAGGATCTT	CTTGAGATCC	TTTTTTTCTG	CGCGTAATCT	GCTGCTTGCA
4801	AACAAAAAAA	CCACCGCTAC	CAGCGGTGGT	TTGTTTGCCG	GATCAAGAGC	TACCAACTCT
4861	TTTTCCGAAG	GTAAGTGGCT	TCAGCAGAGC	GCAGATACCA	AATACTGTCC	TTCTAGTGTA
4921	GCCGTAGTTA	GGCCACCACT	TCAAGAACTC	TGTAGCACCG	CCTACATACC	TCGCTCTGCT
4981	AATCCTGTTA	CCAGTGGCTG	CTGCCAGTGG	CGATAAGTCG	TGTCTTACCG	GGTTGGACTC
5041	AAGACGATAG	TTACCGGATA	AGGCGCAGCG	GTCGGGCTGA	ACGGGGGGTT	CGTGCACACA
5101	GCCCAGCTTG	GAGCGAACGA	CCTACACCGA	ACTGAGATAC	CTACAGCGTG	AGCTATGAGA
5161	AAGCGCCACG	CTTCCCGAAG	GGAGAAAGGC	GGACAGGTAT	CCGGTAAGCG	GCAGGGTCGG
5221	AACAGGAGAG	CGCACGAGGG	AGCTTCCAGG	GGGAAACGCC	TGGTATCTTT	ATAGTCCTGT
5281	CGGGTTTCGC	CACCTCTGAC	TTGAGCGTCG	ATTTTTGTGA	TGCTCGTCAG	GGGGGCGGAG
5341	CCTATGGAAG	AACGCCAGCA	ACGCGGCCTT	TTTACGGTTC	CTGGCCTTTT	GCTGGCCTTT
5401	TGCTCACATG	GCTCGAC				

Legend

- Nucleotides shown in **red** represent the Renilla luciferase region.
- Nucleotides shown in **black bold** represent the inserted sequence of ssRNA-E.
- Nucleotides shown in **green** represent the firefly luciferase region.
- Underlined nucleotides shown in **blue** represent potential binding sites of unmodified PNA P1.

ssRNA-C/p2Luc

5418 bp ds-DNA

1	AGATCGATCT	GCGCAGCACC	ATGGCCTGAA	ATAACCTCTG	AAAGAGGAAC	TTGGTTAGGT
61	ACCTTCTGAG	GCGGAAAGAA	CCAGCTGTGG	AATGTGTGTC	AGTTAGGGTG	TGGAAAGTCC
121	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC	TCAATTAGTC	AGCAACCAGG
181	TGTGGAAAGT	CCCCAGGCTC	CCCAGCAGGC	AGAAGTATGC	AAAGCATGCA	TCTCAATTAG
241	TCAGCAACCA	TAGTCCCGCC	CCTAACTCCG	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC
301	GCCCATTCTC	CGCCCCATGG	CTGACTAATT	TTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC
361	TCGGCCTCTG	AGCTATTCCA	GAAGTAGTGA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC
421	AAAAAGCTTG	ATTCTTCTGA	CACAACAGTC	TCGAACTTAA	GCTGCAGAAG	TTGGTCGTGA
481	GGCACTGGGC	AGGTAAGTAT	CAAGGTTACA	AGACAGGTTT	AAGGAGACCA	ATAGAAATCG
541	GGCTTGTCGA	GACAGAGAAG	ACTCTTGCGT	TTCTGATAGG	CACCTATTGG	TCTTACTGAC
601	ATCCACTTTG	CCTTTCTCTC	CACAGGTGTC	CACTCCAGT	TCAATTACAG	CTCTTAAGGC
661	TAGAGTACTT	AATACGACTC	ACTATAGGCT	AGCCACCATG	ACTTCGAAAG	TTTATGATCC
721	AGAACAAAGG	AAACGGATGA	TAAGTGGTCC	GCAGTGGTGG	GCCAGATGTA	AACAAATGAA
781	TGTTCTTGAT	TCATTTATTA	ATTATTATGA	TTCAGAAAAA	CATGCAGAAA	ATGCTGTTAT
841	TTTTTTACAT	GGTAACGCGG	CCTCTTCTTA	TTTATGGCGA	CATGTTGTGC	CACATATTGA
901	GCCAGTAGCG	CGGTGTATTA	TACCAGACCT	TATTGGTATG	GGCAAAATCAG	GCAAATCTGG
961	TAATGGTTCT	TATAGGTTAC	TTGATCATT	CAAATATCTT	ACTGCATGGT	TTGAACTTCT
1021	TAATTTACCA	AAGAAGATCA	TTTTTGTGCG	CCATGATTGG	GGTGCTTGTT	TGGCATTTCA
1081	TTATAGCTAT	GAGCATCAAG	ATAAGATCAA	AGCAATAGTT	CACGCTGAAA	GTGTAGTAGA
1141	TGTGATTGAA	TCATGGGATG	AATGGCCTGA	TATTGAAGAA	GATATTGCGT	TGATCAAATC
1201	TGAAGAAGGA	GAAAAAATGG	TTTTGGAGAA	TAAGTCTTCT	GTGGAAACCA	TGTTGCCATC
1261	AAAAATCATG	AGAAAGTTAG	AACCAGAAGA	ATTTGCAGCA	TATCTTGAAC	CATTCAAAGA
1321	GAAAGGTGAA	GTTTCGTCGT	CAACATTATC	ATGGCCTCGT	GAAATCCCGT	TAGTAAAAGG
1381	TGGTAAACCT	GACGTTGTAC	AAATTGTTAG	GAATTATAAT	GCTTATCTAC	GTGCAAGTGA
1441	TGATTTACCA	AAAATGTTTA	TTGAATCGGA	CCCAGGATTC	TTTTCCAATG	CTATTGTTGA
1501	AGGTGCCAAG	AAGTTTCCTA	ATACTGAATT	TGTCAAAGTA	AAAGGTCTTC	ATTTTTTCGCA
1561	AGAAGATGCA	CCTGATGAAA	TGGGAAAAAT	TATCAAATCG	TTCGTTGAGC	GAGTTCTCAA
1621	AAATGAACAA	ATGTCGACGG	ATCCCTTCTT	AGGGATAGAG	AGAGAAAGTG	AGCTCGAAGA
1681	GCCCAAAAAC	ATAAAGAAAAG	GCCCGGCGCC	ATTCTATCCT	CTAGAGGATG	GAACCGCTGG
1741	AGAGCAACTG	CATAAGGCTA	TGAAGAGATA	CGCCCTGGTT	CCTGGAACAA	TTGCTTTTAC
1801	AGATGCACAT	ATCGAGGTGA	ACATCACGTA	CGCGGAATAC	TTCGAAATGT	CCGTTTCGGT
1861	GGCAGAAGCT	ATGAAACGAT	ATGGGCTGAA	TACAAATCAC	AGAATCGTCG	TATGCAGTGA
1921	AAACTCTCTT	CAATTCCTTA	TGCCGGTGTT	GGGCGCGTTA	TTTATCGGAG	TTGCAGTTGC
1981	GCCCGCGAAC	GACATTTATA	ATGAACGTGA	ATTGCTCAAC	AGTATGAACA	TTTCGCAGCC
2041	TACCGTAGTG	TTTGTTTCCA	AAAAGGGGTT	GCAAAAAATT	TTGAACGTGC	AAAAAAAATT
2101	ACCAATAATC	CAGAAAATTA	TTATCATGGA	TTCTAAAACG	GATTACCAGG	GATTTCAGTC
2161	AGGTACACG	TTCTGACAT	CTCATCTACC	TCCCGGTTTT	AATGAATACG	ATTTTGTACC
2221	GAGTCCCTTT	GATCGTGACA	AAACAATTGC	ACTGATAATG	AATTCCTCTG	GATCTACTGG
2281	GTTACCTAAG	GGTGTGGCCC	TTCCGCATAG	AACTGCCTGC	GTCAGATTCT	CGCATGCCAG
2341	AGATCCTATT	TTTGGCAATC	AAATCATTCC	GGATACTGCG	ATTTTAAAGT	TTGTTCCATT
2401	CCATCACGGT	TTTGGAAATG	TTACTACACT	CGGATATTTG	ATATGTGGAT	TTTCAGTCGT
2461	CTTAATGTAT	AGATTTGAAG	AAGAGCTGTT	TTTACGATCC	CTTCAGGATT	ACAAAATTCA
2521	AAGTGCGTTG	CTAGTACCAA	CCCTATTTTC	ATTCTTCGCC	AAAAGCACTC	TGATTGACAA
2581	ATACGATTTA	TCTAATTTAC	ACGAAATTGC	TTCTGGGGGC	GCACCTCTTT	CGAAAGAAGT
2641	CGGGGAAGCG	GTTGCAAAAC	GCTTCCATCT	TCCAGGGATA	CGACAAGGAT	ATGGGCTCAC
2701	TGAGACTACA	TCAGCTATTC	TGATTACACC	CGAGGGGGAT	GATAAACCGG	GCGCGGTCGG
2761	TAAAGTTGTT	CCATTTTTTG	AAGCGAAGGT	TGTGGATCTG	GATACCGGGA	AAACGCTGGG
2821	CGTTAATCAG	AGAGGCGAAT	TATGTGTCAG	AGGACCTATG	ATTATGTCCG	GTTATGTAAA
2881	CAATCCGGAA	GCGACCAACG	CCTTGATTGA	CAAGGATGGA	TGGCTACATT	CTGGAGACAT
2941	AGCTTACTGG	GACGAAGACG	AACACTTCTT	CATAGTTGAC	CGCTTGAAGT	CTTTAATTAA
3001	ATACAAAGGA	TATCAGGTGG	CCCCGCTGA	ATTGGAATCG	ATATTGTTAC	AACACCCCAA
3061	CATCTTCGAC	GCGGGCGTGG	CAGGTCTTCC	CGACGATGAC	GCCGGTGAAC	TTCCCGCCGC
3121	CGTTGTTGTT	TTGGAGCACG	GAAAGACGAT	GACGAAAAAA	GAGATCGTGG	ATTACGTCGC
3181	CAGTCAAGTA	ACAACGCGCA	AAAAGTTGCG	CGGAGGAGTT	GTGTTTGTGG	ACGAAGTACC
3241	GAAAGGTCTT	ACCGGAAAAC	TCGACGCAAG	AAAAATCAGA	GAGATCCTCA	TAAAGGCCAA
3301	GAAGGGCGGA	AAGTCCAAAT	TGTAACACGT	GTAATTCTAG	AGCGGCCGCT	TCGAGACAGC
3361	ATGATAAAGAT	ACATTGATGA	GTTTGGACAA	ACCACAATA	GAATCAGATG	AAAAAGATGC
3421	TTTATTGTGT	AAATTTGTGA	TGCTATTGCT	TTATTTGTAA	CCATTATTAAG	CTGCAATAAA
3481	CAAGTTAACA	ACAACAATTG	CATTCATTTT	ATGTTTCAGG	TTCAGGGGGA	GGTGTGGGAG

3541	TTTTTTTAAA	GCAAGTAAAA	CCTCTACAAA	TGTGGTAAAA	TCGATAAGTT	TAAACAGATC
3601	CAGGTGGCAC	TTTTTCGGGA	AATGTGCGCG	GAACCCCTAT	TTGTTTATTT	TTCTAAATAC
3661	ATTCAAATAT	GTATCCGCTC	ATGAGACAAT	AACCCTGATA	AATGCTTCAA	TAATATTGAA
3721	AAAGGAAGAG	TATGAGTATT	CAACATTTCC	GTGTCGCCCT	TATTCCTTTT	TTTGCGGCAT
3781	TTTGCTTCC	TGTTTTTGCT	CACCCAGAAA	CGCTGGTGAA	AGTAAAAGAT	GCTGAAGATC
3841	AGTTGGGTGC	ACGAGTGGGT	TACATCGAAC	TGGATCTCAA	CAGCGGTAAG	ATCCTTGAGA
3901	GTTCGCCCC	CGAAGAACGT	TTTCCAATGA	TGAGCACTTT	TAAAGTTCTG	CTATGTGGCG
3961	CGGTATTATC	CCGTATTGAC	GCCGGGCAAG	AGCAACTCGG	TCGCCGCATA	CACTATTCTC
4021	AGAATGACTT	GGTTGAGTAC	TCACCAGTCA	CAGAAAAGCA	TCTTACGGAT	GGCATGACAG
4081	TAAGAGAATT	ATGCAGTGCT	GCCATAACCA	TGAGTGATAA	CACTGCGGCC	AACTTACTTC
4141	TGACAACGAT	CGGAGGACCG	AAGGAGCTAA	CCGCTTTTTT	GCACAACATG	GGGGATCATG
4201	TAACTCGCCT	TGATCGTTGG	GAACCGGAGC	TGAATGAAGC	CATACCAAAC	GACGAGCGTG
4261	ACACCACGAT	GCCTGTAGCA	ATGGCAACAA	CGTTGCGCAA	ACTATTAACT	GGCGAACTAC
4321	TTACTCTAGC	TTCCCGGCAA	CAATTAATAG	ACTGGATGGA	GGCGGATAAA	GTTGCAGGAC
4381	CACTTCTGCG	CTCGGCCCTT	CCGGCTGGCT	GGTTTATTGC	TGATAAATCT	GGAGCCGGTG
4441	AGCGTGGGTC	TCGCGGTATC	ATTGCAGCAC	TGGGGCCAGA	TGTTAAGCCC	TCCCGTATCG
4501	TAGTTATCTA	CACGACGGGG	AGTCAGGCAA	CTATGGATGA	ACGAAATAGA	CAGATCGCTG
4561	AGATAGGTGC	CTCACTGATT	AAGCATTGGT	AACTGTCAGA	CCAAGTTTAC	TCATATATAC
4621	TTTAGATTGA	TTTAAACTT	CATTTTAAAT	TTAAAAGGAT	CTAGGTGAAG	ATCCTTTTTG
4681	ATAATCTCAT	GACCAAATC	CCTTAACGTG	AGTTTTTCGT	CCACTGAGCG	TCAGACCCCG
4741	TAGAAAAGAT	CAAAGGATCT	TCTTGAGATC	CTTTTTTTCT	GCGCGTAATC	TGCTGCTTGC
4801	AAACAAAAAA	ACCACCGCTA	CCAGCGGTGG	TTTGTTTGCC	GGATCAAGAG	CTACCAACTC
4861	TTTTTCCGAA	GGTAACTGGC	TTCAGCAGAG	CGCAGATACC	AAATACTGTC	CTTCTAGTGT
4921	AGCCGTAGTT	AGGCCACCAC	TTCAAGAACT	CTGTAGCACC	GCCTACATAC	CTCGCTCTGC
4981	TAATCCTGTT	ACCAGTGGCT	GCTGCCAGTG	GCGATAAGTC	GTGTCTTACC	GGGTGGACT
5041	CAAGACGATA	GTTACCGGAT	AAGGCGCAGC	GGTCGGGCTG	AACGGGGGGT	TCGTGCACAC
5101	AGCCCAGCTT	GGAGCGAACG	ACCTACACCG	AACTGAGATA	CCTACAGCGT	GAGCTATGAG
5161	AAAGCGCCAC	GCTTCCCGAA	GGGAGAAAGG	CGGACAGGTA	TCCGGTAAGC	GGCAGGGTCG
5221	GAACAGGAGA	GCGCACGAGG	GAGCTTCCAG	GGGGAAACGC	CTGGTATCTT	TATAGTCCTG
5281	TCGGGTTTCG	CCACCTCTGA	CTTGAGCGTC	GATTTTTGTG	ATGCTCGTCA	GGGGGGCGGA
5341	GCCTATGGAA	AAACGCCAGC	AACGCGGCCT	TTTTACGGTT	CCTGGCCTTT	TGCTGGCCTT
5401	TTGCTCACAT	GGCTCGAC				

Legend

- Nucleotides shown in **red** represent the Renilla luciferase region.
- Nucleotides shown in **black bold** represent the inserted sequence of ssRNA-C.
- Nucleotides shown in **purple bold** represent the mutated slippery sequence.
- Nucleotides shown in **green** represent the firefly luciferase region.
- Underlined nucleotides shown in **blue** represent potential binding sites of unmodified PNA P1.

rHP1-E/p2Luc

5435 bp ds-DNA

1	AGATCGATCT	GCGCAGCACC	ATGGCCTGAA	ATAACCTCTG	AAAGAGGAAC	TTGGTTAGGT
61	ACCTTCTGAG	GCGGAAAGAA	CCAGCTGTGG	AATGTGTGTC	AGTTAGGGTG	TGGAAAGTCC
121	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC	TCAATTAGTC	AGCAACCAGG
181	TGTGGAAAGT	CCCCAGGCTC	CCCAGCAGGC	AGAAGTATGC	AAAGCATGCA	TCTCAATTAG
241	TCAGCAACCA	TAGTCCCGCC	CCTAACTCCG	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC
301	CCCCATTCTC	CGCCCCATGG	CTGACTAATT	TTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC
361	TCGGCCTCTG	AGCTATTCCA	GAAAGTAGTA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC
421	AAAAAGCTTG	ATTCTTCTGA	CACAACAGTC	TCGAACTTAA	GCTGCAGAAG	TTGGTCGTGA
481	GGCACTGGGC	AGGTAAGTAT	CAAGGTTACA	AGACAGGTTT	AAGGAGACCA	ATAGAAACTG
541	GGCTTGTCGA	GACAGAGAAG	ACTCTTGCGT	TTCTGATAGG	CACCTATTGG	TCTTACTGAC
601	ATCCACTTTG	CCTTTCTCTC	CACAGGTGTC	CACTCCAGT	TCAATTACAG	CTCTTAAGGC
661	TAGAGTACTT	AATACGACTC	ACTATAGGCT	AGCCACCATG	ACTTCGAAAG	TTTATGATCC
721	AGAACAAAGG	AAACGGATGA	TAAGTGGTCC	GCAGTGGTGG	GCCAGATGTA	AACAAATGAA
781	TGTTCTTGAT	TCATTTATTA	ATTATTATGA	TTCAGAAAAA	CATGCAGAAA	ATGCTGTTAT
841	TTTTTTACAT	GGTAACGCGG	CCTCTTCTTA	TTTATGGCGA	CATGTTGTGC	CACATATTGA
901	GCCAGTAGCG	CGGTGTATTA	TACCAGACCT	TATTGGTATG	GGCAAATCAG	GCAAACTTGG
961	TAATGGTTCT	TATAGGTTAC	TTGATCATT	CAAATATCTT	ACTGCATGGT	TTGAACTTCT
1021	TAATTTACCA	AAGAAGATCA	TTTTTGTCGG	CCATGATTGG	GGTGCTTGTT	TGGCATTTC
1081	TTATAGCTAT	GAGCATCAAG	ATAAGATCAA	AGCAATAGTT	CACGCTGAAA	GTGTAGTAGA
1141	TGTGATTGAA	TCATGGGATG	AATGGCCTGA	TATTGAAGAA	GATATTGCGT	TGATCAAATC
1201	TGAAGAAGGA	GAAAAAATGG	TTTTGGAGAA	TAAGTCTTTC	GTGGAACCA	TGTTGCCATC
1261	AAAAATCATG	AGAAAGTTAG	AACCAGAAGA	ATTTGCAGCA	TATCTTGAAC	CATTCAAAGA
1321	GAAAGGTGAA	GTTCGTCGTC	CAACATTATC	ATGGCCTCGT	GAAATCCCGT	TAGTAAAAGG
1381	TGGTAAACCT	GACGTTGTAC	AAATTGTTAG	GAATTATAAT	GCTTATCTAC	GTGCAAGTGA
1441	TGATTTACCA	AAAATGTTTA	TTGAATCGGA	CCCAGGATTC	TTTCCAATG	CTATTGTTGA
1501	AGGTGCCAAG	AAGTTTCCTA	ATACTGAATT	TGTCAAAGTA	AAAGGTCTTC	ATTTTTCGCA
1561	AGAAGATGCA	CCTGATGAAA	TGGGAAAATA	TATCAAATCG	TTCGTTGAGC	GAGTTCTCAA
1621	AAATGAACAA	ATGTCGACGG	ATCCTTTTTT	AGGGTAGAGA	GAGAAAAGTT	CGACTTTCTC
1681	TCTCTAGAGC	TCGAAGACGC	CAAAAACATA	AAGAAAGGCC	CGGCGCCATT	CTATCCTCTA
1741	GAGGATGGAA	CCGCTGGAGA	GCAACTGCAT	AAGGCTATGA	AGAGATACGC	CCTGGTTCCT
1801	GGAAACAATTG	CTTTTACAGA	TGCACATATC	GAGGTGAACA	TCACGTACGC	GGAATACTTC
1861	GAAATGTCCG	TTTCGTTGGC	AGAAGCTATG	AAACGATATG	GGCTGAATAC	AAATCACAGA
1921	ATCGTCGTAT	GCAGTGAAAA	CTCTCTTCAA	TTCTTTATGC	CGGTGTTGGG	CGCGTTATTT
1981	ATCGGAGTTG	CAGTTGCGCC	CGCGAACGAC	ATTTATAATG	AACGTGAATT	GCTCAACAGT
2041	ATGAACATTT	CGCAGCCTAC	CGTAGTGTTT	GTTTCCAAAA	AGGGGTTGCA	AAAAATTTTG
2101	AACGTGCAAA	AAAAATTACC	AATAATCCAG	AAAATTATTA	TCATGGATTG	TAAAACGGAT
2161	TACCAGGGAT	TTTTCAGTCG	GTACACGTTT	GTCACATCTC	ATCTACCTCC	CGGTTTTAAT
2221	GAATACGATT	TTGTACCAGA	GTCCTTTGAT	CGTGACAAAA	CAATTGCACT	GATAATGAAT
2281	TCCTCTGGAT	CTACTGGGTT	ACCTAAGGGT	GTGGCCCTTC	CGCATAGAAC	TGCCTGCGTC
2341	AGATTCTCGC	ATGCCAGAGA	TCCTATTTTT	GGCAATCAAA	TCATTCCGGA	TACTGCGATT
2401	TTAAGTGTTG	TTCCATTCCA	TCACGGTTTT	GGAATGTTTA	CTACACTCGG	ATATTTGATA
2461	TGTGGATTTT	GAGTCGTCTT	AATGTATAGA	TTTGAAGAAG	AGCTGTTTTT	ACGATCCCTT
2521	CAGGATTACA	AAATTCAAAG	TGCGTTGCTA	GTACCAACCC	TATTTTCATT	CTTCGCCAAA
2581	AGCACTCTGA	TTGACAAATA	CGATTTCGTT	AATTTACACG	AAATTGCTTC	TGGGGGCGCA
2641	CCTCTTTTCG	AAGAAGTCGG	GGAAGCGGTT	GCAAAACGCT	TCCATCTTCC	AGGGATACGA
2701	CAAGGATATG	GGCTCACTGA	GACTACATCA	GCTATTCTGA	TTACACCCGA	GGGGGATGAT
2761	AAACCGGGCG	CGGTGCGTAA	AGTTGTTCCA	TTTTTTGAAG	CGAAGGTTGT	GGATCTGGAT
2821	ACCGGGAAAA	CGCTGGGCGT	TAATCAGAGA	GGCGAATTAT	GTGTCAGAGG	ACCTATGATT
2881	ATGTCCGGTT	ATGTAAACAA	TCCGGAAGCG	ACCAACGCCT	TGATTGACAA	GGATGGATGG
2941	CTACATTCTG	GAGACATAGC	TTACTGGGAC	GAAGACGAAC	ACTTCTTCAT	AGTTGACCGC
3001	TTGAAGTCTT	TAATTAAATA	CAAAGGATAT	CAGGTGGCCC	CCGCTGAATT	GGAATCGATA
3061	TTGTTACAAC	ACCCCAACAT	CTTCGACGCG	GGCGTGGCAG	GTCTTCCCGA	CGATGACGCC
3121	GGTGAACCTT	CCGCCGCCGT	TGTTGTTTTG	GAGCACGGAA	AGACGATGAC	GGAAAAAGAG
3181	ATCGTGGATT	ACGTGCGCCG	TCAAGTAACA	ACCGCGAAAA	AGTTGCGCGG	AGGAGTTGTG
3241	TTTGTGGACG	AAGTACCGAA	AGGTCTTACC	GGAAAACTCG	ACGCAAGAAA	AATCAGAGAG
3301	ATCTCTCATAA	AGGCCAAGAA	GGGCGGAAAG	TCCAAATTGT	AACACGTGTA	ATTCTAGAGC
3361	GGCCGCTTCG	AGCAGACATG	ATAAGATACA	TTGATGAGTT	TGGACAAACC	ACAAC TAGAA
3421	TGCAGTGAAA	AAAATGCTTT	ATTTGTGAAA	TTTGTGATGC	TATTGCTTTA	TTTGTAAACCA
3481	TTATAAGCTG	CAATAAACAA	GTTAAACAACA	ACAATTGCAT	TCATTTTATG	TTTCAGGTTT
3541	AGGGGGAGGT	GTGGGAGGTT	TTTTAAAGCA	AGTAAAACCT	CTACAAATGT	GGTAAATTCG

3601	ATAAGTTTAA	ACAGATCCAG	GTGGCACTTT	TCGGGGAAAT	GTGCGCGGAA	CCCCTATTTG
3661	TTTATTTTTTC	TAAATACATT	CAAATATGTA	TCCGCTCATG	AGACAATAAC	CCTGATAAAT
3721	GCTTCAATAA	TATTGAAAAA	GGAAGAGTAT	GAGTATTCAA	CATTTCCGTG	TCGCCCTTAT
3781	TCCCTTTTTTT	GCGGCATTTT	GCCTTCCTGT	TTTTGCTCAC	CCAGAAACGC	TGGTGAAAGT
3841	AAAAGATGCT	GAAGATCAGT	TGGGTGCACG	AGTGGGTAC	ATCGAACTGG	ATCTCAACAG
3901	CGGTAAGATC	CTTGAGAGTT	TTCGCCCCGA	AGAACGTTTT	CCAATGATGA	GCACTTTTAA
3961	AGTTCTGCTA	TGTGGCGCGG	TATTATCCCG	TATTGACGCC	GGGCAAGAGC	AACTCGGTCG
4021	CCGCATACAC	TATTCTCAGA	ATGACTTGGT	TGAGTACTCA	CCAGTCACAG	AAAAGCATCT
4081	TACGGATGGC	ATGACAGTAA	GAGAATTATG	CAGTGCTGCC	ATAACCATGA	GTGATAACAC
4141	TGCGGCCAAC	TTACTTCTGA	CAACGATCGG	AGGACCGAAG	GAGCTAACCG	CTTTTTTGCA
4201	CAACATGGGG	GATCATGTAA	CTCGCCTTGA	TCGTTGGGAA	CCGGAGCTGA	ATGAAGCCAT
4261	ACCAAACGAC	GAGCGTGACA	CCACGATGCC	TGTAGCAATG	GCAACAACGT	TGCGCAAACCT
4321	ATTAACCTGGC	GAACACTTTA	CTCTAGCTTC	CCGGCAACAA	TTAATAGACT	GGATGGAGGC
4381	GGATAAAGTT	GCAGGACCAC	TTCTGCGCTC	GGCCCTTCCG	GCTGGCTGGT	TTATTGCTGA
4441	TAAATCTGGA	GCCGGTGAGC	GTGGGTCTCG	CGGTATCATT	GCAGCACTGG	GGCCAGATGG
4501	TAAGCCCTCC	CGTATCGTAG	TTATCTACAC	GACGGGGAGT	CAGGCAACTA	TGGATGAACG
4561	AAATAGACAG	ATCGCTGAGA	TAGGTGCCTC	ACTGATTAAG	CATTGGTAAC	TGTCAGACCA
4621	AGTTTACTCA	TATATACTTT	AGATTGATTT	AAAACCTCAT	TTTTAATTTA	AAAGGATCTA
4681	GGTGAAGATC	CTTTTTGATA	ATCTCATGAC	CAAAATCCCT	TAACGTGAGT	TTTCGTTCCA
4741	CTGAGCGTCA	GACCCCGTAG	AAAAGATCAA	AGGATCTTCT	TGAGATCCTT	TTTTTCTGCG
4801	CGTAATCTGC	TGCTTGCAAA	CAAAAAAACC	ACCGCTACCA	GCGGTGGTTT	GTTTGCCGGA
4861	TCAAGAGCTA	CCAACCTCTT	TTCCGAAGGT	AACCTGGCTC	AGCAGAGCGC	AGATACCAAA
4921	TACTGTCTCT	CTAGTGTAGC	CGTAGTTAGG	CCACCACTTC	AAGAACTCTG	TAGCACCGCC
4981	TACATACCTC	GCTCTGCTAA	TCCTGTTACC	AGTGGCTGCT	GCCAGTGGCG	ATAAGTCGTG
5041	TCTTACCGGG	TTGGACTCAA	GACGATAGTT	ACCGGATAAG	GCGCAGCGGT	CGGGCTGAAC
5101	GGGGGGTTTCG	TGCACACAGC	CCAGCTTGGA	GCGAACGACC	TACACCGAAC	TGAGATACCT
5161	ACAGCGTGAG	CTATGAGAAA	GCGCCACGCT	TCCCGAAGGG	AGAAAGGCGG	ACAGGTATCC
5221	GGTAAGCGGC	AGGGTCGGAA	CAGGAGAGCG	CACGAGGGAG	CTTCCAGGGG	GAAACGCCTG
5281	GTATCTTTAT	AGTCCTGTCG	GGTTTCGCCA	CCTCTGACTT	GAGCGTCGAT	TTTTGTGATG
5341	CTCGTCAGGG	GGGCGGAGCC	TATGGAAAAA	CGCCAGCAAC	GCGGCCTTTT	TACGGTTCCT
5401	GGCCTTTTGC	TGGCCTTTTG	CTCACATGGC	TCGAC		

Legend

- Nucleotides shown in **red** represent the Renilla luciferase region.
- Nucleotides shown in **black bold** represent the inserted sequence of rHP1-E.
- Nucleotides shown in **green** represent the firefly luciferase region.
- Underlined nucleotides shown in **blue** represent potential binding sites of unmodified PNA P1.

rHP1-C/p2Luc

5436 bp ds-DNA

1	AGATCGATCT	GCGCAGCACC	ATGGCCTGAA	ATAACCTCTG	AAAGAGGAAC	TTGGTTAGGT
61	ACCTTCTGAG	GCGGAAAGAA	CCAGCTGTGG	AATGTGTGTC	AGTTAGGGTG	TGGAAAGTCC
121	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC	TCAATTAGTC	AGCAACCAGG
181	TGTGGAAAGT	CCCCAGGCTC	CCCAGCAGGC	AGAAGTATGC	AAAGCATGCA	TCTCAATTAG
241	TCAGCAACCA	TAGTCCCGCC	CCTAACTCCG	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC
301	GCCCATTCTC	CGCCCCATGG	CTGACTAATT	TTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC
361	TCGGCCTCTG	AGCTATTCCA	GAAAGTAGTA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC
421	AAAAAGCTTG	ATTCTTCTGA	CACAACAGTC	TCGAACTTAA	GCTGCAGAA	TTGGTCGTGA
481	GGCACTGGGC	AGGTAAGTAT	CAAGGTTACA	AGACAGGTTT	AAGGAGACCA	ATAGAAATCG
541	GGCTTGTCGA	GACAGAGAAG	ACTCTTGCGT	TTCTGATAGG	CACCTATTGG	TCTTACTGAC
601	ATCCACTTTG	CCTTTCTCTC	CACAGGTGTC	CACTCCAGT	TCAATTACAG	CTCTTAAGGC
661	TAGAGTACTT	AATACGACTC	ACTATAGGCT	AGCCACCATG	ACTTCGAAAG	TTTATGATCC
721	AGAACAAAGG	AAACGGATGA	TAAGTGGTCC	GCAGTGGTGG	GCCAGATGTA	AACAAATGAA
781	TGTTCTTGAT	TCATTTATTA	ATTATTATGA	TTCAGAAAAA	CATGCAGAAA	ATGCTGTTAT
841	TTTTTTACAT	GGTAACGCGG	CCTCTTCTTA	TTTATGGCGA	CATGTTGTGC	CACATATTGA
901	GCCAGTAGCG	CGGTGTATTA	TACCAGACCT	TATTGGTATG	GGCAAAATCAG	GCAAATCTGG
961	TAATGGTTCT	TATAGGTTAC	TTGATCATT	CAAATATCTT	ACTGCATGGT	TTGAACTTCT
1021	TAATTTACCA	AAGAAGATCA	TTTTTGTGCG	CCATGATTGG	GGTGCTTGTT	TGGCAATTTCA
1081	TTATAGCTAT	GAGCATCAAG	ATAAGATCAA	AGCAATAGTT	CACGCTGAAA	GTGTAGTAGA
1141	TGTGATTGAA	TCATGGGATG	AATGGCCTGA	TATTGAAGAA	GATATTGCGT	TGATCAAATC
1201	TGAAGAAGGA	GAAAAAATGG	TTTTGGAGAA	TAAGTCTTTC	GTGGAAACCA	TGTTGCCATC
1261	AAAAATCATG	AGAAAGTTAG	AACCAGAAGA	ATTTGCAGCA	TATCTTGAAC	CATTCAAAGA
1321	GAAAGGTGAA	GTTTCGTCGC	CAACATTATC	ATGGCCTCGT	GAAATCCCGT	TAGTAAAAGG
1381	TGGTAAACCT	GACGTTGTAC	AAATTGTTAG	GAATTATAAT	GCTTATCTAC	GTGCAAGTGA
1441	TGATTTACCA	AAAATGTTTA	TTGAATCGGA	CCCAGGATTC	TTTTCCAATG	CTATTGTTGA
1501	AGGTGCCAAG	AAGTTTCCTA	ATACTGAATT	TGTCAAAGTA	AAAGGTCTTC	ATTTTTTCGCA
1561	AGAAGATGCA	CCTGATGAAA	TGGGAAAAAT	TATCAAATCG	TTGCTTGAGC	TAGTCTCTCA
1621	AAATGAACAA	ATGTCGACGG	ATCCACTTCT	TAGGGTAGAG	AGAGAAAGTT	TGCAATTTCT
1681	CTCTCTAGAG	CTCGAAGACG	CCAAAAACAT	AAAGAAAGGC	CCGCGCCCAT	TCTATCTCTT
1741	AGAGGATGGA	ACCGCTGGAG	AGCAACTGCA	TAAGGCTATG	AAGAGATACG	CCCTGGTTCC
1801	TGGAACAATT	GCTTTTACAG	ATGCACATAT	CGAGGTGAAC	ATCACGTACG	CGGAATACTT
1861	CGAAATGTCC	GTTTCGTTGG	CAGAAGCTAT	GAAACGATAT	GGGCTGAATA	CAAATCACAG
1921	AATCGTCGTA	TGCAGTGAAA	ACTCTCTTCA	ATTCTTTATG	CCGGTGTGG	GCGCGTTATT
1981	TATCGGAGTT	GCAGTTGCGC	CCGCGAACGA	CATTTATAAT	GAACGTGAAT	TGCTCAACAG
2041	TATGAACATT	TCGCAGCCTA	CCGTAGTGTT	TGTTTCCAAA	AAGGGGTTGC	AAAAAATTTT
2101	GAACGTGCAA	AAAAAATTAC	CAATAATCCA	GAAAATTATT	ATCATGGATT	CTAAAACGGA
2161	TTACCAAGGA	TTTCAGTCGA	TGTACACGTT	CGTCACATCT	CATCTACCTC	CCGGTTTTAA
2221	TGAATACGAT	TTTGTACCAG	AGTCCTTTGA	TCTGTGACAA	ACAATTGCAC	TGATCATGAA
2281	TTCTCTGGA	TCTACTGGGT	TACCTAAGGG	TGTGGCCCTT	CCGCATAGAA	CTGCCTGCGT
2341	CAGATTCTCG	CATGCCAGAG	ATCCTATTTT	TGGCAATCAA	ATCATTCCGG	ATACTGCGAT
2401	TTTAAGTGTT	GTTCCATTCC	ATCACGGTTT	TGGAATGTTT	ACTACACTCG	GATATTTGAT
2461	ATGTGGATTT	CGAGTCGTCT	TAATGTATAG	ATTTGAAGAA	GAGCTGTTTT	TACGATCCCT
2521	TCAGGATTAC	AAAATTCAAA	GTGCGTTGCT	AGTACCAACC	CTATTTTCAT	TCTTCGCCAA
2581	AAGCACTCTG	ATTGACAAAT	ACGATTTATC	TAATTTACAC	GAAATTGCTT	CTGGGGGCGC
2641	ACCTCTTTTC	AAAGAAGTCG	GGGAAGCGGT	TGCAAAACGC	TTCCATCTTC	CAGGGATACG
2701	ACAAGGATAT	GGGCTCACTG	AGACTACATC	AGCTATTCTG	ATTACACCCG	AGGGGGATGA
2761	TAAACCGGGC	GCGGTCGGTA	AAGTTGTTCC	ATTTTTTGAA	GCGAAGGTTG	TGGATCTGGA
2821	TACCGGGAAA	ACGCTGGGCG	TTAATCAGAG	AGGCGAATTA	TGTGTCAGAG	GACCTATGAT
2881	TATGTCCGGT	TATGTAAACA	ATCCGGAAGC	GACCAACGCC	TTGATTGACA	AGGATGGATG
2941	GCTACATTCT	GGAGACATAG	CTTACTGGGA	CGAAGACGAA	CACCTCTTCA	TAGTTGACCG
3001	CTTGAAGTCT	TTAATTAAAT	ACAAAGGATA	TCAGGTGGCC	CCCGCTGAAT	TGGAATCGAT
3061	ATTGTTACAA	CACCCCAACA	TCTTCGACGC	GGGCGTGGCA	GGTCTTCCCG	ACGATGACGC
3121	CGGTGAACCT	CCCGCCGCCG	TTGTTGTTTT	GGAGCACGGA	AAGACGATGA	CGGAAAAAGA
3181	GATCGTGGAT	TACGTCGCCA	GTCAGTAAC	AACCGCGAAA	AAGTTGCGCG	GAGGAGTTGT
3241	GTTTGTGGAC	GAAGTACCGA	AAGGTCTTAC	CGGAAAACCT	GACGCAAGAA	AAATCAGAGA
3301	GATCCCTCAT	AAGGCCAAGA	AGGGCGGAAA	GTCCAAATTG	TAACACGTGT	AATTCTAGAG
3361	CGGCCGCTTC	GAGCAGACAT	GATAAGATAC	ATTGATGAGT	TTGGACAAAC	CACAACATGA
3421	ATGCAGTGAA	AAAAATGCTT	TATTTGTGAA	ATTTGTGATG	CTATTGCTTT	ATTTGTAACC
3481	ATTATAAGCT	GCAATAAACA	AGTTAACAAC	AACAATTGCA	TTCATTTTAT	GTTTCAGGTT

3541	CAGGGGGAGG	TGTGGGAGGT	TTTTTAAAGC	AAGTAAAACC	TCTACAAATG	TGGTAAAATC
3601	GATAAGTTTA	AACAGATCCA	GGTGGCACTT	TTCGGGGAAA	TGTGCGCGGA	ACCCCTATTT
3661	GTTTATTTTT	CTAAATACAT	TCAAATATGT	ATCCGCTCAT	GAGACAATAA	CCCTGATAAA
3721	TGCTTCAATA	ATATTGAAAA	AGGAAGAGTA	TGAGTATTCA	ACATTTCCGT	GTCGCCCTTA
3781	TTCCCTTTTT	TGCGGCATTT	TGCCTTCCTG	TTTTTGCTCA	CCCAGAAACG	CTGGTGAAAG
3841	TAAAAGATGC	TGAAGATCAG	TTGGGTGCAC	GAGTGGGTTA	CATCGAACTG	GATCTCAACA
3901	GCGGTAAGAT	CCTTGAGAGT	TTTCGCCCCG	AAGAACGTTT	TCCAATGATG	AGCACTTTTA
3961	AAGTTCTGCT	ATGTGGCGCG	GTATTATCCC	GTATTGACGC	CGGGCAAGAG	CAACTCGGTC
4021	GCCGCATACA	CTATTCTCAG	AATGACTTGG	TTGAGTACTC	ACCAGTCACA	GAAAAGCATC
4081	TTACGGATGG	CATGACAGTA	AGAGAATTAT	GCAGTGCTGC	CATAACCATG	AGTGATAACA
4141	CTGCGGCCAA	CTTACTTCTG	ACAACGATCG	GAGGACCGAA	GGAGCTAACC	GCTTTTTTGC
4201	ACAACATGGG	GGATCATGTA	ACTCGCCTTG	ATCGTTGGGA	ACCGGAGCTG	AATGAAGCCA
4261	TACCAAACGA	CGAGCGTGAC	ACCACGATGC	CTGTAGCAAT	GGCAACAACG	TTGCGCAAAC
4321	TATTAAGTGG	CGAACTACTT	ACTCTAGCTT	CCCGGCAACA	ATTAATAGAC	TGGATGGAGG
4381	CGGATAAAGT	TGCAGGACCA	CTTCTGCGCT	CGGCCCTTCC	GGCTGGCTGG	TTTATTGCTG
4441	ATAAATCTGG	AGCCGGTGAG	CGTGGGTCTC	GCGGTATCAT	TGCAGCACTG	GGGCCAGATG
4501	GTAAGCCCTC	CCGTATCGTA	GTTATCTACA	CGACGGGGAG	TCAGGCAACT	ATGGATGAAC
4561	GAAATAGACA	GATCGCTGAG	ATAGGTGCCT	CACTGATTAA	GCATTGGTAA	CTGTCAGACC
4621	AAGTTTACTC	ATATATACTT	TAGATTGATT	TAAAACTTCA	TTTTTAATTT	AAAAGGATCT
4681	AGGTGAAGAT	CCTTTTTGAT	AATCTCATGA	CCAAAATCCC	TTAACGTGAG	TTTTTCGTCC
4741	ACTGAGCGTC	AGACCCCGTA	GAAAAGATCA	AAGGATCTTC	TTGAGATCCT	TTTTTTCTGC
4801	GCGTAATCTG	CTGCTTGCAA	ACAAAAAAAC	CACCGCTACC	AGCGGTGGTT	TGTTTGCCGG
4861	ATCAAGAGCT	ACCAACTCTT	TTTCCGAAGG	TAAGTGGCTT	CAGCAGAGCG	CAGATACCAA
4921	ATACTGTCCT	TCTAGTGTAG	CCGTAGTTAG	GCCACCACTT	CAAGAACTCT	GTAGCACC GC
4981	CTACATACCT	CGCTCTGCTA	ATCCTGTTAC	CAGTGGCTGC	TGCCAGTGCC	GATAAGTCGT
5041	GTCTTACCGG	GTTGGACTCA	AGACGATAGT	TACCGGATAA	GGCGCAGCGG	TCGGGCTGAA
5101	CGGGGGGTTC	GTGCACACAG	CCCAGCTTGG	AGCGAACGAC	CTACACCGAA	CTGAGATACC
5161	TACAGCGTGA	GCTATGAGAA	AGCGCCACGC	TTCCCGAAGG	GAGAAAGGCG	GACAGGTATC
5221	CGGTAAGCGG	CAGGGTCGGA	ACAGGAGAGC	GCACGAGGGA	GCTTCCAGGG	GGAAACGCCT
5281	GGTATCTTTA	TAGTCCTGTC	GGGTTTCGCC	ACCTCTGACT	TGAGCGTCGA	TTTTTTGTGAT
5341	GCTCGTCAGG	GGGGCGGAGC	CTATGGAAAA	ACGCCAGCAA	CGCGGCCTTT	TTACGGTTCC
5401	TGGCCTTTTTG	CTGGCCTTTT	GCTCATATGG	CTCGAC		

Legend

- Nucleotides shown in **red** represent the Renilla luciferase region.
- Nucleotides shown in **black bold** represent the inserted sequence of rHP1-C.
- Nucleotides shown in **purple bold** represent the mutated slippery sequence.
- Nucleotides shown in **green** represent the firefly luciferase region.
- Underlined nucleotides shown in **blue** represent potential binding sites of unmodified PNA P1.

rHP2-E/p2Luc

5435 bp ds-DNA

1	AGATCGATCT	GCGCAGCACC	ATGGCCTGAA	ATAACCTCTG	AAAGAGGAAC	TTGGTTAGGT
61	ACCTTCTGAG	GCGGAAAGAA	CCAGCTGTGG	AATGTGTGTC	AGTTAGGGTG	TGGAAAGTCC
121	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC	TCAATTAGTC	AGCAACCAGG
181	TGTGGAAAGT	CCCCAGGCTC	CCCAGCAGGC	AGAAGTATGC	AAAGCATGCA	TCTCAATTAG
241	TCAGCAACCA	TAGTCCCGCC	CCTAACTCCG	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC
301	GCCCATTCTC	CGCCCCATGG	CTGACTAATT	TTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC
361	TCGGCCTCTG	AGCTATTCCA	GAAGTAGTGA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC
421	AAAAAGCTTG	ATTCTTCTGA	CACAACAGTC	TCGAACTTAA	GCTGCAGAAG	TTGGTCGTGA
481	GGCACTGGGC	AGGTAAGTAT	CAAGGTTACA	AGACAGGTTT	AAGGAGACCA	ATAGAAACTG
541	GGCTTGTCGA	GACAGAGAAG	ACTCTTGCGT	TTCTGATAGG	CACCTATTGG	TCTTACTGAC
601	ATCCACTTTG	CCTTTCTCTC	CACAGGTGTC	CACTCCAGT	TCAATTACAG	CTCTTAAGGC
661	TAGAGTACTT	AATACGACTC	ACTATAGGCT	AGCCACCATG	ACTTCGAAAG	TTTATGATCC
721	AGAACAAAGG	AAACGGATGA	TAAGTGGTCC	GCAGTGGTGG	GCCAGATGTA	AACAAATGAA
781	TGTTCTTGAT	TCATTTATTA	ATTATTATGA	TTCAGAAAAA	CATGCAGAAA	ATGCTGTTAT
841	TTTTTTACAT	GGTAACGCGG	CCTCTTCTTA	TTTATGGCGA	CATGTTGTGC	CACATATTGA
901	GCCAGTAGCG	CGGTGTATTA	TACCAGACCT	TATTGGTATG	GGCAAAATCAG	GCAAATCTGG
961	TAATGGTTCT	TATAGGTTAC	TTGATCATT	CAAATATCTT	ACTGCATGGT	TTGAACTTCT
1021	TAATTTACCA	AAGAAGATCA	TTTTTGTGCG	CCATGATTGG	GGTGCTTGTT	TGGCAATTTCA
1081	TTATGCTAT	GAGCATCAAG	ATAAGATCAA	AGCAATAGTT	CACGCTGAAA	GTGTAGTAGA
1141	TGTGATTGAA	TCATGGGATG	AATGGCCTGA	TATTGAAGAA	GATATTGCGT	TGATCAAATC
1201	TGAAGAAGGA	GAAAAAATGG	TTTTGGAGAA	TAAGTTCTTC	GTGGAAACCA	TGTTGCCATC
1261	AAAAATCATG	AGAAAGTTAG	AACCAGAAGA	ATTTGCAGCA	TATCTTGAAC	CATTCAAAGA
1321	GAAAGGTGAA	GTTTCGTCGC	CAACATTATC	ATGGCCTCGT	GAAATCCCCT	TAGTAAAAGG
1381	TGGTAAACCT	GACGTTGTAC	AAATTGTTAG	GAATTATAAT	GCTTATCTAC	GTGCAAGTGA
1441	TGATTTACCA	AAAATGTTTA	TTGAATCGGA	CCCAGGATTC	TTTTCCAATG	CTATTGTTGA
1501	AGGTGCCAAG	AAGTTTCCTA	ATACTGAATT	TGTCAAAGTA	AAAGGTCTTC	ATTTTTTCGCA
1561	AGAAGATGCA	CCTGATGAAA	TGGGAAAATA	TATCAAATCG	TTGTTGAGC	GAGTTCTCAA
1621	AAATGAACAA	ATGTCGACGG	ATCCTTTTTT	AGGGTAGAGA	GACAAAGTTT	CGACTTTGTC
1681	TCTCTAGAGC	TCGAAGACGC	CAAAAAACATA	AAGAAAGGCC	CGCGCCCAT	CTATCCTCTA
1741	GAGGATGGAA	CCGCTGGAGA	GCAACTGCAT	AAGGCTATGA	AGAGATACGC	CCTGGTTCCCT
1801	GGAACAATTG	CTTTTACAGA	TGCACATATC	GAGGTGAACA	TCACGTACGC	GGAATACTTC
1861	GAAATGTCCG	TTTCGGTTGG	AGAAGCTATG	AAACGATATG	GGCTGAATAC	AAATCACAGA
1921	ATCGTCGTAT	GCAGTGAAAA	CTCTCTTCAA	TTCTTTATGC	CGGTGTTGGG	CGCGTTATTT
1981	ATCGGAGTTG	CAGTTGCGCC	CGCGAACGAC	ATTTATAATG	AACGTGAATT	GCTCAACAGT
2041	ATGAACATTT	CGCAGCCTAC	CGTAGTGTTT	GTTTCCAAAA	AGGGGTTGCA	AAAAATTTTG
2101	AACGTGCAAA	AAAAATTACC	AATAATCCAG	AAAATTATTA	TCATGGATT	TAAAACGGAT
2161	TACACGGGAT	TTTTCAGTCG	GTACACGTTT	GTCACATCTC	ATCTACCTCC	CGGTTTTAAT
2221	GAAATGAGAT	TTGTACCAGA	GTCCTTTGAT	CGTGACAAAA	CAATTGCAAT	GATAATGAAT
2281	TCCTCTGGAT	CTACTGGGTT	ACCTAAGGGT	GTGGCCCTTC	CGCATAGAAC	TGCCTGCGTC
2341	AGATTCTCGC	ATGCCAGAGA	TCCTATTTTT	GGCAATCAAA	TCATTCCGGA	TACTGCGATT
2401	TTAAGTGTTG	TTCCATTCCA	TCACGGTTTT	GGAATGTTTA	CTACACTCGG	ATATTTGATA
2461	TGTGGATTTT	GAGTCGTCTT	AATGTATAGA	TTTGAAGAAG	AGCTGTTTTT	ACGATCCCTT
2521	CAGGATTACA	AAATTCAAAG	TGCGTTGCTA	GTACCAACCC	TATTTTCATT	CTTCGCCAAA
2581	AGCACTCTGA	TTGACAAATA	CGATTTATCT	AATTTACACG	AAATTGCTTC	TGGGGGCGCA
2641	CCTCTTTTGA	AAGAAGTCGG	GGAAGCGGTT	GCAAAACGCT	TCCATCTTCC	AGGGATACGA
2701	CAAGGATATG	GGCTCACTGA	GACTACATCA	GCTATTCTGA	TTACACCCGA	GGGGGATGAT
2761	AAACCGGGCG	CGGTCGGTAA	AGTTGTTTCA	TTTTTTGAAG	CGAAGGTTTG	GGATCTGGAT
2821	ACCGGGAAAA	CGCTGGGCGT	TAATCAGAGA	GGCGAATTAT	GTGTCAGAGG	ACCTATGATT
2881	ATGTCCGGTT	ATGTAAACAA	TCCGGAAGCG	ACCAACGCCT	TGATTGACAA	GGATGGATGG
2941	CTACATTCTG	GAGACATAGC	TTACTGGGAC	GAAGACGAAC	ACTTCTTCAT	AGTTGACCGC
3001	TTGAAGTCTT	TAATTAAATA	CAAAGGATAT	CAGGTGGCCC	CCGCTGAATT	GGAATCGATA
3061	TTGTTACAAC	ACCCCAACAT	CTTCGACGCG	GGCGTGGCAG	GTCTTCCCGA	CGATGACGCC
3121	GGTGAACCTT	CCGCCGCCGT	TGTTGTTTTG	GAGCACGGAA	AGACGATGAC	GGAAAAAGAG
3181	ATCGTGGAAT	ACGTCGCCAG	TCAAGTAACA	ACCGCGAAAA	AGTTGCGCGG	AGGAGTTGTG
3241	TTTGTGGACG	AAGTACCGAA	AGGTCTTACC	GGAAAACTCG	ACGCAAGAAA	AATCAGAGAG
3301	ATCCTCATAA	AGGCCAAGAA	GGGCGGAAAG	TCCAAATTGT	AACACGTGTA	ATTCTAGAGC
3361	GGCCGCTTCG	AGCAGACATG	ATAAGATACA	TTGATGAGTT	TGGACAAACC	ACAATAAGAA
3421	TGCAGTGAAA	AAAATGCTTT	ATTTGTGAAA	TTTGTGATGC	TATTGCTTTA	TTTGTAAACCA
3481	TTATAAGCTG	CAATAAACAA	GTTAAACAACA	ACAATTGCAT	TCATTTTATG	TTTCAGGTTT

3541	AGGGGGAGGT	GTGGGAGGTT	TTTTAAAGCA	AGTAAAACCT	CTACAAATGT	GGTAAAATCG
3601	ATAAGTTTAA	ACAGATCCAG	GTGGCACTTT	TCGGGGAAAT	GTGCGCGGAA	CCCCTATTTG
3661	TTTATTTTTTC	TAAATACATT	CAAATATGTA	TCCGCTCATG	AGACAATAAC	CCTGATAAAT
3721	GCTTCAATAA	TATTGAAAAA	GGAAGAGTAT	GAGTATTCAA	CATTTCCGTG	TCGCCCTTAT
3781	TCCCTTTTTTT	GCGGCATTTT	GCCTTCCTGT	TTTTTGCTCAC	CCAGAAACGC	TGGTGAAAGT
3841	AAAAGATGCT	GAAGATCAGT	TGGGTGCACG	AGTGGGTTAC	ATCGAACTGG	ATCTCAACAG
3901	CGGTAAGATC	CTTGAGAGTT	TTCGCCCCGA	AGAACGTTTT	CCAATGATGA	GCACTTTTAA
3961	AGTTCTGCTA	TGTGGCGCGG	TATTATCCCG	TATTGACGCC	GGGCAAGAGC	AACTCGGTCTG
4021	CCGCATACAC	TATTCTCAGA	ATGACTTGGT	TGAGTACTCA	CCAGTCACAG	AAAAGCATCT
4081	TACGGATGGC	ATGACAGTAA	GAGAATTATG	CAGTGCTGCC	ATAACCATGA	GTGATAACAC
4141	TGCGGCCAAC	TTACTTCTGA	CAACGATCGG	AGGACCGAAG	GAGCTAACCG	CTTTTTTGCA
4201	CAACATGGGG	GATCATGTAA	CTCGCCTTGA	TCGTTGGGAA	CCGAGCTGTA	ATGAAGCCAT
4261	ACCAAACGAC	GAGCGTGACA	CCACGATGCC	TGTAGCAATG	GCAACAACGT	TGCGCAAACCT
4321	ATTAAGTGGC	GAAGTACTTA	CTCTAGCTTC	CCGGCAACAA	TTAATAGACT	GGATGGAGGC
4381	GGATAAAGTT	GCAGGACCAC	TTCTGCGCTC	GGCCCTTCCG	GCTGGCTGGT	TTATTGCTGA
4441	TAAATCTGGA	GCCGGTGAGC	GTGGGTCTCG	CGGTATCATT	GCAGCACTGG	GGCCAGATGG
4501	TAAGCCCTCC	CGTATCGTAG	TTATCTACAC	GACGGGGAGT	CAGGCAACTA	TGGATGAACG
4561	AAATAGACAG	ATCGCTGAGA	TAGGTGCCTC	ACTGATTAAG	CATTGGTAAC	TGTCAGACCA
4621	AGTTTACTCA	TATATACTTT	AGATTGATTT	AAAACCTCAT	TTTTAATTTA	AAAGGATCTA
4681	GGTGAAGATC	CTTTTTGATA	ATCTCATGAC	CAAAATCCCT	TAACGTGAGT	TTTCGTTCCA
4741	CTGAGCGTCA	GACCCCGTAG	AAAAGATCAA	AGGATCTTCT	TGAGATCCTT	TTTTTCTGCG
4801	CGTAATCTGC	TGCTTGCAAA	CAAAAAAACC	ACCGCTACCA	GCGGTGGTTT	GTTTGCCGGA
4861	TCAAGAGCTA	CCAACCTCTT	TTCCGAAGGT	AACTGGCTTC	AGCAGAGCGC	AGATACCAAA
4921	TACTGTCCCT	CTAGTGTAGC	CGTAGTTAGG	CCACCACTTC	AAGAACTCTG	TAGCACCGCC
4981	TACATACCTC	GCTCTGCTAA	TCCTGTTACC	AGTGGCTGCT	GCCAGTGGCG	ATAAGTCGTG
5041	TCTTACCGGG	TTGGACTCAA	GACGATAGTT	ACCGGATAAG	GCGCAGCGGT	CGGGCTGAAC
5101	GGGGGGTTTC	TGCACACAGC	CCAGCTTGGA	GCGAACGACC	TACACCGAAC	TGAGATACCT
5161	ACAGCGTGAG	CTATGAGAAA	GCGCCACGCT	TCCCGAAGGG	AGAAAGGCGG	ACAGGTATCC
5221	GGTAAGCGGC	AGGGTCGGAA	CAGGAGAGCG	CACGAGGGAG	CTTCCAGGGG	GAAACGCCTG
5281	GTATCTTTAT	AGTCCTGTCT	GGTTTCGCCA	CCTCTGACTT	GAGCGTCGAT	TTTTGTGATG
5341	CTCGTCAGGG	GGGCGGAGCC	TATGGAAAAA	CGCCAGCAAC	GCGGCCTTTT	TACGGTTCCT
5401	GGCCTTTTTG	TGGCCTTTTG	CTCACATGGC	TCGAC		

Legend

- Nucleotides shown in **red** represent the Renilla luciferase region.
- Nucleotides shown in **black bold** represent the inserted sequence of rHP2-E.
- Nucleotides shown in **green** represent the firefly luciferase region.
- Underlined nucleotides shown in **blue** represent potential binding sites of unmodified PNA P1.

rHP2-C/p2Luc

5436 bp ds-DNA

1	AGATCGATCT	GCGCAGCACC	ATGGCCTGAA	ATAACCTCTG	AAAGAGGAAC	TTGGTTAGGT
61	ACCTTCTGAG	GCGGAAAGAA	CCAGCTGTGG	AATGTGTGTC	AGTTAGGGTG	TGGAAAGTCC
121	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC	TCAATTAGTC	AGCAACCAGG
181	TGTGGAAAGT	CCCCAGGCTC	CCCAGCAGGC	AGAAGTATGC	AAAGCATGCA	TCTCAATTAG
241	TCAGCAACCA	TAGTCCCGCC	CCTAACTCCG	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC
301	GCCCATTCTC	CGCCCCATGG	CTGACTAATT	TTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC
361	TCGGCCTCTG	AGCTATTCCA	GAAGTAGTGA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC
421	AAAAAGCTTG	ATTCTTCTGA	CACAACAGTC	TCGAACTTAA	GCTGCAGAAG	TTGGTCGTGA
481	GGCACTGGGC	AGGTAAGTAT	CAAGGTTACA	AGACAGGTTT	AAGGAGACCA	ATAGAAACTG
541	GGCTTGTCGA	GACAGAGAAG	ACTCTTGCGT	TTCTGATAGG	CACCTATTGG	TCTTACTGAC
601	ATCCACTTTG	CCTTTCTCTC	CACAGGTGTC	CACTCCAGT	TCAATTACAG	CTCTTAAGGC
661	TAGAGTACTT	AATACGACTC	ACTATAGGCT	AGCCACCATG	ACTTCGAAAG	TTTATGATCC
721	AGAACAAAGG	AAACGGATGA	TAAGTGGTCC	GCAGTGGTGG	GCCAGATGTA	AACAAATGAA
781	TGTTCTTGAT	TCATTTATTA	ATTATTATGA	TTCAGAAAAA	CATGCAGAAA	ATGCTGTTAT
841	TTTTTTACAT	GGTAACGCGG	CCTCTTCTTA	TTTATGGCGA	CATGTTGTGC	CACATATTGA
901	GCCAGTAGCG	CGGTGTATTA	TACCAGACCT	TATTGGTATG	GGCAAAATCAG	GCAAATCTGG
961	TAATGGTTCT	TATAGGTTAC	TTGATCATT	CAAATATCTT	ACTGCATGGT	TTGAACTTCT
1021	TAATTTACCA	AAGAAGATCA	TTTTTGTGCG	CCATGATTGG	GGTGCTTGTT	TGGCAATTTCA
1081	TTATAGCTAT	GAGCATCAAG	ATAAGATCAA	AGCAATAGTT	CACGCTGAAA	GTGTAGTAGA
1141	TGTGATTGAA	TCATGGGATG	AATGGCCTGA	TATTGAAGAA	GATATTGCGT	TGATCAAATC
1201	TGAAGAAGGA	GAAAAAATGG	TTTTGGAGAA	TAAGTCTTTC	GTGGAAACCA	TGTTGCCATC
1261	AAAAATCATG	AGAAAGTTAG	AACCAGAAGA	ATTTGCAGCA	TATCTTGAAC	CATTCAAAGA
1321	GAAAGGTGAA	GTTTCGTCGC	CAACATTATC	ATGGCCTCGT	GAAATCCCGT	TAGTAAAAGG
1381	TGGTAAACCT	GACGTTGTAC	AAATTGTTAG	GAATTATAAT	GCTTATCTAC	GTGCAAGTGA
1441	TGATTTACCA	AAAATGTTTA	TTGAATCGGA	CCCAGGATTC	TTTTCCAATG	CTATTGTTGA
1501	AGGTGCCAAG	AAGTTTCCTA	ATACTGAATT	TGTCAAAGTA	AAAGGTCTTC	ATTTTTTCGCA
1561	AGAAGATGCA	CCTGATGAAA	TGGGAAAAAT	TATCAAATCG	TTCTGTGAGC	GAGTTCTCAA
1621	AAATGAACAA	ATGTCGACGG	ATCCACTTCT	TAGGGTAGAG	AGACAAAGTT	TGCAATTTGT
1681	CTCTCTAGAG	CTCGAAGACG	CCAAAAACAT	AAAGAAAGGC	CCGCGGCCAT	TCTATCTCTC
1741	AGAGGATGGA	ACCGCTGGAG	AGCAACTGCA	TAAGGCTATG	AAGAGATACG	CCCTGGTTCC
1801	TGGAACAATT	GCTTTTACAG	ATGCACATAT	CGAGGTGAAC	ATCACGTACG	CGGAATACTT
1861	CGAAATGTCC	GTTTCGGTTGG	CAGAAGCTAT	GAAACGATAT	GGGCTGAATA	CAAATCACAG
1921	AATCGTCGTA	TGCAGTGAAA	ACTCTCTTCA	ATTCTTTATG	CCGGTGTGG	GCGCGTTATT
1981	TATCGGAGTT	GCAGTTGCGC	CCGCGAACGA	CATTTATAAT	GAACGTGAAT	TGCTCAACAG
2041	TATGAACATT	TCGCAGCCTA	CCGTAGTGTT	TGTTTCCAAA	AAGGGGTTGC	AAAAAATTTT
2101	GAACGTGCAA	AAAAAATTAC	CAATAATCCA	GAAAATTAT	ATCATGGATT	CTAAAACGGA
2161	TTACCAGGGA	TTTCAGTCGA	TGTACACGTT	CGTCACATCT	CATCTACCTC	CCGGTTTTAA
2221	TGAATACGAT	TTTGTACCAG	AGTCCTTTGA	TCTGTGACAA	ACAATTGCAC	TGATCATGAA
2281	TTCTCTGGA	TCTACTGGGT	TACCTAAGGG	TGTGGCCCTT	CCGCATAGAA	GTCCTGCGT
2341	CAGATTCTCG	CATGCCAGAG	ATCCTATTTT	TGGCAATCAA	ATCATTCGGG	ATACTGCGAT
2401	TTTAAGTGTT	GTTCCATTCC	ATCACGGTTT	TGGAATGTTT	ACTACACTCG	GATATTTGAT
2461	ATGTGGATTT	CGAGTCGTCT	TAATGTATAG	ATTTGAAGAA	GAGCTGTTTT	TACGATCCCT
2521	TCAGGATTAC	AAAATTCAAA	GTGCGTTGCT	AGTACCAACC	CTATTTTCAT	TCTTCGCCAA
2581	AAGCACTCTG	ATTGACAAAT	ACGATTTATC	TAATTTACAC	GAAATTGCTT	CTGGGGGCGC
2641	ACCTCTTTTC	AAAGAAGTCG	GGGAAGCGGT	TGCAAAACGC	TTCCATCTTC	CAGGGATACG
2701	ACAAGGATAT	GGGCTCACTG	AGACTACATC	AGCTATTCTG	ATTACACCCG	AGGGGGATGA
2761	TAAACCGGGC	GCGGTCGGTA	AAGTTGTTCC	ATTTTTTGAA	GCGAAGGTTG	TGGATCTGGA
2821	TACCGGGAAA	ACGCTGGGCG	TTAATCAGAG	AGGCGAATTA	TGTGTCAGAG	GACCTATGAT
2881	TATGTCCGGT	TATGTAAACA	ATCCGGAAGC	GACCAACGCC	TTGATTGACA	AGGATGGATG
2941	GCTACATTCT	GGAGACATAG	CTTACTGGGA	CGAAGACGAA	CACCTCTTCA	TAGTTGACCG
3001	CTTGAAGTCT	TTAATTAAAT	ACAAAGGATA	TCAGGTGGCC	CCCGCTGAAT	TGGAATCGAT
3061	ATTGTTACAA	CACCCCAACA	TCTTCGACGC	GGGCGTGGCA	GGTCTTCCCG	ACGATGACGC
3121	CGGTGAACCT	CCCGCCGCCG	TGTTGTTTTT	GGAGCACGGA	AAGACGATGA	CGGAAAAAGA
3181	GATCGTGGAT	TACGTCGCCA	GTCAGTAAC	AACCGCGAAA	AAGTTGCGCG	GAGGAGTTGT
3241	GTTTGTGGAC	GAAGTACCGA	AAGGTCTTAC	CGGAAAACCT	GACGCAAGAA	AAATCAGAGA
3301	GATCCTCAT	AAGGCCAAGA	AGGGCGGAAA	GTCCAAATTG	TAACACGTGT	AATTCTAGAG
3361	CGGCCGCTTC	GAGCAGACAT	GATAAGATAC	ATTGATGAGT	TTGGACAAAC	CACAACATGA
3421	ATGCAGTGAA	AAAAATGCTT	TATTTGTGAA	ATTTGTGATG	CTATTGCTTT	ATTTGTAACC
3481	ATTATAAGCT	GCAATAAACA	AGTTAAACAAC	AACAATTGCA	TTCATTTTAT	GTTTCAGGTT

3541	CAGGGGGAGG	TGTGGGAGGT	TTTTTAAAGC	AAGTAAAACC	TCTACAAATG	TGGTAAAATC
3601	GATAAGTTTA	AACAGATCCA	GGTGGCACTT	TTCGGGGAAA	TGTGCGCGGA	ACCCCTATTT
3661	GTTTATTTTT	CTAAATACAT	TCAAATATGT	ATCCGCTCAT	GAGACAATAA	CCCTGATAAA
3721	TGCTTCAATA	ATATTGAAAA	AGGAAGAGTA	TGAGTATTCA	ACATTTCCGT	GTCGCCCTTA
3781	TTCCCTTTTT	TGCGGCATTT	TGCCTTCCTG	TTTTTGCTCA	CCCAGAAACG	CTGGTGAAAG
3841	TAAAAGATGC	TGAAGATCAG	TTGGGTGCAC	GAGTGGGTTA	CATCGAACTG	GATCTCAACA
3901	GCGGTAAGAT	CCTTGAGAGT	TTTCGCCCCG	AAGAACGTTT	TCCAATGATG	AGCACTTTTA
3961	AAGTTCTGCT	ATGTGGCGCG	GTATTATCCC	GTATTGACGC	CGGGCAAGAG	CAACTCGGTC
4021	GCCGCATACA	CTATTCTCAG	AATGACTTGG	TTGAGTACTC	ACCAGTCACA	GAAAAGCATC
4081	TTACGGATGG	CATGACAGTA	AGAGAATTAT	GCAGTGCTGC	CATAACCATG	AGTGATAACA
4141	CTGCGGCCAA	CTTACTTCTG	ACAACGATCG	GAGGACCGAA	GGAGCTAACC	GCTTTTTTGC
4201	ACAACATGGG	GGATCATGTA	ACTCGCCTTG	ATCGTTGGGA	ACCGGAGCTG	AATGAAGCCA
4261	TACCAAACGA	CGAGCGTGAC	ACCACGATGC	CTGTAGCAAT	GGCAACAACG	TTGCGCAAAC
4321	TATTAAGTGG	CGAACTACTT	ACTCTAGCTT	CCCGGCAACA	ATTAATAGAC	TGGATGGAGG
4381	CGGATAAAGT	TGCAGGACCA	CTTCTGCGCT	CGGCCCTTCC	GGCTGGCTGG	TTTATTGCTG
4441	ATAAATCTGG	AGCCGGTGAG	CGTGGGTCTC	GCGGTATCAT	TGCAGCACTG	GGGCCAGATG
4501	GTAAGCCCTC	CCGTATCGTA	GTTATCTACA	CGACGGGGAG	TCAGGCAACT	ATGGATGAAC
4561	GAAATAGACA	GATCGCTGAG	ATAGGTGCCT	CACTGATTAA	GCATTGGTAA	CTGTCAGACC
4621	AAGTTTACTC	ATATATACTT	TAGATTGATT	TAAAACTTCA	TTTTTAATTT	AAAAGGATCT
4681	AGGTGAAGAT	CCTTTTTTGAT	AATCTCATGA	CCAAAATCCC	TTAACGTGAG	TTTTTCGTCC
4741	ACTGAGCGTC	AGACCCCGTA	GAAAAGATCA	AAGGATCTTC	TTGAGATCCT	TTTTTTCTGC
4801	GCGTAATCTG	CTGCTTGCAA	ACAAAAAAAC	CACCGCTACC	AGCGGTGGTT	TGTTTGCCGG
4861	ATCAAGAGCT	ACCAACTCTT	TTTCCGAAGG	TAAGTGGCTT	CAGCAGAGCG	CAGATACCAA
4921	ATACTGTCCT	TCTAGTGTAG	CCGTAGTTAG	GCCACCACTT	CAAGAACTCT	GTAGCACC GC
4981	CTACATACCT	CGCTCTGCTA	ATCCTGTTAC	CAGTGGCTGC	TGCCAGTGCC	GATAAGTCGT
5041	GTCTTACCGG	GTTGGACTCA	AGACGATAGT	TACCGGATAA	GGCGCAGCGG	TCGGGCTGAA
5101	CGGGGGGTTC	GTGCACACAG	CCCAGCTTGG	AGCGAACGAC	CTACACCGAA	CTGAGATACC
5161	TACAGCGTGA	GCTATGAGAA	AGCGCCACGC	TTCCCGAAGG	GAGAAAGGCG	GACAGGTATC
5221	CGGTAAGCGG	CAGGGTCGGA	ACAGGAGAGC	GCACGAGGGA	GCTTCCAGGG	GGAAACGCCT
5281	GGTATCTTTA	TAGTCCTGTC	GGGTTTCGCC	ACCTCTGACT	TGAGCGTCGA	TTTTTTGTGAT
5341	GCTCGTCAGG	GGGGCGGAGC	CTATGGAAAA	ACGCCAGCAA	CGCGGCCTTT	TTACGGTTCC
5401	TGGCCTTTTTG	CTGGCCTTTT	GCTCATATGG	CTCGAC		

Legend

- Nucleotides shown in **red** represent the Renilla luciferase region.
- Nucleotides shown in **black bold** represent the inserted sequence of rHP2-C.
- Nucleotides shown in **purple bold** represent the mutated slippery sequence.
- Nucleotides shown in **green** represent the firefly luciferase region.
- Underlined nucleotides shown in **blue** represent potential binding sites of unmodified PNA P1.

Secondary Structure: small subunit ribosomal RNA

Secondary Structure: small subunit ribosomal RNA

Secondary Structure: large subunit ribosomal RNA - 5' half

Secondary Structure: large subunit ribosomal RNA - 3' half