Supporting Information

Template-directed Chemical Ligation to Obtain 3'-3' and 5'-5' Phosphodiester DNA Linkages

Haodong Chen[1,2], Feng Du[1], Gangyi Chen[1], Frank Streckenbach[1], Afshan Yasmeen[1], Yun Zhao[2] & Zhuo Tang[1,2*]

1. Natural Products Research Center, Chengdu Institution of Biology, Chinese Academy of Science, Chengdu, Sichuan, 610041, China

2. The College of Life Science, Sichuan University, Chengdu, Sichuan, 610065, China

*Corresponding author (email: tangzhuo@cib.ac.cn)

Table of Contents

1 DNA Sequences

Table 1 Oligonucleotides (presented 5'-3')

Name/Sequence
P-a': AAAAAATTTTTAAAATTTAAT
P-a: TAATTTAAAATTTTTAAAAAA*
P-b: GTCGACGCATTGAGA
P-b1: GTCGACGCATTGAGAA
Template-33: ATTAAATTTTAAAAATTTTTTCAGCTGCGTAACTCT
Template-33-short: ATTAAATTTTAAAAATTTTTTCAGCTGCGTAA
Template-55: CAGCTGCGTAACTCTATTAAATTTTAAAAATTTTTT
Template-55-short: TGCGTAACTCTATTAAATTTTAAAAATTTTTT
9aa-Template: TAATACGACTCACTATAGGAGGACGAAATGGACTACAAGGACGATGACGATAAG
16aa-Template: TAATACGACTCACTATAGGAGGACGAAATGGACTACAAGGACGATGACGATAAGCAGCTGCGTAACTCTAAAAAA
PCR-P: TAATACGACTCACTATA
PCR-RP-9aa: CTTATCGTCATCGTCC
PCR-RP-16aa: TTTTTTAGAGTTACGCAGC
Tail: AAAAAAAAAAAAAAATCCAAGAGCGCAGGACGCAGCTGCGTAACTCT
Splint-9: GGATTTTTTTTTTTTTTCTTATCGTCA
Splint-16: TTT TTT TTT TTT TTT AGA GT
28P: CCTGCGCTCTTGGAAAAAAAAAAAAACC-Puromycin
55-Primer splint: TTTTTTTTTTTTTTTTTTTTTAAAAATTTTAAATTA
55: 3'- AAAAAATTTTTAAAATTTAAT-5'-5'AAGAGTTACGCAGCTG-3'
55-Primer: 3'-AAAAAATTTTTAAAATTTAAT-5'-5'AAGAGTTACGCAGCTG-3'-5'-CCTGCGCTCTTGGAAAAAAAAAAAAACC-Puromycin
RC-a: AGAGTTACGCTCTAG
RC-b: CTAGATACGCAGCTG
RC-Template-33: ATTAAATTTTAAAAATTTTTTCTAGAGCGTAACTCT
RC-Template-55: CAGCTGCGTATCTAGATTAAATTTTAAAAATTTTTT

Asterisk represents the oligonucleotides is modified with 3'-phosphorylated when used for 3'-3' ligation. 28P and 55-Primer: oligonucleotides modified with 3'-Puromycin. Sequence in red respresents the parallel domain of ligation substrate. Sequence in purple showed the parallel domain in template or ligation substrate

2 Optimization of chemical ligation reaction

2.1 Effect of divalent metal ions on the ligation reaction

Different divalent metal ions solutions were prepared and added into the reactions of 3'-3' ligation. Reaction results were analyzed by Results were analyzed by electrophoresis on 10% polyacrylamide gels with 7 M urea and Tris-borate-EDTA (TBE) buffer and showed as **Figure S1**: Lane 1, ligation reaction was performed in 20 μL volumes incubated for 12 h at 20 ℃ containing 20 μM 3'-phosphorylated **P-a** (5'-TAATTTAAAATTTTAAAA A-3'), 20 μM **template-33** (5'-ATTAAATTTTAAAAATTTTTTCAGCTGCGTAACTCT-3'), 20 μM 5'-phosphorylated **P-b** (5'-GTCGACGCATTGAGA-3'), 1 mM *N*-Cyanoimidazole and 1 mM $ZnCl_2$; Lane 2-6, reactions were carried out in the same way as the one loaded in Lane 1, except for replacing $ZnCl_2$ with 1 mM $CaCl_2$ (lane 2), $MgCl_2$ (lane 3), $NiSO_4$ (lane 4), $MnCl_2$ (lane 5) and $FeCl_2$ (lane 6), respectively.

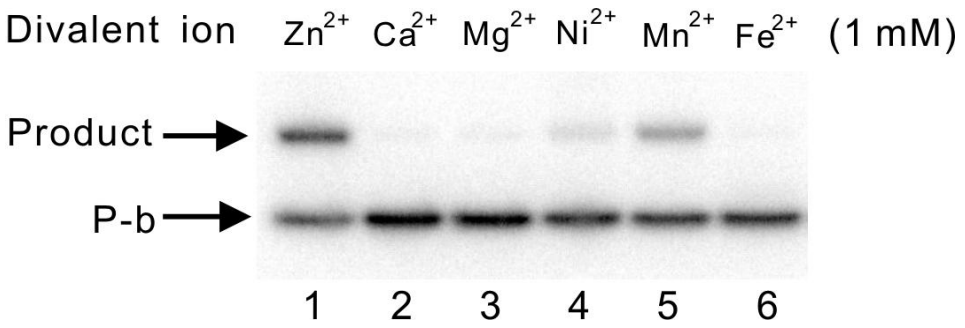

Figure S1 Effect of different divalent metal ions on the 3'-3' ligation reaction.

2.2 Optimization of Zinc ion concentration

Different concentrations of zinc ions solutions were prepared respectively and added into the reactions of 3'-3' ligation. Ligation reaction in lane 5**(Figure S2)** was carried out in 20 μL volumes incubated for 12 h at 20 ℃ containing 20 μM 3'-phosphorylated **P-a**, 20 μM **template-33**, 20 μM 5'-phosphorylated **P-b**, 1 mM *N*-Cyanoimidazole and 1 mM $ZnCl_2$. Other reactions were carried out in the same way as lane 5, except for that the final concentration of Zinc Chloride was kept at 0 μM (lane 1), 1 μM (lane 2), 10 μM (lane 3), 100 μM (lane 4), 10 mM (lane 6) and 100 mM (lane 7) instead.

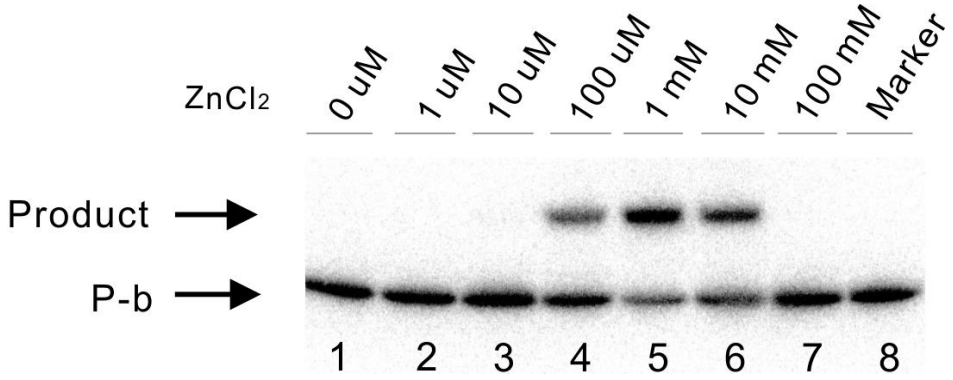

Figure S2 Optimization of Zinc ion concentration.

2.3 Optimization of N-Cyanoimidazole concentration

The concentration gradient of N-Cyanoimidazole was optimized between 10 μM and 100 mM. Ligation reaction in lane 4 **(Figure S3)** was carried out in 20 μL volumes incubated for 12 h at 20 ℃ containing 20 μM 3'-phosphorylated **P-a**, 20 μM **template-33**, 20 μM 5'-phosphorylated **P-b**, 1 mM *N*-Cyanoimidazole and 1 mM $ZnCl_2$. Other reactions were carried out in the same way as lane 4, except for that the final concentration of N-Cyanoimidazole was kept at 0 mM (lane 1), 100 mM (lane 2), 10 mM (lane 3), 1 mM (lane 4), 100 μM (lane 5) and 10 μM (lane 6) instead.

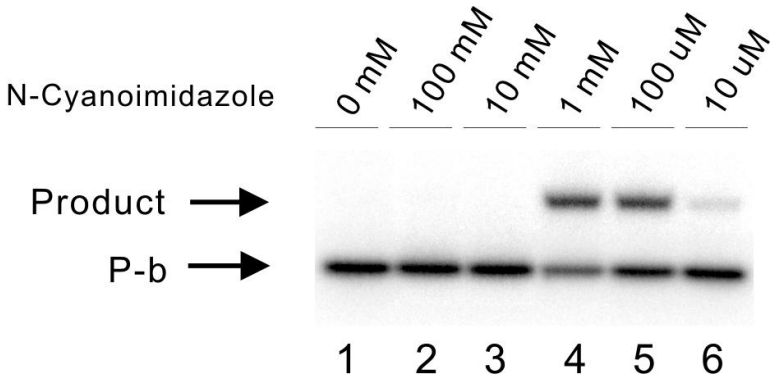

Figure S3 Optimization of N-Cyanoimidazole concentration.

3 Verification experiments

3.1 Verification experiments of 5'-5' chemical ligation

Template-directed chemical ligation reactions were performed in 20 μL volumes incubated for 12 h at 20 ℃ containing 1 mM *N*-Cyanoimidazole and 1 mM $ZnCl_2$. Ligation results was showed in **Figure S4**. Lane 1: isotope labeled **P-a** ligated to sequence **P-b1** (5'-GTCGACGCATTGAGAA-3') in the presence of **Template-55** (5'-CAGCTGCGTAACTCTATTAAATTTTAAAAATTTTTT-3'); Lane 2: same as lane 1 but without sequence **P-b1**; Lane 3: without **Template-55**; Lane 4: without **P-b1** and **Template-55**; Lane 5: isotope labeled sequence **P-b1** ligated to **P-a** in the presence of **Template-55**; Lane 6: same as lane 5 but without **P-a**; Lane 7: without **Template-55**; Lane 8: without **P-a** and **Template-55**. P^{32} in red denotes a labeled phosphate.

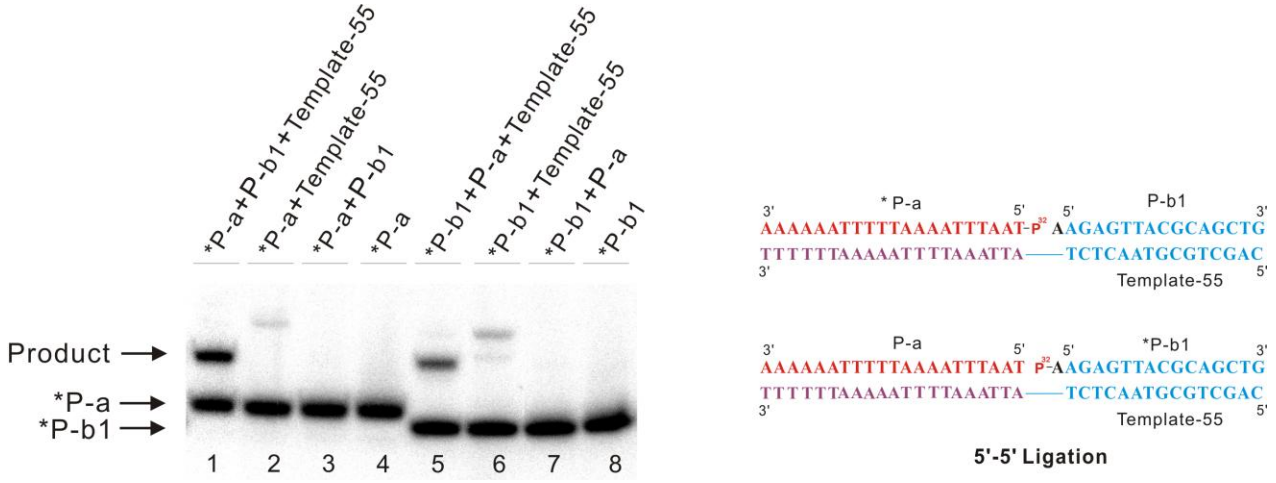

Figure S4 Verification experiments of 5'-5' chemical ligation.

3.2 DNA melting experiments

To investigate the DNA binding ability, DNA melting experiments were carried out by thermal denaturation studies in 2 mL volumes containing 0.5 μM 3'-3' ligation product **33** (5'-AGAGTTACGCAGCTG-3'-3'-AAAAAATTTTTAAAATTTAAT-5'), 0.5 μM **Template-33** in reaction buffer. Melting temperatures of 5'-5' ligation product **55** (3'- AAAAAATTTTTA AAATTTAAT-5'-5'AAGAGTTACGCAGCTG-3') and **Template-55** were detected under the same conditions as **33** with its corresponding template (**Figure S5**). Curve in green represents the detection results between **55** and **Template-55.** Curve in red represents **33** and **Template-33.** The detection of corresponding 5'-3' melting temperatures were also carried out in similar system as 3'-3' and 5'-5' at the same time. Curve in black represents **53-1 (**5'-AGAGTTACGCAGCTGAAAAAATTTTTAAAATTTAAT-3') and **Template-33** (control). Curve in blue represents **53-2** (5'-AAAAAATTTTTAAAATTTAATAAGAGTTACGCAGCTG-3') and **Template-55** (control). Alphabet A represents additional adenine base. Alphabet P represents phosphate group.

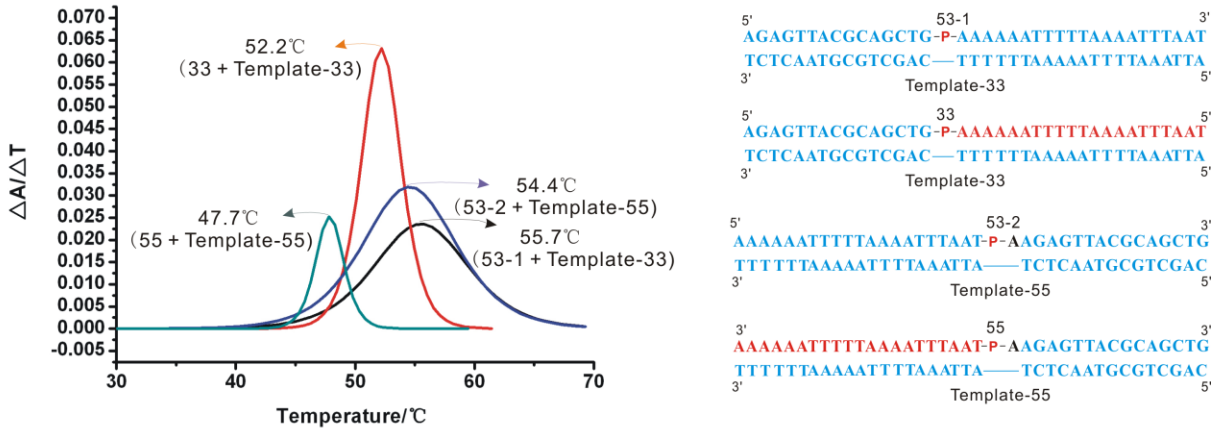

Figure S5 Melting temperature detections.

3.3 Restriction endonuclease verification experiment

Firstly, 3'-3', 5'-5' and their corresponding 5'-3' ligation reactions were performed, and the ligation products were separated and purified from denaturing PAGE. Purified ligation products were then used for the verification of digestion by restriction endonuclease *Xba* I. Cleavage reactions were performed in the mixture containing 20 μM ligation product, 20 μM corresponding template and 1×*Xba* I digestion buffer (1 mM Tris–HCl (pH 7.9), 0.7 mM $MgCl_2$, 10 mM NaCl, 0.7 mM 2-Mercaptoethanol and 0.01% BSA) at 37 ℃ for 30 min. Results of 3'-3', 5'-5' and 5'-3' chemical ligation products with corresponding template digested by restriction endonuclease were analyzed by 10% denaturing PAGE (**Figure S6A**). Lane 1: 5'-3' chemical ligation product similar to 3'-3' sequence treated with endonuclease. Lane 2: 3'-3' chemical ligation product treated with endonuclease. Lane 3: 5'-3' chemical ligation product similar to the sequence of 5'-5' treated with endonuclease. Lane 4: 5'-5' chemical ligation product treated with endonuclease. Corresponding structures of lane 1-4 were shown in **Figure S6B**. dsDNA in purple represents restriction enzyme cutting sites or similar restriction enzyme cutting sites include 3'-3' and 5'-5' structure. **Figure S6C** shows the recognition site or similar recognition site sequences of *Xba*I . SM is the abbreviation of starting material. P^{32} in red denotes a labeled phosphate.

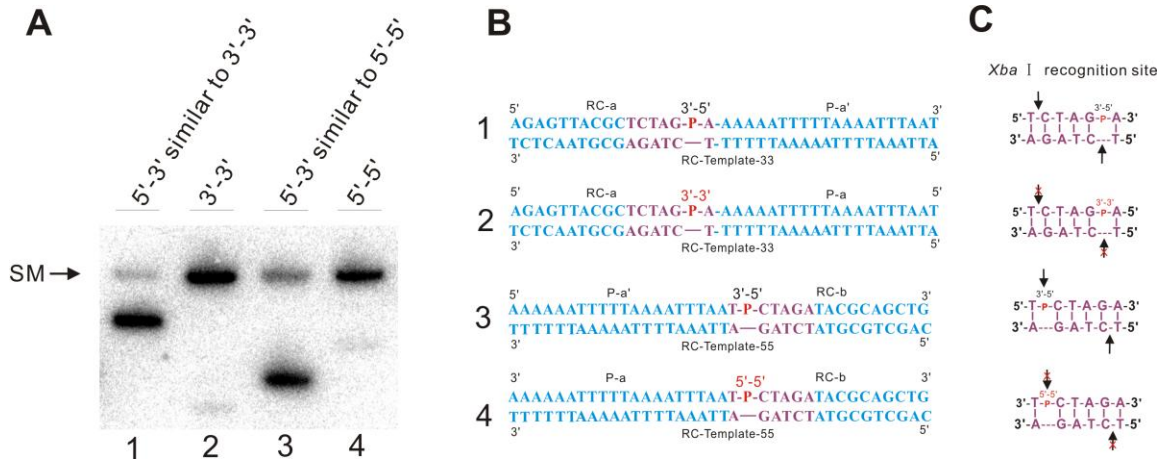

Figure S6 Restriction endonuclease digestions.

3.4 Mass spectrometry analysis of Ligation product

Initially, the lengths of the 3'-3' and 5'-5' ligation products were same with their respective templates, and the products could not be separated from the templates when analyzed by PAGE. To solve this problem, shorter templates were designed and a corresponding high yield of both 3'-3' and 5'-5' ligations in different directions were obtained. Chemical catalyzed ligation of 3'-3' was done in 400 μL volumes with 20 μM phosphorylation **P-a**, 20 μM **Template-33-short** (5'-ATTAAATTTTAAAAATTTTTTCAGCTGCGTAA-3'), 20 μM **P-b** (5'-labeled with radioisotope ^{32}P or non-radioactive phosphorus), 1 mM N-Cyanoimidazole, 1 mM $ZnCl_2$ and incubated at 20℃ for 12 h (**Lane 1, Figure S7**). 5'-5' ligation was carried out with the same system as 3'-3' except by using **Template-55-short** (5'-TGCGTAACTCTATTAAATTTTAAAAATTTTTT-3') as splint (**Lane 3, Figure S7**), **P-a** and phosphorylation **P-b** as ligation substrates. 3'-3' and 5'-5' ligation reactions with their corresponding same length templates **Template-33** and **Template-55** were also carried out as controls (**Lane 2 and Lane 4, Figure S7**). Products **33** and **55** were extracted from the PAGE gel and the mass spectrum analysis was conducted (**Figure S8**). (**A**) HRMS(ESI) m/z of **33** calcd for $[M+H]^+$: 11113.4, found:11113.5; (**B**) HRMS(ESI) m/z of **55** calcd for $[M+H]^+$: 11426.2, found:11426.3. ^{32}P denotes a labeled phosphate. P in black represents phosphate group.

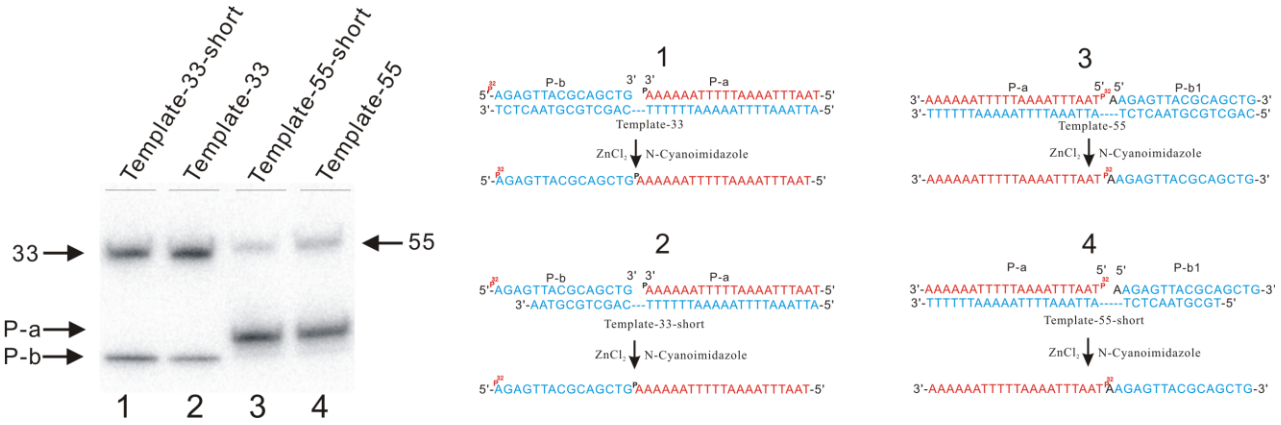

Figure S7 Chemical ligation with short splints.

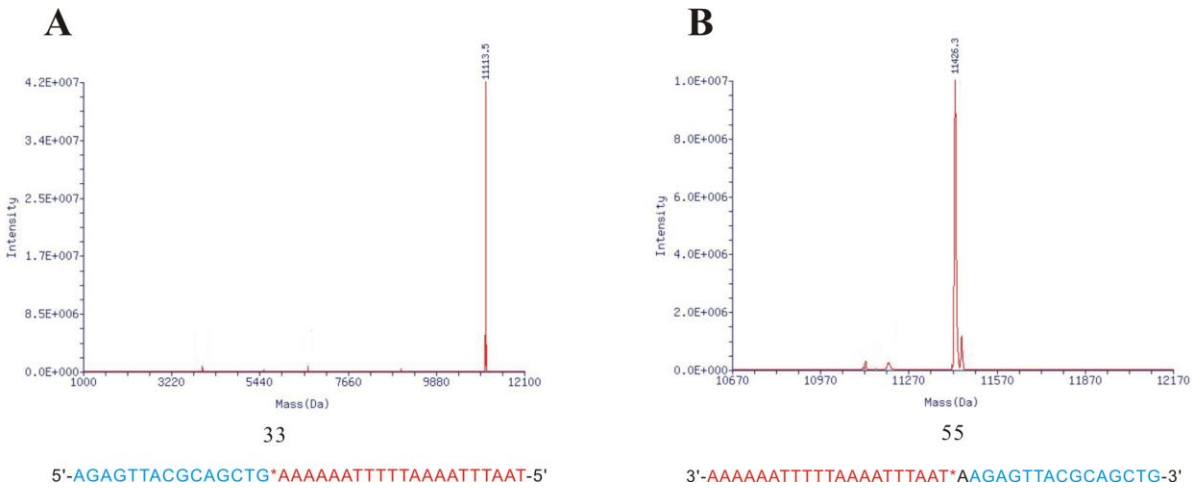

Figure S8 Mass spectrums of **33** and **55**.

4 *In vitro* transcription and translation

4.1 *In vitro* transcription

mRNAs were transcribed by incubating 1 μM of double stranded DNA template from PCR, 10 mM of NTP, 1×T7 RNA polymerase buffer (40 mM Tris–HCl (pH 7.9), 10 mM NaCl, 10 mM dithiothreitol, 6 mM $MgCl_2$ and 2 mM spermidine) and 1.25 unit per ml RNA polymerase at 37℃ for 4 h. After transcription, ethanol precipitation was carried out and the mRNAs were then purified by PAGE (**data not shown**).

4.2 *In vitro* translation

mRNAs were transcripted by using T7 RNA Polymerase and used as the templates for translation after ligate to **Tail** (**Tail**: 5'-AAAAAAAAAAAAAAATCCAAGAGCGCAGGACGCAGCTGCGTAACTCT-3') in the presence of **splint-16** (5'-TTTTTTTTTTTTTTAGAGT-3'). The translation template **16aa-RNA-tail** (5'-GGAG GACGAAAUGGACUACAAGGACGAUGACGAUAAGCAGCUGCGUAACUCUAAAAAAAAAAAAAAAA AAAAAATCCAAGAGCGCAGGACGCAGCTGCGTAACTCT-3') was hybridized to the puromycin and 5'-5' structure-included **55-primer** (3'-GTCGACGCATTGAGAA-5'-5'-TAATTTAAAATTTTTAAAAAACCTGCGC TCTTGGAAAAAAAAAAAAACC-3'-puromycin), which was obtained by the ligation reaction between **55** and phosphorylated **28P** (5'-CCTGCGCTCTTGGAAAAAAAAAAAAACC-3'-Puromycin) at the present of **55-Primer splint** (5'-TTTTTTTTTTTTTTTTTTTTTAAAAATTTTAAATTA-3') and catalyzed by T4 DNA ligase at 25℃ for 1 h. Translations were performed in Rabbit Reticulocyte Lysate (RRL; Promega) at 20℃ for 30 min under standard conditions. After cooling on ice for 40 min, KCl and $MgCl_2$ were added to the final concentration of 500 mM and 50 mM respectively. Translation results were analyzed on 10% denaturing PAGE after keeping the reaction tubes at room temperature for 1 h followed by long incubations at -20℃ for more than 10 h(**Figure 4C**).

4.3 Peptides verification experiment

Before perfoming the peptides degradation experiments, phenol/chloroform extraction followed by ethanol precipitation were carried out to purify the sample of translation. Proteinase K (0.6 U/μl) was added into the translation sample which was redissolved in nuclease-free water. Afterwards, the mixture was incubated at 37℃ for 30 min. Digestion results were analyzed by 10% denaturing PAGE. Results were showed in **Figure S9**, Lane 1: Translation sample without treated with Proteinase K; Lane 2: Translation sample treated with Proteinase K. P in red represents phosphate group.

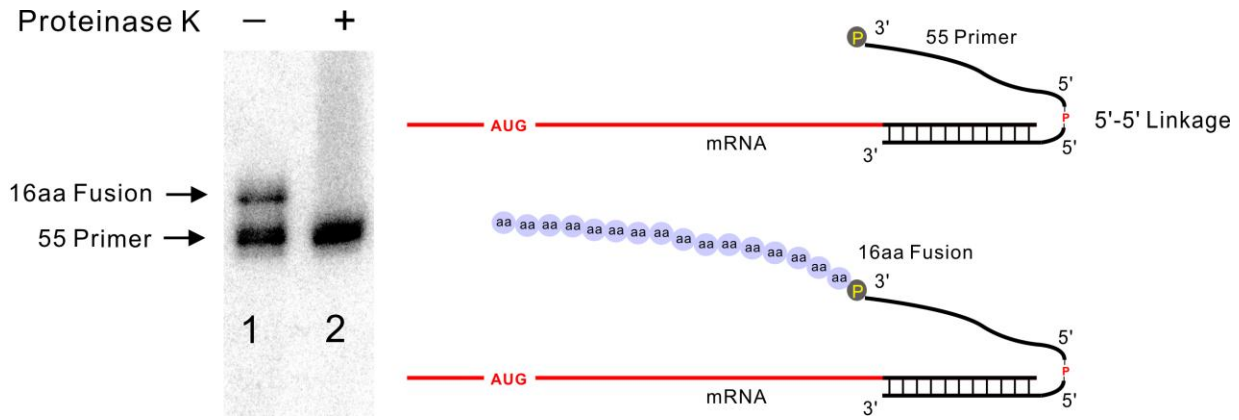

Figure S9 Preliminary validation of DNA-peptide fusions.

4.4 Reverse transcription

Reverse transcription was performed in 10 μL volumes incubated for 30 min at 42℃ containing 20 μM **16aa-RNA-Tail**, 20 μM **16aa-Fusion**, 0.5 mM dNTPs, 1×ES RT buffer, 10 U per μL M-MLV reverse transcriptase (Transgen) and 0.8 U per μL Bsm DNA Polymerase (Thermo). The result was analyzed by 10% denaturing PAGE after ethanol precipitation (**Figure 4D**).

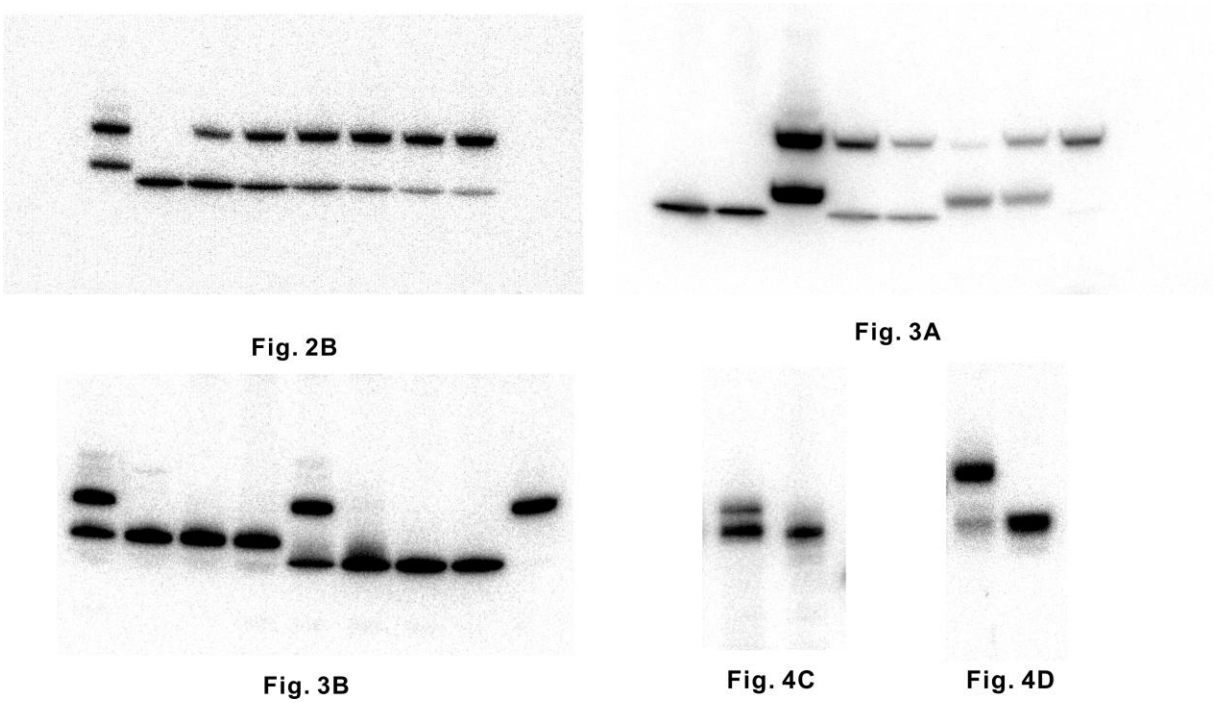

Figure S10 Full-length gels and blots images. Full-length gels and blots images Fig. 2B, 3A, 3B, 4C and 4D.