Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

Efficient construction of stable gene nanoparticles through polymerase chain reaction with flexible branched primers for gene delivery

Jianbing Liu, Runyu Wang, Dejun Ma, Di Ouyang and Zhen Xi*

Department of Chemical Biology, State Key Laboratory of Elemento-Organic Chemistry, National Engineering Research Center of Pesticide (Tianjin), Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Nankai University, Tianjin 300071, China

E-mail: zhenxi@nankai.edu.cn

Table of CONTENTS:

1. General Information	S2
2. DNA sequences of the templates and primers	S2
3. Synthesis of 3M and flexible branched primers	S4
4. PCR process with the special primer pairs	S5
5. Restriction endonuclease AvaI and T4 ligase analysis	S6
6. SP6 cell-free expression	S7
7. Ase I and Afl II digestion	S7
8. Structure analysis for the PCR products F^3+R^3 of CMV-EGFP template	S8
9. Exonuclease III stability assay	S8
10. Cell culture and transfection	S9
11. NMR and HRMS spectra of 3M	S9
12. MS characterization of F^3 and R^3	S11

1. General Information

All chemicals and solvents used in organic synthesis were in analytic level. And all solvents were purified or dried according to standard procedures. Millipore water was used to prepare all aqueous solutions. Simple unmodified oligonucleotides were synthesized and purified by TaKaRa (Dalian, China). The 5' thiol-modified C6 S-S DNA sequences were synthesized and purified by BioSune (Shanghai, China). The PCR products of the template SP6-EGFP and CMV-EGFP were purified by the GeneJET PCR Purification Kit from Thermo (Beijing, China) before the biological applications. Restriction endonuclease and T4 ligase were provided by TianGen (Beijing, China). The SP6 high-yield wheat germ protein expression system was purchased from Promega (Beijing, China). Human embryo kidney cell line (HEK-293A) was purchased from ATCC (Beijing, China).

NMR spectra were recorded on a Bruker AV 400 spectrometer at 400 MHz (^{1}H NMR) and 101 MHz (^{13}C NMR). High resolution mass spectral analyses (HRMS) were measured on a Varian QFT-ESI Mass Spectrometer with ESI resource for small molecules and Waters LCT Premier XE for modified oligonucleotides. Atomic force microscopic (AFM) images were obtained using a NanoScope IV multimode atomic force microscope (Veeco Instruments, USA) in tapping mode in air at room temperature. Dynamic laser scattering (DLS) data were obtained by (BI-200SM, USA) at room temperature. Cell-free system, EGFP fluorescence intensity was recorded on Safire Microplate Reader (Tecan, Switzerland). HEK-293A EGFP fluorescence image was detected by confocal microscope (Olympus FV1000).

2. DNA sequences of the templates and primers

Table S1: DNA sequences of 50 bp template and primers.

Name	Sequence (5'-3')
S_{50}	CCTCCCAAGATGGCCCGATATTCCCGAGGTTTTGGACAGGGAAGAGGAGG
A_{50}	CCTCCTCTTCCCTGTCCAAAACCTCGGGAATATCGGGCCATCTTGGGAGG
F^{1}	CCTCCCAAGATGGCCCGATA
R^{1}	CCTCCTCTTCCCTGTCCAAA

Table S2: DNA sequences of reconstructed SP6-EGFP template and primers.

CCTCCCAAGATGGCCCGATAATTTAGGTGACACTATAGAACAGACCACCATGGTGAGCAAGGG

F^{1} SP6 promoter Kozak region

CGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCA
CAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTT
CATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGG
CGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCAT
GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCG
CGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTT
CAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTA
TATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGA
GGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT
GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAA
GCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGA
GCTGTACAAGTAAAAAAAAAAAAAAAAAAAATTTGGACAGGGAAGAGGAGG-3'

Reconstructed SP6-EGFP template was amplified from pEGFP-N1 by the following primer pair:

Name	Sequence (5'-3')
Forward primer	**CCTCCCAAGATGGCCCGATA**ATTTAGGTGACACTATAGAACAGACCACC **ATG**GTGAGCAAGGGC
Reverse primer	**CCTCCTCTTCCCTGTCCAAA**TTTTTTTTTTTTTTTTTTTT **TTA**CTTGTACAGCTCGTCCA

Table S3: DNA sequences of reconstructed CMV-EGFP template and primers.

CCTCCCAAGATGGCCCGATAATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATA

F[1] **Ase I**

TATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCC
CCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG
ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTA
CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAT
GGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCC
AAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCC
AAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGT
CTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGCGCTACCGGACTCAGATCTC
GAGCTCAAGCTTCGAATTCTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACC
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGC
GACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAG
CTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACC
ACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTC
TTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTG
AAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAAC
AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATC
CGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAA
GACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACT
CTCGGCATGGACGAGCTGTACAAG**TAA**AGCGGCCGCGACTCTAGATCATAATCAGCCATACCA
CATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATA
AAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCA
ATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCA
AACTCATCAATGTATCTTAAG**TTTGGACAGGGAAGAGGAGG**–3'

Afl II

Reconstructed CMV-EGFP template was amplified from pEGFP-N1 by the following primer pair:

Name	Sequence (5'-3')
Forward primer	**CCTCCCAAGATGGCCCGATA**ATTAATAGTAATCAATTACG
Reverse primer	**CCTCCTCTTCCCTGTCCAAA**CTTAAGATACATTGATGAGT

3. Synthesis of 3M and flexible branched primers

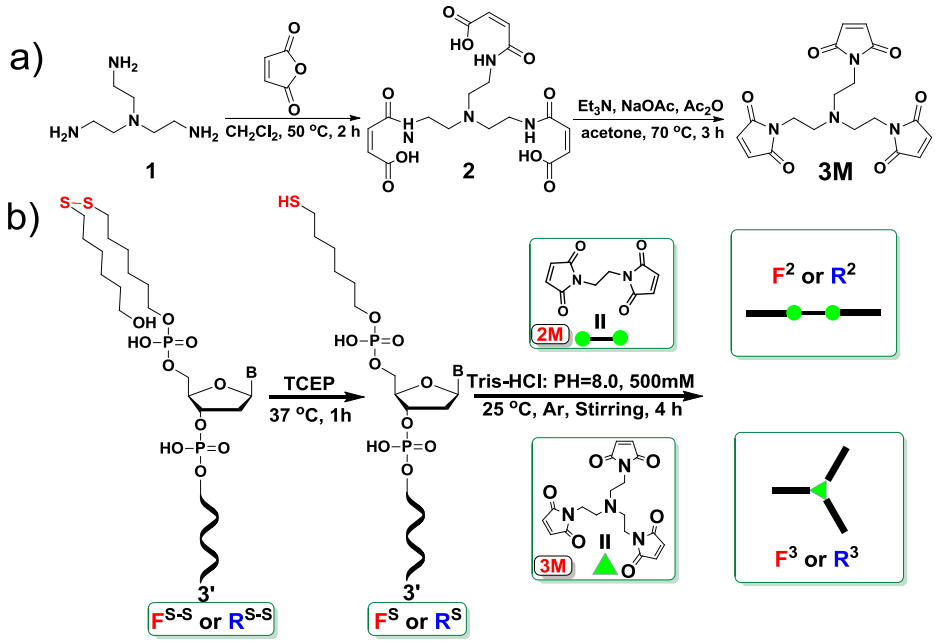


Fig. S1 Synthetic method for **3M** (a) and the flexible branched primers (b).

Synthesis of 3M: (Synthetic method for **2M** was showed in ref 1)

To a solution of N,N-bis(2-aminoethyl)ethane-1,2-diamine **1** (2.016 g, 13.79 mmol) in dry CH_2Cl_2, cis-Butenedioic anhydride (4.435 g, 45.21 mmol) was added in three batches. The reaction mixture was refluxed at 50 °C with stirring for 2 hours and then concentrated in vacuo. Compound **2** thus obtained as yellow solid, and could be used directly for the next step without further purification.

In the next step, Acetic anhydride (3.294 g, 32.26 mmol) was added to a solution of **2** (2.204 g, 5.00 mmol), triethylamine (420 mg, 4.15 mmol) and sodium acetate (335 mg, 4.09 mmol) in dry acetone and refluxed at 70 °C with stirring for 3 hours. The reaction mixture was concentrated into a crude residue, which was applied to a column of silica gel to yield compound **3M** 694 mg as yellow solid, TLC (Petroleum ether : ethyl acetate, 1:2 v/v): Rf = 0.4 and Yield 30% for 2 steps.

3M: ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 6.66 (s, 6H), 3.50 (t, J = 6.4 Hz, 6H), 2.69 (t, J = 6.4 Hz, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 170.7, 134.1, 51.7, 35.7. HRMS (ESI) calcd for $C_{18}H_{18}N_4O_6$, $[M + H]^+$: 387.1299, Found: 387.1295.

Synthesis of the flexible branched primers:

Single DNA with C6-S-S-C6 modification in 5' end (60 μg) ($\mathbf{F}^{S\text{-}S}$ or $\mathbf{R}^{S\text{-}S}$) was dissolved in 57 μL ddH_2O. Then 3 μL 100 mM TCEP was added, with a final concentration of 5 mM. The mixture was incubated at 37 °C for next 1 h. After that, the reaction mixture was precipitated twice in 0.3 M NaOAc and 75% ethanol to wash the TCEP residue which may affect the following reaction. Next, the produced C6-SH DNA solid ($\mathbf{F}^{S}$ or $\mathbf{R}^{S}$) was dissolved in 40 μL ddH_2O. Under argon atmosphere, 40 μL Tris-HCl (pH 8.0) was added with a final concentration of 500 mM in 25 °C. Then, **2M** (1/2 eq) or **3M** (1/3 eq) in DMF was added carefully in two batches every 30 min. With stirring, the reaction was monitored by denatured PAGE (7 M Urea) until no more products generated, which may take 4 h. Finally, $\mathbf{F}^2$ or $\mathbf{R}^2$ and $\mathbf{F}^3$ or $\mathbf{R}^3$ were purified by denatured PAGE (7 M Urea).

4. PCR process with the special primer pairs

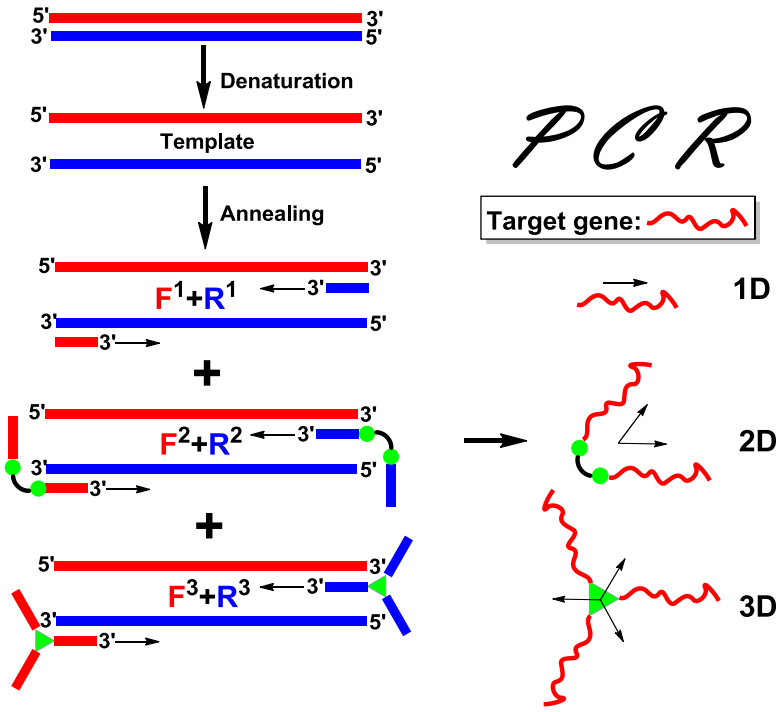


Fig. S2 PCR model with three different primer pairs.

A、50 bp template PCR

PCR content: 10.0 ng (0.324 pmol) 50 bp template, 0.25 μM (single DNA sequences contained) each primer (F^1 and R^1, F^2 and R^2, F^3 and R^3), 1 U La-Taq DNA polymerase, 5.0 μL 5 ×PCR reaction buffer and added ddH_2O into 25 μL PCR system.

PCR program: 95 °C 5 min - (95 °C 20 S - 60 °C 20 S - 72 °C 20 S) ×30 cycles - 72 °C 5 min - 4 °C 1 h.

B、SP6-EGFP template PCR

PCR content: 50.0 ng (0.100 pmol) 809 bp template, 0.25 μM (single DNA sequences contained) each primer (F^1 and R^1, F^2 and R^2, F^3 and R^3), 1 U La-Taq DNA polymerase, 5.0 μL 5 ×PCR reaction buffer and added ddH_2O into 25 μL PCR system.

PCR program: 95 °C 5 min - (95 °C 30 S - 60 °C 30 S - 72 °C 90 S) ×30 cycles - 72 °C 5 min - 4 °C 1 h.

C、CMV-EGFP template PCR

PCR content: 50.0 ng (0.048 pmol) 1679 bp template, 0.25 μM (single DNA sequences contained) each primer (F^1 and R^1, F^2 and R^2, F^3 and R^3), 1 U La-Taq DNA polymerase, 5.0 μL 5 ×PCR reaction buffer and added ddH_2O into 25 μL PCR system.

PCR program: 95 °C 5 min - (95 °C 30 S - 60 °C 30 S - 72 °C 120 S) ×30 cycles - 72 °C 5 min - 4 °C 1 h.

5. Restriction endonuclease AvaI and T4 ligase analysis

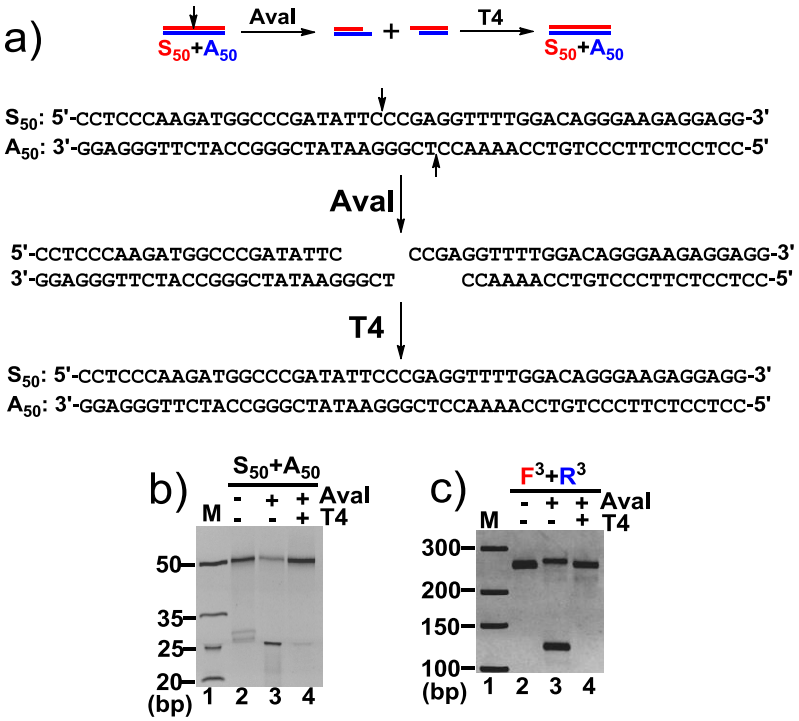


Fig. S3 Restriction endonuclease AvaI and T4 ligase analysis. (a) The model of AvaI and T4 ligase reaction. (b) 15% native PAGE analysis of the AvaI and T4 ligase reaction result of S_{50}+A_{50}. (c) 4% native PAGE analysis of the Ava I and T4 ligase reaction result of PCR product F^3+R^3.

Restriction endonuclease AvaI and T4 analysis was carried out in this way:

25 μL 0.25 μM PCR product based on the tri-branched primer pair F^3 and R^3, 5 U AvaI, 3 μL 10 × AvaI reaction buffer, and added ddH_2O into 30 μL system. The mixture was incubated at 37 °C for 1 h, then 2.0 μL (20 mM) ATP and 1.0 μL (400 U) T4 ligase were added, which was incubated at 16 °C for another 24 h.

6. SP6 cell-free expression

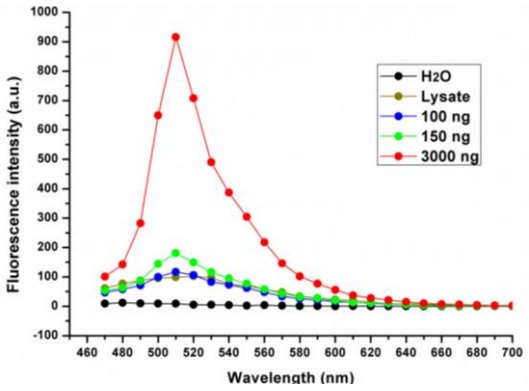


Fig. S4 The EGFP fluorescence detection of the SP6 cell-free expression with different amount of the template and PCR products F^1+R^1 (3000 ng).

About 50 ng SP6-EGFP template was involved in each of the three kinds of PCR products (3000 ng). So we employed 100 and 150 ng SP6-EGFP template in the cell-free system to compare the expression results with 3000 ng PCR product F^1+R^1. We found that the high fluorescence intensity of PCR product F^1+R^1 was not mainly caused by the template residue in these PCR products.

7. Ase I and Afl II digestion

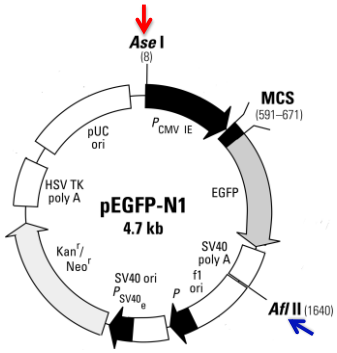


Fig. S5 Backbone of the pEGFP-N1 plasmid with Ase I and Afl II showed by the arrows.

5 μL 0.25 μM PCR product with three primer groups, 20 U Ase I and Afl II for each, 1 μL 10 × NEB 2.1 reaction buffer, and added ddH_2O into 10 μL system. The mixture was incubated at 37 ^oC for 4 h.

8. Structure analysis for the PCR products F^3+R^3 of CMV-EGFP template

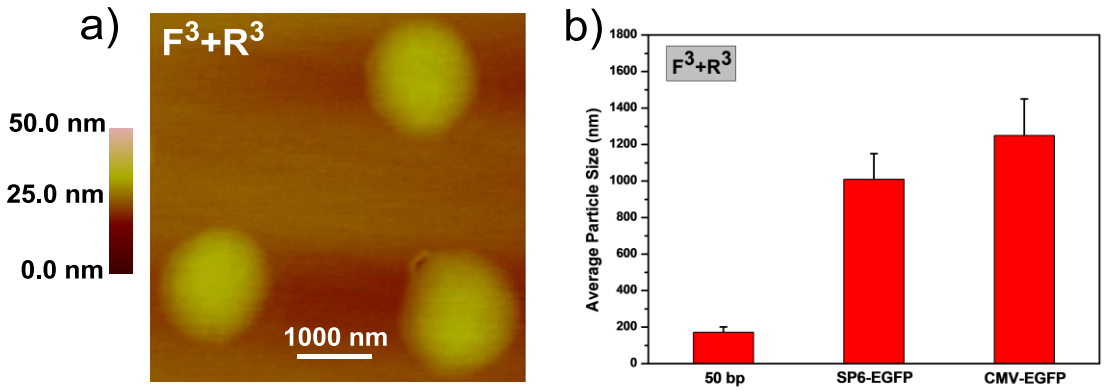


Fig. S6 Structure analysis for the PCR products F^3+R^3. (a) AFM analysis of the PCR products F^3+R^3 based on primer pair F^3 and R^3 of CMV-EGFP template, about 1300-1400 nm in diameter. Scale bar: 1000 nm. (b) DLS analysis of the PCR products F^3+R^3 based on primer pair F^3 and R^3 of three kinds of templates. (50 bp: 171±30 nm, SP6-EGFP: 1010±140 nm and CMV-EGFP: 1250±200 nm. The obtained diameters of DLS are larger than AFM, which may be caused by the shrinkage of gene nanoparticles during the drying process in the AFM samples preparation.)

AFM imaging:

0.25 μM of purified PCR product was diluted into 2.5 nM by 2 mM $Mg(OAc)_2$ and 3 μL of this sample was spotted on a freshly cleaved mica surface. After adsorbing for 5 min, 20 μL 2 mM $Mg(OAc)_2$ was used to wash the sample for once. Then the sample was dried at room temperature before imaging.

DLS analysis:

0.25 μM of purified PCR product was used directely for DLS analysis with a laser wavelength of λ= 678 nm and a scattering angle of 90°.

9. Exonuclease III stability assay

5 μL 0.25 μM PCR product, 3, 6 and 9 U Exonuclease III for the three kinds of PCR products, 1 μL 10 × Exonuclease III reaction buffer, and added ddH_2O into 10 μL system. The mixture was incubated at 37 °C for 1 h. Then, incubated at 70 °C for 20 min was needed to inactivate the Exonuclease III before gel analysis.

Table S4: The statistic results of the residual PCR products by the software of ImageJ.

Exonuclease III	3 U	6 U	9 U
F^1+R^1	20%	1%	0%
F^2+R^2	27%	17%	0%
F^3+R^3	61%	47%	27%

10. Cell culture and transfection

Human embryo kidney cells (HEK-293A) was cultured at 37 ^{o}C, 5% CO_2 in DMEM/HIGH GLUCOSE (GIBICO) supplemented with 10% fetal bovine serum (GIBICO), 100 U/mL penicillin, 100 μg/mL streptomycin, and 4 mM L-glutamine. The cells were maintained in exponential growth, and then seeded in 6-well plates (~30000 cells/well) to reach about 70% confluence for transfection.

The cells were transfected with 250 ng plasmid and three kinds of purified PCR products respectively with Lipofectamine 2000 (Invitrogen, USA) (total transfection volume, 2.0 mL) according to the manufacturer's instructions. Medium was replaced by 2 mL fresh DMEM medium supplemented with 10% FBS 4 hours later. After 20 hours, cells were imaged on a confocal microscope (Olympus FV1000) with 480 nm excitation.

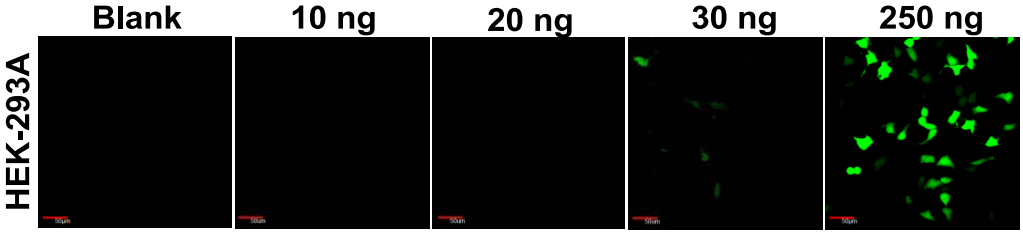
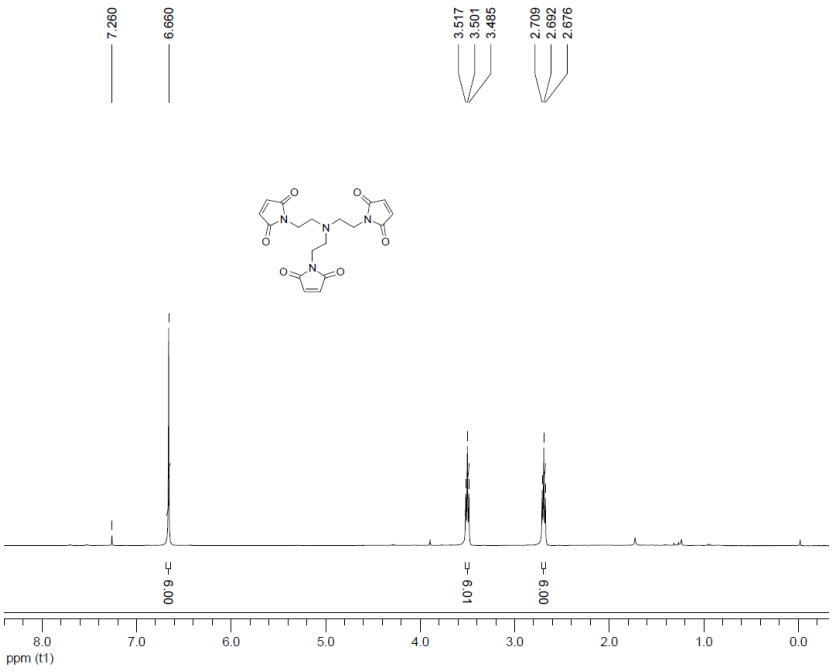


Fig. S7 Confocal microscopy images of the expression of EGFP in eukaryotic cells HEK-293A transfected with different amount of the CMV-EGFP template and PCR products F^1+R^1 (250 ng) respectively. Scale bar: 50 μm.

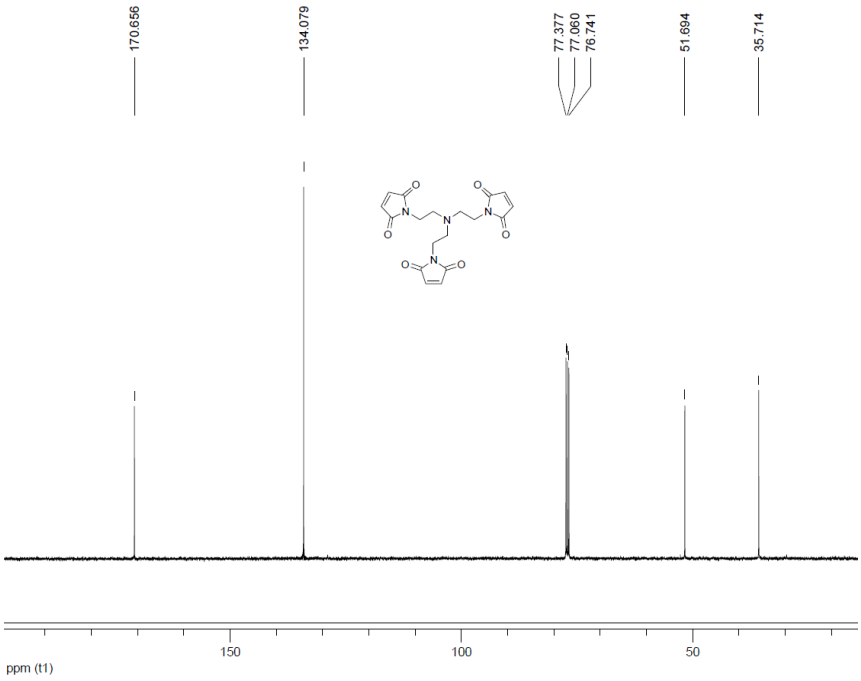
About 10 ng CMV-EGFP template was involved in the three kinds of PCR products (250 ng) respectively. So we transfected 10, 20 and 30 ng CMV-EGFP template into the HEK-293A to compare the cellular expression results with 250 ng PCR product F^1+R^1. We found that the high fluorescence intensity of PCR product F^1+R^1 was not mainly caused by the template residue in these PCR products.

11. NMR and HRMS Spectra of 3M

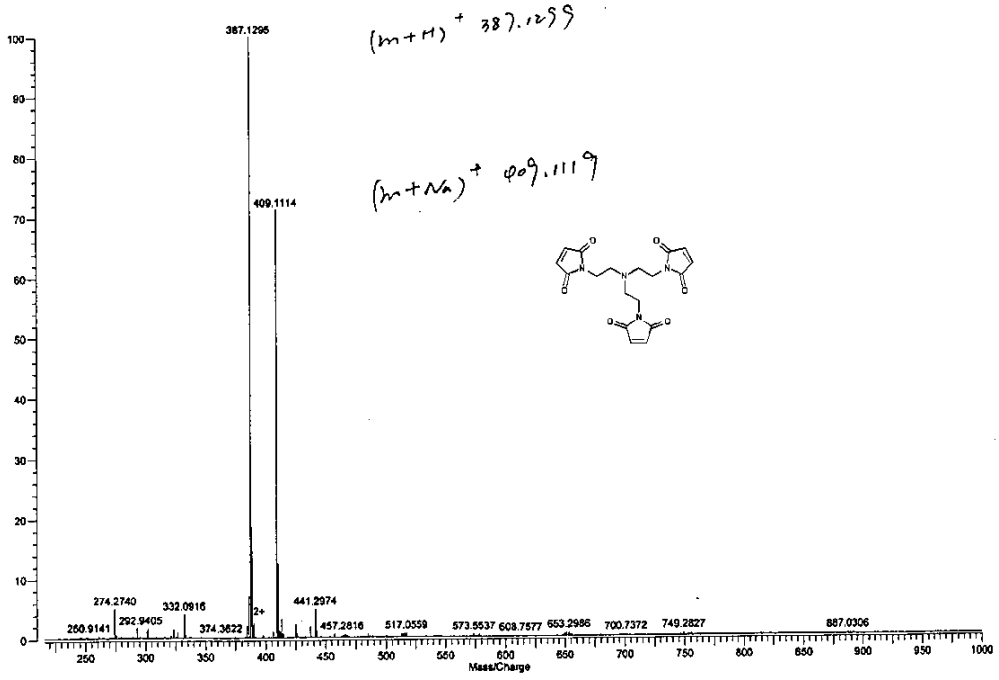
^{1}H NMR (400 MHz, $CDCl_3$):



^{13}C NMR (101 MHz, $CDCl_3$):

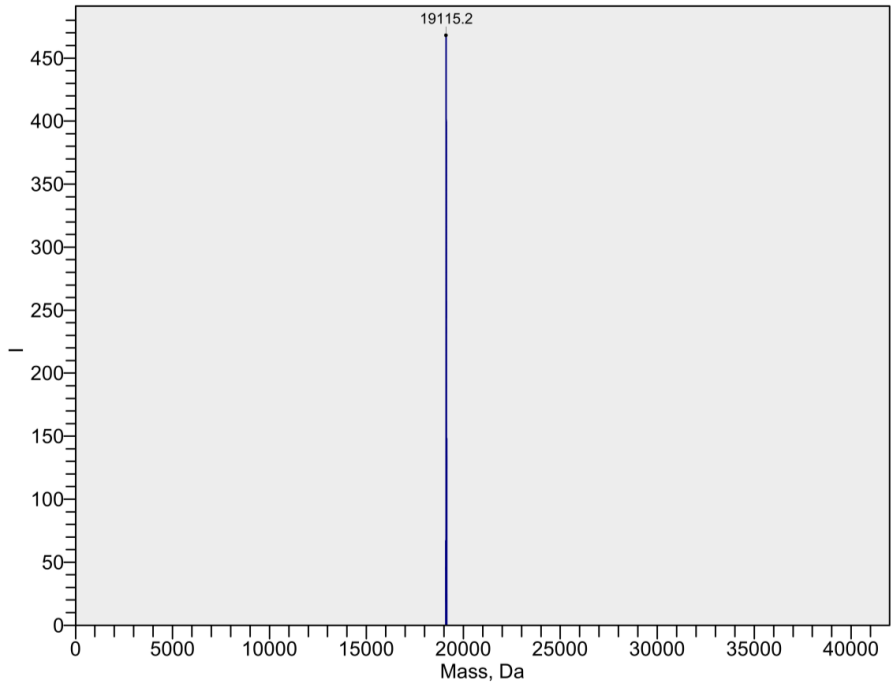


HRMS: $[M + H]^+$: MW. Calcd 387.1299 MW. Found 387.1295

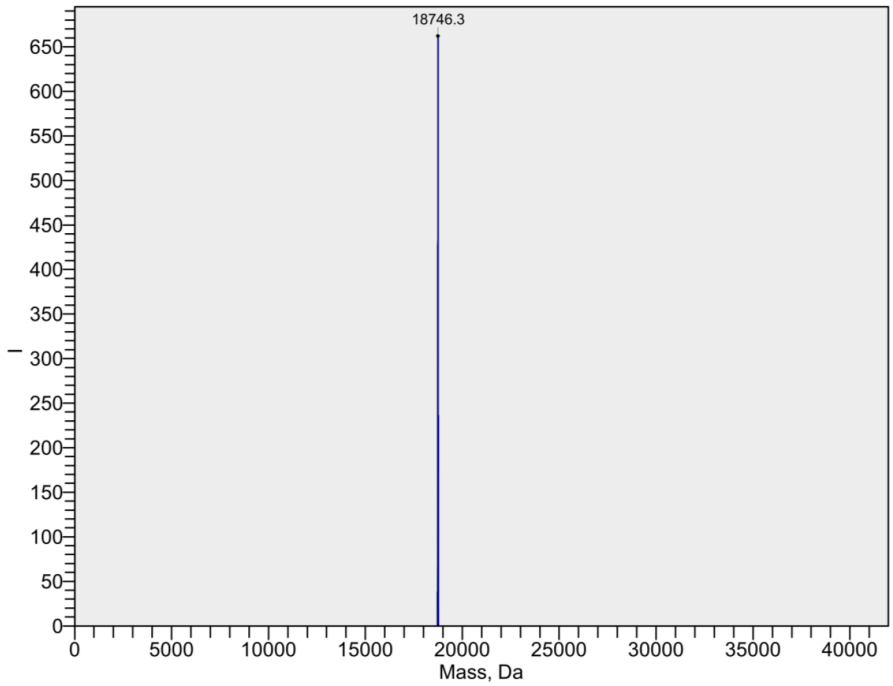


12. MS characterization of F^3 and R^3

F^3: MW.Calcd 19115.2 MW.Found 19115.2



R^3: MW.Calcd 18745.9 MW.Found 18746.3



The mass spectra of F^3 and R^3, with a single major peak, matched very well with the calculated values.

(1) L. Wei, L. Cao and Z. Xi, *Angew*. *Chem*., *Int*. *Ed*., 2013, **125**, 6629.