Electronic Supplementary Information

Proximity ligation-transcription circuits-powered exponential amplifications for single-molecule monitoring of telomerase in human cells

Li-juan Wang,[a,b,‡] Meng-meng Lv,[a,‡] Jin-ping Hu,[b,‡] Meng Liu,[a] and Chun-yang Zhang [a,*]

[a] College of Chemistry, Chemical Engineering and Materials Science, Collaborative Innovation Center of Functionalized Probes for Chemical Imaging in Universities of Shandong, Key Laboratory of Molecular and Nano Probes, Ministry of Education, Shandong Provincial Key Laboratory of Clean Production of Fine Chemicals, Shandong Normal University, Jinan 250014, PR China.

[b] School of Chemistry and Chemical Engineering, Southeast University, Nanjing, 211189, China.

* Corresponding author. Tel.: +86 0531-86186033; Fax: +86 0531-82615258. E-mail: cyzhang@sdnu.edu.cn.

‡ These authors contributed equally.

MATERIALS AND METHODS

Chemicals and materials.

All oligonucleotides (Table S1) were synthesized by Sangong Biotechnology (Shanghai, China). Taq DNA ligase, 10 × Taq DNA ligase reaction buffer (200 mM Tris-HCl, 250 mM KCl, 100 mM MgCl$_2$, 10 mM NAD$^+$, 100 mM DTT, 1% Triton X-100, pH 7.6), deoxynucleotide triphosphates (dNTPs), T7 RNA polymerase, 10× RNAPol reaction buffer (400 mM Tris-HCl, 60 mM MgCl$_2$, 20 mM spermidine, 10 mM DTT, pH 7.9), ribonucleotide solution set (i.e., ATP, UTP, GTP and CTP), formamidopyrimidine [fapy]-DNA glycosylase (FpG), 10 × NEBuffer 1 (100 mM Bis-Tris-Propane-HCl, 100 mM MgCl$_2$, 10 mM DTT), human apurinic/apyrimidinic endonuclease

(APE1), 10 × NEBuffer 4 (500 mM KAc, 200 mM Tris-HAc, 100 mM Mg(Ac)$_2$, 10 mM DTT), and bovine serum albumin (BSA) were purchased from New England Biolabs (Beverly, MA, USA). Duplex-specific nuclease (DSN) and 10 × DSN master buffer (500 mM Tris-HCl, 50 mM MgCl$_2$ and 10 mM DTT, pH 8.0) were obtained from Evrogen Joint Stock Company (Moscow, Russia). Magnesium chloride (MgCl$_2$), potassium chloride (KCl), Tween-20, ethylene glycol tetraacetic acid (EGTA) and ethylenediaminetetraacetic acid (EDTA) were obtained from Sigma-Aldrich Company (St. Louis, MO, USA). *N*, *N*'-1,3-phenylenebis-[2,3-dihydroxy-benzamide] (MST-312) and TRAPeze 1 × CHAPS lysis buffer (10 mM Tris-HCl, 1 mM MgCl$_2$, 5 mM β-mercaptoethanol, 1 mM [ethylenebis (oxyethylenenitrilo)] tetraacetic acid, 0.1 mM benzamidine, 0.5% 3-[(3-cholamidopropyl)-dimethylammonio] propanesulfonate, 10% glycerol, pH 7.5) were purchased from Millipore Corporation (Bedford, MA, USA). Human fetal pulmonary fibroblasts cell line (MRC-5), human cervical cancer cell line (HeLa), human lung cancer cell line (A549) and human breast adenocarcinoma cell line (MCF-7) were purchased from Cell Bank of Chinese Academy of Sciences (Shanghai, China). Diethylpyrocarbonate (DEPC)-treated water and SYBR Gold were obtained from Invitrogen Corporation (California, CA, USA).

Table S1. Sequences of oligonucleotides [α]

Note	Sequences (5'-3')
TS primer	AAT CCG TCG AGC AGA GTT
TPC4	AAT CCG TCG AGC AGA GTT **AGG GTT AGG GTT AGG GTT AGG GTT**
Hairpin probe	P-CCT AAC CCT CCC TAT AGT GAG TCG TAT TAT TTT TAA TAC GAC TCA CTA TAG GG

Linear probe	GCG CCC TAA CCC TAA C
Signal probe	BHQ2-GCG CCC TAA CCC <u>T</u>AA CCC TAT TTT
Cy5-TS primer	Cy5-AAT CCG TCG AGC AGA GTT
Reporter RNA	GGG AGG GUU AGG GUU AGG GUU AGG GCG C

[a] In TPC4, the bold letter represents the telomeric repeats $(TTAGGG)_4$. In hairpin probe, the "P" represents the phosphate group (PO_4) modification. In signal probe, the underlined "T" indicates the modification with a Cy5 fluorophore.

Cells culture and telomerase extracts preparation.

Human fetal pulmonary fibroblasts cell line (MRC-5 cells), human lung cancer cell line (A549 cells), human cervical cancer cell line (HeLa cells), and human breast adenocarcinoma cell line (MCF-7 cells) were cultured in Dulbecco's modified Eagle's medium (DMEM, Invitrogen, USA) supplemented with 1% penicillin-streptomycin and 10% fetal bovine serum at 37 °C in a humidified atmosphere containing 5% CO_2. The cells were collected at the exponential growth phase with trypsinization, washed twice with ice-cold PBS (pH 7.4), and then centrifuged at 800 rpm at 4 °C for 5 min. The numbers of cells were measured by Countstar automated cell counter (IC1000, Wilmington, DE, USA). Then five million cells were suspended in 200 μL of ice-cold 1 × CHAPS lysis buffer, incubated on ice for 30 min, followed by centrifugation at 12000 g at 4 °C for 20 min. The obtained supernatants were stored at -80 °C for further use.

Telomerization-activated proximity ligation-transcription circuits-powered cascade exponential amplifications.

All synthetic oligonucleotides were dissolved in 1 × Tris-EDTA buffer (10 mM Tris, 1 mM EDTA, pH 8.0) for stock solutions preparation. Hairpin probes were diluted to 10 μM in 1 × annealing buffer (10 mM Tris-HCl, 1.5 mM MgCl$_2$, pH 8.0), and then incubated at 95 °C for 5 min, followed by slowly cooling to room temperature to fold into hairpin structures. The obtained hairpin probes were stored at 4 °C for the subsequent experiments. The telomeric extension was performed in 20 μL of reaction solution containing 400 nM TS primer, 500 μM dNTPs, different-concentration telomerase extracts, and 2 μL of 10 × extension solution (30 mM Tris-HCl, 3 mM MgCl$_2$, 70 mM KCl, 1 mM EGTA, 0.05% (v/v) Tween-20, pH 8.3) at 37 °C for 15 min. Then 100 nM hairpin probe, 100 nM linear probe, 500 μM NTPs, 20 U of Taq DNA ligase, 1 U of T7 RNA polymerase, 20 U of RNase inhibitor, and 2 μL of 10 × RNAPol reaction buffer were added into the reaction mixture with a final volume of 30 μL, followed by incubation at 50 °C for 60 min to perform the proximity ligation and T7 transcription amplification circuits. Subsequently, 300 nM signal probe, 0.3 U of DSN, and 2 μL of 10 × DSN master buffer were added into the digestion solution with a total volume of 40 μL, followed by incubation at 55 °C for 15 min to carry out the DSN-assisted cyclic cleavage of signal probes.

Gel electrophoresis and fluorescence measurement.

The telomeric extension products and transcription products were analyzed by 12% nondenaturing polyacrylamide gel electrophoresis (PAGE). The telomerization-induced ligation products were analyzed by 12% denaturing PAGE. The gels were run in 1 × TBE buffer (89 mM Tris-boric acid, 2 mM EDTA, pH 8.2) at a 110 V constant voltage for 50 min at room temperature and visualized by a ChemiDoc MP Imaging system (Hercules, CA, USA) with SYBR Gold as the fluorescence

indicator. The fluorescence emission spectra of Cy5 were analyzed by a Hitachi F-7000 fluorescence spectrophotometer (Tokyo, Japan) with an excitation wavelength of 630 nm, and the fluorescence intensity at the emission wavelength of 670 nm were recorded for further data analysis.

Single-molecule imaging and data analysis.

The reaction products were diluted 200-fold with an imaging buffer (67 mM glycine-KOH, 1 mg/mL Trolox, 50 µg/mL BSA, and 2.5 mM $MgCl_2$, 1 mg/mL glucose oxidase, 0.4 mg/mL catalase, and 4 mg/mL D-glucose, pH 9.4). The 10 µL of samples were directly pipetted to the coverslip for imaging. The Cy5 fluorophores were excited by 640 nm lasers and were imaged onto an Andor Ixon Ultra 897 EMCCD camera. For data analysis, Image J software was used for counting the Cy5 molecules in an area of 600 × 600 pixels. The number of Cy5 molecules was the average of ten frames.

Inhibition assay. HeLa cells were incubated with various concentrations of MST-312 for 48 h, and the telomerase activity (equivalent to 1×10^5 HeLa cells) was measured. The relative activity (*RA*) of telomerase was calculated according to eq. S1.

$$RA\ (\%) = (C_i / C_t) \times 100\% = 10^{(N_i - N_t)/42.53} \times 100\% \qquad \text{(S1)}$$

where N_t and N_i represent the Cy5 counts in the absence and presence of MST-312, respectively. C_i and C_t were obtained according to the linear correlation equation in Figure 2D, respectively.

$$N_t = 86.68 + 42.53 \log X_t \qquad \text{(S2)}$$

$$N_i = 86.68 + 42.53 \log X_i \qquad \text{(S3)}$$

The IC$_{50}$ was estimated based on the fitting curve of *RA* against the MST-312 concentration.

Design of functional probes.

We designed one telomerase substrate (TS) primer, two ligation probes, and one signal probe (Table S1 and Scheme 1). The TS primer functions as the catalytic substrate for telomerase, initiating the incorporation of repetitive units (TTAGGG)$_n$ to its 3'-terminal to generate long-chain telomeric extension product (TEP). Two ligation probes contain the 5'-phosphorylated hairpin probe and 3'-hydroxylated linear probe. The 5'-phosphorylated hairpin probe consists of a double-stranded T7 promoter sequences whose sense and antisense strands are connected by a T$_4$ (5'-TTTT-3') loop, and a 9-nt single-stranded overhang that is complementary to the downstream of TEP. The 3'-hydroxylated linear probe is complementary to the upstream of TEP. With the TEP as the template, two ligation probes will be splinted together by Taq DNA ligase to catalyze the ligation reaction, generating the transcription templates. Signal probe is a 24-nt linear DNA sequence modified with a quencher (BHQ2) at its 5'-terminal and a fluorophore (Cy5) at thymine positioned 11-base away from its 3'-terminal, and it can completely hybridize with reporter RNA to form a RNA/DNA heteroduplex in which signal probe can be digested by DSN to generate an enhanced Cy5 signal.[1]

Evaluation of the proximity ligation-transcription circuits-powered cascade amplifications.

In this research, the core amplification system of this strategy is the proximity ligation-transcription circuits-powered cascade amplifications. Because the BIO-RAD CFX Connect TM Real-Time System (Hercules, CA, USA) is not suitable for Cy5 fluorescence signal

collection, we monitor the variance of Cy5 fluorescence signal with reaction time using a Hitachi F-7000 fluorescence spectrophotometer (Tokyo, Japan) to prove the exponential amplification reaction process. Since the reaction time and temperature of the DSN-digestion step are different from those of the proximity ligation-transcription circuits step, we alternatively monitor the variance of Cy5 fluorescence signal with the reaction time of proximity ligation-transcription circuits with other conditions remaining unchanged (i.e., the last-step reaction: DSN digestion at 55 °C for 15 min). As a proof-of-concept, we utilize the synthetic TPC4 (TS primer extended with $(TTAGGG)_4$) as a telomerase product model, and monitor the variance of Cy5 fluorescence signal with the reaction time of proximity ligation-transcription (Fig. S1). In the absence of TPC4, no distinct Cy5 fluorescence signal is detected in either the control with one tube (i.e., proximity ligation and transcription occur in one tube) (Fig. S1, black curve) or the control with two tubes (i.e., proximity ligation and transcription occur in two tubes) (Fig. S1, green curve), consistent with that obtained in the presence of signal probe + DSN (Fig. S1, orange curve), suggesting that no TS primer is telomerized to activate the proximity ligation and transcription reactions, and consequently no signal probe is digested by DSN to generate a Cy5 fluorescence signal. While in the presence of TPC4, high fluorescence signals are detected in response to both one-tube and two-tube (Fig. S1, red and blue curves), indicating that the proximity ligation and transcription cascades are activated and the signal probes in RNA/DNA heteroduplexes are digested by DSN to liberate abundant Cy5 molecules. Notably, a relatively low Cy5 fluorescence signal is detected in response to two-tube, suggesting the occurrence of proximity ligation and transcription in two tubes prevents the continuation of two circuits (Fig. S1, blue curve). In contrast, a strong Cy5 fluorescence signal is detected when proximity ligation and transcription occur in one tube (Fig.

S1, red curve). Importantly, the Cy5 fluorescence signal enhances with reaction time in a typical sigmoidal response, indicating that the proximity ligation-transcription circuits-powered cascade amplifications proceed in an exponential amplification manner.[2-5]
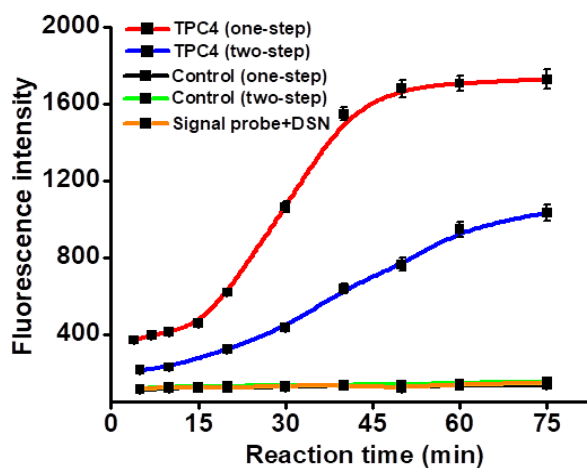

Fig. S1 Variance of Cy5 fluorescence signal with reaction time of the telomerization-activated proximity ligation-transcription circuits-powered cascade exponential amplifications. The TPC4 concentration is 10 nM.

To estimate the amplification efficiency of this strategy, we utilize the synthetic TPC4 as a telomerase product model. Since the core amplification system of this strategy is the proximity ligation-transcription circuits-powered cascade amplifications, we monitor the fluorescence signal of proximity ligation and transcription circuits under different conditions with SYBR Gold as the fluorescence indicator (Fig. S2A and S2B), and then calculate the fold amplification of this reaction process. To quantitatively evaluate the amplification reaction, the fold amplification is estimated based on eq. S4. [5, 6]

Fold amplification = N_1 / N_0 (S4)

where N_0 is the numbers of TPC4 (pmol) and N_1 is the numbers of the NTP incorporation (pmol).

The $F - F_0$ value is used to calculate the numbers of the incorporated NTP, where F and F_0 are the fluorescence intensities in the presence and absence of TPC4, respectively. For simple calculation, we use the fluorescence signal of the synthesized reporter RNA (28 nt) to estimate the fluorescence signal of the total reporter RNA products. The numbers of the reporter RNA products are measured to be 21.0 pmol according to the linear correlation equation (Fig. S2C). The numbers of the NTP incorporation are measured to be 588.5 pmol. Fold amplification is calculated to be 1961 (i.e., each individual TPC4 can initiate 1961 NTP incorporation during proximity ligation and transcription circuits).
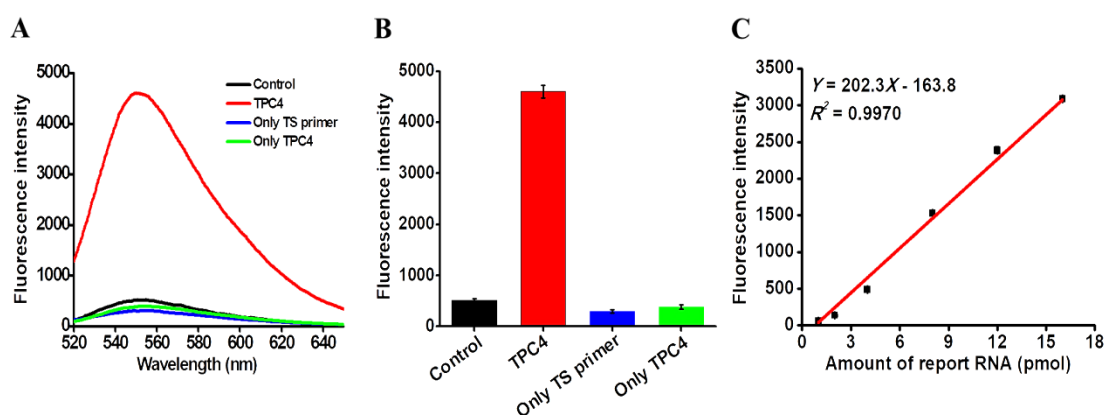

Fig. S2 (A) Fluorescence emission spectra in response to the control (black color), TPC4 (red color), only TS primer (blue color), and only TPC4 (green color). (B) Comparison of fluorescence intensity in response to the control (black color), TPC4 (red color), only TS primer (blue color), and only TPC4 (green color). (C) Variance of fluorescence intensity with different amounts of the synthesized reporter RNA. Error bars show the standard deviation of three independent experiments. The TPC4 concentration is 10 nM.

Optimization of the concentrations of ligation probes.

Hairpin probe contains the T7 promoter and part of transcription template, and linear probe

contains the other part of transcription template. With the telomeric extension product (TEP) as the template, hairpin probe and linear probe will be splinted together by Taq DNA ligase to initiate the proximity ligation-transcription circuits-powered cascade exponential amplifications. Thus, the concentrations of hairpin probe and linear probe should be optimized. We optimized the concentration of hairpin probe when the ratio of hairpin probe to linear probe is 1 : 1 (Fig. S3). When the concentration of hairpin probe increases from 20 nM to 400 nM, the $N - N_0$ value enhances and reaches a plateau at the concentration of 100 nM, followed by the decrease beyond 100 nM. Therefore, 100 nM hairpin probe and 100 nM linear probe are used in the subsequent research.
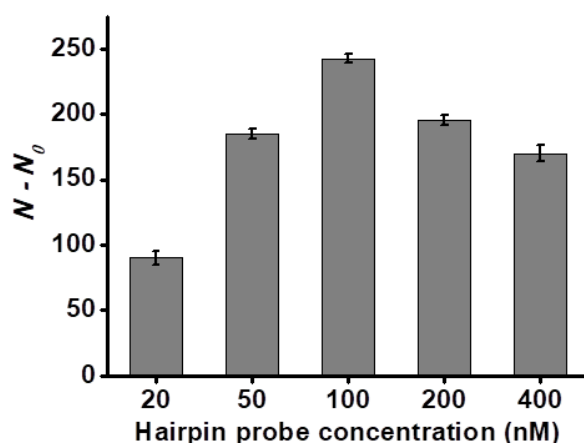

Fig. S3 Variance of $N - N_0$ value with different concentrations of hairpin probe. The N and N_0 represent the Cy5 counts in the presence and absence of telomerase, respectively. Error bars show the standard deviation of three independent experiments.

Optimization of the amount of T7 RNA polymerase.

In this strategy, T7 RNA polymerase is employed to mediate the transcription amplification circuit for the generation of numerous reporter RNAs that can be digested by DSN in reporter RNA/signal probe heteroduplexes to release large amounts of Cy5 molecules. Thus, T7 RNA

polymerase will directly affect the transcription efficiency, and the amount of T7 RNA polymerase should be optimized. As shown in Fig. S4, the $N - N_0$ value enhances with the increasing amount of T7 RNA polymerase from 0.5 U to 1 U, followed by the decreases beyond the amount of 1 U. Therefore, 1 U of T7 RNA polymerase is used in the subsequent research.
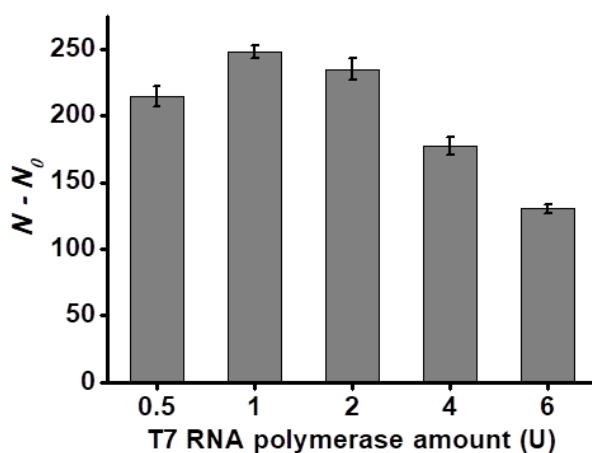

Fig. S4 Variance of $N - N_0$ value with different amounts of T7 RNA polymerase. The N and N_0 represent the Cy5 counts in the presence and absence of telomerase, respectively. Error bars show the standard deviation of three independent experiments.

Optimization of the reaction time and reaction temperature of proximity ligation-transcription amplification circuits.

To achieve the high efficiency of the proximity ligation and T7 transcription amplification circuits, the reaction time should be optimized to reduce the assay time. As shown in Fig. S5A, the $N - N_0$ value improves with the reaction time from 20 min to 60 min, and reaches the highest value at 60 min. Thus, 60 min of reaction time is used in the subsequent experiments. Taking into account the reaction temperature requirements of both enzymes and hybridization,[7] we investigated the effect of reaction temperature upon the $N - N_0$ value. The $N - N_0$ value enhances when the reaction temperature increases from 30 °C to 50 °C, followed by the derease beyond the reaction

temperature of 50 °C (Fig. S5B). The decrease of $N - N_0$ value may be ascribed to the denaturation

of duplex helixes at high temperature. Therefore, 50 °C is used as the appropriate reaction

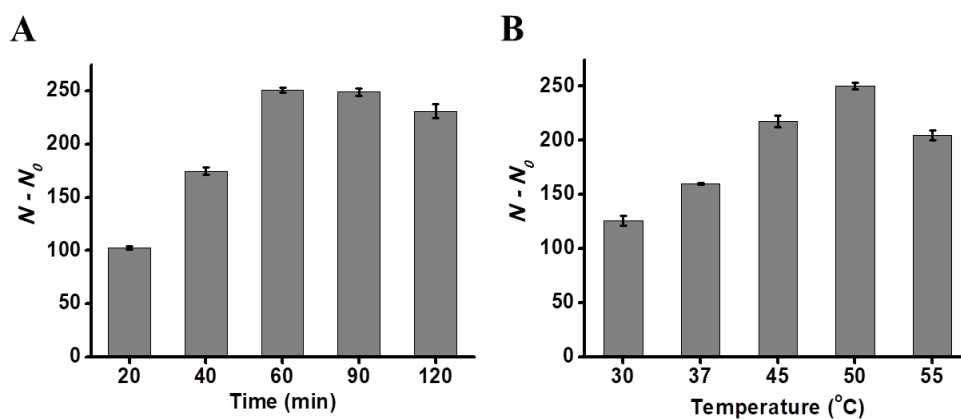temperature in the following experiments.

Fig. S5 (A) Variance of $N - N_0$ value with the reaction times. (B) Variance of $N - N_0$ value with

different reaction temperatures. The N and N_0 represent the Cy5 counts in the presence and

absence of telomerase, respectively. Error bars show the standard deviation of three independent

experiments.

Optimization of the amount of DSN.

In this assay, DSN is responsible for the cleavage of signal probes in reporter RNA / signal probe

heteroduplexes to dissociate Cy5 molecules[8] and simultaneously release the reporter RNA to

initiate continuous hybridization-cleavage-release for the generation of an enhanced Cy5 signal.

Therefore, the amount of DSN should be optimized. As shown in Fig. S6, the $N - N_0$ value

improves with the increasing amounts of DSN from 0.1 to 0.3 U, followed by the decrease beyond

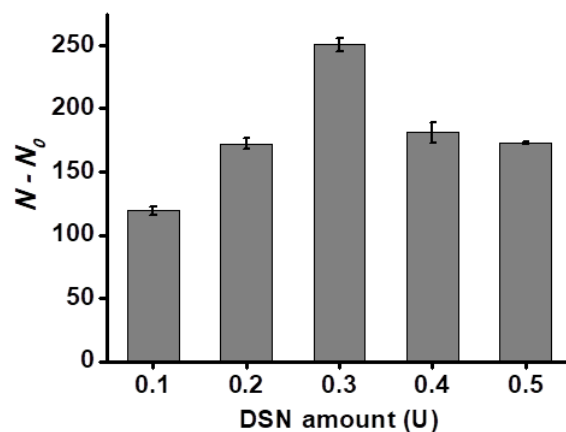the amount of 0.3 U. Therefore, 0.3 U of DSN is used in the subsequent experiments.

Fig. S6 Variance of $N - N_0$ value with different amounts of DSN. The N and N_0 represent the Cy5 counts in the presence and absence of telomerase, respectively. Error bars show the standard deviation of three independent experiments.

References

1. B. C. Yin, Y. Q. Liu and B. C. Ye, *J. Am. Chem. Soc.*, 2012, **134**, 5064-5067.

2. L.-j. Wang, Y. Zhang, C.-y. Zhang, *Analy. Chem.*, 2013, **85**, 11509-11517.

3. D.-M. Zhou, W.-F. Du, Q. Xi, J. Ge, J.-H Jiang, *Analy. Chem.*, 2014, **86**, 6763-6767.

4. L.-j. Wang, Y.-y. Lu, C.-y Zhang, *Chem. Sci.*, 2020, **11**, 587-595.

5. J. Hu, W. Liu, J. Wang, J.-G. Qiu, C.-y Zhang, *Analy. Chem.*, 2021, **93**, 6913-6918.

6. F. B. Dean, J. R. Nelson, T. L. Giesler, R. S. Lasken, *Genome Res.*, 2001, **11**, 1095-1099.

7. Y.-p. Zeng, J. Hu, Y. Long and C.-y. Zhang, *Analy. Chem.*, 2013, **85**, 6143-6150.

8. L.-j. Wang, L. Liang, B.-j. Liu, B. Jiang and C.-y. Zhang, *Chem. Sci.*, 2021, **12**, 5544-5554.