Supporting Information

Headpiece-assisted DNA data storage in solution and solid

Chunjie Hu,[a†] Qingya Wen,[a†] Qiuyang Lai,[a] Ziyi Xie,[a] Kaiyue Zhang,[a] Lu Zhou,[a] and Zhi-bei Qu*[a]

[a]*Department of Medicinal Chemistry, School of Pharmacy, Fudan University, Shanghai 201203, China.*

† These authors contributed equally.

*Email: quzhibei@fudan.edu.cn

DNA and primer sequences

DNA headpiece HP-NH2(5′-/5Phos/GAGTCA/iSp9/iUniAmM/iSp9/TGACTCCC-3′, Figure 1) was obtained from Biosearch Technologies, Novato, CA.

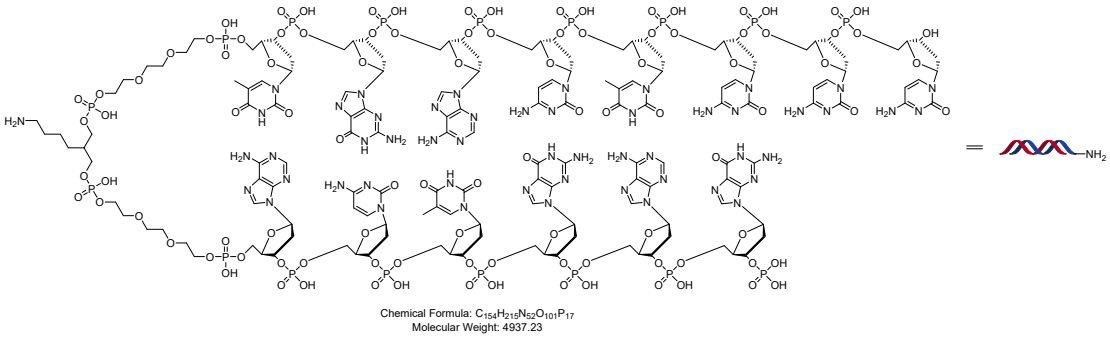


Figure S1. Headpiece: 5’-/5phos/GAGTCA/iSp9/iUniAmM/iSp9/TGACTCCC-3’, MW = 4937.2

Code1-1：

5’-AATCCCTTTCTCAGTGGGTAATTTCTGTCCTAAAGG-3’

5’-TTTAGGACAGAAATTACCCACTGAGAAAGGGATTGG-3’

Code1-2：

5’- AATCCCTTTCTCAGTGGGTAATTTCTGTCCTAAAGG -3’

5’-TTTAGGACAGAAATTACCCACTGAGAAAGGGATT-3’

Code2-1:

5’-AAATCGATGTGAATAAATACTCCCTTTCTCAGTGGGTAATTTCTGTCCT

AAAAGAAAGAAAGAAAGAACATAGCTAATGAGCAGC-3’

5’-GCTGCTCATTAGCTATGTTCTTTCTTTCTTTCTTTTAGGACAGAAATTAC

CCACTGAGAAAGGGAGTATTTATTCACATCGATTTGG-3’

Code2-2:

5’-AAATCGATGTGAATAAATACTCCCTTTCTCAGTGGGTAATTTCTGTCCT

AAAAGAAAGAAAGAAAGAACATAGCTAATGAGCAGCGG-3’

5’-GCTGCTCATTAGCTATGTTCTTTCTTTCTTTCTTTTAGGACAGAAATTAC

CCACTGAGAAAGGGAGTATTTATTCACATCGATTTGG-3’

Code2 primer

Forward primer:

5’-CAAGTCCGCCACCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTT

CTTCAAGGACGACGGCAACTATGACTCCCAAATCGATGTGAATAAATAC-

3’

Reverse primer:

5’-TGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCAC

CTCGGCGCGGGTCTTGCTGCTCATTAGCTATGTTCTTTCTT -3’

HP biotin-Code2 primer

Forward primer:

5’-CAAGTCCGCCACCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTT

CTTCAAGGACGACGGCAACTATGACTCCCAAATCGATGTGA-3’

Reverse primer:

5’-TGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCAC

CTCGGCGCGGGTCTTGCTGCTCATTAGCTATGTTCTTTCTT -3’

HP biotin-Code2-HP biotin primer

Forward primer:

5’-TTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT

TCAAGGACGACGGCAACTACTGACTCCCAAATCGATGTGA-3’

Reverse primer:

5’-GATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTC

ACCTCGGCGCGGGTCTTTGACTCCCGCTGCTCATTAGCTAT -3’

qPCR Forward primer:

5’-TAAATACTCCCTTTCTCAGTGGGT-3’

qPCR Reverse primer:

5’-GCTGCTCATTAGCTATGTTCT-3’

Abbreviated Terms

DMA: dimethylacetamide

HATU: 1-bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b] pyridinium 3-oxid hexafluorophosphate

DIPEA: N, N-diisopropylethylamine

Biotin: 5-((3aS,4S,6aR)-2-Oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoic acid

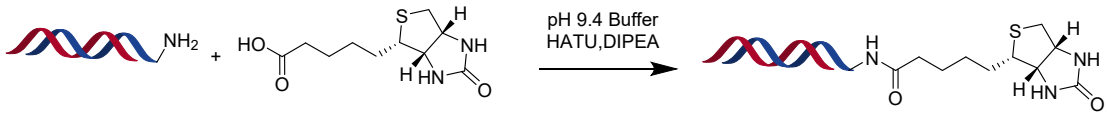
pH=9.4 Buffer: 500 mM sodium borate in water

N, N-dimethylacetamide (DMAc) and EtOH were purchased from Sigma-Aldrich. HATU (CAS: 148893-10-1), N, N-Diisopropylethylamine (DIPEA), NaCl were purchased from TCI. The ddH_2O was obtained by passing the Milli-Q Direct. The buffer was purchased from Bidepharm. Salmon sperm DNA was purchased from Thermo Fisher Scientific.

Analysis for DNA Compounds

The conversion rate was used to calculate the effect of one-step reaction according to the proportion of peak height of mass spectrum. Analysis was performed by HPLC/ESI-MS. After reaction, an aliquot of the reaction mixture solution was diluted (typically a 1 μL aliquot diluted with 20 μL of water) for LC/MS. Reversephase chromatography column (XBridge C18, 3.5um 3.0*50mm) with monitoring at 260 nm. Solvent A: 2.25% hexafluoroisopropanol(v/v) and 0.114% Triethylamine(v/v) in deionized water. Solvent B: 2.25% hexafluoroisopropanol(v/v) and 0.114% Triethylamine(v/v) in 90/10 methanol/water. Flow rate: 0.450 mL/min; Time: 6.00 min.

Experimental Procedure for on-DNA Reaction



Materials.

Headpiece: 1 mM in water

Biotin-COOH: 200 mM in DMA

DIPEA: N,N-Diisopropylethylamine 200 mM in DMA

HATU: 2-(7-Azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate 200mM in DMA

pH 9.4 Buffer: 500 mM sodium borate in water (pH 9.4)

Procedure.

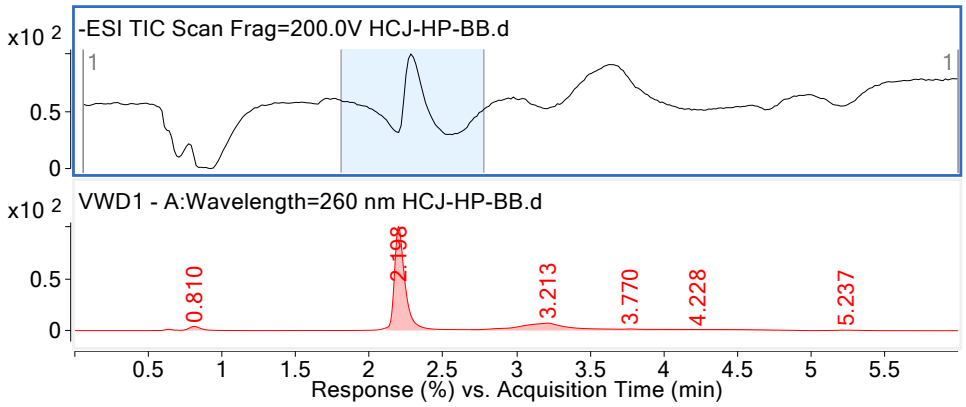
1) Mixing Biotin-COOH with equal volume HATU and DIPEA to prepare the solution I.

2) To the headpiece solution (200 nmol, 200 μL), was added 200 μL pH 9.4 sodium borate buffer solution, 40 equiv. of solution I (120 μL), and mix.

3) React at room temperature for 2 hours.

4) The product was obtained by ethanol precipitation as described above.

Mass Spectrum of on-DNA Product HP biotin



conversion: 96 %

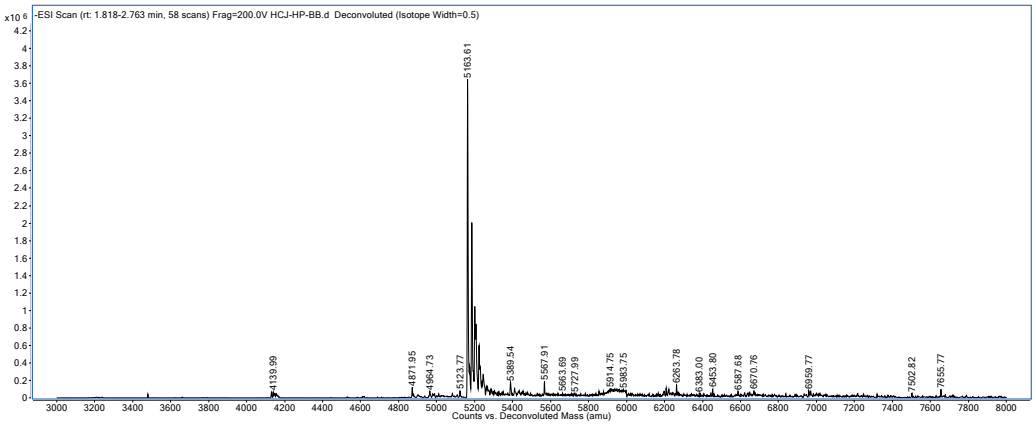


Figure S2. Deconvoluted mass spectra for biotinylated HP-Code.

Conversion Determination

The conversions were determined by examination of the UV traces, TIC traces and MS abundance of the LCMS chromatograms.

Ethanol Precipitation for DNA

To the reaction vessel, 10% volume of 5M NaCl solution and 2.5 volumes of cold ethanol (stored at -80 °C) were added. The solution was mixed and stood at -80 °C for 30 min. Then the reaction vessel was centrifuged at 10,000 RPM for 30 minutes. After that, the supernatant was removed and the pellet was dried by vacuum.HPLC Purification for DNA-linked compounds Preparative-HPLC was performed using InfinityLab Poroshell 120 EC-C18, 4.6*150mm 4-Micron, flow rate: 1 mM/min;

Solvent A: 50 mM triethylamine acetate aqueous solution, pH = 7.5; Solvent B: 1% water in acetonitrile; UV collection: at 260 nm.

Quantification for DNA

The on-DNA Intermediate were purified by a 3 K Spinfilter tube after centrifugation for 30 minutes at 4 ℃ in a microcentrifuge at 10 K rpm three times. Use the Thermo ScientificTM NanoDrop TM One to quantify concentration of products under the dsDNA mode.

Oligonucleotide Ligation Experiment

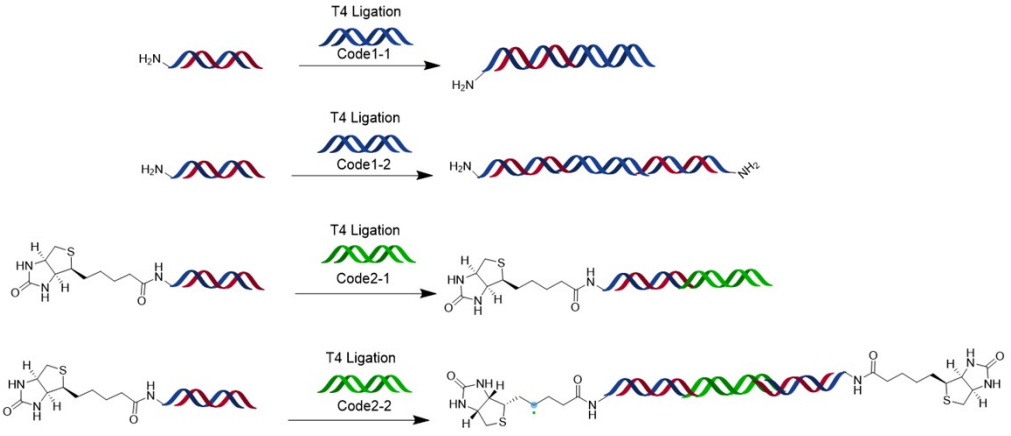


Figure S3. DNA Code ligation procedure.

Procedure.

1) For the ligation producer, the reaction headpiece was ligated with Code1-1 (100μM,1.5 equiv.) in the presence of T4 ligase (300U) for 16 h at 16℃.

2) For the ligation producer, the reaction HP-biotin was ligated with Code2-1 (100μM,1.5 equiv.) in the presence of T4 ligase (300U) for 16 h at 16℃.

3) For the ligation producer, the reaction headpiece was ligated with Code1-2 (100μM,3 equiv.) in the presence of T4 ligase (300U) for 16 h at 16℃.

4) For the ligation producer, the reaction HP-biotin was ligated with Code2-2 (100μM,3 equiv.) in the presence of T4 ligase (300U) for 16 h at 16℃.

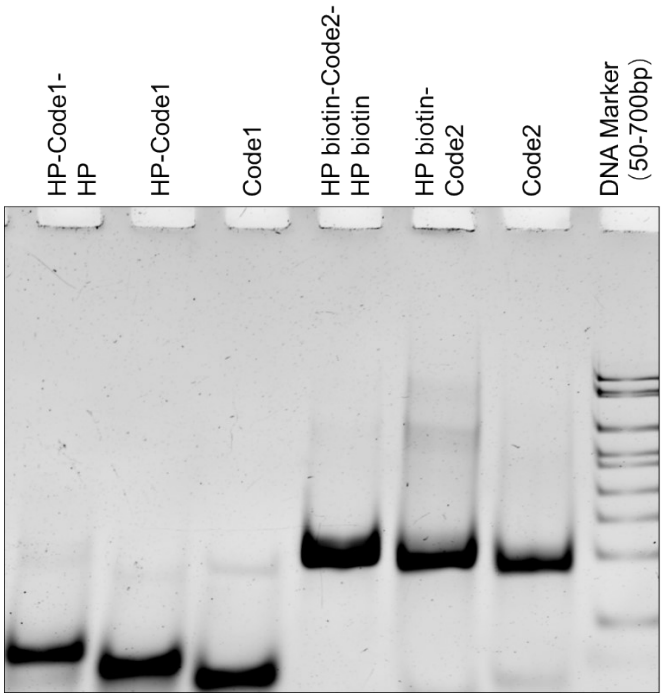


Figure S4. Construction of HP-conjugated DNA strands using 5% polyacrylamide electrophoresis gel.

Encapsulation in silica beads.

Silica beads preparation: 18 mL ethanol, 0.8 mL 25% (wt/wt) ammonia solution and 0.5 mL H_2O were mixed evenly and added 0.8 mL TEOS. Stirred the mixture at 900 r.p.m. for 6 h at room temperature. The particles were collected by centrifugation and washed with isopropanol. Resuspended the particles with 20 mL isopropanol (50 mg/ mL).

Ammonium functionalization of silica beads: 10 μL TMAPS (50% (wt/wt) in methanol) solution was added to 1 mL 50 mg mL−1 SiO_2 particles. Stirred the mixture at 900 rpm. for 12 h at room temperature. The particles were collected by centrifugation and washed with isopropanol. Resuspended the particles with 1 mL isopropanol.

DNA encapsulation: 35 μL functionalized particle solution and 700 μL water were mixed evenly, and then added 50 μg/mL 320 μL DNA solution. Mixed well and waited for DNA to be fully adsorbed. The particles were collected by centrifugation and washed with deionized water. Resuspended the particles in 500 μL H_2O. Added 0.5 μL 50% (wt/wt) TMAPS solution and then added 0.5 μL TEOS with stirring at 900 rpm. After 4 h, added 4 μL TEOS with stirring successively. Stirred the mixture for 4 d at 900 rpm. at room temperature. The particles were collected by centrifugation and washed with deionized water. Finally, the particles was freeze dried and generated white powder.

DNA release: Centrifuge 100 μl of previously synthesized DNA/SiO_2 particles at 21,500g for 3 minutes at room temperature and discard the supernatant. Add 300 μL of the NH_4FHF/NH_4F etching solution for 40 minutes. Add Streptavidin Agarose Resin retain 1h for maximum adsorption. After centrifugation to discard the supernatant, it is followed by two washes with water. After adding an appropriate amount of water to the resin species, take 5 μL of the resin solution and enter the PCR program.

DNA PCR amplification: Each sample well was filled with a total volume of 50 μL containing: 5 μL sample volume, 25 μL 2 × Phanta Max Master Mix, 18 μL mQ water,

1 μl (1 μM) forward primer and 1 μl (1 μM) reverse primer. The protocol consisted of a 1 minutes activation step at 95°C, 3-step amplification (33 cycles, 95°C for 15 s, 56°C for 30 s and 72°C for 20 s), extend 5 minutes at 72°C. Purify PCR products by using FastPure Gel DNA Extraction Mini Kit according to the manufacturer's instructions. The purified product undergoes first-generation Sanger sequencing.

DNA extraction and purification: HP biotin-Code1 at concentrations of 1 ng/ μL, 10 ng/ μL and 100 ng/ μL were recovered by cold ethanol precipitation and streptavidin agarose resin respectively. Measure the DNA concentration with the Qubit dsDNA HS assay.

Environmental DNA damage experiments.

Test 1 is for the solution of the synthesized Code1-1, HP-Code1, HP-Code1-HP. Test 2 is for unencapsulated Code1-1, encapsulated HP-Code1, encapsulated HP-Code1-HP. The qPCR and sequencing for the unencapsulated and encapsulated DNA were subjected to 120 °C for 12 min and enzymes, acid, base, ROS conditions in Code2 test1. The qPCR for the DNA pool also consisted of a 3-step amplification protocol (95°C for 10 s, 54°C for 30 s and 72°C for 30 s), and each sample was analyzed in duplicates.

Heating test1: The temperature treatment stability assay is designed to evaluate the quality for introduction of HP process. To test the protective properties for introduction of HP, the following solutions of the synthesized Code1-1,HP-Code1,HP-Code1-HP are treated with a solution for 120 ℃ 10 minutes, Agarose gel electrophoresis characterizes its loss.

Heating test2: The temperature treatment stability assay is designed to evaluate the quality for introduction of HP process and encapsulation. To evaluate the quality of the encapsulation process. To test the protective properties of the particles, they are treated with 120 ℃ 15 minutes. The encapsulated DNA is released. Agarose gel electrophoresis characterizes its loss.

ROS test1: The radical treatment stability assay is designed to evaluate the quality for introduction of HP process. To test the protective properties for introduction of HP, they are treated with a solution producing ROS. Add the following solutions to 10 μL of the synthesized Code1-1,HP-Code1,HP-Code1-HP, Compound Volume

Concentration, l-Sodium ascorbate 1.25 μL 20 mM,H_2O_2 6.25 μL 20 mM,$CuCl_2$ 8.75 μL 500 μM. Wait for 7 minutes. Add 10 μL of 100 mM EDTA for quench the reaction. And then load the sample for agarose gel electrophoresis to characterize its loss.

ROS test 2: The radical treatment stability assay is designed to evaluate the quality of the encapsulation process and encapsulation. To test the protective properties of the particles, they are treated with a solution producing ROS. The concentrations of the reactants are chosen so that they would completely disintegrate nonencapsulated DNA. l-Ascorbic acid 2.5 μL 20 mM,H_2O_2 12.5 μL 20 mM,$CuCl_2$ 17.5 μL 500 μM,Wait for 10 min. Add 17.5 μL of 100 mM EDTA and 20 μL of NH_4FHF/NH_4F etching solution. The encapsulated DNA is released. And then load the sample for agarose gel electrophoresis to characterize its loss.

DNaseI test 1: The DNaseI treatment stability assay is designed to evaluate the quality for introduction of HP process.To test the protective properties for introduction of HP, they are treated with 1μL DNaseI test，wait for 15 min at 37℃，and then load the sample for agarose gel electrophoresis to characterize its loss.

ExonucleasesIII test 1: The ExonucleasesIII treatment stability assay is designed to evaluate the quality for introduction of HP process.To test the protective properties for introduction of HP, they are treated with 1μL ExonucleasesIII test，wait for 15 min at 37℃，and then load the sample for agarose gel electrophoresis to characterize its loss.

DNaseI test 2: The DNaseI treatment stability assay is designed to evaluate the quality for introduction of HP process and encapsulation. to evaluate the quality of the encapsulation process. To test the protective properties of the particles, they are treated with 10 μL DNaseI test，wait for 60 minutes at 37℃. The encapsulated DNA is released. And then load the sample for agarose gel electrophoresis to characterize its loss.

ExonucleasesIII test 2: The ExonucleasesIII treatment stability assay is designed to evaluate the quality for introduction of HP process and encapsulation. To evaluate the quality of the encapsulation process. To test the protective properties of the particles,

they are treated with 1 μL ExonucleasesIII test， wait for 20 minutes at 37℃. The encapsulated DNA is released. And then load the sample for agarose gel electrophoresis to characterize its loss.

Acid test 1: The acid treatment stability assay is designed to evaluate the quality for introduction of HP process. Add 20 μL of 1 M HCl to the 10 μL DNA solution, wait for 25 minutes, add water to dilute, and then load the sample for agarose gel electrophoresis to characterize its loss.

Acid test 2: The acid treatment stability assay is designed to evaluate the quality for introduction of HP process and encapsulation. Add 30 μL of 1 M HCl to the 10 μL DNA solution, wait for 40 minutes, The encapsulated DNA is released. Add water to dilute these samples, and then load the sample for agarose gel electrophoresis to characterize its loss.

Base test 1: The base treatment stability assay is designed to evaluate the quality for introduction of HP process. Add 20 μL of 4 M NaOH to the 10 μL DNA solution, wait for 60 minutes, add water to dilute, and then load the sample for agarose gel electrophoresis to characterize its loss.

Base test 2: The base treatment stability assay is designed to evaluate the quality for introduction of HP process and encapsulation. Add 30 μL of 4 M NaOH to the 10 μL DNA solution, wait for 80 minutes, The encapsulated DNA is released. Add water to dilute, and then load the sample for agarose gel electrophoresis to characterize its loss.

External DNA contamination experiment.

Excess salmon sperm was added to Code2 and HP-Code2 overnight, and the salmon sperm DNA mixed with salmon sperm DNA was entered into the PCR program for first-generation sequencing, and HP-Code2 was extracted by biotin streptavidin binding method, and then PCR program was followed by first-generation sequencing.

5% Polyacrylamide gel electrophoresis. 9.4 mL of water,2.5 mL of acrylamide:bis-acrylamide (Acr:Bis=29:1),3 mL of 5×TBE,110 μL of 10% ammonium persulfate, and 10 μL of TEMED were placed in a 50 mL centrifuge tube. Inject into a glass plate, insert a mold and allow it to solidify, and remove the comb. In polyacrylamide gel electrophoresis, we used 1×TBE as the buffer, and then set the operating parameters of

the electrophoresis apparatus to a voltage of 110 V for 30 minutes. Remove the gel at the end of the run. We used the bubble staining method to stain the gels. Add 15 μL of 4S Gelred and 5 mL of 1M NaCl to 45 mL of H_2O to make the stain and place the gel in the stain. After staining for 30 minutes, the gel was placed into the gel imaging system for imaging to see the relative position of the bands.

2% **Agarose gel electrophoresis.** Measure 30 mL of 1×TAE with a graduated cylinder and weigh 600 mg of agarose in an Erlenmeyer flask. The Erlenmeyer flask is then placed in the microwave and heated at high temperature until the agarose is completely dissolved. Wait until the heated solution has cooled to a palpable temperature, then add 3 μL of 4S Gelred to the solution. The mixture is then quickly poured into the gel mold and a plastic comb is inserted into the sample tank. Allow to allow the gel solution to solidify. Gels of 2% agarose (dissolved in 1× TAE buffer) was used for electrophoresis. A sample was loaded in the gel and electrophoresis separation was performed with a Bio-Rad electrophoresis system at 110 V for 20 min. The gels were photographed with a digital camera.

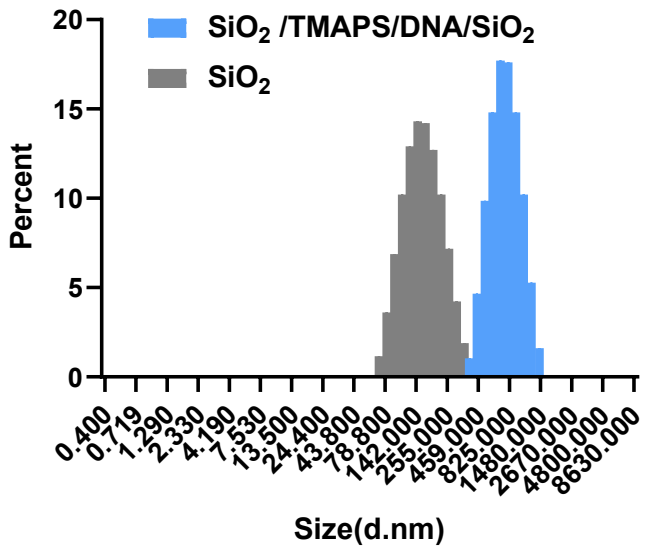


Figure S5. Size distribution of silica beads (SiO_2), SiO_2/TMAPS/DNA/SiO_2.

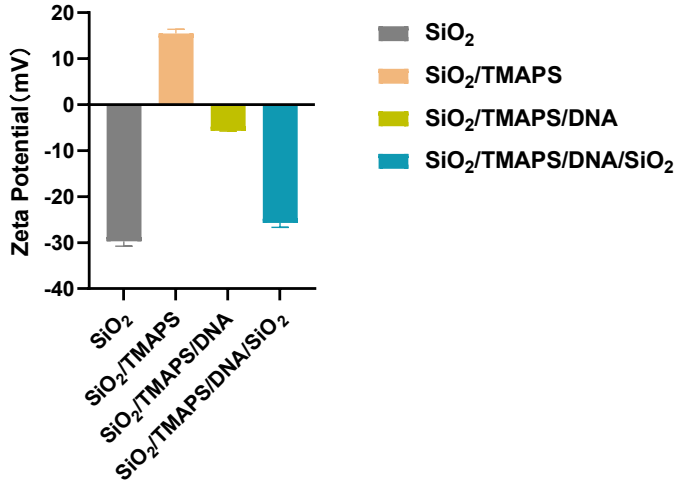


Figure S6. Zeta potential of silica beads (SiO_2), SiO_2/TMAPS, SiO_2/TMAPS/DNA, SiO_2/TMAPS/DNA/SiO_2.

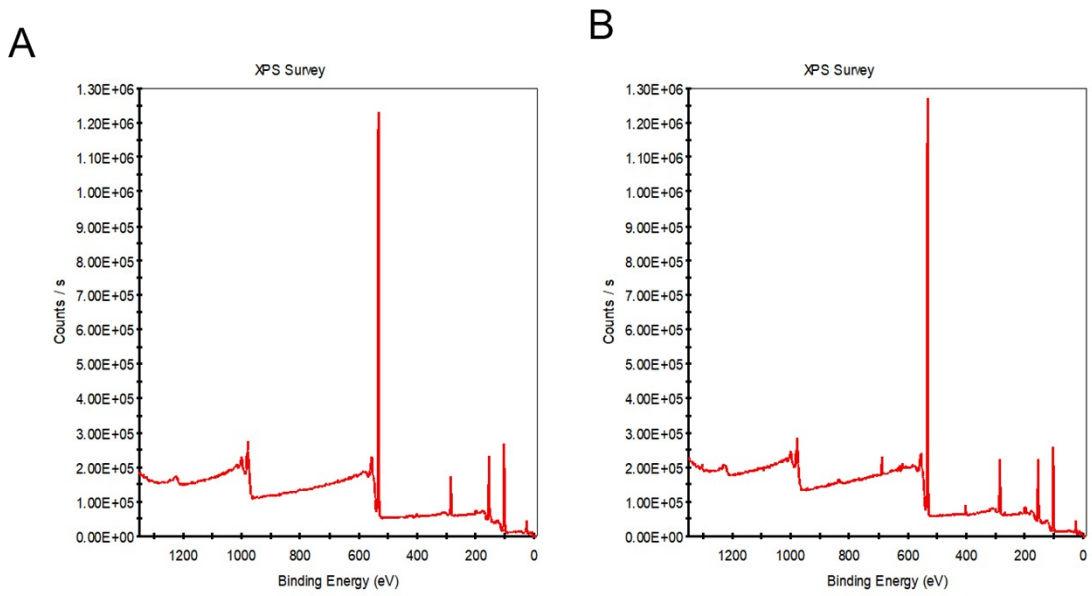


Figure S7. X-ray photoelectron spectra (XPS) of core/shell DNA-silica particles. (A) SiO_2 (B) SiO_2-DNA.

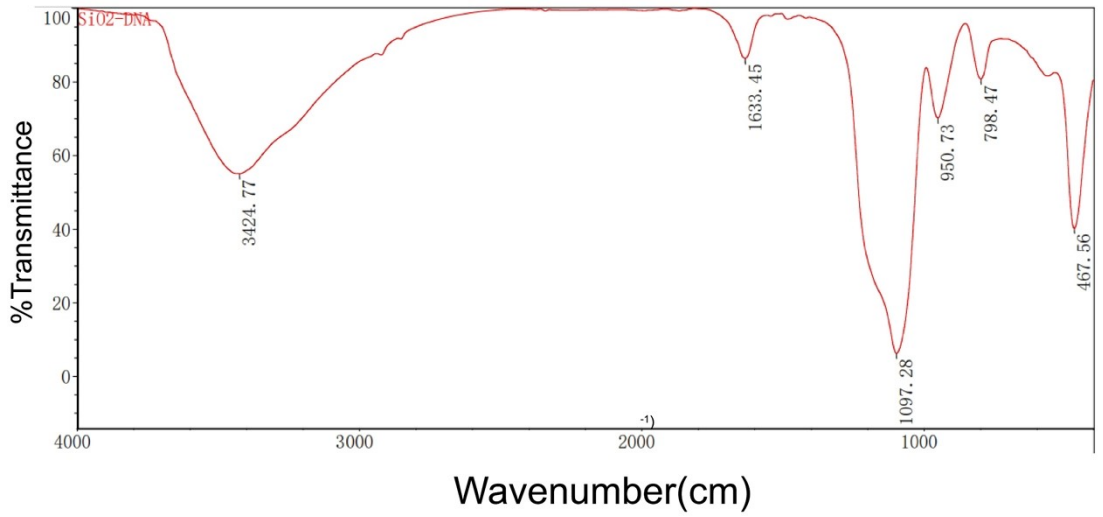


Figure S8. FTIR spectra of SiO_2-DNA.

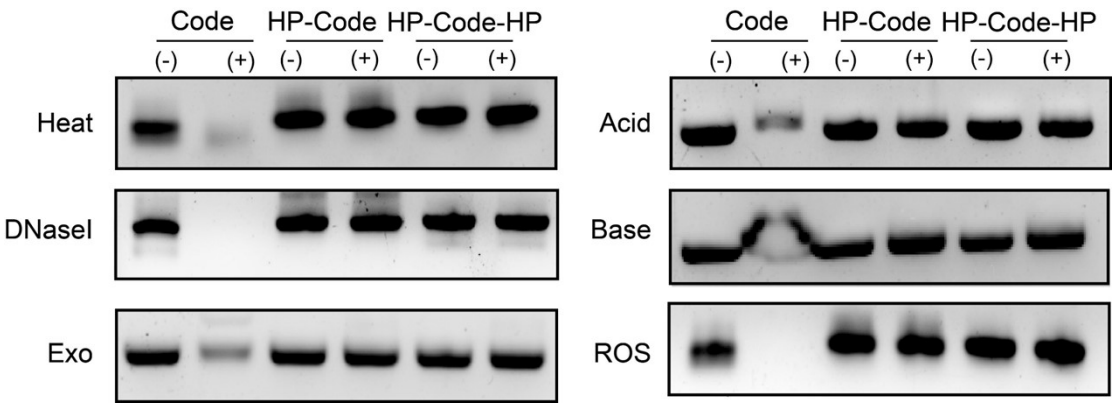


Figure S9. Unencapsulated Code and silica encapsulated HP-Code1, HP-Code1-HP agarose gel after base, acid, enzyme treatments.

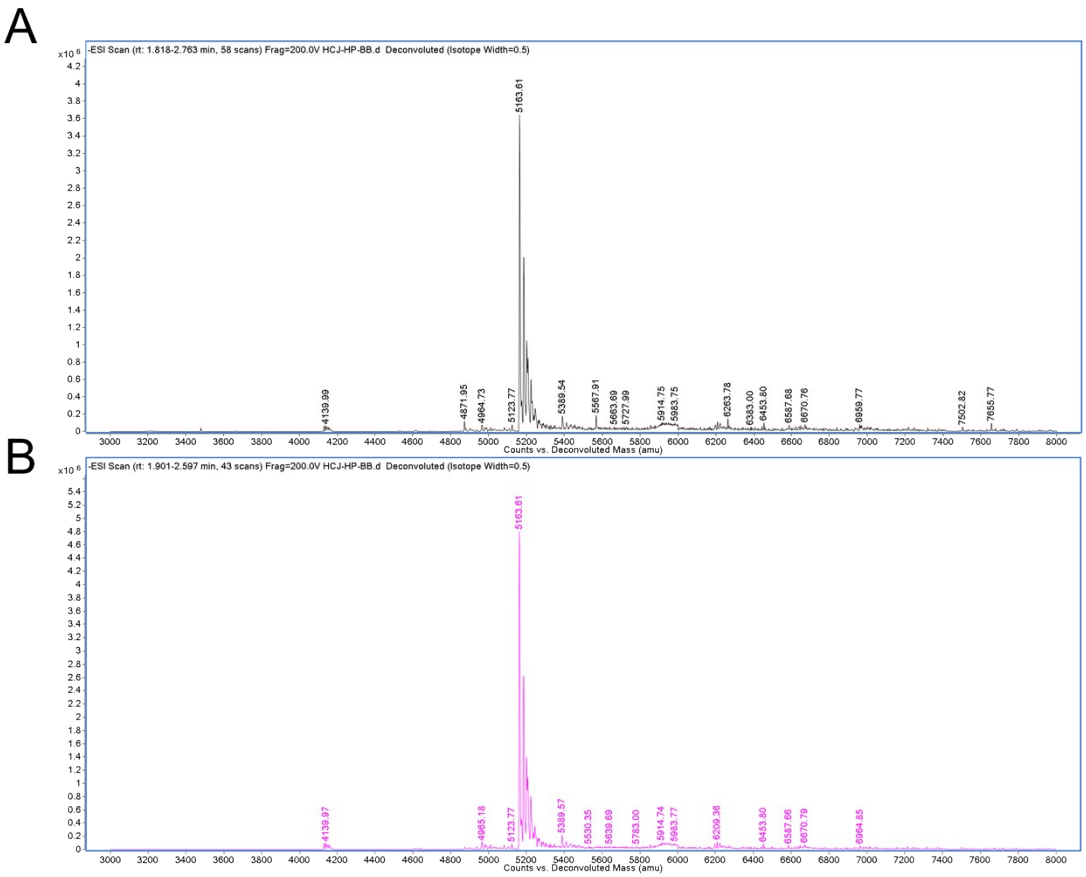


Figure S10. MALDI-TOF-MS spectroscopy: (a) the released HP biotin DNA strands. (b) the released biotin-HP DNA strands.

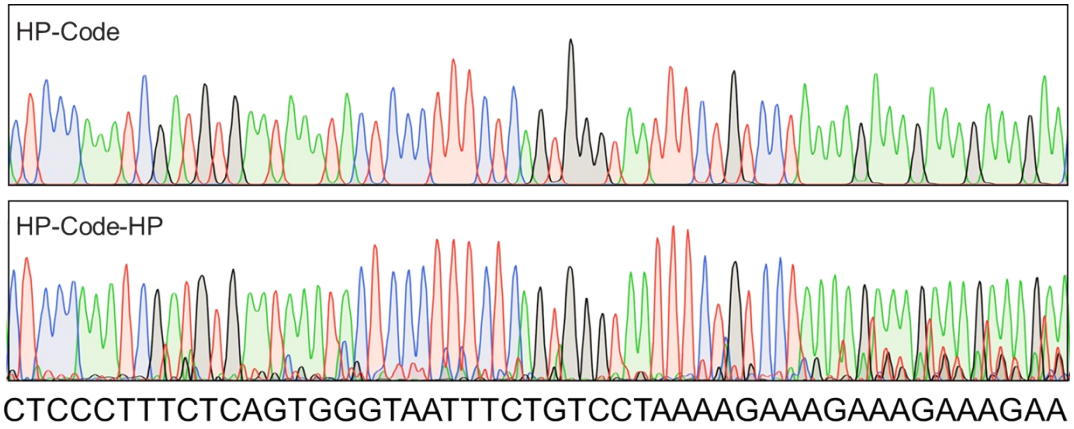


Figure S11. Sequencing results of HP-Code2 and HP-Code2-HP Under the same system at the same amount of templates.

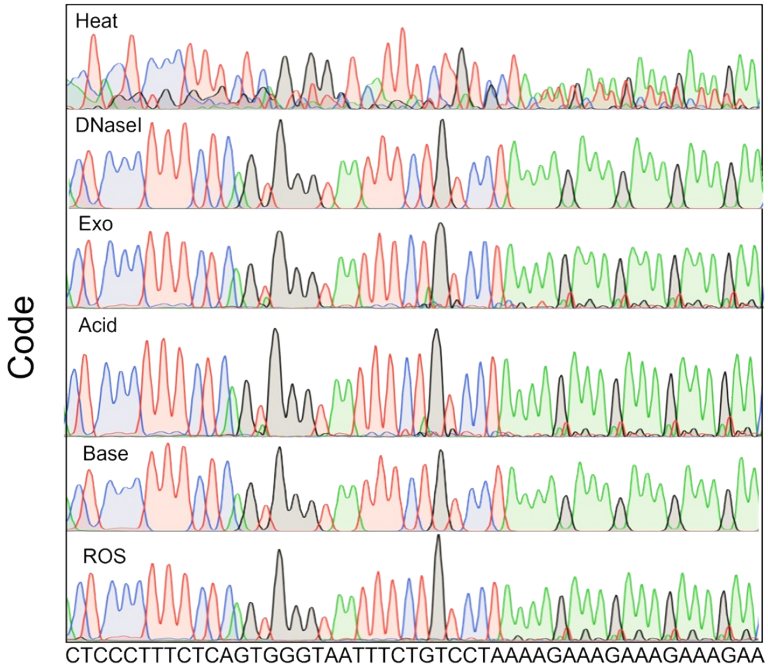


Figure S12. Sequencing results of unencapsulated Code2 after being treated with 120 ℃ heating, enzymes, acid, base, ROS conditions etc.

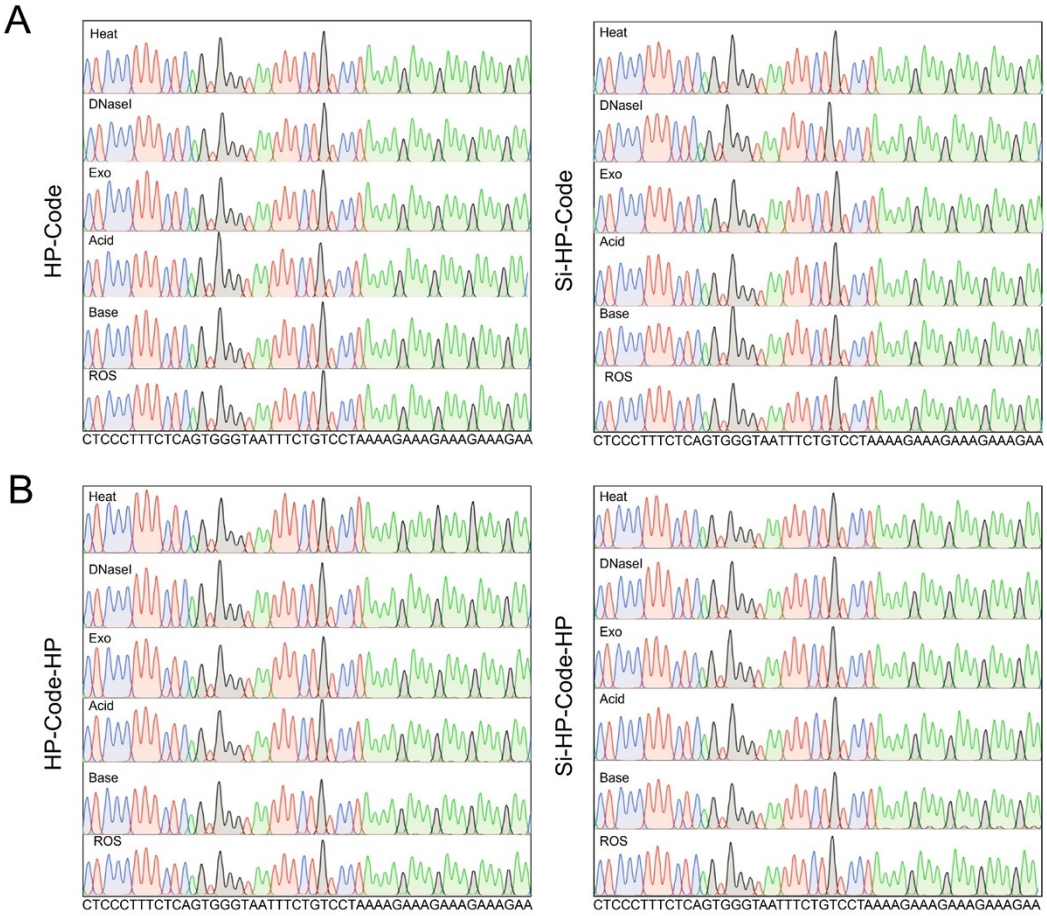


Figure S13. Sequencing results of unencapsulated and encapsulated HP-Code2 and HP-Code2-HP after being treated with 120 ℃ heating, enzymes, acid, base, ROS conditions etc.

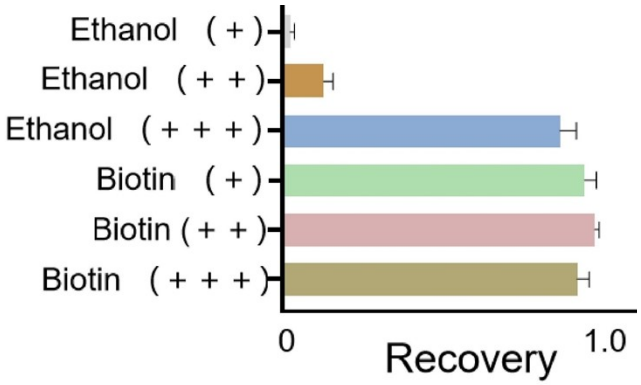


Figure S14. Recoveries of encoded DNAs using cold ethanol condensation and biotin-streptavidin affinity enrichment in various DNA concentration. +: 1 ng/μL, ++: 10 ng/μL, +++: 100 ng/μL.

Table S1. Zeta potential of SiO_2 microsphere, SiO_2/TMAPS, SiO_2/TMAPS/DNA, SiO_2/TMAPS/DNA/SiO_2 .

Sample Name	SiO_2	SiO_2/TMAPS	SiO_2/TMAPS/DNA	SiO_2/TMAPS/DNA/SiO_2
Zeta potential (mV)	-28.7	14.8	-5.76	-25.1
	-30.8	15.2	-5.69	-25.3
	-29.7	16.5	-5.65	-26.8
Average (mV)	-30.1	15.5	-5.70	-25.7

Table S2. Comparison of the properties of the four storage forms of DNA data storage systems with conventional one in this work.

	Code	HP-Code	HP-Code-HP	Si-HP-Code	Si-HP-Code-HP
stability	10%	15%	27%	47%	69%
recovery	27%	85%	96%	63%	75%
amplification	99%	95%	94%	92%	87%
accuracy	35%	57%	45%	87%	54%