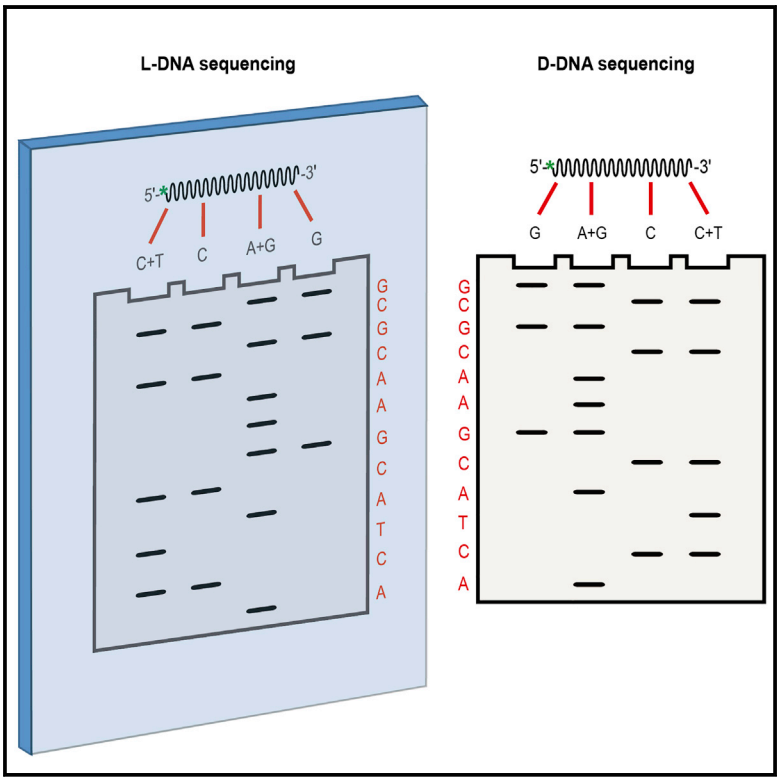
Brief Communication

Cell Chemical Biology

Sequencing Mirror-Image DNA Chemically

Graphical Abstract



Highlights

- Developing a practical method for L-DNA sequencing
- Applying the classic Maxam-Gilbert sequencing method to mirror-image biology systems
- Sequencing nuclease-resistant L-DNA aptamers for therapeutic applications

Authors

Xianyu Liu, Ting F. Zhu

Correspondence

tzhu@tsinghua.edu.cn

In Brief

A key technical hurdle in the development of mirror-image biology systems is the lack of a practical L-DNA sequencing approach. Liu and Zhu developed a method for sequencing mirror-image DNA by adopting the classic Maxam-Gilbert approach through cleavage of end-labeled L-DNA by achiral chemicals.

Liu & Zhu, 2018, Cell Chemical Biology *25*, 1151–1156
September 20, 2018 © 2018 Elsevier Ltd.
https://doi.org/10.1016/j.chembiol.2018.06.005

CellPress

Cell Chemical Biology
Brief Communication

Sequencing Mirror-Image DNA Chemically

Xianyu Liu[1] and Ting F. Zhu[1,2,*]
[1]School of Life Sciences, Tsinghua-Peking Center for Life Sciences, Center for Synthetic and Systems Biology, Ministry of Education Key Laboratory of Bioinformatics, Tsinghua University, Beijing 100084, China
[2]Lead Contact
*Correspondence: tzhu@tsinghua.edu.cn
https://doi.org/10.1016/j.chembiol.2018.06.005

SUMMARY

The development of mirror-image biology systems faces a crucial barrier of lacking an L-DNA sequencing technique. Here, we developed a practical method for sequencing mirror-image DNA by adopting the Maxam-Gilbert sequencing approach, through which specific nucleobases in an end-labeled L-DNA are cleaved by achiral chemicals. This technique may facilitate the therapeutic application of nuclease-resistant L-aptamer drugs, and bring the vision of building an alternative, mirror-image self-replicating system closer to reality.

INTRODUCTION

Mirror-image biology systems hold promise for many exciting applications in biomedicine. For example, nuclease-resistant L-DNA aptamers have been selected to bind natural protein targets, and therefore may become a new category of plasma-stable L-aptamer drugs (Klussmann et al., 1996; Nolte et al., 1996; Williams et al., 1997; Yatime et al., 2015). Most of the L-DNA aptamers are produced by solid-phase oligonucleotide (oligo) synthesis, during which synthesis errors do occur, particularly for long sequences. Our inability to sequence the synthetic L-DNA aptamer drugs to ensure their quality has hindered their safe clinical use. In addition, a direct selection scheme for L-DNA aptamers against biological targets through a proposed mirror-image systematic evolution of ligands by exponential enrichment (MI-SELEX) approach is on the horizon (Xu et al., 2017; Jiang et al., 2017), but also requires the sequencing of enriched L-aptamer sequences from a randomized L-DNA pool. The rapid development of this field has created a need for a practical method to sequence mirror-image DNA.

Despite the remarkable advancements in DNA sequencing technologies, no method for sequencing L-DNA has been reported. Most of the commonly used sequencing-by-synthesis methods are currently unsuitable because they require a particular polymerase capable of incorporating labeled L-di-deoxynucleotide triphosphates (L-ddNTPs) or L-deoxyribonucleotide triphosphates (dNTPs). Although a couple of mirror-image polymerase systems based on enzymes small enough for total chemical synthesis, such as the African swine fever virus polymerase X (ASFV pol X) (Wang et al., 2016) and the *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4) (Xu et al., 2017; Jiang et al., 2017; Pech et al., 2017), have been developed, they still suffer from issues such as poor fidelity or inability to incorporate labeled ddNTPs or dNTPs. While the next-generation nanopore DNA sequencing approach could be applied to sequencing mirror-image DNA in principle, it also requires a particular D-amino acid polymerase or helicase (which is not currently available) to help slow down the passing DNA (Derrington et al., 2015).

Maxam-Gilbert sequencing was the first widely used DNA sequencing method, which relies on non-enzymatic chemical modifications of specific nucleobases, followed by strand scission adjacent to the modified sites by treatment with strong alkali (Banaszuk et al., 1983; Friedmann and Brown, 1978; Maxam and Gilbert, 1977; Rosenthal et al., 1985; Rubin and Schmid, 1980). We reasoned that the chemicals used for Maxam-Gilbert sequencing are achiral and thus should also work with L-DNA. Here, we show that the sequences of end-labeled mirror-image DNA molecules, including several L-DNA oligos and a 55-nucleotide (nt) L-DNA aptamer, can be reliably determined using this approach.

RESULTS

Chemical Sequencing of Short L-DNA Molecules

We first tested the chemical sequencing approach on a 12-nt L-DNA oligo with fluorescein amidite (FAM) label at the 5′ end (Key Resources Table). We used fluorescent end-labeling instead of radioactive labeling, in part because it is impractical to radioactively label L-DNA without a mirror-image polynucleotide kinase (Wang et al., 2016). The C + T reaction was carried out by hydrazine at 45°C, the C-specific reaction with hydroxylamine hydrochloride (pH 6.0) at room temperature, the A + G reaction by formic acid at room temperature, and the G-specific reaction with methylene blue under UV irradiation, followed by strand scission adjacent to the modified sites by treatment with strong alkali (Method Details and Figure 1A). Here, because the UV absorption spectrum of methylene blue is different from that of FAM (Figure S1), we used UV at 254 nm to specifically excite methylene blue. Additionally, we used unlabeled *Escherichia coli* genomic DNA as carrier DNA during the sequencing (Pichersky, 1996). The final products were analyzed by 20% polyacrylamide gel electrophoresis (PAGE) in 8 M urea (Figure 1B).

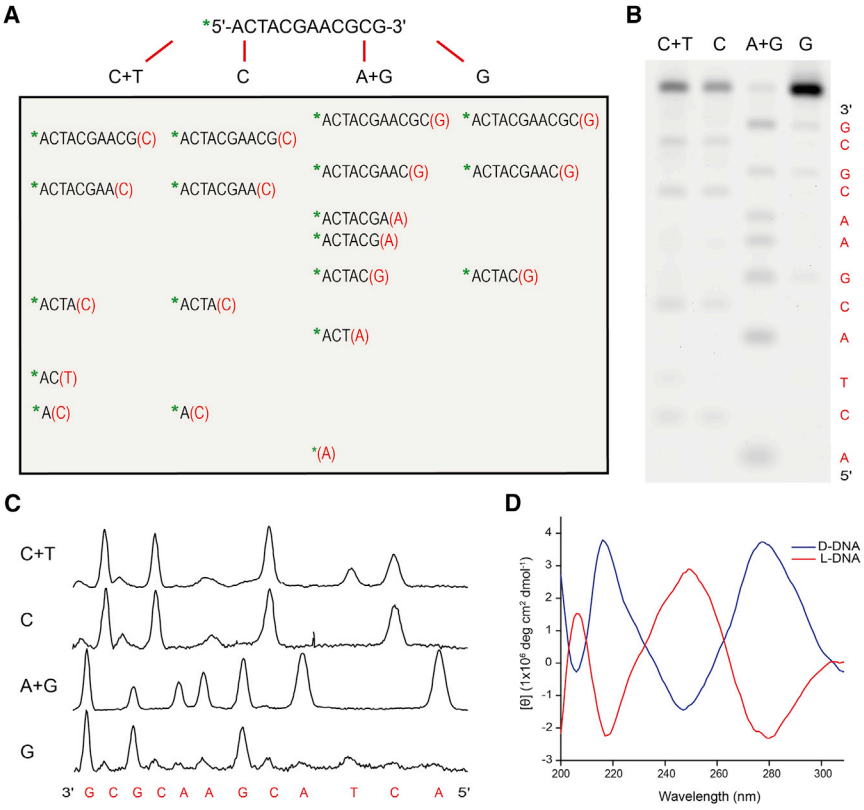


Figure 1. Chemical Sequencing of a 12-nt L-DNA Oligo

(A) Sequence and predicted chemical degradation pattern of the 12-nt L-DNA oligo, with cleaved nucleobases highlighted in parentheses, and fragments separated by PAGE corresponding to the positions of those in (B). Asterisks indicate 5′-FAM label.

(B) Chemical sequencing and denaturing PAGE analysis of the L-DNA oligo: C + T reaction was performed by treatment with 50% hydrazine at 45°C for 18 min; C-specific reaction with 4 M NH_2OH-HCl (pH 6.0) at room temperature for 20 min; A + G reaction with 80% formic acid at room temperature for 40 min; G-specific reaction with 0.1% (m/v) methylene blue under UV for 2 min. The products were analyzed by 20% PAGE in 8 M urea.

(C) Sequencing chromatogram of the 12-nt L-DNA oligo.

(D) CD spectra of 12-nt D- and L-DNA oligos of the same sequence.

During the sequencing, several faint, non-specific bands were observed, particularly with the C + T and C-specific reactions, which were also observed in previous studies on D-DNA chemical sequencing (Banaszuk et al., 1983). Additionally, photo-oxidation in the G-specific reaction tends to be less selective owing to the highly active singlet oxygens (Saito et al., 1984). To overcome potential misreading of the L-DNA sequences, we optimized the C + T and C-specific reactions by carefully adjusting the pH in reaction systems, which is key to the reduction of non-specific bands (Method Details) (Pichersky, 1996; Rubin and Schmid, 1980). Moreover, the major bands in the A + G reaction are known to be highly reliable and can help to minimize the possibility of misreading the sequences (Pichersky, 1996). We show that with optimization by testing different reaction conditions (Table S1), the sequences of the 12-nt L-DNA oligo can be reliably determined by PAGE analysis (Figure 1B) and sequencing chromatography (Figure 1C). Similar degradation patterns were observed with a 12-nt D-DNA oligo of the same sequence (Figure S2) but a mirror-image circular dichroism (CD) spectrum (Figure 1D). We also show that when the method was applied to a non-PAGE-purified 12-nt L-DNA oligo, more non-specific bands were observed, especially toward the 3′ end (Figure S3, the non-specific bands are indicated by red arrows). Thus the L-DNA sequencing method could be useful for the quality control of synthetic L-DNA oligos as a replenishment to mass spectrometry and high-performance liquid chromatography data, which are less sequence specific.

Next, to test the method on L-DNA oligos of other sequences and lengths, we performed the sequencing of two FAM-labeled L-DNA oligos of 11 and 25 nt, respectively (Key Resources Table). Again the C + T, C-specific, A + G, and G-specific reactions were carried out, and the final products were analyzed by 20% PAGE in 8 M urea. We show that with optimization to the reaction conditions by adjusting the reagent concentration and reaction time to reduce the non-specific bands (Table S1), one can correctly read the sequences of the 11- and 25-nt L-DNA oligos by PAGE analysis and sequencing chromatography (Figures 2 and 3).

Chemical Sequencing of an L-DNA Aptamer Molecule

Encouraged by the successful sequencing of short L-DNA oligos, we set out to test the sequencing approach on a longer L-DNA molecule with therapeutic applications. We chose a previously reported 55-nt L-DNA aptamer as a model (Figure 4A and Key Resources Table), which has been shown to bind natural vasopressin and thus has potential to be used as a nuclease-resistant vasopressin antagonist (Williams et al., 1997). Since the efficiency of the G-specific reaction is prone to be affected by the formation of secondary structures (Friedmann and Brown, 1978), we applied A > C reaction by NaOH at 90°C to replace the G-specific reaction (if bands are present in both A + G and A > C lanes, the corresponding base is an A; if a band is present in A + G lane but absent in A > C lane, the corresponding base is a G) (Okawa et al., 1985). We optimized the reaction conditions by adjusting the reagent concentration, temperature, and reaction time to reduce the non-specific bands (Table S1). Because different concentrations of polyacrylamide gels as well as running time provide different optimal separation of fragments at different lengths (Figure S4), we applied a multiple loading strategy through which the full-length (without being

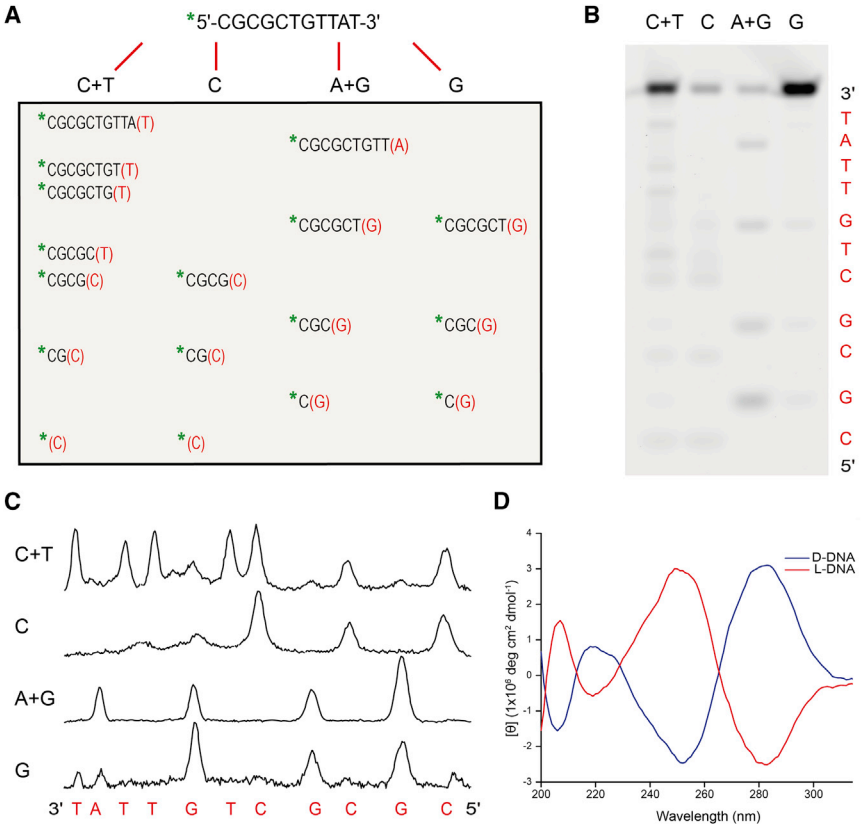


Figure 2. Chemical Sequencing of an 11-nt L-DNA Oligo

(A) Sequence and predicted chemical degradation pattern of the 11-nt L-DNA oligo, with cleaved nucleobases highlighted in parentheses, and fragments separated by PAGE corresponding to the positions of those in (B). Asterisks indicate 5′-FAM label.

(B) Chemical sequencing and denaturing PAGE analysis of the L-DNA oligo: C + T reaction was performed by treatment with 50% hydrazine at 45°C for 18 min; C-specific reaction with 4 M NH_2OH-HCl (pH 6.0) at room temperature for 20 min; A + G reaction with 80% formic acid at room temperature for 40 min; G-specific reaction with 0.1% (m/v) methylene blue under UV at room temperature for 2 min. The products were analyzed by 20% PAGE in 8 M urea.

(C) Sequencing chromatogram of the 11-nt L-DNA oligo.

(D) CD spectra of 11-nt D- and L-DNA oligos of the same sequence.

fragmented) L-DNA was sequenced (Maxam and Gilbert, 1977, 1980), but was loaded into and separated by four polyacrylamide gels of different concentration (10% or 20%), under different voltages (2,000 or 3,000 V) and running time (1.5–5 hr) to ensure optimal separation and reliable reading of the bands (Method Details). With these modifications to the methodology, we show that the full-length L-DNA aptamer sequence can be validated by combining the results from four sequencing gels (Figure 4C).

DISCUSSION

In an effort to realize a practical L-DNA sequencing approach, we have successfully adopted the classic Maxam-Gilbert sequencing method and applied it to the mirror-image system. Using this approach, the sequences of therapeutically useful L-DNA aptamer drugs can now be validated, which is key to the safety of their clinical use. In addition, this approach can be used for confirming the L-DNA gene sequences enzymatically copied and assembled by synthetic mirror-image polymerases (Wang et al., 2016; Xu et al., 2017; Jiang et al., 2017; Pech et al., 2017), toward the realization of a mirror-image central dogma of molecular biology capable of genetic replication, transcription, and translation, and ultimately leading to the synthesis of mirror-image life. It may also facilitate the audacious endeavor to search for life forms with an opposite chirality on exoplanets. Furthermore, since DNA is an excellent medium for storing information, mirror-image DNA sequencing may enable L-DNA to be used as a bio-orthogonal information carrier for L-DNA-based food barcoding and long-term information storage.

So far, we have demonstrated the chemical sequencing of synthetic L-DNA molecules of up to 55 nt. The sequencing of L-RNA is not performed but is potentially feasible when a mirror-image reverse transcriptase becomes available. Alternatively, liquid chromatography-mass spectrometry methods have been used for sequencing non-canonical RNAs of both chiralities (Bjorkbom et al., 2015; Turner et al., 2011). In comparison, the sequencing of longer L-DNA molecules (typically of up to a couple of hundred nucleotides) is more feasible by the Maxam-Gilbert approach, albeit ultimately limited by the resolution of sequencing gel and occurrence of incomplete reactions (Franca et al., 2002). One of the disadvantages of the current approach is its low throughput compared with other modern sequencing techniques, which could be important for its application in a proposed MI-SELEX scheme (Xu et al., 2017; Jiang et al., 2017). Other drawbacks found in the classic Maxam-Gilbert sequencing approach, such as the technical complexity of using hazardous chemicals and requirement of end-labeling, may also hinder the convenient use of this approach. In practice, the 5′ FAM labels can be introduced by the recently established mirror-image polymerase chain reaction (MI-PCR) system (Xu et al., 2017; Jiang et al., 2017; Pech et al., 2017) with an FAM-labeled primer. Alternatively, if a mirror-image polynucleotide kinase could be chemically synthesized, it could also be used to radiolabel the 5′ end of L-DNA molecules for sequencing. Other labeling methods could also be used in principle (Ansorge et al., 1988), as long as the labeling motifs are inert to the chemical cleavage reactions. The development of other alternative L-DNA sequencing techniques, such as adaptations of the Sanger sequencing method (Sanger et al., 1977) or nanopore sequencing (Derrington et al., 2015), may help to address these issues.

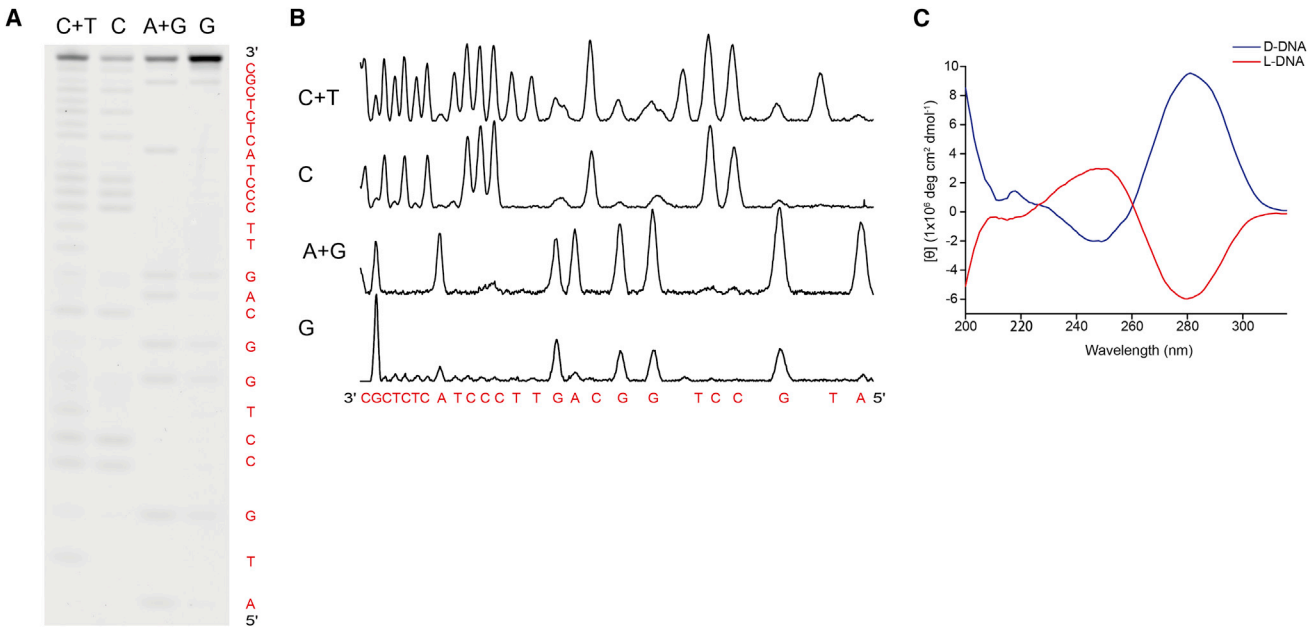


Figure 3. Chemical Sequencing of a 25-nt L-DNA Oligo

(A) Chemical sequencing and denaturing PAGE analysis of the L-DNA oligo: C + T reaction was performed by treatment with 50% hydrazine at 45°C for 10 min; C-specific reaction with 4 M NH_2OH-HCl (pH 6.0) at room temperature for 10 min; A + G reaction with 66% formic acid at room temperature for 10 min; G-specific reaction with 0.1% (m/v) methylene blue under UV at room temperature for 4 min. The products were analyzed by 20% PAGE in 8 M urea.

(B) Sequencing chromatogram of the 25-nt L-DNA oligo.

(C) CD spectra of 25-nt D- and L-DNA oligos of the same sequence.

SIGNIFICANCE

Mirror-image biology systems may enable many exciting applications in biomedicine. In our work, we developed an L-DNA sequencing technique by adopting the classic Maxam-Gilbert approach. This technique may facilitate a number of future studies on mirror-image biology systems such as using L-DNA aptamers as therapeutic tools, as well as realizing a mirror-image central dogma of molecular biology capable of genetic replication, transcription, and translation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and one table and can be found with this article online at https://doi.org/10.1016/j.chembiol.2018.06.005.

ACKNOWLEDGMENTS

We thank C. Fan, M. Wang, B. Wei, and J. Yan for assistance with the experiments. This work was supported in part by funding from the National Natural Science Foundation of China (21750005, 31470532, 91543102, and 31711530153), the Ministry of Science and Technology of China (2015CB553402), the Beijing Nova Program (Z171100001117011), the Tsinghua University Initiative Scientific Research Program (20161080152), the Tsinghua University-Peking University Center for Life Sciences (CLS), and the Beijing Advanced Innovation Center for Structural Biology.

AUTHOR CONTRIBUTIONS

X.L. performed the experiments. X.L. and T.F.Z. designed the experiment and wrote the manuscript.

DECLARATION OF INTERESTS

The authors have filed a provisional patent application related to this work.

Received: December 12, 2017
Revised: March 15, 2018
Accepted: May 1, 2018
Published: July 12, 2018

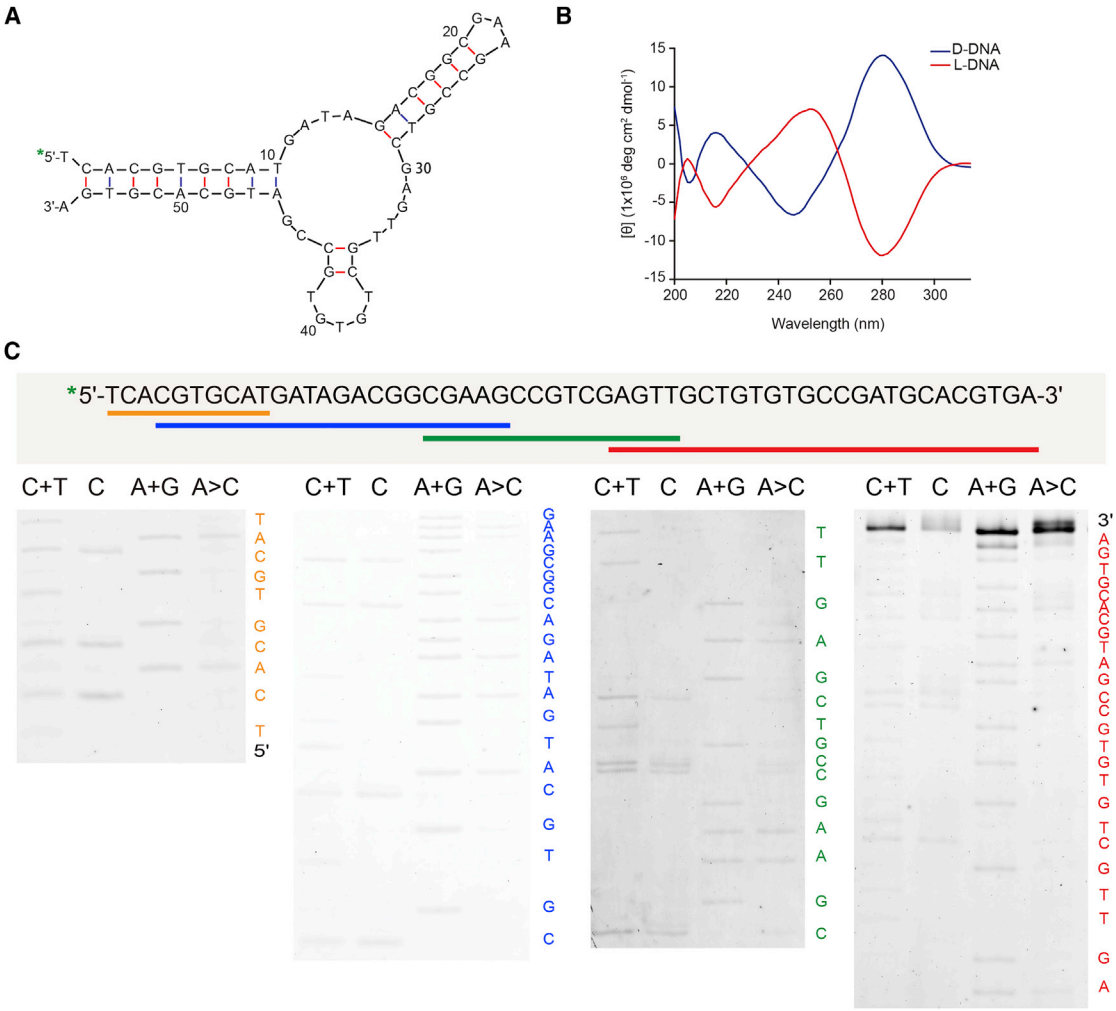


Figure 4. Chemical Sequencing of a 55-nt L-DNA Aptamer

(A) Mfold-predicted secondary structure of the 55-nt L-DNA aptamer (Zuker, 2003). Asterisk indicates 5′-FAM label.

(B) CD spectra of 55-nt D- and L-DNA aptamers of the same sequence.

(C) Multiple loading strategy for sequencing the L-DNA in four sections, indicated by four different colors that correspond to those of the determined sequences. Chemical sequencing and denaturing PAGE analysis of the L-DNA aptamer: C + T reaction was performed by treatment with 50% hydrazine at 45°C for 5 min; C-specific reaction with 4 M NH_2OH-HCl (pH 6.0) at 90°C for 50 s; A + G reaction with 66% formic acid at room temperature for 3 min; A > C reaction with NaOH at 90°C for 12 min. The products were analyzed by 10% or 20% PAGE in 8 M urea.

REFERENCES

Ansorge, W., Rosenthal, A., Sproat, B., Schwager, C., Stegemann, J., and Voss, H. (1988). Non-radioactive automated sequencing of oligonucleotides by chemical degradation. Nucleic Acids Res. *16*, 2203–2206.

Banaszuk, A.M., Deugau, K.V., Sherwood, J., Michalak, M., and Glick, B.R. (1983). An efficient method for the sequence analysis of oligodeoxyribonucleotides. Anal. Biochem. *128*, 281–286.

Bjorkbom, A., Lelyveld, V.S., Zhang, S., Zhang, W., Tam, C.P., Blain, J.C., and Szostak, J.W. (2015). Bidirectional direct sequencing of noncanonical RNA by two-dimensional analysis of mass chromatograms. J. Am. Chem. Soc. *137*, 14430–14438.

Derrington, I.M., Craig, J.M., Stava, E., Laszlo, A.H., Ross, B.C., Brinkerhoff, H., Nova, I.C., Doering, K., Tickman, B.I., Ronaghi, M., et al. (2015). Subangstrom single-molecule measurements of motor proteins using a nanopore. Nat. Biotechnol. *33*, 1073–1075.

Franca, L.T., Carrilho, E., and Kist, T.B. (2002). A review of DNA sequencing techniques. Q. Rev. Biophys. *35*, 169–200.

Friedmann, T., and Brown, D.M. (1978). Base-specific reactions useful for DNA sequencing: methylene blue-sensitized photooxidation of guanine and osmium tetraoxide modification of thymine. Nucleic Acids Res. *5*, 615–622.

Jiang, W., Zhang, B., Fan, C., Wang, M., Wang, J., Deng, Q., Liu, X., Chen, J., Zheng, J., Liu, L., et al. (2017). Mirror-image polymerase chain reaction. Cell Discov. *3*, 17037.

Klussmann, S., Nolte, A., Bald, R., Erdmann, V.A., and Furste, J.P. (1996). Mirror-image RNA that binds D-adenosine. Nat. Biotechnol. *14*, 1112–1115.

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA *74*, 560–564.

Maxam, A.M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. *65*, 499–560.

Nolte, A., Klussmann, S., Bald, R., Erdmann, V.A., and Furste, J.P. (1996). Mirror-design of L-oligonucleotide ligands binding to L-arginine. Nat. Biotechnol. *14*, 1116–1119.

Okawa, N., Suyama, Y., and Kaji, A. (1985). Reduced cleavage by sodium hydroxide of methyladenine in DNA sequencing. Nucleic Acids Res. *13*, 7639–7645.

Pech, A., Achenbach, J., Jahnz, M., Schulzchen, S., Jarosch, F., Bordusa, F., and Klussmann, S. (2017). A thermostable d-polymerase for mirror-image PCR. Nucleic Acids Res. *45*, 3997–4005.

Pichersky, E. (1996). DNA sequencing by the chemical method. Methods Mol. Biol. *58*, 447–451.

Rosenthal, A., Schwertner, S., Hahn, V., and Hunger, H.D. (1985). Solid-phase methods for sequencing of nucleic acids I. Simultaneous sequencing of different oligodeoxyribonucleotides using a new, mechanically stable anion-exchange paper. Nucleic Acids Res. *13*, 1173–1184.

Rubin, C.M., and Schmid, C.W. (1980). Pyrimidine-specific chemical reactions useful for DNA sequencing. Nucleic Acids Res. *8*, 4613–4619.

Saito, I., Sugiyama, H., Matsuura, T., Ueda, K., and Komano, T. (1984). A new procedure for determining thymine residues in DNA sequencing. Photoinduced cleavage of DNA fragments in the presence of spermine. Nucleic Acids Res. *12*, 2879–2885.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Turner, J.J., Hoos, J.S., Vonhoff, S., and Klussmann, S. (2011). Methods for L-ribooligonucleotide sequence determination using LCMS. Nucleic Acids Res. *39*, e147.

Wang, Z., Xu, W., Liu, L., and Zhu, T.F. (2016). A synthetic molecular system capable of mirror-image genetic replication and transcription. Nat. Chem. *8*, 698–704.

Williams, K.P., Liu, X.-H., Schumacher, T.N.M., Lin, H.Y., Ausiello, D.A., Kim, P.S., and Bartel, D.P. (1997). Bioactive and nuclease-resistant l-DNA ligand of vasopressin. Proc. Natl. Acad. Sci. USA *94*, 11285–11290.

Xu, W., Jiang, W., Wang, J., Yu, L., Chen, J., Liu, X., Liu, L., and Zhu, T.F. (2017). Total chemical synthesis of a thermostable enzyme capable of polymerase chain reaction. Cell Discov. *3*, 17008.

Yatime, L., Maasch, C., Hoehlig, K., Klussmann, S., Andersen, G.R., and Vater, A. (2015). Structural basis for the targeting of complement anaphylatoxin C5a using a mixed L-RNA/L-DNA aptamer. Nat. Commun. *6*, 6481.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
E. coli str. K12 substr. MG1655	Tiandz	Cat#12-211
Chemicals, Peptides, and Recombinant Proteins		
Hydroxylamine hydrochloride	Sigma-Aldrich	Cat#159417
Glycogen	Thermo Fisher Scientific	Cat#R0561
Triethylamine	Sinopharm Chemical Reagent	Cat#80134318
Hydrazine hydrate	Sinopharm Chemical Reagent	Cat#80070418
Formic acid	Amresco	Cat#VWRV0916
Methylene blue	Amresco	Cat#97062-276
Sodium hydroxide	Sigma-Aldrich	Cat#S5881
Ethnol	Sinopharm Chemical Reagent	Cat#10009159
Yeast tRNA	Solarbio	Cat#T8630-100
L-deoxynucleoside phosphoramidites	ChemGenes	N/A
Formamide	Amresco	Cat#97061-392
Urea	Sigma-Aldrich	Cat#U5378
Critical Commercial Assays		
PAGE DNA Purification Kit	Tiandz	Cat#70908-30
Oligonucleotides		
D-/L-FAM-*primer12* 5′-FAM-ACTACGAACGCG-3′	This paper	N/A
L-FAM-*primer11* 5′-FAM-CGCGCTGTTAT-3′	This paper	N/A
L-FAM-*primer25* 5′-FAM-ATGCCTGGCAGTTCCCTACTCTCGC-3′	This paper	N/A
L-FAM-*primer55* 5′-FAM-TCACGTGCATGATAGACGGCGAAGCC GTCGAGTTGCTGTGTGCCGATGCACGTGA-3′	This paper	N/A
Software and Algorithms		
ImageQuant TL 8.1	GE Healthcare Life Sciences	https://www.gelifesciences.com/en/us/shop/protein-analysis/molecular-imaging-for-proteins/imaging-software/imagequant-tl-8-1-p-00110?current=29000605
Microsoft Excel 2016	Microsoft	https://products.office.com/en-us/?legRedir=default&CorrelationId=062cb0fc-b73f-4e95-90a8-8f921441afbb
Pro-Data Viewer 4.2.5	Applied Photophysics	https://www.photophysics.com/systems/chirascan-systems/chirascan/system-information/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ting F. Zhu (tzhu@tsinghua.edu.cn).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Carrier DNA Preparation from *E. coli*

The *E. coli* str. K12 substr. MG1655 cells were used for genomic DNA extraction. Briefly, the bacteria were inoculated in 5 ml of Lysogeny broth (LB) medium and grown overnight at 37°C under shaking at 200 RPM. The genomic DNA was purified by phenol/chloroform extraction and ethanol precipitation, and resuspended in RNase-free water.

METHOD DETAILS

C+T Reaction by Hydrazine

An aliquot of 2 μl FAM-labelled primer (10 μM) was mixed with 3 μg carrier *E. coli* genomic DNA and kept on ice. The mixture was denatured by heating to 95°C for 2 min followed by quick chilling on ice. An aliquot of 40 μl 80% hydrazine hydrate was added and the mixture was incubated at 45°C for 18 min (reduced to 10 min for the 25-nt sequence and to 5 min for the 55-nt sequence, respectively). The reaction was quenched by adding 200 μl 0.3 M sodium acetate, 2 μl glycogen (10 mg/ml), 2 μl EDTA (10 mM, pH 8.0), 5 μl yeast tRNA (10 mg/ml), and 1 ml absolute ethanol, and the mixture was chilled in liquid nitrogen for 10 min. The processed DNA was precipitated by centrifugation at 12,000 rpm for 10 min and washed by 1 ml absolute ethanol. The residual ethanol was removed by evaporation, and the pellet was dissolved into 120 μl 1 M piperidine and incubated at 90°C for 50 min. After lyophilization, the remaining pellet was dissolved in a denaturation buffer containing 98% formamide, 0.25 mM EDTA, and 0.0125% SDS. The products were analyzed by 10% or 20% PAGE in 8 M urea and scanned by a Typhoon Trio+ system operated under Cy2 mode.

C-Specific Reaction by NH_2OH-HCl

An aliquot of 2 μl FAM-labelled primer (10 μM) was mixed with 3 μg carrier *E. coli* genomic DNA and kept on ice. The mixture was denatured by heating to 95°C for 2 min followed by quick chilling on ice. An aliquot of 40 μl 4 M NH_2OH-HCl (pH adjusted to 6.0 by trimethylamine) was added and the mixture was incubated at RT for 20 min (reduced to 10 min for the 25-nt sequence and to 50 s at 90°C for the 55-nt sequence, respectively). The reaction was quenched by adding 200 μl 0.3 M sodium acetate, 2 μl glycogen (10 mg/ml), 2 μl EDTA (10 mM, pH 8.0), 5 μl yeast tRNA (10 mg/ml), and 1 ml absolute ethanol, and the mixture was chilled in liquid nitrogen for 10 min. The processed DNA was precipitated by centrifugation at 12,000 rpm for 10 min and washed by 1 ml absolute ethanol. The residual ethanol was removed by evaporation, and the pellet was dissolved into 100 μl 1 M piperidine and incubated at 90°C for 30 min. After lyophilization, the remaining pellet was dissolved in a denaturation buffer containing 98% formamide, 0.25 mM EDTA, and 0.0125% SDS. The products were analyzed by 10% or 20% PAGE in 8 M urea and scanned by a Typhoon Trio+ system operated under Cy2 mode.

A+G Reaction by Formic Acid

An aliquot of 2 μl FAM-labelled primer (10 μM) was mixed with 3 μg carrier *E. coli* genomic DNA and kept on ice. An aliquot of 40 μl 80% formic acid was added and the mixture was incubated at RT for 30 min (formic acid concentration reduced to 66% and incubation time reduced to 10 min for the 25-nt sequence and to 3 min for the 55-nt sequence, respectively). The reaction was quenched by adding 200 μl 0.3 M sodium acetate, 2 μl glycogen (10 mg/ml), 5 μl yeast tRNA (10 mg/ml), and 1 ml absolute ethanol, and the mixture was chilled in liquid nitrogen for 10 min. The processed DNA was precipitated by centrifugation at 12,000 rpm for 10 min and washed by 1 ml absolute ethanol. The residual ethanol was removed by evaporation, and the pellet was dissolved into 100 μl 1 M piperidine and incubated at 90°C for 30 min. After lyophilization, the remaining pellet was dissolved in a denaturation buffer containing 98% formamide, 0.25 mM EDTA, and 0.0125% SDS. The products were analyzed by 10% or 20% PAGE in 8 M urea and scanned by a Typhoon Trio+ system operated under Cy2 mode.

G-Specific Reaction by UV with Methylene Blue

An aliquot of 2 μl FAM-labelled primer (10 μM) was mixed with 3 μg carrier *E. coli* genomic DNA and kept on ice. The mixture was denatured by heating to 95°C for 2 min followed by quick chilling on ice. An aliquot of 20 μl 0.1% (m/v) methylene blue was added and the mixture were exposed to a handheld UV lamp at a distance of ~10 cm for 2 min (exposure time increased to 4 min for the 25-nt sequence). The reaction was quenched by adding 200 μl 0.3 M sodium acetate, 2 μl glycogen (10 mg/ml), 5 μl yeast tRNA (10 mg/ml), and 1 ml absolute ethanol, and the mixture was chilled in liquid nitrogen for 10 min. The processed DNA was precipitated by centrifugation at 12,000 rpm for 10 min and washed by 1 ml absolute ethanol. The residual ethanol was removed by evaporation, and the pellet was dissolved into 100 μl 1 M piperidine and incubated at 90°C for 30 min. After lyophilization, the remaining pellet was dissolved in a denaturation buffer containing 98% formamide, 0.25 mM EDTA, and 0.0125% SDS. The products were analyzed by 20% PAGE in 8 M urea and scanned by a Typhoon Trio+ system operated under Cy2 mode.

A>C Reaction by NaOH

An aliquot of 2 μl FAM-labelled primer (10 μM) was mixed with 3 μg carrier *E. coli* genomic DNA and kept on ice. An aliquot of 20 μl 1.5 M NaOH/1 mM EDTA was added and the mixture was incubated at 90°C for 12 min. The reaction was quenched by adding 100 μl 1 M sodium acetate, 2 μl glycogen (10 mg/ml), 5 μl yeast tRNA (10 mg/ml), and 1 ml absolute ethanol, and the mixture was chilled in

liquid nitrogen for 10 min. The processed DNA was precipitated by centrifugation at 12,000 rpm for 10 min and washed by 1 ml absolute ethanol. The residual ethanol was removed by evaporation, and the pellet was dissolved into 100 μl 1 M piperidine and incubated at 90°C for 30 min. After lyophilization, the remaining pellet was dissolved in a denaturation buffer containing 98% formamide, 0.25 mM EDTA, and 0.0125% SDS. The products were analyzed by 10% or 20% PAGE in 8 M urea and scanned by a Typhoon Trio+ system operated under Cy2 mode.

Polyacrylamide Gel Electrophoresis (PAGE) and Chromatogram Analysis

The reaction products from sequencing short DNA oligos (≤25 nt) were loaded on slabs of 1 mm × 200 mm × 200 mm, separated by 20% polyacrylamide gel in 8 M urea. The gel was pre-run at 450 V (average) for 0.5 h until being heated to 30-40°C. After loading, the gel was run at 450 V (average) for 2.5 h. The reaction products from sequencing longer DNA oligos (>25 nt) were loaded on slabs of 0.4 mm × 200 mm × 550 mm, separated by 10%-20% polyacrylamide gel in 8 M urea. The gel was pre-run at 2,000-3,000 V (depending on the gel concentration, e.g., 2,000 V for 10% polyacrylamide gel and 3,000 V for 20% polyacrylamide gel) for 1 h until being heated to 50-55°C. After loading, the gel was run at 2,000-3,000 V for 3-5 h. For the sequencing of 55-nt L-DNA aptamer, the four sections were run at 3,000 V for 1.5 h, 2.5 h, 5 h in 20% polyacrylamide gel (the orange, blue, and green sections in Figure 4C, respectively), and at 2,000 V for 3 h in 10% polyacrylamide gel (red section). The polyacrylamide gels were scanned by a Typhoon Trio+ system operated under Cy2 mode. Gel quantitation and chromatogram analysis were performed by the ImageQuant TL 8.1 software with 1D gel analysis package. The data were exported to and plotted by the Microsoft Excel 2016 software.

Circular Dichroism (CD) Spectroscopy

The CD spectra were measured using an Applied Photophysics Chirascan-Plus CD spectrometer. The D- and L- DNA oligos were diluted to 25 μM in ddH_2O and scanned from 200 nm to 320 nm. Pure ddH_2O was also scanned and data used as background. The data were background-subtracted and averaged from three independent measurements by the Pro-Data Viewer 4.2.5 software. The data were exported to and plotted by the Microsoft Excel 2016 software.

QUANTIFICATION AND STATISTICAL ANALYSIS

The chromatogram data were normalized according to the maximum peak value of the sequenced bases in the chromatogram. The CD data were background-subtracted and averaged from three independent measurements by the Pro-Data Viewer 4.2.5 software. The data were exported to and plotted by the Microsoft Excel 2016 software.