Electronic Supplementary Material (ESI) for Analytical Methods.

Supporting Information

Characterization of DNA nanostructure stability by size exclusion chromatography

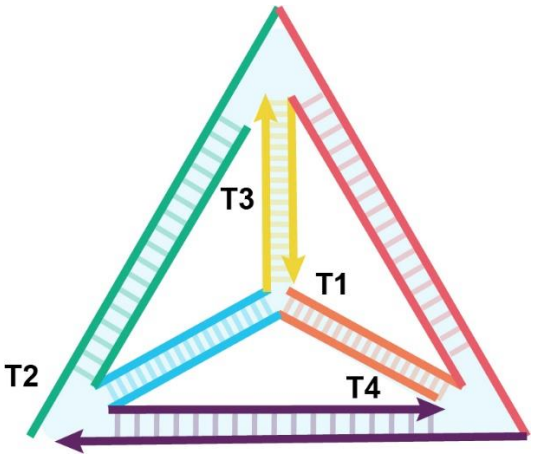
Nicole I. Langlois[†] and Heather A. Clark*[†‡]

[†]Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA, United States
[‡]Department of Bioengineering, Northeastern University, Boston, MA, United States

*Corresponding author: Heather A. Clark
Email: h.clark@northeastern.edu

TABLE OF CONTENTS

1. DNA Tetrahedron (TDN) Design



ID	Sequence (5' → 3')	Length (nt)
T1	ACATTCCTAAGTCTGAAACATTACAGCTTGCTACACGAGAAGAGCCGCCATAGTA	55
T2	TATCACCAGGCAGTTGACAGTGTAGCAAGCTGTAATAGATGCGAGGGTCCAATAC	55
T3	TCAACTGCCTGGTGATAAAACGACACTACGTGGGAATCTACTATGGCGGCTCTTC	55
T4	TTCAGACTTAGGAATGTGCTTCCCACGTAGTGTCGTTTGTATTGGACCCTCGCAT	55

Figure S1. Design of the DNA tetrahedron (TDN) used in this study. Four oligonucleotides are annealed to form this TDN with 17-bp edges. The design of the structure and all sequences are used directly from the original report by Tuberfield and colleagues[1]. The colors indicate regions of the oligonucleotides with complementary domains, while the black two-base regions are located at the TDN vertices.

2. Analysis of TDN and Aggregate Peak Profiles

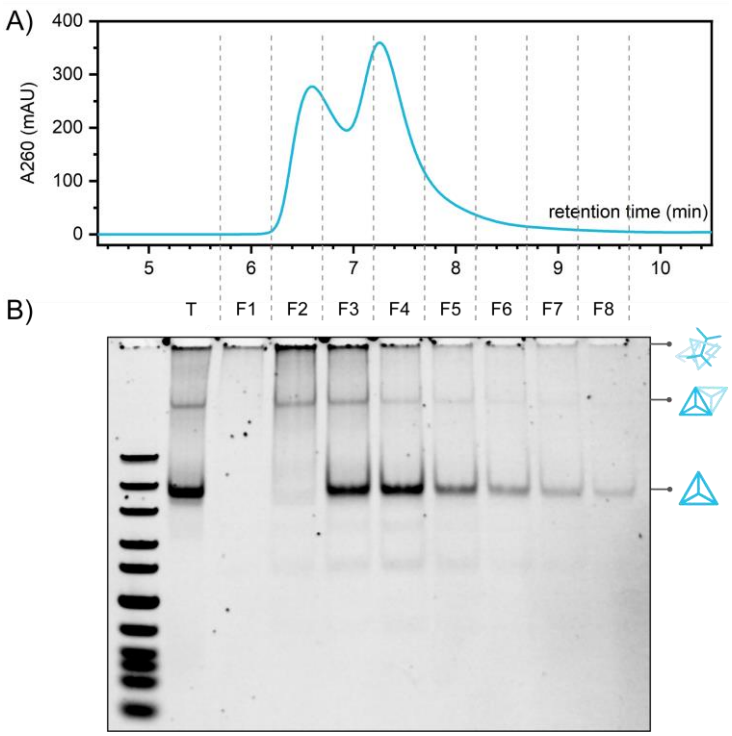


Figure S2. Fraction collection across the annealed TDN chromatogram. A sample of TDNs was annealed using equimolar additions (2 μM) of each component strand, then 50 μL was injected onto the SuperSW2000 column with a mobile phase of 1X PBS at 0.300 mL^{-1} min^{-1}, the same HPLC conditions as shown in ***Figure 1***. Fractions were collected manually at 0.5 min intervals (150 μL), as indicated by dotted lines in (A), then analyzed by 6.0% native PAGE (B) to assess the composition of each fraction. Gel lanes: ultra low DNA ladder, TDN annealed sample before injection (T), collection fractions (F1-F8). On the chromatogram, the elution order follows as high MW aggregates, followed by multimer side products[2], then the assembled TDN starting from fraction 3. As such, the second peak appearing in the chromatogram ~7.3 min can be identified as the assembled TDN structures. The ultra-low DNA ladder bands are as follows: 300, 200, 150, 100, 75, 50, 35, 25, 20, 15, 10-bp.

3. Purification of TDN Samples by SEC

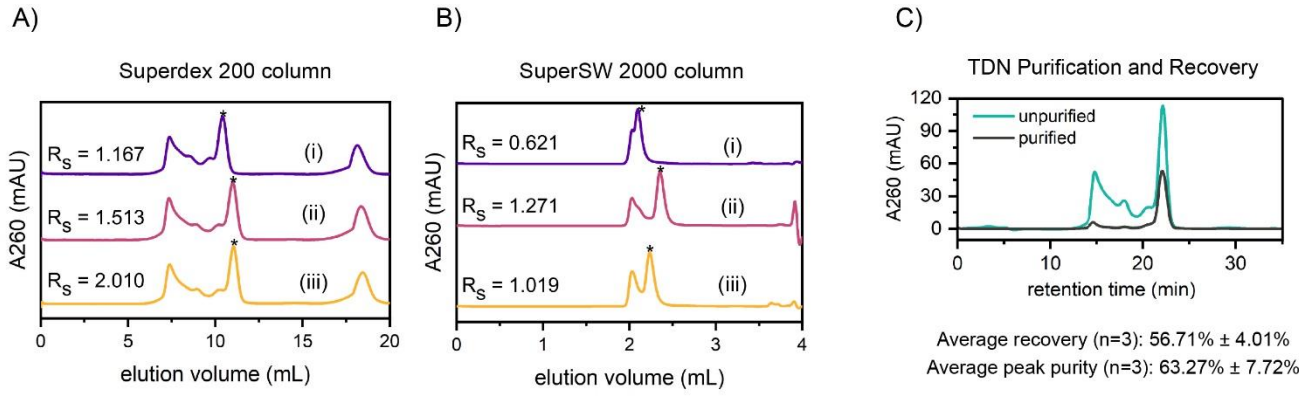


Figure S3. Purification of annealed TDNs. Following a reported approach[3], we evaluated the impact of mobile phase composition and flow rate on the resolution of annealed TDNs from aggregates using a semi-preparative (A) and an analytical (B) column. The targeted TDN peak is marked with an asterisk. The resolution is determined by the retention time distance between the TDN peak and the highest abundance aggregate peak, divided by the average width at baseline of the two sections. Example conditions shown for the Superdex 200 column include (i) 0.3X TN buffer at 0.75 mL min^{-1}, (ii) 1X TN buffer at 0.75 mL min^{-1}, and (iii) 1X TN buffer at 0.5 mL min^{-1}. Example conditions shown for the SuperSW 2000 column include (i) 0.3X TN buffer at 0.4 mL min^{-1}, (ii) 1.0X TN buffer at 0.4 mL min^{-1}, and (iii) 1X PBS at 0.3 mL min^{-1}. The highest resolution condition shown in A (iii) was used for all further purifications. The conditions shown in B (iii) were used for the TDN and HS stability studies, where TDNs are separated from the aggregates with a Rs value of 1.019. C) Purification and recovery analysis of TDNs showed an average of 56.7% recovery with 63.3 ± 7.7% peak purity.

4. Chromatographic Data Analysis Method

For stability analysis experiments, the peak integration range for all time points was determined by first calculating the peak width at the baseline for a pure TDN sample. The baseline-subtracted chromatograms for absorbance at 260 nm over time were exported from ChemStation and analyzed in Origin 2021b software. The TDN peak corresponding to retention time of ~7.3 – 7.4 min was selected with the Multiple Peak Fit program and fitted to a Gaussian distribution, using the "Gauss" function.

The resulting equation in OriginPro follows the form of Eq 1:

$$y = y_0 + \frac{A}{w\sqrt{\frac{\pi}{2}}} e^{-2\frac{(x-x_c)^2}{w^2}}$$

Transforming the output to fit the traditional notation of a standard Gaussian curve equation, we arrive at Eq. 2:

$$y = y_0 + \frac{A}{(2\sigma)\sqrt{\frac{\pi}{2}}} e^{-2\frac{(x-\mu)^2}{(2\sigma)^2}}$$

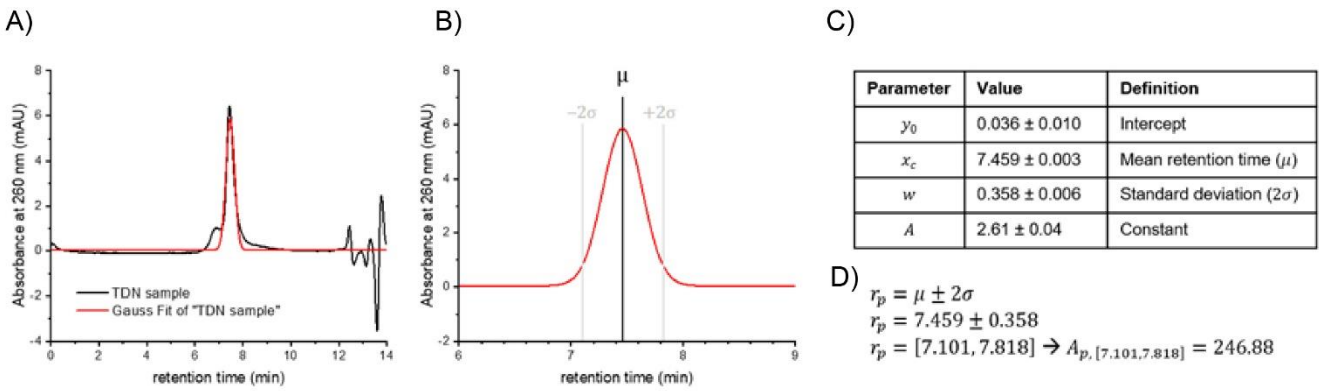
$w = 2\sigma$ and σ is the standard deviation of the normal distribution

$x_c = \mu$ and μ is the mean of the normal distribution, i.e the peak retention time

Eq. 2 can be simplifed to fit the standard formula of a standard Gaussian curve equation as Eq. 3:

$$y = y_0 + \frac{A}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The fit parameters can be used to calculate the peak width at baseline ($w_b = 4\sigma$)[4] and used to determine the retention time range (r_p) over which the peak width will encompass 95.4% of the total peak area ($r_p = \mu \pm 2\sigma$). Using the Peak Analysis program, each chromatogram for a data set was then integrated over the specified retention time range (r_p), with the integrated peak area (A_p) serving as the metric for amount of intact TDN present for that incubation time point. This process can be illustrated below in ***Figure S4***



Parameter	Value	Definition
y_0	0.036 ± 0.010	Intercept
x_c	7.459 ± 0.003	Mean retention time (μ)
w	0.358 ± 0.006	Standard deviation (2σ)
A	2.61 ± 0.04	Constant

Figure S4. Chromatographic analysis of peaks using OriginPro 2021b software. A) An example chromatogram of purified TDNs (79 ng) with Gaussian curve fit of the TDN peak. B) Illustration of the mean (μ) and standard deviation (±2σ) parameters for the corresponding TDN curve fit. C) OriginPro exported parameters for the Gaussian curve fit correlated to the standard notation. D) An example calculation of the retention time range for integration, and output for the integrated peak area.

5. Comparative DNaseI Digestion of TDN Intermediates

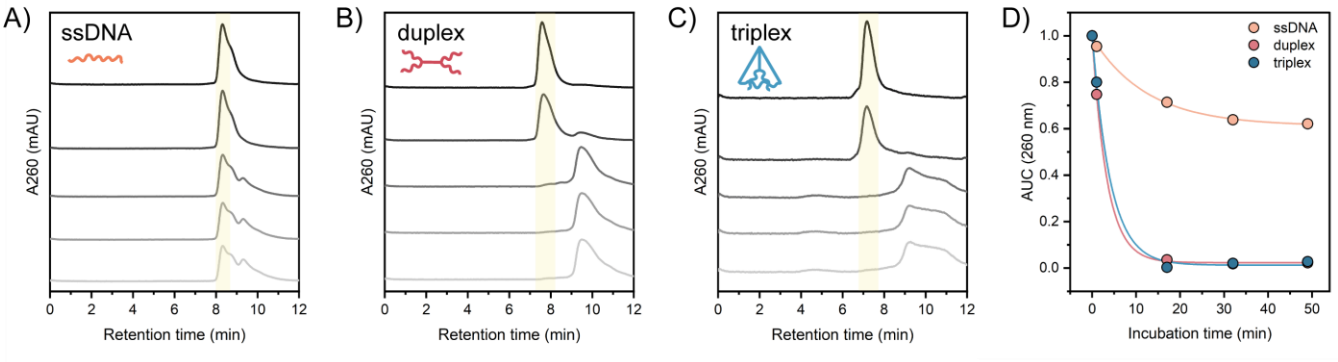


Figure S5. Digestion of (A) ssDNA, (B) duplex, and (C) triplex intermediates from TDN formation by 1.56 U mL^{-1} of DNaseI. The samples were prepared as in ***Figure 3***, using 500 ng of each species. First order exponential decay profiles in (D) for ssDNA (T1), duplex (T1+T2), and triplex (T1+T2+T3) were used to approximate half-lives of 8.72 min, 2.33 min, and 2.85 min, respectively for one replicate digestion. The duplex and triplex dsDNA domains were completely digested into short oligo products in under 15 minutes. As expected from numerous reports, the TDN with a half-life of ~11.93 min at this DNaseI concentration, shows higher stability due to increased nuclease resistance. Notably, the T1 ssDNA exhibited a much slower rate of digestion compared to duplex and triplex intermediates, due to the preferential DNaseI digestion of dsDNA at a rate 100–500 times faster than that of ssDNA.

6. Control Samples for TDN and Serum Stability Study

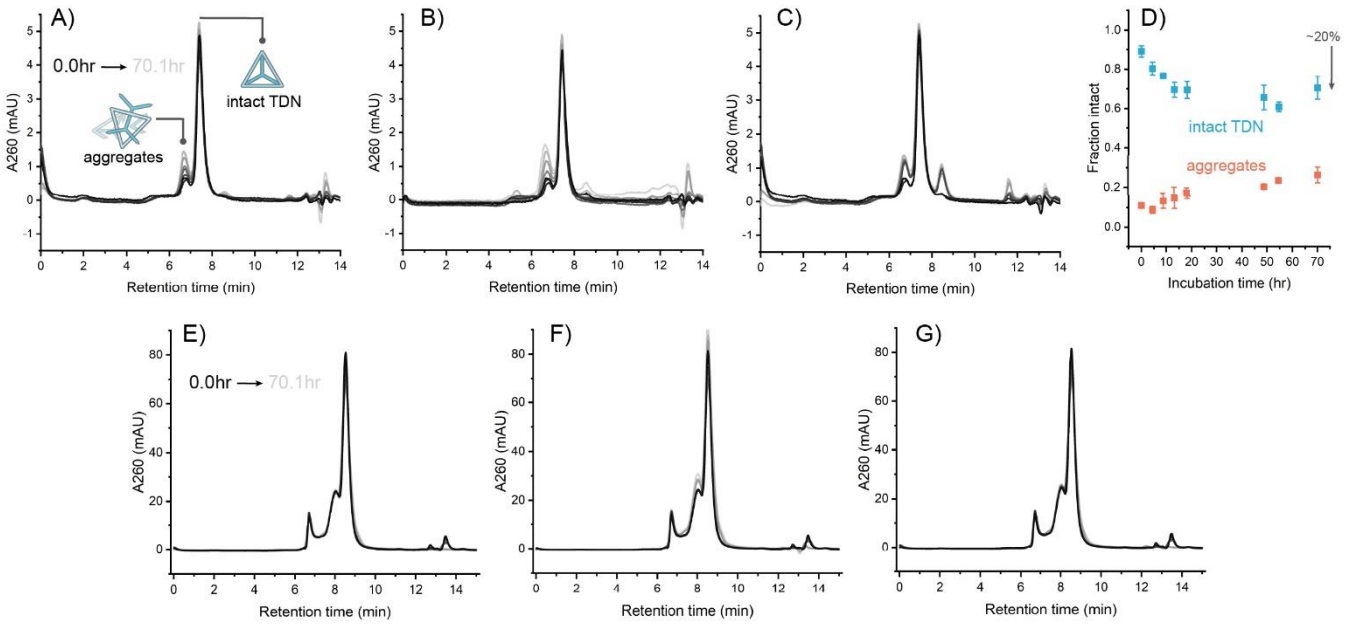


Figure S6. Control data for the TDN and HS incubations. Three independent replicates of TDN (500 ng in PBS) (A–C) were analyzed over a three-day incubation period by the direct-injection method. The relative areas of the TDN peak and the aggregate peak were used to determine the fraction of TDNs remaining intact. After 70.1 hours of incubation at 37°C in PBS, there was a decrease of intact TDNs by ~20% primarily as a result of aggregation. Three independent replicates of 5% HS in PBS (E–G) were also analyzed over the incubation period by direct-injection. The chromatographic profile of the HS samples exhibited negligible changes to peak area over the analysis period.

7. Serum Protein Background Interference in Gel Electrophoresis

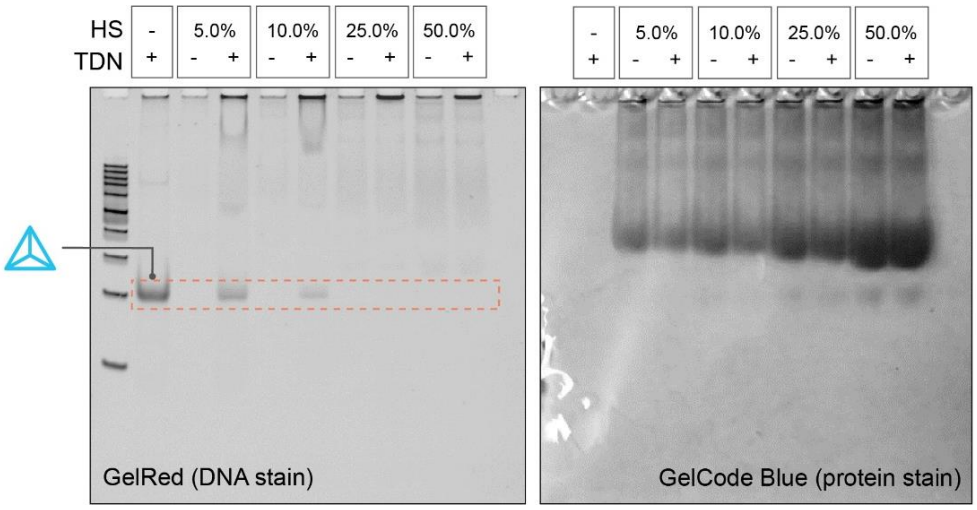


Figure S7. 6% native PAGE analysis of TDN samples spiked into diluted human serum. All samples were prepared on ice immediately before gel characterization to prevent degradation. The gel was stained in GelRed DNA staining solution, documented, then stained using GelCode Blue protein staining solution. As the percentage of HS increases in the sample preparation, the apparent amount of TDN (band in dotted box) decreases. This is a result of the TDNs interacting with the serum proteins and forming aggregate complexes, evidenced by the increasing staining intensity in the wells at the top of each TDN-containing lane. Additionally, background signal from the serum proteins can be seen all HS-containing lanes, even at as low as 5% HS.

8. References

1. Goodman, R. P.; Berry, R. M.; Turberfield, A. J., The single-step synthesis of a DNA tetrahedron. Chemical Communications 2004, (12), 1372-1373.
2. Gao, L.; Liu, L.; Tian, Y.; Yang, Q.; Wu, P.; Fan, C.; Zhao, Q.; Li, F., Probing the Formation Kinetics and Thermodynamics with Rationally Designed Analytical Tools Enables One-Pot Synthesis and Purification of a Tetrahedral DNA Nanostructure. Analytical Chemistry 2021, 93 (18), 7045-7053.
3. Xing, S.; Jiang, D.; Li, F.; Li, J.; Li, Q.; Huang, Q.; Guo, L.; Xia, J.; Shi, J.; Fan, C.; Zhang, L.; Wang, L., Constructing Higher-Order DNA Nanoarchitectures with Highly Purified DNA Nanocages. ACS Applied Materials & Interfaces 2015, 7 (24), 13174-13179.
4. Harris, D. C., Quantitative chemical analysis. Eighth ed. ed.; W.H. Freeman and Co.: New York, 2010.