SUPPLEMENTARY MATERIAL

Design of Tetrahedron Structures

For each construct, ssDNA strands were designed to anneal to form dsDNA that formed three-dimensional tetrahedra. The strands for the 20T tetrahedron were designed as described. [1] Additionally, a 3x20bp, 3x30bp (20x30T) and 30 bp (30T) were designed and characterized in this work. Schematics indicating the overall structures for each tetrahedron are shown in Figure S1.

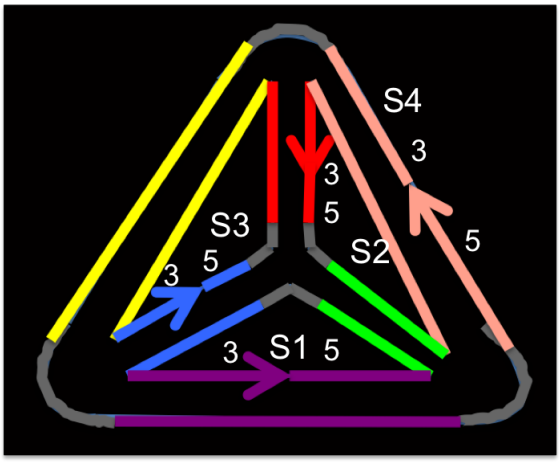
Single-stranded oligonucleotides for 20T Tetrahedron

S1 –
AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCACCCGCCATAGTAGACGTATCACC

S2 –
CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGAGGGTCCAATACCGACGATTACAG

S3 –
GGTGATAAAACGTGTAGCAAGCTGTAATCGACGGGAAGAGCATGCCCATCCACTACTATGGCG

S4 -
CCTCGCATGACTCAACTGCCTGGTGATACGAGGATGGGCATGCTCTTCCCGACGGTATTGGAC



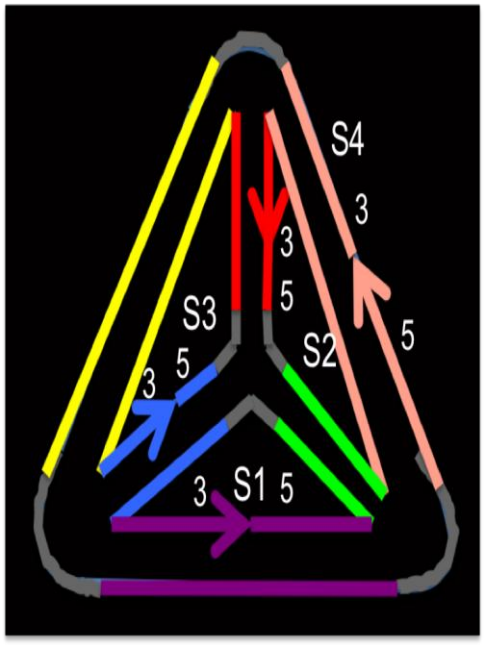
20bp on each edge (20T). The “3” and “5” on each end of each of the four strands indicates 3’ and 5’ ends of the strands, but the (‘) was eliminated for clarity in the figure. Each strand spans one entire face of the tetrahedron, and there are four faces and six edges. The gray areas represent a single adenine unpaired nucleotide.

Single-stranded oligonucleotides for 20x30T Tetrahedron

S1 -
AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCACCCGCCATAGTAGACGTATCACC
S2 -
TGCGCCTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGAGGAGGAAATGAAGTCCAAT ACCGACGATTACAGGCCTT

S3 -
GGTGATAAAACGTGTAGCAAGGCGCAAAGGCCTGTAATCGACTCTACGGGAAGAGCATGCCCAT CCGGCTCACTACTATGGCG

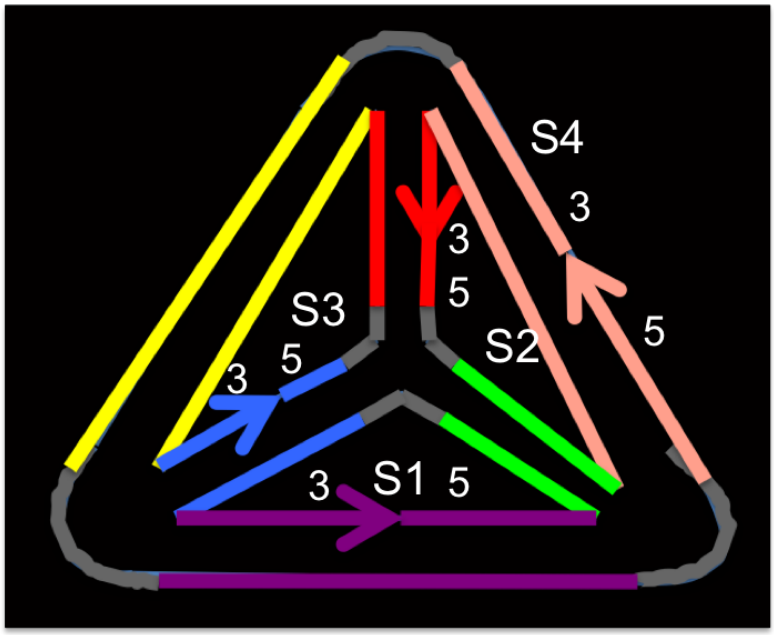
S4 –
TTCCTCCTCGCATGACTCAACTGCCTGGTGATACGAGAGCCGGATGGGCATGCTCTTCCCGTAG AGACGGTATTGGACTTCAT



3x20, 3x30T. The same principles apply as described in the 20T figure above, except there are three edges that have 30 bp forming an asymmetric tetrahedron.

Single-stranded oligonucleotides for 30T Tetrahedron

S1 -
AGGCAGTTGAGACGAACATTCCTAAGTCTGAATTAGCATAACATTTATCACCCGCCATAGTAGCCATAAGTCCATATCAGC ATTCGTATCACC
S2 -
CTTGCTACACGATTCAGACTTAGGAATGTTCGAATCGTATTGACATGCGAGGGTCCAATACCGTTACGATGCAAGCTTAGC TATCGATTACAG
S3 -
GGTGATAAAACGTGTAGCAAGCTGTAATCGATAGCTAAGCACGGGAAGAGCATGCCCATCCGGATTCAATAAGGACTTAT GGCTACTATGGCG
S4 –
CCTCGCATGACTCAACTGCCTGGTGATACGAATGCTGATAAGGATGGGCATGCTCTTCCCGCCTAAGTTATATGCATCGTA ACGGTATTGGAC



30bp on each edge (30T). The same principles apply as described in the 20T structure, except all edges in this construct are 30 bp in length.

MATERIALS AND METHODS

Oligonucleotide Purification

The four single-stranded DNA oligonucleotides for all three constructs[1] were purchased from Integrated DNA Technologies (Coralville, IA) and purified using 8% denaturing-PAGE. The amount equal to 4 O.D. at 260 nm of each strand were loaded in each lane of the gel. The bands were excised from the gel and placed into 0.45 μm filtered Spin-X columns (Corning, Lowell, MA) with 0.5 mL of gel elution buffer (500 mM ammonium acetate, 10 mM magnesium acetate, and 2 mM EDTA-Na_2 pH 8.0). The filter trapped the gel pieces from eluting along with the DNA. The DNA was extracted from the gel matrix at room temperature using a shaker for 1-2 hours before samples were centrifuged at 8000 rpm for 6 minutes. The top of the Spin-X column containing the gel pieces was discarded and 1 mL of 1-butanol was added to the liquid remaining in the bottom of the column, which resulted in the formation of an aqueous layer at the bottom (containing the DNA, and an organic layer containing ethidium bromide and gel tracking dyes). Samples were centrifuged at 2000 rpm for one minute, and the upper organic layer containing ethidium bromide and loading dyes was discarded. 1 mL of 100% ethanol was added to the remaining DNA and placed in the freezer at -20°C for 2 hours, to precipitate the DNA. Samples were centrifuged at 13,000 rpm for 30 minutes at 4°C. The ethanol served as an additional precipitation step to ensure purity of the DNA. A white pellet was obtained and after removing the ethanol supernatant another precipitation followed, consisting of 70% ethanol. The samples were centrifuged at 13,000 rpm for 10 minutes at 4°C, the ethanol supernatant was discarded and the pellet was allowed to dry in a vacuum centrifuge for 2 hours with the temperature set to 30°C. After drying the DNA, 50 uL of nanopure water was added to each of the tubes and the pellet was resuspended. The DNA was quantified by measuring the absorbance at 260 nm.

Tetrahedron Purification and Characterization

The DNA strands were annealed in 50 nM concentrations in 45 mL Falcon tubes and concentrated to approximately 0.5 mL before loading the sample onto a gel filtration column. The tetrahedron was purified in 120 mL of 1X TAE-Mg^{2+} buffer pH 8.0 at room temperature using a Superdex SD200 prep-grade size-exclusion column (GE Healthcare, Piscataway, NJ) with a column volume of 120 mL. The tetrahedron eluted after approximately 65 minutes at a flow rate of 1 mL min^{-1} and the samples were collected and concentrated to at least 1.5 mg DNA/mL or 12 μM for crystallization experiments. Native gel electrophoresis was used to analyze the homogeneity of the 20T tetrahedron. Samples were prepared for native gel electrophoresis by mixing 20T at a concentration of 2 mg ml^{-1} with 10X native tracking dye [0.2% (*w/v*) bromophenol blue, 0.2% (*w/v*) xylene cyanol, and 50% glycerol in 1X TAE-Mg^{2+}]. Samples were run on a 5% polyacrylamide native gel for approximately 3 hours at 23°C. Unpurified and purified tetrahedron samples were both analyzed using native gel electrophoresis. In another attempt to verify the homogeneity of the sample, Dynamic Light Scattering (DLS) was used. (Molecular Dimensions, Apopka, FL, USA) Samples were prepared for DLS by placing 3 μL of 2 mg ml^{-1} 20T tetrahedron in the center of a siliconized cover slide, inverting, and placing on the top of a well containing 0.9 mL of buffer solution (1X TAE-Mg^{2+} pH 8.0). A 60 mW 785 nm laser was aligned to penetrate the drop, and data was collected at 20 second intervals 10 times. Parameters for the DLS instrument were adjusted for the 20T to account for its lower density compared to that of proteins. The 20T has a theoretical solvent content of ~80%, whereas proteins typically have a solvent content around 50%.[2] Following DLS data collection, the resulting hydrodynamic radius was calculated, as shown in the results. The

parameters used were the following: refractive index 1, shape factor 1, hydration shell 0.2, exponent 2.3 (a calculation which would provide the molecular weight given the size of the molecule), and molecular density 0.4. The fraction of single tetrahedrons vs aggregates were calculated by integration of the peaks and divided by scattering intensity increases by r^6. The amount of aggregates was calculated by dividing the peak integration area by $radius^6$. DLS revealed that aggregates formed with samples with concentrations higher than 3 mg ml^{-1}; therefore, 2.5 mg ml^{-1} was the concentration used for both DLS and crystallization experiments.

RESULTS

Tetrahedron Designs

In addition to the 20T described in this work, 20x30T and 30T constructs were also purified and characterized. A schematic view of the tetrahedron constructs investigated in this study is shown in Figure S1, which include a 20T (20 base pairs on each edge), [1] 20x30T (asymmetrical with 20 base pairs on each edge for one face and 30 bp on each edge for the other faces), [1] and a 30T (30 base pairs on each edge) tetrahedron. The size variations were designed to increase the cavity size which would in a later stage of the project allow for the incorporation of larger proteins and macromolecules inside the tetrahedrons. The sequences for all constructs is shown at the beginning of the Supplemental Information. While the 20T tetrahedron has an estimated inner volume of 32,364 $Å^3$, the inner volume is significantly increased to 58,955 $Å^3$ for the 20x 30 T construct and 125,064 $Å^3$ for the 30T tetrahedron. Assuming a sphere fitting inside of each symmetrical tetrahedron, it would have a diameter of 70 Å for the 20T and 100 Å for the 30T tetrahedron. The increased size of the larger tetrahedron constructs is confirmed by native gel electrophoresis as shown in Figure S2c-d.

20T, 20x30T, and 30T Characterization

In addition to the 20T tetrahedron, two other constructs were designed, annealed, and purified—the 20x30T tetrahedron and a 30T tetrahedron which were designed to further increase the cavity size (see Fig S1 and S2). Figure S2 shows the purification of the 20x30T and 30T by size exclusion chromatography and verification of the purification procedure by native gel-electrophoresis (see Figs S2c and d).

20T Purification and Characterization

The 20T tetrahedron was successfully purified using size-exclusion chromatography, as shown in Figure 1 in the article. The figure shows a schematic overview of the assembly and purification indicating the annealing of the four single-stranded DNA oligonucleotides, the purification via size-exclusion chromatography. The purification procedure is verified by native gel electrophoresis of the sample before and after size exclusion chromatography (Fig 1 bottom right) which shows that the aggregates (band "a") and partially assembled structures (band "c") are successfully removed from the fully assembled tetrahedron monomer (band 'b') by the purification procedure. 80% of the DNA could be recovered as monomeric tetrahedron from the column runs.

Optimization of the Annealing Steps: Concentration Dependence

In prior research[1], annealing concentrations of 5 μM or more were used to form DNA tetrahedron structures from for the single-stranded DNA oligonucleotides. However, at these concentrations, large qualities of aggregates are detected by size-exclusion chromatography and

native gel electrophoresis. As aggregation is a concentration dependent process, we tested if we can eliminate or reduce the formation of aggregates by varying the annealing concentration. During the annealing step, different concentrations of oligonucleotides were tested—ranging from 50 nM to 5 μM, and also explored different cooling times. With the variation of the cooling time we wanted to explore if aggregate formation can be reduced by slower cooling rates that provides more time for the tetrahedron structures to form. The results are shown in Figure S3 for cooling times of 30 seconds and 90 minutes and DNA concentrations of 50nM, 0.2 μM and 0.5μM. The higher concentration during annealing of the 20T and 20x30T tetrahedra resulted in a significant amount of aggregation, as shown in Figure S3a, lanes 1 and 3, while significantly lower amounts of aggregates are observed at the 0.2 μM concentration shown in Figure S3a, lanes 2 and 4. Figure S3b compares the results for the 50 nM and 0.2μM concentrations of DNA. The results shows that aggregates could be further reduced for the 20T and 20x30T construct by decrease of the annealing DNA concentration to 50nM (compare lanes 3/4 and 7/8 (50nM) to lanes 1/2 and 4/5 (0.2 μM). In contrast to the clear concentration dependence, the annealing time had no detectable influence on the formation of aggregates as shown in Fig S3b where very similar ratios of monomers and aggregates are detected at annealing times of 30 seconds (odd numbered lines) and 90 minutes (even-numbered lines in Fig S3b). As the lower concentration of oligonucleotides at the time of annealing reduced the amount of aggregates formed, all samples used for characterization and crystallization experiments were annealed at a concentration of 50 nM and further purified by size exclusion chromatography.

Dynamic Light Scattering (DLS) Characterization of 20T Tetrahedron

X-ray structure determination requires crystals, and monodisperse samples are highly desired prior to crystallization. Traces of oligomers formed during the annealing step at low DNA concentrations are qualitatively removed by size exclusion chromatography, but concentrating of the DNA tetrahedrons prior to crystallization could lead to reformation of aggregates. Dynamic Light Scattering (DLS) was used as a method to quantify the amount of aggregation present in the purified, concentrated DNA tetrahedron sample prior to crystallization. The DLS experiments revealed a hydrodynamic radius of approximately 4 nm (diameter 8 nm) for the monomeric 20T Tetrahedron, which confirms the expected size of our construct and these results are also consistent with the dimensions determined by cryo-EM [3]. The DLS results show a bi-modal size distribution of the DNA before size exclusion purification, with a monomer peak featuring a hydrodynamic radius of approx 35 ± 5 nm and an aggregate peak at radius of approx 370nm. DLS data are shown in most publications in form of "radius-corrected" scattering intensities (scattering intensity is increased with r^6) [4], which leads to suppression of aggregate peaks. However, we aimed for the detection of even small traces of aggregates and therefore we present the raw DLS data without "radius corrections" in Figure S4. Figure S4a shows the purified 20T tetrahedron at a concentration of 2.5 mg ml $^{-1}$ in 1X TAE-Mg^{2+} buffer. Two main peaks indicating a hyrodynamic radius of 4 nm and 100 nm appear. S4b shows the unpurified sample at a concentration of 2.5 mg ml $^{-1}$, indicating a more prominent aggregate peak. Figures S4c and S4d show the distribution plot of unpurified 20T (5c) showing more aggregates than the purified sample in S4d. The quantification of aggregates is summarized in Table 1. The evaluation of the data is based on peak integration, size of the aggregates and number of tetrahedrons per aggregate. The evaluation shows that the ratio of proteins that are present in from of aggregates versus monomers decreased from 3 10^{-4} before purification to 3.2 10^{-6} after

purification and concentration of the samples prior to crystallization. DLS can also be used to further study the crystallization process of the tetrahedron – from initial set up of the hanging drop, nucleation, and the growth of the crystals.[5]

Table S1. Data evaluation of dynamic light scattering results

	before SEC purification	purified, concentr. sample
∫ aggregate peak {au}	0.5	0.28
∫ of monomer peak {au}	0.09	0.22
radius of aggregate peak {nm}	38	143
radius monomer peak{nm)	3.5± 0.5	4.6± 0.5
$\int / r^6$ aggregate	$1.8\ 10^{-10}$	$3.17\ 10^{-14}$
$\int / r^6$ monomer	$5.4\ 10^{-5}$	$2.26\ 10^{-5}$
$\int / r^6$ aggregate/$\int / r^6$ monomer	$3.3\ 10^{-6}$	$1.4\ 10^{-9}$
MW of aggregate [6]	$8\ 10^6$	$1.84\ 10^8$
MW of monomer {Da}	$8\ 10^4$	$80\ 10^4$
no of monomers per aggregate	100	2300
$\int / r^6$ aggregate/$\int / r^6$ monomer x no of monomers per aggregate = ratio of number of DNA molecules present in aggregates/ no of DNA molecules in monomers	$3.3\ 10^{-4}$	$3.22\ 10^{-6}$

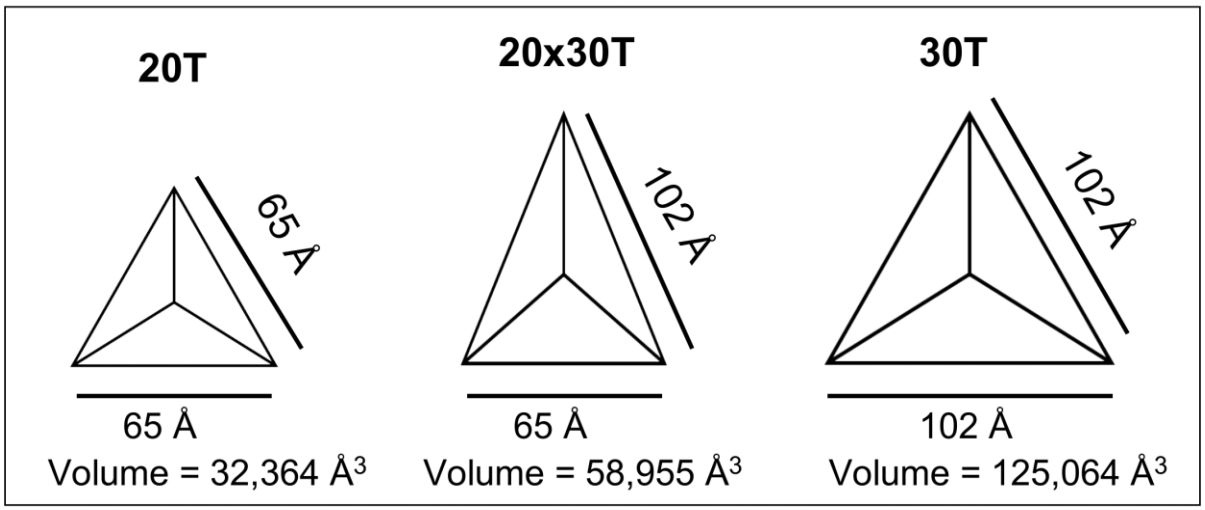


Figure S1. Dimensions for all three constructs.

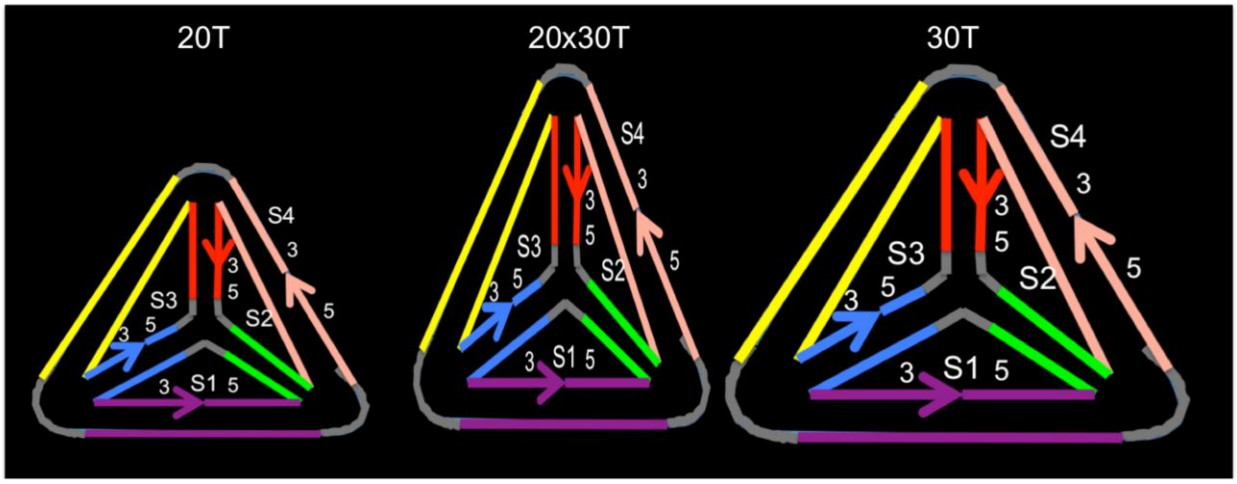


Figure S2. Assembly of all three constructs (20T, 20x30T, and 30T) with colors for each sequence corresponding to the formed double-stranded edge colors.

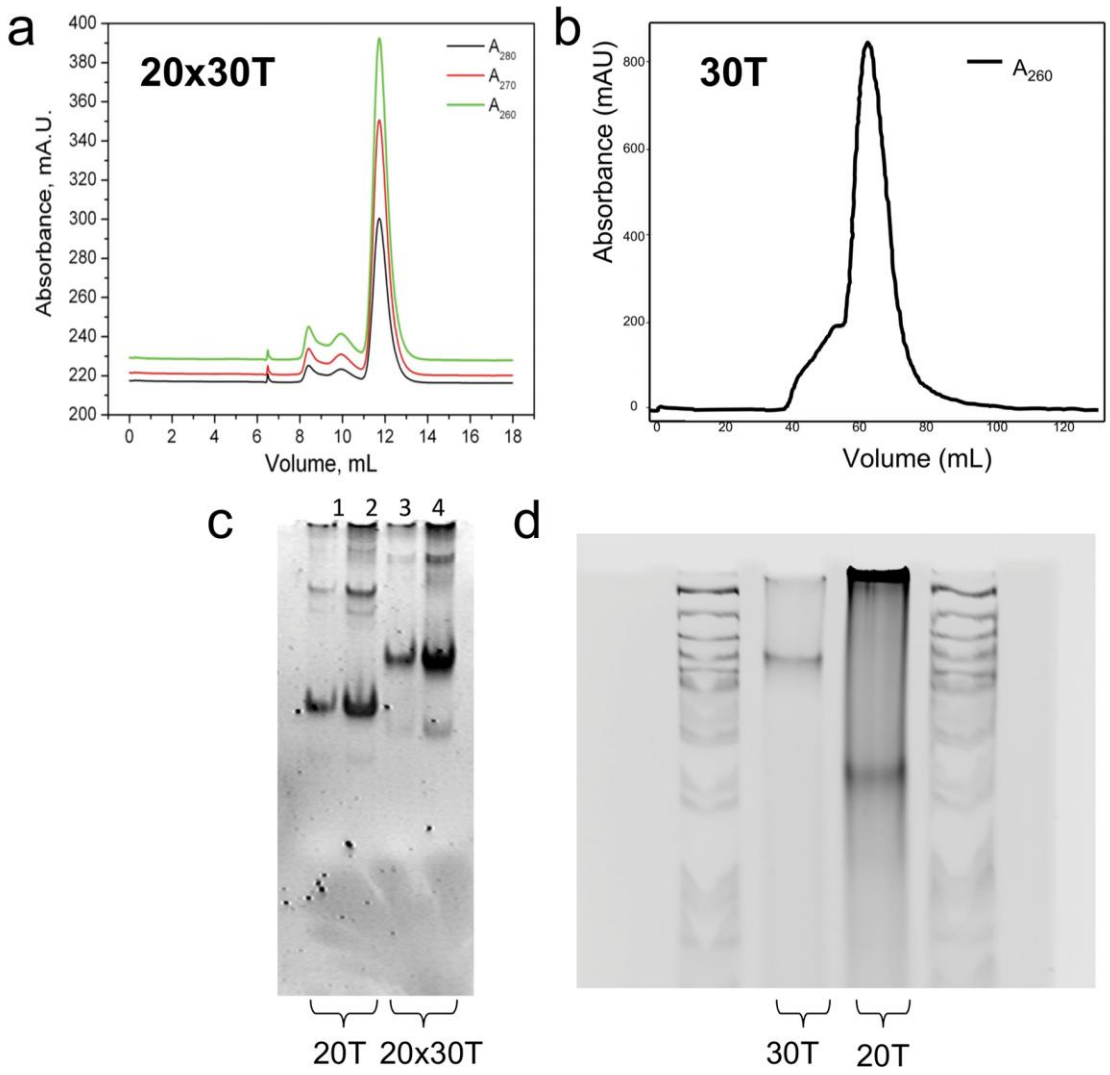


Figure S3. Purification of other constructs – 20x30T and 30T

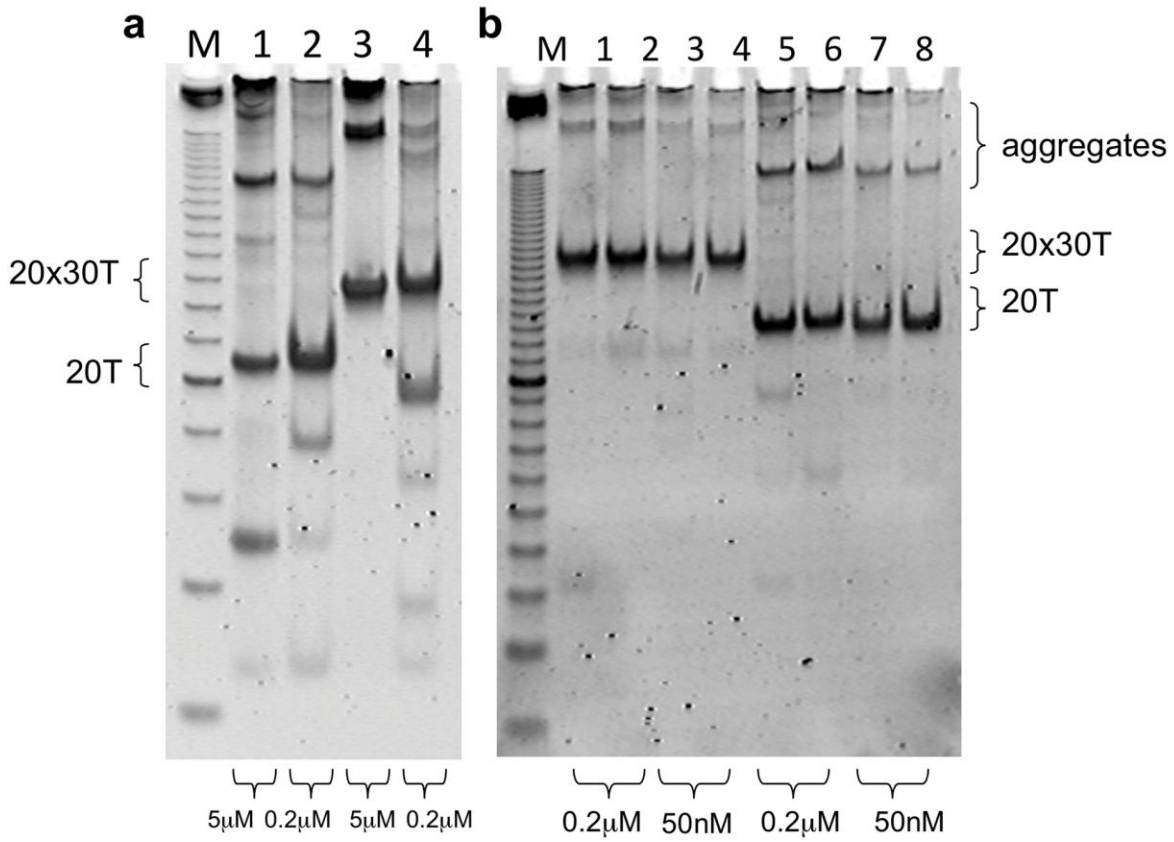


Figure S4. Concentration dependence during the annealing step

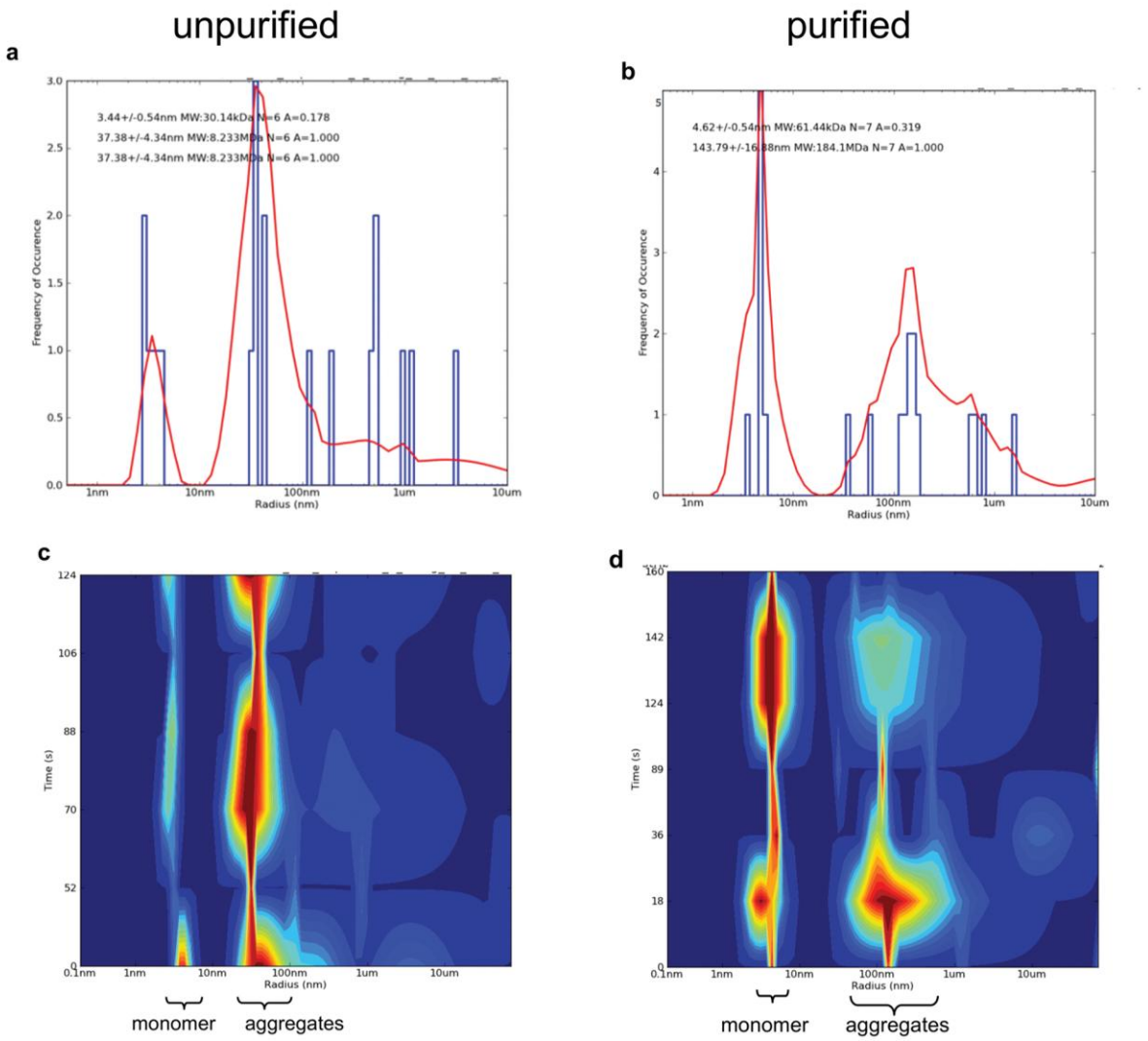


Figure S5. DLS results comparing unpurified and purified 20T tetrahedron

References

1. Goodman, R. P., Schaap, A. T., Tardin, C. F., Erben, C. M., Berry, R. M. et al. (2005). *Science*, **310**, 1661-1665.
2. Rupp, B. & Kantardjieff, K. A. (2008). *Molecular Biomethods Handbook*, edited by J. M. Walker & R. Rapley, pp. 821-849. Totowa: Humana Press.
3. Kato, T., Goodman, R. P., Erben, C. M., Turberfield, A. J. & Namba, K. (2009.) *Nano Letters*, **9**, 2747-2750.
4. Kok, C. M. & Rudin, A. (1981). *Makromol. Chem. Rapid Commun.* **2**, 655-659.
5. Dierks, K., Meyer, A., Einspahr, H. & Betzel, C (2008). *Crystal Growth Des.* **8**, 1628-1634.
6. Robertson, S. A., Harada, K., Frankel, A. D. & Wemmer, D. E. (2000). *Biochemistry*, **39**, 946-954.