CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/119001/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Schwenger, Alexander, Jurkowski, Tomasz P and Richert, Clemens 2018. Capturing and stabilizing folded proteins in lattices formed with branched oligonucleotide hybrids. ChemBioChem 19 (14) , pp. 1523-1530. 10.1002/cbic.201800145

Publishers page: https://doi.org/10.1002/cbic.201800145

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Capturing and Stabilizing Folded Proteins in Lattices Formed with Branched Oligonucleotide Hybrids

Alexander Schwenger[a], Tomasz P. Jurkowski[b] and Clemens Richert*[a]

Dedicated to Lothar Schwenger

Abstract: The encapsulation of folded proteins in stabilizing matrices is one of the challenges of soft matter material sciences. Capturing such fragile biomacromolecules from aqueous solution, and embedding them in a lattice that stabilizes them against denaturation and decomposition is difficult. Here we report that tetrahedral oligonucleotide hybrids as branching elements and connecting DNA duplexes with sticky ends can assemble into materials. The material-forming property was used to capture DNA-binding proteins selectively from aqueous protein mixtures. The three-dimensional networks also encapsulate guest molecules in size-selective manner, accommodating proteins up to a molecular weight of approx. 10 kDa for the connecting duplex lengths tested. Exploratory experiments with green fluorescent protein showed that, when embedded in the DNA-based matrix, the protein is more stable toward denaturation than in free form, retaining its luminescent properties for at least 90 days in dry form. The non-crystalline biohybrid matrices presented here may be used for capturing other proteins or for producing functional materials.

Introduction

Protein-nucleic acid complexes are important for the cell.[1,2,3] Nature uses proteins for packaging DNA or RNA to prevent its decomposition, e.g. in sperm cells [4,5] or virus assemblies, such as tobacco mosaic virus.[6] In the complex, the genetic material sometimes survives for years. Packaging proteins is an interesting challenge for which different approaches are being proposed.[7,8,9] It is important to meet this challenge, e.g. for capturing enzymes when they have finished catalyzing the biotransformation that they were supposed to induce. Often, remaining enzymes will destabilize a bioproduct, if not removed or captured. Likewise, recycling the enzyme can facilitate production and can reduced the cost of biotechnological processes.[10,11] So, there is a need for capturing, packaging, protecting and releasing proteins. Ideally, an approach for doing so should involve a non-toxic material that assembles into an encapsulating lattice of tunable dimension. If one was to use DNA as packing material to stabilize proteins, this would be an "inverse" approach to the packaging of genetic material in nature, mentioned above.

Given the incentives, it is not surprising that encapsulating proteins in a designed three-dimensional matrix is a long-standing goal of DNA nanotechnology. Encapsulation was expected to facilitate structure elucidation. The use of DNA lattices for this purpose was first proposed by Seeman in 1982,[12] and Seeman's proposal to use DNA for arraying proteins three-dimensionally is often seen as the starting point of DNA nanotechnology as a field. Possible solutions to the protein encapsulation challenge include DNA nanocages[13] and channels in DNA crystals. Different DNA crystals with continuous lattices formed by base pairing have been described. Crystals with hexagonal lattice geometry were found to form through self-assembly of a DNA tridecamer that engages in both canonical and parallel base pairing.[14] Another type of crystals contains designed lattices, with cohesion of the folding motifs via sticky ends.[15,16,17] Crystals of the former type have been shown to act as molecular sieves for proteins, with an adsorption cut-off of approx. 45 kDa.[18] Further, a small enzyme, ribonuclease A, was found to hydrolyze dinucleotides as substrates when captured in DNA crystals.[19] Finally, DNA hydrogels with encapsulated insulin have been shown to act as slow-release formulations.[20]

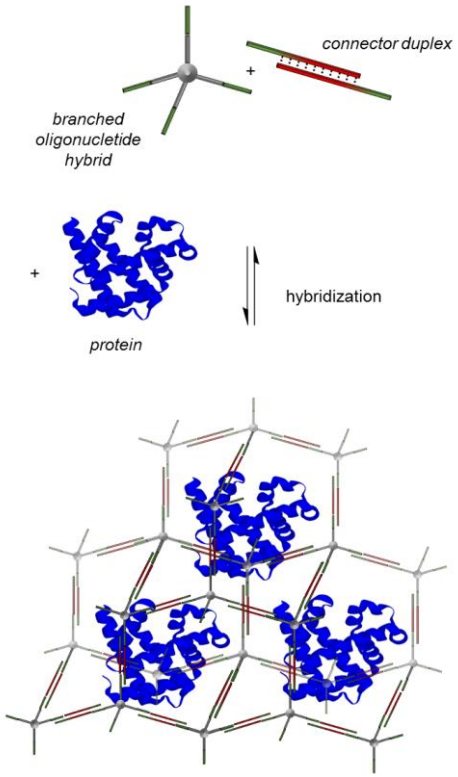


Figure 1. Cartoon representation of a process for capturing proteins in designed three-dimensional lattices, formed by oligonucleotides and connecting DNA duplexes. No highly ordered, crystalline structure is required for capturing proteins in such reticular networks.

[a] Dipl.-Chem. A. Schwenger, Prof. C. Richert,
Institut für Organische Chemie,
Universität Stuttgart, 70569 Stuttgart, Germany,
E-Mail: lehrstuhl-2@oc.uni-stuttgart.de
[b] Jun.-Prof. Tomasz Jurkowski,
Institut für Biochemie, Universität Stuttgart,
70569 Stuttgart, Germany
E-Mail tomasz.jurkowski@ibc.uni-stuttgart.de
Supporting information for this article is given via a link at the end of the document.

While the known nucleic acid-protein complexes indicate that DNA is suitable for protein encapsulation, no products that make use of the encapsulation approaches mentioned above have yet entered the market. Since crystallization is slow, and long DNA is expensive to produce on a large scale, there is room for alternative approaches. Synthetic compounds that utilize rigid cores and the predictable pairing properties of short DNA arms and that can be produced via efficient syntheses are one attractive approach.

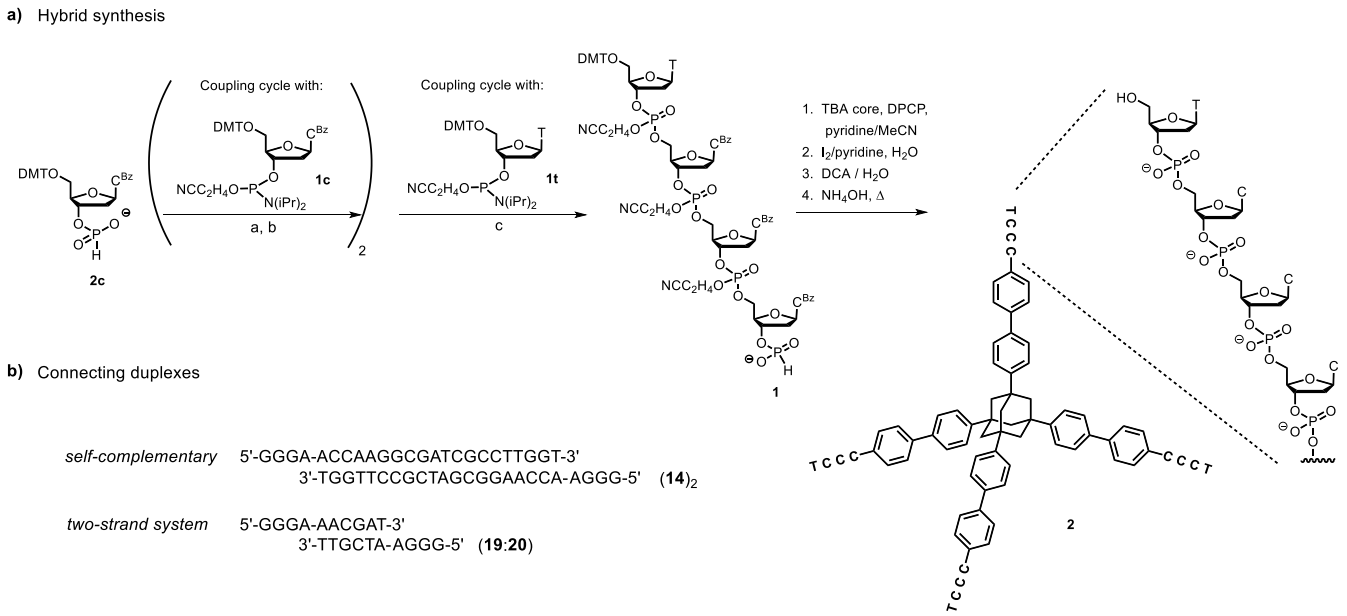
We have recently reported branched oligonucleotide hybrids with tetraaryladamantane cores that readily assemble into materials when their CG dimer arms are allowed to hybridize in dilute aqueous solution.[21,22] Solution-phase syntheses that are based on phosphoramidites[23,24] or *H*-phosphonates[25] were established that make these compounds inexpensive alternatives to constructs that are based on linear DNA strands. The propensity of hybrids to self-assemble can be tuned through changes in the structure of the organic core.[26] Expanded versions of such hybrids containing triplex binding motifs have been developed for capturing cofactors.[27] However, the pore size of the materials formed by the self-assembling hybrids with short sticky ends are too small to accommodate proteins.

Here we report the encapsulation of proteins in self-assembling matrices made up of branched oligonucleotide hybrids and connecting duplexes (Figure 1). The connecting duplexes allow for an adjusting of the size of the cavities in the three-dimensional lattices. Capture of DNA-binding proteins with affinity for the connecting duplex, and size-specificity for non-DNA-binding protein guest molecules were demonstrated. Macroscopic quantities of hybrid biomaterials are readily obtained and stabilization against denaturation was demonstrated by encapsulating green fluorescent protein (GFP).

Results and Discussion

Our encapsulating material was to be formed through self-assembly, with the adamantane-based branched oligonucleotide hybrids as the junction nodes that link connecting DNA duplexes in the three-dimensional lattices. For this, we first synthesized a new hybrid with tetramer arms, building upon our methodology for the synthesis of dimer *H*-phosphonate strands.[25] An overview of the synthetic elaboration of the *H*-phosphonate-terminated tetramer chain is shown in Scheme 1. The DNA sequence TCCC was chosen in order to obtain sufficiently stable duplexes with GGGA overhangs of connecting duplexes. Protected tetramer **1** was obtained in an overall yield of 61% on a gram scale without chromatography over 6 synthetic steps, as detailed in the Supporting Information (SI). With this tetramer *H*-phosphonate, we then assembled hybrid **2**, using tetrakis(*p*-hydroxybiphenyl)adamantane (TBA) as core and a modification of the coupling methodology described for hybrids with shorter arms.[25,26]

We then proceeded to testing whether hybrid **2** and DNA duplexes with complementary single-stranded overhangs form detectable assemblies and/or a macroscopically visible material. To induce assembly, the self-complementary DNA strands were first denatured using heat and alkaline conditions to break down any existing secondary structures, and subsequently allowed to pair by neutralizing with AcOH, addition of phosphate buffer (final concentration 10 mM, pH 7) and 1 M NaCl, and cooling at a rate of 0.5 °C/min. Duplex formation was monitored by UV-absorption (Figure S5, Supporting Information).



Scheme 1. Solution-phase synthesis of oligonucleotide hybrid $(TCCC)_4TBA$ (**2**) and sequences of selected connecting duplexes, other sequences are shown in Table 1. a) ACN, pyr-TFA, *t*-BuOOH and DCA/H_2O; b) DMF/ACN, pyr-TFA, *t*-BuOOH and DCA/H_2O; c) DMF/ACN, pyr-TFA and *t*-BuOOH.

After reaching the final temperature of 5 °C, hybrid **2** was added at a molar ratio of 1:2, (hybrid/duplex). At defined time intervals, absorption spectra were recorded at 5 °C for **2**:(**9**)$_2$, **2**:(**12**)$_2$, and **2**:(**17**)$_2$ (Figures S5 and S6; SI). Duplex-mediated lattice

formation was usually complete after 48 h. The kinetics depended on the length of the duplex in a non-trivial way, which may be due to topological issues. The absorption measurements in the supernatants showed that on average approx. 57% of the compounds precipitated under the conditions chosen.

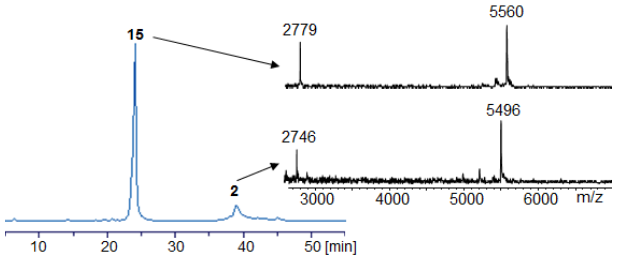


Figure 2. Compositional analysis of solid formed from **2** and connecting duplex (**15**)$_2$ upon annealing of the linear strands, adding **2**, assembling at 4 °C, isolating the precipitate, and washing. A sample of the solid was dissolved in hot deionized water and analyzed by HPLC (shown on the left) and MALDI-TOF mass spectrometry of the fractions (shown on the right).

5'-GGGACTGAACGATTAG-3' (**23**)
3'-GACTTGCTAATCAGGG-5' (**24**)

5'-GGGATACTGAACGATTAGTT-3' (**25**)
3'-ATGACTTGCTAATCAAAGGG-5' (**26**)

5'-GGGAACTGAACGATTAGT-3' (**27**)
3'-TGACTTGCTAATCAAGGG-5' (**28**)

5'-GGGATGAACGATTA-3' (**29**)
3'-ACTTGCTAATAGGG (**30**)

5'-GGGATGAA^{m}CGATTA-3' (**31**)
3'-ACTTG^{m}CTAATAGGG-3' (**32**)

5'-GGGACTGAAGCATTAG-3' (**37**)
3'-GACTTCGTAATCAGGG-3' (**38**)

5'-GGGACTGAA^{m}CGATTAG-3' (**33**)
3'-GACTTG^{m}CTAATCAGGG-3' (**34**)

5'-CCCTCTGAACGATTAG-3' (**39**)
3'-GACTTGCTAATCTCCC-3' (**40**)

5'-GGGAACTGAA^{m}CGATTAGT-3' (**35**)
3'-TGACTTG^{m}CTAATCAAGGG-3' (**36**)

5'-CCCTCTGAAGCATTAG-3' (**41**)
3'-GACTTCGTAATCTCCC-3' (**42**)

5'-GGGACTATGATCATAG-3' (**43**)

5'-GGGAACTGTAATTACAGT-3' (**44**)

5'-GGGAGATTATGATCATAATC-3' (**45**)

5'-GGGACGACTATAATTATAGTCG-3' (**46**)

5'-GGGAACCATATAATATTATATGGT-3' (**47**)

Figure 3. Additional sequences used for connecting duplexes.

Connecting duplexes with lengths ranging from 6 to 18 base pairs, equipped with GGGA overhangs were used. The duplexes are all at least 2 nanometers in length, as required for nanoscale lattices. All gave visible material at the end of the assays (Table 1). When samples were isolated and analyzed by HPLC and MALDI-TOF mass spectrometry, the analysis confirmed that the solids were composed of the hybrid and the duplex. Figure 2 shows a representative result for the solid formed from **2** and connecting duplex (**15**)$_2$, and Table 1 lists other data.

Assays were also carried out at a high concentration of duplex and hybrid (up to 330/165 µM). In each case, the material formed was isolated, washed, and analyzed via HPLC to determine its composition. A perfect lattice should contain duplexes and hybrids at a ratio of 2:1. The chromatograms showed that the materials consisted of the components at a ratio of 1.8-2.5/1 (duplex/hybrid), in reasonable agreement with the ideal stoichiometry of the lattice. The formation of large assemblies manifested itself in a hysteresis between heating and cooling transitions in UV-monitored analyses (see Figure S8, SI for a representative example). Taken together, these results suggested that assembly of hybrids and duplexes into three-dimensional networks is a robust process.

The DNA-based lattices were designed to offer cavities for harboring guest molecules of the size of protein enzymes. We therefore proceeded to studying the encapsulation of proteins in those lattices. For this, the salt concentrations in the assembly assay were lowered to be closer to physiological ionic strength. Here, assays were performed at 150 mM NaCl and ≤ 10 mM $MgCl_2$. Initially, we used small DNA-binding protein domains to test the capturing capabilities of the lattices. In order to further reduce the likelihood of possible hairpin structures that could inhibit the growth of the lattice, complementary DNA strands were used, rather than one self-complementary sequence. Four different connecting duplexes were used, namely **23**:**24**, **25**:**26**, **27**:**28**, and **29**:**30** (Figure 3). Lattice formation was again initiated by addition of hybrid **2** and was monitored by absorption measurements. Figure 4a shows representative results. When both **2** and a full connecting duplex were present, the formation of a lattice was observed in a time-dependent manner. When **2** and just one of the two strands, without the complementary strand, were used, no material formed.

Table 1. Formation of solids upon assembly of hybrid **2** and connecting duplexes of different length into networks, as monitored by UV-absorption.

Entry No	DNA strand	Duplex [µM]	Hybrid [µM]	buffer	Fraction precipitated (%)	Compound ratio in solid (duplex/hybrid)
1	--	-	35		--	--
2	5'-GGGA-CGATCG-3' (**9**)	69	35		54	1.4 / 1
3	5'-GGGA-GCGATCGC-3' (**10**)	36	18		28	1.8 /1
4	5'-GGGA-TGCGATCGCA-3' (**11**)	63	31	A[a]	45	1.8 / 1
5	5'-GGGA-CTGCGATCGCAG-3' (**12**)	39	19		57	2.5 / 1
6	5'-GGGA-GACTGCGATCGCAGTC-3' (**13**)	32	16		34	2.1 / 1
7	5'-GGGA-ACCAAGGCGATCGCCTTGGT-3' (**14**)	69	35		55	2.3 / 1
8	5'-GGGA-ACTGTAATTACAGT-3' (**15**)	263	132		66	1.5/1
9	5'-GGGA-GATTATGATCATAATC-3' (**16**)	310	155		66	2.2/1
10	5'-GGGA-CGACTATAATTATAGTCG-3' (**17**)	317	158		67	2.5/1
11	5'-GGGA-ACCATATAATATTATATGGT-3' (**18**)	167	83	B[b]	58	1/1
12	5'-GGGA-AACGAT-3' (**19**), 5'-GGGA-ATCGTT-3' (**20**)	330	165		85	2.1/1
13	**19**:**20**	333	--		--	--
14	5'-CCCT-AACGAT-3' (**21**), 5'-CCCT-ATCGTT-3' (**22**)	333	167		--	--
15	**21**:**22**	333	--		--	--

[a] Carried out in Assembly Buffer A; 10 mM phosphate buffer, 1 M NaCl); [b] Assembly Buffer B; 10 mM HEPES, 1 M NaCl.

At the salt concentration chosen, slightly less material precipitated than in the experiments with the self-complementary sequences. Still, lattice formation was strong enough to proceed to encapsulation tests with proteins. For this, after annealing of the linear DNA duplexes, either the DNA-binding protein

CxxC,[28,29,30] or aldolase, a small metabolic enzyme were added, followed by the addition of **2**.

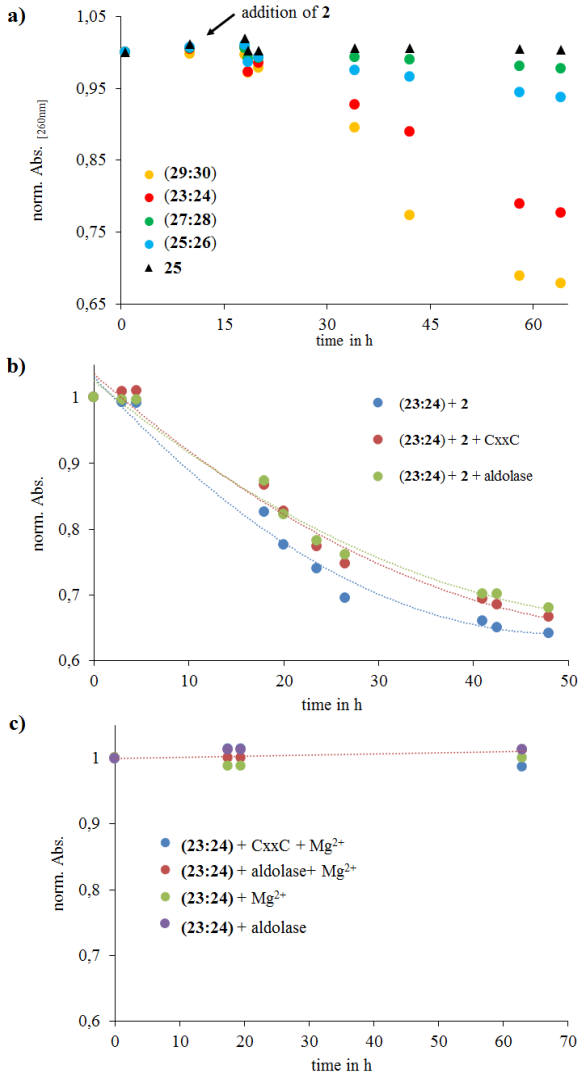


Figure 4. Kinetics of the formation of solids upon assembly into three-dimensional lattices, as detected by UV-absorption at 260 nm, after addition of hybrid **2**. a) Hybrid **2** and connecting duplexes formed from pairs of complementary strands of different length. The total duplex length is as defined in Figure 5. b) Kinetics of formation of solids upon assembly into three-dimensional networks, without any protein, with CxxC, or with aldolase, as monitored by the decrease in absorbance; c) absorption of solutions of controls. Conditions: 3 µM strands, 1.5 µM **2**, 0.3 µM protein in 20 mM HEPES buffer, 150 mM NaCl and 10 mM $MgCl_2$.

The solutions showed a drop in absorption of up to 35% for both CxxC and aldolase under these condition (1.5 µM hybrid **2**, 3 µM duplex **23**:**24** and 0.3 µM protein) after an incubation time of 48 h (Figure 4b). The kinetics of assembly were similar to those found for the assembly of the DNA components alone, suggesting that interactions driving network formation are DNA-DNA interactions, not interactions between oligonucleotides and proteins. The control experiments, in which the hybrid was left out, showed no change in absorption (Figure 4c). After the composite material had assembled, as shown in Figure 4b, some samples were slowly heated to 85°C. The results from this thermal analysis indicate that the lattices start to re-dissolve at approx. 35 °C (Figure S7, SI). The melting curves also showed more than one transition when proteins were encapsulated, as expected for assemblies with a more complex morphology than that of duplexes and hybrid alone.

The next experiments on protein encapsulation used CxxC (11 kDa) and Mbd2[31,32,33] (10 kDa), another small DNA-binding protein with a different specificity. Connecting duplexes consisted of strands **31**-**47** (Figure 3). This group of duplexes included sequences containing a CG dinucleotide featuring either 5-methylcytosine (^{m}CG) or unmethylated cytosine (CG). The former is a recognition motif for Mbd2, and the latter for CxxC. Material formation in the presence of increasing concentrations of the proteins was studied. The solids obtained were separated from the mother liquor by centrifugation and extensively washed with buffer to remove unentrapped proteins. The composition of the liquid and the solid fractions was determined by gel electrophoresis. We observed high yields of encapsulation of either of the DNA binding proteins in the solid (Figure S9, SI). The gel analysis of the solids showed that the DNA-binding proteins were encapsulated (Table S1 and Figure S9, SI). Uptake of up to 90% was achieved under these conditions (20 mM HEPES, 150 mM NaCl and 10 mM $MgCl_2$). The selectivity of CxxC and Mbd2 for their cognate sequences was modest, though, suggesting that at the concentrations chosen, both proteins had sufficient affinity for the DNA duplexes.

This prompted us to test whether the lattice formed by connecting duplexes and **2** can capture proteins without significant affinity for DNA, by simple physical entrapment. This entrapment was expected to be size-limited, and proteins too large to fit into the pores of the three-dimensional network were expected to be selected against. To test this, we used a set of proteins with sizes ranging from 12 kDa (cytochrome C) to 529 kDa for (ferritin complex), as shown in Figure 5. The diameters of the proteins range from 3.8 nm to 13 nm. The proteins used in the entrapment experiments were expected not to interact strongly with DNA electrostatically or via a specific complex.

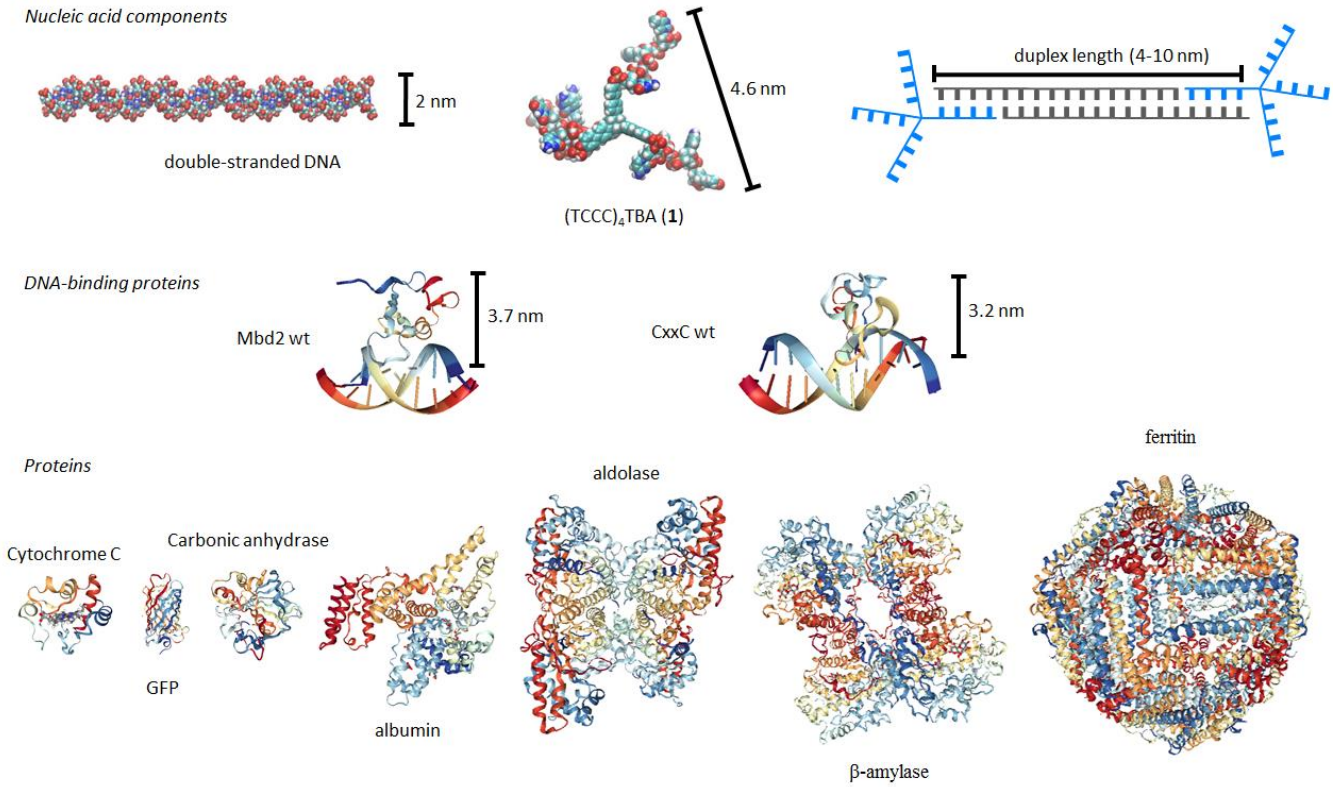


Figure 5. Structure of components employed in encapsulation assays, roughly drawn to scale. For proteins, the structures were generated from pdb entries 2KY8[34] (Mbd2, 3.7 nm /10 kDa), 4NW3[35] (CxxC, 3.2 nm /11 kDa), 5TY3[36] (cytochrome c, 3.8 nm /12 kDa), 1GFL[37] (GFP, 3,8 nm /27 kDa), 2CAB[38] (carbonic anhydrase, 4.3 nm /29 kDa), 4LUF[39] (albumin, 8.3 nm /68 kDa), 1ALD[40] (alodlase, 11 nm /159 kDa), 1FA2[41] (amylase, 12 nm /226 kDa), and 1LB3[42] (ferritin, 13 nm /529 kDa). For double-stranded DNA, a canonical B-form duplex was generated in Maestro, version 7.5.106, for hybrid **2**, an energy minimized structure generated in Chem3D Pro, version 14.0 is shown. The latter two graphics were generated in VMD.[43]

Table 2. Results from encapsulation assays with proteins of different size.

protein	diameter (nm)	Molecular weight (kDa)	total duplex length[b] conc. **2** [µM][c]	Encapsulation (%)[a] 14 bp (**19**:**20**)		20 bp (**23**:**24**)	
				27	20	20	30
cyctochrome C	3.8	12		28	29	26	29
GFP	3.8	27		n. d.		16	30[e]
carbonic anhydrase	4.3	29		12	16	8	15
albumin	8.3	68		23	28	17	19
aldolase	11	159		22	12	46	34
β-amylase	12	226		14	5	4	< 5
ferritin	13	529		<7	<2	<2	<10

a) As determined by integration of band intensities in SDS PAGE of supernatant and material fractions; b) as defined in Figure 5; c) connecting duplex strands at twice the concentration of the hybrid, as required by the stoichiometry of the 3D lattice. d) GFP encapsulation was performed with the self-complementary strand GGGAACTGTAATTACAGT (**15**) as connecting duplex; e) at a hybrid concentration of 39 µM.

Encapsulation of the proteins showed a size dependence reminiscent of that of gel permeation or size exclusion chromatography, albeit at much larger sizes than in GPC. Small proteins were entrapped, whereas proteins with sizes above the lattice dimensions were excluded from the matrix formed upon assembly. The size exclusion correlated with the length of the connecting DNA duplexes. The upper limit for the lattice with **23**:**24** as connecting duplex was approx. 159 kDa, and with the duplex **19**:**20** it was approx. 68 kDa, with aldolase no longer being well accepted into the three-dimensional network. The comparison of the trapping efficiency of CxxC and Mbd2 versus those of the non-DNA binding proteins, shows that the DNA binding proteins are more efficiently encapsulated in the DNA material in this experiments, as expected due to their affinity for the duplexes.

For biotechnological or biomedical applications, it is interesting to stabilize enzymes against denaturation and degradation, and to preserve their integrity and function during storage and transport. To see whether such an effect can be achieved with our lattices, we performed exploratory tests with purified recombinant green fluorescent protein (GFP). Here, the integrity of the protein can be monitored by its fluorescence, which simplifies analysis. When GFP was encapsulated in the DNA lattices, it remained strongly fluorescent. In contrast to the uncoated control sample, which turned non-fluorescent within

one day, the solid with encapsulated GFP remained fluorescent over more than three months, when left on the bench under ambient condition, suggesting that the DNA lattice protects the protein against unfolding and decomposition (Figure 6).

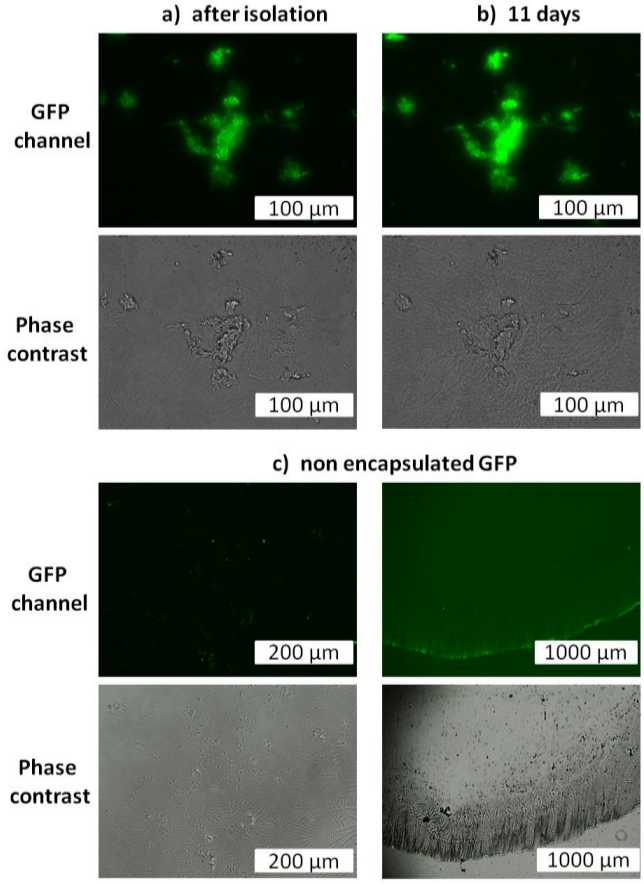


Figure 6. Microscopic photographs of GFP encapsulated in the DNA lattice, a) shortly after isolation of the material, and b) after 11 days at room temperature, and c) non encapsulated GFP at the same time points under the same storage conditions.

Most probably, the composites formed from the DNA-based lattice and the proteins trap substantial amounts of hydration water, helping to preserve the native state, even when allowed dried. The DNA scaffold may inhibit unfolding, together helping to explain the high stability of the encapsulated GFP, compared to an unentrapped one, which probably denatured quickly when the sample was left to dry. Taken together, the results suggest that our encapsulation method produces a solid that has the potential to stabilize sensitive folded proteins, even when allowed to dry, upon exposure to air at room temperature.

The elucidation of structural details of our reticular networks is a challenge that will not be easy to match. We do not assume that the DNA-based lattices have crystalline order. But, the formation of macromolecular assemblies manifests itself readily in a hysteresis between heating and cooling curves in UV-melting transitions, as shown for a sample of hybrid **2** and connecting duplex (**9**)$_2$ in Figure S8 of the Supporting Information. This type of analysis can be performed on the time scale of hours, and does not require special equipment or crystallization set-ups.

Conclusions

Here we report a novel DNA-based, nano-structured material with cavities large enough to harbor proteins as molecular guests. The size of the cavities in this material can be tuned by using appropriate connecting duplexes. Both those duplexes and the branched oligonucleotide hybrid **2** are readily accessible by organic synthesis. When the lattices form, they have a modest and tunable propensity to precipitate from aqueous buffer, which allows for the capturing of proteins. Both DNA-binding and non-DNA binding proteins were encapsulated in our study, the former with higher efficiency, as expected for interactions beyond simple physical entrapment. Protein entrapment depends on size, making the lattices selective and the process potentially useful for purification. Under optimized conditions, the DNA binding proteins CxxC and Mbd2 were harvested almost quantitatively from solution. Through encapsulation, proteins may be protected from thermal denaturation and other processes that render them inactive when exposed to air, as demonstrated for GFP. Since the lattice can be dissociated by gentle warming, the material also has the potential to act as controlled release medium. Given the versatility of the design and its unique properties, we expect the hybrid lattices with their tunable lattice constants to be useful for practical applications.

Experimental Section

Encapsulation. The following protocol is representative. A mixture of solutions of strand **23** (25 µL, 8 nmol, 5'GGGA-CTGAACGATTAG-3'), strand **24** (27 µL, 8 nmol, 5'GGGA-CTAATCGTTCAG-3'), water (82 µL), HEPES buffer (20 mM, 6 µL of 0.5 M stock solution) and NaCl (150 mM, 4.5 µL of 5 M stock) was placed in a polyropylene microtube. Hybridization was induced by heating to 85 °C and cooling to 4 °C in 8 h. Then, the protein Mbd2 (0.3 µL, 0.3 nmol) in elution buffer was added, mixed, and the sample was incubated for 30 min at 4 °C. A solution of hybrid **2** (4.2 µL, 4 nmol) was added, and the mixture was kept at 4 °C for 30 min, followed by the addition of 10 mM $MgCl_2$ (1.5 µL of a 1M stock solution). The sample was stored at 4 °C for 48 h. The solid was then separated from the supernatant by centrifugation (4 °C, 30 min, 21500 g). The supernatant was removed, and the solid was washed with cold buffer (100 µL; 20 mM HEPES, 150 mM NaCl, 10 mM $MgCl_2$), and centrifuged again. The supernatant was again aspirated and the material harvested.

Acknowledgements

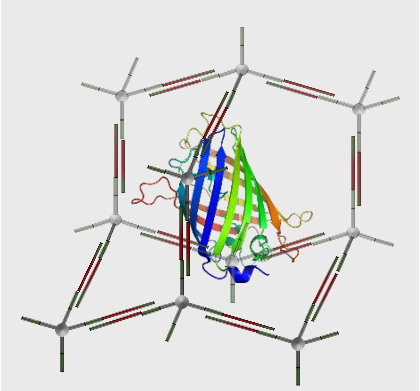
The authors thank D. Jovanovic for help with melting curves and H. Griesser for a review of the manuscript. This work was supported by DFG, grant No. RI 1063 15-1 to C.R.

Keywords: Proteins • DNA • soft matter • oligonucleotide hybrids • nanoscale lattices

1. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell, 4th ed.*; New York: Garland Science; 2002.
2. T. Schalch, S. Duda, D. F. Sargent, T. J. Richmond, *Nature* **2005**, *436*, 138-141.
3. A. Sarai, H. Kono, *Ann. Rev. Biophys.* **2005**, *34*, 379-398.
4. J. Erenpreiss, M. Spano, J. Erenpreisa, M. Bungum, A. Giwercman, *Asian J. Androl.* **2006**, *8*, 11–29.
5. A. Zini, O. Albert, B. Robaire, *Andrology* **2014**, *2*, 322-325.
6. P. Goelet, G. P. Lomonossoff, P. J. Butler, M. E. Akam, M. J. Gait, J. Karn, *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5818-5822.
7. J. A. Trelles, M. J. Lapponiab, *Curr. Pharm. Des.* **2017**, ASAP, doi: 10.2174/1381612824666171204102204.
8. Q. Sun, C. W. Fu, B. Aguila, J. Perman, S. Wang, H. Y. Huang, F. S. Xiao, S. Ma, *J. Am. Chem. Soc.* **2018**, *140*, 984–992.
9. B. J. Jones, H. Y. Lim, J. Huang, R. J. Kazlauskas, *Biochemistry* **2017**, *56*, 6521-6532.
10. M. Jin, Y. Liu, L. da Costa Sousa, B.E. Dale, V. Balan, *Biotechnol Bioeng.* **2017**, *114*, 1713-1720.
11. J. Cui, Z. Tan, P. Han, C. Zhong, S. Jia, *J. Agric. Food Chem.* **2017**, *17*, 3883-3890.
12. N. C. Seeman, *J. Theor. Biol.* **1982**, *99*, 237-247.
13. C. M. Erben, R. P.Goodman, A. J. Turberfield, *Angew. Chem.* **2006**, *118*, 7574-7577.
14. P. J. Paukstelis, J. Nowakowski, J. J. Birktoft, N. C. Seeman, *Chem. Biol.* **2004**, *11*, 1119–1126.
15. J. Zheng, J. J. Birktoft, Y. Chen, T. Wang, R. Sha, P. E. Constantinou, S. L. Ginnell, C. Mao, N. C. Seeman, *Nature* **2009**, *461*, 74-77.
16. R. Sha, J. J. Birktoft, N. Nguyen, A. R. Chandrasekaran, J. Zheng, X. Zhao, C. Mao, N. C. Seeman, *Nano Lett.* **2013**, *13*, 793-797.
17. C. R. Simmons, F. Zhang, J. J. Birktoft, X. Qi, D. Han, Y. Liu, R. Sha, H. O. Abdallah, C. Hernandez, Y. P. Ohayon, N. C. Seeman, and H.Yan, *J. Am. Chem. Soc.* **2016**, *138*, 10047–10054.
18. P. J. Paukstelis, *J. Am. Chem. Soc.* **2006**, *128*, 6794-6795.
19. Chun Geng and Paul J. Paukstelis, *J. Am. Chem. Soc.* **2014**, *136*, 7817–7820.
20. Soong Ho Um, Jong Bum Lee, Nokyoung Park, Sang Yeon Kwon, Christopher C. Umbach and Dan Luo, *Nat. Materials* **2006**, *5*, 797- 801.
21. M. Meng, C. Ahlborn, M. Bauer, O. Plietzsch, S. A. Soomro, A. Singh, T. Muller, W. Wenzel, S. Bräse, C. Richert, *ChemBioChem* **2009**, *10*, 1335-1339.
22. A. Singh, M. Tolev, M. Meng, K. Klenin, O. Plietzsch, C. I. Schilling, T. Muller, M. Nieger, S. Bräse, W. Wenzel, C. Richert, *Angew. Chem.* **2011**, *123*, 3285-3289; *Angew. Chem., Int. Ed.* **2011**, *50*, 3227-3231.
23. H. Griesser, M. Tolev, A. Singh, T. Sabirov, C. Gerlach, C. Richert,. *J. Org. Chem.* **2012**, *77*, 2703-2717.
24. A. Schwenger, N. Birchall, C. Richert, *Eur. J. Org. Chem.* **2017**, *39*, 5852-5864.
25. A. Singh, M. Tolev, C. Schilling, S. Bräse, H. Griesser, C. Richert, *J. Org. Chem.* **2012**, *77*, 2718-2728.
26. A. Schwenger, C. Gerlach, H. Griesser, C. Richert, *J. Org. Chem.* **2014**, *79*, 11558-11566.
27. M. Kalinowski, R. Haug, H.Said, S. Piasecka, M. Kramer, C. Richert, *ChemBioChem* **2016**, *17*,1150-1155.
28. T. Cierpicki, L. E. Risner, J. Grembecka, S. M. Lukasik, R. Popovic, M. Omonkowska, D. S. Shultis, N. J. Zeleznik-Le, J. H. Bushweller, *Nat. Struct. Mol. Biol.* **2010**, *17(1)*, 62-68.
29. P.M. Ayton, E. H. Chen, M. L. Cleary*, Mol Cell Biol.* **2004**, *24(23)*, 10470-10478.
30. J.-H. Lee, K. S. Voo, D. G. Skalnik, *J. Biol. Chem.* **2001**, *276*, 44669-44676.
31. Y. Yu, S. Blair, D. Gillespie, R. Jensen, D. G. Myszka, A. H. Badran, I. Ghosh, A. Chagovetz. *Anal Chem.* **2010**, *82*(12), 5012-5019.
32. T. Clouaire, I.Stancheva *Cell Mol Life Sci.* **2008**, *65*(10), 1509-1522.
33. X. Zou, W. Ma, I. A. Solov'yov, C. Chipot, K. Schulten. *Nucleic Acids Res.* **2012**, *40*(6), 2747-2758.
34. J. N. Scarsdale, H. D. Webb, G. D. Ginder, D. C. Williams, *Nucleic Acids Res.* **2011**, *39*, 6741-6752.
35. L. Tang, T. M. Gamal El-Din, J. Payandeh, G. Q. Martinez, T. M. Heard, T. Scheuer, N. Zheng, W. A. Catterall, *Nature* **2014**, *505*, 56-61.
36. S. M. Nold, H. Lei, T. C. Mou, B. E. Bowler, *Biochemistry* **2017**, *56*, 3358-3368.
37. F. Yang, L. G. Moss, Jr. G. N. Phillips, *Nat. Biotechnol*. **1996**, *14*, 1246-1251.
38. K. K. Kannan, M. Ramanadham, T. A. Jones, *Ann. N. Y. Acad. Sci.* **1984**, *429*, 49-60.
39. A. Bujacz, J. A. Talaj, A. J. Pietrzyk, DOI: 10.2210/pdb4luf/pdb
40. S. J. Gamblin, G. J. Davies, J. M. Grimes, R. M. Jackson, J. A. Littlechild, H. C. Watson, *J. Mol. Biol.* **1991**, *219*, 573-576.
41. C. G. Cheong, S. H. Eom, C. Chang, D. H. Shin, H. K. Song, K. Min, J. H. Moon, K. K. Kim, K. Y. Hwang, S. W. Suh, *Proteins* **1995**, *21*, 105-117.
42. T. Granier, B. Langlois D'Estaintot, B. Gallois, J.-M. Chevalier, G. Precigoux, P. Santambrogio, P. Arosio*, J. Biol. Inorg. Chem.* **2003**, *8*, 105-111.
43. W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33–38.

Entry for the Table of Contents

Text for Table of Contents



*A. Schwenger, T. P. Jurkowski, C. Richert**

Page No. – Page No.

Capturing and Stabilizing Folded Proteins in Lattices Formed with Branched Oligonucleotide Hybrids

1. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell, 4th ed.*; New York: Garland Science; 2002.
2. T. Schalch, S. Duda, D. F. Sargent, T. J. Richmond, X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **2005**, *436*, 138-141.
3. A. Sarai, H. Kono, Protein-DNA recognition patterns and predictions. *Ann. Rev. Biophys.* **2005**, *34*, 379-398.
4. J. Erenpreiss, M. Spano, J. Erenpreisa, M. Bungum, A. Giwercman. Sperm chromatin structure and male fertility: biological and clinical aspects. *Asian J. Androl.* **2006**, *8*, 11–29.
5. A. Zini, O. Albert, B. Robaire, Assessing sperm chromatin and DNA damage: clinical importance and development of standards. *Andrology* **2014**, *2*, 322-325.
6. P. Goelet, G. P. Lomonossoff, P. J. Butler, M. E. Akam, M. J. Gait, J. Karn, Nucleotide sequence of tobacco mosaic virus RNA. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5818-5822.
7. J. A. Trelles, M. J. Lapponiab, Immobilization techniques applied to the development of biocatalysts for the synthesis of nucleoside analogue derivatives. *Curr. Pharm. Des.* **2017**, ASAP, doi: 10.2174/1381612824666171204102204.
8. Q. Sun, C. W. Fu, B. Aguila, J. Perman, S. Wang, H. Y. Huang, F. S. Xiao, S. Ma, Pore Environment Control and Enhanced Performance of Enzymes Infiltrated in Covalent Organic Frameworks. *J. Am. Chem. Soc.* **2018**, *140*, 984–992.
9. B. J. Jones, H. Y. Lim, J. Huang, R. J. Kazlauskas, Comparison of Five Protein Engineering Strategies for Stabilizing an α/β-Hydrolase. *Biochemistry* **2017**, *56*, 6521-6532.
10. M. Jin, Y. Liu, L. da Costa Sousa, B.E. Dale, V. Balan, Development of rapid bioconversion with integrated recycle technology for ethanol production from extractive ammonia pretreated corn stover. *Biotechnol Bioeng.* **2017**, *114*, 1713-1720.
11. J. Cui, Z. Tan, P. Han, C. Zhong, S. Jia, Enzyme Shielding in a Large Mesoporous Hollow Silica Shell for Improved Recycling and Stability Based on CaCO3 Microtemplates and Biomimetic Silicification. *J. Agric. Food Chem.* **2017**, *17*, 3883-3890.
12. N. C. Seeman, Nucleic Acid Junctions and Lattices *J. Theor. Biol.* **1982**, *99*, 237-247.
13. Erben, C. M.; Goodman, R. P.; Turberfield, A. J. Single-molecule protein encapsulation in a rigid DNA cage. *Angew. Chem.* **2006**, *118*, 7574-7577.
14. P. J. Paukstelis, J. Nowakowski, J. J. Birktoft, N. C. Seeman, Crystal structure of a continuous three-dimensional DNA lattice. *Chem. Biol.* **2004**, *11*, 1119–1126.
15. J. Zheng, J. J. Birktoft, Y. Chen, T. Wang, R. Sha, P. E. Constantinou, S. L. Ginnell, C. Mao, N. C. Seeman From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009**, *461*, 74-77.
16. R. Sha, J. J. Birktoft, N. Nguyen, A. R. Chandrasekaran, J. Zheng, X. Zhao, C. Mao, N. C. Seeman Self-assembled DNA crystals: the impact on resolution of 5'-phosphates and the DNA source. *Nano Lett.* **2013**, *13*, 793-797.
17. C. R. Simmons, F. Zhang, J. J. Birktoft, X. Qi, D. Han, Y. Liu, R. Sha, H. O. Abdallah, C. Hernandez, Y. P. Ohayon, N. C. Seeman, and H.Yan, Construction and Structure Determination of a Three-Dimensional DNA Crystal. *J. Am. Chem. Soc.* **2016**, *138*, 10047–10054.
18. Paukstelis, P. J. Three-dimensional DNA crystals as molecular sieves. *J. Am. Chem. Soc.* **2006**, *128*, 6794-6795.
19. Chun Geng and Paul J. Paukstelis, DNA Crystals as Vehicles for Biocatalysis. *J. Am. Chem. Soc.* **2014**, *136*, 7817–7820.
20. Soong Ho Um, Jong Bum Lee, Nokyoung Park, Sang Yeon Kwon, Christopher C. Umbach and Dan Luo, Enzyme-catalysed assembly of DNA hydrogel. *Nature Materials* **2006**, *5*, 797- 801.
21. M. Meng, C. Ahlborn, M. Bauer, O. Plietzsch, S. A. Soomro, A. Singh, T. Muller, W. Wenzel, S. Bräse, C. Richert, *ChemBioChem* **2009**, *10*, 1335-1339.
22. A. Singh, M. Tolev, M. Meng, K. Klenin, O. Plietzsch, C. I. Schilling, T. Muller, M. Nieger, S. Bräse, W. Wenzel, C. Richert, *Angew. Chem.* **2011**, *123*, 3285-3289; *Angew. Chem., Int. Ed.* **2011**, *50*, 3227-3231.

23. H. Griesser, M. Tolev, A. Singh, T. Sabirov, C. Gerlach, C. Richert,. *J. Org. Chem.* **2012**, *77*, 2703-2717.
24. A. Schwenger, N. Birchall, C. Richert, Solution-phase synthesis of branched oligonucleotides with up to 32 nucleotides and the reversible formation of materials. *Eur. J. Org. Chem.* **2017**, *39*, 5852-5864.
25. A. Singh, M. Tolev, C. Schilling, S. Bräse, H. Griesser, C. Richert, *J. Org. Chem.* **2012**, *77*, 2718-2728.
26. A. Schwenger, C. Gerlach, H. Griesser, C. Richert, *J. Org. Chem.* **2014**, *79*, 11558-11566.
27. M. Kalinowski, R. Haug, H.Said, S. Piasecka, M. Kramer, C. Richert, *ChemBioChem* **2016**, *17*,1150-1155.
28. T. Cierpicki, L. E. Risner, J. Grembecka, S. M. Lukasik, R. Popovic, M. Omonkowska, D. S. Shultis, N. J. Zeleznik-Le, J. H. Bushweller „*Structure of the MLL CXXC domain – DNA complex and its functional role in MLL-AF9 leukemia" Nat Struct Mol Biol*. **2010**, *17(1)*, 62-68.
29. P.M. Ayton, E. H. Chen, M. L. Cleary *"Binding to Nonmethylated CpG DNA Is Essential for Target Recognition, Transactivation, and Myeloid Transformation by an MLL Oncoprotein" Mol Cell Biol.* **2004**, *24(23)*, 10470-10478.
30. J.-H. Lee, K. S. Voo, D. G. Skalnik *"Identification and Characterization of the DNA Binding Domain of CpG-binding protein" J. Biol. Chem.* **2001**, *276*, 44669-44676.
31. Y. Yu, S. Blair, D. Gillespie, R. Jensen, D. G. Myszka, A. H. Badran, I. Ghosh, A. Chagovetz "Direct DNA Methylation Profiling Using Methyl Binding Domain Proteins" *Anal Chem.* **2010**, *82*(12), 5012-5019.
32. T. Clouaire, I.Stancheva "Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin?" *Cell Mol Life Sci.* **2008**, *65*(10), 1509-1522.
33. X. Zou, W. Ma, I. A. Solov'yov, C. Chipot, K. Schulten *"Recognition of methylated DNA through methyl-CpG binding domain proteins" Nucleic Acids Res.* **2012**, *40*(6), 2747-2758.
34. J. N. Scarsdale, H. D. Webb, G. D. Ginder, D. C. Williams, *Nucleic Acids Res.* **2011**, *39*, 6741-6752.
35. L. Tang, T. M. Gamal El-Din, J. Payandeh, G. Q. Martinez, T. M. Heard, T. Scheuer, N. Zheng, W. A. Catterall, *Nature* **2014**, *505*, 56-61.
36. S. M. Nold, H. Lei, T. C. Mou, B. E. Bowler, *Biochemistry* **2017**, *56*, 3358-3368.
37. F. Yang, L. G. Moss, Jr. G. N. Phillips, *Nat. Biotechnol*. **1996**, *14*, 1246-1251.
38. K. K. Kannan, M. Ramanadham, T. A. Jones, *Ann. N. Y. Acad. Sci.* **1984**, *429*, 49-60.
39. A. Bujacz, J. A. Talaj, A. J. Pietrzyk, DOI: 10.2210/pdb4luf/pdb
40. S. J. Gamblin, G. J. Davies, J. M. Grimes, R. M. Jackson, J. A. Littlechild, H. C. Watson, *J. Mol. Biol.* **1991**, *219*, 573-576.
41. C. G. Cheong, S. H. Eom, C. Chang, D. H. Shin, H. K. Song, K. Min, J. H. Moon, K. K. Kim, K. Y. Hwang, S. W. Suh, *Proteins* **1995**, *21*, 105-117.
42. T. Granier, B. Langlois D'Estaintot, B. Gallois, J.-M. Chevalier, G. Precigoux, P. Santambrogio, P. Arosio, *J. Biol. Inorg. Chem.* **2003**, *8*, 105-111.
43. W. Humphrey, A. Dalke, K. Schulten, VMD – visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

Supporting Information

for manuscript entitled

Capturing and Stabilizing Folded Proteins in Lattices Formed with Branched Oligonucleotide Hybrids

by

Alexander Schwenger, Tomasz Jurkowski and Clemens Richert

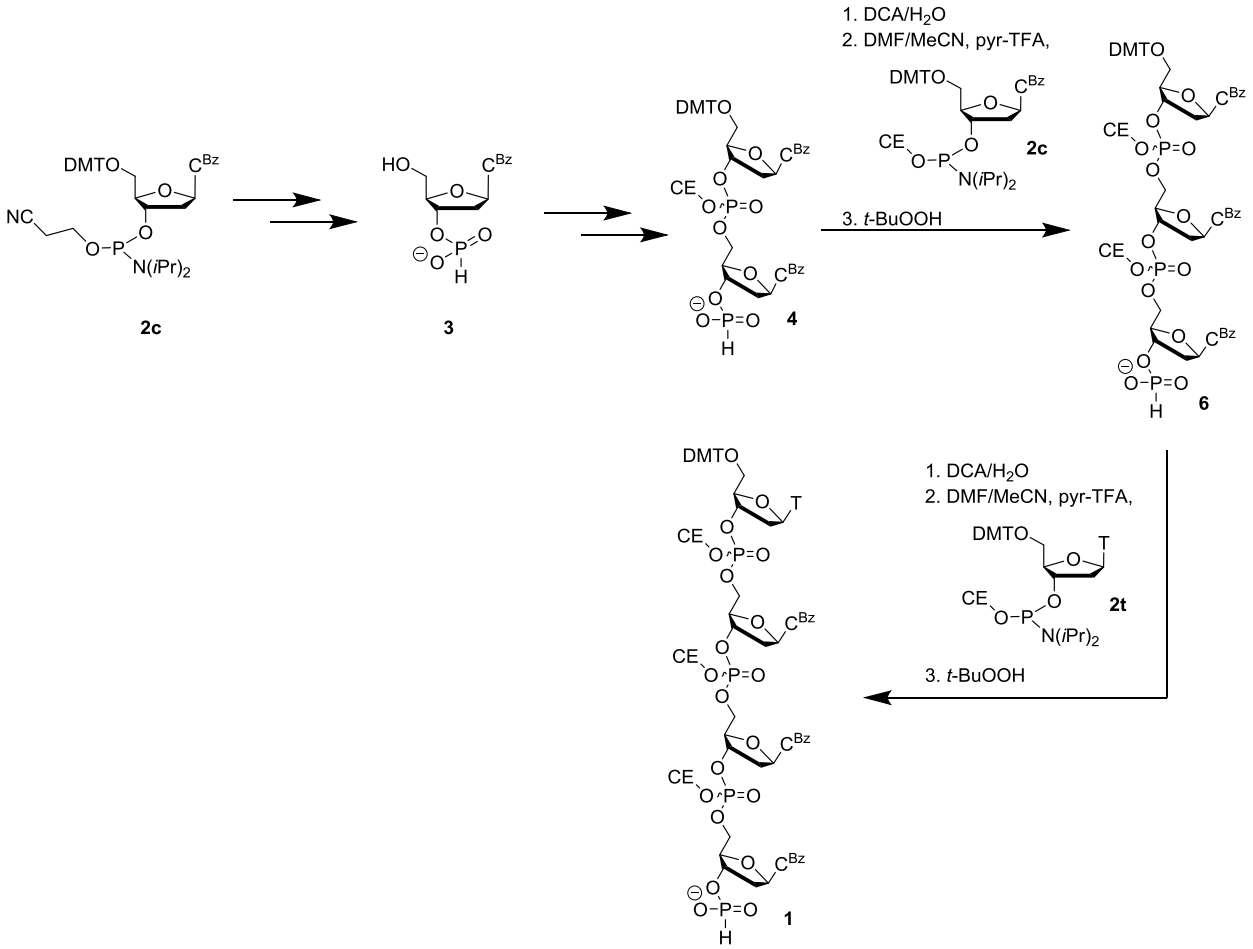
Table of Contents

1. General

Chemicals. Unless otherwise noted, reagents and solvents were obtained from commercial suppliers and used without further purification. Unmodified oligonucleotides were purchased from Biomers (Ulm, Germany), Eurofins (Ebersberg, Germany), or IDT (Leuven, Belgium) in HPLC-purified form, and their integrity was confirmed by MALDI-TOF mass spectrometry. ***Chromatography***. RP-HPLC chromatography was performed on a Nucleosil C_{18} column (5 µm, 250 x 4.6 mm) from Macherey-Nagel (Düren, Germany). Column chromatography was performed on silica gel 60 (0.040-0.063 mm particle size). Thin layer chromatography was performed on aluminum sheets, coated with silica gel 60 F_{254} and visualized with UV light (254 nm). ***Mass spectrometry***. MALDI-TOF mass spectra were measured in linear negative mode with a matrix/comatrix mixture of 2,4-6-trihydroxyacetophenone/diammonium citrate on a microFlex mass spectrometer (Bruker Daltonics). ***Other analytica methods***. UV-melting curve experiments were performed on a Lambda 25 spectrophotometer (Perkin Elmer), using the software TempLab 2.0. NMR spectra were recorded on Bruker AVANCE 300, 400 or 700 spectrometers.

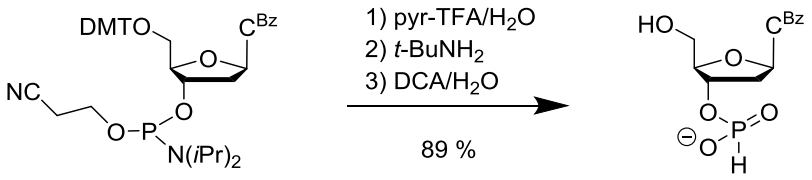
2. Synthetic route to 1

Overview over the synthesis of tetramer *H*-phosphonate **1**



Scheme S1. Synthesis of *H*-phosphonate tetramer **1**. CE = cyanoethyl.

2.1 Synthesis of N^4-benzoyl-2′-deoxycytidine-3′-yl *H*-phosphonate (3).[S1]



2c **3**

The following is a slight modification of a published procedure.[S1] A solution of phosphoramidite **2c** (5 g, 6 mmol) in CH_3CN (20 mL), H_2O (216 µL, 12 mmol, 2 equiv) and pyridinium trifluoroacetate (2.54 g, 13.2 mmol, 1.1 equiv) was shaken for 10 min at room temperature. Subsequently, *tert*-butylamine (10 mL) was added, and the reaction mixture was stirred for 15 min. The mixture was then concentrated in vacuo. The residue was dissolved in CH_2Cl_2 (5 mL) and concentrated again to give a foam. This process was repeated twice. For the detritylation, the residue was dissolved in CH_2Cl_2 (24 mL), followed by the addition of H_2O (600 µL) and dichloroacetic acid in CH_2Cl_2 (35 mL, 6% w/w). After 10 min, the reaction was quenched with CH_3OH (15 mL) and the solution was concentrated *in vacuo* to a small volume (10 mL). The crude product was then precipitated by addition of *tert*-butyl methyl ether (40 mL) followed by centrifugation, and the supernatant was aspirated. The solid was redissolved in CH_2Cl_2/CH_3OH (2/1, *v*/*v*, 5 mL), and precipitated again by the addition of *tert*-butyl methyl ether (45 mL). After separation by, this process was repeated twice. The remaining solid was dissolved in H_2O (20 mL) and washed with ethyl acetate (30 mL). The organic phase was back extracted with water (2 × 10 mL) and the combined aqueous phases were evapored *via* lyophilization. The *H*-phosphonate **3** was thus obtained in a yield of 2.11 g (4.5 mmol, 75%) as an off-white solid. The analytical data were in agreement with the literature.[S1] ^{31}P NMR (121.5 MHz, DMSO-d_6): δ = 0.13 ppm.

2.2 Synthesis of *H*-phosphonate dimer 4.

The *H*-phosphonate **3** (2.1 g, 4.5 mmol, 1.0 equiv.) was dried together with **2c** (4.5 g, 5.4 mmol, 1.2 equiv) and molecular sieves 3 Å (5-7 beads) for 2 h at 45 °C and 0.001 mbar. The flask was flushed with argon, sealed with a septum, and dry CH_3CN (15 mL) was added, followed by pyridinium trifluoroacetate solution (1 M in CH_3CN, 6.5 mL, 1.2 equiv), previously dried over molecular sieves, 3 Å, 10 beads). The mixture was placed in an ultrasonic bath for 1 min to obtain a clear solution of the starting materials. The reaction mixture was then shaken for 45 min at 25 °C, and then cooled to 0 °C. After the addition of *tert*-butyl hydroperoxide (2.7 mL, 5.5 M in decane, 15.1 mmol, 2.8 equiv.), the reaction mixture was allowed to reach room temperature within 30 min. Subsequently, a mixture of aqueous phosphate buffer (0.2 M, pH 7) and brine (60 mL, 1:1, *v/v*) was added to the mixture, and the solution was extracted with CH_2Cl_2 (80 mL). The aqueous phase was separated, and back-extracted with CH_2Cl_2 three times (3 × 30 mL). The combined organic phases were dried over molecular sieves (3 Å, 20 beads), filtered and concentrated under reduced pressure. The residue was then dissolved in dioxane (20 mL) and separated from the remaining solid, which was washed twice with dioxane (2 x 5 mL). The combined organic phases were concentrated under reduced pressure to approx. 10 mL, and split into 5 mL portions, followed by precipitation by adding MTBE (45 mL) and centrifugation (3500 rpm, 5 min). The supernatant solution was aspirated , and this procedure was repeated four times. The resulting solid was dissolved in dioxane (10 mL), frozen with liquid N_2, and then reduced to dryness under vacuum (<0.001 mbar). Compound **4** was obtained as a colorless solid in a yield of 5 g (4.4 mmol, 97%).

^{31}P NMR (121.5 MHz, DMSO-d_6): δ = 0.17, -2.50, -2.66 ppm (mixture of diastereomers); ESI-TOF *m/z* calcd for $C_{56}H_{55}N_7O_{16}P_2$ $[M-H]^-$ 1142.3, obsd 1142.3.

2.3 General Protocol A (Detritylation of *H*-phosphonate-terminated strands).

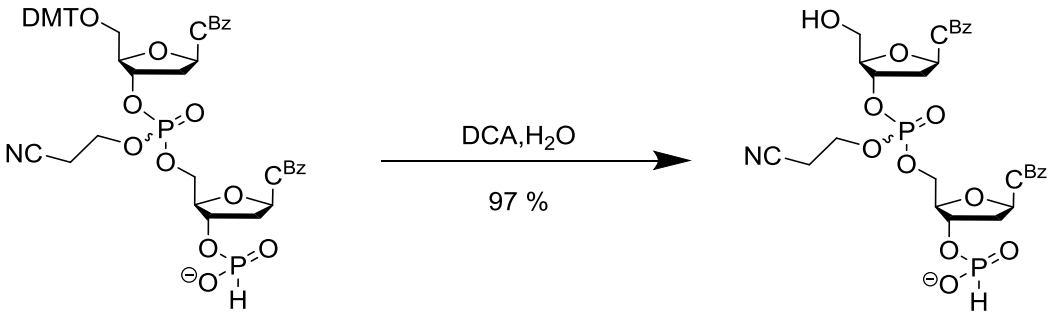
For the detritylation of **4** or **6** (0.34 mmol), the starting material was dissolved in CH_2Cl_2 (4 or 6 mL, respectively), followed by the addition of H_2O (100 µL) and a solution of dichloroacetic acid in CH_2Cl_2 (6 mL, 6% w/w). The reaction mixture was shaken for 10-15 min at 25 °C. After 10-15 min, the reaction was quenched with MeOH (4 mL), and the mixture was concentrated *in vacuo* to a volume of approx. 5 mL. The product was then precipitated by addition of *tert*-butyl methyl ether (40 mL), followed by centrifugation (rpm/min), and the supernatant was aspirated. The solid was redissolved in CH_2Cl_2/CH_3OH (1/1, *v/v*, 2 mL), and precipitated again by the addition of *tert*-butyl methyl ether (45 mL). This purification procedure was repeated one more time, and the resulting solid was suspended in dioxane (6 mL) and then frozen with liquid N_2 and dried under high vacuum (<0.001 mbar).

2.4 General Protocol B (Synthesis of trimer or tetramer *H*-phosphonates).

The detritylated *H*-phosphonate building block (**5** or **7**, 2.1 g, 1.62 mmol) was combined with **2c** or **2t** (2.1 mmol, 1.3 equiv.) and dried together with molecular sieves 3 Å (10 beads) for 2 h at 45 °C and 0.001 mbar. The flask was flushed with argon, sealed, and DMF (2.3 mL) was added, followed by pyridinium trifluoroacetate solution (1 M in CH_3CN, 2.74 mL, 1.3 equiv, previously dried over molecular sieves). The mixture was exposed to an ultrasonic bath for 2 min, resulting in a clear solution. The reaction mixture was then shaken for 50 min at 25 °C and then cooled to 0 °C, followed by addition of *tert*-butyl hydroperoxide (1.07 mL, 5.5 M in decane, 5.9 mmol, 2.8 equiv). The reaction mixture was shaken for 30 min at 25 °C. A mixture of aqueous phosphate buffer (0.2 M, pH 7) and brine (60 mL, 1:1, *v/v*) was added to the reaction mixture, followed by extraction with CH_2Cl_2 (60 mL). The phases were separated after centrifugation (2500 rpm, 5 min.). The aqueous phase was back-extracted with CH_2Cl_2 (3 × 30 mL). The combined organic phases were dried over molecular sieve 3 Å

(20 beads), filtered and concentrated under reduced pressure. The residue was dissolved in CH_2Cl_2 (20 mL), split into four 5 mL portions, followed by precipitation through addition of *tert*-butyl methyl ether (45 mL). After centrifugation (3500 rpm, 5 min), the precipitate was harvested, and the precipitation was repeated four times. The resulting solid was washed by treating with MeOH (10 mL) for 5 min in an ultrasonic bath. The supernatant was aspirated after centrifugation (3500 rpm, 5 min). The solid was dissolved in dioxane (5-10 mL) and treated in an ultrasonic bath for 10 min. The supernatant was then separated from the remaining solid which was washed twice with dioxane (10 mL). The supernatant was divided into 5 mL portions and the compound precipitated through addition of *tert*-butyl methyl ether (45 mL). This purification step was repeated three times. The resulting solid was dissolved in dioxane (10 mL) and frozen with liquid N_2, followed by drying at <0.001 mbar.

2.4 Synthesis of 5'deprotected dimer 5.

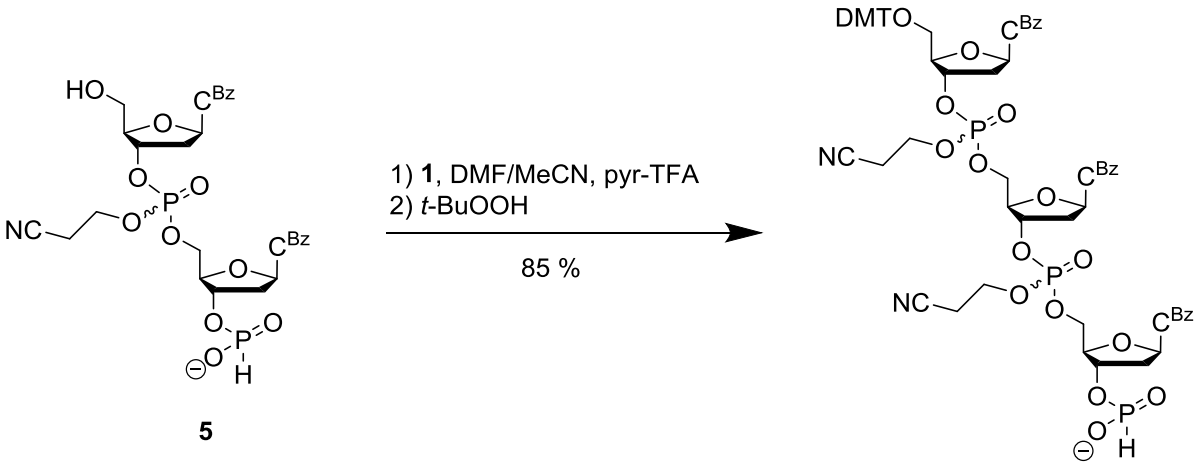


4 **5**

Compound **5** was prepared following General Protocol A, starting from **4** (5 g, 4.37 mmol), and was isolated in a yield of 3.65 g (4.35 mmol, 99%).

^{31}P NMR (121.5 MHz, DMSO-d_6): δ = 0.59, 0.46, -2.73, -2.95 ppm (mixture of diastereomers); (ESI-MS) *m/z* calcd for $C_{35}H_{37}N_7O_{14}P_2$ $[M-H]^-$ 840.18, obsd 840.18.

2.5 Synthesis of DMT-protected trimer 6.

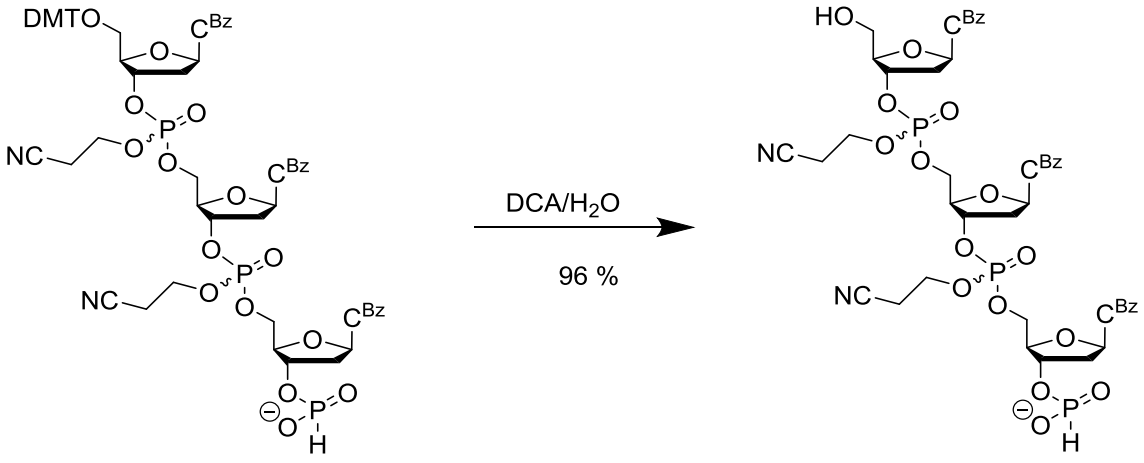


5 **6**

Following the general protocol B, compound **6** was prepared starting from **5** (0.54 g, 0.64 mmol) and was obtained as a colorless solid in a yield of 0.87 g (0.55 mmol, 85%).

^{31}P NMR (121.5 MHz, DMSO-d_6): δ = 0.17, -2.18, -2.51, -2.53, -2.65, -2.70 ppm (mixture of diastereomers); (ESI-MS) *m/z* calcd for $C_{75}H_{74}N_{11}O_{23}P_3$ $[M-H]^-$ 1588.40, obsd 1588.40.

2.6 Synthesis of trimer 7.

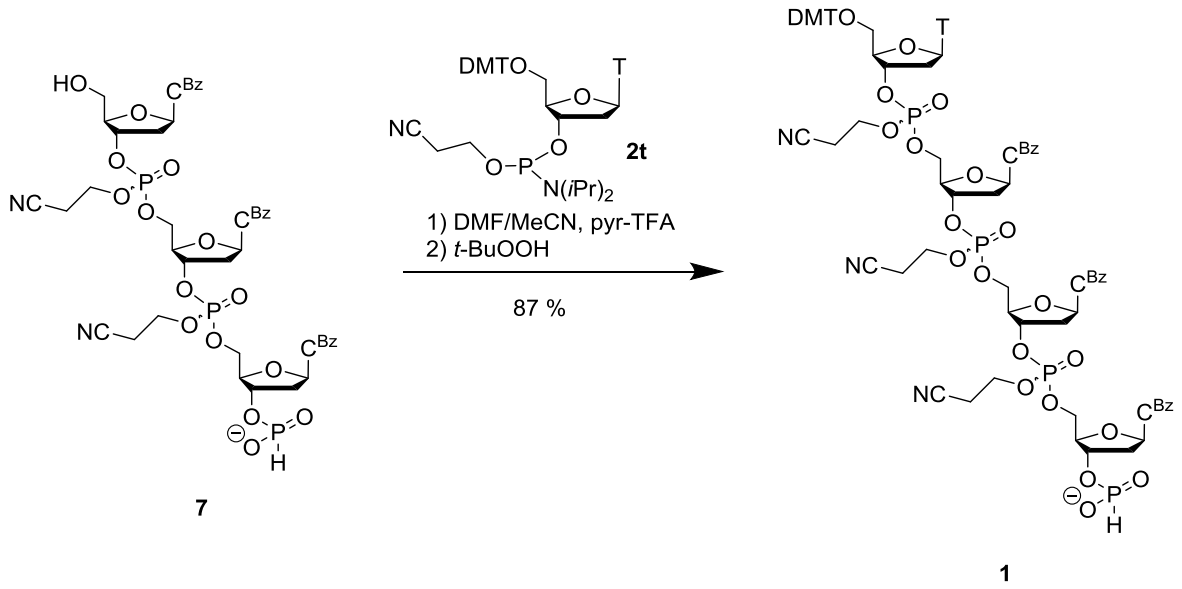


6 **7**

Compound **7** was prepared following the General Protocol A, starting from **6** (1.87 g, 1.17 mmol). The product was isolated in a yield of 1.45 g (1.13 mmol, 96%).

^{31}P NMR (121.5 MHz, DMSO-d_6): δ = 0.58, -2.20, -2.55, -2.72, -2.83 ppm (mixture of diastereomers); (ESI-MS) *m/z* calcd for $C_{54}H_{56}N_{11}O_{21}P_3$ $[M-H]^-$ 1286.27, obsd 1286.27.

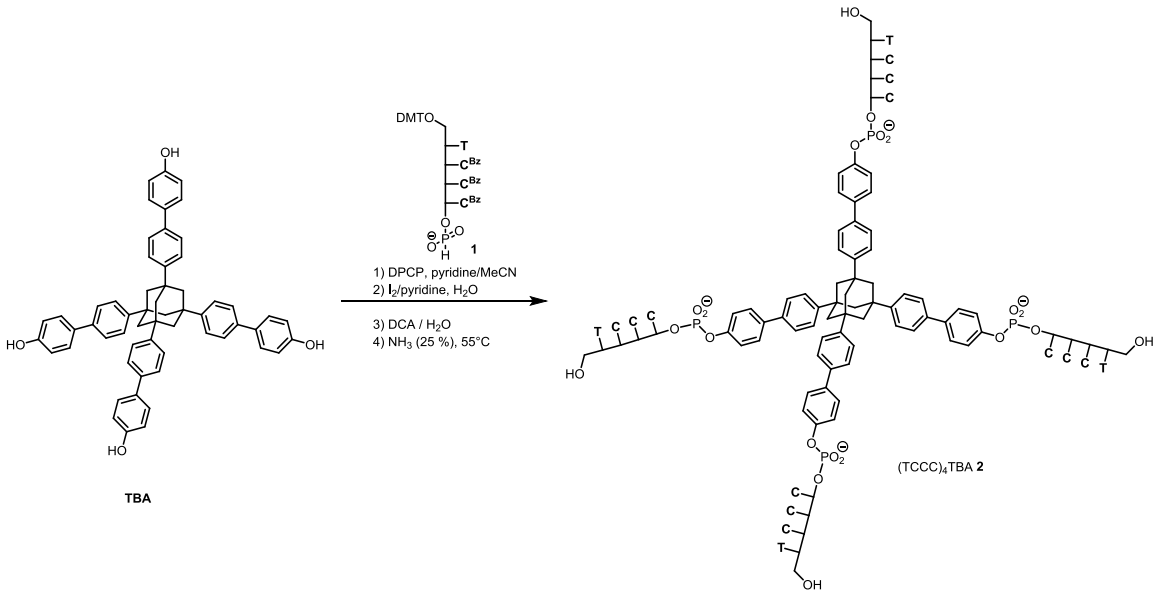
2.7 Synthesis of DMT-protected tetramer 1.



Following General Protocol B, compound **1** was prepared, starting from **7** (1.5 g, 1.17 mmol) and was obtained as an off-white solid in a yield of 1.99 g (1.02 mmol, 87%).

^{31}P NMR (121.5 MHz, DMSO-d_6): δ = 0.14, -2.03, -2.52, -2.60, -2.63, -2.70 ppm (mixture of diastereomers); (MALDI-TOF-MS) *m/z* calcd for $C_{88}H_{90}N_{14}O_{30}P_4$ $[M-H]^-$ 1945.4, obsd 1944.6.

2.8 Synthesis of $(TCCC)_4TBA$ (2).



Hybrid **2** was assembled from the core and the *H*-phosphonate tetramer using a modification of a known method.[S1,S2] Starting from a mixture of TBA (12.4 μmol, 1 equiv), *H*-phosphonate tetramer **1** (337 mg, 0.17 mmol, 14 equiv) and molecular sieves (3 Å, ten beads), which was dried for 2 h at 45 °C and 0.001 mbar, followed by flushing with argon. A solution of pyridine in CH_3CN (3.5 mL, 4:1, *v/v*) was added, and the resulting mixture was cooled to -40 °C. Diphenyl chlorophosphate (54 μL, 0.26 mmol) was added, and the mixture was stirred at -40 °C for 2 h. Then, a solution of iodine in pyridine (260 μL, 1 M) was added, followed by addition of water (16 μL, 0.87 mmol) after 1 min. The resulting mixture was stirred for 10 min at -40 °C and then for 30 min at room temperature. After addition of CH_2Cl_2 (10 mL), the solution was washed with aqueous sodium thiosulfate (10 mL, 10%, *w/w*), and then with phosphate buffer (9 mL, 0.2 M, pH 7). The aqueous phase was extracted with CH_2Cl_2 (3 × 10 mL). The combined organic layers were concentrated *in vacuo*, the residue was coevaporated three times with toluene, and then redissolved in CH_2Cl_2 (3 mL), followed by precipitation with MTBE (45 mL) and isolation by centrifugation. This precipitation was repeated twice, followed by dissolving in CH_2Cl_2 (5 mL). The supernatant was separated from remaining solid, which was washed with CH_2Cl_2 (2× 2 mL). The

combined organic phases were concentrated under reduced pressure, and the residue was washed with MeOH (5 mL). After the washing step, the crude was used for deprotection. For this, the protected hybrid was dissolved in a mixture of CH_2Cl_2 (2 mL) and H_2O (5 µL, 0.26 mmol), followed by addition of dichloroacetic acid (4 mL DCA, 6% v/v in CH_2Cl_2). After 15 min, methanol (1.5 mL) was added. The solution was then concentrated, and the detritylated hybrid was isolated by precipitation with MTBE (30 mL), followed by centrifugation, redissolving of the separated solid in CH_2Cl_2/CH_3OH (4 mL, 2/1, *v/v*) and precipitation with MTBE (30 mL). This process was repeated twice. The remaining protecting groups were removed by treating the resulting solid with ammonium hydroxide (25%, 12 mL) for 5 h at 55 °C. Excess ammonia was removed by passing a stream of N2 over the surface until the sample was odourless. Evaporation of the remaining solution yielded crude **2** (140 mg). Cartridge purification on a Sep-Pak C18 cartridge with a gradient of CH_3CN (0-25%) in 10 mM NH_4Ac buffer led to elution of **2** at 12-16% CH_3CN. After lyophilization, the title compound was obtained in a yield of 54 mg (9.7 µmol, 78%). A small sample was purified by HPLC, using a gradient of CH_3CN (5-30% in 45 min) in 10 mM TEAA buffer, with elution of **2** at t_R = 27.5 min.

MALDI-TOF-MS m/z calcd for $C_{206}H_{244}N_{44}O_{104}P_{16}$ $[M-H]^-$ 5495, obsd 5496.

3. Additional HPLC chromatogram

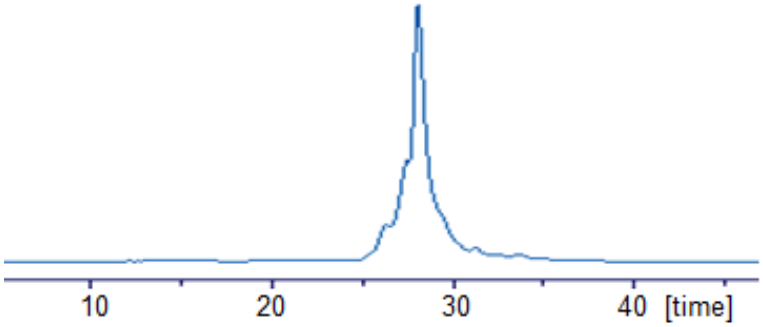


Figure S1. HPLC trace of $(TCCC)_4TBA$ (**2**), as obtained after cartridge purification. Conditions: gradient of CH_3CN (5-30%) in 45 min; **2** was detected at t_R = 27.5 min.

4. Additional MALDI-TOF mass spectra

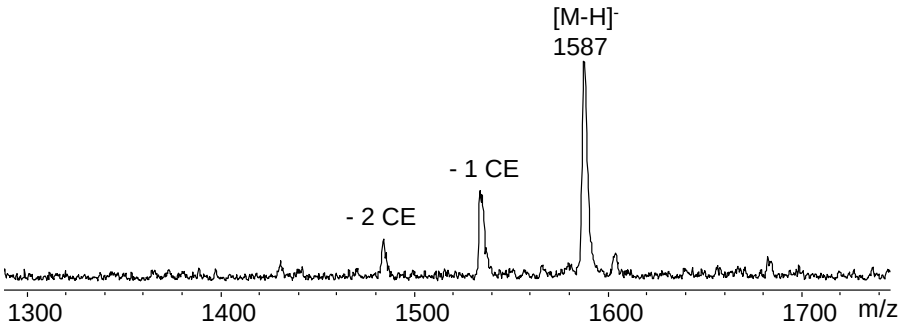


Figure S2. MALDI-TOF mass spectrum (linear negative mode, TC matrix) of fully protected intermediate **6**, as used for the subsequent transformation; -CE denotes loss of cyanoethyl groups, which is believed to occur during sample preparation.

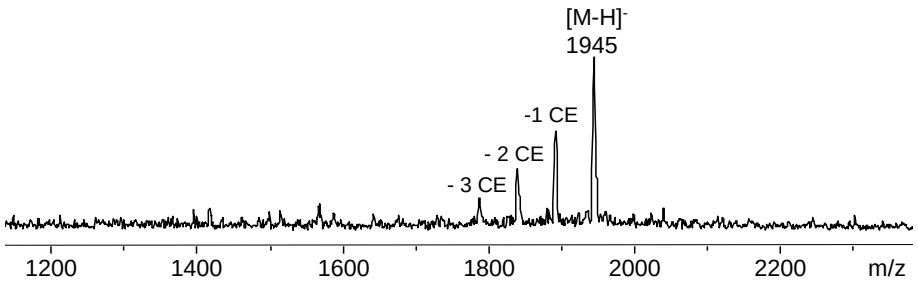


Figure S3. MALDI-TOF mass spectrum of fully protected tetramer strand **1**; -CE denotes loss of cyanoethyl groups (lin. neg., TC matrix). See legend of Figure S2 for other details.

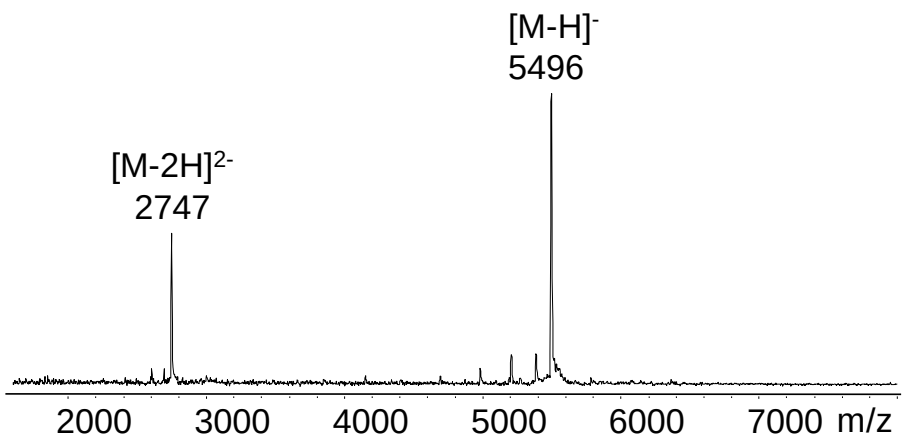


Figure S4. MALDI-TOF mass spectrum of hybrid $(TCCC)_4TBA$ (**2**).

5. Additional data from assembly studies

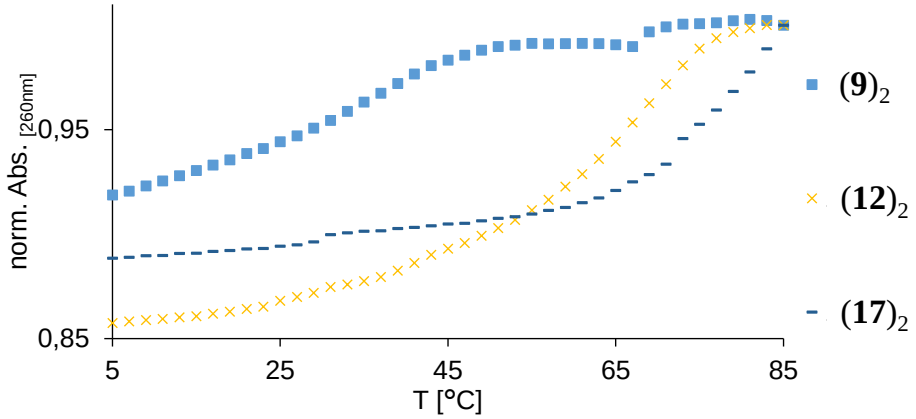


Figure S5. UV-melting curve, obtained by cooling the solution of strands forming connecting duplexes (compare Table 1 in the main manuscript).

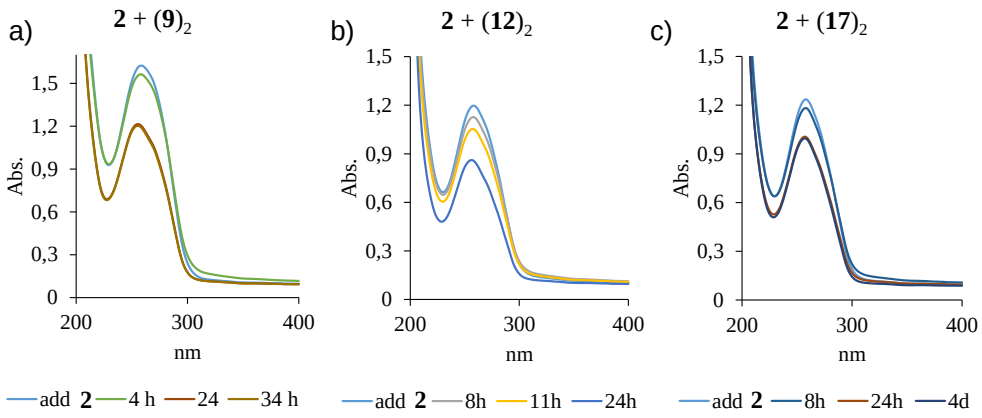


Figure S6. Decrease of UV-absorption recorded after different time points, subsequent to the addition of **2**.

6 UV-Melting curves

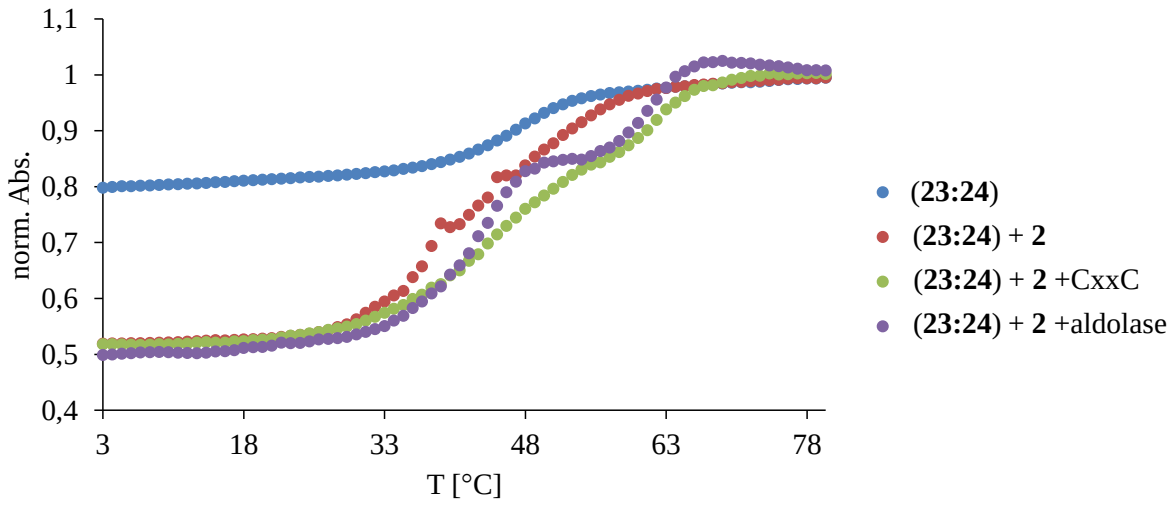


Figure S7. UV-Melting curves of assemblies as obtained after 48 h assembly time, measured at a heating rate of 1°C/min; (blue: duplex (**23:24**); red: of (**23:24**) and **2**; green: duplex/hybrid plus CxxC; and purple: duplex-hybrid network plus aldolase.

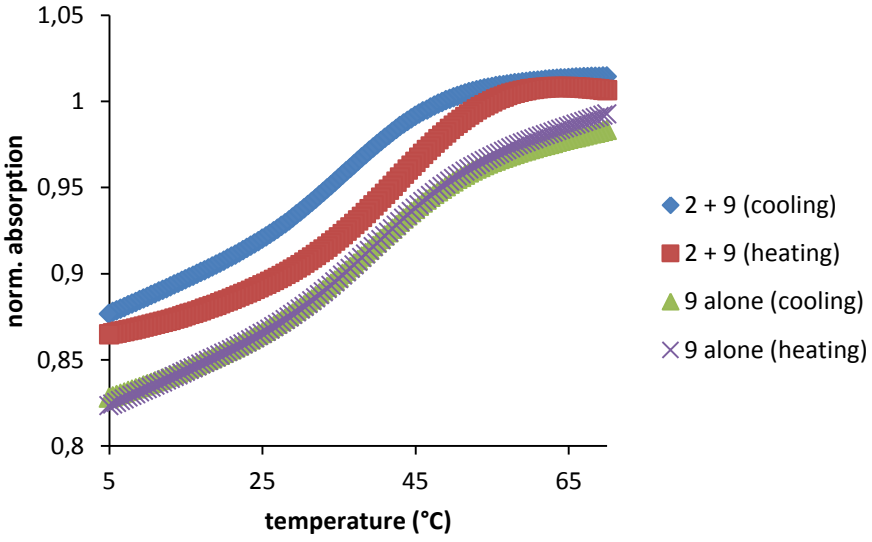


Figure S8. The assembly into three-dimensional networks manifests itself in the hysteresis of heating and cooling curves. Overlay of 90 point-smoothed UV-melting curves of connector duplex $(\mathbf{9})_2$ alone and the assembly of this duplex and hybrid **2**, as measured in 10 mM triethylammonium acetate buffer, 150 mM NaCl, and 100 mM $MgCl_2$ at 2 µM hybrid concentration and the required four equivalents (8 µM) of the strand forming the connecting duplexes. Note that the assemblies are slower form and slower to disassemble than the connecting duplex alone.

7. Data on encapsulation of DNA-binding proteins

Duplexes formed from sequences **31**-**47** (Figure 3 of the manuscript) were employed to generate connecting duplexes. Table S1, below, lists results obtained with or without either of the two small DNA-binding proteins and Figure S8 shows gel analyses of the fractions (solid/liquid).

Table S1. Formation of solids upon assembly of hybrid **2**, connecting duplexes of different length, and DNA-binding proteins CxxC or Mbd2 into three-dimensional lattices.

Entry No	duplex	duplex [μM]	hybrid [μM]	CxxC [μM]	Mbd2 [μM]	observation	encapsulation (%)
1	**23**:**24**	40	20	-	1	solid formation	≥ 90
2	**23**:**24**	40	20	1	-		≥ 70
3	**29**:**30**	40	20	-	1		≥ 90
4	**29**:**30**	40	20	1	-		≥ 70
5	**27**:**28**	40	20	-	1		≥ 90
6	**27**:**28**	40	20	1	-		≥ 70
7	**31**:**32**	40	20	-	1		≥ 90
8	**31**:**32**	40	20	1	-	partially gel-like, particle-like or solid-like	≥ 70
9	**33**:**34**	40	20	-	1		≥ 90
10	**35**:**36**	40	20	-	1		≥ 90
11	**23**:**24**	40	20	0.25			n. d.
12	**31**:**32**	40	20	-	0.25		quant.
13	**37**:**38**	40	20	0.5	-		--
14	**39**:**40**	40	20	0.5	-	no formation of solids	--
15	**41**:**42**	40	20	0.5	-		
16	(**43**)$_2$	50	25	-	1		
17	(**44**)$_2$	50	25	-	1		
18	(**45**)$_2$	50	25	-	1	solid formation	90-80
19	(**46**)$_2$	50	25	-	1		
20	(**47**)$_2$	50	25	-	1		

Conditions: 20 mM HEPES, 150 mM NaCl and 10 mM $MgCl_2$

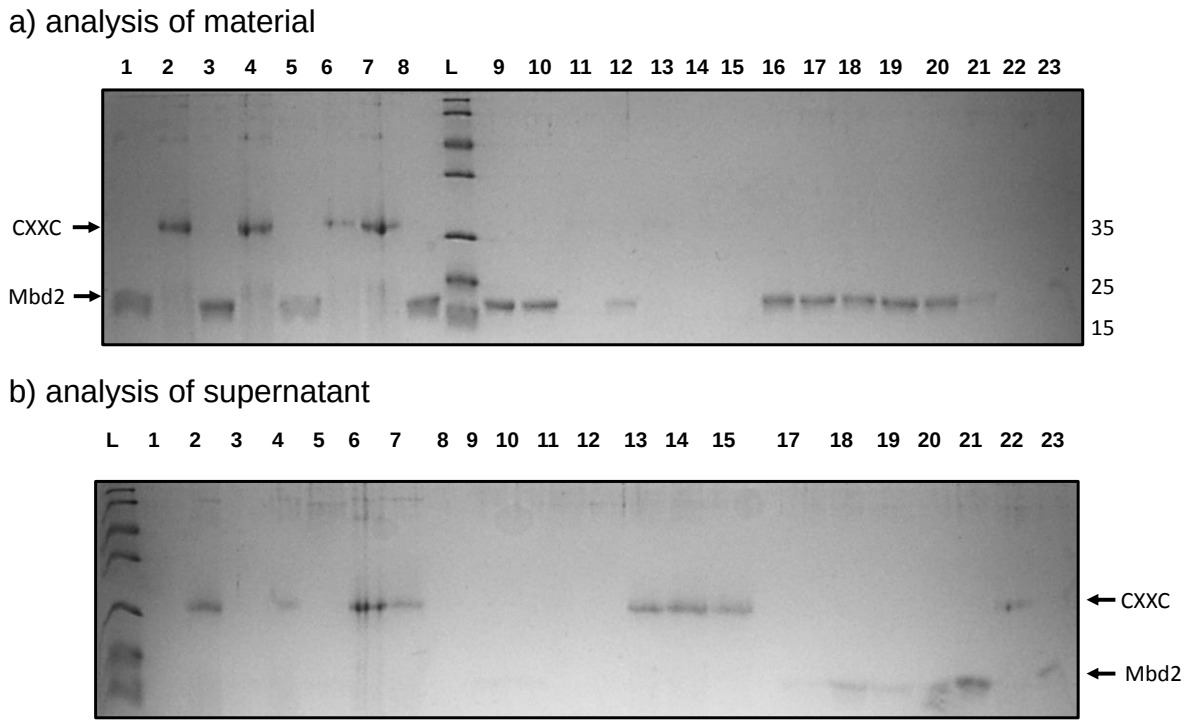


Figure S9. SDS PAGE gels of the solids formed in the assays of Table S 1 and the corresponding supernatant solutions. The numbers are according to the entries of Table S1, L = calibration ladder.

8. References for Supporting Information

S1. A. Singh, M. Tolev, C. I. Schilling, S. Bräse, H. Griesser, C. Richert, *J. Org. Chem.* **2012**, *77*, 2718-2728.

S2. A. Schwenger, C. Gerlach, H. Griesser, C. Richert, *J. Org. Chem.* **2014**, *79*, 11558-11566.