

ARTICLE

Received 9 Jul 2015 | Accepted 15 Jun 2016 | Published 28 Jul 2016

DOI: 10.1038/ncomms12294

OPEN

Peptide–oligonucleotide conjugates as nanoscale building blocks for assembly of an artificial three-helix protein mimic

Chenguang Lou¹, Manuel C. Martos-Maldonado², Charlotte S. Madsen², Rasmus P. Thomsen³, Søren Roi Midtgaard⁴, Niels Johan Christensen², Jørgen Kjems³, Peter W. Thulstrup⁵, Jesper Wengel¹ & Knud J. Jensen²

Peptide-based structures can be designed to yield artificial proteins with specific folding patterns and functions. Template-based assembly of peptide units is one design option, but the use of two orthogonal self-assembly principles, oligonucleotide triple helix and a coiled coil protein domain formation have never been realized for *de novo* protein design. Here, we show the applicability of peptide–oligonucleotide conjugates for self-assembly of higher-ordered protein-like structures. The resulting nano-assemblies were characterized by ultra-violet-melting, gel electrophoresis, circular dichroism (CD) spectroscopy, small-angle X-ray scattering and transmission electron microscopy. These studies revealed the formation of the desired triple helix and coiled coil domains at low concentrations, while a dimer of trimers was dominating at high concentration. CD spectroscopy showed an extraordinarily high degree of α -helicity for the peptide moieties in the assemblies. The results validate the use of orthogonal self-assembly principles as a paradigm for *de novo* protein design.

¹ Department of Physics, Chemistry and Pharmacy, Biomolecular Nanoscale Engineering Center, University of Southern Denmark, Campusvej 55, Odense M 5230, Denmark. ² Department of Chemistry, Biomolecular Nanoscale Engineering Center, University of Copenhagen, Thorvaldsensvej 40, Frederiksberg 1871, Denmark. ³ Biomolecular Nanoscale Engineering Center and Interdisciplinary Nanoscience Center (iNANO), University of Aarhus, Gustav Wieds Vej 14, Aarhus C 8000, Denmark. ⁴ Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, Copenhagen Ø 2100, Denmark. ⁵ Department of Chemistry, University of Copenhagen, Universitetsparken 5, Copenhagen Ø 2100, Denmark. Correspondence and requests for materials should be addressed to J.W. (email: jwe@sdu.dk) or to K.J.J. (email: kjj@chem.ku.dk).

Designed peptide-based structures have been shown to yield artificial proteins¹ and even in a few cases nanoscale objects^{2–4}. Simultaneously, oligonucleotides (ONs) have been intensively used for nanotechnology^{5,6}, including so-called DNA origami⁷, to create structurally advanced objects such as DNA-based nanoboxes⁸. In contrast, formation of nanoscale objects and artificial proteins from peptide-ON conjugates (POCs) as building blocks with combined involvement of two separate self-assembly principles have, to our knowledge, not yet been realized. Protein *de novo* design involves rational design of peptide or protein molecules to fold into a target protein or protein-like structure, rather than the use or re-design of a naturally occurring sequence. Protein *de novo* design offers a test of our understanding of the factors controlling protein structure, folding and stability. The approach also offers the prospect of access to tailor-made proteins^{9–17}. One way to overcome the complexity of protein folding is the concept of template-assembled synthetic proteins (TASPs)^{18,19}. Several groups have explored this approach with a diverse set of templates^{20–25}, and we have reported the first low-resolution structure of a TASP in the form of a carboprotein, using small-angle X-ray scattering (SAXS)²⁶.

Well-defined DNA secondary structures such as double helices²⁷, triple helices²⁸, multi-way junctions²⁹ and quadruplexes³⁰ are potential scaffolds in designs of TASPs, requiring POCs as building blocks. Different methods have previously been applied to couple peptides or proteins with ONs³¹, including azide-alkyne cycloaddition reactions due to their remarkable compatibility with diverse functional groups and their high second-order rate constant under mild conditions^{32,33}.

However, to the best of our knowledge, only three examples have been reported on peptide and protein assembly driven by the formation of DNA secondary structures^{34–36}. In one report it was demonstrated that quadruplex formation could orient two very short peptides to form an adjacent two-loop protein-like surface³⁴, in another that three- and four-way DNA junctions positioned one recombinant eADF4(C16) element at each terminal³⁵, and in a more recent report that two model substrates (a maltose-binding protein and an antibody fragment against urokinase plasminogen activator receptor) positioned on DNA junctions imitated the geometry of an antibody³⁶. Previously, POCs have thus not been explored in *de novo* protein design using two orthogonal self-assembly principles.

For the present proof-of-concept study we relied on a sequence derived from CoilV_aL_d^{37,38} which was N-terminally extended with an azidohexanoyl-Tyr linker. The α -helical coiled coil is a ubiquitous protein motif that exists in 5–10% of all protein sequences³⁹. One of the main characteristics of a coiled coil is the simplicity of its sequence, as it consists of a motif that repeats itself every seven residues, (abcdefg)_n. In coiled coils, two or more helices wrap around each other in a left- or right-handed helical twist conformation. These helices can adopt different topologies, as they can assemble in parallel or antiparallel orientation. The solution structure analysis of CoilV_aL_d also revealed a cooperative monomer-dimer-trimer equilibrium, with the dimer state being an intermediate³⁷. Later the crystal structure of CoilV_aL_d revealed a parallel triple-helical structure³⁸.

Herein, we report an efficient and high-yielding preparation of POCs by copper-free alkyne-azide cycloaddition reactions, and we show that ON triple helix-formation can be used to

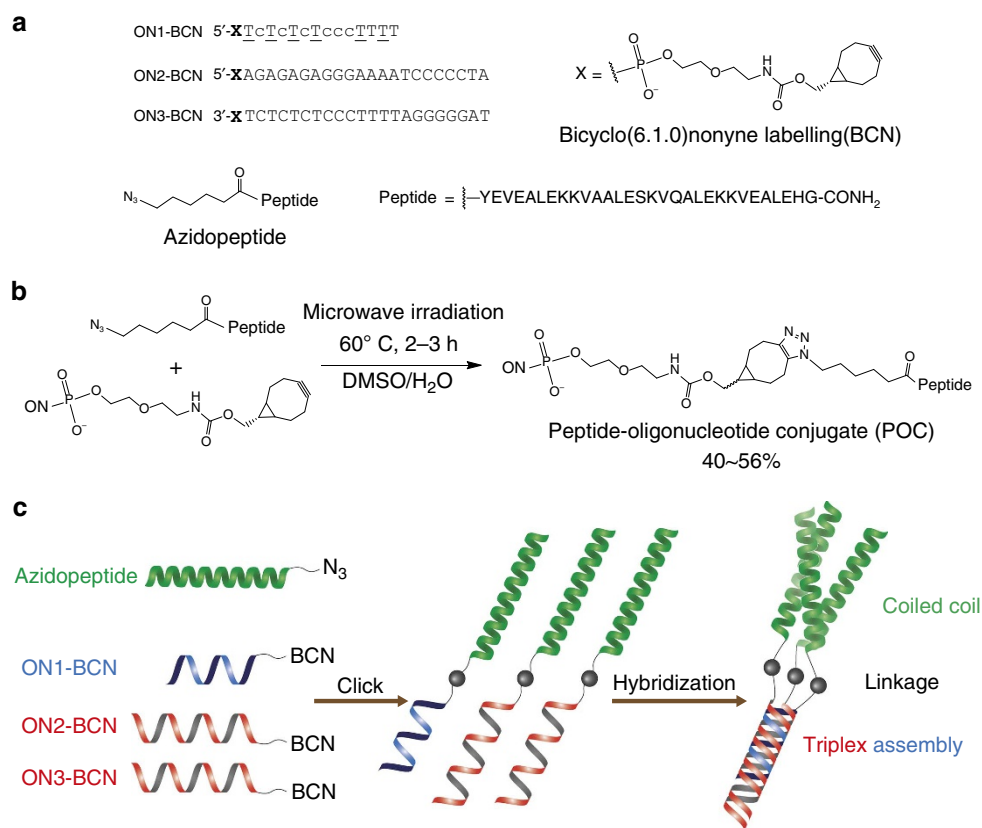


Figure 1 | Design and assembly of the desired three-helix protein mimic. (a) Sequence of the azidopeptide and ON1-BCN, ON2-BCN and ON3-BCN: T = LNA-T and c = 5-methyl-2'-deoxycytidine. A, C, G and T are DNA monomers. (b) Illustration of ring-strain promoted azide-alkyne cycloaddition chemistry between the azidopeptide and a BCN-containing ON. (c) Schematic illustration of the synthesis of POCs through azide-alkyne chemistry and formation of a triple-helical protein mimic.

organize peptide strands leading to the formation of a highly stable three-helix bundle protein mimic that dimerizes at higher concentrations. Locked nucleic acid (LNA) was central to our design⁴⁰.

Results

Synthesis of POCs. A convenient and efficient method using copper-free ring-strain promoted azide-alkyne coupling reactions was developed for synthesis of POCs (Fig. 1a). Among strained cyclooctynes^{41,42}, bicyclo[6.1.0]nonyne (BCN) was chosen due to its high second-order rate constant and commercial availability as a phosphoramidite monomer ready for use on an automated DNA synthesizer (Fig. 1b)^{43,44}. In this way the ON building blocks ON1-BCN, ON2-BCN and ON3-BCN were obtained. The azido-labelled peptide was synthesized by automated Fmoc chemistry and N-terminal attachment of a 6-azido hexanoyl group. Efficient and high-yielding preparation of POCs containing long peptide fragments (≥ 30 -mers) or proteins is challenging and only very few successful examples have been published via chemical synthesis^{35,36,45}, but this method furnished the desired conjugates POC1, POC2 and POC3 in satisfactory yields for milligram-scale preparation (further details are included in Supplementary Discussion, Supplementary Methods and Supplementary Figs 1–6).

Ultraviolet-melting studies and gel electrophoresis analysis.

As a first step in studying the applicability of nucleic acid self-assembly towards organizing peptide bundles (Fig. 1c), the stability of duplexes and triple helices involving one or more POC(s) was evaluated (see also details in Supplementary Methods), with their ON-based counterparts as references (Table 1a). One aspect was to study if the helical structures were stabilized by a synergistic effect involving both the ON and the peptide moieties. A wavelength of 275 nm was used for ultraviolet thermal denaturation ('melting') studies where the ultraviolet absorption difference between the triple helix and the underlying duplex is maximal (Supplementary Fig. 7). An ultraviolet-scan showed that the peptide unit did not display significant absorption at this wavelength (Supplementary Fig. 8). For interpretation of these studies it is assumed that any higher-order structures possibly formed (see sections below on gel electrophoretic, SAXS and TEM analysis) have no significant effect on annealing or denaturation of duplexes or triple helices.

In the case of duplex systems (Table 1b), the POC2 + POC3 assembly (Table 1b, entry 4) showed substantially higher thermal stability than the ON2 + ON3 duplex (Table 1b, entry 1), which we attribute to an additional interaction between the peptide moieties acting cooperatively with the Watson–Crick interactions of the ON units (Table 1b, entry 4). The melting curve of the POC2 + POC3 duplex is well superimposed on the corresponding annealing curve (Supplementary Fig. 9), indicating that the peptide modification does not perturb the thermal reversibility of the duplex structure. Importantly, no stabilization was seen for assemblies containing only one peptide unit (Table 1b, entries 2 and 3), strongly supporting that interactions between the peptide units play an indispensable role in stabilizing the POC2 + POC3 structure.

Eight triple-helical complexes were formed using ON1 and POC1 as triple-helix-forming ON (TFO) designed to bind in the major groove of the four duplexes (Table 1b, entries 5–12). A significant increase in melting temperature was observed for the POC1 + POC2 + POC3 triple helix (Table 1b, entry 6) when compared with the corresponding ON1 + ON2 + ON3 triple helix (Table 1b, entry 5). We ascribe this increased stability to the formation of a three-helix parallel peptide bundle, which for the non-conjugated peptides in solution was reported as the predominant structure^{37,46}. Accordingly, the annealing curve for

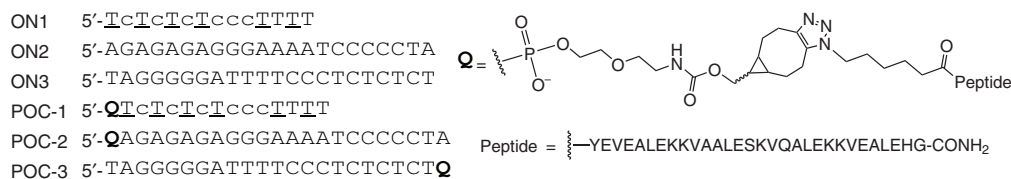
POC1 + POC2 + POC3 gave a clear triple-helix annealing transition, whereas that transition was not detected for the ON1 + ON2 + ON3 triple helix (Supplementary Fig. 10). As expected, there was no or only little improvement of triple-helix stability when ON1 was bound to the three duplexes containing one or two peptide units (Table 1b, entries 7–9, Supplementary Fig. 11). The remaining assemblies displayed in Table 1b (entries 10–12) involve POC1, where increased triple-helix stability was also observed when only one strand of the duplex contained a peptide moiety (Table 1b, entries 10 and 11, Supplementary Fig. 12). This observation again strongly indicates cooperative effects between nucleic acid hybridization and peptide assembly. This was further evidenced by the lack of triple-helix annealing transitions for the assemblies with only one peptide unit (Table 1b, entries 7, 8 and 12) or with the two peptide units on the two duplex-forming strands (Table 1b, entry 9).






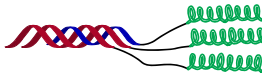






For POC1 + POC2 + ON3 and POC1 + ON2 + POC3 where the triple-helix transitions were overlapping with duplex transitions (Table 1b, entries 10 and 11), the experiments were repeated at pH 7.5. The stability of parallel nucleic acid triple helices is pH dependent, and in near physiological conditions the triple helices become less stable when the pH increases, mainly due to de-protonation of N3 of cytosine nucleobases⁴⁷. Under these conditions, both hybrid triple helices exhibited distinctive triple-helix melting and annealing transitions, whereas the transition temperatures for the underlying duplexes were unchanged (Supplementary Fig. 13).

Non-denaturing polyacrylamide gel electrophoresis (see Supplementary Methods for details) was used to study the assemblies discussed above (Table 1). Relatively low-voltage and low-temperature (4 °C) conditions minimized the possible dissociation of any triple helix during electrophoresis.

The reference ON2 + ON3 assembly (Fig. 2, lane 1) appeared as a single band, consistent with the length of the duplex (22-mers). The ON1 + ON2 + ON3 assembly (Fig. 2, lane 3) appeared as one dominant band (triple helix), and a faint band corresponding to the ON2 + ON3 duplex as a result of TFO dissociation (for room temperature conditions see Supplementary Fig. 14). When the pH was increased to 8.5 (Supplementary Figs 15 and 16) only the lower band was observed, consistent with the melting studies where no triple-helix transition was observed at pH 7.5 (Supplementary Fig. 13). The POC1 + POC2 + POC3 assembly (Fig. 2, lane 4) appeared as a dominating band with a far slower mobility than for the corresponding ON-based triple-helix control, confirming an increase in molecular size and mass-to-charge ratio due to the peptide moieties. In addition, a faint band moving close to the 150-mer DNA mark was observed which could be explained by dimerization (or multimerization) of the POC1 + POC2 + POC3 triple helix, possibly induced by additional intermolecular interactions. Similar behaviour was observed at three different conditions (pH 7.0, room temperature in Supplementary Fig. 14; pH 8.5, 4 °C in Supplementary Fig. 15; and pH 8.5, room temperature in Supplementary Fig. 16), demonstrating the existence of POC1 + POC2 + POC3 as a triple helix even at pH 8.5 at room temperature (Supplementary Fig. 16, lane 4) and the remarkable peptide-induced stabilization of the otherwise rather unstable ON triple helix. In contrary to the POC1 + POC2 + POC3 assembly, the POC2 + POC3 as well as other bimolecular assemblies appeared as several bands with a lower mobility than the band for the POC1 + POC2 + POC3 triple helix (Fig. 2, lanes 2, 5 and 6) probably reflecting the formation of various multimeric higher-order structures induced by a drive towards formation of three-stranded peptide domains.

Circular dichroism spectroscopy. Figure 3 shows the circular dichroism (CD) spectra of POC1 + POC2 + POC3, azido-peptide,

Table 1 | Ultraviolet-melting studies.**A****B**

Entry	Duplex/Triplex	Melting	Annealing	Structure
1	ON2+ON3	61.9 ± 0.5 °C	62.0 ± 0.5 °C	
2	POC2+ON3	61.7 ± 0.3 °C	61.9 ± 0.2 °C	
3	ON2+POC3	61.8 ± 0.3 °C	61.9 ± 0.4 °C	
4	POC2+POC3	70.0 ± 0.6 °C	70.0 ± 0.3 °C	
5	ON1+ON2+ON3	40.3 ± 0.3 °C (61.6 ± 0.1 °C)	n.d. (61.9 ± 0.2 °C)	
6	POC1+POC2+POC3	61.0 ± 0.6 °C (70.7 ± 0.1 °C)	59.8 ± 0.3 °C (70.3 ± 0.1 °C)	
7	ON1+POC2+ON3	41.7 ± 0.1 °C (61.4 ± 0.1 °C)	n.d. (61.6 ± 0.0 °C)	
8	ON1+ON2+POC3	42.3 ± 0.6 °C (61.7 ± 0.2 °C)	n.d. (61.9 ± 0.1 °C)	
9	ON1+POC2+POC3	42.3 ± 0.6 °C (70.3 ± 0.4 °C)	n.d. (69.8 ± 0.4 °C)	
10	POC1+POC2+ON3	p.o. (61.8 ± 0.2 °C)	p.o. (61.7 ± 0.2 °C)	
11	POC1+ON2+POC3	p.o. (60.6 ± 0.3 °C)	p.o. (61.9 ± 0.5 °C)	
12	POC1+ON2+ON3	41.2 ± 0.8 °C (61.7 ± 0.2 °C)	n.d. (61.8 ± 0.2 °C)	

ND, not detected; PO, partly overlapping.

A: POC and ON sequences: T = thymine LNA monomer, c = 5-methylcytosine DNA monomer; A, C, G and T are DNA monomers. **B:** Thermal denaturation and annealing temperatures (T_m and T_a values) of POC-including and ON-reference duplexes (entries 1–4) and triple helices (entries 5–12) at pH 7.0 measured as an average of three independent melting temperature determinations shown with the corresponding standard deviations. For entries 5–12, the values in brackets are T_m and T_a values measured for the corresponding underlying duplexes. The experiments were recorded at 275 nm in 5.8 mM NaH₂PO₄/Na₂HPO₄ buffer (pH 7.0) containing 100 mM NaCl and 0.10 mM EDTA. The concentration of the individual duplex components was 1.0 μM, while the TFO component was used in 1.5 μM concentration. The peptide moiety is marked in green, the TFO moiety in blue and the DNA duplex moiety in red.

ON1 + ON2 + ON3 and a sample containing azidopeptide and triple-helix ON1 + ON2 + ON3 at 20 °C (see also Supplementary Methods for details). As expected, the far-ultraviolet CD spectrum of azidopeptide exhibited an intense α -helix signal with negative peaks at 222 and 208 nm, respectively. ON1 + ON2 + ON3 showed a wide positive signal between 305 and 251 nm with a maximum at 284 nm and a shoulder around 260 nm, a wide negative signal between 251 and 206 nm with a shoulder around

240 nm and a maximum around 211 nm, which corresponds to the spectrum of a hybridized B-form DNA⁴⁸. Similarly, the spectral features in the high-wavelength region also indicated the formation of an ON triple helix (Supplementary Fig. 18). In Fig. 3, the higher wavelengths of the spectrum of POC1 + POC2 + POC3 were characterized by the contribution of the triple-helix ON signals with a wide positive feature between 305 and 251 nm with a maximum around 284 nm. At lower

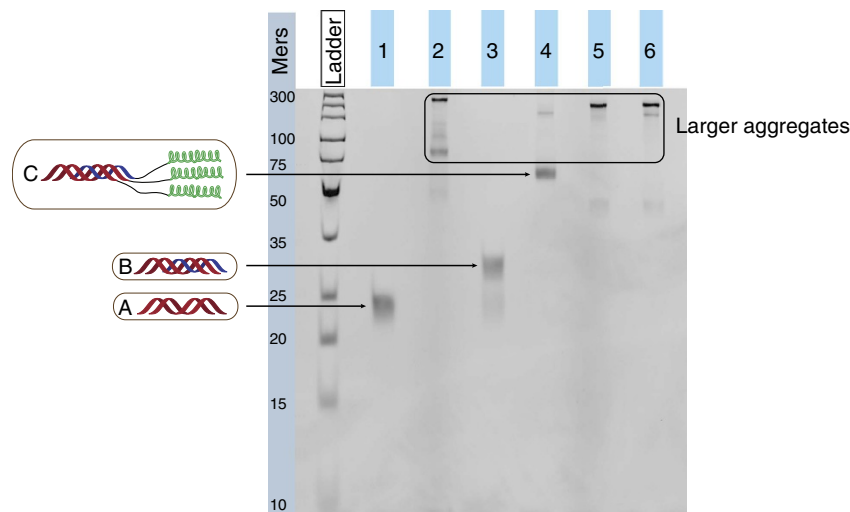


Figure 2 | Gel analysis. Non-denaturing 13% PAGE at pH 7.0 and 4 °C: lane 1, ON2 + ON3 (structure shown as 'A'); lane 2, POC2 + POC3; lane 3, ON1 + ON2 + ON3 (structure shown as 'B'); lane 4, POC1 + POC2 + POC3 (structure shown as 'C'); lane 5, ON2 + POC3; lane 6, POC2 + ON3. The gel was visualized by ultraviolet excitation at 260 nm after ethidium bromide staining. The O'GeneRuler Ultra Low Range DNA Ladder from bottom to top: 10, 15, 20, 25, 35, 50, 75, 100, 150, 200 and 300 mers.

wavelengths, an intense negative signal between 251 and 201 nm was attributed to the peptide α -helix element with only a minor contribution by the ON region.

The potential intermolecular interaction between non-conjugated peptides and ONs were tested in a control experiment, where it was found that the spectrum of azidopeptide and ON1 + ON2 + ON3 matches the mathematical sum of the spectra of its separated components (Supplementary Fig. 19). Thus, conformational changes induced by interactions between unconjugated peptides and ONs can be excluded.

Importantly, the spectrum of POC1 + POC2 + POC3 was significantly different from the sum of the azidopeptide and ON1 + ON2 + ON3 spectra, but only in the far-ultraviolet region where the peptide transitions dominate (Fig. 3). In the near-ultraviolet region the positive signals observed for the ON samples between 305 and 251 nm were practically identical to those of the POC1 + POC2 + POC3 sample. The latter, however, exhibited a significantly more intense α -helix signal, which indicated a higher degree of α -helicity. The $\theta_{222\text{ nm}}/\theta_{208\text{ nm}}$ ratios, assumed to be ≥ 1 for coiled coil peptides, were calculated^{49,50}. Azidopeptide and POC1 + POC2 + POC3 (after subtracting the spectrum of ON1 + ON2 + ON3 to remove the contribution of the ON part) showed $\theta_{222\text{ nm}}/\theta_{208\text{ nm}}$ ratios equal to 1.03 and 1.05, respectively. These values are consistent with the existence of stable α -helical coiled coils in solution and POC1 + POC2 + POC3 having a markedly increased helical content (see discussion below).

The concentration dependence of the helical content of POC1 + POC2 + POC3 and azidopeptide were tested by CD spectroscopy (Supplementary Discussion). Compared with the dilute samples shown in Fig. 3, a further change in the mean residue ellipticity was observed when the concentration was increased, both for azidopeptide and POC1 + POC2 + POC3. However, the latter exhibited a markedly increased CD signal magnitude in the concentration range tested. For the azidopeptide a plateau was reached at $\sim 15\ \mu\text{M}$ with mean residue ellipticity values at 222 nm around $-31,000\ \text{deg cm}^{-2}\ \text{dmol}^{-1}$, whereas further decrease to values below $-50,000\ \text{deg cm}^{-2}\ \text{dmol}^{-1}$ was observed for POC1 + POC2 + POC3 (Supplementary Fig. 20 and Supplementary Table 1). These observations indicate that the interactions between the ON moieties of POC1 + POC2 + POC3 induced a singularly higher degree of coiled coil in comparison with the unconjugated peptide. The observed CD signal is

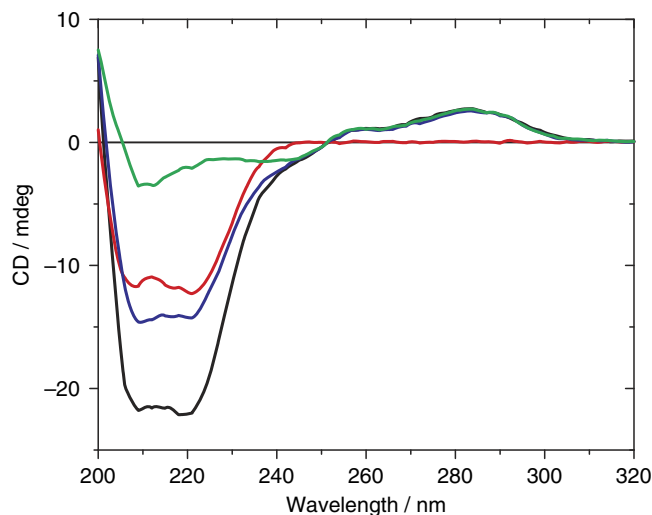


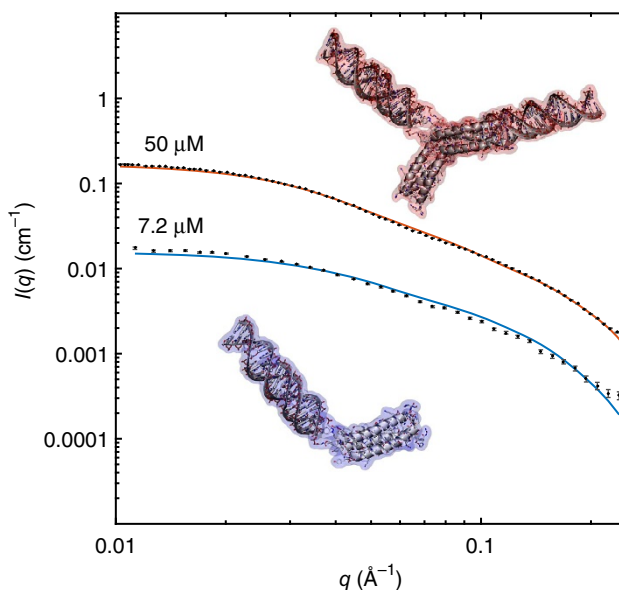
Figure 3 | CD analysis. Far-ultraviolet CD spectra of POC1 + POC2 + POC3 (3:2:2 μM ; black), azidopeptide (7 μM ; red), ON1 + ON2 + ON3 (3:2:2 μM ; green) and a control containing azidopeptide (7 μM) and ON1 + ON2 + ON3 (3:2:2 μM ; blue) at 20 °C in 5.8 mM phosphate buffer pH 7.0 with 100 mM NaCl and 0.10 mM EDTA. The spectra were recorded using a 0.2 cm path-length cell. The corresponding reference (buffer solution) was subtracted from each spectrum but no other manipulation was applied.

noteworthy as negative ellipticity values of such a magnitude to the best of our knowledge have never been reported in a purely aqueous solvent at ambient temperature. To evaluate the contribution of the ON region of POC1 + POC2 + POC3 to the ellipticity at 222 nm, we measured samples of unconjugated ON1 + ON2 + ON3 at the same concentrations. In correspondence with the spectra shown in Fig. 3, the signal was in every case $\sim 10\%$ of that of POC1 + POC2 + POC3 (Supplementary Table 1). CD temperature denaturation studies provided results in agreement with the ultraviolet-melting studies (Supplementary Discussion and Supplementary Figs 21–24).

Small-angle X-ray scattering. SAXS analysis was used to provide information on overall structure and oligomeric state of the

Table 2 | SAXS studies and computational modelling.

A Samples	Maximum distance (nm)*	Oligomeric state [†]	R _g (Å)	Experimental I(Q)	Calculated I(Q)
ON1 + ON2 + ON3	7.5	Trimer	22.11 ± 0.05	3.95 × 10 ⁻²	3.97 × 10 ⁻²
Azidopeptide	6.7	Trimer	19.93 ± 0.15	5.13 × 10 ⁻³	5.19 × 10 ⁻³
POC1 (TFO)	12.5	Dimer	36.59 ± 0.15	1.75 × 10 ⁻²	1.46 × 10 ⁻²
POC2	12.6	Trimer	38.49 ± 0.12	3.91 × 10 ⁻²	3.49 × 10 ⁻²
POC3	13.1	Trimer	41.68 ± 0.15	3.57 × 10 ⁻²	3.40 × 10 ⁻²
POC1 + POC2 + POC3 (7.2 μM)	12.7	Trimer	39.81 ± 0.23	1.81 × 10 ⁻²	2.13 × 10 ⁻²
POC1 + POC2 + POC3 (50 μM)	16.3	Dimer of trimers	48.01 ± 0.07	1.88 × 10 ⁻¹	1.81 × 10 ⁻¹

BR_g, radius-of-gyration.**A:** SAXS analysis of ON1 + ON2 + ON3 (50 μM each), azidopeptide (150 μM), individual POCs (50 μM) and POC1 + POC2 + POC3 (7.2 and 50 μM).**B:** Molecular models and experimental (black dots with error bars) and calculated (red and blue lines) SAXS data for POC1 + POC2 + POC3 measured at high (50 μM) and low (7.2 μM) concentration. The molecular models and their predicted scattering curves are coloured red (dimer of POC1 + POC2 + POC3 trimer) and blue (POC1 + POC2 + POC3 trimer). Alternative visualization of the dimer of POC1 + POC2 + POC3 trimer model is found in Supplementary Fig. 29.

*As determined from the indirect Fourier transformation (pair distance distribution).

†On the basis of the Guinier fit to the data. Experimental I(Q) and R_g were derived from this.

self-assembled complexes (Table 2a). Performing an indirect Fourier transformation gave the overall physical dimensions of the complexes in solution, while the total scattering provided the oligomeric state of the molecules (see details in Supplementary Methods and Supplementary Figs 25 and 26). Analysis of the data assisted by molecular modelling then provided more detailed solution structures (see below and Supplementary Figs 27–29). SAXS measurements typically require a relatively high concentration; in this study 50 μM of each ON/POC unit was used, while the azidopeptide was measured at 150 μM (3 × 50 μM). POC1 + POC2 + POC3 was measured at 50 μM, as well as at the lower concentrations of 3.6 and 7.2 μM.

The size of the self-assembled structures and the oligomeric state of ON1 + ON2 + ON3, azidopeptide and individual POCs corresponded well with the expected values. Both ON1 + ON2 + ON3 and the azidopeptide formed trimers, the latter with dimensions of the longest axis at ~7 nm and a cross section at ~2 nm, almost identical to the CoilV_aL_d and carbohydrate-templated peptide self-assembled structures^{26,51}. ON1 + ON2 + ON3 also formed a trimer with a longest axis of ~7.5 nm and a cross section of ~2 nm, corresponding with a triple-helical DNA structure.

Studying the POCs individually, it was clear that POC1 formed a homo-dimer, while POC2 and POC3 formed homo-trimers, all with overall lengths of 13 nm, corresponding well with a DNA and peptide moiety. POC1 contains a rather short and rigid triplex-forming pyrimidine oligonucleotide moiety with half of the sequence modified by LNA nucleotides, while POC2 and POC3 contain a substantial proportion of purine nucleotide but no LNA. Some unexpected interactions and secondary structures may be formed for POC1, but not POC2 and POC3, which inhibit trimer formation. Furthermore, a thinner rod was observed for POC1 (short and rigid TFO sequence), which may also contribute to the dimeric appearance of POC1, although the dimensions are shown to be approximately as for POC2/POC3. Furthermore, a small feature was observed for the POCs in the SAXS data around 0.1 Å⁻¹ (Supplementary Fig. 25), which arose from the formed two-domain two-contrast structure, again supporting the formation of the expected structure.

For the POC1 + POC2 + POC3 (1:1:1) combination, SAXS data were acquired at 3.6, 7.2 and 50 μM to study self-assembly at different concentrations (Supplementary Fig. 25). Measuring SAXS at low concentrations can be challenging, but was achieved at the beam line BM-29 at the European Synchrotron Radiation

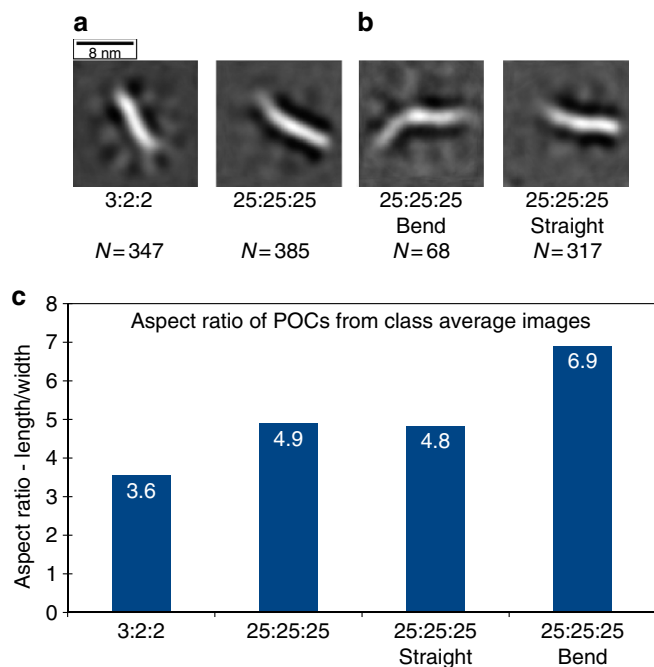


Figure 4 | TEM analysis. (a) TEM class average images of POC1+POC2+POC3 at low ‘3:2:2’ (3 μ M in POC1 (TFO) and 2 μ M in POC2 and POC3) and high ‘25:25:25’ (25 μ M in each POC). (b) Average images of the two subcategories (straight and bend) at high concentration. (c) The length and width was measured of the class average image and the aspect ratio was calculated. TEM class average images are 176 \times 176 \AA .

Facility, while POC samples at 50 μ M were measured at beamline B21 at the Diamond light Source (see Supplementary Information). Rewardingly, at the low concentrations, scattering corresponding to the POC1+POC2+POC3 trimer was observed. Interestingly, at high concentrations an apparent hexamer, interpreted as a dimer of trimers, assembled as the major component. Despite severe dissociation, both size-exclusion chromatography and gel electrophoresis showed the presence of both trimer and multimers, even at concentrations down to 0.5 μ M (Supplementary Figs 17 and 40).

The scattering of the POC1+POC2+POC3 assembly and the pair distance distribution function (Supplementary Figs 25 and 26) showed that at low concentrations (3.6 and 7.2 μ M) both the scattering extrapolated to zero angle and the maximum distances (13.0 and 12.7 nm) corresponded with a POC1+POC2+POC3 monomer. In contrast, at high concentration the mass (\sim 60 kDa) and overall size (\sim 16.5 nm) corresponded with a dimer of POC1+POC2+POC3. The longest distance was not double the distance found for the individual POCs, which indicated that either the dimer of trimers was not a fully linear structure, or that at least parts of the trimers interact and overlap.

Molecular modelling. We used molecular modelling to propose detailed POC1+POC2+POC3 models consistent with the SAXS measurements. Our approach (see Supplementary Methods and Supplementary Discussion section) yielded an ensemble of models for which the monomeric POC1+POC2+POC3 structure in Table 2b gave the best fit against the experimental SAXS data at low concentrations (3.6 and 7.2 μ M), while the data recorded at 50 μ M corresponded very well with the dimer of POC1+POC2+POC3 trimers also shown in Table 2b.

The calculated and experimental SAXS curves showed a very convincing agreement in the experimentally robust 0.01–0.2 \AA^{-1} region (see Supplementary Figs 27 and 28). Importantly, the

trimer itself can be ruled out as a significant species in solution at the high concentration, due to its characteristic SAXS curve shape deviation in the low q range calculated for the MD structures (Supplementary Fig. 28c). Significant amounts of higher-order oligomers, for example, trimers of trimers, can also be excluded since the balancing amount of trimer required to yield the effective molecular weight consistent with SAXS would change the shape of the SAXS curve (Supplementary Figs 28c and 30–31). This point is illustrated with MultiFoXS calculations yielding an inferior curve shape for the composite trimer-of-trimer and trimer curve (Supplementary Fig. 32). A key feature of the POC1+POC2+POC3 dimer of trimers model in Table 2b (Supplementary Fig. 29) is that the two trimers interact hydrophobically via their linkers. This positions Lys residues in one trimer close to the DNA backbone of another, enabling electrostatic interactions between these moieties. The hydrophobic linker interaction was also found for other dimer of trimer models (not shown) with generally low χ (\leq 20), whereas other types of interaction dominated in dimer of trimers with higher χ , suggesting that the actual solution shape (or ensemble of solution shapes) for the POC1+POC2+POC3 dimer of trimers is well-represented by the model in Table 2b.

Transmission electron microscopy. The POC1+POC2+POC3 assembly was analysed by TEM (Fig. 4 and Supplementary Fig. 33). See also Supplementary Discussion and Supplementary Methods for details) at two different concentrations, namely ‘3:2:2’ (3 μ M in POC1 (TFO) and 2 μ M in POC2 and POC3) and ‘25:25:25’ (25 μ M in each POC). The dimensions of the complexes were at the resolution limit of the microscope; hence detailed structural information was not extracted from the data. Class average images of the assemblies show an average higher aspect ratio of the complex at 25 μ M than at lower concentration indicative of longer assemblies (Fig. 4). Most of the particles distinguishable at the lower concentration appear to be sized as trimers with a minor proportion appearing as dimerized trimers. In contrast, at 25 μ M most of the POCs appear to be dimers of trimers (Fig. 4). It should be noted that drying and staining during the grid preparation procedure could alter the structure, but the increase in aspect ratio in the average images (Fig. 4) and the raw data (Supplementary Figs 34–39) is compatible with a transition from trimers to dimers of trimers at increased concentration. The TEM image analysis also suggests that the POCs have some internal flexibility, especially the 25:25:25, which make the class averaging more difficult. We note that though the TEM data are only indicative, they do give additional support to the observations from gel electrophoretic, CD and SAXS experiments that dimers of trimers dominate at higher concentration and single trimers at lower concentration.

Discussion

We envisioned that the combination of two fully orthogonal biomolecular self-assembly principles would enable the design of new protein-like structures and higher-order nanoscale-assemblies. Three POCs with molecular weights of 8.5–10.5 kDa were prepared by Cu-free alkyne azide coupling reactions. These POCs formed nano-assemblies which were characterized by ultraviolet thermal denaturation and annealing studies, gel electrophoresis, CD spectroscopy, SAXS and TEM. Ultraviolet and CD melting studies showed that duplexes and triple helices involving POCs are more stable than the non-peptide controls testifying to the importance of interactions between the peptide strands. In contrast, the apparent melting temperatures of the coiled coil peptide remained the same both in triple-helix POCs and as unconjugated peptide. CD spectroscopy further revealed

that the triple-helical POC1 + POC2 + POC3 assembly displayed a high degree of α -helicity, higher than that of the unconjugated peptide. To the best of our knowledge, such α -helicity values have never been reported in a purely aqueous solvent at ambient temperature which points to a helicity-transfer effect from the oligonucleotide triple-helix domain to the coiled coil peptide domain. SAXS performed at high POC concentration revealed the formation of a dimer of heterotrimers of triple-helix POCs, which was corroborated by computational simulation on the ensemble of POCs. At lower concentrations, the obtained SAXS data corresponded to the monomeric POC1 + POC2 + POC3 assembly. A higher-order assembly was also observed by gel electrophoresis. TEM analysis pointed to two different assembly morphologies, mostly trimers at low concentration and mostly dimers of trimers at high concentration. The structure obtained for the dimer of trimers points towards the possible use of the POC building blocks in creation of larger nanoscale assemblies. The successful construction of a trimeric protein-like structure and its dimer clearly demonstrates the potential of POCs for *de novo* design of proteins, including the potential to control the oligomeric state of peptide self-assemblies.

Methods

Synthesis of POC1. ON1-BCN (90 nmol, 0.25 ml, dissolved in Milli-Q water) was added to a solution of azidopeptide (113 nmol) in DMSO (0.8 ml) before H₂O (0.55 ml) was added. The resulting solution was transferred to a Biotage microwave reaction vial (2 ml) and sealed under an atmosphere of nitrogen. The reaction was carried out on a Biotage Initiator microwave synthesizer at 60 °C for 2 h, whereupon all solvents were removed *in vacuo* and the residue was re-dissolved in Milli-Q water. The synthesis procedure was repeated four times. The four crude solutions (from 360 nmol ON1-BCN) were combined, filtered using a GHP Acrodisc 13 mm syringe filter with 0.45 μ m GHP membrane and the solution was heated at 90 °C for 2 min to denature secondary structures before slowly cooling down to room temperature to give a solution of crude product ready for further purification.

Cyclic preparation protocol for synthesis of POC2 and POC3. ON2-BCN (53 nmol, 0.75 ml) was added into a solution of azidopeptide (66 nmol) in DMSO (0.8 ml) before H₂O (0.05 ml) was added. The solution was transferred into a Biotage microwave reaction vial (2 ml) and sealed under nitrogen atmosphere. The reaction was carried out on a Biotage Initiator microwave synthesizer at 60 °C for 3 h. After all solvents were removed *in vacuo*, a stepwise wash process was undertaken to separate the desired POC from unreacted ON. The white residue was first rinsed with Milli-Q water (2 \times 1 ml) to recover intact ON for next round reaction. The residue was subsequently dissolved in buffer A (2 \times 1 ml, 0.025 M Tris-HCl, 0.01 M sodium perchlorate, pH 7.6) to obtain the desired POCs. The synthesis procedure was repeated four times and all crude solutions in buffer A (from 212 nmol ON2-BCN) were combined, filtered using GHP Acrodisc 13 mm syringe filter with 0.45 μ m GHP membrane and heated to 90 °C to denature secondary structures before slowly cooling down to room temperature to give a solution of crude product for further purification. The same preparation cycle was applied on synthesis of POC3 (Supplementary Fig. 6): ON3-BCN (48 nmol, 0.735 ml) was added into a solution of azidopeptide (60 nmol) in DMSO (0.735 ml). The synthesis procedure was repeated four times. All crude solutions in buffer A (from 192 nmol ON3-BCN) were combined, filtered using GHP Acrodisc 13 mm syringe filter with 0.45 μ m GHP membrane and heated 90 °C to denature secondary structures before slowly cooling down to room temperature to give a solution of crude product for further purification.

Purification and analysis. The three crude POCs were purified by IE-HPLC using a GE PURIFIER 10 system equipped with a SOURCE 15Q 4.6/100 PE column (15 μ m, 100 mm \times 4.6 mm). Elution was performed with an isocratic hold of buffer A for 5 min, followed by a linear gradient to 35% buffer B in 45 min at a flow rate of 2.5 ml min⁻¹ (buffer A: 0.025 M Tris-HCl, 0.01 M sodium perchlorate, pH 7.6; buffer B: 0.025 M Tris-Cl, 1.0 M sodium perchlorate, pH 7.6). After IE-HPLC purification, the resulting solutions were de-salted by precipitation of the POC products by first adding an aqueous solution of sodium acetate (3 M, 15 μ l) followed by addition of cold ethanol (1 ml, 99% w/w; -20 °C). The resulting suspensions were stored at -20 °C for 1 h, and after centrifugation (13,200 r.p.m., 5 min, 4 °C) the supernatants were removed and the pellet further washed with cold ethanol (2 \times 1 ml; -20 °C), dried for 30 min under a flow of nitrogen and then dissolved in Milli-Q water (1.0 ml) to give POC1 (108 nmol, 56% in total), POC2 (98 nmol, 46%) and POC3 (76 nmol, 40%), respectively. Mass spectra of POCs were recorded on a Bruker micrOTOF II focus ESI-TOF MS instrument in ES⁻ mode

(representative MS in Supplementary Fig. 5). Analytical IE-HPLC traces were recorded on a Merck-Hitachi Lachrom system equipped with a DNAPac PA100 analytical column (13 μ m, 250 mm \times 4 mm) heated to 60 °C. Elution was performed with an isocratic hold of buffer B (10%), starting from 2 min hold on 2% Buffer A in Milli-Q water, followed by a linear gradient to 30% buffer A in 23 min at a flow rate of 1.1 ml min⁻¹ (buffer A: 1.0 M sodium perchlorate; buffer B: 0.25 M Tris-Cl, pH 8.0; representative IE-HPLC traces in Supplementary Fig. 5). Concentrations of purified POCs were determined by ultraviolet absorbance at 260 nm (extinction coefficients in Supplementary Methods).

Data availability. All relevant data are available from the corresponding authors on request.

References

- Nanda, V. & Koder, R. L. Designing artificial enzymes by intuition and computation. *Nat. Chem.* **2**, 15–24 (2010).
- Hamley, I. W. Peptide nanotubes. *Angew. Chem. Int. Ed.* **53**, 6866–6881 (2014).
- Gradiar, H. & Jerala, R. Self-assembled biostructures: proteins following the lead of DNA nanostructures. *J. Nanobiotechnol.* **12**, 4 (2014).
- Fletcher, J. M. *et al.* Self-assembling cages from coiled-coil peptide modules. *Science* **340**, 595–599 (2013).
- Yin, H. *et al.* Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* **15**, 541–555 (2014).
- Bandy, T. J. *et al.* DNA as supramolecular scaffold for functional molecules: progress in DNA nanotechnology. *Chem. Soc. Rev.* **40**, 138–148 (2011).
- Jones, M. R., Seeman, N. C. & Mirkin, C. A. Programmable materials and the nature of the DNA bond. *Science* **347**, 840 (2015).
- Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459**, 73–76 (2009).
- Hecht, N. H., Richardson, J. S., Richardson, D. C. & Ogden, R. C. *De novo* design, expression, and characterization of Felix: a 4-helix bundle protein of native-like sequence. *Science* **249**, 884–891 (1990).
- Baltzer, L., Nilsson, H. & Nilsson, J. *De novo* design of proteins—what are the rules? *Chem. Rev.* **101**, 3153–3164 (2001).
- DeGrado, W. F., Summa, C. M., Pavone, V., Nastro, F. & Lombardi, A. *De novo* design and structural characterization of proteins and metalloproteins. *Annu. Rev. Biochem.* **68**, 779–819 (1999).
- Hill, R. B., Raleigh, D. P., Lombardi, A. & DeGrado, W. F. *De novo* design of helical bundles as models for understanding protein folding and function. *Acc. Chem. Res.* **33**, 745–754 (2000).
- Dahiyat, B. I. & Mayo, S. L. *De novo* protein design: fully automated sequence selection. *Science* **278**, 82–87 (1997).
- DeGrado, W. F. Proteins from scratch. *Science* **278**, 80–81 (1997).
- DeGrado, W. F., Wasserman, Z. R. & Lear, J. D. Protein design, a minimalist approach. *Science* **243**, 622–628 (1989).
- Grell, D., Richardson, J. S. & Mutter, M. Protein design and folding: template trapping of self-assembled helical bundles. *J. Pept. Sci.* **7**, 146–151 (2001).
- Jensen, K. J. *in Peptide and protein design for biopharmaceutical applications* 1st edn. (Wiley, Copenhagen, 2009).
- Mutter, M. *et al.* The construction of new proteins. Part III. Artificial folding units by assembly of amphiphilic secondary structures on a template. *Helv. Chim. Acta* **71**, 835–847 (1988).
- Mutter, M. & Vuilleumier, S. A chemical approach to protein design—template-assembled synthetic proteins (TASP). *Angew. Chem. Int. Ed.* **28**, 535–554 (1989).
- Dawson, P. E. & Kent, S. B. H. Convenient total synthesis of a 4-helix template-assembled synthetic protein (TASP) molecule by chemoselective ligation. *J. Am. Chem. Soc.* **115**, 7263–7266 (1993).
- Freeman, J. O., Lee, W. C., Murphy, M. E. P. & Sherman, J. C. X-ray crystal analysis of a TASP: structural insights of a cavitein dimer. *J. Am. Chem. Soc.* **131**, 7421–7429 (2009).
- Kwak, J., De Capua, A., Locardi, E. & Goodman, M. TREN (tris(2-aminoethyl) amine): an effective scaffold for the assembly of triple helical collagen mimetic structures. *J. Am. Chem. Soc.* **124**, 14085–14091 (2002).
- Mezo, A. R. & Sherman, J. C. Cavitands are effective templates for inducing stability and nativelike structure in *de novo* four-helix bundles. *J. Am. Chem. Soc.* **121**, 8983–8994 (1999).
- Sasaki, T. & Kaiser, E. T. Helichrome: synthesis and enzymic activity of a designed hemeprotein. *J. Am. Chem. Soc.* **111**, 380–381 (1989).
- Wong, A. K., Jacobsen, M. P., Winzor, D. J. & Fairlie, D. P. Template assembled synthetic proteins (TASPs). Are template size, shape, and directionality important in formation of four-helix bundles? *J. Am. Chem. Soc.* **120**, 3836–3841 (1998).
- Høiberg-Nielsen, R. *et al.* 3-Instead of 4-helix formation in a *de novo* designed protein in solution revealed by small-angle X-ray scattering. *ChemBiochem.* **9**, 2663–2672 (2008).
- Hartmann, B. & Lavery, R. DNA structural forms. *Q. Rev. Biophys.* **29**, 309–368 (1996).

28. Frank-Kamenetskii, M. D. & Mirkin, S. M. Triplex DNA structures. *Annu. Rev. Biochem.* **64**, 65–95 (1995).
29. Lilley, D. M. J. Structures of helical junctions in nucleic acids. *Q. Rev. Biophys.* **33**, 109–159 (2000).
30. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **34**, 5402–5415 (2006).
31. Lu, K., Duan, Q. P., Ma, L. & Zhao, D. X. Chemical strategies for the synthesis of peptide-oligonucleotide conjugates. *Bioconjugate Chem.* **21**, 187–202 (2010).
32. Stephanopoulos, N. *et al.* Bioactive DNA-peptide nanotubes enhance the differentiation of neural stem cells into neurons. *Nano Lett.* **15**, 603–609 (2015).
33. Tang, W. & Becker, M. L. ‘Click’ reactions: a versatile toolbox for the synthesis of peptide-conjugates. *Chem. Soc. Rev.* **43**, 7013–7039 (2014).
34. Ghosh, P. S. & Hamilton, A. D. Noncovalent template-assisted mimicry of multiloop protein surfaces: assembling discontinuous and functional domains. *J. Am. Chem. Soc.* **134**, 13208–13211 (2012).
35. Humenik, M. & Scheibel, T. Nanomaterial building blocks based on spider silk-oligonucleotide conjugates. *ACS Nano.* **8**, 1342–1349 (2014).
36. Liang, S. I., McFarland, J. M., Rabuka, D. & Gartner, Z. J. A modular approach for assembling aldehyde-tagged proteins on DNA scaffolds. *J. Am. Chem. Soc.* **136**, 10850–10853 (2014).
37. Boice, J. A., Dieckmann, G. R., DeGrado, W. F. & Fairman, R. Thermodynamic analysis of a designed three-stranded coiled coil. *Biochemistry* **35**, 14480–14485 (1996).
38. Ogihara, N. L., Weiss, M. S., DeGrado, W. F. & Eisenberg, D. The crystal structure of the designed trimeric coiled coil-V(a)L(d): implications for engineering crystals and supramolecular assemblies. *Protein Sci.* **6**, 80–88 (1997).
39. Walshaw, J. & Woolfson, D. N. Socket: a program for identifying and analysing coiled-coil motifs within protein structures. *J. Mol. Biol.* **307**, 1427–1450 (2001).
40. Højland, T. *et al.* LNA (locked nucleic acid) and analogs as triplex-forming oligonucleotides. *Org. Biomol. Chem.* **5**, 2375–2379 (2007).
41. Jewett, J. C. & Bertozzi, C. R. Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* **39**, 1272–1279 (2010).
42. King, M. & Wagner, A. Developments in the field of bioorthogonal bond forming reactions-past and present trends. *Bioconjugate Chem.* **25**, 825–839 (2014).
43. Dommerholt, J. *et al.* Readily accessible bicyclonynes for bioorthogonal labeling and three-dimensional imaging of living cells. *Angew. Chem. Int. Ed.* **49**, 9422–9425 (2010).
44. Hodges, J. C., Van Delft, F. L., Van Berkel, S. S., Verkade, J. & Berry, D. A. Preparation of bicyclo[6.1.0]non-4-yne reagents for chemical modification of oligonucleotides. Patent WO2013036748A1 (2013).
45. Lee, M.-K. & Lim, Y.-B. Facile synthesis, optical and conformational characteristics, and efficient intracellular delivery of a peptide-DNA conjugate. *Bioorg. Med. Chem.* **22**, 4204–4209 (2014).
46. Dieckmann, G. R. *et al.* The role of protonation and metal chelation preferences in defining the properties of mercury-binding coiled coils. *J. Mol. Biol.* **280**, 897–912 (1998).
47. Gowers, D. M. & Fox, K. R. Towards mixed sequence recognition by triple helix formation. *Nucleic Acids Res.* **27**, 1569–1577 (1999).
48. Kypr, J., Kejnovska, I., Renciuik, D. & Vorlickova, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37**, 1713–1725 (2009).
49. Lau, S. Y. M., Taneja, A. K. & Hodges, R. S. Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils. *J. Biol. Chem.* **259**, 13253–13261 (1984).
50. Zhou, N. E., Kay, C. M. & Hodges, R. S. Synthetic model proteins. Positional effects of interchain hydrophobic interactions on stability of two-stranded alpha-helical coiled-coils. *J. Biol. Chem.* **267**, 2664–2670 (1992).
51. Malik, L. *et al.* Self-assembly of designed coiled coil peptides studied by small-angle X-ray scattering and analytical ultracentrifugation. *J. Pept. Sci.* **19**, 283–292 (2013).

Acknowledgements

We thank VILLUM FONDEN for funding The Biomolecular Nanoscale Engineering Center, a VILLUM center of excellence. The Tesla K40 used for this research was donated by the NVIDIA Corporation. We thank ATDBio Ltd. and Professor Tom Brown for help on ESI-MS. Dr Kasper K. Sørensen is acknowledged for his assistance in peptide synthesis; we thank Joan Hansen, Louise Johansson and Tina Grubbe Hansen for technical assistance on oligonucleotide synthesis and purification. Dr Nicola Derbyshire is acknowledged for assistance on gel electrophoresis.

Author contributions

The present work was designed and supervised by J.W. and K.J.J., C.L. performed synthesis and characterization of ONs and POCs, ultraviolet-melting studies and gel electrophoresis experiments; C.S.M. performed the synthesis of azido-peptide; S.R.M. carried out SAXS analysis; M.C.M.-M. and P.W.T. performed the CD experiments; N.J.C. carried out the molecular modelling calculations; and R.P.T. and J.K. performed the TEM studies. C.L., C.S.M., J.W. and K.J.J. co-wrote the manuscript draft. All authors contributed to data analysis and edited the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Lou, C. *et al.* Peptide-oligonucleotide conjugates as nanoscale building blocks for assembly of an artificial three-helix protein mimic. *Nat. Commun.* **7**:12294 doi: 10.1038/ncomms12294 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016