ARTICLE
PUBLISHED ONLINE: 28 APRIL 2013 | DOI: 10.1038/NCHEMBIO.1248

nature chemical biology

Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments

Helena Gradišar[1,2], Sabina Božič[1], Tibor Doles[1,2], Damjan Vengust[3], Iva Hafner-Bratkovič[1], Alenka Mertelj[3,4], Ben Webb[5–7], Andrej Šali[5–7], Sandi Klavžar[4,8] & Roman Jerala[1,2,9]*

Protein structures evolved through a complex interplay of cooperative interactions, and it is still very challenging to design new protein folds *de novo*. Here we present a strategy to design self-assembling polypeptide nanostructured polyhedra based on modularization using orthogonal dimerizing segments. We designed and experimentally demonstrated the formation of the tetrahedron that self-assembles from a single polypeptide chain comprising 12 concatenated coiled coil–forming segments separated by flexible peptide hinges. The path of the polypeptide chain is guided by a defined order of segments that traverse each of the six edges of the tetrahedron exactly twice, forming coiled-coil dimers with their corresponding partners. The coincidence of the polypeptide termini in the same vertex is demonstrated by reconstituting a split fluorescent protein in the polypeptide with the correct tetrahedral topology. Polypeptides with a deleted or scrambled segment order fail to self-assemble correctly. This design platform provides a foundation for constructing new topological polypeptide folds based on the set of orthogonal interacting polypeptide segments.

The base complementarity of DNA can now be used to rationally design artificial nanostructures that can adopt versatile two- and three-dimensional shapes, such as different polyhedra[1–4]. In contrast, the design of new polypeptide folds is substantially more complicated because of the contribution of many cooperative and long-range interactions. Yet in nature, polypeptides perform the most complex tasks because they can form intricate tertiary structures with versatile chemical properties and functionalities. *De novo* protein fold design has been, with few exceptions, successful only for folds similar to those existing in nature[5–8]. Self-assembled symmetrical polypeptide assemblies have been designed on the basis of natural protein oligomerization domains[9,10], and the interacting surfaces of natural trimerizing protein domains have been optimized for self-assembly into tetrahedral and octahedral nanostructures[11,12]. Peptides have also been used for the self-assembly of fibrils, nanotubes and nanospheres (reviewed in ref. 13). The use of modular building elements could facilitate the design of new protein folds. One type of modular element is the coiled-coil dimer, whose structure, comprising intertwined helices, was first proposed in 1952 (ref. 14). Interactions that govern the specificity of coiled-coil dimerization are fairly well understood[15] as coiled coils are characterized by a heptad repeat that forms two turns of a helix spanning approximately 1 nm. The formation of coiled-coil dimers is governed by hydrophobic and electrostatic interactions involving residues at positions a and d of the periodic heptad repeat and e and g, respectively (**Supplementary Results**, **Supplementary Fig. 4**). The specificity and orthogonality of the desired coiled-coil combination can be further improved by negative design, for example, by introducing an asparagine at a hydrophobic coiled-coil interaction position[16]. Staggered arrangement of self-complementary coiled coil–forming segments has been used to prepare polypeptide nanofibrils[17–19]. However, combining two types of interacting coiled-coil elements can only lead to one-dimensional fibrillar assemblies[18], and at least three different elements are required for two- or three-dimensional assemblies.

We surmised that the concatenated orthogonal polypeptide interacting segments could be used to guide the self-assembly of complex polypeptide polyhedra. This strategy was tested on the self-assembly of a tetrahedral polypeptide fold designed from 12 coiled coil–forming peptides interspersed with short flexible peptide linkers and validated by determining the hydrodynamic diameter and stability by transmission electron microscopy (TEM) and atomic force microscopy (AFM) imaging and by confirming the coincidence of the N- and C-terminal segments.

RESULTS

Design of polyhedral fold from a single chain

We reasoned that it should be possible to self-assemble the designed three-dimensional objects from a single polypeptide chain composed of concatenated coiled coil–forming segments. This strategy relies on the ability of those segments to pair in a selected orientation with their complementary interacting segments within the same polypeptide chain, thus driving the self-assembly. Each of those coiled coil–forming segments is, in isolation, unstructured and forms a coiled-coil helix only when it dimerizes with the corresponding complementary segment. Each edge of the constructed polyhedron should be traversed by a polypeptide chain exactly twice to form a rigid coiled-coil dimer. The topology of the self-assembled polypeptide chain should therefore be defined by the orientation and sequential arrangement of each coiled-coil pair. A tetrahedron is the simplest three-dimensional geometric object consisting of six edges; thus, a tetrahedron could be constructed from six pairs, that is, 12 concatenated coiled-coil dimer–forming segments. In topological terms, the polypeptide path along the edges

[1]Department of Biotechnology, National Institute of Chemistry, Ljubljana, Slovenia. [2]Excellent NMR–Future Innovation for Sustainable Technologies Centre of Excellence, Ljubljana, Slovenia. [3]Jožef Stefan Institute, Ljubljana, Slovenia. [4]Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia. [5]Department of Bioengineering and Therapeutic Sciences, University of California–San Francisco, San Francisco, California, USA. [6]Department of Pharmaceutical Chemistry, University of California–San Francisco, San Francisco, California, USA. [7]California Institute for Quantitative Biosciences, University of California–San Francisco, San Francisco, California, USA. [8]Faculty of Natural Sciences and Mathematics, University of Maribor, Maribor, Slovenia. [9]Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia. *e-mail: roman.jerala@ki.si

of a polyhedron corresponds to a Eulerian trail, an oriented path that connects each of the two vertices by exactly two connections. Such a trail is guaranteed to exist by the graph theory because all vertices of a double tetrahedral graph have an even degree. We formally proved that polyhedra are, in principle, realizable by a single multisegmental polypeptide chain approach (**Supplementary Note**). Additional constraints on the topology of the polypeptide trail were imposed to ensure the stability of the polyhedron by interlocking the edges at each vertex (**Supplementary Note** and **Supplementary Fig. 1**). Unexpectedly, we proved by graph theory analysis that a tetrahedron cannot be constructed exclusively from either parallel or antiparallel dimeric segments; it requires a combination of both types. The ability of the polypeptide chain to form either parallel- or antiparallel-oriented dimers is therefore an additional advantage over nucleic acid–based designed nanostructures because a nucleic acid duplex can only assume an antiparallel orientation. We showed that the tetrahedron could self-assemble from a single polypeptide chain comprising either three parallel and three antiparallel coiled-coil pairs or four parallel and two antiparallel pairs, altogether encompassing three distinct topological solutions (**Supplementary Note** and **Supplementary Figs. 2** and **3**). The number of different path topologies increased rapidly with the number of polyhedral edges; however, the number of parallel- or antiparallel-only topologies is very low (**Supplementary Table 1**). Additionally, we note that for all of the polyhedra traversed by a single chain, the beginning and end of the polypeptide trail must coincide at the same vertex, which can be tested experimentally.

Design of a single-chain polypeptide tetrahedron

For the experimental test of the design of the tetrahedral topology, we chose to construct the topology comprising four parallel and two antiparallel coiled-coil dimeric edges (**Fig. 1**), which is the topology with the largest number of parallel segments. Although both parallel and antiparallel coiled-coil dimeric arrangements are represented in nature[20], a substantially higher number of isolated parallel coiled-coil dimers have been characterized and designed[15,21,22]. We selected the pairs for the formation of three edges of the polyhedron from the orthogonal coiled coil–forming pairs P3-P4, P5-P6 and P7-P8. Peptides P3 to P8, each composed of four heptad repeats (amino acid sequences are in Online Methods), were designed on the basis of known coiled-coil stability and selectivity principles[21], and the orthogonality of the peptide pairs was demonstrated (**Supplementary Fig. 6**). None of these peptides share three or more consecutive heptads with the same interacting motif at positions a, d, e and g. In addition to using the designed parallel heterodimers, we used one parallel homodimer (based on GCN4 (ref. 23)) and two antiparallel homodimers (APH[24] and BCR peptide[25]; amino acid sequences are in Online Methods) to demonstrate that not only heterodimers but also homodimers can be used for the self-assembly of polypeptide structures (**Supplementary Fig. 6b**). The tetrapeptide Ser-Gly-Pro-Gly was selected as the flexible linker to connect the consecutive coiled coil–forming segments. This linker introduces the strong helix-interrupting residues proline and glycine to prevent formation of an extended helix spanning several coiled coil–forming segments and to provide hinge flexibility at the vertices, which is required for the formation of acute angles between the edges of a tetrahedron. The synthetic gene encoding the designed polypeptide, TET12 (**Fig. 1a** and **Supplementary Table 2**), was codon optimized for *Escherichia coli*, and the polypeptide was expressed in bacteria, purified (**Supplementary Fig. 7a**) and refolded by dialysis. The secondary structure of the self-assembled polypeptide, assessed by CD (**Fig. 2a**), was consistent with the designed molecular model with high helical content. Each of the six isolated peptide pairs forming the edges of the tetrahedron unfolded between 0.5 M and 1.8 M guanidine hydrochloride (GdnHCl) (**Fig. 2b**). Although the cooperativity of

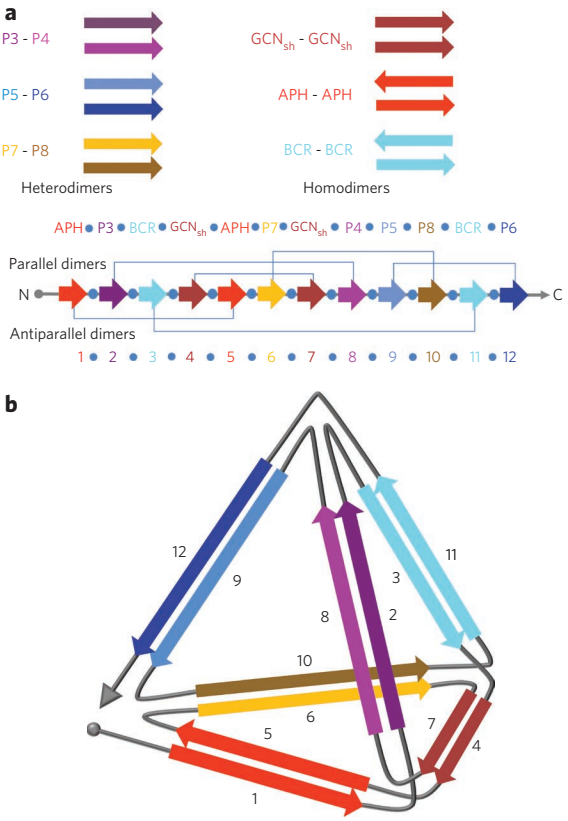

Figure 1 | Design of the self-assembling polypeptide tetrahedron. (**a**) The designed orthogonal peptide pairs P3-P4, P5-P6 and P7-P8 and homodimeric peptides APH, BCR and shortened natural GCN (GCN_{sh}) are used to construct a tetrahedron-forming polypeptide chain, TET12. Twelve coiled coil–forming elements are concatenated in a defined order, separated by the tetrapeptide linker Ser-Gly-Pro-Gly (blue circle), wherein the coiled coil–forming segments numbered 1 to 12 form parallel heterodimers (2-8, 6-10, 9-12), a parallel homodimer (4-7) and antiparallel homodimers (1-5, 3-11). (**b**) Schematic representation of the polypeptide path forming a tetrahedron. Arrows denote the orientation of the interacting coiled coil–forming elements in the assembly.

the designed tetrahedral fold is provided by the sequential order of concatenated coiled coil–forming elements rather than by the hydrophobic core, as in the native protein folds, the annealed tetrahedral polypeptide had substantially greater stability than its coiled-coil segment constituents, with a midpoint of unfolding at 3 M GdnHCl (**Fig. 2b**). The self-assembled polypeptide particles had an average hydrodynamic diameter of 6.9 ± 0.4 nm, as determined by dynamic light scattering (DLS) (**Fig. 2c**), a value close to the calculated dimensions[26] of the molecular model of the designed tetrahedron (**Supplementary Fig. 5**). Its diameter was also substantially more compact than that of TET12 polypeptide denatured in 6 M GdnHCl, which was 20.0 ± 2.9 nm (**Fig. 2c**). At a polypeptide concentration above 4 μM in the refolding mixture, coiled coil–forming elements of TET12 formed intermolecular interactions with the complementary segments from other molecules during the refolding process, which led to formation of aggregates with hydrodynamic size above 80 nm (**Supplementary Fig. 8**).

Visualization of the assembled structures by microscopy

AFM was used to visualize the self-assembled polypeptide structures that were bound to the mica substrate. We found small, discrete, homogeneous objects that were produced under the slow annealing conditions by dialysis at low polypeptide concentrations

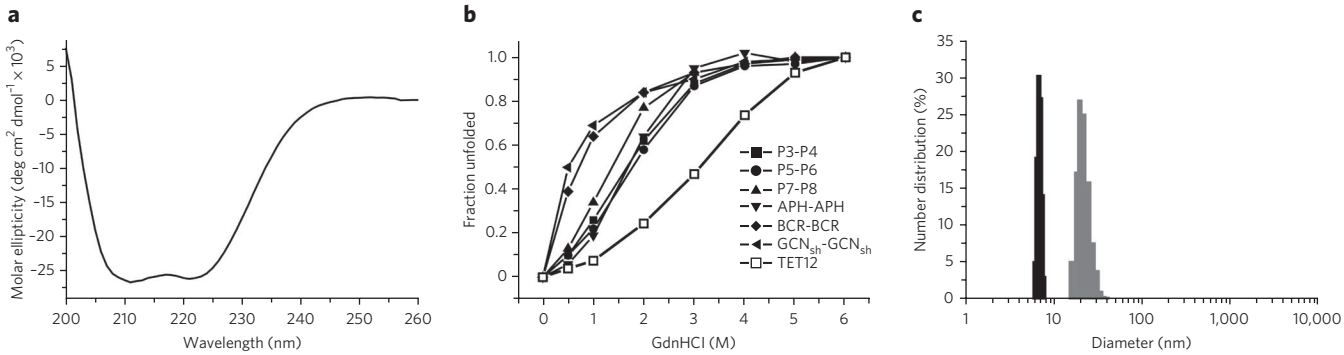

Figure 2 | Self-assembly of the tetrahedron-forming polypeptide TET12. (**a**) Analysis of the secondary structure of self-assembled TET12 polypeptide was performed by CD measurement. The average curve of three scans of the representative sample is shown. The analysis of this far-UV CD spectrum indicates 82% of α-helical structure. (**b**) The thermodynamic stability of TET12 was increased in comparison to its dimeric coiled-coil peptide-building elements, as determined by the CD signal at 222 nm at different concentrations of GdnHCl. The representative curves are shown. (**c**) The hydrodynamic diameter of TET12 unfolded in 6 M GdnHCl (gray) was 20.0 ± 2.9 nm and that of self-assembled TET12 (black) was 6.9 ± 0.4 nm, as determined by DLS (values are mean ± s.d. of three independent experiments). A representative histogram is shown.

(100 nM), wherein structural features were arranged in a triangular shape with dimensions below 25 nm (**Fig. 3a**). AFM images of polypeptide particles were similar to the images of DNA tetrahedra formed from trigonal DNA star components[3]. As observed in DNA tetrahedra, we observed a collapse of our tetrahedral three-dimensional structure probably due to interactions with mica and dehydration. Additionally, the inherent characteristic of AFM is a tip-sample convolution, which makes features appear wider and is especially prominent when features are of a similar size as the tip diameter. Further, we used positive uranyl staining to obtain the sufficient contrast of small polypeptide particles for TEM visualization. The observed particles, which had a triangular shape or other projections of a tetrahedron, measured approximately 5 nm along inner edges (**Supplementary Fig. 9a**), in agreement with hydrodynamic size measurement (**Fig. 2c**) and consistent with the length of the designed coiled-coil dimers (**Supplementary Fig. 5**). We observed an even more convincing topological arrangement of the vertices of the particles by using a prolonged uranyl staining procedure, which resulted in an increased concentration of the stain at vertices, where the three linkers converge (**Fig. 3b** and **Supplementary Fig. 9b**). In those images, we observed the arrangement of vertices corresponding to different projections of a tetrahedron. Like polypeptide refolding by dialysis, slow temperature annealing also produced objects of similar size and shape (**Supplementary Fig. 9e**). Annealing at the higher, aggregation-inducing 10-μM concentration led to formation of a connected network (**Supplementary Fig. 10**). Polypeptide TET12 comprised an N-terminal His_6 polypeptide tag that can bind Ni^{2+} ions complexed with nitrilotriacetic acid (NTA). By adding 1.8-nm nanogold beads coated with Ni-NTA, we were able to locate the polypeptide N terminus in the self-assembled polypeptide TET12 tetrahedron. Clearly recognizable triangles and other tetrahedral projections were visible by TEM, and a dense nanogold sphere attached to one vertex (**Fig. 3c,d** and **Supplementary Fig. 9c,d**) improved the contrast and provided additional evidence of the dimensions of the formed nanostructures, corroborating the success of the design.

Verification of a closed polypeptide topological path

The N and C terminus of the TET12 polypeptide must, by design, coincide at the same vertex of the self-assembled tetrahedron (**Fig. 1b** and **Supplementary Note**). We confirmed this property by grafting the N- and C-terminal segment of a split YFP to each terminus of the polypeptide. Therefore, YFP fluorescence could only be reconstituted by correct self-assembly (**Fig. 4a**). This is

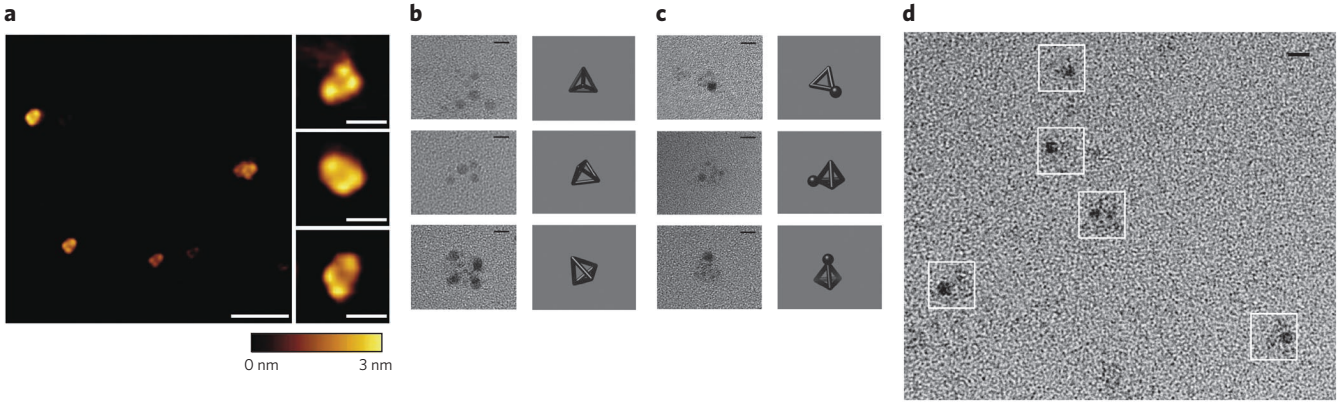

Figure 3 | Imaging of the assembled TET12 reveals tetrahedral topology. (**a**) A representative AFM image of the self-assembled TET12 polypeptide (left; scale bar, 100 nm) and close-up views of three tetrahedral particles (right; scale bars, 20 nm). A height scale bar is shown at the bottom. (**b**–**d**) TEM images of self-assembled TET12 tetrahedral structures. Representative particles from TEM images and projections of a tetrahedron in the matching orientation are shown, where the samples on grids were stained either with uranyl for 3 min (**b**) or first with 1.8-nm Ni-NTA–nanogold beads followed by 1 min uranyl staining (**c**). A representative TEM image field with tetrahedral structures is shown in **d**. White boxes indicate the identified particles. Scale bars in **b**–**d** represent 5 nm.

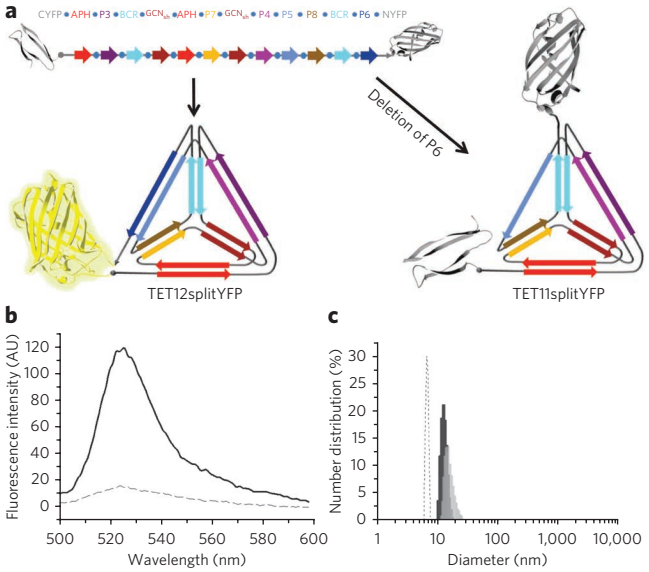

Figure 4 | Determination of the sequence-dependent topology of the self-assembled polypeptide. (**a**) Scheme of the polypeptide chains TET12splitYFP and TET11splitYFP, used to establish the coincident N and C termini of the polypeptide chain and the requirement for a defined order of segments for the assembly of a tetrahedron. CYFP, C-terminal segment of YFP; NYFP, N-terminal segment of YFP. (**b**) Reconstituted YFP fluorescence measurements reveal the correct self-assembly of TET12splitYFP (solid line), whereas the deletion of a single building element (TET11splitYFP) prevents formation of complete tetrahedron and YFP reconstitution (dashed line). The average curves of three scans of the representative samples are shown. Fluorescence intensity is shown in arbitrary units (AU). (**c**) DLS measurements show increased hydrodynamic diameters of two assembled polypeptides, TET11 with a single deleted segment (13.9 ± 1.1 nm, dark gray) and TET12scr with two swapped segments (15.0 ± 1.5 nm, light gray), in comparison to the diameter of the assembled TET12 (6.9 ± 0.4 nm, dotted). Values are mean ± s.d. of three independent experiments. A representative histogram is shown.

a strong topological constraint because the first and the last segments of TET12 do not interact directly, and the proximity of both termini of the self-assembled tetrahedron is a consequence of the complex global topology, encoded in the sequential order of coiled coil–forming segments (**Fig. 1a**,**b**). The recombinant fusion polypeptide TET12splitYFP, isolated from bacteria (**Supplementary Fig. 7b**), had high α-helix content (data not shown), yet fluorescence was not observed, suggesting misfolding of the tetrahedron in bacteria presumably as a result of kinetically trapped intermediates or formation of intermolecular aggregates. This finding demonstrates that YFP does not reconstitute in nonspecific aggregates. However, dialysis of the fusion polypeptide TET12splitYFP solubilized in 6 M GdnHCl resulted in high fluorescence intensity of the reconstituted split YFP (**Fig. 4b**), confirming formation of the correct circular path of the chain and the coincidence of both termini of the polypeptide chain at the same vertex of the structure. TEM analysis additionally confirmed the formation of tetrahedron-like particles from TET12splitYFP polypeptide (**Supplementary Fig. 9f**).

Tetrahedral self-assembly depends on the segment order

Polypeptide self-assembly into the tetrahedral topology depended on the order of coiled coil–forming segments within the polypeptide, just as folding of the native protein relies on the precise amino acid sequence. Thus, the self-assembly should be sensitive to the disruption of the order of coiled-coil building elements. To this end, we tested two additional polypeptide constructs: one in which a single coiled coil–forming segment was deleted (TET11) and another scrambled construct that retained all coiled coil–forming segments but swapped two segments (TET12Scr; segment order 4, 5, 1, 2, 3, 6, 7, 8, 9, 10, 11, 12; amino acid sequences are in Online Methods). We predicted that both polypeptides would fail to adopt the tetrahedral topology. Indeed, both constructs had increased aggregation during refolding, increased hydrodynamic size of the soluble monomers (13.9 ± 1.1 nm and 15.0 ± 1.5 nm for TET11 and TET12Scr, respectively) (**Fig. 4c**) and decreased stability (**Supplementary Fig. 11**). We did not observe formation of any regular structures by TEM in either the truncated or scrambled constructs (data not shown), in contrast to that observed for TET12. Moreover, in the TET11 polypeptide, which had a deletion of a single coiled coil–forming segment (P6), the N and C terminus no longer coincided, and, as expected, the tetrahedral polypeptide (TET11splitYFP) could not reconstitute YFP, in contrast to the integral tetrahedral construct TET12splitYFP (**Fig. 4b**).

DISCUSSION

We have demonstrated that the rational design of tertiary structures from natural polymers, as pioneered by the DNA-based self-assembly, can be extended to polypeptides using the specificity of pairwise interactions between the coiled-coil building elements. Modularization of the design, based on the well-understood building blocks, allowed us to use topological design rather than full atomistic modeling. The ability of coiled coil–forming peptide segments to adopt either a parallel or antiparallel orientation expanded the number of reachable topological arrangements and facilitated the formation of a complex polypeptide scaffold. This principle enabled us to design a tetrahedral topology that self-assembles from a single polypeptide chain, whereas the tetrahedron based on DNA nanoscale folding has to be composed of several chains or include a four-stranded edge[27]. The method described above allowed us to create a new topological protein fold, which has not been observed in nature. The design, based on concatenated coiled coil–forming segments, can produce asymmetric structures, which are difficult to achieve by the assembly of protein domains or for other nanoparticles composed of organic or inorganic building blocks. We demonstrated attachment of nanogold or protein domain cargo to the N and/or C termini of the polypeptide and suggested that other vertices of the polypeptide structures could similarly be decorated with selected functional groups. Owing to their variable length and high length/width aspect ratio, coiled-coil dimers seem to be an almost ideal type of building element, allowing formation of large cavities within the polypeptide polyhedra, unlike interacting protein oligomerization domains, which occupy a major fraction of the volume. Interactions between coiled-coil segments can also be used to combine several building blocks into larger assemblies[28]. The presented technology could therefore be used for a range of applications, such as using designed cavities and arranging amino acid side chains for drug delivery or the creation of artificial catalytic sites. Polypeptides can be produced cost effectively in the recombinant form, which has clear advantages for the technological implementation of designed polypeptide structures. The current experimental limitation is the number of available orthogonal coiled coil–forming segments. More than 2,000 coiled coil–containing proteins in the human proteome[29] and hundreds of natural coiled coils[30,31] represent the source of potential building elements, whereas the rational design of new orthogonal sets may extend the variability in size or oligomerization state or other features. The principles of 'polypeptide origami' demonstrated in the design of a tetrahedron could be applied to construct other polyhedra, such as those listed in **Supplementary Table 1**, or, in general, any shape that can be tessellated.

Received 14 February 2013; accepted 28 March 2013; published online 28 April 2013

METHODS

Methods and any associated references are available in the online version of the paper.

References

1. Chen, J.H. & Seeman, N.C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **350**, 631–633 (1991).
2. Rothemund, P.W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
3. He, Y. *et al.* Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **452**, 198–201 (2008).
4. Han, D. *et al.* DNA origami with complex curvatures in three-dimensional space. *Science* **332**, 342–346 (2011).
5. Kuhlman, B. *et al.* Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364–1368 (2003).
6. Regan, L. & DeGrado, W.F. Characterization of a helical protein designed from first principles. *Science* **241**, 976–978 (1988).
7. Fleishman, S.J. *et al.* Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816–821 (2011).
8. Hecht, M.H., Richardson, J.S., Richardson, D.C. & Ogden, R.C. *De novo* design, expression, and characterization of Felix: a four-helix bundle protein of native-like sequence. *Science* **249**, 884–891 (1990).
9. Padilla, J.E., Colovos, C. & Yeates, T.O. Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proc. Natl. Acad. Sci. USA* **98**, 2217–2221 (2001).
10. Doles, T., Bozic, S., Gradisar, H. & Jerala, R. Functional self-assembling polypeptide bionanomaterials. *Biochem. Soc. Trans.* **40**, 629–634 (2012).
11. King, N.P. *et al.* Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science* **336**, 1171–1174 (2012).
12. Lai, Y.T., Cascio, D. & Yeates, T.O. Structure of a 16-nm cage designed by using protein oligomers. *Science* **336**, 1129 (2012).
13. Lakshmanan, A., Zhang, S. & Hauser, C.A. Short self-assembling peptides as building blocks for modern nanodevices. *Trends Biotechnol.* **30**, 155–165 (2012).
14. Crick, F.H. Is α-keratin a coiled coil? *Nature* **170**, 882–883 (1952).
15. Woolfson, D.N. The design of coiled-coil structures and assemblies. *Adv. Protein Chem.* **70**, 79–112 (2005).
16. Fong, J.H., Keating, A.E. & Singh, M. Predicting specificity in bZIP coiled-coil protein interactions. *Genome Biol.* **5**, R11 (2004).
17. Potekhin, S.A. *et al. De novo* design of fibrils made of short α-helical coiled coil peptides. *Chem. Biol.* **8**, 1025–1032 (2001).
18. Gribbon, C. *et al.* MagicWand: a single, designed peptide that assembles to stable, ordered α-helical fibers. *Biochemistry* **47**, 10365–10371 (2008).
19. Gauba, V. & Hartgerink, J.D. Self-assembled heterotrimeric collagen triple helices directed through electrostatic interactions. *J. Am. Chem. Soc.* **129**, 2683–2690 (2007).
20. Moutevelis, E. & Woolfson, D.N. A periodic table of coiled-coil protein structures. *J. Mol. Biol.* **385**, 726–732 (2009).
21. Gradišar, H. & Jerala, R. *De novo* design of orthogonal peptide pairs forming parallel coiled-coil heterodimers. *J. Pept. Sci.* **17**, 100–106 (2011).
22. Bromley, E.H., Sessions, R.B., Thomson, A.R. & Woolfson, D.N. Designed α-helical tectons for constructing multicomponent synthetic biological systems. *J. Am. Chem. Soc.* **131**, 928–930 (2009).
23. Lumb, K.J., Carr, C.M. & Kim, P.S. Subdomain folding of the coiled coil leucine zipper from the bZIP transcriptional activator GCN4. *Biochemistry* **33**, 7361–7367 (1994).
24. Gurnon, D.G., Whitaker, J.A. & Oakley, M.G. Design and characterization of a homodimeric antiparallel coiled coil. *J. Am. Chem. Soc.* **125**, 7518–7519 (2003).
25. Taylor, C.M. & Keating, A.E. Orientation and oligomerization specificity of the Bcr coiled-coil oligomerization domain. *Biochemistry* **44**, 16246–16256 (2005).
26. Ortega, A., Amoros, D. & Garcia de la Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898 (2011).
27. Li, Z. *et al.* A replicable tetrahedral nanostructure self-assembled from a single DNA strand. *J. Am. Chem. Soc.* **131**, 13093–13098 (2009).
28. Boyle, A.L. *et al.* Squaring the circle in peptide assembly: from fibers to discrete nanostructures by *de novo* design. *J. Am. Chem. Soc.* **134**, 15457–15467 (2012).
29. Rackham, O.J. *et al.* The evolution and structure prediction of coiled coils across all genomes. *J. Mol. Biol.* **403**, 480–493 (2010).
30. Reinke, A.W., Grant, R.A. & Keating, A.E. A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. *J. Am. Chem. Soc.* **132**, 6025–6031 (2010).
31. Newman, J.R. & Keating, A.E. Comprehensive identification of human bZIP interactions with coiled-coil arrays. *Science* **300**, 2097–2101 (2003).

Acknowledgments

This research was supported by grants from the Slovenian Research Agency (J2-2131, P4-0176 to R.J.) and the Excellent NMR–Future Innovation for Sustainable Technologies Centre of Excellence, which is financed in part by the European Union regional development funds. B.W. and A.Š. acknowledge US National Institutes of Health grants R01 GM083960 and R01 GM54762 (both to A.Š.). We thank the rest of the members of the 2009 Slovenian International Genetically Engineered Machine (iGEM) competition team (students M. Verce, A. Lukan, N. Debeljak, Š. Miklavič and U. Jelerčič and mentors O. Fekonja, J. Pohar, R. Bremšak and M. Benčina) for their pioneering work on the development of concatenated coiled coil–based nanostructures (http://2009.igem.org/Team:Slovenia), underlying the development of the polypeptide polyhedra; R. Bremšak for excellent technical support; J. Rus for calculations regarding the number of polyhedral topology; the Centre for Electron Microscopy at the Jožef Stefan Institute for the use of electron microscopes; and K. Djinović Carugo for comments on the manuscript.

Author contributions

R.J. conceived the idea, designed the tetrahedral polypeptide, participated in the mathematical analysis of the polyhedral topology, coordinated the project, discussed the results and wrote the manuscript. H.G. performed TEM, DLS, CD and fluorescence experiments; coordinated the project; discussed the results; and wrote the manuscript. S.B. prepared, purified and analyzed polypeptides; performed TEM, CD and fluorescence experiments; discussed the results; and wrote the manuscript. T.D. prepared, purified, and analyzed polypeptides; performed AFM measurements; discussed the results; and wrote the manuscript. D.V. performed TEM measurements. I.H.-B. performed AFM measurements. A.M. performed DLS measurements. B.W. prepared the molecular model. A.Š. prepared the molecular model and wrote the manuscript. S.K. solved the mathematical analysis of the polyhedral topology and wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to R.J.

ONLINE METHODS

Peptide and linker sequences. The sequences of the coiled coil–forming peptides used in the TET12 are artificial sequences P3–P8 and APH, as well as natural BCR and a homodimer-forming peptide, shortened natural GCN (GCN_{sh}).

P3: SPEDEIQQLEEEIAQLEQKNAALKEKNQALKYG;

P4: SPEDKIAQLKQKIQALKQENQQLEEENAALEYG;

P5: SPEDENAALEEKIAQLKQKNAALKEEIQALEYG;

P6: SPEDKNAALKEEIQALEEENQALEEKIAQLKYG;

P7: SPEDEIQALEEKNAQLKQEIAALEEKNQALKYG;

P8: SPEDKIAQLKEENQQLEQKIQALKEENAALEYG;

APH: MKQLEKELKQLEKELQAIEKQLAQLQWKAQARKKKLAQLKKKLQA;

BCR: DIEQELERAKASIRRLEQEVNQERSRMAYLQTLLAK;

GCN_{sh}: QLEDKVEELLSKNYHLENEVARLKKLVG

The tetrapeptide linker Ser-Gly-Pro-Gly was positioned between every two consecutive coiled coil–forming peptide segments to interrupt the sequence with helical propensity and enable the flexibility at the vertices.

Molecular modeling of a TET12 tetrahedron. A molecular model of the self-assembled TET12 was constructed using MODELLER comparative modeling software, version 9.10 (http://salilab.org/modeller/)[32,33]. Structural features, such as atom-atom distances and dihedral angles, were extracted from known three-dimensional structures of coiled-coil dimers (Protein Data Bank codes 1ZIK for parallel dimers, and 1K1F for antiparallel dimers). Candidate structures of the designed sequence were restrained to reproduce these features along each tetrahedron edge, as were correct stereochemistry and excluded volume and known properties of proteins such as main chain and side chain dihedral angle distributions. To reduce the conformational search space, the linkers were initially placed at the tetrahedron vertices, and the remaining residues were simply interpolated between them. Two thousand candidate models were generated, and the model that best satisfied all restraints was selected. The structure was verified using MolProbity[34].

Construction of plasmids. The plasmid BB-NIC-II-HisN-TET12 used for the production of TET12 polypeptide was constructed by ligation of the NgoMIV- and SpeI-digested synthetic gene (1,392 bp) encoding polypeptide TET12 synthesized by Geneart (Regensburg, GE) into plasmid BB-NIC-II-HisN, previously cut with NgoMIV and SpeI (2,261 bp, http://partsregistry.org/wiki/index.php?title=Part:BBa_K245005). BB-NIC-II-HisN contains the T7 promoter and a His tag coding sequence upstream of the gene. TET12scr was constructed with ligation of three PCR products. Regions encoding GCN_{sh}-APH, APH-P3-BCR and P7-GCN_{sh}-P4-P5-P8-BCR-P6 were amplified from the TET12 gene and first ligated into BB-NIC-II-HisN separately through NgoMIV and SpeI restriction sites. Afterward, BB-NIC-II-HisN-APH-P3-BCR was cut with EcoRI and BspEI and ligated as a front insert into EcoRI- and NgoMIV-digested BB-NIC-II-HisN-P7-GCN_{sh}-P4-P5-P8-BCR-P6. BB-NIC-II-HisN-GCN_{sh}-APH was cut with EcoRI and BspEI and ligated as a front insert into BB-NIC-II-HisN-APH-P3-BCR-P7-GCN_{sh}-P4-P5-P8-BCR-P6 previously cut with EcoRI and NgoMIV. The resulting order of the coiled coil–forming segments was GCN_{sh}-APH-APH-P3-BCR-P7-GCN_{sh}-P4-P5-P8-BCR-P6 (TET12scr). TET11 was obtained by PCR amplification of the APH-P3-BCR-GCN_{sh}-APH-P7-GCN_{sh}-P4-P5-P8-BCR encoding region from the TET12 gene. The PCR product was then digested by NgoMIV and SpeI and subcloned into the BB-NIC-II-HisN plasmid previously cut with NgoMIV and SpeI. Split YFP fusions were constructed by joining the C-terminal segment of YFP (CYFP) at the 5′ end and the segment encoding the N-terminal portion of YFP (NYFP) to the 3′ end of the TET12- or TET11-coding regions. First, BB-NIC-II-HisN-CYFP (this plasmid was obtained by amplifying CYFP from a synthetic gene synthesized by Geneart and subcloning it into BB-NIC-II-HisN) was digested with EcoRI and BspEI, and the resulting insert was ligated into EcoRI- and NgoMIV-digested BB-NIC-II-HisN-TET12 or BB-NIC-II-HisN-TET11. Afterward, BB-NIC-II-HisN-CYFP-TET12 or BB-NIC-II-HisN-CYFP-TET11 was cut with EcoRI and BspEI, and the resulting front insert was ligated into BB-NIC-II-HisN-NYFP (this plasmid was obtained by amplifying NYFP from a synthetic gene synthesized by Geneart and subcloning it into BB-NIC-II-HisN) previously cut with EcoRI and NgoMIV. All constructs were verified by DNA sequencing.

Polypeptide production and isolation. Peptides for testing the orthogonality of the building blocks were synthesized by small-scale solid-state chemical synthesis the Keck Biology Resource Lab of the Yale School of Medicine; mass spectra of the purified products were in agreement with expected values. *E. coli* BL21(DE3)pLysS cultures transformed with proper plasmids were grown at 37 °C and 160 r.p.m. in LB medium containing ampicillin (0.1 mg/ml). After induction of polypeptide production with 1 mM IPTG, incubation proceeded for additional 4 h. Cell pellets were resuspended in the lysis buffer (10 mM Tris, pH 8.0, 0.1% deoxycholate) and sonicated. Bacterial inclusion bodies were solubilized in 6 M guanidine hydrochloride (GdnHCl) and purified by chelating chromatography (Ni-NTA agarose column, Qiagen, GE) followed by size-exclusion chromatography (Agilent, USA) on Proteema 300 Å 5 µM, 8 × 300 mm column (PSS,GE). Eluted fractions that contained polypeptides were dialyzed in Milli-Q water, and the precipitate was dissolved in 6 M GdnHCl.

Polypeptide self-assembly. Purified polypeptides were diluted to a concentration of 100 nM in 6 M GdnHCl in 20 mM Tris buffer, pH 8.5, and 150 mM NaCl. Refolding proceeded by dialysis against 20 mM Tris buffer, pH 8.5, and 150 mM NaCl for 20 h at 4 °C. After dialysis, the samples were prepared for TEM and AFM characterization. Then the sample was concentrated with a 30-kDa MWCO concentrator (Amicon, USA), and the 4-µM solution was used for CD, DLS and fluorescence measurements. Slow temperature annealing was performed by cooling the sample from 90 °C to 20 °C over 20 h.

CD. CD measurements were performed on a Chirascan CD spectrometer (Applied Photophysics, Leatherhead, UK) in far-UV in a 1-mm quartz cell (Hellma, Mullheim, Germany) at 20 °C using 1-nm step, 1-nm bandwidth and 1-s sampling. The secondary structure of 4 µM polypeptide solution was analyzed by scan measurement from 200 nm to 260 nm. A chemical denaturation study was performed at different concentrations of GdnHCl from 0 M to 6 M, and the thermodynamic stability of the structures was determined from the ellipticity at 222 nm. The sample measurements were performed in triplicate, and the mean values ± s.d. were calculated.

DLS. The size of self-assembled structures in samples of 4-µM concentration was measured on a Zetasizer Nano (Malvern, UK) at 20 °C using an angle of 173° and a 633-nm laser. The size distribution of particles was recorded, and the hydrodynamic diameter was calculated using the software provided by the manufacturer.

Fluorescence measurements. Fluorescence measurements were performed with an LS55 spectrofluorometer (PerkinElmer Life Sciences, UK) with a double monochromator. All measurements were done at 20 °C in 3 mm × 3 mm quartz cell. A slit width of 5 nm was used for both excitation and emission. The samples of 4-µM concentration were excited at 485 nm, and the emission spectra from 500 nm to 600 nm were recorded.

TEM. Polypeptide solutions of 100-nM concentration were applied to holey carbon-coated copper grids (SPI, West Chester, PA, USA). The grid was placed on a droplet of sample solution (15 µL) for 2 min. Then, the grid was rinsed three times with distilled water. The uranyl staining was performed using 2% (w/v) uranyl acetate or formate by incubating the grid on a drop of stain solution for 1 min (short staining) or for extended time of 3 min (long staining). Excess solution was removed by rinsing the grid with water. Alternatively, before positive staining, 1.8-nm Ni-NTA–nanogold beads (Nanoprobes, NY, USA) were added to the samples, according to the manufacturer's instructions, to form a complex with the His_6 tag of the polypeptide. The structures were observed on a TEM microscope (JOEL JEM 2100 LaB6 microscope, Japan), operated at 180 kV. Images were acquired on a Gatan Orius CCD camera in underfocus.

AFM. A drop (10 µl) of 100-nM polypeptide solution was spotted onto a freshly cleaved mica surface and held for 15 s to achieve strong adsorption. The sample drop was then washed twice with filtered Milli-Q water and dried under a nitrogen stream. Protein samples were imaged in the acoustic alternating current mode on an Agilent Technologies 5500 scanning probe microscope using silicon cantilevers (Arrow-NCR-50) with a force constant of 42 N/m.

32. Eswar, N. *et al.* Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics* **15**, 5.6 (2006).
33. Šali, A. & Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
34. Chen, V.B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

doi:10.1038/nchembio.1248