DNA Structures

DOI: 10.1002/ange.200603392

Single-Molecule Protein Encapsulation in a Rigid DNA Cage**

*Christoph M. Erben, Russell P. Goodman, and Andrew J. Turberfield**

Herein, we demonstrate the encapsulation of a single molecule of cytochrome *c* within a rigid tetrahedral cage made of DNA. Encapsulation can be used to exert control over a guest molecule:[1] confinement within chaperone complexes promotes protein folding,[2] while synthetic cages have been used to stabilize reactive intermediates,[3] catalyze reactions,[4] and influence the conformation of peptides.[5] Containers on all scales down to the molecular level have been used to protect and target the delivery of drugs.[6] Capsules that could contain single proteins include multimeric fusion proteins,[7] viral capsids,[8] and DNA polyhedra.[9–11] DNA polyhedra are cagelike structures with edges formed from rigid double helices connected by branch junctions;[12] their dimensions are comparable to those of proteins and protein complexes. DNA tetrahedra are particularly attractive candidates for the encapsulation of other molecules: they can be made rapidly and in high yield by self-assembly, their braced architecture confers structural stability, and they can be opened by breaking bars of the cage.[11]

DNA tetrahedra were made by self-assembly of four oligonucleotides as described by Goodman et al.,[11a] and their structure is illustrated in Figure 1a. Each edge of a tetrahedron is a 20-base-pair double helix and is linked to the two edges that it meets at a vertex through unpaired single-nucleotide spacers: this double connection prevents rotation of the edge about its axis. The assembly process is stereoselective and forms the diastereomer that has the major grooves facing inwards at the vertices.[11a] Each oligonucleotide runs around one face. Four of the six edges of the tetrahedron contain nicks where the ends of an oligonucleotide meet: these nicks can be ligated or used as sites for chemical modification. We estimate that the central cavity of the tetrahedron could accommodate a sphere of radius of

[*] C. M. Erben, R. P. Goodman, Prof. A. J. Turberfield
Clarendon Laboratory
Department of Physics
University of Oxford
Parks Road, Oxford, OX1 3PU (UK)
Fax: (+44) 1865-272-400
E-mail: a.turberfield@physics.ox.ac.uk

[**] We thank Jonathan Bath for helpful discussions. This work was supported by the UK research councils BBSRC, EPSRC, MRC, and the MoD through the UK Bionanotechnology Interdisciplinary Research Collaboration, by the Rhodes Trust and the Oxford Life Sciences Interface Doctoral Training Centre.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

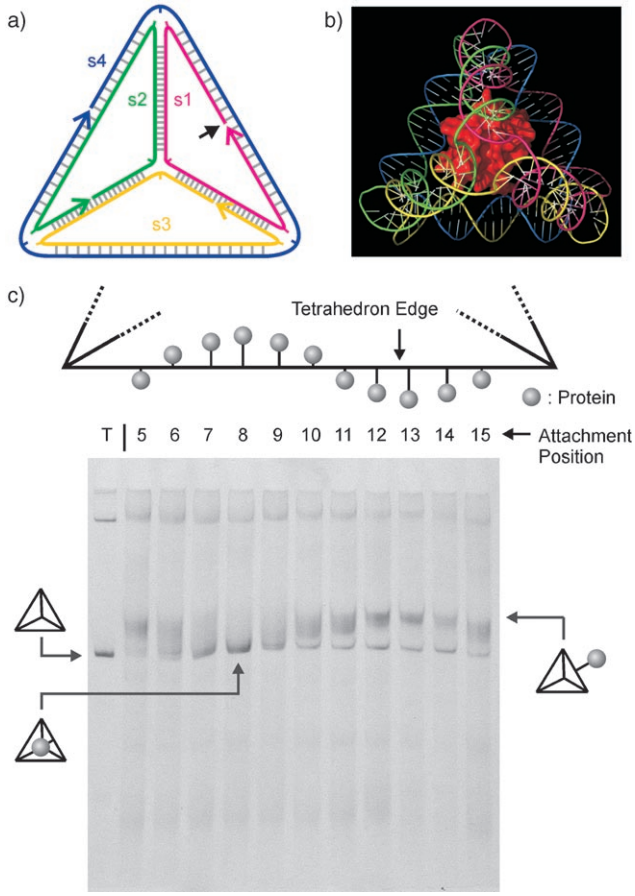

Figure 1. Positioning a protein within a DNA tetrahedron. a) Tetrahedron made from four oligonucleotides s1–s4. The black arrow indicates the protein attachment site at the 5′ end of s1. b) Molecular model which shows that cytochrome *c* can fit inside the tetrahedron. c) Native polyacrylamide gel showing the variation in electrophoretic mobility with the position of the attachment site. Control T: unmodified tetrahedron.

approximately 2.6 nm, corresponding to a globular protein with a molecular weight of roughly 60 kDa.[13] A model of the tetrahedron with a molecule of cytochrome *c* (12.4 kDa)[14] inside the cavity is shown in Figure 1b.

Horse-heart holo-cytochrome *c* was conjugated to the 5′ end of oligonucleotide s1 through a surface amine before assembly of the tetrahedron. This protein–DNA conjugate was then combined with oligonucleotides s2–s4 to form tetrahedra with the protein attached to one edge (see Experimental Section and Supporting Information). The position of the protein relative to the cage was controlled by altering the sequence of s1: when one nucleotide is transferred from the 3′ end of s1 to the 5′ end, then the attachment point of the protein—the nick in the edge formed by s1 and s4 where the ends of s1 meet—moves 0.34 nm along the double-helical edge of the tetrahedron and rotates clockwise by about 35° about its axis. As the edge is not free to rotate as a result of the double connection to neighboring edges, the attachment point of the protein on the edge determines the position of the protein relative to the tetrahedron cage. This makes it possible to choose whether the protein is held on the inside or on the outside of the tetrahedron. For the diastereomer that is selected by the assembly process,[11a] the attachment point is on the inside of the tetrahedron for conjugation at nucleotide 8 and on the outside at nucleotide 13 (counting in the 5′ to 3′ direction, with the unpaired nucleotide at the vertex as zero). The stereoselectivity of the synthesis[11a] is essential to the success of our strategy, as an attachment point that is on the inside of one diastereomer would be on the outside of the other.

Polyacrylamide gel electrophoresis (PAGE) was used to compare a group of 11 tetrahedra with cytochrome *c* conjugated at consecutive positions along one edge from the 5th to the 15th nucleotide (see Figure 1c). The control lane labeled "T" contains a tight band corresponding to a DNA tetrahedron without protein: this band is present in all other lanes (as a result of the presence of residual unconjugated oligonucleotides s1) and serves as a reference marker. A slower band corresponding to protein-conjugated tetrahedra is also visible; this band is broader, consistent with a distribution of conformations of the flexible DNA–protein linker. The mobility of the conjugate varies approximately sinusoidally with attachment position: the mobility difference between the conjugate and the unmodified tetrahedron is maximal for conjugation at nucleotide 13 and minimal for conjugation at nucleotide 8, that is, five bases further along the helix corresponding to approximately 180° rotation about the edge. When the protein is attached at nucleotide 8, the conjugate band is not separated from the nonconjugated control. The mobilities of the constructs with protein attachment at nucleotides 5 and 15, which are separated by one turn of the double helix, are approximately equal. This pattern is consistent with the expected helical trajectory of the attachment point along the edge.

To confirm this interpretation tetrahedron $T13^+$, with cytochrome *c* conjugated at nucleotide 13, the corresponding unmodified tetrahedron, $T13^-$ and $T8^+/T8^-$ (not resolved on non-denaturing gels), were purified on gel and analyzed on both non-denaturing and denaturing gels (Figure 2a and b). These tetrahedra differ from those shown in Figure 1c only in that the unmodified oligonucleotides s2–s4 had been ligated after assembly of the tetrahedron to form linked circles. Denaturation of $T13^-$ (band C) produces the expected fragments of unmodified tetrahedra: linked circles and single unmodified linear oligonucleotides (unligated s2–s4 or unmodified s1; see Supporting Information). Tetrahedron $T13^+$ (band B) contains one additional band corresponding to s1 conjugated to cytochrome *c*, thus confirming that this slow-migrating construct is the protein conjugate as identified above. Denaturing gel analysis of $T8^+/T8^-$ (band D) also reveals the presence of s1 conjugated to cytochrome *c*, thus confirming that $T8^+/T8^-$, which is unresolved from an empty tetrahedron on a non-denaturing gel, is not simply an empty tetrahedron but does contain the protein.

We conclude that when the protein is attached at nucleotide 13, it is positioned on the outside of the tetrahedron where it decreases the electrophoretic mobility of the construct. When it is attached at nucleotide 8, it has minimal effect on mobility as it is encapsulated within the tetrahedral cage.

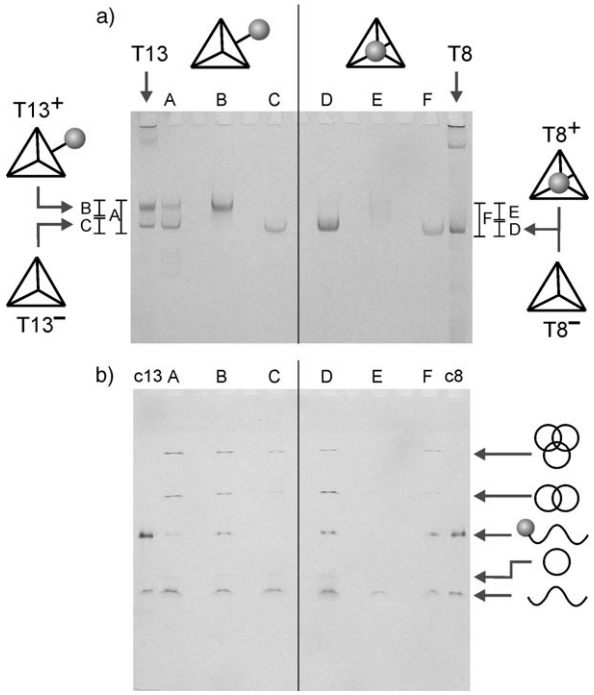

Figure 2. Analysis of protein–tetrahedron conjugates confirms the presence of the protein in $T8^+$ and $T13^+$. a) Native polyacrylamide gel showing constructs $T8^+$ and $T13^+$ with cytochrome *c* attached at the 8th and 13th nucleotides of one edge, that is, inside and outside the tetrahedron, respectively. Nonconjugated tetrahedra $T8^-$ and $T13^-$ are also present: $T8^+$ and $T8^-$ are not resolved. The outer lanes labeled T13 and T8 contain unpurified products of the synthesis; contents of lanes A–F were purified from the indicated regions of separate gels. b) Denaturing gel to analyze contents of gel-purified constructs. Controls c8, c13: cytochrome *c* conjugated to oligonucleotides s1 used to construct T8 and T13, respectively.

A DNA cage whose bars could be extended or broken might allow functional control of the encapsulated protein; for example, a cage that contains cytochrome *c* could control access by Apaf-1 (130 kDa), which forms a complex with cytochrome *c* to initiate an apoptotic protease cascade.[15] We have already demonstrated opening of the DNA tetrahedron cage by nucleases[11] and, through incorporation of a hairpin motif in one edge, by binding of a specific oligonucleotide sequence.[16] Conformation changes that could reveal the encapsulated protein could also be triggered by the following signals: binding of specific DNA[17] or RNA[18] sequences or, by incorporation of aptamer domains, proteins or small molecules;[19] changes in ionic concentrations[20] including pH;[21] ultraviolet[22] or radiofrequency radiation.[23] In contrast to conventional "caging" techniques that achieve functional control by means of photolabile protecting groups,[24] our strategy for geometric encapsulation does not require chemical modification or even knowledge of the active site: all that is necessary is a link between the encapsulated molecule and an oligonucleotide during assembly of the cage. Use of a DNA cage may also allow techniques for nonviral gene delivery[25] to be adapted to deliver the encapsulated molecule within a target cell.

In summary, we have demonstrated that a protein can be encapsulated within a DNA cage. Our assembly strategy requires control of the position of the attachment site on the cage; this was made possible by the rigid design and stereochemical purity of the DNA tetrahedron. Responsive DNA cages may allow functional control of proteins.

Experimental Section

All oligonucleotides were supplied by Integrated DNA Technologies Inc. (USA); see Supporting Information for sequences. Oligonucleotide s1 was synthesized with a C_6-amino modification at the 5′ end. Holo-cytochrome *c* from equine heart (Sigma) was conjugated to s1 by means of two heterobifunctional cross-linkers, sulfosuccinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (sulfo-SMCC) and *N*-succinimidyl 3-(2-pyridyldithio)propionate (SPDP; Pierce Biotechnology). Conjugates were gel-purified to remove proteins bound to more than one oligonucleotide. Tetrahedra were assembled by combining equal amounts of the s1–cytochrome *c* conjugate and oligonucleotides s2–s4 to a final concentration of 250 nM of each oligonucleotide in Tris buffer with 5 mM of divalent cations (either TM buffer: 10 mM Tris-HCl (pH 8), 5 mM $MgCl_2$; or TC buffer: 20 mM Tris-HCl (pH 8), 5 mM $CaCl_2$). The reaction mixture was maintained at 54 °C for 3 min then cooled to 4 °C over approximately 30 s. Tetrahedra can also be formed by incubation at room temperature, although the yield of the correctly self-assembled structure is reduced (see Supporting Information). When required, oligonucleotides s2–s4 were ligated by incubating assembled tetrahedra overnight with T4 DNA ligase (New England Biolabs) at room temperature. See Supporting Information for a detailed description of the experimental procedures used.

Received: August 18, 2006

Keywords: cytochrome *c* · DNA structures · host–guest systems · molecular devices · self-assembly

[1] J. C. Sherman, D. J. Cram, *J. Am. Chem. Soc.* **1989**, *111*, 4527–4528.
[2] F. U. Hartl, *Nature* **1996**, *381*, 571–580.
[3] D. J. Cram, M. E. Tanner, R. Thomas, *Angew. Chem.* **1991**, *103*, 1048–1051; *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1024–1027.
[4] J. M. Kang, J. Rebek, *Nature* **1997**, *385*, 50–52.
[5] S. Tashiro, M. Tominaga, Y. Yamaguchi, K. Kato, M. Fujita, *Angew. Chem.* **2006**, *118*, 247–250; *Angew. Chem. Int. Ed.* **2006**, *45*, 241–244.
[6] D. A. LaVan, T. McGuire, R. Langer, *Nat. Biotechnol.* **2003**, *21*, 1184–1191.
[7] J. E. Padilla, C. Colovos, T. O. Yeates, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2217–2221.
[8] F. P. Seebeck, K. J. Woycechowsky, W. Zhuang, J. P. Rabe, D. Hilvert, *J. Am. Chem. Soc.* **2006**, *128*, 4516–4517.
[9] a) J. H. Chen, N. C. Seeman, *Nature* **1991**, *350*, 631–633; b) Y. W. Zhang, N. C. Seeman, *J. Am. Chem. Soc.* **1994**, *116*, 1661–1669.
[10] W. M. Shih, J. D. Quispe, G. F. Joyce, *Nature* **2004**, *427*, 618–621.
[11] a) R. P. Goodman, I. A. T. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, A. J. Turberfield, *Science* **2005**, *310*, 1661–1665; b) R. P. Goodman, R. M. Berry, A. J. Turberfield, *Chem. Commun.* **2004**, 1372–1373.
[12] N. C. Seeman, *J. Theor. Biol.* **1982**, 99, 237–247.
[13] a) P. G. Squire, M. E. Himmel, *Arch. Biochem. Biophys.* **1979**, *196*, 165–177; b) K. Gekko, H. Noguchi, *J. Phys. Chem.* **1979**, *83*, 2706–2714.
[14] G. W. Bushnell, G. V. Louie, G. D. Brayer, *J. Mol. Biol.* **1990**, *214*, 585–595.

[15] P. Li, D. Nijhawan, I. Budihardjo, S. M. Srinivasula, M. Ahmad, E. S. Alnemri, X. D. Wang, *Cell* **1997**, *91*, 479 – 489.
[16] R. P. Goodman, M. Heilemann, A. N. Kapanidis, A. J. Turberfield, unpublished results.
[17] B. Yurke, A. J. Turberfield, A. P. Mills, F. C. Simmel, J. L. Neumann, *Nature* **2000**, *406*, 605 – 608.
[18] Y. Benenson, B. Gil, U. Ben-Dor, R. Adar, E. Shapiro, *Nature* **2004**, *429*, 423 – 429.
[19] J. Tang, R. R. Breaker, *Chem. Biol.* **1997**, *4*, 453 – 459.
[20] C. D. Mao, W. Q. Sun, Z. Y. Shen, N. C. Seeman, *Nature* **1999**, *397*, 144 – 146.
[21] T. Liedl, F. C. Simmel, *Nano Lett.* **2005**, *5*, 1894 – 1898.
[22] P. Ordoukhanian, J. S. Taylor, *J. Am. Chem. Soc.* **1995**, *117*, 9570 – 9571.
[23] K. Hamad-Schifferli, J. J. Schwartz, A. T. Santos, S. G. Zhang, J. M. Jacobson, *Nature* **2002**, *415*, 152 – 155.
[24] a) K. Curley, D. S. Lawrence, *Curr. Opin. Chem. Biol.* **1999**, *3*, 84 – 88; b) G. Mayer, A. Heckel, *Angew. Chem.* **2006**, *118*, 5020 – 5042; *Angew. Chem. Int. Ed.* **2006**, *45*, 4900 – 4921.
[25] X. Gao, L. Huang, *Biochem. Biophys. Res. Commun.* **1991**, *179*, 280 – 285.

