LETTER

doi:10.1038/nature10500

Self-replication of information-bearing nanoscale patterns

Tong Wang[1], Ruojie Sha[1], Rémi Dreyfus[2]†, Mirjam E. Leunissen[2]†, Corinna Maass[2], David J. Pine[2], Paul M. Chaikin[2] & Nadrian C. Seeman[1]

DNA molecules provide what is probably the most iconic example of self-replication—the ability of a system to replicate, or make copies of, itself. In living cells the process is mediated by enzymes and occurs autonomously, with the number of replicas increasing exponentially over time without the need for external manipulation. Self-replication has also been implemented with synthetic systems, including RNA enzymes designed to undergo self-sustained exponential amplification[1–5]. An exciting next step would be to use self-replication in materials fabrication, which requires robust and general systems capable of copying and amplifying functional materials or structures. Here we report a first development in this direction, using DNA tile motifs that can recognize and bind complementary tiles in a pre-programmed fashion. We first design tile motifs so they form a seven-tile seed sequence; then use the seeds to instruct the formation of a first generation of complementary seven-tile daughter sequences; and finally use the daughters to instruct the formation of seven-tile granddaughter sequences that are identical to the initial seed sequences. Considering that DNA is a functional material that can organize itself and other molecules into useful structures[6–13], our findings raise the tantalizing prospect that we may one day be able to realize self-replicating materials with various patterns or useful functions.

Nucleic acids comprise the genetic material of all living organisms, with the ordered pattern of bases in the DNA double helix readily copied by an enzymatic process that leads to semi-conservative replication inside the cell[14,15]. The replication process exploits the complementarity of the four-letter code of bases in the linear DNA double helix, where adenine (A) pairs with thymine (T) and guanine (G) pairs with cytosine (C)[16]. We transfer this concept to DNA materials, and use as an analogue of the nucleotide letter the bent triple crossover motif (BTX)[17], which displays four single strands (of seven nucleotides each) to connect with a second, similar BTX molecule. Information is encoded using these four strands, which ensure a specific interaction between two BTX molecules and therefore the transfer of information from a seed structure to later generations. As shown in Fig. 1a, the structure formed by two paired BTX letters is a paired 6-helix bundle motif (P6HB). It contains six helices connected to each other by four small double helices that extend from one duplex to another.

Once specific recognition has been set and BTX motifs have been paired on an initial seed as desired, they must be bound together linearly, analogous to the formation of a DNA backbone. This is achieved using 18 nucleotides per duplex (9 from each side of the three BTX double-helical domains) and binding them with a solution-derived DNA strand. We use the same DNA sequence in every case (except in the seed), so that any pattern can be replicated. The cohesive strength of the longitudinal bonds linking our tiles to form longer arrays is necessarily of higher melting temperature than the lateral binding between seed and daughter arrays, both because of greater length and because of end-stacking interactions[18]; this is analogous to the bases in the double helix being strongly attached (covalently bonded) in the direction parallel to the helix axis and weakly attached (hydrogen-bonded) perpendicular to it. The difference in bonding interactions allows a complementary sequence to be assembled on a seed template, internally fused, and then separated from the template to form an autonomous copy. We note that, in contrast to the DNA code, the BTX-based code is not limited to four letters: in principle, we can design as many as 4^{28} different combinations of strands involved in lateral pairing.

Experimentally, we start with the simplest case of using two different logical BTX species: an A tile, a B tile, and their complementary A′ and B′ tiles. Tiles are labelled for identification by atomic force microscopy (AFM), with either biotinylated nucleotides that bind streptavidin or

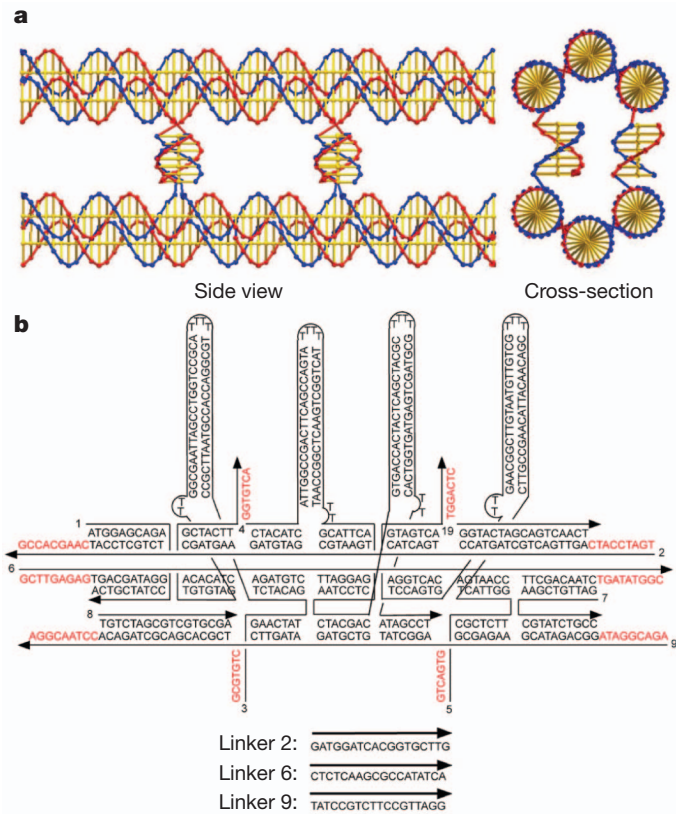

Figure 1 | DNA tile sequences and structures. **a**, The P6HB motif, drawn with GIDEON[24], comprising two BTX domains paired by four lateral connections. The cross-section view shows two of the four helices that are formed by the lateral cohesive interactions. The connections at the rear are eclipsed in this projection. **b**, Sequence and structure of the B′ BTX tile. Four helical domains, hairpins, are attached perpendicular to the BTX motif, to create a topographic feature that can be detected by AFM. Four lateral and six longitudinal sticky ends are shown in red. Other tiles are shown in Supplementary Information section 1.

[1]Department of Chemistry, New York University, New York, New York 10003, USA. [2]Center for Soft Matter Research, New York University, New York, New York 10003, USA. †Present addresses: Complex Assemblies of Soft Matter, CNRS-Rhodia-UPenn UMI 3254, Bristol, Pennsylvania 19007, USA (R.D.); FOM Institute AMOLF, Sciencepark 104, 1098 XG Amsterdam, The Netherlands (M.E.L.).

with large hairpin features (see Fig. 1b for a labelled B′ tile). We can thus read with AFM any sequence created from A and B components, and check the fidelity of our replication from the initial seed to the new generations of daughters and granddaughters. The labelling also allows us to verify whether the initial seed pattern has formed accurately. The first tile of the seed is an initiator I (an A-like tile attached to a magnetic bead), followed by a series of A and B tiles to give the sequence IBBABAB. In the seed it is the A tiles that contain labels, whereas in the complements it is the B′ tiles. Thus, using L and N to represent, respectively, the presence and absence of labels, the seed array should have the sequence LNNLNLN and the newly replicated sequence should appear as NLLNLNL. (Note, however, that any array can be labelled uniquely in any generation.) In addition to creating the seed and its replica, we have also replicated the replica and thereby obtained a direct copy of the information in the seed array itself.

The starting point for each stage in the replication process is the self-assembly of BTX tiles. (See Supplementary Information section 1 for the sequences of every tile used; section 2 for non-denaturing gels that demonstrate that the BTX tiles form properly and only associate as complementary pairs; and section 3 for electrophoretic mobility (Ferguson) analyses of the individual tiles and their P6HB complexes.)

Figure 2a shows a schematic of the seed assembly process. In step 1, the seven different seed tiles are formed separately, by annealing their constituent strands. In step 2, the tiles are mixed so they can assemble into seven-tile seed arrays, with each tile placed in the proper order as a result of specific inter-tile interactions provided by unique 16-mer sticky ends on each of its helices. In the figure, the sticky ends are labelled to indicate position within the array and binding partner; for example, Y1 on the A1 (I1) tile to Y1′ on the B2 tile, leading to an A tile at the first position of the seed and a B tile at the second position. There is a single sticky end on the left of the A1 tile, labelled S, which is used in a later step to bind to a magnetic dynabead. The A tiles contain accessible biotin groups, and in step 3 are labelled with streptavidin for AFM identification. The images in Fig. 2b illustrate the successful formation of the seeds: a number of incorrectly formed, multimeric or aggregated complexes are visible, but most of the images are straight linear arrays containing bumpy features, as we expect. The three higher-magnification images in Fig. 2b show that the target arrays have been assembled. For replication, we use unpurified seeds that have not been streptavidin-labelled.

The generation of complementary daughter arrays, schematically outlined in Fig. 3a, starts similarly to the generation of seed arrays, with

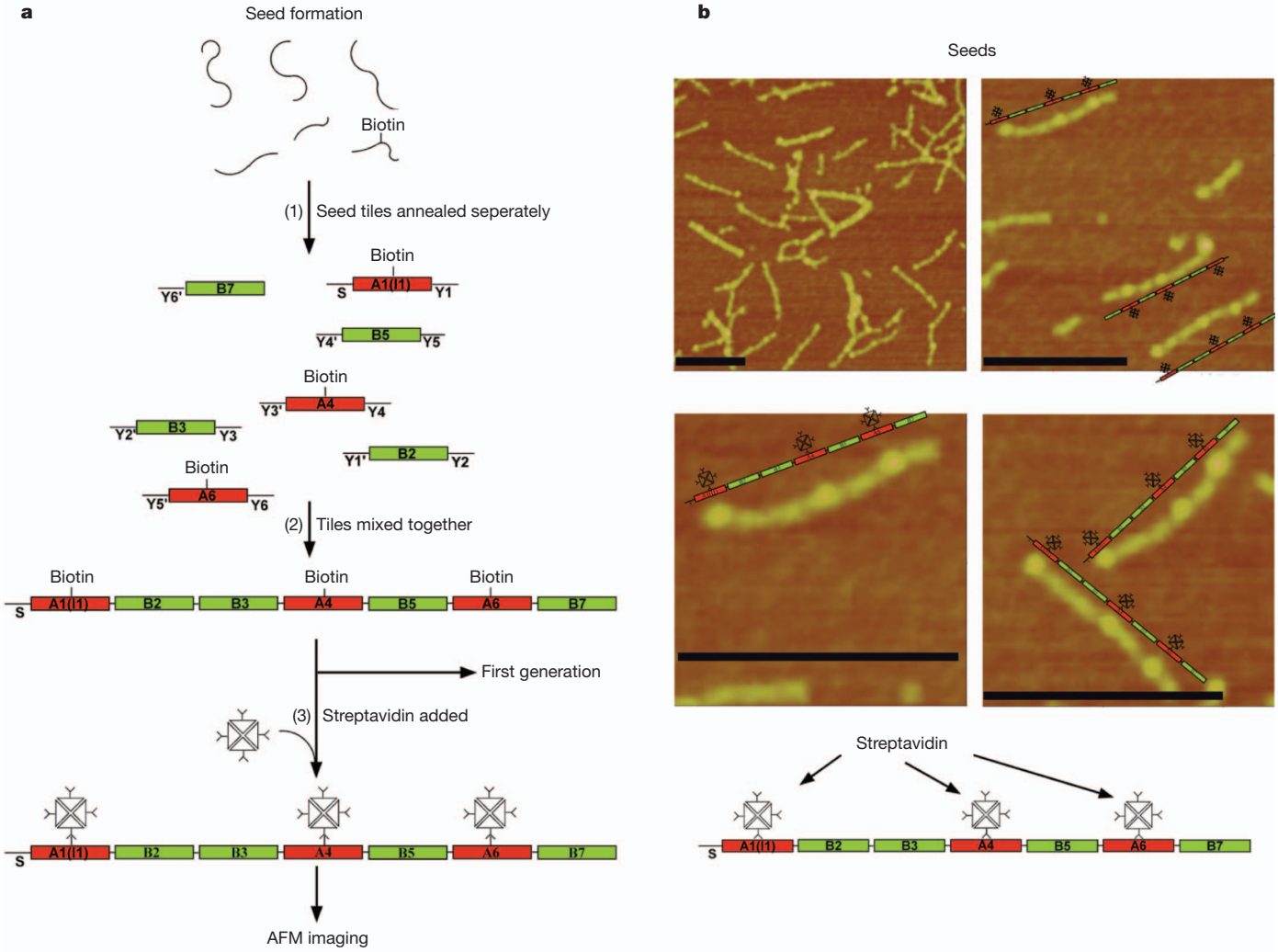

Figure 2 | DNA seeds. **a**, Seed formation. In step 1, the individual strands of the seed tiles self-assemble in separate vessels to produce seven different BTX tiles, each flanked by unique sticky ends labelled Y and a number; primed numbers are complementary to unprimed numbers. The red tiles are the A tiles and the green tiles are the B tiles. The A tiles contain a biotin group to enable their decoration by streptavidin. The tile labelled A1(I1) is the initiator tile. The strand labelled S on its left can bind to a dynabead during the replication process. In step 2, the tiles are mixed together, producing 7-unit seeds when they are mixed together. In step 3, the tiles are prepared for AFM imaging by the addition of streptavidin. **b**, AFM images of seeds. The upper left panel shows a typical field of view, of slightly less than 1 μm^2. A large number of seeds are present, along with some multimeric complexes. The other three panels are zoomed images. (Black scale bars, 200 nm.) A schematic image of each seed is shown next to the seed.

the self-assembly of A′ and B′ motifs, annealed in different vessels (step 1), and then added to a solution of seed arrays and mixed (step 2). The initiator tile of the daughter generation contains the same S end that was on the initiator tile of the seed, but this strand is protected by a cover strand[19] to keep it free of any beads. Otherwise, the tiles are all flanked by the same set of sticky-end connectors (nonamers denoted Y and Z in the figure). The B′ tiles contain a group of covalently attached hairpins as labels (Fig. 1b).

Following mixing, the BTX tiles with complementary lateral cohesive sequences are paired together to form an extended P6HB motif; the result is a complementary daughter array attached to the seed array, with the daughter tiles not yet bound to each other. At this point, a dynabead is introduced, containing a linker complementary to the S strand on the seed, so that the entire construct is attached to the bead. Following a wash, 18-mer linkers (Fig. 1b), complementary to both of the nonamer sticky ends extending between tiles, are added to the solution. These 18-mers are annealed to the construct, connecting the tiles (magnified inset following step 7 in Fig. 3a). Following a wash to remove excess linkers, seeds and daughters are separated by heating the system to 37 °C, at which temperature lateral pairs

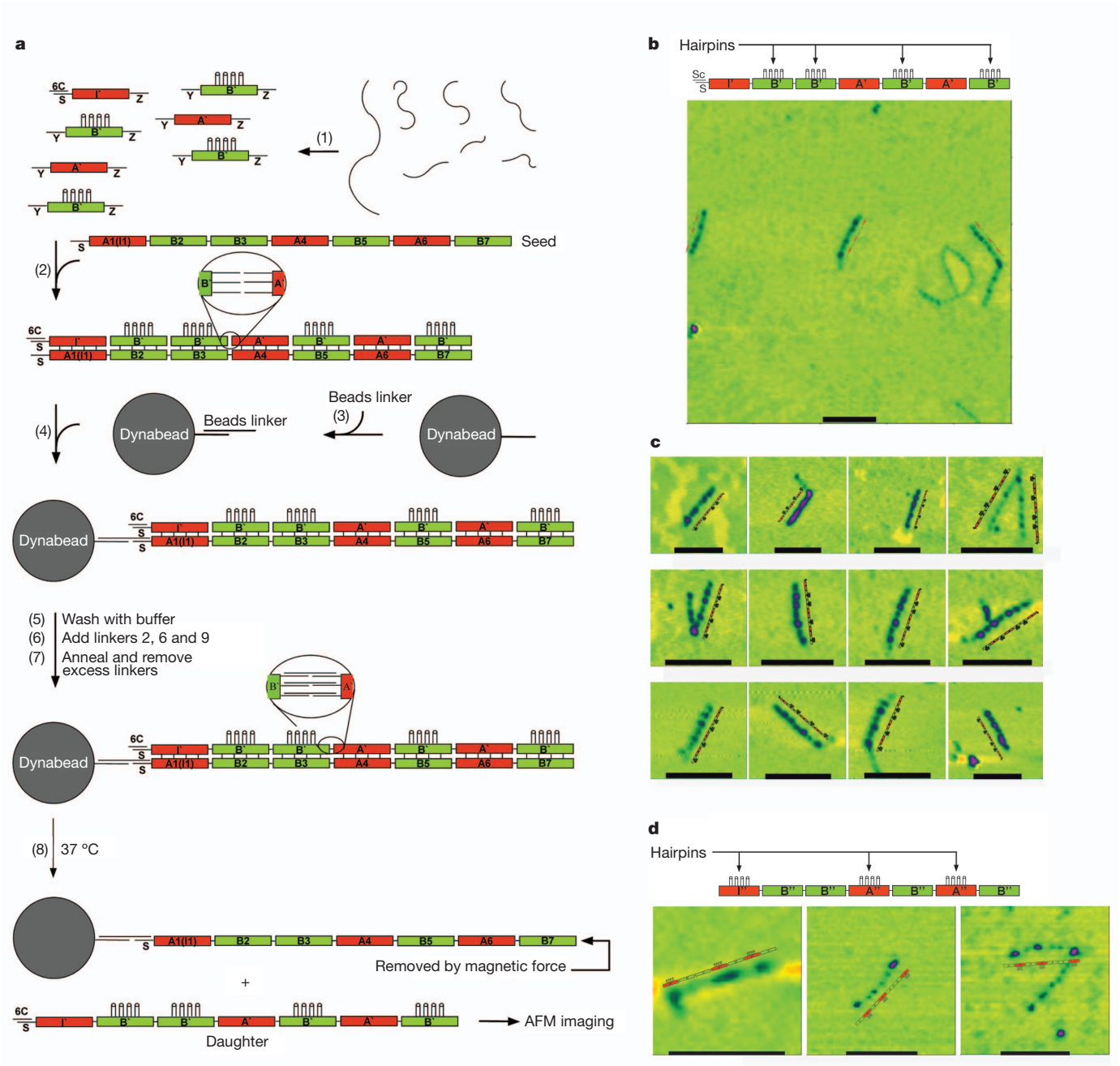

Figure 3 | DNA generations. **a,** Replication of the seven-tile seed pattern in the first generation. In step 1, strands are annealed into tiles, all flanked by the same connectors, designated Y and Z. The initiator tile (I′) contains a protected S-strand, paired with a cover strand, 6C. The B′ tiles (Fig. 1b) contain the 4-hairpin markers for AFM imaging. In the presence of the seed (step 2), the strands assemble into a pattern mimicking the seed pattern. The magnetic dynabead (large grey circle) is prepared in step 3, and attached to the seed (step 4). This is followed by a wash step, the addition of linkers, and their annealing (steps 5–7). Heating the system to 37 °C results in the separation of the daughter 7-tile complex and the seed (removed magnetically, using the dynabead). **b**, **c**, AFM images of first-generation daughter complexes, showing a typical field of view (**b**) and a collection of zoomed images (**c**). (Black scale bars, 200 nm.) These images, adjacent to explanatory schematics, demonstrate that the ABBABAB pattern has been replicated successfully. **d**, AFM images of second-generation molecules, showing the pattern that was programmed in the original seed tile.

are broken but the longitudinal bonds remain intact (step 8, bottom of Fig. 3a).

Despite the large number of manipulations required to prepare the daughters, the large-field AFM image of this system (Fig. 3b) still contains multiple copies of the molecules in an area measuring a little over 1 μm^2, along with some flawed products. Comparison with the image in Fig. 2b reveals that the daughters are significantly sparser than the seeds. The zoomed images in Fig. 3c show daughter tile arrays with four clear features (the hairpin labels), showing that the arrays appear to have formed correctly, although some of them are parts of aggregates. The percentage of heptamers with the correct sequence is estimated to be 70% (23 molecules sampled).

Granddaughters (replicas of the seed arrays) were produced directly from an initial preparation, rather than from purified seeds or purified daughters (see Supplementary Information section 4). Cover strands were removed from the initiator tile, and A″ and B″ tiles were added to the solution. In this second generation, A″ tiles contain the hairpin labels. The same sequence of steps is then performed as in the generation of the daughter strands (see Methods and Supplementary Information, section 4), with daughters and granddaughters separated by heating to 37 °C.

Figure 3d shows AFM images of granddaughters, of which there are fewer molecules and fewer cases of unambiguously correct images than seen with daughters. We estimate the percentage of second-generation heptamers with a correct tile sequence to be 31% (18 molecules sampled), corresponding to 55% ($\sim 0.31^{1/2}$) per generation, similar to the 70% seen for the daughters (See Supplementary Information section 5). We note that a net increase of molecules, and hence a growing system, would be obtained if we were to retain the original seeds in the solution containing the granddaughter molecules.

At present, our replication method is rather cumbersome, as it requires multiple chemical and thermal processing cycles; and, unlike the work of ref. 1 on RNA enzymes[1], it does not yet achieve exponential amplification. However, we have demonstrated that it is possible to replicate not just molecules such as DNA or RNA, but discrete tertiary structures that could in principle assume many different shapes and functional features. We also note that, in the same way as the first cars, airplanes and computers were clunky relative to current systems, it should be possible to make the present procedure smoother and more sophisticated. For example, yield is obviously affected by the removal of seed molecules to produce the daughters, and likewise by the removal of daughters to produce the granddaughters so as to simplify analysis. This could be avoided through elimination of bead removal steps, by using self-protected hairpins[20] or photoactive molecules for the longitudinal interactions. We expect that this and other improvements will deliver a robust replication method that is applicable to molecular, nanometre-sized and colloidal systems (such as patchy[21] and lock and key[22] particles) displaying programmed recognition.

METHODS SUMMARY

All strands were designed using the program SEQUIN[23]. Following PAGE purification, strands for the seeds, daughter and granddaughter tiles were mixed stoichiometrically as estimated by OD_{260} and dissolved to 0.5 μM in TAE/Mg^{2+} buffer (40 mM Tris-HCl, 20 mM acetic acid, 2 mM EDTA, 12.5 mM magnesium acetate, pH 8.0). The solutions were slowly annealed from 90 °C to 23 °C over 48 h in a 2-litre water bath insulated in a Styrofoam box. Stoichiometric quantities of seven seed tiles were mixed and annealed from 45 °C to 23 °C over 24 h to make seeds. To form the first generation, three first-generation tiles (I′, A′ and B′) were mixed with annealed first-generation tiles (seeds:I′:A′:B′ = 1:2:4:8), and slowly annealed from 45 °C to 23 °C Dynabeads were washed with double-distilled H_2O and TAE/Mg buffer, mixed with beads linker in TAE/Mg buffer, slowly annealed from 55 °C to 23 °C, washed with buffer, and mixed with DNA solution. The solution containing dynabeads was annealed from 33 °C to 23 °C, placed on a magnetic stand and washed with TAE/Mg buffer. Linking strands 2, 6 and 9 (Fig. 1b) were then added, the solution cooled from 33 °C to 23 °C, placed on a magnetic stand and washed with TAE/Mg buffer to remove excess linkers. Dynabeads in TAE/Mg buffer were kept at 37 °C for one hour, placed on the magnetic stand, and the solution was removed from dynabeads and stored in a clean tube for AFM imaging. Formation of the second generation is similar to the first: it starts from initial seed preparation, followed by formation of the first generation, and adding second-generation tiles (I″, A″ and B″). Steps 2–8 described in formation of the first generation were repeated.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 October 2010; accepted 18 August 2011.

1. Lincoln, T. A. & Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science* **323,** 1229–1232 (2009).
2. Wintner, E. A., Conn, M. M. & Rebek, J. Jr. Studies in molecular replication. *Acc. Chem. Res.* **27,** 198–203 (1994).
3. Schulman, R. & Winfree, E. Synthesis of crystals with a programmable kinetic barrier to nucleation. *Proc. Natl Acad. Sci. USA* **104,** 15236–15241 (2007).
4. Lin, C. *et al.* In vivo cloning of artificial DNA nanostructures. *Proc. Natl Acad. Sci. USA* **105,** 17626–17631 (2008).
5. Lee, D. H., Severin, K., Yokobayashi, Y. & Ghadiri, M. R. Emergence of symbiosis in peptide self-replication through a peptide hypercyclic network. *Nature* **390,** 591–594 (1997).
6. Seeman, N. C. DNA in a material world. *Nature* **421,** 427–431 (2003).
7. Zhu, L., Lukeman, P. S., Canary, J. W. & Seeman, N. C. Nylon/DNA: single-stranded DNA with covalently stitched nylon lining. *J. Am. Chem. Soc.* **125,** 10178–10179 (2003).
8. Zheng, J. *et al.* 2D nanoparticle arrays show the organizational power of robust DNA motifs. *Nano Lett.* **6,** 1502–1504 (2006).
9. Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. DNA-guided crystallization of colloidal nanoparticles. *Nature* **451,** 549–552 (2008).
10. Park, S. Y. *et al.* DNA-programmable nanoparticle crystallization. *Nature* **451,** 553–556 (2008).
11. Kim, A. J., Biancaniello, P. L. & Crocker, J. C. Engineering DNA-mediated colloidal crystallization. *Langmuir* **22,** 1991–2001 (2006).
12. Valignat, M. P., Theodoly, O., Crocker, J. C., Russel, W. B. & Chaikin, P. M. Reversible self-assembly and directed assembly of DNA-linked micrometer-sized colloids. *Proc. Natl Acad. Sci. USA* **102,** 4225–4229 (2005).
13. Leunissen, M. E. *et al.* Towards self-replicating materials of DNA-functionalized colloids. *Soft Matter* **5,** 2422–2430 (2009).
14. Watson, J. D. & Crick, F. H. C. The genetical implications of the structure of deoxyribonucleic acid. *Nature* **171,** 964–967 (1953).
15. Lehman, I. R., Bessman, M. J., Simms, E. S. & Kornberg, A. Enzymatic synthesis of deoxyribonucleic acids 1. *J. Biol. Chem.* **233,** 163–170 (1958).
16. Watson, J. D. & Crick, F. H. C. The molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171,** 737–738 (1953).
17. Kuzuya, A., Wang, R., Sha, R. & Seeman, N. C. Six-helix and eight-helix DNA nanotubes assembled from half-tubes. *Nano Lett.* **7,** 1757–1763 (2007).
18. Wang, R., Kuzuya, A., Liu, W. & Seeman, N. C. Blunt-ended DNA stacking interactions in a 3-helix motif. *Chem. Commun.* **46,** 4905–4907 (2010).
19. Zhong, H. & Seeman, N. C. RNA used to control a DNA rotary machine. *Nano Lett.* **6,** 2899–2903 (2006).
20. Leunissen, M. E. *et al.* Switchable self-protected attractions in DNA-functionalized colloids. *Nature Mater.* **8,** 590–595 (2009).
21. Zhang, Z. & Glotzer, S. C. Self-assembly of patchy particles. *Nano Lett.* **4,** 1407–1413 (2004).
22. Sacanna, S., Irvine, W. T. M., Chaikin, P. M. & Pine, D. J. Lock and key colloids. *Nature* **464,** 575–578 (2010).
23. Seeman, N. C. *De novo* design of sequences for nucleic acid structure engineering. *J. Biomol. Struct. Dyn.* **8,** 573–581 (1990).
24. Birac, J. J., Sherman, W. B., Kopatsch, J., Constantinou, P. E. & Seeman, N. C. GIDEON, A program for design in structural DNA nanotechnology. *J. Mol. Graph. Model.* **25,** 470–480 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was partially supported by the W.M. Keck Foundation (P.M.C., D.J.P., N.C.S.), the National Science Foundation (MRSEC Program award DMR-0820341, P.M.C., D.J.P. and grants CTS-0608889 and CCF-0726378, to N.C.S.), NASA (NNX08AK04G, to P.M.C.), the National Institute of General Medical Sciences (grant GM-29554, to N.C.S.), the Army Research Office (grants 48681-EL and W911NF-07-1-0439, to N.C.S.), the Office of Naval Research (grants N000140910181 and N000140911118, to N.C.S.), the Netherlands Organization for Scientific Research (Rubicon grant, to M.E.L.) and DAAD (post-doctoral grant, to C.M.).

Author Contributions T.W., R.S. and C.M. designed experiments, performed experiments, analysed data and wrote the paper; R.D. analysed data and wrote the paper; M.E.L. initiated the project, analysed data and wrote the paper; D.J.P. and P.M.C. initiated and directed the project, analysed data and wrote the paper; N.C.S. initiated and directed the project, designed experiments, analysed data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.M.C. (chaikin@nyu.edu) or N.C.S. (ned.seeman@nyu.edu).

METHODS

DNA strand design, synthesis, and purification. DNA sequences were designed using the program SEQUIN[23]. The DNA strands with modifications—for example, biotinylated—were synthesized on an Applied Biosystems 394, removed from the support, and deprotected using routine phosphoramidite procedures. Other strands were purchased from Integrated DNA Technologies, Inc. All strands were purified by denaturing gel electrophoresis (PAGE); bands were cut out of 15–20% denaturing gels and eluted in a solution containing 500 mM ammonium acetate, 11 mM magnesium acetate and 1 mM EDTA. This is the only purification step in the procedure.

Formation of hydrogen-bonded BTX complexes. The strands of each tile were mixed stoichiometrically, as estimated by OD_{260}, and dissolved to 0.5 μM in TAE/Mg^{2+} buffer (40 mM Tris-HCl, 20 mM acetic acid, 2 mM EDTA, 12.5 mM magnesium acetate, pH 8.0). The solutions were slowly annealed from 90 °C to 23 °C over 48 h in a 2-l water bath insulated in a Styrofoam box.

Formation of self-replication seeds. Seven individual seed tiles were prepared using the protocol described above and were mixed stoichiometrically. The mixed solution was slowly annealed from 45 °C to room temperature over 24 h in a 2-l water bath insulated in a Styrofoam box. Seeds (seven tiles in length, in specific sequence) with biotin and without biotin were prepared separately in the same conditions. Seeds with biotin were used for AFM imaging only; seeds without biotin were used as starting material in the first step of self-replication. Before AFM imaging, seeds with biotin were incubated with streptavidin (biotin:streptavidin = 1:1) at room temperature for at least 6 h.

First step of self-replication (formation of first generation). (1) Three first-generation tiles (I′, A′ and B′) were prepared as described above. (2) Annealed seeds were mixed with annealed first-generation tiles (seeds:I′:A′:B′ = 1:2:4:8), slowly annealed from 45 °C to room temperature over 24 h in a 2-l water bath insulated in a Styrofoam box. (3) Dynabeads (Invitrogen) were washed with ddH_2O and TAE/Mg buffer, mixed with beads linker in TAE/Mg buffer, slowly annealed from 55 °C to room temperature at 5 °C h^{-1} on a rotator in a programmable incubator, and mixed with DNA solution prepared in step (2). (4) The above solution containing dynabeads was annealed from 33 °C to 23 °C at 1 °C h^{-1} on a rotator in a programmable incubator. (5) Solution from (4) was placed on a magnetic stand for 2 min and washed with TAE/Mg buffer. (6) Linkers 2, 6 and 9 were added. (7) Solution was cooled from 33 °C to 23 °C at 1 °C h^{-1} on a rotator in a programmable incubator, placed on a magnetic stand for 2 min, and washed with TAE/Mg buffer to eliminate excess linkers. (8) Dynabeads in TAE/Mg buffer were kept at 37 °C for 1 h, then placed on a magnetic stand for 2 min at 37 °C. The solution was removed from dynabeads and stored in a clean tube for AFM imaging.

Second step of self-replication (formation of second generation). Formation of the second generation starts from initial seed preparation, followed by formation of the first generation, but with steps (6) and (7), described above, performed before step (5). (1) Three second-generation tiles (I″, A″ and B″) were prepared as described above. A strand fully complementary to 6C was added to the solution of first-step self-replication and incubated at 25 °C for 2 h. (2) Steps (2)–(8) were performed as described for formation of first generation, except with steps (6) and (7) preceding step (5).

AFM imaging. All AFM imaging was performed in tapping-mode AFM in air. 5–7 μl of DNA sample was spotted on freshly cleaved mica (Ted Pella, Inc.) and was left for 1 min to be absorbed. Mica was washed with 3–5 drops of ddH_2O three times, and excess water was removed by blotting the mica with filter paper. The mica was then blown dry using compressed air. All AFM imaging was performed on a NanoScope IV MultiMode SPM (Digital Instruments), using commercial cantilevers with silicon tips (Veecoprobes).