

# Cryo-EM structure of a 3D DNA-origami object

Xiao-chen Bai<sup>a</sup>, Thomas G. Martin<sup>b</sup>, Sjors H. W. Scheres<sup>a,1</sup>, and Hendrik Dietz<sup>b,1</sup>

<sup>a</sup>Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom; and <sup>b</sup>Physics Department, Walter Schottky Institute, Technische Universität München, 85748 Garching near Munich, Germany

Edited by David DeRosier, Brandeis University, Waltham, MA, and approved October 18, 2012 (received for review September 10, 2012)

A key goal for nanotechnology is to design synthetic objects that may ultimately achieve functionalities known today only from natural macromolecular complexes. Molecular self-assembly with DNA has shown potential for creating user-defined 3D scaffolds, but the level of attainable positional accuracy has been unclear. Here we report the cryo-EM structure and a full pseudoatomic model of a discrete DNA object that is almost twice the size of a prokaryotic ribosome. The structure provides a variety of stable, previously undescribed DNA topologies for future use in nanotechnology and experimental evidence that discrete 3D DNA scaffolds allow the positioning of user-defined structural motifs with an accuracy that is similar to that observed in natural macromolecules. Thereby, our results indicate an attractive route to fabricate nanoscale devices that achieve complex functionalities by DNA-templated design steered by structural feedback.

Natural macromolecular machines have complex 3D shapes with subnanometer-precise structural features that enable executing tasks such as signal transduction, molecular transport, and enzymatic catalysis (1). A key goal for nanotechnology is to fabricate synthetic objects with similarly precise features to ultimately control tasks known today only from natural “nanomachines.” Molecular self-assembly with DNA is considered a candidate route to achieve this goal (2–13). Densely packed 3D DNA-origami objects (14–18) seem particularly suited for use as rigid scaffolds to position reactive groups at target locations in space, and to implement features known from natural macromolecular complexes, such as shape complementarity and controlled domain movement. Designing such objects to meet precise structural specifications will benefit strongly from detailed 3D structural feedback. Here we present a cryo-EM map of an asymmetric, densely packed DNA object that comprises 15,238 nt in 164 chains. Combined with prior knowledge about the topology of chain connectivity, this map provided sufficient detail to construct a full pseudoatomic model for this particle. Thereby, our results provide structural feedback on DNA-templated design that will prove valuable for the design of complex functionalities.

## Results and Discussion

**DNA-Templated Design and Synthesis.** Our DNA object was designed to be suitable for structure determination by cryo-EM single-particle analysis, because the choice of a distinctly asymmetric shape facilitated the recognition of different particle orientations in electron micrographs (Fig. 1A). The object was designed to assemble as 82 parallel dsDNA helices of varying length in a square lattice (17) of 10 columns and 12 rows and comprising a total of 15,238 nt with a molecular mass of 4.8 MDa (Fig. S1). A total of 740 nt were distributed among flexible loops at the ends of the helices to prevent aggregation by blunt-end stacking (19). Synthesis of the object was based on templated molecular self-assembly (3) with a 7,249-nt-long “scaffold” DNA strand derived from M13 bacteriophage and 163 shorter “staple” strands (Table S1). The object formed overnight in a one-pot reaction with high yield and few byproducts, such as higher-molecular-weight aggregates. It was purified from excess staple strands by molecular-weight cutoff filtration, which yielded particle solutions free from nucleic acid stains and gel matrix residues (Fig. S2).

**Cryo-EM Single-Particle Analysis and Model Generation.** Imaging of the sample by cryo-EM revealed monodisperse particles with the expected dimensions and shape (Fig. S3). The particles showed suitable contrast in ice and adopted multiple orientations on the supporting continuous carbon film (Figs. 1B and C). This allowed us to calculate a 3D reconstruction (20) from 28,502 individual particles with an overall resolution of 11.5 Å (Fig. S4). The resolution of the map varies with position, with an estimated range of 9.7 Å at the core to 14 Å at the periphery (Fig. 1D; *Materials and Methods*). With the exception of a region comprising helices in row 11 at the periphery of the object and a single cross-over in row 10, all helices and connecting cross-overs are well resolved in the map. The global shape of the reconstructed density is in agreement with the designed topology of the object. The overall dimensions of the object and the extent of global twist that results from a mismatch of B-form DNA helicity and the imposed square-lattice packing also agree with a prediction (21, 22) when using 2.6-nm effective helix diameter and 10.44 bp/turn reciprocal twist density (17) (Fig. S5).

Aided by knowledge of the designed topology of the 164 strands in the object, we constructed an initial atomic model that consisted of canonical B-form DNA segments (23) for all expected dsDNA helices and cross-overs. Flexible fitting of this model inside the cryo-EM density (24) yielded a pseudoatomic model for the entire object (Fig. 2 and Movie S1), except for the single-stranded loops at the boundaries, for which no density was observed. Throughout the object, the fitted model is in good agreement with the cryo-EM density, and nearly all residues form expected base pairs. Note that in the EM density, connections between neighboring helices are often visible at positions that do not coincide with covalent strand cross-over (Fig. 3A–C), suggesting additional contacts beyond the two covalent strand cross-overs. However, low-pass filtering of our atomic model to the same resolution as the EM map shows similar contacts (Fig. S6), indicating that the additional connections are due to the limited resolution of our map. The quality of the pseudoatomic model is highest at the core of the object, where it reproduces the observed major and minor grooves resolved in the map (Fig. 3B). According to the frequency-dependent decay of the power in our EM reconstruction (25), the rmsd of the atoms at the core of the object was estimated to be in the range of 2–3 Å (*SI Text*), thus providing experimental evidence that structural order that approaches that of natural nanomachines may be attained in discrete DNA objects.

Author contributions: S.H.W.S. and H.D. designed research; X.-c.B. and T.G.M. performed research; X.-c.B., T.G.M., S.H.W.S., and H.D. analyzed data; and S.H.W.S. and H.D. wrote the paper.

The authors declare no conflict of interest.

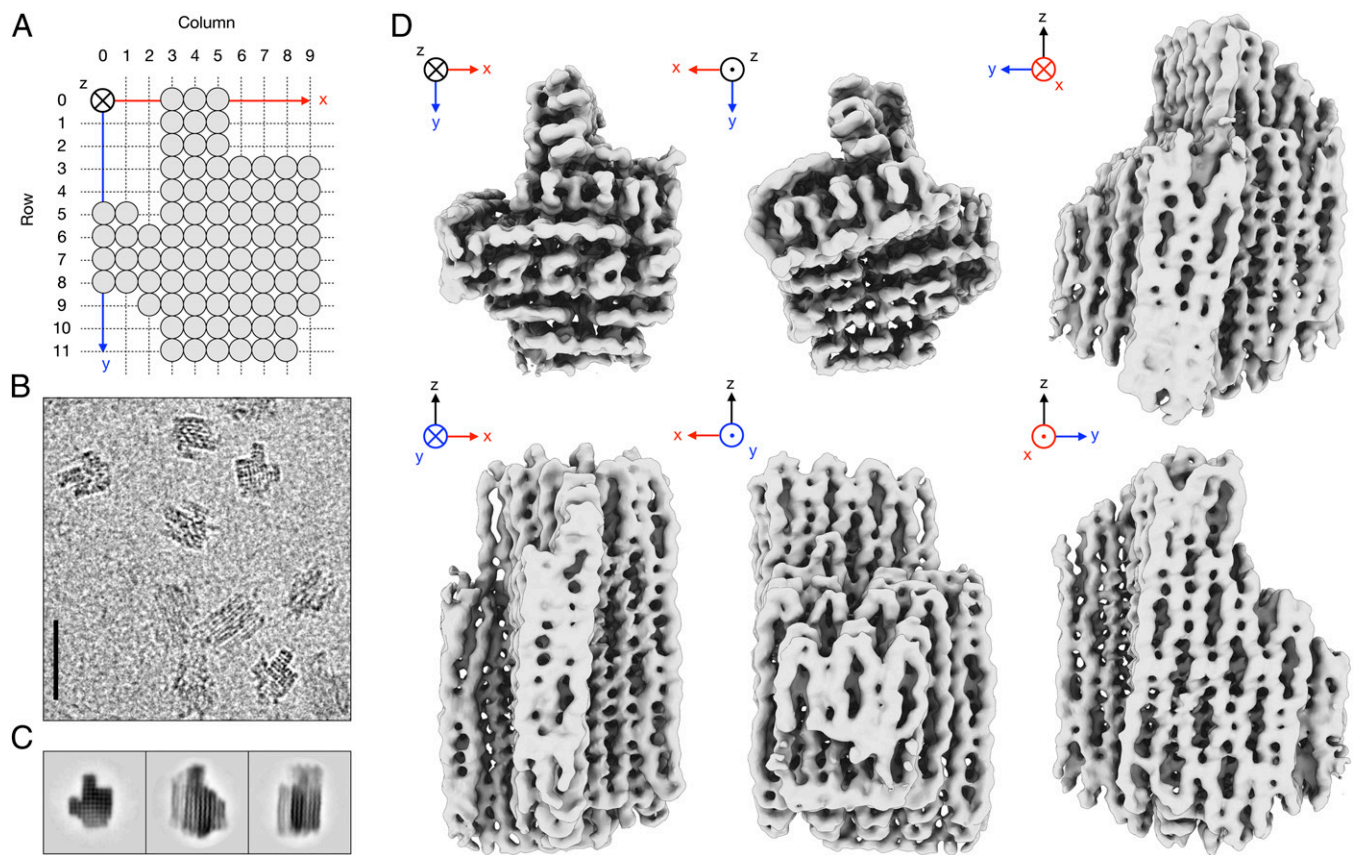
This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The cryo-EM map and the atomic coordinates have been deposited in the Electron Microscopy Data Bank (EMDB accession no. 2210) and in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) (PDB ID codes: 2ymf–2ymi and 2ymr).

<sup>1</sup>To whom correspondence may be addressed. E-mail: [scheres@mrc-lmb.cam.ac.uk](mailto:scheres@mrc-lmb.cam.ac.uk) or [dietz@tum.de](mailto:dietz@tum.de).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1215713109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1215713109/-DCSupplemental).



**Fig. 1.** Cryo-EM reconstruction of a designed, densely packed DNA object. (A) Schematic representation of the designed rectangular lattice comprising 82 parallel dsDNA helices (gray circles). (B) Representative part of an electron micrograph. (Scale bar, 50 nm.) (C) Reference-free 2D class averages. (D) Six orthogonal views of the 3D reconstruction, shown as iso-density surface at density level 0.1.

**Analysis of Structural Features.** The pseudoatomic model enables the type of geometrical analysis required for designing objects with precise structural specifications. For example, in agreement with default B-form DNA geometry, we find that the average distance from base pair to base pair midpoints is 3.35 Å as derived from the helical distance between consecutive cross-overs that are resolved in the cryo-EM density. The design of the object assumes parallel double helices packed on a square lattice. Although the cross-sectional square lattice is indeed realized, we find that the helices are not parallel: on either side of the majority of cross-overs they enclose a nonzero angle within, but also out of the plane normal to the cross-over direction (Figs. 1D and 3A–D). Globally, this results in a 3D chickenwire-like pattern whereby individual dsDNA helices form diamond-shaped cavities in between cross-overs (Fig. 3E and F). The 3D chickenwire pattern is also correlated with an unusual geometry of the 377 Holliday junctions that connect the helices in the object (Fig. 3G and H). As established by FRET (26, 27) and atomic force microscopy (AFM) measurements (28), Holliday junctions adopt stacked conformations in the presence of cations where the two opposite helices are at an angle  $\gamma$  of  $\sim 60^\circ$ . In crystal structures, Holliday junctions show similar features (29–31). However, both the in-plane and the out-of-plane bending that we find in our structure is absent in the crystal structures and was also not resolved in the FRET and AFM measurements. Moreover, the junctions in our object exhibit a left-handed, rather than a right-handed, interhelical conformation (Fig. S7). These large conformational differences suggest that Holliday junctions are easily deformed by any ordering scheme, be it through crystal contacts or scaffolding interactions inside a densely packed

DNA object. In the light of these results, future design of DNA objects to more precise structural specifications will benefit from the consideration of more appropriate cross-over models (Fig. 3I and J).

We designed the termini of the 163 linear strands in the object to be mainly located 2 or 3 base pairs away from cross-overs. The termini result in covalent phosphodiester backbone nicks in dsDNA helices. We hypothesized that Holliday junction cross-overs with such nearby nicks might not fully form or that nicked helices might differ in shape from nonnicked helices. Our map reveals that this is not the case. Cross-overs with nearby nicks do not differ in geometry from nonnicked cross-overs (Fig. 3H). We also did not observe systematic differences when comparing nicked and nonnicked dsDNA helical stretches in the map. In addition, we investigated whether the nature of the local sequences at DNA cross-over positions leads to discernible differences in cross-over shape but could find no evidence for this.

**Previously Unobserved DNA Topologies.** The cryo-EM map also provides structural insights into a library of highly unusual DNA topologies, some of which can only exist because of the scaffolding provided by the densely packed design. In row 4, columns 4–9 of the object, we created a vertical stack of five Holliday junctions (Fig. 4A; Fig. 3A shows a vertical stack of three Holliday junctions). This motif is stabilized only by the stacking of base pairs at the junctions in addition to two covalent phosphodiester bonds at either end of the junction stack. In our map this motif is well resolved, thus pointing at a surprisingly high degree of structural order, contrary to what one might have intuited. We speculate that the systematic use of many of these junction stacks in future



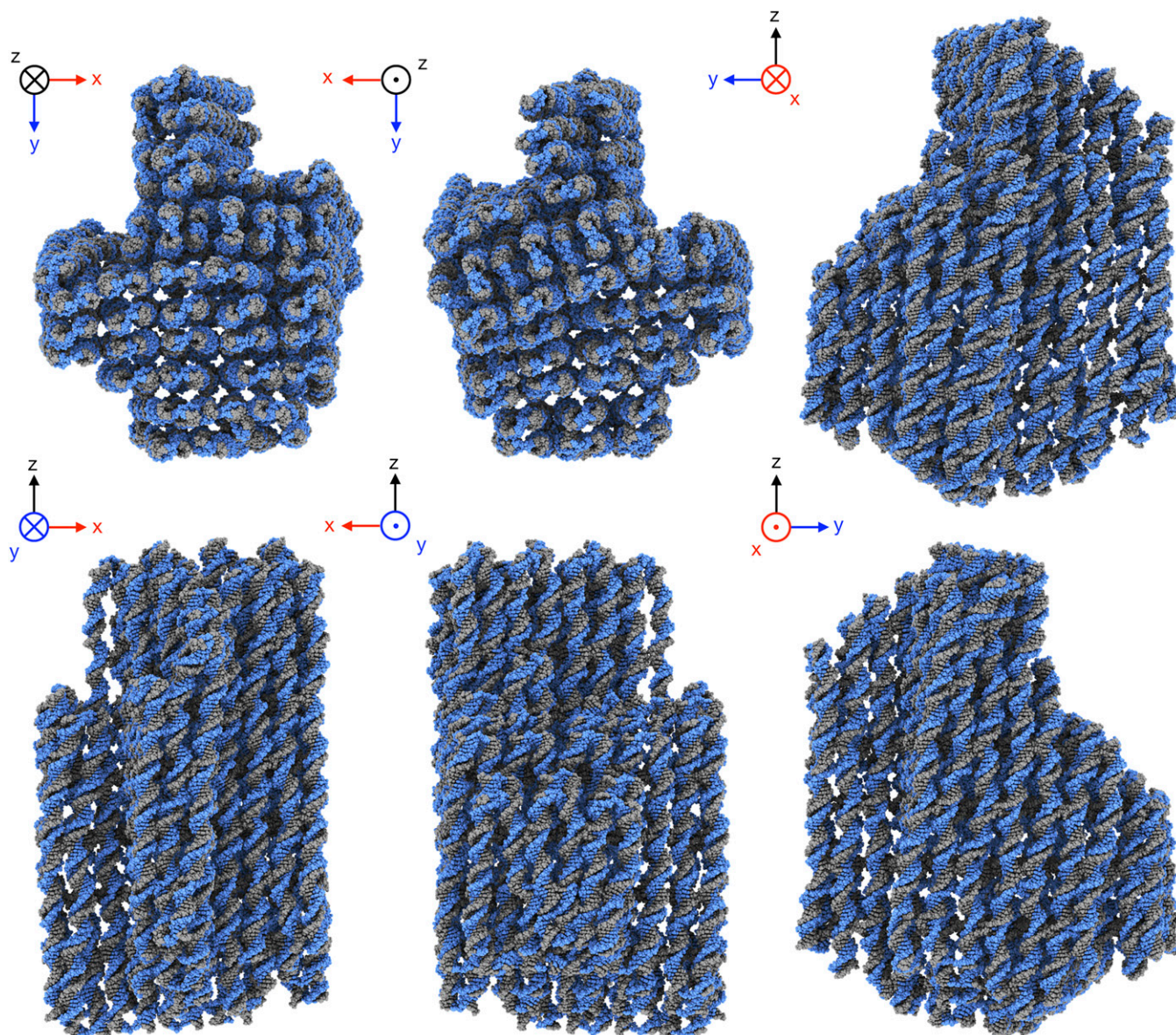


Fig. 2. Pseudoatomic model. Six orthogonal views of the pseudoatomic model that was fitted into the EM density map.

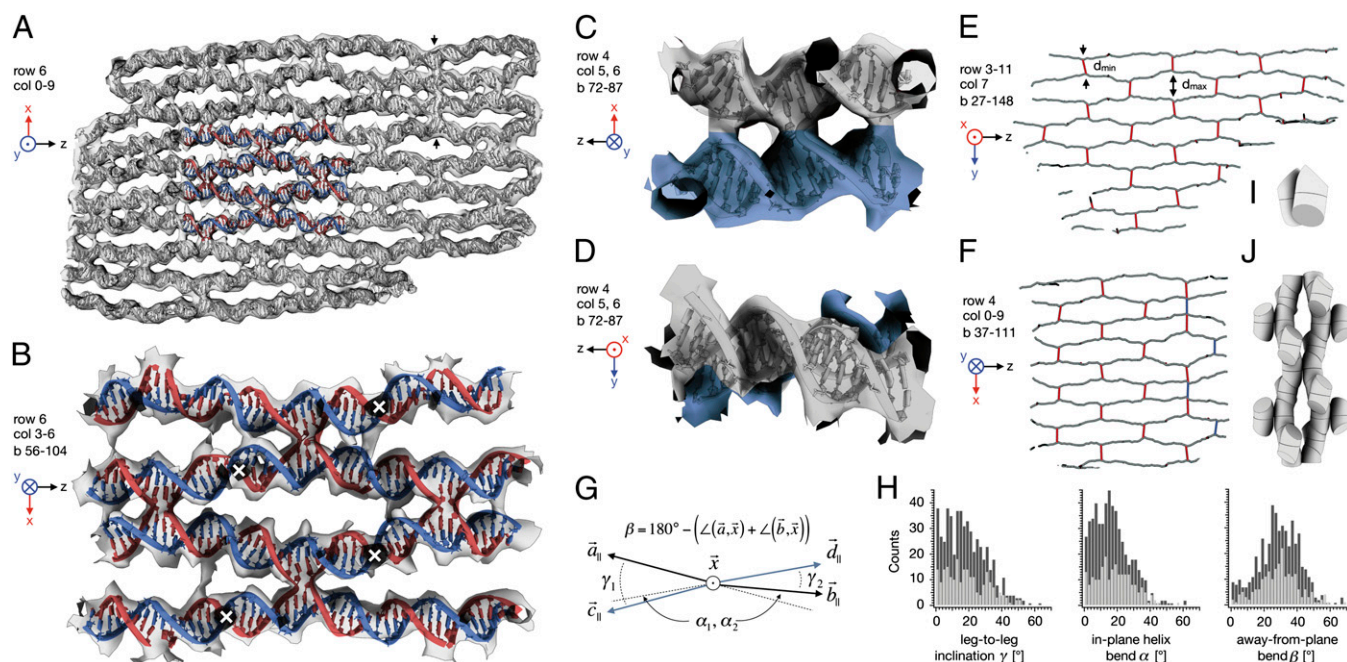
designs will lead to compaction and confer increased mechanical rigidity. In row 6, columns 1–6, we created a motif (Fig. 4B) that may be considered as a synthetic pseudohelix in which one strand follows a left-handed helical path with base pairs pointing along the helical direction, rather than orthogonally to the helical axis as in B-form dsDNA. The left-handed pseudohelix is stabilized by six short counter-strands of only 5- or 6-nt length. Again, the motif is well resolved in the EM density, albeit with less resolution because this motif occurs at the more flexible periphery of the object. In row 7, column 6, we omitted a single DNA base pair, which leads to a mismatch in backbone orientation at a cross-over to a neighboring helix in column 7. Previously, the introduction of many such omissions was used to design globally curved structures (15). In our map, a single base pair omission already results in a well-resolved bending deformation of both helices and a cross-over direction rotation (Fig. 4C), confirming that fine-tuning at the level of individual base pairs may afford precise control over orientation and location of structural features. Finally, column 0 is illustrative for how custom crevices may be created by rationally distributing cross-

overs, because dsDNA helices tend to splay away from each other in the absence of stabilizing cross-overs between them (Fig. 4D).

### Conclusions

Our results demonstrate that designed, densely packed DNA objects are amenable to detailed structural characterization and may provide near atomic-level positional control. The range of topological motifs in our structure illustrates the versatility of DNA as a material for nanotechnology and highlights the existence of many unexplored design options for creating richer, more complex, but also more precise objects. Thereby, our results support a perspective in which chemical motifs may be arranged with precise structural specifications through an iterative strategy of DNA-templated design and 3D structural feedback. By using chemical groups attached to DNA strands or even reactive motifs formed by DNA itself, this strategy offers an attractive route to achieving complex functionalities known today only from natural nanomachines.





**Fig. 3.** Analysis of internal geometry. (A) Central slice through the object at row 6, showing the cryo-EM density map (transparent gray) and the fitted pseudoatomic model as ribbon/slab model. For part of the model, the scaffold is shown in blue and the staples in red. Vertical arrows indicate a vertical stack of three cross-overs. (B) Close-up of colored area in A, but rotated around the z axis by  $180^\circ$ . White crosses indicate out-of-plane cross-over positions. (C) Side view of a single cross-over. (D) Top view of the same cross-over as in C. (E and F) Schematic representation of the 3D chickenwire-like pattern found in the structure, depicting dsDNA helical stretches in gray and cross-overs in red. The pattern was computed using the coordinates of base pair midpoints in the pseudoatomic model. The midpoints of neighboring dsDNA helices move on average from a minimum distance  $\langle d_{\min} \rangle = 18.5 \text{ \AA}$  at the cross-over to a maximum distance of  $\langle d_{\max} \rangle = 36 \text{ \AA}$  away from each other. Cross-overs marked in blue indicate scaffold-based strand cross-overs. (G) Definition of the angles enclosed by the four helical legs of a cross-over. Vectors are computed using the coordinates of base pair midpoints at the cross-over position and 2 bp away from the cross-over in each leg. The cross-over vector  $\vec{x}$  is computed from the coordinates of the midpoints between the two base pairs in each of the two helices at the cross-over position and is normal to what we call the cross-over plane. The subscript “||” indicates vectorial projections into the cross-over plane. The angle  $\beta$  is also computed as indicated for vectors C and D. (H) Observed distribution of angles between the legs for cross-overs with (light gray) and without (dark gray) nearby nicks for 377 cross-overs in the structure, where each cross-over contributed two  $\gamma$ ,  $\alpha$ , and  $\beta$  angle values each. (I) Schematic representation of a revised cross-over model for DNA nanostructure design. (J) 3D chickenwire-like pattern formed by multiple instances of the revised cross-over scheme when rotated and translated according to the square-lattice packing connectivity scheme.

## Materials and Methods

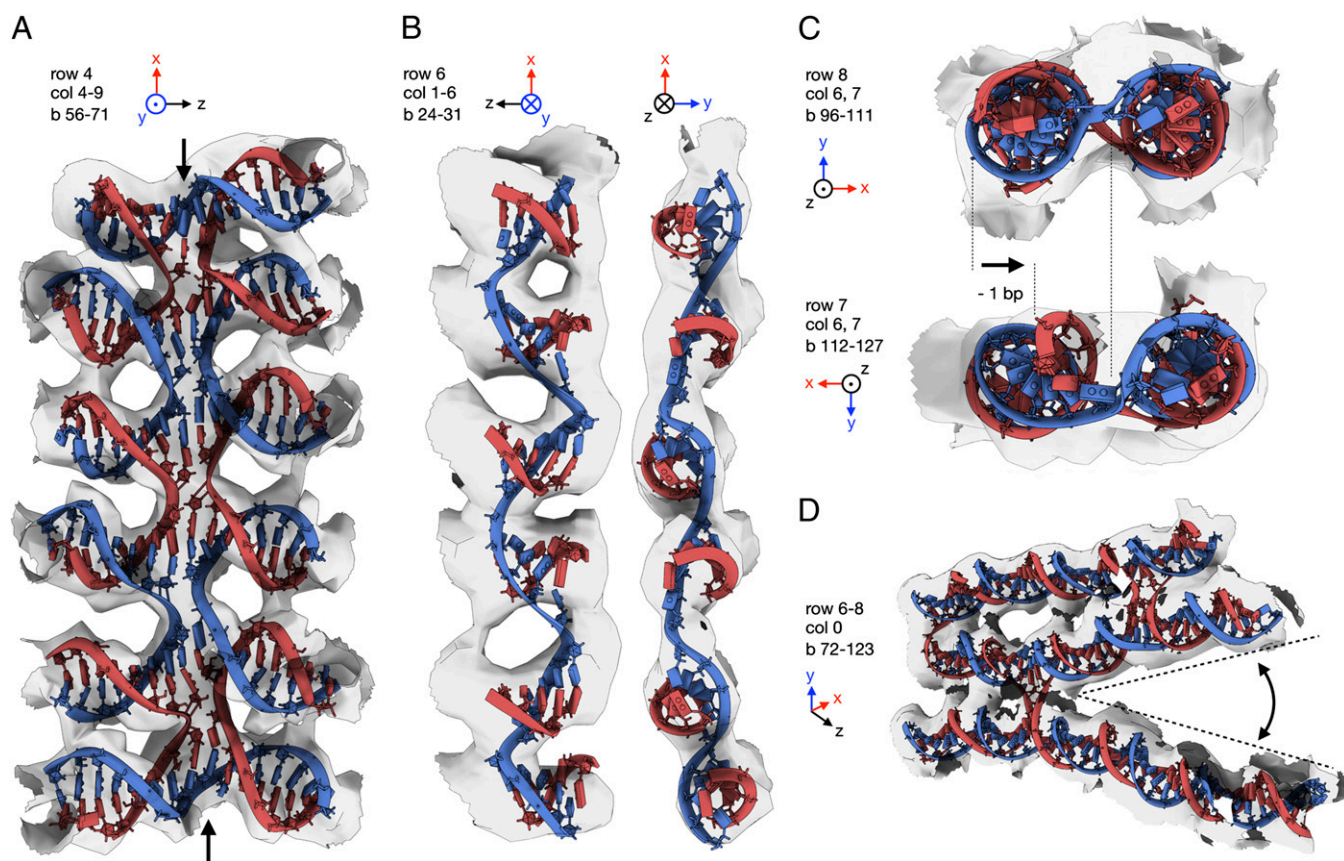
**DNA-Templated Design and Synthesis.** Our object was designed in iterative cycles of using caDNAo v0.2 (16) for strand routing and CanDo (21, 22) for estimation of rigidity and global twist. DNA scaffold strands were prepared as previously described (32). DNA staple oligonucleotide strands (Table S1) were prepared by solid-phase chemical synthesis (Eurofins MWG) with Eurofins MWG high purity salt free purification grade. The objects were synthesized in a one-pot mixture containing 20 nM 7,249-bases-long M13mp18 phage DNA, 200 nM oligonucleotide staple in a pH = 8 buffer that included 5 mM Tris-base, 1 mM EDTA, 20 mM MgCl<sub>2</sub>, and 5 mM NaCl. The mixture was incubated at  $65^\circ\text{C}$  for 15 min, then annealed from  $56^\circ\text{C}$  to  $44^\circ\text{C}$  over the course of 12 h, and then stored at  $4^\circ\text{C}$ . Analysis of the reaction products by agarose gel electrophoresis (Fig. S2) showed that the object assembled with high yield with negligible formation of byproducts. Purification from excess staple strands could therefore be performed using simple molecular weight cutoff filtration using 100 kDa Amicon filters (Millipore).

**Electron Cryo-Microscopy.** The purified sample at a concentration of  $\sim 10 \text{ nM}$  in 10 mM Tris buffer (pH 7.6) with 20 mM MgCl<sub>2</sub>, 5 mM NaCl, and 1 mM EDTA was heated in a water bath at  $37^\circ\text{C}$  for 1 h before preparing the grids to resolve multimers formed by blunt-end stacking. Aliquots of  $3 \mu\text{L}$  were incubated for 3 min on glow-discharged holey carbon grids with an ultrathin carbon film on top (Agar Scientific, catalog no. S187-4), blotted, and plunge-frozen in liquid ethane using a Vitrobot (FEI Company). Grids were transferred to an FEI Polara G2 microscope that was operated at 300 kV. Images were recorded on a back-thinned FEI Falcon detector at a calibrated magnification of  $42,277\times$  (yielding a pixel size of  $3.55 \text{ \AA}$ ). The defocus was varied from 1 to  $4 \mu\text{m}$ , using a dose of  $\sim 10$  or  $20 \text{ e}^-/\text{\AA}^2$ . Electron micrographs were evaluated for astigmatism and drift, and 456 micrographs were selected for further analysis.

**Image Processing.** Contrast transfer function parameters were estimated using CTFIND3 (33), and 28,502 particles were selected manually using the boxer program in the EMAN package (34). Reference-free 2D class averages and 3D reconstructions were calculated using RELION (20). The initial model for 3D refinement was calculated using CanDo (21). To avoid model bias against false high-resolution features, the initial model was low-pass filtered to  $60 \text{ \AA}$  before refinement. The data set was split into two separate halves at the outset of refinement, and two independent models were refined simultaneously. This procedure allowed us to prevent overfitting and thereby obtain a reliable estimate of the resolution based on so-called “gold-standard” Fourier shell correlation (FSC) (35). The overall resolution of the final model was estimated to be  $11.5 \text{ \AA}$ , based on the FSC = 0.143 criterion (35) (Fig. S4A). Low-pass filtering of the atomic model to the same resolution yielded a density map with similar features as the ones observed in our experimental map.

Inspection revealed that the density at the core of the object showed more high-resolution features than at the periphery. FSC calculations that used soft masks at different positions throughout the object indicated that the resolution varied from  $9.7 \text{ \AA}$  at the core to  $\sim 14 \text{ \AA}$  at the periphery. To optimally represent the information content throughout the object, we calculated a composite map by low-pass filtering a sharpened map (with a B-factor of  $-2,000 \text{ \AA}^2$ ) at different resolution cut-offs in the range of  $10\text{--}14 \text{ \AA}$  and combining these maps, depending on the estimated resolution at each position.

We confirmed the overall correctness of our structure by performing tilt-pair validation (36) (Fig. S4B). Using similar grids and microscope settings as the ones described above, 50 pairs of images were recorded at tilt angles of  $0^\circ$  and  $10^\circ$ , from which 342 particle pairs were selected manually. Alignment of the particles against the final composite map and analysis of the tilt-pair transformations were performed in the XMIPP package (37). Apart from



**Fig. 4.** Motifs beyond B-form DNA for nanotechnology. (A) A vertical stack of five Holliday junctions. (B) A pseudohelical structure that runs along the direction of the base pairs. (C) A bent helix and distorted cross-over due to an omitted base pair (*Lower*) and a typical cross-over without omission (*Upper*). (D) Crevices formed by splayed-out helices due to the absence of stabilizing cross-overs.

validating our map, this analysis also indicated that the individual particles may be aligned with an accuracy of  $\sim 2^\circ$ .

**Model Building.** An initial atomic model with the designed topology of parallel dsDNA helices in a rectangular lattice and with the known nucleotide sequence was calculated using custom software that relied on 3DNA (23) to calculate canonical B-DNA for all designed helical fragments. Separate rigid body fitting of each of the 82 helices in the EM density map yielded a model that was suitable for flexible fitting using MDFF (24). Because the MDFF approach was not capable of handling the entire structure (which, including hydrogens, consists of 460,641 atoms), we split the structure into multiple substructures, consisting of three neighboring columns of the rectangular lattice each. Before molecular dynamics (MD) simulation, the geometry of these models was improved by up to 5,000 steps of energy minimization, which included hydrogen bond restraints but did not include restraints toward the EM density. Subsequent MD simulations were performed at 300 K for 200,000 steps using both hydrogen bond and dihedral angle restraints and a weight on the EM density (GSCALE) of 0.3 kcal/mol. After MD, another 2,000 steps of energy minimization were performed

using hydrogen bonds and dihedral restraints and a weight on the EM density of 1.0 kcal/mol. A final model for the entire object was obtained by combining the central columns of all substructures. Low-pass filtering of the atomic model density yielded a map with similar features as the ones observed in the experimental reconstruction (Fig. S6 A and B). Additionally, FSC calculations between the final model and the reconstruction showed significant correlation up to 12 Å, indicating that the quality of the atomic model reflects the quality of the reconstructed density (Fig. S6C).

**ACKNOWLEDGMENTS.** We thank Lori Passmore, Andrew Carter, and Friedrich Simmel for discussions; Jake Grimmet, Shaoxia Chen, and Monika Rusp for technical support; Wei Zhao and Carlos Olguin (Autodesk Research) for support with graphics rendering; and Do-Nyun Kim and Mark Bathe for support with CanDo structure predictions. This work was supported by the Deutsche Forschungsgemeinschaft through the Excellence Cluster Center for Integrated Protein Science, Nano Initiative Munich, the Technische Universität München Institute for Advanced Study, European Research Council Starting Grant 256270 (to H.D.), and Medical Research Council Grant MC\_UP\_A025\_1013 (to S.H.W.S.).

- Alberts B, et al. (2002) *Molecular Biology of the Cell* (Garland Science, New York).
- Seeman NC (2010) Nanomaterials based on DNA. *Annu Rev Biochem* 79:65–87.
- Rothemund PWK (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* 440(7082):297–302.
- Shih WM, Quispe JD, Joyce GF (2004) A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature* 427(6975):618–621.
- Zheng J, et al. (2009) From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* 461(7260):74–77.
- Han D, et al. (2011) DNA origami with complex curvatures in three-dimensional space. *Science* 332(6027):342–346.
- Yurke B, Turberfield AJ, Mills AP, Jr., Simmel FC, Neumann JL (2000) A DNA-fueled molecular machine made of DNA. *Nature* 406(6796):605–608.
- Omabegho T, Sha R, Seeman NC (2009) A bipedal DNA Brownian motor with coordinated legs. *Science* 324(5923):67–71.
- Wickham SF, et al. (2012) A DNA-based molecular motor that can navigate a network of tracks. *Nat Nanotechnol* 7(3):169–173.
- Zhang C, et al. (2012) DNA-directed three-dimensional protein organization. *Angew Chem Int Ed Engl* 51(14):3382–3385.
- Andersen ES, et al. (2009) Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* 459(7243):73–76.
- Voigt NV, et al. (2010) Single-molecule chemical reactions on DNA origami. *Nat Nanotechnol* 5(3):200–203.
- Kato T, Goodman RP, Erben CM, Turberfield AJ, Namba K (2009) High-resolution structural analysis of a DNA nanostructure by cryoEM. *Nano Lett* 9(7):2747–2750.
- Douglas SM, et al. (2009) Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* 459(7245):414–418.
- Dietz H, Douglas SM, Shih WM (2009) Folding DNA into twisted and curved nanoscale shapes. *Science* 325(5941):725–730.

16. Douglas SM, et al. (2009) Rapid prototyping of 3D DNA-origami shapes with caDNA. *Nucleic Acids Res* 37(15):5001–5006.
17. Ke Y, et al. (2009) Multilayer DNA origami packed on a square lattice. *J Am Chem Soc* 131(43):15903–15908.
18. Ke Y, Voigt NV, Gothelf KV, Shih WM (2012) Multilayer DNA origami packed on hexagonal and hybrid lattices. *J Am Chem Soc* 134(3):1770–1774.
19. Woo S, Rothmund PWK (2011) Programmable molecular recognition based on the geometry of DNA nanostructures. *Nat Chem* 3(8):620–627.
20. Scheres SH (2012) A Bayesian view on cryo-EM structure determination. *J Mol Biol* 415(2):406–418.
21. Castro CE, et al. (2011) A primer to scaffolded DNA origami. *Nat Methods* 8(3):221–229.
22. Kim DN, Kilcherr F, Dietz H, Bathe M (2012) Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures. *Nucleic Acids Res* 40(7):2862–2868.
23. Lu XJ, Olson WK (2003) 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* 31(17):5108–5121.
24. Trabuco LG, Villa E, Mitra K, Frank J, Schulten K (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16(5):673–683.
25. Henderson R, et al. (2011) Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J Mol Biol* 413(5):1028–1046.
26. Murchie AI, et al. (1989) Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of antiparallel molecules. *Nature* 341(6244):763–766.
27. Eis PS, Millar DP (1993) Conformational distributions of a four-way DNA junction revealed by time-resolved fluorescence resonance energy transfer. *Biochemistry* 32(50):13852–13860.
28. Mao C, Sun W, Seeman NC (1999) Designed two-dimensional DNA Holliday junction arrays visualized by atomic force microscopy. *J Am Chem Soc* 121:5437–5443.
29. Ortiz-Lombardía M, et al. (1999) Crystal structure of a DNA Holliday junction. *Nat Struct Biol* 6(10):913–917.
30. Nowakowski J, Shim PJ, Prasad GS, Stout CD, Joyce GF (1999) Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme. *Nat Struct Biol* 6(2):151–156.
31. Eichman BF, Vargason JM, Mooers BH, Ho PS (2000) The Holliday junction in an inverted repeat DNA sequence: Sequence effects on the structure of four-way junctions. *Proc Natl Acad Sci USA* 97(8):3971–3976.
32. Douglas SM, Chou JJ, Shih WM (2007) DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proc Natl Acad Sci USA* 104(16):6644–6648.
33. Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334–347.
34. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128(1):82–97.
35. Scheres SHW, Chen S (2012) Prevention of overfitting in cryo-EM structure determination. *Nat Methods* 9(9):853–854.
36. Rosenthal PB, Henderson R (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* 333(4):721–745.
37. Scheres SHW, Núñez-Ramírez R, Sorzano COS, Carazo JM, Marabini R (2008) Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc* 3(6):977–990.



# Supporting Information

Bai et al. 10.1073/pnas.1215713109

## SI Text

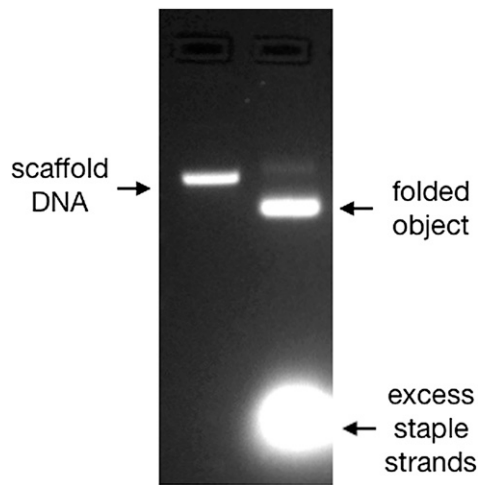
The frequency-dependent fall-off of signal power in a cryo-EM reconstruction may be used to estimate the intrinsic variability in the underlying structure. After correcting for the modulation transfer function of the detector, a B-factor sharpening of approximately  $-1,000 \text{ \AA}^2$  gave a density map at the core of the object, as expected from the atomic model. This B-factor still comprises various components, among which errors in the rotational and translational assignments and intrinsic variability in the structure are assumed to be the most significant ones. From the tilt-pair validation plot in Fig. S4, we estimated the errors in the rotational assignments to be on the order of  $2^\circ$ . According to Henderson et al. (1) this estimate results in a B-factor component

of  $B_{\text{rot}} = (\Delta\theta D)^2/2200 = \sim 400 \text{ \AA}^2$ . On the basis of signal-to-noise considerations, we estimated the errors in the translational assignments to be on the order of  $\sim 0.5$  pixel, or  $\sim 1.8 \text{ \AA}$ , leading to a B-factor component of  $B_{\text{trans}} = 8 \pi^2 (\text{rmsd})^2 = 250 \text{ \AA}^2$ . That then leaves a remaining B-factor component for the intrinsic structural variability (at the core of the object) of  $B_{\text{structural}} = 1,000 - 400 - 250 = 350 \text{ \AA}^2$ , which corresponds to an rmsd in the atom positions of  $\sim 2 \text{ \AA}$ . The mentioned range of up to  $3 \text{ \AA}$  in the main text is an attempt to reflect uncertainties in our B-factor estimation and in the empirical calculations related to the orientational errors. Additionally, as mentioned in the main text, the structural variability is larger at the periphery of the object than at its core.

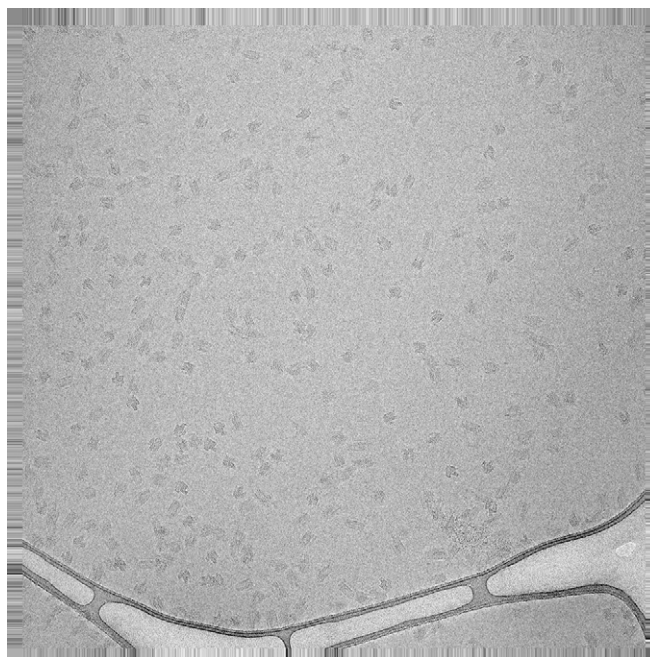
1. Henderson R, et al. (2011) Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J Mol Biol* 413(5):1028–1046.



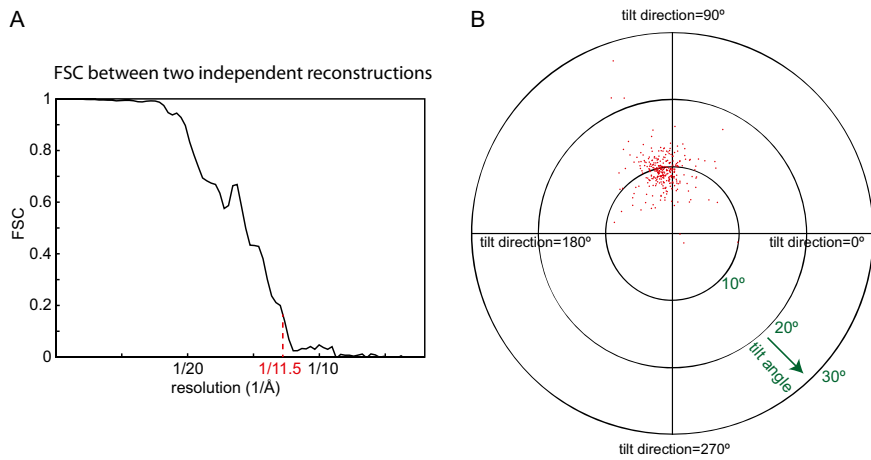




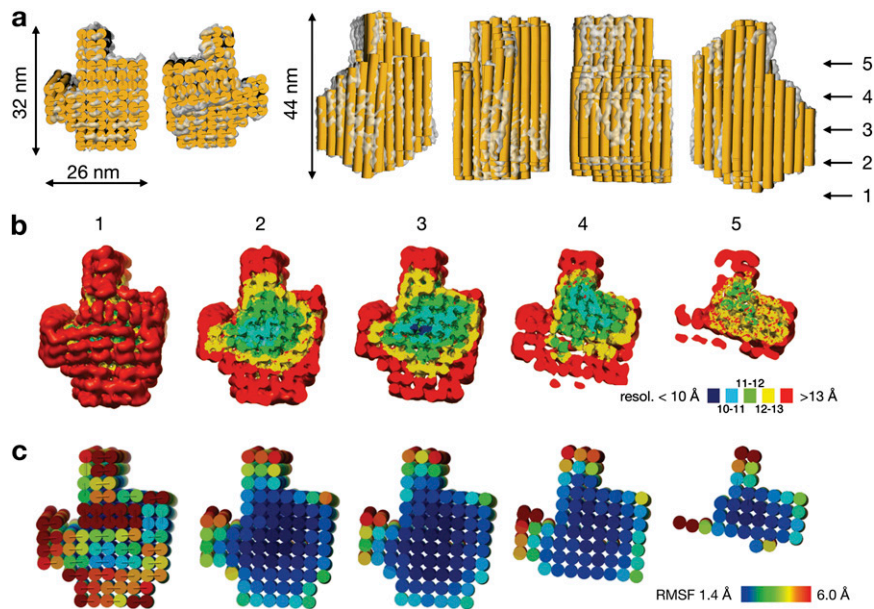
**Fig. S2.** Gel electrophoretic folding quality assessment. Photograph of an ethidium bromide-stained 2% agarose gel on which reaction products were electrophoresed.



**Fig. S3.** Exemplary cryo-EM micrograph. The field of view is 1.45  $\mu\text{m}$   $\times$  1.45  $\mu\text{m}$ .

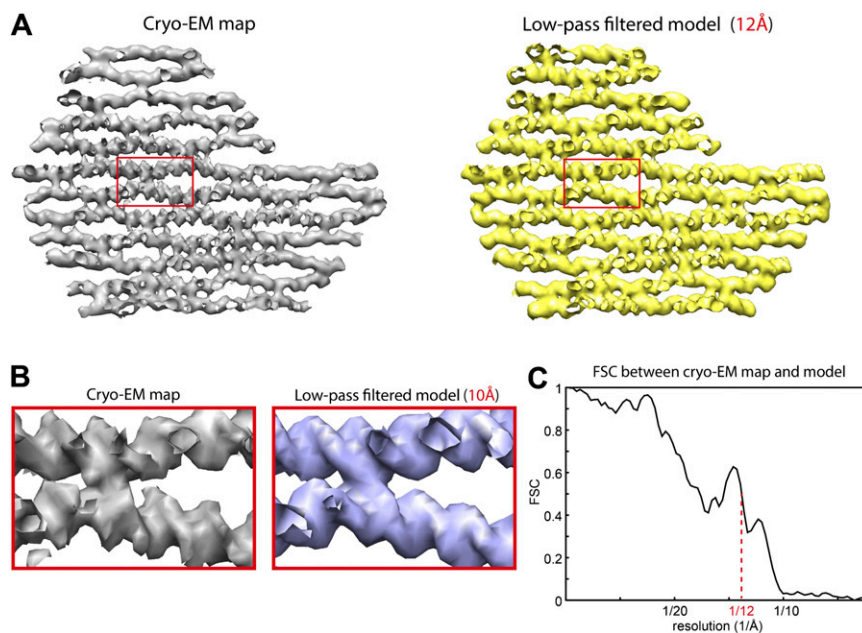


**Fig. 54.** Quality assessment of the cryo-EM map. (A) Fourier shell correlation (FSC) curve. (B) Tilt-pair validation plot. Note that 32 of the 342 particle pairs had tilt differences larger than 30° and are therefore not depicted.

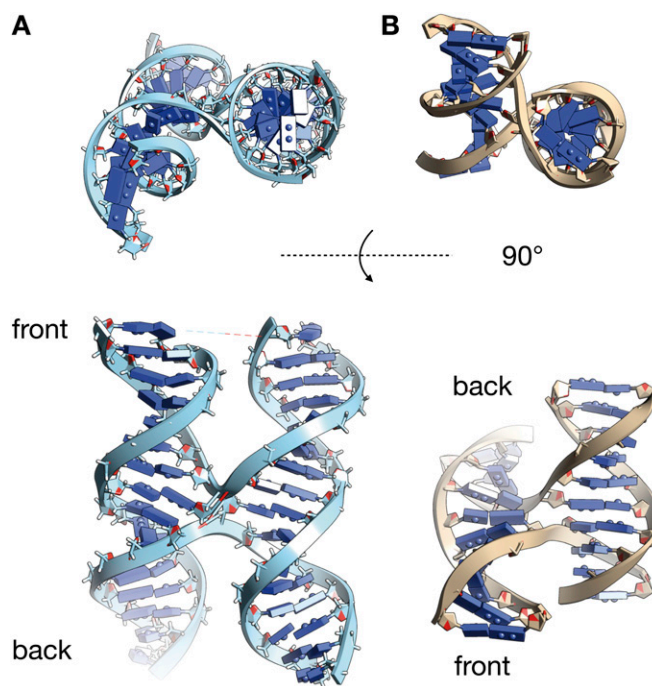


**Fig. 55.** Comparison of cryo-EM map and CanDo predictions. (A) Cryo-EM map (gray) and CanDo predicted shape (ocres) agree in overall dimensions and twist when assuming a reciprocal twist density of 10.44 bp/turn and an effective helical diameter of 2.6 nm. (B and C) The variation in resolution as estimated in the cryo-EM map (B) and the estimated root mean square fluctuation (rmsf) of the atoms in the shape prediction (C). Cross-sections are shown at five different slices (1–5) through the object, the positions of which are indicated with vertical arrows in A. Note that CanDo does not output a color scale bar for the estimated rmsf values. The scale bar shown reflects that colors are reported to range from dark blue (for an rmsf of 1.4 Å) to dark red (for an rmsf of 6 Å).





**Fig. 56.** Quality assessment of the pseudoatomic model. (A) Comparison of a slice (column 4) through the cryo-EM reconstruction and a 12-Å low-pass filtered map obtained from the pseudoatomic model. (B) As in A, but for a close-up of an area at the core of the object and a 10-Å low-pass filtered map obtained from the pseudoatomic model. (C) FSC between the cryo-EM map and the pseudoatomic model.



**Fig. 57.** Holliday junction conformation. (A) A typical Holliday junction in our structure (row 4, columns 4 and 5, base pairs 72–87). (B) Protein Data Bank ID code 1DCW. Both structures are depicted as ribbon-slab models. (Upper) The right helix in both junctions is pointing normally into the figure plane. (Lower) The right helix in both junctions is parallel to the figure plane.

**Table S1. Sequence information**

Start	End	Sequence	Length (nt)
0[79]	22[78]	CAAAGGGCGGGTGCCTCGCTCACTGCTTAATGTTGCATGCGGTACAGTTG	50
1[78]	23[77]	ATTGTTGTTCCGGCCAACCAGGGTGGTCCAAAGCGACAGGAAGCCGGA	48
3[43]	14[38]	TTTTTTTTTGGGTGAGACGGGCGAGTTTTTTTTTAAAGGGTCGTGCCGAGGATCCCCGGGTTTTT	67
3[78]	21[74]	AGCGAAATCGCCTGGCCCTGCCCCAGAGTAAAATAACATCACTT	45
4[69]	16[86]	AAAATCAGCCCCGGCAGGCGAAAATCCTGACGCTGGTTGAGAGAGGG	48
5[62]	6[38]	GGCCGTAAGCACTAAATTTTTTTTTTCGGAACCGGAAAGCCGGCGATTTTT	54
6[100]	8[82]	TTTTTAGCGGTCCCTTTTTTTTTTTTTGATGGTGGTTCAATAGCCCTTGGGC	51
7[38]	22[54]	TTTTTACGTGGCAACAGCTGAGAAGTTTTCCCTTTTTTTTTTTCAGTCACGACGTTGT	56
8[81]	32[82]	CGCGCGGGGACGTGCTTCGCGCTACAGCTTTCATAGGAATTGCGG	48
9[90]	10[86]	AGCGGTTTTTTTTTTTTTTTCGCTAGAGTTTTTTTTTTATAGGGTTGAGTAAAGAACTT	58
10[52]	13[53]	AACCTAAGGGTTTTTTTTTAAAGAAAGCGAAAGGTCACGCTAGCTCTTTTTTTTTTGAATTCGTAAT	67
10[85]	30[86]	GCGTTGAATGAGTGTAAAGTGTAAATTGTTAAGGAAGATGATAATCATG	48
12[65]	3[77]	AACCCGCCGCGCCCGCTTAAATGAATCAGTTTGGCCTTATAAATCAAA	48
12[109]	32[106]	TTTAACTCGCGTCTTTTTTTTTTGAATGGATCACCATCAACTGTTTTTTTTTATAGCAAACAT	66
13[54]	51[45]	CAAAACGTTAAAACCTAGCTTGAGAGATCTGGAGCTCATTGAATCCCCCTCGAATCG	56
14[61]	10[53]	ACTCTAAGCTGCATTCCAGTCGCGTGAACCGTCTATCAGGGCGATGTTTTTTTTTGGCCACTAGGA	67
15[38]	33[53]	TTTTTACC GGCGCTAAACAACCCGAAATTTTTTAAAAATCTGAGT	48
15[62]	1[77]	CAGTGTAGCGGAGCGGGCGCTAGGGCAAAAACCAACAAGAGTCCACT	48
16[85]	44[86]	ATTTTATCCTCGTTCGTGCATCTGGTGTAGCACCAGCACTCAGAGCAA	48
16[111]	36[106]	TTTTTAAACAGAGGAATTTTTTTTTTAAACGCTCCATCACCTCATTAG	48
18[64]	20[53]	GAGTAGAAGGTTGGGTAACGCCTTTTTTTTTAGGGTGTTTTTATCCA	48
18[93]	37[81]	TCTTTGATTAGTAAGAGTCTGTGGCGATCCAGGCAACAAATAACCCA	48
20[52]	56[54]	GCTGGCAGCTTTCGCGCGAGCAACACCGCTCCTGATTTAGAACCTGA	47
20[61]	2[43]	ATTACGAATCAGTGCAGAATCCGATTGCCTTACCAGGTCGAGGTATCCCCAAATCAAGTTTTTT	67
20[116]	41[106]	TTTTTGCAACGCTAGTTTGACATATCTTTTTTTTTTTGGTCAGTTGCCACCG	52
21[48]	59[53]	GCGATTAAATAGCTCAACTGGAAGTCTTTTGATCTATTATACCATCCTAGTCCCTG	54
21[75]	60[86]	GCTGGTAATATCCGAATTGAGCCAATCTAGGTCAGAAAAAGATGC	47
22[53]	52[54]	ATCGCCGGCGGATTTCAGTATTGGCAAAGAGATGATGATTAACAATAA	48
22[65]	5[61]	GACTTCTGGTTCGGTACGCAGGCCACCATTTAGAGCTTGACGGCTAAAG	48
22[77]	63[73]	CCAGTTCGTGAGAGCTGTTTAGCATTGCATCATTAGAGAAA	40
23[46]	64[54]	GCCGAAAGGGTACGGTGTCTATGTTTTGCGGATGGCTTAGAGCTT	44
23[78]	64[82]	AAGGTGCGGGTTGATTCGAAGTTACTTTAGGAGCACTAACGTTTTTA	48
23[102]	21[116]	TTAGGAGGCCCAAATTAACCGTTTTTTTTTTTGCAGCCATTTTTTT	47
24[69]	4[70]	AATGTGATGGGATACTGCAGGTACGGCCAGTTTTTCTTCTCACCGGC	48
24[77]	60[78]	CAACATAATAAGCGAAGATAATTGCACGTGATGATGGCATCGGAAT	48
24[89]	0[80]	GCCAGGGCGCCGAAGCAAGCTAACTTTTTTTTTTTCACATTAAGTGTTTTTTTTTTGACTCCAACGT	68
24[116]	42[106]	TTTTTTCATCGTTGTACCAACCTAATTTTTTTTTTAAATCGCCATTAAGATTTTCGAGCGG	63
26[57]	24[78]	GTTGGTCATAGCTGTTTCCGTGTGAAAAGCCTGTTTTAACCAT	44
26[77]	56[78]	TATAAGCAAAAGGGGAAGCCTTTTTTAATATTCAATCATATGCGGG	48
26[97]	25[116]	ACAAAATAATATACGAGCTACTATAGTTTTTTTTTTTTTTGCTTTTGACGCTTTTT	55
27[90]	12[110]	CCATCACCAAAAACCTCCGCTCACAATFCCACACTTTTTTT	40
28[80]	46[78]	GGCCAACAGATACGTGTAATGCGGGGAAGAAG	35
28[105]	47[101]	CGATTAGTCTTTAATTTTTTTTTTTCGCGAATATGATAACGGAACCTGACGAGAAACACCAGTCAATA	68
30[61]	31[45]	CTATTTATGTCAATCCCTTCTGAAACAGAGAATCGATCCTGAG	44
30[85]	51[73]	ATAAATCACAGACACCACATTCAACTAATGAT	32
31[27]	32[46]	TTTTTGGTCATTGGAATTTTTTTTTTCGGTAATCGTAATATTTGTGCAATGCTT	55
31[46]	46[22]	AGFCTACAAAAGAAGCTGTTGTAATTACCTTATAGAAATTTTT	44
31[91]	28[106]	TTTTTGAATGGCTACCAGTAAATTTGGCAGATTCACCATTTTTTTTTTGTACACA	53
32[45]	56[35]	TAGAACAATTACATAAACAACAATCATAATAGTACCGACAAAA	43
32[81]	51[93]	GATGAGAAAGGTAAATTGAAATCTACAAAAGAAGAGCAACACTATCAT	48
32[105]	58[110]	TATGTGAGTGAAGTTACAAGCCAACGATTTAACATAAATTTGTA	44
33[38]	45[41]	AATTATTTTTTTTTTAAATTCGCATTTTCGGATTACAGGCAATAGCATTAGG	52
33[54]	51[63]	AATTAACTCAGCTCATTAATGTTTAAATAAATAAAGGAAT	42
35[27]	36[42]	TTTTTTAAATCATCTCTTTTTTTTTTTGTGGGAACAACTCAGGAAAATAGTAGTGAAAA	59
36[41]	65[53]	GGTACCCTGAAAAGAGGTCTAAACCAATTTATTTTTTTTTTAAATCAAGAT	48
36[93]	9[89]	AATGGTAGCGCCATATCGTAACAGAAATCAGCAGTATAAG	40
36[105]	65[116]	ATAGACTTCAACCAGACCACCGGCCCTCCGGTATCTAACGAGCGTCT	47
37[58]	15[61]	GCAACATGAGGCGGTGACCCTAAGCGAGTACCACCA	36
37[82]	80[80]	GCAGCAGATTATCAAAAACAGATAGGCAGATTATACAGACCTAAACTATATGTATCAATAG	62
37[122]	56[118]	CAAAATTTTTTTTTTTTTTCAAACCCCAATCATGCTGAACACCAGAAGAGGTTTTAAAT	59
38[92]	7[100]	GAAAGAGAACAATGTTGGGAACCATCAGGATTAAGTTGCAGCATTTTT	50
40[39]	21[47]	ATTTTTTTTTTTTTTAAATATGCTTTTTTTTTTAACTAAAGGATTTTTTTTTTTGTGCTGCAAG	64
40[61]	14[62]	TTGCTCTTCATCCCATTTGGGCGCACCCGACGACAGTAAACGCG	48





