Integrated computer-aided engineering and design for DNA assemblies

Authors: Chao-Min Huang[1], Anjelica Kucinic[2], Joshua A. Johnson[3], Hai-Jun Su*[1], and Carlos E. Castro*[1,3]

Affiliations:

[1]Department of Mechanical and Aerospace Engineering, The Ohio State University, Columbus, Ohio 43210, USA

[2]Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus, Ohio 43210, USA

[3]Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210, USA

* corresponding author emails: castro.39@osu.edu, su.298@osu.edu

Abstract:

Recently, DNA has been used to make nanodevices for a myriad of applications across fields including medicine, nanomanufacturing, synthetic biology, biosensing, and biophysics. However, current DNA nanodevices rely primarily on geometric design, and it remains challenging to rationally design functional properties such as force-response or actuation behavior. Here we report an iterative design pipeline for DNA assemblies that integrates computer-aided engineering (CAE) based on coarse-grained molecular dynamics with a versatile computer-aided design (CAD) approach that combines top-down automation with bottom-up control over geometry. This intuitive framework allows for rapid construction of large, multicomponent assemblies from 3D models with finer control over the geometrical, mechanical and dynamical properties of the DNA structures in an automated manner. This approach expands the scope of structural complexity and enhances mechanical and dynamic design of DNA assemblies.

maintext

Introduction

Combining computer-aided design (CAD) with computer-aided engineering (CAE)[1] (i.e. iterative design guided by simulation) into Integrated Computational Materials Engineering (ICME) frameworks[2,3] is essential to coalesce material properties and geometric design across

multiple length scales. ICME is well-established for tailoring performance metrics of engineering materials like alloys and composites[4]. In contrast, integrating CAD and CAE for biomolecular self-assembly has remained elusive. Computationally-guided design of proteins is well-established[5], but the complexity of structures and interactions governing self-assembly impede the development of geometric CAD. On the other hand, CAD tools that describe the structure and interactions of DNA have been essential to facilitating structural DNA nanotechnology[6–9], but currently these approaches rely purely on geometric design. The emergence of high fidelity coarse-grained molecular dynamics (MD) simulation tools for DNA nanostructures[10–13] provides an opportunity to realize CAE for DNA-based design to enable systems with new levels of structural complexity that can also be tailored for functional properties such as reconfiguration, mechanical properties, or stimulus response. Here we present an ICME approach for DNA assemblies that leverages a custom CAD tool that enhances the scope of geometric design and facilitates tight integration with coarse-grained MD simulations[10–12] to enable CAE for DNA assemblies.

Precise control over geometry makes DNA assemblies[14–16] attractive for applications such as drug delivery[17], templating nanomaterials or molecules[18–21], molecular measurement tools[20,22], and nanorobotics[23–25]. However, as the complexity of design and function increases, DNA-based design must be integrated with computational modeling to efficiently tune functional properties and digitally validate system performance. Until the last few years, DNA assemblies were primarily designed using bottom-up approaches[6,7] where strands are manually arranged and connected. In particular, caDNAno[6] was transformative in simplifying the design process and broadening use of DNA origami, but the largely manual routing is a slow process that is challenging for non-experts and limits designs to a small number of components with simple connectivity. While bottom-up design approaches remain widely used, recent efforts have

established top-down approaches that lower the barrier and speed up the design process. These top-down approaches take convenient geometric inputs (e.g. line models, 2D or 3D polyhedra) and utilize routing algorithms to partially[8] or fully[9,26–28] automate the design of wireframe structures. These tools are simple and fast, but they lack user interfaces to tune design parameters at multiple structural levels that regulate dynamic behavior and actuated reconfiguration; hence, they are primarily focused on design of static wireframe structures. In summary, current bottom-up methods provide user control over geometry but limited complexity and relatively slow manual design, and top-down methods offer rapid and simple approaches to design complex shapes while sacrificing structural diversity and the ability to design features beyond a static shape (Supplementary Figure 1). Hence, a rapid and versatile design approach is still needed to harness the potential of CAE for DNA assemblies.

Here we introduce a hybrid design methodology that merges bottom-up and top-down methods to provide a high level of structural diversity, expand the scope of complex design, and enable engineering of mechanical and dynamic properties. This hybrid framework leverages design at multiple scales spanning from the single nucleotide level to large and complex DNA assemblies. We implemented this approach through a GUI-based software called Multi-component Assembly in a Graphical Interface guided by Computation for DNA assemblies (MagicDNA) (Supplementary Figure 2) that integrates simple user inputs, intuitive 3D visualization and manipulation, automated routing algorithms, and straightforward interfacing with other design and modeling tools for rapid CAE. We present 66 designs with simulation results and selected 14 structures with a range of complexity for experimental validation. This framework simplifies and accelerates the design process, significantly expands the design domain for more applications, and enhances the robustness of DNA-based design.

Iterative design process with simulation feedback

Modern CAD software supports geometric modeling for single component design, and assembly modeling for design of machinery with multiple components. Mimicking this framework, we introduce a hybrid top-down and bottom-up design process for DNA assemblies based on scaffolded DNA origami[15,16], where a long scaffold strand is folded into a compact structure via base-pairing with many shorter strands. Expanding on prior approaches[6,8,9] that allow design at the nucleotide (bottom-up) and overall assembly (top-down), we introduce an intermediate component level for design, where components are bundles of two or more dsDNA helices (Supplementary Figure 3 and Supplementary Note 1). Introducing this component level provides a convenient intermediate to design a wide range of static and dynamic assemblies with simple user inputs and interactive 3D visualization. In MagicDNA, GUI tools and algorithms at the nucleotide, component, or assembly level enable seamless data exchange from lower to higher levels (i.e. bottom-up) or higher to lower levels (i.e. top-down). "Part" design is carried out among the nucleotide and component levels, and "assembly" modeling is carried out among the component and assembly levels.

This collection of algorithms and GUI tools enables a systematic design workflow (Fig. 1) where the first step is inputting the overall assembly geometry (Fig. 1a). The top-down approach converts a 3D line model (imported from .STEP file or using built-in sketch interface) to components with user-defined cross-section and length. Alternatively, users can use a bottom-up approach by inserting components from a part library to build the assembly. Either way, users can modify component geometries (Supplementary Figures 4 to 6) and manipulate components into a desired assembly configuration (Supplementary Figure 7). Connections between components (Fig.

1b) can be introduced either at the ends of components or on the surface where the scaffold is at an outward facing helical orientation. By specifying which components are connected (i.e. defining the connectivity matrix), connections can be automatically formed based on minimal distances between potential connection points (Supplementary Figure 8). Alternatively, users can manually specify connections. In either case, each connection is formed by a double-scaffold crossover, and the length of these connections can be adjusted to tune joint geometry or mechanical properties.[29]

This approach enables straightforward design of assemblies with many interconnected components, making manually routing the DNA exceedingly challenging. Recent top-down approaches developed automated routing algorithms[8,9,26–28]; however, those implementations have been limited to static wireframe structures with uniform components. We developed a general scaffold routing algorithm based on double-scaffold crossovers (i.e. two single-stranded scaffold connections) that allows for diverse geometries at the component level including different lattices (e.g. honeycomb[16] or square[30]), varied connectivity between components (e.g. number and length of connections, end-to-end, end-to-side, and side-to-side), and incorporation of multiple scaffolds (Fig. 1c, discussed below), providing versatile control over structural, mechanical, and dynamic properties. Details of the algorithm are provided in Supplementary Figures 9 to 12. Briefly, components are sub-divided into pairs of neighboring helices connected by external scaffold crossovers at their ends to form a scaffold cycle. Then, cycles corresponding to pairs of helices that connect across a joint are merged to reach a total number of N scaffold cycles. We use a spanning tree algorithm[31] to identify N-1 internal crossovers to combine the N cycles into a single cycle. The use of the spanning tree approach builds on prior work that solved a similar scaffold connectivity step for automated design of 3D wireframe geometries made from 2-helix bundle

edges[9], and more recent work on automated design of 2D and 3D wireframe structures with 2-helix or 6-helix bundle edges[26–28]. Similarly, we adopted a staple routing algorithm based on prior work[6] (details in Supplementary Figures 13-15) with added functionality for actuation or higher-order assembly. MagicDNA provides convenient GUIs for defining routing parameters and visualizing scaffold and staple routing in 3D (Fig. 1c and Supplementary Figures 16 to 19). We also added a feature to export .JSON files, which can be modified in caDNAno[6] and uploaded back into MagicDNA. This also enables the use of computational tools such as CanDo[32,33] and COSM[13] (Supplementary Figure 20), which run caDNAno files.

To realize an ICME approach for DNA materials, we incorporated rapid virtual prototyping in MagicDNA with CAE simulation feedback (Supplementary Figures 21 and 22) to fine-tune structural and functional properties. In particular, we incorporated tools to interface with the coarse-grained MD model oxDNA[10,11,34], which is frequently used to predict the shape, mechanical properties, and motion of DNA nanostructures.[35–38] We automated the generation of oxDNA input files and integrated tools that calculate the average shape and fluctuations (Fig. 1e) or track key parameters (e.g. angles). Once desired metrics are achieved, the design can be fabricated and verified experimentally (e.g. by TEM, Fig. 1f).[32,39]

Top-down parametric design of functional devices

To demonstrate the versatility of our hybrid design process and the ability to rapidly adjust parameters used to build out the line model, we present several examples with the design workflow consisting of: 1) sketching the line model, 2) specifying component properties, 3) 3D manipulation of components to arrange a desired assembly configuration, 4) specifying ssDNA connections between components, and 5) running automated routing algorithms (Supplementary Movie 1).

This workflow can be completed within ~10 minutes, and then nucleotide or component parameters can be readjusted within seconds. We designed series of nanopores, nano-rings, nano-hinges, and 4-bar mechanisms (Fig. 2). which have demonstrated applications such as detecting or probing biomolecules[22,40], templating nanoparticles[41], or as platforms for biomedical applications[42]. For each case, we generated multiple designs for simulation (10^7 steps in oxDNA) and chose one for experimental verification.

For the nanopore, we generated four designs (Fig. 2a, Supplementary Figures 23 and 24). The one chosen for fabrication (Fig. 2a, right) consists of a honeycomb lattice platform and a square lattice central pore, demonstrating distinct geometries for individual components rigidly connected in 3D. For the ring, we generated three designs (Fig. 2b, Supplementary Figures 25 and 26) starting with a polygonal line model and approximated the local curvature by incorporating gradients along the ends of bundles (Fig. 2b inset) to form angled vertex connections. Linear end gradients, as in the ring, can be input directly as a component property, and non-linear or discrete end gradients can be specified in a bottom-up manner by extruding helices with base-pair resolution. Both curvature and vertices are useful features in DNA-based design[43,44].

A key goal of this framework is to simplify design of reconfigurable assemblies, since no automated tools address this emerging class of DNA systems. We designed three versions of a dynamic hinge (Fig. 2c, Supplementary Figures 27 and 28). MagicDNA allows visualization of the local 3D helical structure to assign joint connections at desired locations, helical orientations, and ssDNA lengths to form an axis of rotation. The hinge selected for fabrication has a non-uniform cross-section (i.e. hollow in the middle) and exhibits flexible angular motion (Fig. 2c, right). Finally, we designed three mechanisms based on 4-bar linkages (Fig. 2d, Supplementary Figures 29 and 30). We selected one for fabrication and used a longer oxDNA simulation (3×10^8

steps) to track the motion of the top vertex. The simulations closely matched conformations measured by TEM (Fig. 2d, middle right), demonstrating the design of functional properties beyond shape, such as mobility.

Top-down iterative design for complex structures

The ability to create complex multi-component assemblies makes CAE simulation feedback essential to design verification and improvement. We used a modular approach by first optimizing sub-systems consisting of a few components. This reduces simulation time and allows more extensive study on how design parameters affect metrics (e.g. shape, stiffness, configuration, stability, etc., see Supplementary Figures 31 and 32 for examples). Additionally, simulation feedback can be used to create deformed features by connecting components with mismatched length and stiffness (Supplementary Figure 33). While a modular approach is efficient, one can still iterate sub-component designs within a larger assembly (e.g. birthday cake structure in Supplementary Figure 34).

We used this "hierarchical design" approach (i.e. fine-tuning component or sub-assembly geometry and then adding those into larger assemblies) to design four assemblies (Fig. 3, Supplementary Figures 35 to 43): a Stewart platform, a compound compliant joint, a gripper, and a trophy, all inspired by macroscopic counterparts. These structures would not be practical to design with prior software tools due to the complex features (many components, 3D geometry and connectivity, vertices, curvature, hybrid lattice, etc.). The top-down line model input is convenient for assigning approximate geometric parameters to a large number of components in the first iteration. Then component and nucleotide level parameters are specified in MagicDNA to complete or modify a design. The simulation trajectory of the full design was also used to guide

adjustments to enhance mechanical properties. The platform, compound joint, gripper, and trophy, required 2, 4, 13, and 5 iterations, respectively.

Bottom-up and hierarchical design of reconfigurable assemblies

Functional materials are often comprised of many similar structural units, as in structural metamaterials[45], which is well-suited to the hierarchical approach. Here, we demonstrate reconfigurable assemblies comprised of multiple similar units, specifically deployable (Fig. 4a, serial tetrahedron) and rotational mechanisms (Fig. 4b, the butterfly). We generated the structural unit design using the top-down approach and optimized the design with simulation feedback. Then, we used the bottom-up approach to integrate these units into a desired assembly pattern.

The deployable tetrahedron was designed by first sketching three lines and converting them into bundles with end gradients to form a triangular shape verified with oxDNA (Fig. 4a top and Supplementary Figure 44). The verified triangle was duplicated into two instances, which were arranged to form two sides of a tetrahedron. A blade component, which controls the open or closed configuration, was added to complete the tetrahedron. Our scaffold algorithm does not allow connections to multiple components on a single node. Since multi-way junctions are required for the tetrahedron design, we fine-tuned the routing in caDNAno (Supplementary Figure 44). The blade can be folded into a straight or contracted configuration to form the deployed or compact states, both validated by oxDNA. The verified tetrahedron was duplicated into three instances while removing one triangular plate on the two outer tetrahedrons. The final designs in both deployed and compact configurations were verified by simulation and experimental fabrication

(Fig. 4a, Supplementary Figures 45 to 47). We also demonstrated a deployable umbrella mechanism following this hierarchical design (Supplementary Figures 48 to 50).

Another common actuation strategy for dynamic devices is to add strands that connect multiple overhangs to latch two components together.[24,25] We created an overhang design GUI in MagicDNA where users input parameters (e.g. length) and specify locations and connections directly on the 3D model. The staple routing algorithm satisfies these inputs and optimizes overhang sequences to minimize complementarity to the scaffold.[41] We used the bottom-up duplication of two triangles for the butterfly design and fine-tuned the scaffold routing at the joint with caDNAno (Supplementary Figure 51). Next, we specified 28 pairs of overhangs (Fig. 4b, top, Supplementary Figure 51) in the overhang design GUI, including 8 pairs with an identical sequence to close the wings along the upper edges, 8 pairs with a second identical sequence to close the wings along the lower edges, and 12 pairs all with unique sequences to assemble multiple butterflies together along the outer edges. By adding complementary strands to the overhangs in a specific order, we show different actuation/assembly pathways result in two distinct high-order assemblies (Fig. 4b, Supplementary Figures 52-54).

Expanding the design domain of complex DNA assemblies

We have shown the versatility of our hybrid top-down and bottom-up design framework implemented in MagicDNA. Here we generalize to an even wider spectrum of design by integrating wireframe, surface-based, and lattice-based models (Fig. 5, Supplementary Note 2 and Supplementary Figures 55-68) into complex assemblies. These types of structures have been demonstrated individually[46]; however, no current CAD tools integrate these classes of structures into a single assembly.

We demonstrated this capability to integrate lattice, surface, and wireframe components into a DNA airplane assembly (Fig. 5 center). We started with a top-down line model for the whole system to establish approximate sizes for the six sub-systems (Supplementary Figure 69). The sub-systems were individually optimized and then combined into a single assembly consisting of a lattice-based fuselage, wireframe wings and tail, and surface-based turbines. This airplane design totals 28 bundle components, ~33 kbps, and over 800 staples, showing the design framework, algorithms, and software essentially have no limit for scale. Simulation times increase, but continued efforts in coarse-grained MD[12] are addressing this challenge (Supplementary Figure 22).

Multi-scaffold and modular designs

Increasing design complexity and size generally requires more material. There have been multiple recent advances in integrating multiple DNA structures[47,48] to make well-controlled micron-scale assemblies; however, these are usually carried out through multiple reaction steps and have only demonstrated assembly of similar and relatively simple structures. We sought to enable straightforward design of large assemblies with many unique components that could be fabricated in a single-pot reaction. Prior work demonstrated single-pot folding of large DNA structures using up to ~50 kb scaffolds[49], integrating multiple orthogonal sequence scaffolds[39,44], or using exclusively brick strands[50] (similar to staple strands). We focused on using orthogonal scaffolds based on a recent breakthrough demonstrated for simpler assemblies.[39] Here we enable assemblies with many distinct structural components, asymmetric and fully addressable shape, and programmable mechanical and dynamic properties. Also, the use of multiple scaffolds allows for modular design where single scaffold sub-assemblies can be re-used in multiple higher order assemblies.

We implemented two approaches for multi-scaffold design in MagicDNA. The first is based on a spanning forest algorithm and intended for modular design, allowing users to add well-defined interfaces between structures by: 1) adding internal scaffold crossovers to form a desired interface, and 2) ignoring potential scaffold crossovers in this region during automated scaffold routing to ensure formation of separate trees and cycles (details in Supplementary Note 3, Supplementary Figures 9, and 70 to 72). This approach is demonstrated by the robotic manipulator with an exchangeable end effector (Fig. 6a, Supplementary Figures 73-78). The arm is connected to either a claw-like or tweezer-like end effector with the two structures interlocked at the interface, which was shown to improve yield[39]. Our results demonstrate high yield of this robotic manipulator with exchangeable end effectors as an example of modular robots. This approach with defined interfaces also allows for direct design and simulation of multiple structures bound together intended for hierarchical multi-pot assembly (Supplementary Figures 79-80)[48].

In the second multi-scaffold approach, intended for complex asymmetric shapes, we split the full single scaffold cycle into *K* cycles by applying *K*-1 crossovers (Supplementary Figure 81). The algorithm searches internal scaffold crossovers until it finds crossovers that break the initial single scaffold into *K* cycles of the desired lengths. To facilitate finding solutions, we include some tolerance for the cycle lengths to be somewhat shorter than the divided scaffold length (default = 10%). We used this approach to design a wireframe MagicDNA logo using a bottom-up process to assemble a wand with a stick and star (Supplementary Figure 58) and the "DNA" script (Supplementary Figure 65). The MagicDNA logo was folded with an M13-derived p8064[16] scaffold and a CS4-7559 scaffold[39], and verified with TEM (Fig. 6b, and Supplementary Figures 82 and 83). In the case of the airplane (Fig. 6c), there are about ~30 kb and roughly two thousand possible crossovers in the initial single scaffold cycle. We used a custom stochastic search

algorithm to identify three (K-1) crossovers splitting the K=4 scaffolds with the desired scaffold lengths (M13-derived p8064 and CS3_L_7560, CS4_7557, CS5_7559[39]). To ensure stable attachment between the four scaffolds during folding, we developed a heuristic optimization with the objective of maximizing the number of staples that connect at least two scaffolds (Supplementary Figures 84 to 87, and Movie 2). In the resulting four-scaffold routing for the airplane, 69% of staples connect at least two scaffolds (red staples in Fig. 6c).

Outlook

We demonstrated a versatile framework that combines the benefits of top-down, bottom-up, and hierarchical design. This framework evolved from a heuristic design process for dynamic DNA origami mechanisms (DOMs)[23–25,38] and from recent advances in hierarchical multi-structure assembly[47,48], multi-scaffold assembly[39,44], and simulation of DNA nanodevices[10–12,37]. Compared with bottom-up design tools[6,7], MagicDNA has routing algorithms and component and assembly level manipulations to allow for rapid construction of large many-component designs directly from 3D models with simple user inputs. Compared with top-down tools[8,9,26], our framework significantly enhances user control over geometric, mechanical and dynamic properties of assemblies, and enables actuation, higher order assembly, and multi-scaffold capabilities. The expanded control over design does come at the cost of reduced automation due to user inputs and manual component manipulation. Hence, for designing wireframe geometries within the scope of existing automated design tools[8,9,26–28], those tools provide a more appropriate alternative.

Moving forward, this integrated CAD/CAE framework can accelerate the development of next generation molecular robots. In addition, the ability to template gold, silver, or silica on DNA assemblies[18,19,41], and the use of DNA "masks"[21] for lithography provide avenues to exploit these

new levels of size and geometric complexity. Furthermore, with this CAE-ready design approach, the continued development of simulation tools should facilitate rapid design of devices for targeted functions. Finally, this new regime of fabrication opens new questions about folding pathways, kinetics, and thermodynamics for these complex (many-component, multi-scaffold, hybrid lattice, etc.) assemblies.

Data availability

Original data for TEM images and gel electrophoresis are included as Source data. The remaining data supporting the findings of this study are available within the article and its supplementary information files or available from the corresponding author upon reasonable request.

Code availability

The developed design software MagicDNA is available from GitHub at https://github.com/cmhuang2011/MagicDNA.

Acknowledgments: This work was supported by National Science Foundation grants 1536862 to H.-J.S and C.E.C. and grant 1921955 to C.E.C. We acknowledge Floris Engelhardt and Hendrik Dietz for providing custom scaffolds, Tural Aksel and Shawn Douglas for sharing the caDNAno toolkit, Christopher Maffeo and Aleksei Aksimentiev for supporting interface to MrDNA, Tara MacCulloch and Nickolas Stephanopoulos for providing K-10 peptide, and Angela Tran, Phuong Le, and Philip Lukeman for testing MagicDNA Runtime packages. We thank the Campus Microscopy and Imaging Facility (CMIF) of The Ohio State University for imaging support. We also thank Wolfgang Pfeifer and Christopher Maffeo for critiques on the manuscript and supplement material.

Funding: National Science Foundation (NSF CMMI) 1536862, National Science Foundation (NSF CMMI) 1921955.

Author contributions: C.-M.H. developed the software and the algorithm, designed and simulated all the structures, analyzed the data, prepared the tutorial, and supported experiments. A.K. conducted the majority of experiments and analyzed the experimental results. J.A.J was an initial user and provided critical early feedback on software features, interface, and instructions. H.-J.S. supervised the development of the software and interpreted the data. C.E.C supervised the experimental validation and the entire study, supported the development of the software, and interpreted the data. C.-M.H, A.K., H.-J.S, and C.E.C. wrote the manuscript. All authors commented on and edited the manuscript.

Competing interests: The authors declare no competing interests.

Additional information:

Supplementary Notes 1 to 3

Supplementary Figures 1 to 87

Supplementary Table 1

Supplementary References 1 to 30

Supplementary Movies 1 and 2

The excel sheets for the staple list of the 14 structures for fabrication

References:

1. Chang, K.-H. *e-Design: Computer-Aided Engineering Design*. (Academic Press, 2016).
2. Olson, G. B. Computational Design of Hierarchically Structured Materials. *Science* **277**, 1237–1242 (1997).

3. Panchal, J. H., Kalidindi, S. R. & McDowell, D. L. Key computational modeling issues in Integrated Computational Materials Engineering. *Computer-Aided Design* **45**, 4–25 (2013).
4. Backman, D. G. *et al.* ICME at GE: Accelerating the insertion of new materials and processes. *JOM* **58**, 36–41 (2006).
5. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
6. Douglas, S. M. *et al.* Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res* **37**, 5001–5006 (2009).
7. Williams, S. *et al.* Tiamat: A Three-Dimensional Editing Tool for Complex DNA Structures. in *DNA Computing* 90–101 (Springer, Berlin, Heidelberg, 2008).
8. Benson, E. *et al.* DNA rendering of polyhedral meshes at the nanoscale. *Nature* **523**, 441–444 (2015).
9. Veneziano, R. *et al.* Designer nanoscale DNA assemblies programmed from the top down. *Science* **352**, 1534–1534 (2016).
10. Doye, J. P. K. *et al.* Coarse-graining DNA for simulations of DNA nanotechnology. *Phys. Chem. Chem. Phys.* **15**, 20395–20414 (2013).
11. Snodin, B. E. K. *et al.* Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *J. Chem. Phys.* **142**, 234901 (2015).
12. Maffeo, C., & Aksimentiev, A. MrDNA: a multi-resolution model for predicting the structure and dynamics of DNA systems. *Nucleic acids research*, 48(9), 5135-5146 (2020).
13. Reshetnikov, R. V. *et al.* A coarse-grained model for DNA origami. *Nucleic Acids Res* **46**, 1102–1112 (2018).
14. Seeman, N. C. Nucleic acid junctions and lattices. *Journal of Theoretical Biology* **99**, 237–247 (1982).
15. Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
16. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009).
17. Jiang, Q. *et al.* DNA Origami as a Carrier for Circumvention of Drug Resistance. *J. Am. Chem. Soc.* **134**, 13396–13403 (2012).
18. Sun, W. *et al.* Casting inorganic structures with DNA molds. *Science* **346**, 1258361 (2014).

19. Liu, X. *et al.* Complex silica composite nanomaterials templated with DNA origami. *Nature* **559**, 593–598 (2018).
20. Shaw, A. *et al.* Spatial control of membrane receptor function using ligand nanocalipers. *Nature Methods* **11**, 841–846 (2014).
21. Shen, B. *et al.* Plasmonic nanostructures through DNA-assisted lithography. *Science Advances* **4**, eaap8978 (2018).
22. Le, J. V. *et al.* Probing Nucleosome Stability with a DNA Origami Nanocaliper. *ACS Nano* **10**, 7073–7084 (2016).
23. Gerling, T., Wagenbauer, K. F., Neuner, A. M. & Dietz, H. Dynamic DNA devices and assemblies formed by shape-complementary, non–base pairing 3D components. *Science* **347**, 1446–1452 (2015).
24. Marras, A. E., Zhou, L., Su, H.-J. & Castro, C. E. Programmable motion of DNA origami mechanisms. *PNAS* **112**, 713–718 (2015).
25. Zhou, L., Marras, A. E., Huang, C.-M., Castro, C. E. & Su, H.-J. Paper Origami-Inspired Design and Actuation of DNA Nanomachines with Complex Motions. *Small* **0**, 1802580 (2018).
26. Jun, H. *et al.* Automated Sequence Design of 3D Polyhedral Wireframe DNA Origami with Honeycomb Edges. *ACS Nano*, 13(2), 2083-2093 (2019)
27. Jun, H. *et al.* Autonomously designed free-form 2D DNA origami. *Science Advances* **5**, eaav0655 (2019).
28. Jun, H., Wang, X., Bricker, W. P. & Bathe, M. Automated sequence design of 2D wireframe DNA origami with honeycomb edges. *Nature Communications* **10**, 1–9 (2019).
29. Zhou, L., Marras, A. E., Su, H.-J. & Castro, C. E. DNA Origami Compliant Nanostructures with Tunable Mechanical Properties. *ACS Nano* **8**, 27–34 (2014).
30. Ke, Y. *et al.* Multilayer DNA Origami Packed on a Square Lattice. *J. Am. Chem. Soc.* **131**, 15903–15908 (2009).
31. Pandey, S. *et al.* Algorithmic design of self-folding polyhedra. *PNAS* **108**, 19885–19890 (2011).
32. Castro, C. E. *et al.* A primer to scaffolded DNA origami. *Nat Meth* **8**, 221–229 (2011).
33. Kim, D.-N., Kilchherr, F., Dietz, H. & Bathe, M. Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures. *Nucleic Acids Res* **40**, 2862–2868 (2012).

34. Poppleton, E. *et al.* Design, optimization and analysis of large DNA and RNA nanostructures through interactive visualization, editing and molecular simulation. *Nucleic Acids Res* **48**, e72–e72 (2020).
35. Snodin, B. E. K., Schreck, J. S., Romano, F., Louis, A. A. & Doye, J. P. K. Coarse-grained modelling of the structural properties of DNA origami. *Nucleic Acids Res* **47**, 1585–1597 (2019).
36. Shi, Z., Castro, C. E. & Arya, G. Conformational Dynamics of Mechanically Compliant DNA Nanostructures from Coarse-Grained Molecular Dynamics Simulations. *ACS Nano* **11**, 4617–4630 (2017).
37. Sharma, R., Schreck, J. S., Romano, F., Louis, A. A. & Doye, J. P. K. Characterizing the Motion of Jointed DNA Nanostructures Using a Coarse-Grained Model. *ACS Nano,* 11(12), 12426-12435 (2017).
38. Huang, C.-M., Kucinic, A., Le, J. V., Castro, C. E. & Su, H.-J. Uncertainty quantification of a DNA origami mechanism using a coarse-grained model and kinematic variance analysis. *Nanoscale* **11**, 1647–1660 (2019).
39. Engelhardt, F. A. S. *et al.* Custom-Size, Functional, and Durable DNA Origami with Design-Specific Scaffolds. *ACS Nano* **13**, 5015–5027 (2019).
40. Wei, R., Martin, T. G., Rant, U. & Dietz, H. DNA Origami Gatekeepers for Solid-State Nanopores. *Angewandte Chemie International Edition* **51**, 4864–4867 (2012).
41. Johnson, J. A., Dehankar, A., Winter, J. O. & Castro, C. E. Reciprocal Control of Hierarchical DNA Origami-Nanoparticle Assemblies. *Nano Lett.,* 19(12), 8469-8475 (2019).
42. Ponnuswamy, N. *et al.* Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature Communications* **8**, 15654 (2017).
43. Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into Twisted and Curved Nanoscale Shapes. *Science* **325**, 725–730 (2009).
44. Zhang, F. *et al.* Complex wireframe DNA origami nanostructures with multi-arm junction vertices. *Nature Nanotechnology* **10**, 779–784 (2015).
45. Bertoldi, K., Vitelli, V., Christensen, J. & van Hecke, M. Flexible mechanical metamaterials. *Nature Reviews Materials* **2**, 1–11 (2017).
46. Seeman, N. C. & Sleiman, H. F. DNA nanotechnology. *Nat Rev Mater* **3**, 1–23 (2017).
47. Wagenbauer, K. F., Sigl, C. & Dietz, H. Gigadalton-scale shape-programmable DNA assemblies. *Nature* **552**, 78 (2017).

48. Tikhomirov, G., Petersen, P. & Qian, L. Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature* **552**, 67–71 (2017).
49. Marchi, A. N., Saaem, I., Vogen, B. N., Brown, S. & LaBean, T. H. Toward Larger DNA Origami. *Nano Lett.* **14**, 5740–5747 (2014).
50. Ong, L. L. *et al.* Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature* **552**, 72–77 (2017).
51. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
52. Sobczak, J.-P. J., Martin, T. G., Gerling, T. & Dietz, H. Rapid Folding of DNA into Nanoscale Shapes at Constant Temperature. *Science* **338**, 1458–1461 (2012).
53. Tang, G. *et al.* EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology* **157**, 38–46 (2007).
54. Abramoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ. *Biophotonics international* http://localhost/handle/1874/204900 (2004).

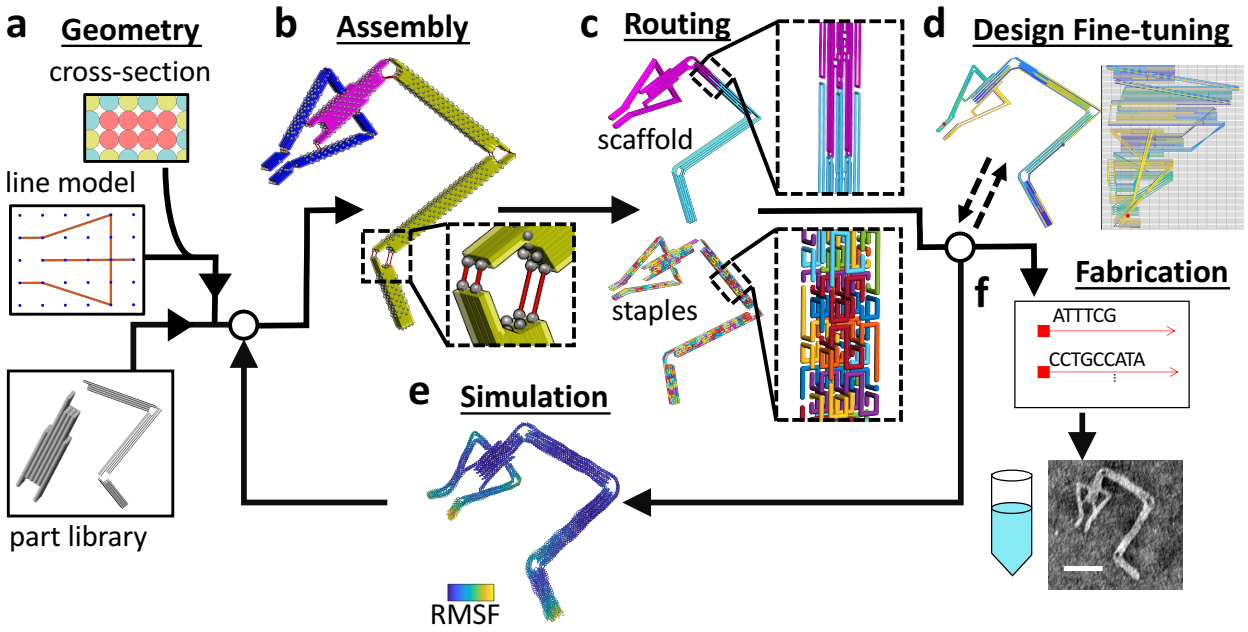

Figure 1. Schematic of the proposed design framework for multi-component DNA origami assemblies. a, To define the initial overall geometry users can take a top-down approach using a line model (.STEP file or MagicDNA sketch GUI) and specifying the length and cross-section of each line to create a full cylinder model of the assembly. Alternatively, individual or groups of components can be imported from a part library to build up assemblies. **b,** For assembly, each component can be subjected to translation or rotation to arrange the desired configuration. Users can specify connectivity between components manually or specify the number and type (e.g. end-to-end, end-to-side) and allow the program to automatically search for the closest sites (potential connection sites indicated by gray dots). **c,** Routing of scaffold and staple strands is automated including the capability to incorporate multiple scaffolds. **d,** Details of the strand routing can be visualized in a 3D structure and 2D diagram, and there is a two-way interface with the software caDNAno[6] for fine modification of routing. **e,** Input files for simulation in oxDNA[10] are automatically generated for virtual prototyping with built-in analysis including calculating the average shape and root-mean-squared fluctuations (RMSF). **f,** Once desired design metrics are achieved, the corresponding DNA sequences are automatically generated for fabrication and verification as shown by TEM. Scale bar = 50 nm.

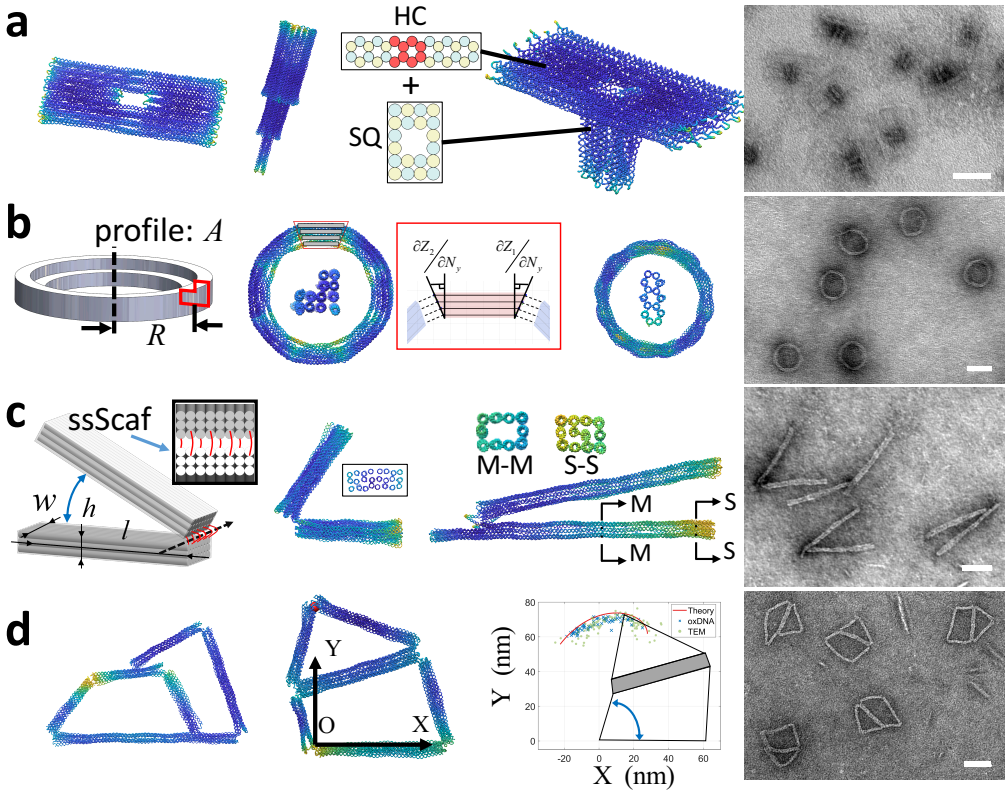

Figure 2. Parametric design of functional nano-devices. Structures depicted are the average oxDNA configuration with color-coded RMSF values. **a,** Horizontal, vertical, and 3D hybrid lattice nanopore structures are presented from left to right. **b,** The nano-ring series are approximated by a polygon of straight bundles with gradients at the ends. **c,** The nano-hinge devices are formed by two stiff arms joined by ssDNA scaffold connections to form a flexible rotational joint. **d,** The linkage designs implement multiple hinge joints to achieve a desired motion path. The simulated motion closely matches the experimental data. TEM images illustrate well-folded structures with high yield. Scale bars = 50 nm.

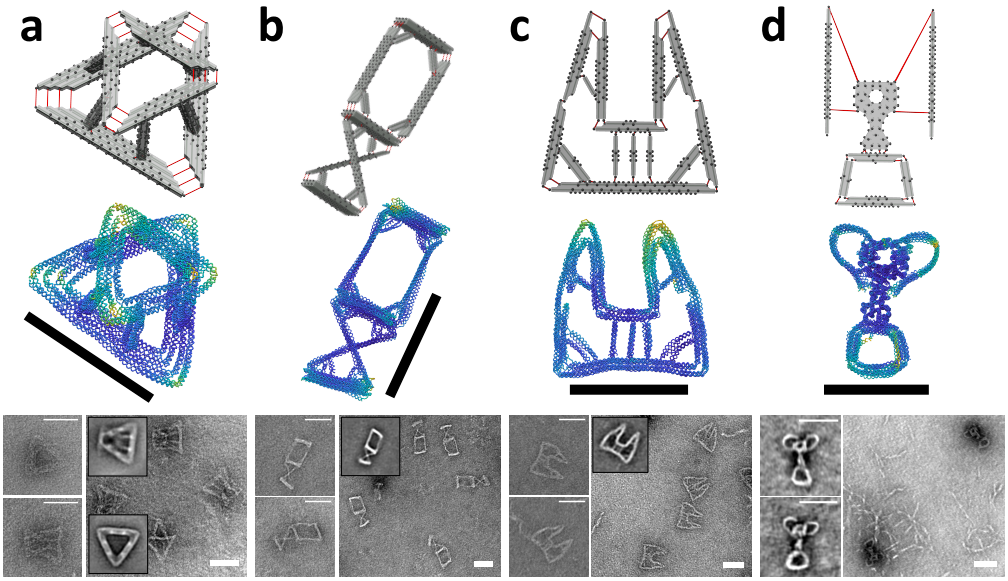

Figure 3. Design of multi-component complex structures. The first row shows assembly models from MagicDNA illustrating potential connection locations (black dots) and joint connections (red lines). The second row shows the average structure from oxDNA simulations with color-coded RMSF. The third row shows TEM images with inset image averages (except for trophy due to low yield, Supplementary Figure 43). **a,** The Stewart platform consists of top and bottom triangular plates and six 2×2 square-lattice connecting limbs. **b,** The compound joint incorporates a compliant sliding joint on top and a compliant rotational joint on bottom with vertices on both joints reinforced with struts. **c,** The gripper has 15 square-lattice bundles with several sharp corners and seven struts to reinforce the overall shape. **d,** The trophy consists of a 62-helix bundle in the center, two single-layer square-lattice bundles, and four double-layer bundles for the base. Connections between the two "handles" and the central component were manually assigned to create the curved shapes. Scale bars = 50nm.

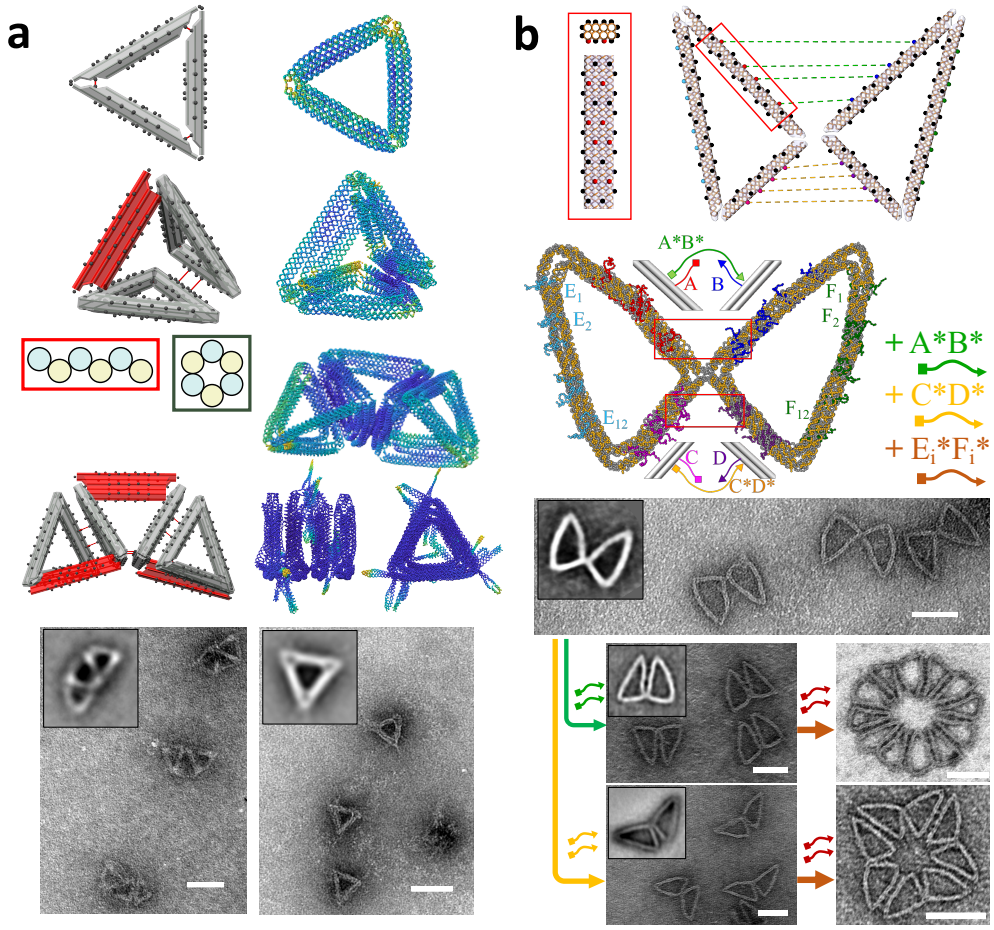

Figure 4. Reconfigurable devices by hierarchical design. a, A deployable mechanism formed by a serial chain of three tetrahedrons. From top to bottom: a triangular plate validated with oxDNA, a tetrahedron obtained from duplicating the triangular plate with an extra blade component (red), and a serial chain of tetrahedral obtained from duplicating the verified tetrahedron. The verified tetrahedral in a serial chain was further validated in deployed and compact configurations with TEM images. Insets show image averages. **b,** The butterfly mechanism is made of two triangles connected by a hinge joint. There are overhangs on the upper, lower, and outer edges to actuate into different configurations and polymerize into distinct circular assemblies. Scale bars = 50nm.

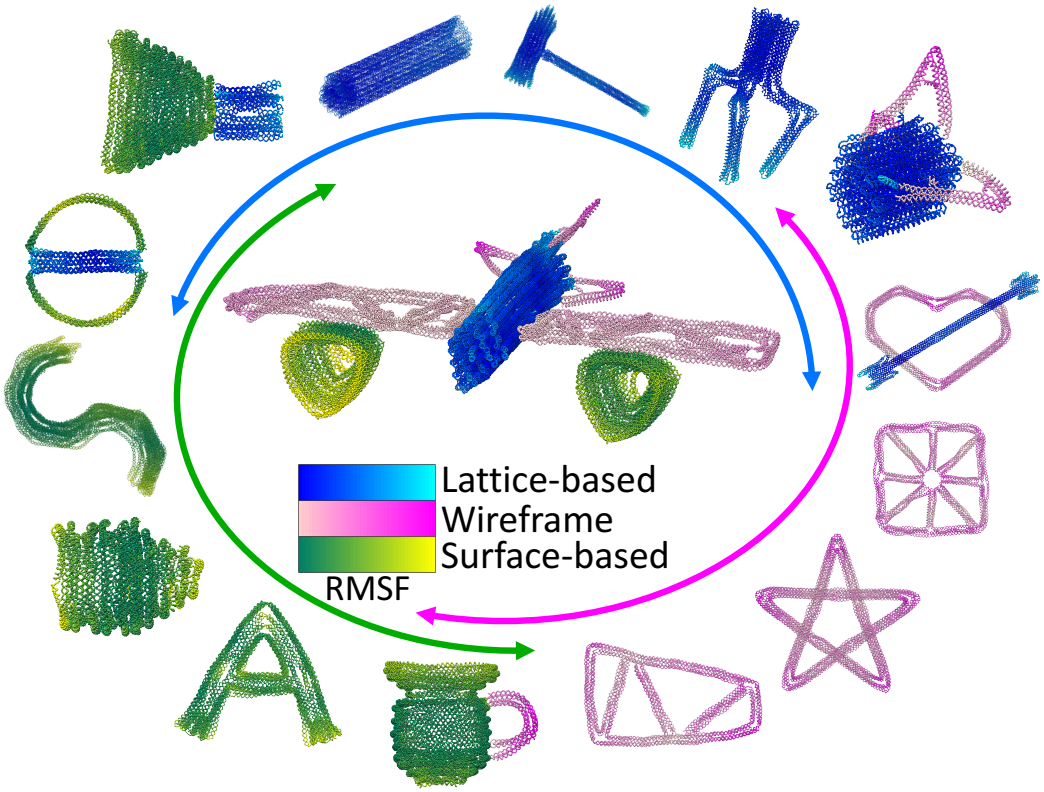

Figure 5. Broadening the design spectrum by integrating wireframe, lattice, and surface-based components. The three classes of geometric modeling are supported by various built-in GUIs in MagicDNA. Stiff lattice-based components (relative RMSF shown blue to cyan) are made by assigning an initial geometry to lines (top-down) and extruding helices in the bundle editor GUI (bottom-up). Surface-based (or shell) components (relative RMSF shown green to yellow) are designed using multiple segments with end gradients to approximate features with curvature. The wireframe models (RMSF shown pink to magenta) are similar to lattice but with small cross-section (e.g. 2×2) and many can be easily connected in space. The airplane (~33 kbps) in the center exploits all three types of geometric modeling, including wireframe wings, surface-based turbine, and lattice-based fuselage.

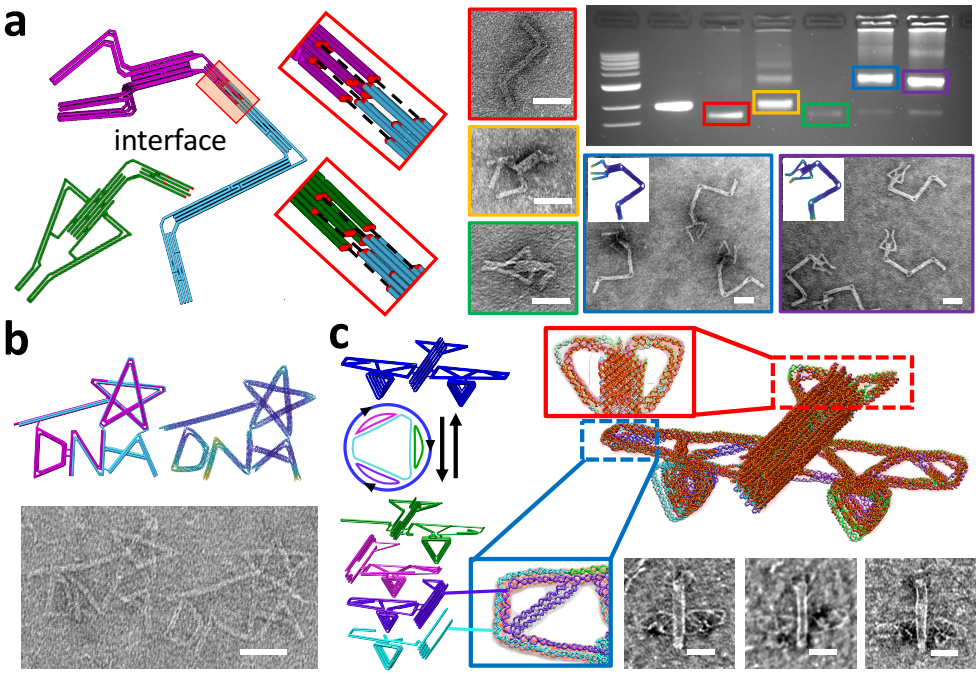

Figure 6. Multi-scaffold designs. a, One of two multi-scaffold design methods to achieve devices with interchangeable parts. The two-scaffold robot arm design has an exchangeable End-Of-Effector, claw (magenta) or tweezer (green), and the insets show the zoom-in of the user-defined interface specified by adding crossovers across the interface and then forming a single scaffold cycle on either side of the interface. Individual and combined structures were folded, each in a single-pot reaction, and validated by gel electrophoresis and TEM. The lanes in the gel are: 1kb DNA ladder, p8064 scaffold, robot arm individual structure, claw individual structure, tweezer individual structure, robot arm with claw multi-scaffold structure, and robot arm with tweezer multi-scaffold structure. **b,** The second multi-scaffold approach applies *K*-1 crossovers to split a single cycle scaffold into *K* cycles, as demonstrated here for the MagicDNA logo where *K*=2 and the TEM image of the logo for experimental validation. **c,** In the left, this approach was used for the airplane to add three crossovers to split the original scaffold into four cycles (*K*=4). The final airplane design comprises 4 orthogonal scaffolds with a total of 682 staples. 462 of these staples connect at least two scaffolds (depicted as red staples, and staples that bind to a single scaffold are grey and transparent), showing most of the structural components are populated by staples that connect at least two scaffolds. A single-pot folding with these scaffold and staple strands to fold the structure and the TEM images show the formation of this airplane in nanoscale. Scale bars = 50nm.

Methods

Multi-component assembly design and software availability

MagicDNA is an open-source software available at https://github.com/cmhuang2011/MagicDNA. It was coded in MATLAB 2017a and is compatible with newer versions of MATLAB. Detailed descriptions for installing the software are in the software user manual. Additional material including tutorial movies can be accessed through the Supplementary Information or the YouTube channel "MagicDNA software". To broaden accessibility of the software, we also compiled MagicDNA into standalone MATLAB Runtime versions in Windows, Linux, and Mac platforms. Output files from MagicDNA include the staple sequence list for ordering staples, caDNAno .JSON files for fine-tuning of strand routing details, and oxDNA topology and configuration files for validating the design profile with coarse-grained simulation.

Typical Design workflow in MagicDNA

The general design workflow consists of four steps (Supplementary Figure 3): 1) Define the overall geometry and the geometries of each component, 2) Assemble the components by forming stiff or flexible joints between them, 3) Use the routing algorithms and fine-tune the routing if necessary, and 4) Generate topology and configuration files for coarse-grained simulations. For the top-down approach in geometry, either sketching lines in MagicDNA or importing a line model through .STEP files is needed to convert lines to bundles with also inputting design parameters like cross-section and lengths. Alternatively, one can remove or insert components to the assembly using a bottom-up approach. The next assembly step includes manipulations of each component or a set of grouped components to arrange a desired 3D assembly configuration, connecting the components by specifying the connectivity matrix and/or using the

optional manual mode to specify locations of connections between components, and finally specifying the single-stranded scaffold lengths. Once the routing algorithms receive the design parameters from the geometry and assembly steps, the scaffold and staple routings are automatically generated with the option of fine-tuning in caDNAno[6]. Lastly, using the automatically generated simulation files to conduct the coarse-grained simulations allows users to evaluate the design and provides feedback to guide modifications in the next iteration if needed. This design process is illustrated in detail for a hinge example in Supplementary Movie 1.

Coarse-grained MD simulation

The topology and initial configuration files were generated directly from MagicDNA. The relaxation algorithm was similar to our previous study[38] adapted from standard oxDNA relaxation protocols[35]. The relaxation is carried out in three steps: oxDNA1, oxDNA2 relaxations with gradually increasing coefficients, and a short simulation, all with mutual traps between paired scaffold and staple bases. After relaxation, the oxDNA2 interaction model was used to conduct coarse-grained simulations without applying any mutual traps. For most simulations, a total of 10^7 steps with GPU acceleration were used. For the 4-bar mechanism, 3×10^8 steps were used to get a better depiction of the motion. The simulation time for each step was set to 15.15 fs. Simulation parameters included an Anderson-like thermostat, temperature at 30 °C, and monovalent salt concentration at 0.5 M, all standard conditions in oxDNA simulations[10,11]. The frequency to save the current configuration into the trajectory was set as either 10^6 or 5×10^5 steps. The processes mentioned above were executed through a shell script for all structures in this study in a Linux computer equipped with a NVIDIA GeForce 1080Ti graphics card. The trajectory file was later analyzed in MATLAB, including visualization of configurations, root-mean-square deviation

(RMSD), and root-mean-squared fluctuations (RMSF). The average configurations were exported to the UCSF Chimera[51] software and rendered to high-quality images.

Assembly and Fabrication of DNA Origami structures

All DNA staple strands were ordered and synthesized with salt-free purification and in 10 nmole scale from Eurofins (Louisville, KY), except the staple strands of the ring structure in 25 nmole scale from IDT (Coralville, IA). Scaffolds for single-scaffold structures were made in-house as described in[32] or purchased from Guild Biosciences (Dublin, OH) for M13mp18 derived scaffolds. Scaffolds for multi-scaffold structures were kindly provided by the Dietz lab at Technische Universität München[39]. Each structure was folded (thermal cycler from Bio-Rad, Hercules, CA) and optimized for solution conditions (i.e. salt, scaffold, and staple concentrations), annealing ramp protocol, and in some cases gel running conditions (i.e. salt concentration in gel). Single-scaffold structures were folded with 200 nM staples and 20 nM scaffold. Multi-scaffold structures and single-components of multi-scaffold structures were folded with 110 nM staples and 10 nM scaffold or folded with 100 nM staples and 10 nM scaffold. Each folding reaction contained a buffer solution consisting of 5 mM Tris, 5 mM NaCl (pH 8), 1 mM EDTA, and varying $MgCl_2$ conditions found in the respecting supplemental figure captions. Folding conditions varied by structure, and specific details for all structures are provided in Supplementary Table 1. Thermal annealing ramps were also tailored for individual structures (details also in Supplementary Table 1). The different annealing ramps used included a two-and-a-half-day fold[32] starting with a 1 hr/°C from 65-61°C melt, followed by, 2 hr/°C from 60-40°C anneal, and a cool step 30 min/°C from 39-4°C ; a four-and-a-half-day fold starting with a 1hr/°C from 65-62°C melt, 2hr/°C from 61-59°C anneal, 5hr/°C from 58-46°C anneal, 2hr/°C 45-40°C cooling , and a final cooling step

1hr/°C from 39-4°C. Single-scaffold structures folded in rapid folds[52] all include a 15min 65°C melt: then an anneal 4hr/°C in non-linear increments from 60-40°C , and a 4hr/°C from 56-50°C anneal. The multi-scaffold structures folded using an annealing protocol described by Engelhardt et al.[39] starting with a 65°C melt for 15 minutes, followed by an anneal 3hr/°C from 60-40°C then a cool at 4°C.

Purification of DNA Origami

Each DNA origami structure was purified and analyzed post-folding reaction via agarose gel electrophoresis. Buffer conditions included 0.5x TBE (45 mM Boric acid, 45 mM Tris base, and 1 mM EDTA) with either 5.5 mM or 11 mM $MgCl_2$ and agarose gels from 1.5-2% agarose and 0.5μg/mL ethidium bromide. 1.5% agarose gels with 0.5x TBE and 5.5 mM $MgCl_2$ buffer[39] were used for all multi-scaffold DNA origami structures and components as well as the umbrella closed configuration and trophy. All other structures were purified with 2% agarose gels and 0.5x TBE and 11 mM $MgCl_2$ running buffer. Each gel was run at 90V for 90-120 minutes in an ice water bath. Gels were imaged on a UV table using a FotoDyne Express FOTO/Analyst system. Details for gel purification are also summarized in Supplementary Table 1.

Actuation and polymerization of DNA Origami

The butterfly mechanism was actuated post-fold and gel purification. The structure concentration was quantified via Nanodrop as ~3 nM. 10μL of gel-purified structure was then combined with 2μL of actuation staples for a final concentration of 2.5 nM structure, 25 nM actuation staples and 10 mM $MgCl_2$ (10x excess concentration of actuation staples relative to the concentration of the structure). The mixture was incubated at 37°C for 2 hours. After actuation of

structures, polymerization staples were added at 150 nM. The final solution contains 10μL of structure at 2 nM, 2 μL of actuation staples at 21 nM, and 2 μL of polymerization staples at 21 nM and ~8 mM $MgCl_2$. The solution was then incubated again at 37C for 2 hours.

DNA Origami Analysis and Imaging via TEM

Structures were suspended in respective running buffer conditions post purification with concentrations between 1-5 nM depending on structure yield. The trophy and Stewart platform structures (see Supplementary Table 1) were additionally incubated with the peptide K10 (kindly provided by the Stephanopoulos Lab at Arizona State University) at a ratio of 0.5N:P for ~30 mins prior to preparing TEM samples to improve contrast[42]. A sample volume of 4 μL was deposited onto a plasma-cleaned Formvar-coated 400 mesh copper grid (Ted Pella, Inc.) with incubation times between 4-8 minutes prior to wicking away the solution with filter paper. For the trophy, umbrella closed configuration, robotic manipulator, logo, and airplane structures (see Supplementary Table 1), a 4μL droplet of 30 mM $MgCl_2$ was added to the plasma-cleaned grid prior to sample incubation and wicked away after 2 minutes followed by adding the sample drop to enhance surface deposition. After wicking away the sample drop, a 10μL droplet of staining solution consisting of 2% uranyl formate + 25 mM NaOH was added to the grid, immediately wicked away, followed by adding a 20μL droplet of the same staining solution incubated for 40 seconds and finally wicking away the stain solution. Samples were allowed to dry for at least at least 20 minutes before imaging. The structures were imaged at the OSU Campus Microscopy and Imaging Facility on a FEI Tecnai G2 Spirit TEM with an acceleration of 80kV.

EMAN2[53] and ImageJ[54] were used for post-processing and analysis of gel images and raw TEM images. We use gel intensity analysis to estimate yields of well-folded structures. Specifically, we

used ImageJ to collect the intensity of the well-folded band, subtracted the background signal, and then divided by the intensity of the entire lane (background also subtracted from the entire lane signal) to calculate the yield of well-folded structures. For multi-scaffold structures, the gel intensity analysis gives an approximation of the total mass of DNA corresponding to the well-folded structure. The yield estimates are included in the captions for each gel in the corresponding supplemental figures. Old Particle Picker (e2boxer_old.py) in EMAN2 was used to select particles from raw TEM TIF files. At least 300 particles and up to 900 particles were used to create particle sets for image averages. Particle sets were then built and 2D analysis with 4 ncls (number of classes) and 3-8 iterations were performed for image averaging. Particles from the 4-bar mechanism were used separately in a MATLAB code for a 5-point analysis with manual selection. ImageJ set scale function was used for scale bars on TIF files and brightness/contrast/FFT bandpass filters were applied in ImageJ analysis.