Uncertainty Quantification of a DNA Origami Mechanism Using a Coarse-Grained Model and Kinematic Variance Analysis

Journal:	*Nanoscale*
Manuscript ID	NR-ART-08-2018-006377.R1
Article Type:	Paper
Date Submitted by the Author:	28-Oct-2018
Complete List of Authors:	Huang, Chao-Min; The Ohio State University, Mechanical and Aerospace Engineering Kucinic, Anjelica; The Ohio State University, Chemical and Biomolecular Engineering Le, Jenny; The Ohio State University, Biophysics Graduate Program Castro, Carlos; The Ohio State University, Mechanical and Aerospace Engineering; The Ohio State University, Biophysics Graduate Program Su, Haijun; The Ohio State University, Mechanical and Aerospace Engineering

SCHOLARONE™
Manuscripts

Uncertainty Quantification of a DNA Origami Mechanism Using a Coarse-Grained Model and Kinematic Variance Analysis

Chao-Min Huang[1], Anjelica Kucinic[2], Jenny V. Le[3], Carlos E. Castro[1, 3*] and Hai-Jun Su[1*]

[1]Department of Mechanical and Aerospace Engineering, [2]Department of Chemical and Biomolecular Engineering and [3]Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210, USA

Emails: castro.39@osu.edu & su.298@osu.edu

Abstract

Significant advances have been made towards the design, fabrication, and actuation of dynamic DNA nanorobots including the development of DNA origami mechanisms. These DNA origami mechanisms integrate relatively stiff links made of bundles of double-stranded DNA and relatively flexible joints made of single-stranded DNA to mimic the design of macroscopic machines and robots. Despite reproducing the complex reconfiguration of macroscopic machines, these DNA origami mechanisms exhibit significant deviations from their intended motion behavior since nanoscale mechanisms are subject to significant thermal fluctuations that lead to variations in the geometry of the underlying DNA origami components. Understanding these fluctuations is critical to assess and improve the performance of DNA origami mechanisms and to enable precise nanoscale robotic functions. Here, we report a hybrid computational framework combining coarse-grained modeling with kinematic variance analysis for predicting uncertainties in the motion

pathway of a multi-component DNA origami mechanism. Coarse-grained modeling was used to evaluate the variation in geometry of individual components due to thermal fluctuations. This variation was incorporated in kinematic analyses to predict the motion pathway uncertainty of the entire mechanism, which agrees well with experimental characterization of motion. We further demonstrated the ability to predict the probability density of DNA origami mechanism conformations based on analyses of mechanical properties of individual joints. This integration of computational analysis, modeling tools, and experimental methods establish a foundation to predict and manage motion uncertainties of general DNA origami mechanisms to guide the design of DNA-based nanoscale machines and robots.

Introduction

While macroscopic machines require energy input to achieve their desired function, biological nanomachines, such as motor proteins, exploit a combination of thermal fluctuations and energy to perform mechanical functions through conformational changes[1,2].Supramolecular[3–5] and DNA-based systems[6–11] have emerged as powerful approaches to mimic aspects of these exquisite natural or macroscopic machines. In particular, scaffolded DNA origami nanotechnology[12,13] has been used to reproduce the geometry; designed 1D, 2D, and 3D motion; and triggered reconfiguration of macroscopic machines by incorporating flexible motion into nanodevices through strategic integration of single-stranded DNA (ssDNA)[7,14–16], compliant components[8,17], or relative sliding or rotation of complementary geometries[7,9,18,19]. However, in contrast to macroscopic machines, these DNA origami mechanisms (DOMs) function at length and energy scales similar to biological motors where thermal fluctuations not only drive motion, but can also cause variations in the geometry of the underlying components. These fluctuations can lead to significant deviations, or

uncertainty, in the designed motion behavior[7,20,21], which is critical to understand and manage to improve device functions. The goal of this work is to understand and establish an integrated computational, theoretical, and experimental framework and user-friendly tools for predicting motion, including uncertainty and conformational distributions, as a critical step to assessing mechanical performance of DOM and enabling designs with improved motion fidelity and dynamic function.

A number of computer-aided design (CAD) tools have been developed to facilitate[22–24] and automate[25,26] the DNA nanostructure design process. For the scaffolded DNA origami[12] approach, cadnano[23] is the most widely used design software. In addition, computer-aided engineering tools have been developed for simulating the shape of designed DNA nanostructures. For example, CanDo[27,28], a finite element modeling framework developed by Bathe and co-workers, enables rapid prediction of the folded shape of DNA origami structures. Pan et al.[29] later extended CanDo's scope to lattice-free designs by adapting it to Tiamat. These tools primarily focus on designing or predicting shapes of static nanostructures.

Recently, structural DNA nanotechnology has demonstrated tremendous potential for engineering dynamic nanomachines[7,8,15,18], which has led to a demand for more advanced modeling tools. Pan et al. introduced Brownian dynamics[30] to capture time-dependent properties, such as trajectories, of structures predicted by CanDo. However, this approach does not capture some molecular scale design features that may influence dynamic behavior. Compared to continuum mechanics-based methods, atomistic molecular dynamics (MD) provides a potential solution for DNA origami, up to ~15k bp in size, by modeling detailed structures and properties such as ionic conductivity[31–34]. However, the high computational demand of a full atomic MD model prevents studies of dynamics beyond nanosecond timescales, which are critical to

understanding and improving device designs. Coarse-grained models[35–37] have been developed to speed up the computational process. In general, depending on the level of simplification, fewer particles or rigid bodies reflect a faster but less accurate algorithm. Among many coarse-grained models, oxDNA[38–42]was developed with a focus on modeling designed DNA nanostructures. In the oxDNA model, the basic unit is a nucleotide base, which is treated as a rigid body with three sites of non-collinear interactions: hydrogen bonding, base stacking, and covalent backbone bonds. Recently, some dynamic DNA origami structures were studied using oxDNA[20,43,44], including compliant hinges, revolute joints, sliders, crank-sliders and Bennett linkages. However, even though GPU acceleration has reduced the computational demand by one to two orders of magnitude[45], it is still difficult to completely capture the motion of a DOM. Furthermore, this incomplete capture would create difficulties in iterating the design and simulation processes to optimize designs for motion fidelity, especially for mechanisms with large configuration changes. In addition to +computational demand, salt-dependent properties in the oxDNA model are based on monovalent sodium cations, but most DNA origami experiments are conducted in the presence of divalent magnesium cations.

To address these challenges, this work seeks to establish integrated computational, theoretical, and experimental methods into efficient tools to model and predict the motion behavior of complex DOMs and ultimately guide future designs. We chose to use oxDNA as the basis of our modeling. To address the computational demand limitations, we integrated theoretical kinematic analysis and Monte Carlo simulation approaches. Furthermore, to validate our models and address salt-dependence limitations, we integrated experimental analysis of DOM with feedback into the modeling. The particular focus of this work was to study thermally driven motion of DOM and understand what design parameters govern deviations from the motion pathway (i.e. motion

uncertainty) and conformational probability distributions along key reaction coordinates. We investigated the sources of motion uncertainty from the underlying components, namely the links, which are usually bundles of double-stranded DNA (dsDNA), and the joints, which usually comprise one or several single-stranded DNA (ssDNA) connections. While links are relatively stiff, they may undergo some changes in shape due to thermal fluctuations, especially fraying at their ends[28]. The joints may also exhibit motion in unintended rotational or translational degrees of freedom due to fluctuations within ssDNA connections.

To develop and test our modeling framework, we constructed a new multi-component DOM that integrates 4 links (each made from 8 bundles of dsDNA) connected by 4 rotational joints. This mechanism is called a straight-line linkage (SLL) as shown in Figure 1 since it is designed to closely approximate a straight-line motion at a tip point of two rigidly connected links. We first developed an approach to speed up the computational relaxation process for multi-component designs. We then integrated coarse-grained MD simulation results with kinematic variance analyses to create a hybrid computational tool to predict and study the motion behavior. Experimental characterization of the energy landscapes of the mechanism and the joints was used to investigate the salt-dependent aspects of motion that govern the conformational distribution. Lastly, we used toehold-mediated actuation to steer the probability distribution and verified our hybrid tool again to predict the uncertainty in the actuated conformation. We also applied this computational framework to another DOM, the crank-slider[7], to demonstrate the general applicability.

Experimental

DNA Origami Structure Design and Fabrication. We designed DNA origami structure in the computer-aided design software cadnano[23]. The Straight-Line Linkage (SLL) is based on a 7249-

nt single-stranded scaffold originating from the M13mp18 bacteriophage virus that was purchased from Guild Biosciences and then prepared following previously described protocols[27]. The oligonucleotide staples were ordered from a commercial vendor (Eurofins, Huntsville, AL). The self assembly reaction follows previously described protocols[27]. The folding reactions contained 20nM scaffold and 200 nM of each staple within a solution consisting of ddH_2O, 5mM Tris, 5 mM NaCl, 1 mM EDTA, and 20 mM $MgCl_2$. An $MgCl_2$ screen revealed that a $MgCl_2$ concentration higher than 14 mM yielded well-folded structures. The folding reaction occurred in a 2.5 day thermal annealing ramp based on a previously published method[27] in a thermal cycler (Bio-Rad, Hercules, CA).

Agarose gel electrophoresis was used to characterize these folding results, as well as to purify out excess staples. Well-folded structures from purified bands were characterized and imaged as negatively stained samples on the Transmission Electron Microscope (TEM).

Sample Analysis and Imaging using Transmission Electron Microscopy. Gel-purified SLL structures were imaged on a TEM for quantification and validation of well-folded structures. A sample volume of 4 µL was deposited onto a Formvar-coated copper grid, stabilized with plasma-cleaned carbon film (Electron Microscopy Sciences, Hatfield, PA). The sample was incubated onto the grid for 4 minutes, and then wicked away. 10 µL of 2% uranyl formate (SPI, West Chester, PA) stain with 25 mM NaOH was immediately added to the grid and wicked away. Twenty µL of the same stain was added, incubated for 40 s, and then wicked away. TEM took place at the OSU Campus Microscopy and Imaging Facility on a FEI Tecnai G2 Spirit TEM at an acceleration of 80 kV at a magnification of 100 000X. MATLAB was used to quantify and process the raw TEM images. The TEM images were organized into galleries of individual SLL structures for further quantification and analysis. A 5-point analysis of the SLL was carried out using a custom

MATLAB code to measure the kinematic motion of the tip position P_3 (Figure 1). Five points were manually selected to obtain lengths and angle measurements for analysis of P_3 motion.

Preparation of the Actuated Straight-line Linkage. The SLL was actuated with ssDNA staples complementary to overhangs connected to links 2, 4, 5, and 6. Folded samples of the SLL with overhangs were purified *via* polyethylene glycol (PEG) centrifugation using a protocol modified from[46] after a 2.5 day folding reaction. The well-folded SLL with overhangs were mixed in a 1:1 volume ratio with 15% PEG8000 (Sigma Life Science) in a 200 mM NaCl solution. The mixture was centrifuged at 16000g for 30 min to pellet the nanostructures. The supernatant was removed and the structures were resuspended in buffer containing 10 mM Tris, 1 mM EDTA, and 20 mM MgCl$_2$. The motion of the tip position P_3 to the left or right (marker-side) is actuated by combining purified SLL structures at 5 nM with the overhang sites and a concentration at 500nM of the complementary ssDNA staples. The solution was incubated at 37°C for 24 hours. Once incubation was complete, the actuated SLL was gel-purifiedand was characterized via TEM.

Preparation of Open-chain Straight-line Linakge. The open-chain SLL was folded with a protocol similar to the closed-loop SLL, however, the staples connecting link 4 were not included in the folding process, purification, or characterization.

Preparation of the Straight-line Linkage Configurations for Molecular Dynamic Simulations. The cadnano design of the SLL was converted to the initial configuration through the python code cadnano_interface.py, provided by the oxDNA group (https://dna.physics.ox.ac.uk), in which all bundles were parallel to each other as shown in Figure 2. Due to forced connections present in cadnano, this initial configuration had many over-stretched bonds across adjacent bundles causing failed starts of the MD simulations (Figure S2). To address

this issue, we developed a custom MATLAB code to enable "virtual assembly" of the mechanism. This code allows the user to manually apply rigid body transformations to the SLL link components or other multi-component structure. The code detects components based on the staple colors. Hence, uniform staple colors were assigned to all staples in a given link, and each link was assigned a different color using the graphic user interfaces (GUI) of the cadnano software. The rigid-body transformations were used to shorten the over-stretched bonds as well as to define the overall initial configuration for the MD simulations. Using the virtual assembly to achieve an appropriate configuration, the computational relaxation step was more efficient and prevented unwanted entanglement.

Next, we ran the relaxation algorithms originating from oxDNA packages. The goal of relaxation is to adjust all neighboring bonds on the backbone site to be within a certain range of the finite extensible nonlinear elastic (FENE) potential curve. This step is composed of two parts: oxDNA1 and oxDNA2 relaxation. Moreover, the latter gradually adapted the structure to have major and minor grooves. The coefficients of the relaxation algorithm were gradually increased to prevent the structure from breaking down. The entire relaxation step took ~20 minutes. To make sure that all staples were completely annealed with the scaffold strand, our code also exported a file indicating the mutual traps of paired staple and scaffold bases according to the cadnano file. This file pertaining to external forces is optional but recommended.

MD Simulation

The coarse-grained MD simulations were performed with the oxDNA2 package[47] without any external forces at a temperature of 303K. The total number of steps was set to 3×10^8 (for the SLL starting at the center configuration) or 1×10^8 (for the other five simulations). The simulation time step for integration was 15.15 fs and the Newtonian step of an Andersen-like thermostat was 103.

A scaling factor α ≈ 330 was used to convert the simulation time to the physical time[43]. The salt condition for all simulations in this paper was set to monovalent NaCl concentration as 0.5 M NaCl. To reduce the computation time, the oxDNA simulation was run with GPU acceleration[45] on a computer with a NVIDIA GPU. For example, a simulation with 1×10^8 steps required ~22 hours to complete the process. The UCSF Chimera[48] software was used to render all 3D coarse-grained models in this paper.

Free Energy Landscape

For the open-chain SLL experiment, the probability distributions of two open joints ϕ_j were fitted by combining three Gaussian distributions as

$$p(\phi_j) = \sum_{i=1}^{3} \frac{a_{ij}}{\sqrt{2\pi}\sigma_{ij}} \mathrm{Exp}\left[-\frac{(\phi_j - \mu_{ij})^2}{2\sigma^2_{ij}} \right] , j = 1 \text{ or } 2 \tag{1}$$

Once probability distribution function $p(\phi_j)$ was obtained, the free energy of each joint was written as follows with a symmetric assumption:

$$\Delta E(\phi_j) = -k_B T \ln\left\{ p(\phi_j) \right\} \quad j = 1,...,4 \tag{2}$$

where k_B is the Boltzmann's constant, T is the absolute temperature. By solving the kinematic equations, each joint angle as function of the input angle θ_3 (the coupler link angle relative to the ground link, Figure 1A) can be obtained as $\phi_j(\theta_3)$. With Eq (2), the probability distribution of input angle θ_3 can be calculated by the Boltzmann distribution again.

$$p(\theta_3) = \frac{1}{Z} \exp(-\frac{\sum_{j=1}^{4} \Delta E(\phi_j)}{k_B T}) \tag{3}$$

9

where Z is a constant to satisfy the condition that $\int_{-\infty}^{\infty} p(\theta) = 1$.

Results & Discussion

Design and Fabrication of the Straight-line Linkage (SLL)

To demonstrate our motion analysis, we consider the Robert's Straight-Line Linkage (SLL)[49] mechanism, shown in Figure 1A. This SLL is a four-bar linkage with four hinge joints (P_1, P_2, P_4, and P_5) and four links (l_1, l_2, l_3-l_5-l_6, and l_4). The macroscopic version of this SLL design has one degree of freedom (DOF) according to Kutzbach-Grübler's mobility equation[50], where the triangle l_3-l_5-l_6 is considered a single rigid body. The hinge joints P_1 and P_5 are anchored on the reference ground link, l_1. When a link is driven, the tip point P_3 traces an approximately straight line (shown in red), hence the name SLL.

The DNA origami SLL was first designed as a solid cylinder model (Figure 1B) with cylinders representing dsDNA helices. The detailed DNA routing and sequence design was performed in cadnano[23] (Figure S1). In this design, the rigid links, l_1 through l_6, consist of bundles of dsDNA helices arranged in a 2×4 square lattice[51] cross-section. Link l_3 (red bar) was designed with a bump feature to make the overall shape asymmetric and enable us to discern the left and right sides of the SLL mechanism. This is critical for proper motion analysis since the mechanism could land on either its front or back side in imaging assays. The four joints (P_1, P_2, P_4, and P_5), which function as revolute joints, are designed with two short and two long scaffold connections (Figure 1B, inset), where the short connections form the axis of rotation as previously established hinge designs[7,52].

We fabricated the SLL design using the scaffolded DNA origami self-assembly process[27] and purified well-folded structures via gel electrophoresis (Figure S33). Figure 1C shows a representative TEM image illustrating a range of SLL conformations with zoomed-in images

depicting three distinct configurations of the SLL. To quantify conformations, we manually selected points P_1 through P_5 in a custom MATLAB code, selecting P_1 as the lower hinge joint on the opposite side of the bump. We describe the conformation of the mechanism in terms of the location of the tip point, P_3, which undergoes the straight-line motion. First, we defined a local coordinate system origin on the midpoint of P_1 and P_5 to describe P_{3y} and P_{3x}, the vertical and horizontal position of the tip P_3. To quantify the motion performance of the SLL, we experimentally measured the position P_3 for 791 samples with respect to the local coordinate system and we depicted the area encompassing their distribution in gray (Figure 1D, more TEM snapshots in MovieS1 middle). We observed that the DNA origami SLLs reasonably follow the designed motion pathway (red line); however, there is significant deviation from the motion pathway of approximately ± 10nm, which is consistent with analysis of prior DOMs[7,21].

We have previously demonstrated the ability to design DOMs that exhibit a programmed motion pathway. Here, we aimed to enable quantitative predictions of two aspects of the motion, namely the uncertainty in motion (i.e. deviation from the designed motion pathway) and the conformational probability distribution. In the case of the SLL, we consider: (1) the motion uncertainty of the tip P_3 (shaded area in top part of Figure 1D) as a quantitative measure of the motion performance of the design, and (2) the probability distribution of the coordinate P_{3x} (bottom part of Figure 1D), which we consider a key reaction coordinate to evaluating the mechanical properties and enabling effective actuation of the device.
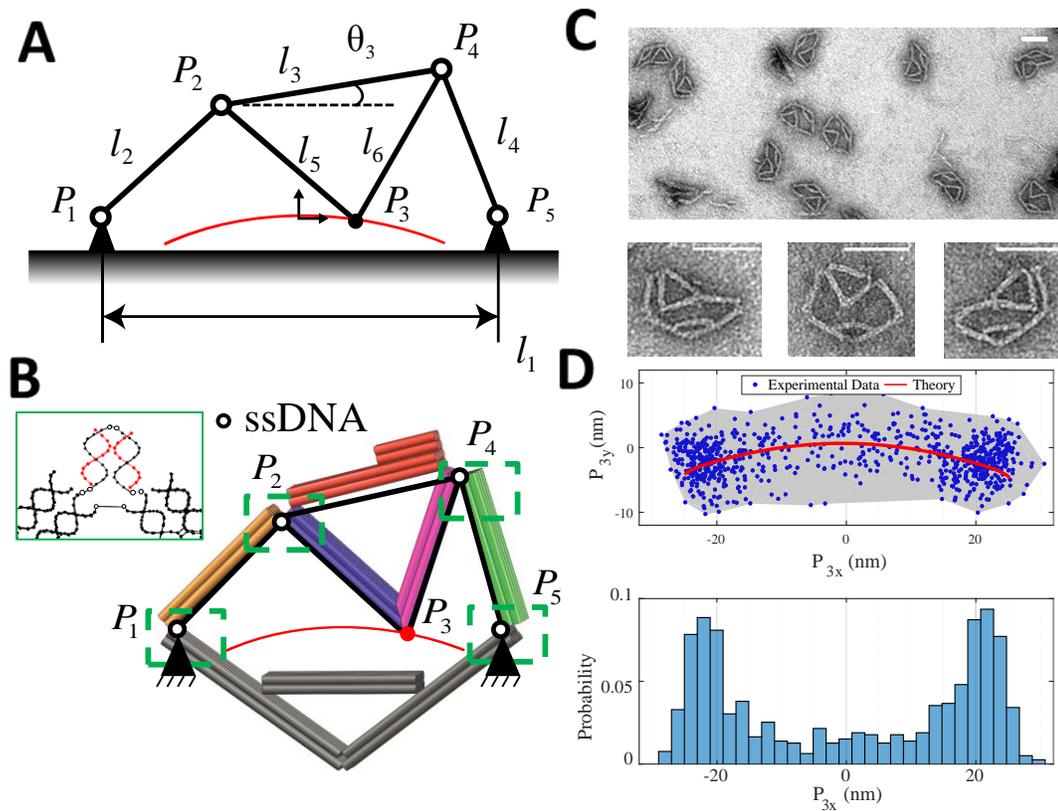
Figure 1: Design and fabrication of the DNA origami straight-line linkage (SLL). (A) Simplified line model for kinematic analysis of the SLL four-bar linkage with the corresponding tip trajectory (red). (B) The SLL consists of 8 square-lattice helix bundles with 4×2 cross sections. The strut on the ground link has a 3×2 cross section. (C) TEM images show the conformational range of the free SLL and zoomed-in TEM images of three representative conformations illustrate the motion pathway along the horizontal direction. Scale bars, 50 nm. (D) (Top) The experimental distribution of the tip position (blue points) deviates from the theoretical curve (red line) that assumes each bundle is a rigid body with perfect planar motion. (Bottom) The conformational distribution of the tip P_3 along the horizontal direction exhibits bimodal behavior.

Coarse-Grained Modeling and Results

To quantify thermal fluctuations, we started by building models of the SLL with existing DNA origami simulation tools. First, CanDo was used to simulate our design (Figure S1). The CanDo result accurately captures the overall geometry, but does not capture the motion distribution, likely because the finite element-based model does not consider aspects such as excluded volume, charge repulsions, base-stacking of blunt ends, and transient base-pairing of ssDNA, all of which could

play important roles in governing the motion. Hence, we focused on building a coarse-grained model that captures more molecular-level details using oxDNA.

oxDNA has been used to model static DNA origami structures[41,42] and more recently compliant[43] and dynamic[20] DNA origami devices. In general, these models undergo a relaxation step, which determines an initial geometry based on the design. This relaxation step can be challenging for multi-component structures that require significant rotation or translation of components of the original cadnano design (Figure 2A), where all helices are in parallel and may be spatially separated when compared to a real configuration (Figure 2B). In these cases, the oxDNA relaxation step may not converge to a feasible configuration (Figure S2) because the displacements required are too large.

To address this problem, we developed a workflow together with a dedicated computational tool that converts multi-component DOMs in the cadnano format (Figure S3-S7) into a coarse-grained model for simulation in oxDNA with a user-determined initial configuration (Movie S2). As shown in Figure 2A-2B, the cadnano routing was first converted to a coarse-grained model with all dsDNA bundles positioned and oriented as in the cadnano design. Then, a set of rigid-body transformations can be conducted on individual components using our computational tool (custom MATLAB code with user-friendly GUI) (Figure S4) to position all bundles into the desired configuration of the closed-loop mechanism. A few steps of these rigid body transformations are illustrated in Figure 2B. We term these transformations "virtual assembly", since the components are laid out and we are placing them into their appropriate position and orientation within the mechanism. The final assembled configuration is shown in Figure 2C. Next, an oxDNA relaxation algorithm (Figure 2D) and molecular dynamic simulation (Figure 2E and 2F) were performed to visualize and quantify conformational changes under thermal fluctuations. In addition to enabling

effective relaxation of the oxDNA simulation, the "virtual assembly" approach also allows for one to run multiple simulations with different initial conditions since bundles can be rearranged into various configurations along the SLL motion pathway (Figure S5). This is particularly useful, in combination with the ability to run different lengths of simulation time, to test whether simulations have converged to an equilibrium distribution of conformations and to probe the mechanics of specific regions of the conformational distribution, which may not be easily thermally accessible.

Here, we ran four simulations of the DNA origami SLL with different initial configurations. Three simulations with initial configurations of the tip P_3 positioned at left, center, and right were performed for ~500 μs and one simulation with the tip P_3 positioned at the center was performed for ~1500 μs. Figure 2G shows the simulation results of the motion of the tip P_3 in terms of the X- and Y-positions through time for the different initial configurations and simulation lengths. Movie S3 shows the SLL configurations for all trajectories in front-facing views.

The simulation results show the horizontal range of motion of P_3 to be roughly between -25 nm and 25 nm (Figure 2G), which is consistent with experimental data. However, despite reproducing some of the basic characteristics of motion, neither the individual simulations (Figure 2G) nor the compiled simulation results (Figure S8) capture several features of the experimentally measured motion including the deviation from the theoretical motion pathway (i.e. motion uncertainty) and the bimodal distribution, which are the primary features of interest for this study. Particularly, the experimental data shows a wider vertical distribution (Figure S8). The range of motion uncertainty (i.e. the deviation from the theoretical motion) observed in our experiments is consistent with prior DOM measurements[7] including measurements based on 3D reconstruction from cryo-EM imaging[21].

While the compiled simulation results capture the full range of motion, they do not capture the strong bimodal behavior we observe in experiments (Figure S8). In addition, conformational distributions from the individual simulations deviate significantly. For example, the simulations starting from the left and the right positions (the orange and yellow traces in Figure 2G) have no overlap. This indicates the simulations do not capture the full conformational distribution, likely due to insufficient simulation time. A recent study[20] required ~50ms simulation time for a two-component DNA origami hinge to converge to an equilibrium distribution. This length of simulation is intractable (our 500 μs simulations took ~22 hr on a standard workstation with GPU acceleration) without using clusters or supercomputers. In addition, the bimodal behavior, which suggests the presence of an energy barrier, could increase the time required for convergence to an equilibrium distribution. The longest simulation (1500 μs) starting from the center position (the purple trace in Figure 2G) explored did explore both sides of the horizontal motion distribution and exhibited some motion pathway deviations similar to our experiments. The compiled simulation results do not capture the bimodal behavior (Figure S8), but the three shorter simulations primarily explored one half of the horizontal motion distribution possible due to the presence of an energy barrier in simulations. This suggests running the simulation for significantly longer times may better capture the full motion distribution and our experimental motion uncertainty (and possibly the bimodal probability distribution of P_{3x}). However, to avoid running excessively long, intractable simulations, we developed an alternative approach that integrates analyses of motion uncertainty sources from the oxDNA model with Monte Carlo simulations and kinematic analyses. This integrated approach can be carried out more efficiently, since the sources of uncertainty come from conformational fluctuations of components, which converge to an equilibrium distribution more rapidly (Figure S9 and S10). Hence, either using compiled or

individual MD simulation data provides design insight into the sources of motion uncertainty within the design.
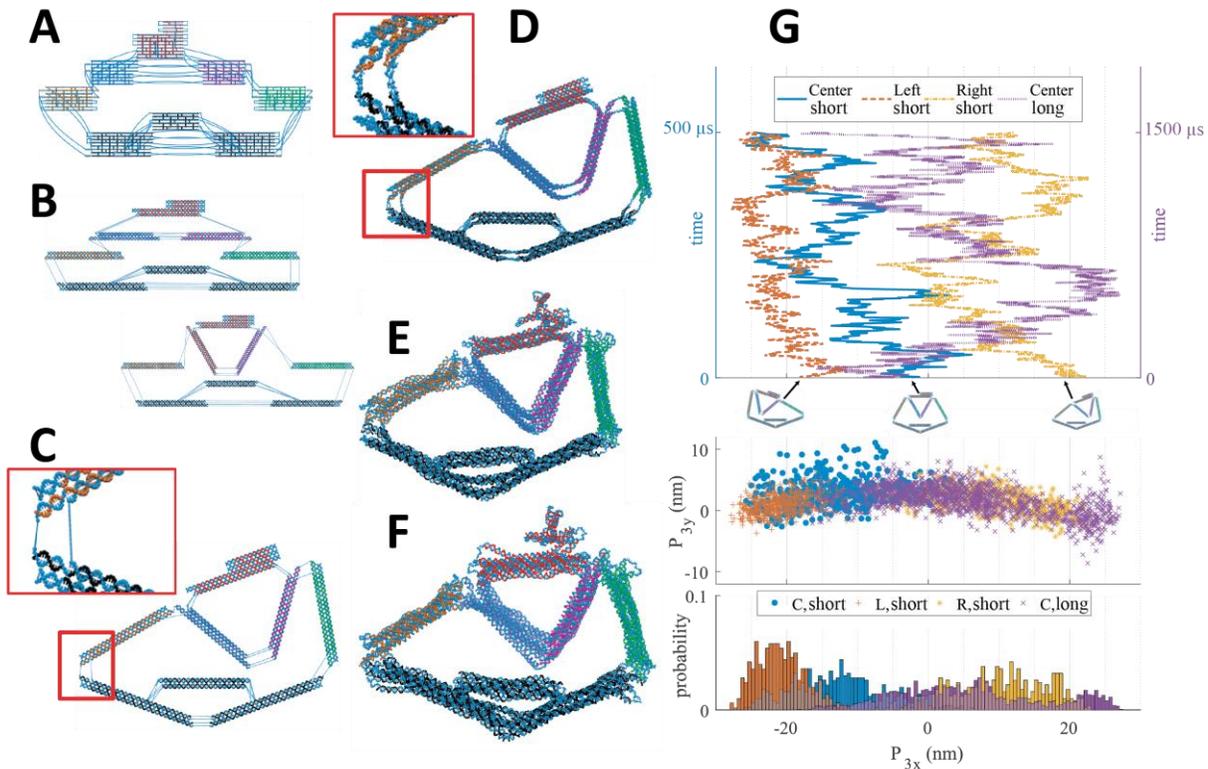

Figure 2 : Virtual DOM assembly and results of the coarse-grained molecular simulation. (A) The design diagram in cadnano software. (B) The initial configuration taken directly from the cadnano design yields an unrealistic initial configuration and topology. Rigid-body transformations are repeatedly applied to transform each bundle to an appropriately assemble initial configuration. (C) After rigid-body transformation, the configuration resembles the cylindrical model, although some bonds may still be overstretched (inset). (D) The relaxation algorithm gradually changes the configuration, including the major and minor grooves and distances between neighboring bases. (E) A fully relaxed configuration prior to starting the oxDNA simulation. (F) After completing the oxDNA simulation, the trajectory file is generated for visualization and post-processing. (G) Results of four repeated simulations for the tip position with different time lengths and initial configurations. Through rigid body transformations, three short ~500 μs simulations starting at the center, left, and right are plotted on the left y-axis on the top. The other long ~1500 μs simulation, corresponding the right y-axis, was shown to vaguely converge even at a longer time scale.

Characterization of Error Sources from oxDNA Simulations

We identified three potential sources of the motion uncertainty of P_3 from the oxDNA simulation results: out-of-plane motion (i.e. motion in unintended degrees of freedom), variations in link

geometry (i.e. changes in link end-to-end distance), and variations in joint geometry (i.e. variations in the geometry of ssDNA connections that form the joints). We first quantified each source of motion uncertainty and then integrated those variations into a kinematic analysis to develop an efficient approach for predicting the motion uncertainty of the overall SLL in terms of the position of P_3.

We introduced one input variable to specify the location along the theoretical motion pathway. Here we chose link 3 as the driver. Once we specify the orientation of link 3, the rest of the theoretical mechanism geometry can be determined by solving vector loop equations based on the geometric parameters of the closed-loop SLL mechanism. The basic framework of the vector loop analysis assumes rigid links with constant length. Here, we specified the dimensions of the links assuming 0.334nm/bp, which resulted in the theoretical motion curve depicted in Figure 1D.

Out-of-plane motion: We noticed some joints rotate out of the linkage plane in all simulations (Movies S4), which we refer to as out-of-plane motion. To quantify the error due to this out-of-plane motion, we built our kinematic model on the joints P_1, P_2, P_4, and P_5 and the tip P_3. In the coarse-grained simulations, we defined the positions of these joints as the center of mass (yellow spheres in Figure 3A and 3B) of the four single-stranded nucleotides that make up the short hinge connections (red points in Figure 3A and 3B). We tracked the 3D relative positions of these joints throughout all the simulations. The link geometries were then defined as the distance spanning two joints in the SLL mechanism. This analysis revealed that the links do not remain perfectly coplanar. Figure 3A shows a top view of a single frame of the simulation to illustrate this out-of-plane motion. The out-of-plane motion of the SLL is also visualized in Movie S4.

Since the links are not coplanar, it is non-trivial to define a reference plane for the mechanism. We used principle component analysis (PCA) to define a mechanism reference plane as a basis for

quantifying out-of-plane motion. The first and second eigenvectors, $\hat{e}_1$ and $\hat{e}_2$, from the PCA analysis represent the maximum position variance to the projected planes. Hence, the plane normal to the third eigenvector, $\hat{e}_3$, is where the SLL exhibits the maximum projected area as illustrated in Figure 3B, and this is the orientation where the SLL is likely to be observed on the TEM grid. Thus, we assumed the plane normal to the third eigenvector $\hat{e}_3$ as the projection and reference plane for the simulated DNA origami SLL. The tilted 3D vectors defining the links (blue cylinders in Figure 3A and 3B) were compared to this plane to quantify the out-of-plane motion. Quantitative results measured from oxDNA simulations are shown in Figure S9 in terms of the tilt angle of all the links relative to the reference plane. The longest, link 1, exhibits the least out-of-plane motion because its length is twice that of the others.

Variations in Link Geometry (End-to-end distance): Although the persistence length of the link components is expected to be much longer than their length[15], we still observed significant fluctuations of the bundle components in simulations (Figure 3B and Figure S10), especially at the ends of the bundles, which is consistent with prior simulation and experimental analysis[28]. This, combined with effects of staining, leads to the ends of dsDNA bundles being less resolved in TEM images, resulting in the link lengths being difficult to accurately measure experimentally. However, coarse-grained modeling provides a way to visualize the thermal fluctuations in the shape of these bundle components. Similar to out-of-plane motion, we tracked the 3D distances between the joints as a measure of variation in link end-to-end distance. The distributions of these link lengths are an indicator of variability in link geometries resulting from thermal fluctuations that occur even in well-folded nanostructures.

Since the TEM images are 2D projections of the nanostructures, it is unclear whether changes in the link dimensions in experiments are due to thermal fluctuations or due to the out-of-plane

motion. To address this challenge, we incorporated the influence of out-of-plane motion with the 3D end-to-end distances computationally. We first projected the 3D lengths of the six links (blue cylinders in Figure 3B inset) onto the reference plane (assumed as the largest projected area from principal component analysis) to obtain 2D-projected end-to-end distances (pink cylinders) for every frame in each simulation as shown in Movie S5. The 2D end-to-end distances, which now contained the effects of out of plane motion, were then separately fitted using the semi-flexible polymer[53] and the skewed Gaussian models. We found that the skewed Gaussian model fits better for these links, made of bundles of dsDNA (Figure S14 and S16). The 2D probability distributions of each link and the skewed Gaussian fits are shown in Figure 3C. All fitting parameters for each link are listed in Table S4.

Joint error: The hinge joints between any two links in our DNA origami SLL were designed as 2-nt ssDNA connections, similar to what we have used in our previous designs[7,52]. Although they are very short, it is possible that the length of these connections may fluctuate. To evaluate this fluctuation, we developed an error model for these hinges joints, illustrated in Figure 3D, in which dashed (black) and dotted (red) lines are the long and short ssDNA connections, respectively. Due to the highly flexible nature of ssDNA, these connections constrain the motion between two links primarily to rotate about an axis of rotation defined as a line between these two short connections (Figure 3D, right). However, the 2-nt length may also introduce additional clearance between links at every joint in the mechanism. To quantify the effect of this joint clearance error, we assumed that the lengths of the joint connections can vary anywhere from zero to a maximum of 1.2nm (0.6nm per ssDNA base) with a uniform distribution using Monte Carlo method.
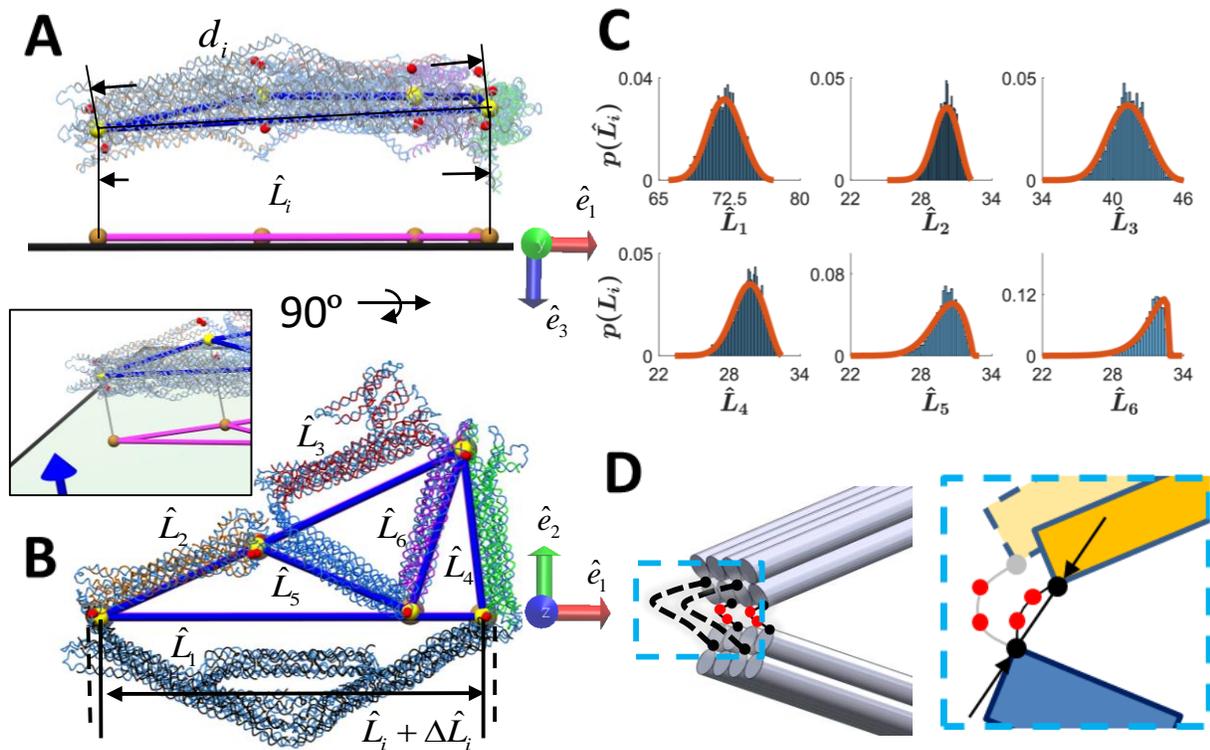
Figure 3: Sources of motion uncertainty for the DNA origami SLL. (A-B) The top and front views of a simulation snapshot show the out-of-plane motion and the fluctuating end-to-end distances. The red, yellow, and brown spheres in the ribbon model indicate the two 2-nt ssDNA connections, the arithmetic mean center in 3D space, and the center projection of each vertex, respectively. The blue and pink cylinders indicate the 3D line model and the projected links, respectively. (C) Probability distribution of the six projected link lengths and fitted curves using the skewed Gaussian distribution. Unit: nm. (D) The two 2-nt connections between layers of the two arms. The 2-nt connections constrain the two arms to a mostly rotational motion with a slight joint displacement error.

Computational Tools for Predicting DOM Motion Uncertainty

In macro-scale precision engineering, error analysis relates machine performance to several varying geometric factors influenced by manufacturing tolerances[54]. While uncertainty in component geometry at the macroscopic scale comes from variation in manufacturing processes, thermal fluctuations in our SLL and other DOM designs cause variations in component geometry that may lead to deviations in motion relative to the theoretical kinematic motion. To predict the motion pathway uncertainty of the SLL, we introduced two computational tools: (1) Taylor series

expansion with an optimization algorithm (TSOA) and (2) Monte Carlo simulation (MC). Both tools are based on the tip position's analytical solution, which in turn, is derived from solving vector loop equations using link 3 as the driver[50] (Figure S19). The TSOA is a deterministic method that estimates the boundary of the tip position with varying lengths assuming link lengths lie somewhere between the 5th and 95th percentiles of length distributions (Figure 3C). On the other hand, the MC method repeatedly calculates the output tip position by randomly sampling one set of the six link lengths from the distribution (Figure 3C). Full details of both methods are provided in section 4(b) and 4(C) in the Supporting Information.

For TSOA, we first used the analytical solution and the median values of link lengths to discretize the curve into several points with different θ_3 values (Figure 4A and Figure S20). For each point on the curve, we took first-order partial derivatives of the six link lengths as input variables. The arrows shown in the inset of Figure 4A correspond to the changes in tip position that occur when individually changing the length of each link. Then, we aimed to maximize the deviation along the direction $\hat{e}_k$ (gray dashed arrow in Figure 4A inset) by summing linear combinations of these vectors from partial derivatives, and repeated this maximization, iterating the direction $\hat{e}_k$ from 10º to 360º to determine a boundary of potential positions for a given theoretical position (black dots in Figure 4A inset). We also tested a second order model that included 36 quadratic terms in addition to the 6 linear terms to increase the accuracy. This boundary discretization was repeated for several theoretical conformations defined by evenly spaced θ_3. We then integrated those boundary points into a shape indicating where the SLL tip position P_3 may be located for the entire motion pathway (Figure 4A, gray shaded area). However, TSOA cannot consider the random joint errors since this method is deterministic.

For the MC algorithm, we aim to evaluate the impact of the joint error by extending the original vector loop equations with four extra terms of joint errors, formulated as modified vector loop equations. In the modified vector loop model (Eq S11), we could separate the joint and link errors with the MC approach using different sampling settings (Eq S13). For example, to specifically consider the effects of joint clearance, we can assume each joint exhibits a maximum clearance of 1.2 nm while each link has a constant length as determined by the median value from the generated link length distributions. Alternatively, to specifically consider the motion uncertainty due to variations in link length, we assumed perfect joints with zero clearance and only considered variations in link lengths. Figure 4B illustrates the results of repeatedly sampling the tip positions P_3 (N=10^6) and separately considering the joint error and the link error, showing the deviation in the link length dominates the overall motion uncertainty. This suggests that making links stiff is critical to minimizing motion uncertainty of DOM performance.

The MC and TSOA computational tools are both capable of predicting the motion pathway of expected experimental results of the SLL. Figure 4C compares the results from the TSOA and MC simulation with TEM data for verification. The MC method can calculate the probability distribution of the tip position to quantitatively compare the TEM-based distribution. The contour levels we set for the MC method are 30%, 50%, 90% and 99% which capture 30%, 45%, 76% and 92% of the TEM data, respectively. 15 (1%) imaged SLLs are outside of the TSOA, and 62 (8%) for the 99% contour of the MC. These small differences may be due to experimental limitations in selecting points or due to limitations of the modeling, for example in not properly accounting for ion interactions. Nevertheless, the results, generally agree well with experimental data.

These two computational tools, TSOA and MC, were also consistent with each other. The TSOA method used the median, lower bound, and higher bound values in a multi-variable model to depict

the possible outcome. Hence, the TSOA method may be more useful when all that is known or can be estimated are geometrical tolerances of the components, such as from TEM images, rather than a full conformational distribution. On the other hand, the MC method is capable of decoupling the link and the joint errors. In addition, it also provides a means to estimate the 2D probability distribution of the tip position.
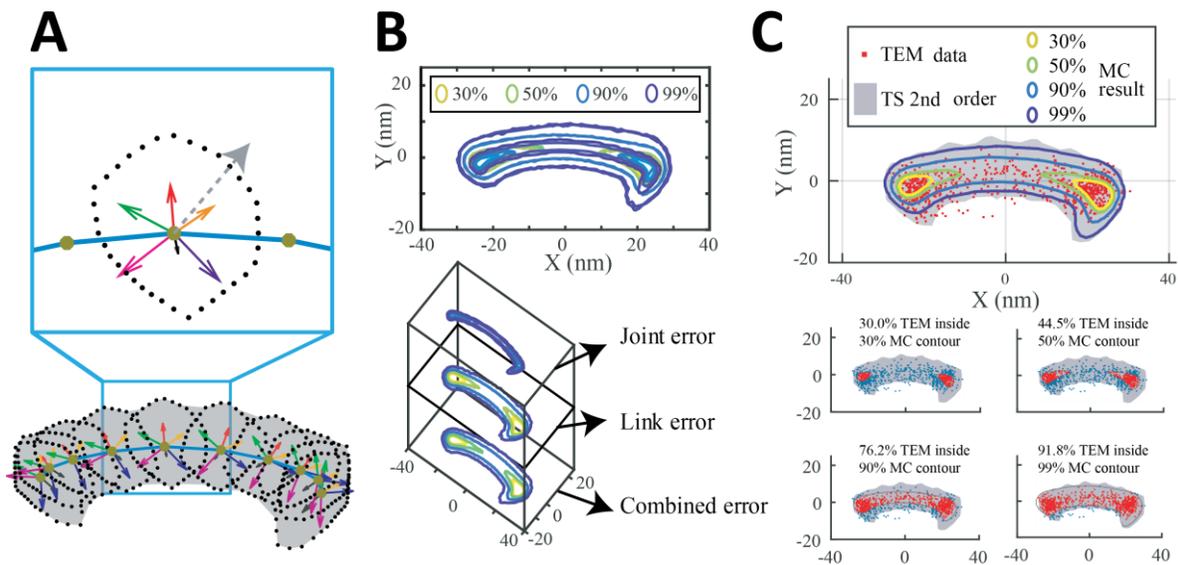

Figure 4: Computational tools and error quantification of the SLL. (A) Schematic of combining the Taylor series expansion and optimization algorithm (TSOA) to calculate the boundary of the tip P_3. The green dots denote regularly spaced positions along the motion pathway where the TSOA was carried out. The different color arrows denote the changes in the tip position with changes of link 1(black), link 2(orange), link 3 (red), link 4 (green), link 5(purple), and link 6(magenta). The boundary for motion uncertainty was determined by maximizing the deviation from the theoretical position along a specific direction (dashed gray arrow). That direction was varied over 360° to obtain the full boundary (black dots). This was then convolved for different theoretical positions to obtain the overall boundary of motion uncertainty (gray shaded area). (B) Contour plots sampling the tip position P_3 with the Monte Carlo (MC) simulation, with the decoupling of joint error and link error (bottom) illustrating that the overall motion uncertainty is dominated by the link error. (C) Combined results from MC and TSOA methods compared to the TEM data. The bottom four plots compare directly the fraction of experimentally measured conformations that lie within various contours predicted by the MC approach.

DOM Conformational Probability Distribution

While the MC method gives more information including estimating the probability distribution for P_3, it does not completely agree with the experimental conformational distribution. In particular, the TEM measurements showed a stronger bimodal behavior compared to the MC predictions. This is likely because the MC predictions assume a uniform distribution of θ_3 as the input for the driver link, while our experimental data (Figure S26) reveals θ_3 actually exhibits a bimodal distribution. In the MC predictions, we simply extracted the lower and the upper bound for θ_3 of -43º and 47º, respectively, and assumed a uniform distribution between the range. Using a uniform distribution for the input created a less bimodal P_{3x} distribution (Figure S27), although some of the bimodal nature of P_{3x} is already seen even with a uniform θ_3 distribution as a result of the kinematics of the SLL.

We hypothesized that the strong bimodal distribution of P_{3x} and the bimodal distribution of θ_3 was due to the mechanical properties of the joints. These mechanical properties are not considered in our computational tools, which only account for variations in link geometry and joint clearance. To test this hypothesis, we converted the SLL to an open-chain mechanism by excluding staples that bridge the scaffold seam on link 4. This cuts link 4 and allows the SLL to fluctuate freely in an open-chain mode (Figure 5A) where individual joints can fluctuate independently. Once in the open-chain mode, we measured the angles ϕ_1 and ϕ_2 of the two joints on the open loop SLL (Figure 5B) and compiled their angular probability distributions (Figure 5C), which we then used to determine joint free energy landscapes (Figure 5D). We observed a strong peak in the probability distributions of ϕ_1 at 185º and ϕ_2 at 167º. While the 185º peak for ϕ_1 is not feasible in the closed-

loop mechanism, the 167° peak for ϕ_2 is, and therefore, the mechanical properties of ϕ_2, and similarly ϕ_3, are likely to play an important role in the closed-loop mechanism.

We used the energy landscapes measured from the open loop SLL to evaluate the free energy of each joint when it is at a given angle in the closed-loop mechanism, assuming that joint 4 and joint 3 behave similarly to joints 1 and 2, respectively, due to the symmetry of the mechanism Figure 5E shows the free energy of the closed-loop SLL in terms of the individual joint contributions and the total sum free energy of all the joints (more details in Figure S28-29). The total free energy local minima exhibited by joints 2 and 3 translate to energy minima of the overall mechanism at θ_3= -31° and 33° (Figure 3E bottom and Figure 3F), which correspond to the peaks in P_{3x} at -22.8 and 23.6 (Figure S26).

The preferred angles of joints 2 and 3 (Figure 5C) occur when the joints are open enough for interactions to occur between the ends of neighboring links, such as in base stacking interactions[55] or base-pairing between ssDNA scaffold loops at the ends of links[56] (Figure S32). These interactions are known to be salt-dependent. Therefore, we hypothesized that lowering salt concentrations would reduce the strength of these interactions and yield a more uniform distribution of θ_3. However, we expect increasing salt concentrations would yield a more prominent bimodal distribution of θ_3. To test this hypothesis, we conducted experiments with a set of two different salt conditions, 5mM $MgCl_2$ (Figure S30) and 20mM $MgCl_2$ (Figure S31) to compare to the previous experiments performed at 10mM $MgCl_2$ (Figure S26). The results supported our hypothesis that the lower salt concentration created a more uniform θ_3 distribution and the higher salt enhanced the bimodal distribution. In addition, the distribution from the SLL in a 20mM $MgCl_2$ buffer shows a bias to the right side. This may be explained by noting that the

sequences of the scaffold loops show preference to form secondary structures at joint 3 (Figure S31, Tables S9, and S10).

From experimental data, we can confirm that the lateral distribution of the SLL is salt dependent. Describing salt-dependence of DNA nanodevices in molecular dynamics simulations is a challenge. oxDNA2 contains salt dependence[47]; and it is possible that running significantly longer simulations in oxDNA may better capture the bimodal distribution. However, the oxDNA2 model includes dependence on NaCl while our experiments were conducted in $MgCl_2$, which could also explain the discrepancies between the experimental and simulation distributions (Figures 1, 2, and S8). Regardless, decoupling the system into joints as sub-systems is beneficial to reduce the computational burden if oxDNA or other computational models become compatible with magnesium conditions in the future. Moreover, since the number of bases in a sub-system is much less than in a full system, we can foresee the possibility of using an all-atomic model to capture the joint properties and integrate those MD simulation results with our computational tools.
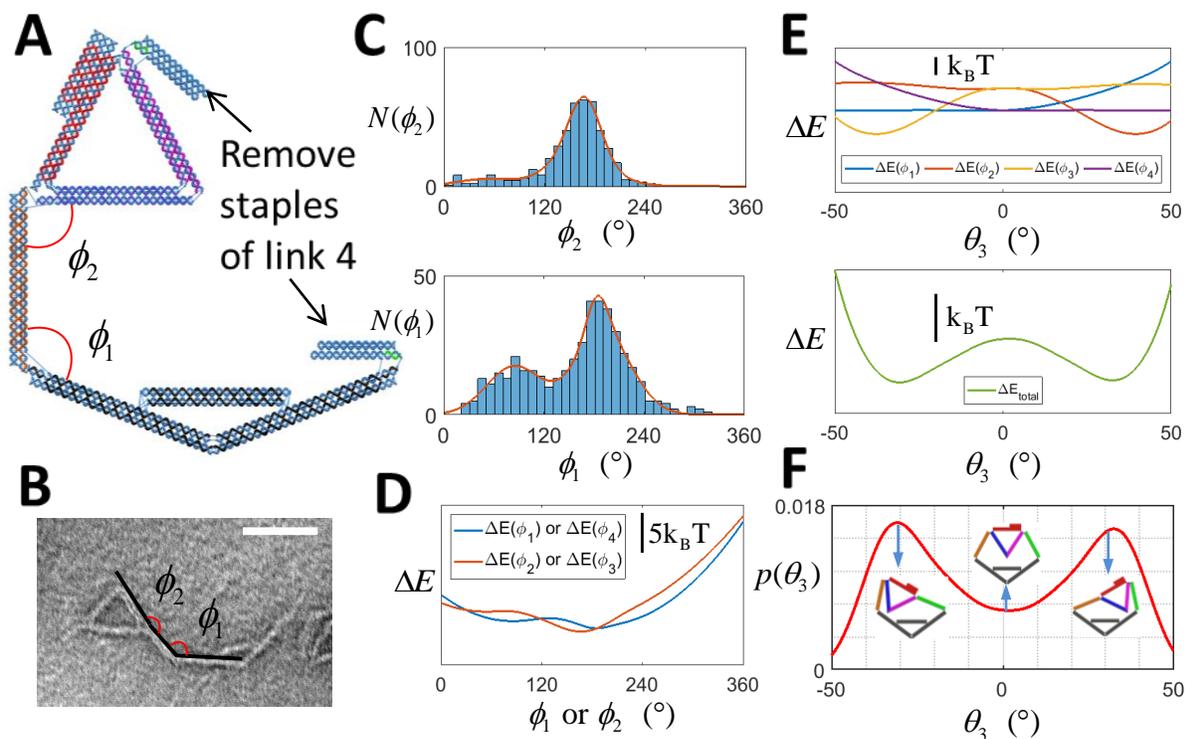
Figure 5: Measuring joint properties to model the SLL bimodal distribution. (A) Schematic of the open-chain SLL, which is obtained by excluding the link 4 staple strands from the folding process. With the open-chain mechanism, we were able to measure the mechanical properties of the free joints in the absence of mechanism constraints. (B) A representative TEM image shows the measured angles of the two joints of interest. (C) The distributions of the measured joints (N = 436). For the waist joint, ϕ_1, two peaks at ~80º and ~185º were observed. The shoulder joint, ϕ_2, peaked at ~165º. Both distributions were fitted with a combination of three Gaussian distributions. (D) The free energy landscapes of all four joints were calculated according to the Boltzmann distribution assuming a symmetric design. (E) By solving kinematic equations, the angles and corresponding free energies of all four joints at different configurations was computed (top), and the summation of energy of the four joints shows the energy landscape of the entire mechanism (bottom). (F) Reconstruction of the θ_3 distribution using the open-chain SLL experiment, which were performed at 10 mM MgCl$_2$.

Actuating DNA Origami SLL Mechanisms

To demonstrate actuation of the SLL and test our capabilities for predicting uncertainty in actuated positions, we folded a version of the SLL mechanism that contained 12-nt ssDNA overhangs to move P_3 to the left or to the right with introduction of unique sequences (Figure 6A).

As shown in Figure 6A and 6D, we successfully actuated the mechanism to either side, depending on the closing strands introduced. We used the asymmetric feature on link 3 to confirm which side the mechanism was actuated toward. Compared to the closed-loop SLL, the two actuated versions of the SLL exhibit relatively narrow configurational distributions tilted to the left or to the right (Figure 6B, Movie S1, S6 and S7). Additional TEM images of the actuated SLL mechanisms are provided in Figures S34 and S35. Agarose gel electrophoresis in Figure 6C also showed the actuated structures exhibited a slight shift, running faster compared to the open-chain SLL while the addition of overhangs did not cause a shift in the absence of actuation.

In order to predict the motion of the actuated SLL mechanism, we applied the oxDNA simulation (Figure 6D) and our TSOA and MC approaches (Figure 6E). This hybrid computational tool captured most errors within the experimental data. Specifically, the 99% contours from the MC approach captured 78% and 89% of TEM samples in the cases of actuating to the left and the right, respectively. We speculated that the slightly lower capture accuracy of the left side actuation came from the binding properties of the actuation strands. The overhang sequence with the 3'end (arrow) for the actuation to the right in Figure 6A has higher melting temperature at 33.1℃ than the other three overhangs (5'end (square) on link 2, 3'end on link 5, and 5'end on link 6 with melting temperatures at 31.6°, 26°, 23.2℃, respectively). Hence, it may be possible that the SLL is more likely to consistently incorporate all of the actuation strands when it is actuated to the right.
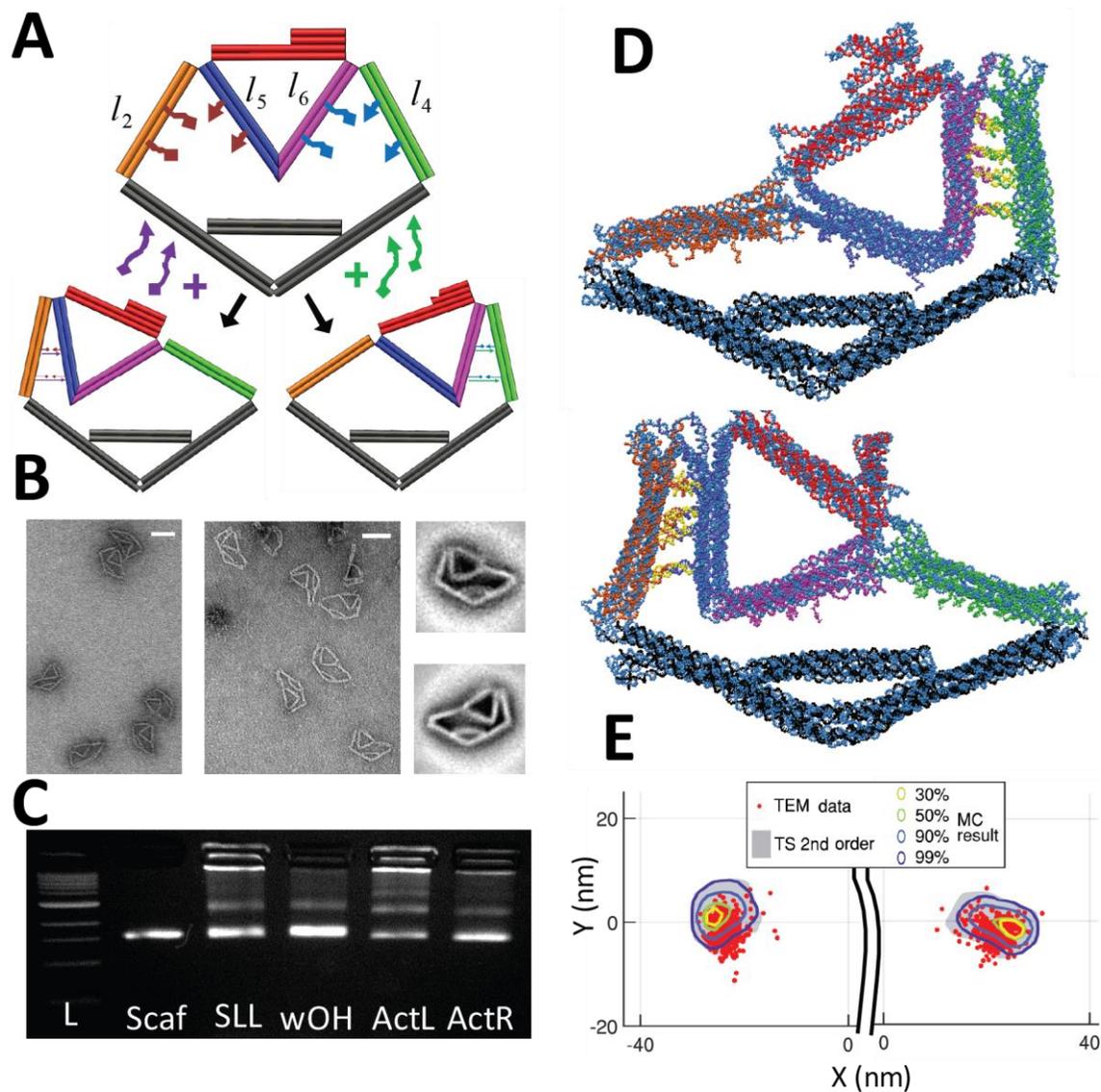
Figure 6: Controlling the distribution of the SLL by actuating with overhangs. (A) With the addition of eight 12-nt overhangs, the SLL can be actuated to the left or the right configuration by adding the associated actuation strands. (B) TEM images of the actuated SLL to the right and left. TEM image averaging was applied to estimate the most likely configurations. (C) Agarose gel electrophoresis image of 1kb ladder (L), scaffold (7249), SLL, SLL with overhangs, SLL actuated to the left, and SLL actuated to the right. (D) Representative snapshots of the actuated oxDNA simulation results. The yellow strands are the closing strands added to one side to actuate the mechanism. (E) Combined results from two computation methods, TSOA and MC, and TEM data. The distribution of the tip P_3 position from TEM images illustrated the uncertainty in the actuated position of the SLL. Sample size = 398 (left) and 469 (right).

Uncertainty Quantification of the Crank-Slider DOM

This proposed integrated approach aims to lower the computational burden of simply using coarse-grained MD simulations for dynamic mechanisms by introducing the kinematic variance analysis. To demonstrate the broader applicability of this approach, we also applied this framework to one of our previously designed complex DOM, the crank-slider mechanism[7]. Our results for this mechanism are summarized in Figure S25. The crank-slider is composed of 4 bundle components, connected by either rotational or slider joints, and converts a crank angle rotations, θ, to slider translations, r. Similar to the SLL, we simulated the mechanism using oxDNA and quantified the geometric parameters that define the kinematic motion from the trajectory (Figure S25C). With the analytical solution, we applied the MC method to repeatedly sample the output extension value r versus the input rotation angle θ. Our results yield a prediction of the motion uncertainty that agrees reasonably with the experimental data presented by Marras *et al.*[7], although there are larger deviations from the experiment compared to our SLL results, perhaps since we only ran one 1500 μs simulation for this mechanism, which may not have explored the full range of mechanism conformations. Hence, we also tested using the crank angle distribution from the previous study Sharma *et al.*[20], which ran much longer oxDNA simulations of this same mechanism. Using their distribution of the input crank angle along with the component length distributions from our simulation yielded good agreement with experiments (Figure S25).

Conclusions

Here, we presented an integrated computational, theoretical, and experimental approach and associated computational tools for quantifying uncertainties and predicting motion of DOM. We introduced two kinematic variance approaches based on the coarse-grained simulation results.

Integrating these kinematic approaches greatly reduced the required time to simulate this large DNA origami structure that exhibits large configurational changes. To validate these tools, we designed and fabricated a DNA origami straight-line linkage (SLL) and tested the tools with a previously reported DOM[7]. For the SLL, we further identified that the overall mechanism conformations can be described in terms of fundamental joint properties by breaking a closed-loop mechanism into an open-chain mechanism. We utilized experiments measuring the conformational distribution of the SLL at varying salt conditions to confirm the bimodal behavior is in part due to ion-sensitive interactions. Lastly, we utilized overhangs to demonstrate the ability of actuating, predicting, and controlling the motion of the mechanism. The results show that our tools can precisely predict the motion uncertainties of the SLL in both freely fluctuating and actuated cases.

Through design, simulation, and experimental validation, we gained fundamental insight into the motion behavior of DOM and the relation to underlying design parameters, which are likely generally applicable to a variety of dynamic DNA origami devices. First, compared to the theoretical cylindrical model design that would exhibit perfect planar motion, we observed significant out-of-plane motion in the oxDNA simulations when using two short ssDNA connections to form a hinge joint. This suggests that effectively constraining DOMs to planar motion requires either placing the ssDNA connections further apart and/or including more ssDNA connections to constrain the motion. Second, we found that although links are generally considered rather stiff, their deformations, in particular the fraying at the ends, contributes significantly to the motion uncertainty. Hence, designing stiffer links and especially minimizing fraying of the ends with the underlying design could improve motion fidelity. Finally, our results illustrate that the conformational distribution of the overall mechanisms can be well described by understanding the mechanical properties of the joints. This suggests that designing mechanisms that can easily be

converted into open-loop versions can be a useful approach to understand the mechanism behavior. In this case, we found the horizontal motion of the SLL was related to salt-dependent interactions that likely occur between the ends of links, either base stacking or base-pairing between single-stranded portions of the scaffold. This could either be used as a design feature, or could potentially be minimized by using poly-T staple overhangs that extend from the ends of links as has previously been done[57].

From a computational viewpoint, our integration of coarse-grained modeling and kinematic analysis enables a rapid, accurate, and systematic tool to predict and evaluate the quality of DNA origami designs. In addition, the ability to predict motion uncertainties with a relatively efficient framework can facilitate an iterative design process to achieve DOM designs with improved motion fidelity. Moving forward, our approach is also amenable to multi-level simulations where components are simulated separately to extract features that affect the motion of the overall design. Although, here we found some bundle components can exhibit different length distributions with a DOM relative to the isolated component by itself. Likely this is more of a concern for closed-loop mechanisms where kinematic constraints can lead to stresses on individual components. To explore the configurational space, another example of multi-level simulation is the virtual-move Monte Carlo (VMMC) algorithm using clusters of particles to avoid kinetic traps[58]. Similarly, by using coarse-grained models to the component level and applying rigid body transformations to components as we did here, one can sample a variety of states for an initial configuration and then switch to a finer coarse-grained model to minimize the energy of the entire structure. Furthermore, all-atomic MD simulations can be added at the finest hierarchy[32] to cooperate with coarse-grained and kinematic analysis, depending on the spectrum of accuracy, time-scale, and computational

costs. We also showed that the computational tool is also applicable to other DNA origami mechanisms such as the crank-slider.

In addition, a significant outcome of this work is the computational tools, which can greatly simplify the use of oxDNA for general multi-component structures. This creates a foundation for simulating general dynamic DNA origami mechanisms, and allows user-defined control over initial configurations for simulation, which can enable studies of low probability conformations or verification of equilibrium distributions. These tools will continue to improve the rational and model-driven design of DNA origami structures and mechanisms with complex mechanical function.

Conflicts of interest

The authors declare no competing financial interests.

Acknowledgements

Author contribution

This project was initiated and designed by C.E.C. and H.-J.S. C.-M.H designed and simulated the DNA origami structure. A.K. conducted the majority of experiments and analyzed experimental results. J.V.L. conducted part of the TEM imaging. C.-M.H., A.K., C.E.C. and H.-J.S. analyzed the data. C.-M.H. drafted the initial manuscript, which was revised by A.K., J.V.L., H.-J.S. and C.E.C..

Corresponding Authors

*E-mail: su.298@osu.edu

*E-mail: castro.39@osu.edu

Reference

1 J. Howard, in *Biological Physics*, Springer, Basel, 2011, pp. 47–59.

2 A. Furuta, M. Amino, M. Yoshio, K. Oiwa, H. Kojima and K. 'ya Furuta, *Nat. Nanotechnol.*, 2017, **12**, 233–237.

3 V. Balzani, A. Credi and M. Venturi, *Molecular Devices and Machines: Concepts and Perspectives for the Nanoworld*, John Wiley & Sons, 2008.

4 G. Ragazzon, M. Baroncini, S. Silvi, M. Venturi and A. Credi, *Nat. Nanotechnol.*, 2015, **10**, 70–75.

5 W. R. Browne and B. L. Feringa, *Nat. Nanotechnol.*, 2006, **1**, 25–35.

6 N. C. Seeman, *J. Theor. Biol.*, 1982, **99**, 237–247.

7 A. E. Marras, L. Zhou, H.-J. Su and C. E. Castro, *Proc. Natl. Acad. Sci.*, 2015, **112**, 713–718.

8 L. Zhou, A. E. Marras, H.-J. Su and C. E. Castro, *Nano Lett.*, 2015, **15**, 1815–1821.

9 P. Ketterer, E. M. Willner and H. Dietz, *Sci. Adv.*, 2016, **2**, e1501209.

10 A. J. Thubagere, W. Li, R. F. Johnson, Z. Chen, S. Doroudi, Y. L. Lee, G. Izatt, S. Wittman, N. Srinivas, D. Woods, E. Winfree and L. Qian, *Science*, 2017, **357**, eaan6558.

11 D. C. Khara, J. S. Schreck, T. E. Tomov, Y. Berger, T. E. Ouldridge, J. P. K. Doye and E. Nir, *Nucleic Acids Res.*, 2018, **46**, 1553–1561.

12 P. W. K. Rothemund, *Nature*, 2006, **440**, 297–302.

13 S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf and W. M. Shih, *Nature*, 2009, **459**, 414–418.

14 H.-J. Su, C. E. Castro, A. E. Marras and M. Hudoba, in *Advances in Reconfigurable Mechanisms and Robots I*, eds. J. S. Dai, M. Zoppi and X. Kong, Springer London, 2012, pp. 487–500.

15 C. E. Castro, H.-J. Su, A. E. Marras, L. Zhou and J. Johnson, *Nanoscale*, 2015, **7**, 5913–5921.

16 T. Gerling, K. F. Wagenbauer, A. M. Neuner and H. Dietz, *Science*, 2015, **347**, 1446–1452.

17 L. Zhou, C. E. Castro, A. E. Marras and H.-J. Su, in *Proceedings of ASME IDETC/CIE 2015*, 2015.

18 A. E. Marras, L. Zhou, V. Kolliopoulos, H.-J. Su and C. E. Castro, *New J. Phys.*, 2016, **18**, 055005.

19 J. List, E. Falgenhauer, E. Kopperger, G. Pardatscher and F. C. Simmel, *Nat. Commun.*, 2016, **7**, 12414.

20 R. Sharma, J. S. Schreck, F. Romano, A. A. Louis and J. P. K. Doye, *ACS Nano*, , DOI:10.1021/acsnano.7b06470.

21 D. Lei, A. E. Marras, J. Liu, C.-M. Huang, L. Zhou, C. E. Castro, H.-J. Su and G. Ren, *Nat. Commun.*, 2018, **9**, 592.

22 S. Williams, K. Lund, C. Lin, P. Wonka, S. Lindsay and H. Yan, in *DNA Computing*, Springer, Berlin, Heidelberg, 2008, pp. 90–101.

23 S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church and W. M. Shih, *Nucleic Acids Res.*, 2009, **37**, 5001–5006.

24 E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, A. Lind-Thomsen, W. Mamdouh, K. V. Gothelf, F. Besenbacher and J. Kjems, *ACS Nano*, 2008, **2**, 1213–1218.

25 R. Veneziano, S. Ratanalert, K. Zhang, F. Zhang, H. Yan, W. Chiu and M. Bathe, *Science*, 2016, **352**, 1534–1534.

26 E. Benson, A. Mohammed, J. Gardell, S. Masich, E. Czeizler, P. Orponen and B. Högberg, *Nature*, 2015, **523**, 441–444.

27 C. E. Castro, F. Kilchherr, D.-N. Kim, E. L. Shiao, T. Wauer, P. Wortmann, M. Bathe and H. Dietz, *Nat. Methods*, 2011, **8**, 221–229.

28 D.-N. Kim, F. Kilchherr, H. Dietz and M. Bathe, *Nucleic Acids Res.*, 2012, **40**, 2862–2868.

29 K. Pan, D.-N. Kim, F. Zhang, M. R. Adendorff, H. Yan and M. Bathe, *Nat. Commun.*, 2014, **5**, 5578.

30 R. S. Sedeh, K. Pan, M. R. Adendorff, O. Hallatschek, K.-J. Bathe and M. Bathe, Computing Nonequilibrium Conformational Dynamics of Structured Nucleic Acid Assemblies, http://pubs.acs.org/doi/suppl/10.1021/acs.jctc.5b00965, (accessed 7 March 2018).

31 K. Göpfrich, C.-Y. Li, M. Ricci, S. P. Bhamidimarri, J. Yoo, B. Gyenes, A. Ohmann, M. Winterhalter, A. Aksimentiev and U. F. Keyser, *ACS Nano*, 2016, **10**, 8207–8214.

32 C. Maffeo, J. Yoo and A. Aksimentiev, *Nucleic Acids Res.*, 2016, **44**, 3013–3019.

33 S. M. Slone, C.-Y. Li, J. Yoo and A. Aksimentiev, *New J. Phys.*, 2016, **18**, 055012.

34 K. Pan, W. P. Bricker, S. Ratanalert and M. Bathe, *Nucleic Acids Res.*, 2017, **45**, 6284–6298.

35 A. Savelyev and G. A. Papoian, *Proc. Natl. Acad. Sci.*, 2010, **107**, 20340–20345.

36 E. J. Wallace and M. S. P. Sansom, *Nano Lett.*, 2007, **7**, 1923–1928.

37 H. Yagyu, J.-Y. Lee, D.-N. Kim and O. Tabata, *J. Phys. Chem. B*, 2017, **121**, 5033–5039.

38 T. E. Ouldridge, A. A. Louis and J. P. K. Doye, *J. Chem. Phys.*, 2011, **134**, 085101.

39 P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye and A. A. Louis, *J. Chem. Phys.*, 2012, **137**, 135101.

40 J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Šulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison and W. P. J. Smith, *Phys. Chem. Chem. Phys.*, 2013, **15**, 20395–20414.

41 J. S. Schreck, F. Romano, M. H. Zimmer, A. A. Louis and J. P. K. Doye, *ACS Nano*, 2016, **10**, 4236–4247.

42 B. E. K. Snodin, F. Romano, L. Rovigatti, T. E. Ouldridge, A. A. Louis and J. P. K. Doye, *ACS Nano*, 2016, **10**, 1724–1737.

43 Z. Shi, C. E. Castro and G. Arya, *ACS Nano*, 2017, **11**, 4617–4630.

44 L. Zhou, A. E. Marras, C.-M. Huang, C. E. Castro and H.-J. Su, *Small*, 2018, **0**, 1802580.

45 L. Rovigatti, P. Šulc, I. Z. Reguly and F. Romano, *J. Comput. Chem.*, 2014, **36**, 1–8.

46 Stahl Evi, Martin Thomas G., Praetorius Florian and Dietz Hendrik, *Angew. Chem. Int. Ed.*, 2014, **53**, 12735–12740.

47 B. E. Snodin, F. Randisi, M. Mosayebi, P. Šulc, J. S. Schreck, F. Romano, T. E. Ouldridge, R. Tsukanov, E. Nir and A. A. Louis, *J. Chem. Phys.*, 2015, **142**, 06B613_1.

48 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

49 R. L. Norton, *Design of machinery: an introduction to the synthesis and analysis of mechanisms and machines*, McGraw-Hill Boston, 1999, vol. 924.

50 K. J. Waldron, G. L. Kinzel and S. K. Agrawal, *Kinematics, Dynamics, and Design of Machinery*, John Wiley & Sons, 2016.

51 Y. Ke, J. Sharma, A. Cheng, A. Leung, Y. Liu, W. M. Shih and H. Yan, *J. Am. Chem. Soc.*, 2009, **131**, 15903–15908.

52 J. V. Le, Y. Luo, M. A. Darcy, C. R. Lucas, M. F. Goodwin, M. G. Poirier and C. E. Castro, *ACS Nano*, 2016, **10**, 7073–7084.

53 D. Thirumalai and B.-Y. Ha, *ArXivcond-Mat9705200*.

54 X. Li, X. Ding and G. S. Chirikjian, *Mech. Mach. Theory*, 2015, **91**, 69–85.

55 F. Kilchherr, C. Wachauf, B. Pelz, M. Rief, M. Zacharias and H. Dietz, *Science*, 2016, **353**, aaf5508.

56 J. SantaLucia, *Proc. Natl. Acad. Sci.*, 1998, **95**, 1460–1465.

57 X. Bai, T. G. Martin, S. H. W. Scheres and H. Dietz, *Proc. Natl. Acad. Sci.*, 2012, **109**, 20012–20017.

58 S. Whitelam and P. L. Geissler, *J. Chem. Phys.*, 2007, **127**, 154101.

$$\vec{P}_3 = f(l_1, l_2, l_3, l_4, l_5, l_6, \theta_3)$$