bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Allosteric modulation of local reactivity in DNA origami

Antonio Suma,[1,2,3] Alex Stopar,[2,3] Allen W. Nicholson,[2] Matteo Castronovo,[4,2,3] and Vincenzo Carnevale[1,2]

[1] *Institute for Computational Molecular Science, Temple University, Philadelphia, PA 19122, USA*
[2] *Department of Biology, Temple University, Philadelphia PA, 19122, USA*
[3] *Department of Chemical Science and Technologies, University of Rome, Tor Vergata, Via della Ricerca Scientifica, 00133, Rome, Italy*
[4] *School of Food Science and Nutrition, University of Leeds, Leeds, UK*

(Dated: May 30, 2019)

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ABSTRACT

We used extensive coarse-grained molecular dynamics simulations of a sharp triangular DNA origami nanostructure to characterize its global and local mechanical properties. Globally, the structure presents two metastable conformations separated by a free energy barrier, which is lowered upon removing four staples from the structure. Locally, the fluctuations are monitored to estimate the endonuclease reactivity towards DNA cleaving and show good agreement with experimental data. We finally show that global fluctuations interfere with local fluctuations, provided that the active site is 40nm distant from the staple defect, proving that an allosteric control of reactivity is possible.

INTRODUCTION

The concept of allostery describes how conformational transitions in distant regions of a biomolecule can be thermodynamically coupled. Due to allostery, a localized perturbation can affect the conformational ensemble of the entire molecule [1]. Examples of allosteric biomolecules include ion channels, which open and close a gate in response to a conformational change occurring in a functionally specialized transduction domain, and many enzymes for which the catalytic site is activated by binding partners and by co-factors that are recognised by regions distant from the active site [2, 3].

Nanometer-sized mechanical meta-materials are artificial assemblies that can mimic the behaviour of allosteric biomolecules by responding to mechanical perturbations as inputs through global reshaping of their free energy landscape. Such meta-materials can be programmed by appropriately arranging their domains [4]. An appropriate design can make this *allosteric metamaterials* poised to change abruptly structural properties (*e.g.* surface roughness) or mechanical response (*e.g.* bending stiffness) [5, 6], thereby paving the way for futuristic technologies such as mechanologic gates [7], adaptive computing materials [8] and soft nanorobots [9, 10].

Owing to their unique self-assembly properties, nucleic acids represent mechanical meta-materials of growing importance. As an example, DNA origami objects are comprised of a long single-stranded (ss)DNA (scaffold) , folded into programmed 2D or 3D shapes by the pairing with multiple short DNA strands (staples), forming defined patterns of double-stranded (ds) domains [11–14]. In turn, patterns of soft (ss) and stiff (ds) domains can be generated with single nucleotide accuracy within nanostructures that are able to adopt interchangeable conformations [15–18].

Very recently, our group discovered unexpected allosteric effects in a sharp triangle [19], a DNA origami first introduced by Rothemund, and subsequently involved in DNA nanotechnology studies [20–22], by experimentally investigating the action of several sequence-specific, DNA-cutting enzymes (restriction endonucleases or REases) towards their recognition sites present in the M13 scaffold sequence. The study demonstrated that REases act in a binary fashion, as certain sites cannot be cut, or, for the HhaI REase, become reactive by introducing a distant structural defect in the sharp triangle (i.e. by omitting in the self-assembly process four staples at a distance of approx. 40 nm from the REase site).

In this study, we computationally examine this phenomenon to obtain insight on the global mechanical properties of the DNA origami triangle and their allosteric connection to restriction site reactivity. In particular, we studied two variants of the triangle: the first is a well-formed triangle having the complete complement of staples; and the second is a defective triangle with a hole generated by removing four staples (referred as bottom triangle in ref. [19]). We compared the global mechanical properties of the two triangles, then correlated the properties with the reactivity of individual sites towards cut by HhaI REase.

Our findings are as follows. First, removal of four staples affects the amplitude of the global conformational fluctuations of the triangle. Two metastable states are observed in both triangles; however, the removal of the staples lowers the free energy barrier between the two states. Second, we devised a computational method that estimates the restriction site accessibility to a restriction endonuclease, and obtained good agreement between the experimental and the theoretical results; these results revealed that site accessibility depends on local conformational fluctuations. Third, we correlated the site accessibility to the global conformational fluctuations of the triangle, thereby showing how, in a strongly coupled mechanical system, the restriction site reactivity can be modulated in an allosteric fashion.

RESULTS AND DISCUSSION

System setup.

To determine the conformational dynamics of two sharp triangles that differ slightly in the staple composition, we carried out extensive Langevin molecular dynamics simulations. The first structure, hereafter referred to as "triangle", has the scaffold organized by the full complement of staples (Fig. 1a). The second structure lacks four staples that

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

bridge the seam of one of the three trapezoidal elements forming the triangle, and thus is referred to as "defective triangle" (Fig. 1c). Both structures were previously studied experimentally in ref. [19].

Each structure is constituted by about $N = 14000$ nucleotides, and was modelled using oxDNA, a coarse-grained DNA model that has been shown to accurately reproduce the mechanical properties of double-stranded and single-stranded DNA molecules [23, 24]. The time of evolution of the two structures was monitored for 23ms; the mapping between the intrinsic and standard time was previously found to yield $\tau_{LJ} \sim 0.7ns$ [25, 26] and the configurations are sampled every $7\mu s$ (for a total of $\sim 3,000$ configurations).

Global mechanical properties and metastable conformations.

As a first characterization, we computed average structure and the root mean square fluctuation (RMSF) of each particle around the mean position, see Fig. 1b-d. To this end, we applied optimal rotations and translations (so as to minimize the root mean square deviation with respect to the initial configuration) to each frame of the trajectory using the Kabsch algorithm [27].

For both triangles, the average structure is globally over-twisted over the three trapeziums in the same manner, due to the helical pitch of 10.67 bp/turn imposed by the crossovers arrangements [11, 28]. The fluctuations are concentrated on the tips, pointing all in the same direction, and with an RMSF greater than 6.5 nm.

At a first glance, the triangle and the defective triangle do not show evident differences that might rationalize the experiments of ref. [19], except for the area bearing the defect and the upper edges of the trapeziums, which show slight variations. A deeper analysis of the trajectories, however, show that the two triangles fluctuate in different ways over time (see Movies S1-S2).

To distinguish between the different global fluctuation modes, we performed principal component analysis [29] over the trajectories using GROMACS [30]. We constructed the covariance matrix for the two different triangles as $C_{nm} = \frac{1}{T}\sum_{t=0}^{T} x_n(t)x_m(t)$, where t is the frame index and x_n is the n-th component of the 3N-dimensional vector of deviations from the average position ($n = 1, \ldots 3N$), and we computed eigenvalues and eigenvectors ($\{\lambda_n, \mathbf{p}_n\}$ and $\{\lambda_n^d, \mathbf{p}_n^d\}$) for both the triangle and the defective triangle.

We projected the first 15 normalized eigenvectors of the triangle on the corresponding eingenvectors of the defective triangle, $\langle \mathbf{p}_n | \mathbf{p}_n^d \rangle$, and observed perfect alignment for the first 9 modes (Fig. 1e). For this reason, we can quantitatively compare the first 9 eigenvalues of the two triangles (Fig. 1f). We observe a significant difference in the first eigenvalue, with $\lambda_1 - \lambda_1^d = 101nm^2$, implying that the defect increases the fluctuation amplitude along the first mode direction $\mathbf{p}_1$, and leaves almost unchanged the remaining ones.

To further analyze the first fluctuation mode, we projected the trajectories over $\mathbf{p}_1$, see Fig. 2a. This allows us to analyze the fluctuations of the first mode as a function of time. What emerges is the existence of two metastable conformations, with transitions observed in both trajectories. In particular, we observe that the triangle has a much longer dwell time than the defective triangle. By reconstructing the free energy profiles as a function of $\langle \mathbf{x} | \mathbf{p}_1 \rangle$(Fig. 2b), we estimate an energy barrier between the two minima of about $5k_BT$ for the triangle, while for the defective triangle this barrier decreases to about $2.5k_BT$. The smaller energy barrier enhances the rate of exchange between the two basins.

The two dominant metastable conformations associated with the two free energy minima are shown in Fig. 2b. In both the vertices of the triangle point in the same direction, but with opposite orientations in the two states. Thus, the introduction of a defect can change the rate and duration of triangle tip orientation, and thus the possibility of manipulating a DNA origami so that it adopts different transitional conformations.

Computing the theoretical site accessibility.

We developed a method to determine the local accessibility of the triangle towards an REase. We defined a structural proxy by docking the enzyme crystal structure to the DNA site under study, then evaluated if, for each time frame, the enzyme clashes with adjacent DNA molecules, or has enough space in its immediate surroundings to dock, see Fig. 3b and methods for the exact procedure. We then define the fraction of frames where the site is accessible as the theoretical site accessibility.

Since the HhaI REase used in ref. [19] lacks a PDB structure, we used HinP1I as a proxy enzyme, as it recognizes the same site (GCGC), has a similar mass (HinP1I, 28.7 kDa; HhaI, 27.8 kDa), and a crystal structure of HinP1I docked to a duplex DNA has been reported (PDB 2FL3 [31]). Moreover, Hinp1I functions as a monomer [31], thus site accessibility has to be computed independently from either the scaffold strand or the staple strand.

The sharp triangle has twenty GCGC sites, see Fig. 3a. We computed the accessibility of each site, distinguishing the cases in which only the scaffold side, or only the staple side, or both sides are simultaneously accessible (see left side of Fig. 3c. We excluded sites 10, 11, 15 (shaded) as they overlap with a DNA crossover junction and therefore are inherently unreactive.

The emerging picture is that four sites exhibit high accessibility from the scaffold side (which comprises the accessibility only from the scaffold and the one from both sides at the same time), i.e. an accessibility higher than $\sim 50\%$: 12, 7, 19, 16 (in ascending order). The right side of Fig. 3c, shows the experimental assessment of the reactivity of every GCGC site in the triangle towards HhaI, involving extended (1 hr) incubation. The individual site reactivity

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

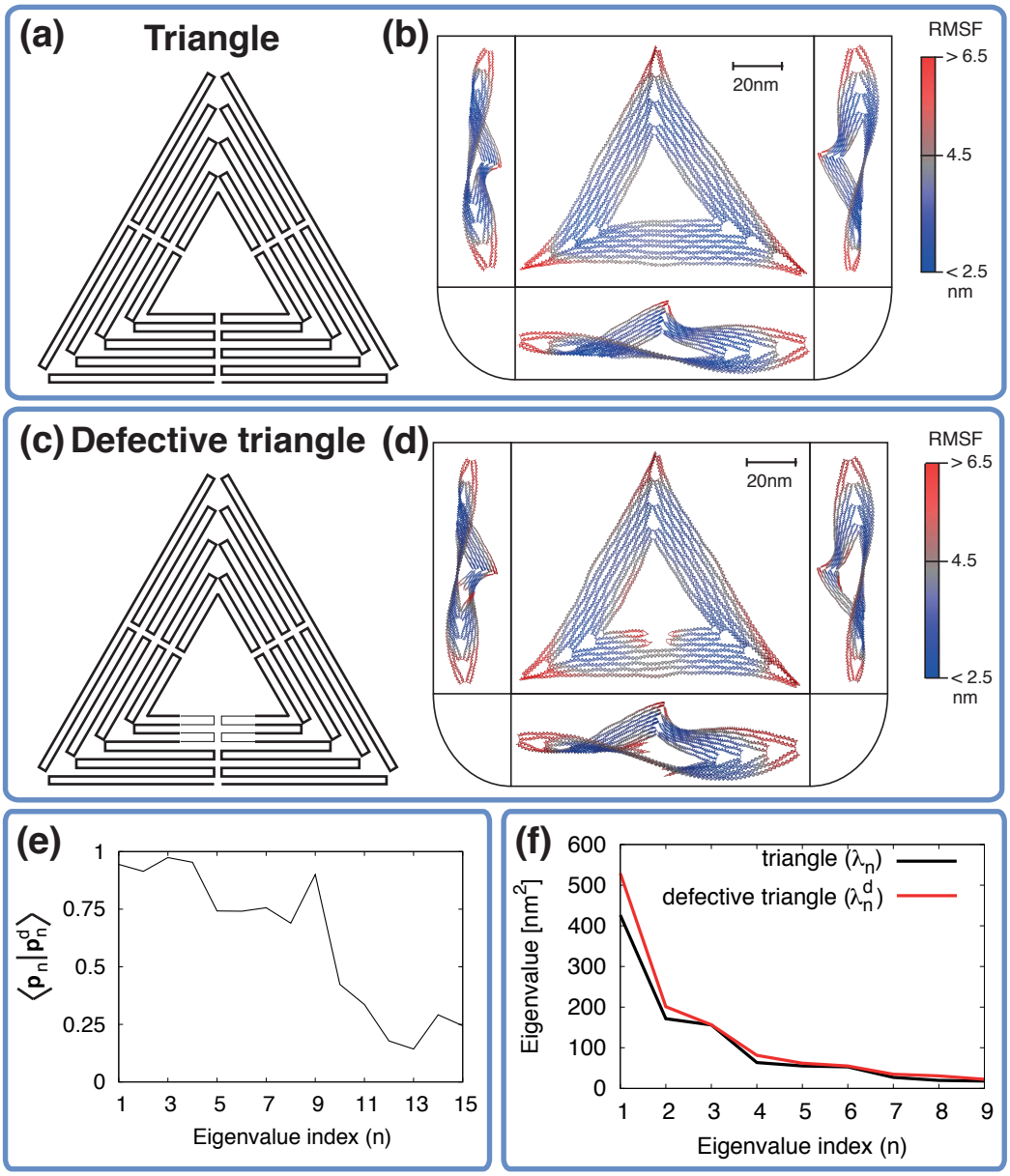

FIG. 1. a) Schematic representation of the scaffold routing for the sharp DNA origami triangle, and c) for the defective triangle. The thinner lines in the latter represent the scaffold portion left unpaired after the removal of four staples (defective region). b) The triangle and d) defective triangle average structure computed across the trajectory, viewed from the top and from the three sides. The S-shape is similar across the three sides and is due to the imposed over twisting. Color scale indicate the root mean square fluctuations (RMSF) values. Larger fluctuations concentrate on the tips, which point to the same side, and close to the defect. e) Projection of the first 15 eigenvectors (normalized) of the triangle over the first 15 of the defective triangle. The first 9 eigenvectors have similar directions, implying the fluctuations modes are comparable to each other. f) First 9 eigenvalues for triangle and defective triangle.

was reported by the fragmentation of the scaffold strand. The results show that HhaI can access and cut from the scaffold side sites 4, 7, 16, 19. Although the probing proxy is different, there is strong agreement for three out of four sites between the modeled accessibility to HinP1I and the measured reactivity of HhaI.

A quantitative comparison between the theoretical accessibility of HinP1I and the experimental reactivity of HhaI is shown in Figure 3d. In this comparison, we assumed that HhaI, like HinP1I, acts as a monomer, thus we are considering only the site accessibility from the scaffold. Aside from sites 4 and 12, which are also found to have

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

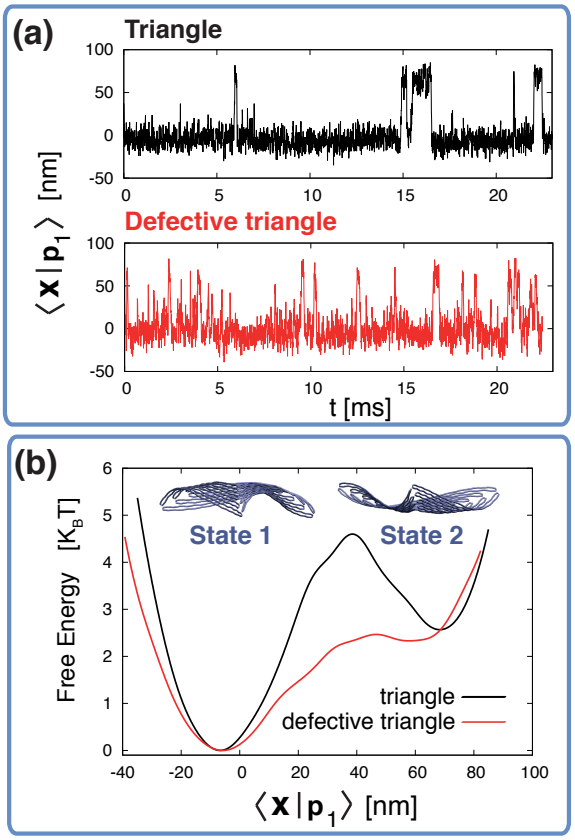

FIG. 2. a) Projection of trajectories of triangle and defective triangle over the first eigenvector of the triangle, $\langle \mathbf{x}|\mathbf{p}_1\rangle$. The structure hop in both cases between two states with a different rate. b) Free energy as a function of $\langle \mathbf{x}|\mathbf{p}_1\rangle$. An energy barrier is present between the two states and is lowered upon introducing the defect. The metastable states associated to the two minima are displayed on the top, and represent the triangle with the tip pointing altogether on one side or the opposite.

an intermittent reactivity across different defective triangles [19], we observed the two values match within an error margin of 20%, reflecting the noise level of the gel electrophoretic data. In particular, site 7 is cut more slowly than sites 16 and 19, which is predicted by the model. These results support the model assumption that the reactivity of a given GCGC site is related to its accessibility.

To better understand why some sites are locally available and others are not, we projected the center of mass of the adjacent DNA double-stranded molecules over a plane perpendicular to the GCGC helical axis, using as a reference frame the direction of one of the GCGC nucleotides. Figure 3e shows the projection in three different situations: a non-accessible site (site 6); a two-side-accessible site (from both the scaffold side and the staple side) (site 7); and a one-side-only accessible site, in this case accessible only from the scaffold strand (site 8).

The accessibility is determined locally by how much the enzyme HinP1I projection overlaps with the adjacent DNA molecules: a complete overlap prevents cut, while a partial overlap allows accessibility, with a cut rate proportional to the degree of overlap. This approach rationalizes the differences observed for different sites (e.g. sites 7 and 16), and the degree of overlap also can explain the different reactivity observed in experiments (see Fig. 3c).

We note that the degree of overlap between the enzyme and the DNA molecules depends on several variables, such as the orientation of the bases in the GCGC sites (from which the orientation of the docking protein depends), and how much the adjacent DNA structures fluctuate, which in turn is dictated by the distance from a crossover junction. These correlations may be useful in engineering a structure having a specific staple design that allows a specific enzyme to access and act.

Allosteric effect in DNA origami affects site reactivity.

We next compared the differences in local site reactivity between the triangle and the defective triangle, and their possible connection to changes in global mechanical fluctuations. A variation from 5 to 10% is observed for site 8 and site 12 for the accessibility from the scaffold, see Fig. 4a, which mirrors the fact that these sites become reactive when a defect is introduced at a distal area of the triangle, 40 nm distant [19].

A link between global fluctuation and local reactivity was determined by computing the accessibility of site 8 on the

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

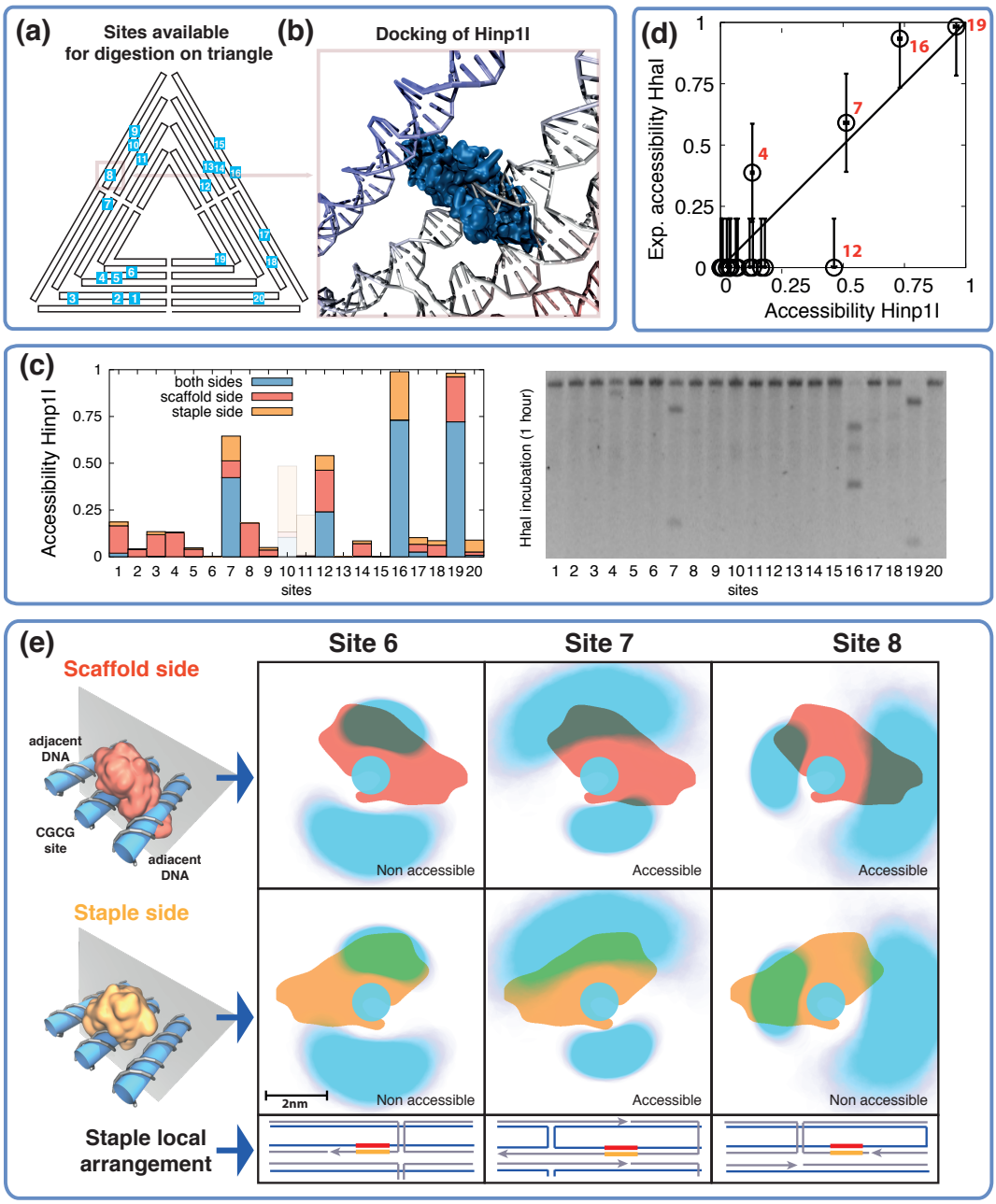

FIG. 3. a) Mapping of the 20 GCGC sites present in triangle. b) Detail of Hinp1I (in blue) docking on site 8 from the scaffold strand side. c) On the left, theoretical Hinp1I accessibility of every site of the triangle, decomposing the total probability in cases where the site may be cut only from the scaffold side (red), staple side (orange), or both sides at the same time. Site 10, 11 and 15 (shadowed) were excluded as they overlap with a crossover and are not a substrate for the enzyme. On the right, the experimental analysis of the triangle GCGC site reactivity in towards HhaI (data taken from ref. [19]) after 1 hour of incubation. d) Scatter plot of the theoretical and experimental accessibility percentages, from the scaffold side, the latter estimated from the gel bands. The Pearson's correlation between points is 0.89. e) On the left, the representation of the complex formed from the HinP1I enzyme docked on the GCGC site and the adjacent DNA molecules strands. On the center to the right, the latter complex projected (taking their center of mass) over a plane perpendicular to the site DNA molecule, for sites 6, 7 and 8 and for both scaffold and side docking. The reference frame is the position of one nucleotide of the GCGC site, thus the enzyme position is fixed, while the adjacent strands positions can vary. For site 6, in either sides the adjacent strands strongly overlap with the protein, suppressing the cut. For site 7, there is enough space instead for HinP1I to dock from both sides, due to higher fluctuations of the adjacent strand overlapping. Site 8 has the scaffold side accessible, while the staple side is not. On the bottom, the local staple arrangement geometry corresponding to each site, shows that the shorter is the distance between the GCGC site and the crossover, the smaller are local fluctuations of the corresponding adjacent DNA strand, and vice versa.

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

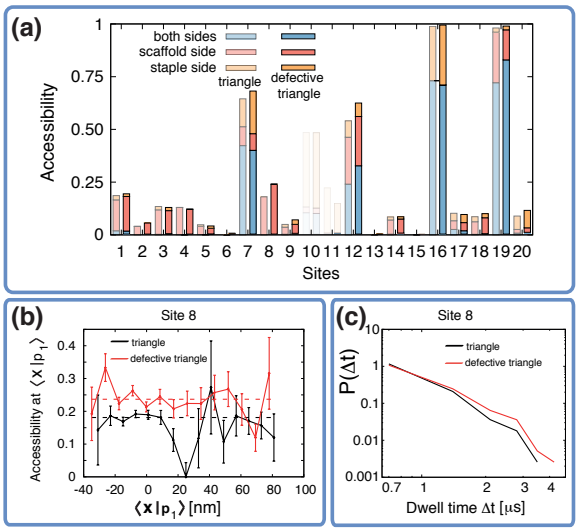

FIG. 4. a) Side by side comparison between theoretical accessibility for the triangle and defective triangle (see Fig. 3 for more details). b) Cleaving probability at fixed $\langle \mathbf{x}|\mathbf{p}_1\rangle$ for the two structures. Dropping in the probability are observed around $\langle \mathbf{x}|\mathbf{p}_1\rangle = 20nm$ for the triangle, but not for the defective one. c) Dwell time for site 8 (from the scaffold side).

scaffold side as a function of the projection of the instantaneous displacement on the first eigenvector of the covariance matrix $\langle \mathbf{x}|\mathbf{p}_1\rangle$, see Fig. 4b. While for both triangles there is a baseline value which is related to the accessibility (global value of Fig. 4a), some of the states close to the energy barrier at ~ 20 nm (see also Fig. 2b), are not accessible in the triangle, yet become so after introducing the defect. Importantly, a significant increase in the dwell time was observed for site 8 in the defective triangle, see Fig. 4c. These results suggest that allosteric effects might be relevant in DNA origami: while HhaI REase site 8 is distant from the defective region, a change in the latter region can modulate the dynamics of site 8 and control its reactivity.

Conclusions and perspectives

DNA nanostructures, and in particular DNA origami, constitute an important class of materials with features yet to be fully understood and applied. Here, we used extensive coarse-grained molecular dynamics simulations to probe the global and local mechanical properties of a triangular DNA origami.

We found that two metastable conformations are present, with the triangle tips pointing all in one direction or in the opposite one. Upon removal of several staples, the transition rate between the two states increases due to a decrease in the intervening free energy barrier.

We devised a method to quantify the local accessibility of GCGC sites to a restriction endonuclease, and showed that reactivity correlates with predicted accessibility. In particular, reactivity of a site can be rationalized in terms of the degree of overlap between the enzyme docked to a site and the DNA strands adjacent to it.

For site 8, which becomes experimentally accessible only when a distant defect is introduced, the accessibility is modulated by the global concerted displacement entailed by the first eigenvector of the covariance matrix. Strikingly, the amplitude of this mode changes dramatically on passing from the intact triangle to the defective one. As a result, both the population and the dwell time of the accessible states increase, suggesting that the enzyme may be subject to greater steric hindrances in the intact triangle compared to the defective one.

Our work is the initial step towards a better understanding of the allosteric properties of DNA nanostructures. Future studies should explore the relationship between global fluctuation modes and the design arrangement of staples and analyze quantitatively the kinetic behavior of enzymes acting on these nanostructures. These advancements ought to enable rational design of allosteric metamaterials whose shape and conformational fluctuations can be controlled by appropriate localized stimuli, a property that could be used, for instance, in nanodetectors and/or nanosensors.

Acknowledgments. This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 645684, the National Institutes of Health Grant S10OD020095 (to V.C.) and the National Science Foundation Grant ACI-1614804 (to V.C.). This research involved calculations carried out using Temple University's HPC resources, and was supported in part by the National Science Foundation through major research instrumentation grant number 1625061, and by the US Army Research

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Laboratory under contract number W911NF-16-2-0189.

Supporting Information

- Movie of trajectory of triangle (MovieS1.mp4)
- Movie of trajectory of defective triangle (MovieS2.mp4)
- Cadnano blueprint of DNA origami, further details of the methods (SI.pdf)

METHODS

System setup and simulation details The model used to parametrize the DNA is oxDNA [23, 24, 32], a coarse-grained model in which each nucleotide is represented by three rigidly connected interaction centers, developed to quantitatively reproduce the structural and kinetic experimental properties of DNA [24, 33–36]. Simulations were carried out with $T = 300K$ and at $1M$ monovalent salt concentration.

To obtain the initial triangular configurations, we started from the cadnano [37] constructs of ref [19], see Supplementary Figures S1-2, and converted them to oxDNA using a built-in script available in the oxDNA package. The initial structure presented the three trapeziums in sequential order; thus they were first roto-translated to obtain the final geometry. Over-stretched bonds were relaxed using MD by setting a maximum cutoff value of the oxDNA backbone potential to 4.8pN.

Each simulation was run on 140 processors for 1.5 months using LAMMPS [38, 39].

Docking of Hinp1I to GCGC sites and accessibility measurement

The first step consisted in superimposing an ideal oxDNA B-DNA helix of 10 bases to the atomistic DNA structure bound to HinP1I in the PDB structure (PBD code 2FL3 [31]), using a rigid body motion [27]. The ideal helix was then superimposed to each site of the DNA origami for each frame of the trajectory. The same transformation was applied to the protein.

To determine whether or not the protein docked to the restriction site overlaps with an adjacent DNA strand, two distance maps were built and compared. These maps, $max_{i,\theta}$ and $min_{i,\theta}$ are functions of the site base-pairs ($i = 1, \dots 10$) and of the angle around each of them $\theta \in [-\pi, \pi]$, see Supplementary Figure S3 for more details. The maximum distance map, $max_{i,\theta}$, reports the distance between the site and the protein particle that is farthest from it. The minimum distance map, $min_{i,\theta}$, reports the distance between the site and the closest particle from an adjacent DNA strand. Clearly, if $max_{i,\theta} < min_{i,\theta}$ for every i and θ, the enzyme satisfies the accessibility requirement and the configuration is classified as "accessible".

Gel analysis

In order to have an unbiased measurement of the experimentally determined percent of cut substrate of Fig. 3d, we measured for each band the intensity of the highest peak (i_{un}), representing uncut scaffolds, and of the two peaks representing the cut scaffold, split in two parts (i_{c1}, i_{c2}). Afterwards, we measured the cleavage percentage as $\frac{i_{c1}+i_{c2}}{i_{un}+i_{c1}+i_{c2}}$. For site 4, where the third band has an intensity below the background noise, i_{c2} was estimated through the ratio $\frac{i_{c2}}{i_{c1}+i_{c2}}$, which is equal to the scaffold fraction (this is also true for sites 12, 16, 19 within an error of 15%). We assigned an error bar at 20% using as reference the difference in estimations for the intensities considering or not the background noise.

[1] Yan, L.; Ravasio, R.; Brito, C.; Wyart, M. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2526–2531.
[2] Laskowski, R. A.; Gerick, F.; Thornton, J. M. *FEBS lett.* **2009**, *583*, 1692–1698.
[3] Changeux, J.-P.; Edelstein, S. J. *Science* **2005**, *308*, 1424–1428.
[4] Bertoldi, K.; Vitelli, V.; Christensen, J.; van Hecke, M. *Nat. Rev. Mater.* **2017**, *2*, 17066.
[5] Zhai, Z.; Wang, Y.; Jiang, H. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2032–2037.
[6] Yang, N.; Silverberg, J. L. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3590–3595.
[7] Treml, B.; Gillman, A.; Buskohl, P.; Vaia, R. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6916–6921.
[8] Wilhelm, D.; Bruck, J.; Qian, L. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 903–908.
[9] Silverberg, J. L.; Evans, A. A.; McLeod, L.; Hayward, R. C.; Hull, T.; Santangelo, C. D.; Cohen, I. *Science* **2014**, *345*, 647–650.
[10] McEvoy, M. A.; Correll, N. *Science* **2015**, *347*, 1261689.
[11] Rothemund, P. W. *Nature* **2006**, *440*, 297.

bioRxiv preprint doi: https://doi.org/10.1101/640847; this version posted May 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[12] Castro, C. E.; Kilchherr, F.; Kim, D.-N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H. *Nat. Methods* **2011**, *8*, 221.

[13] Douglas, S. M.; Dietz, H.; Liedl, T.; Högberg, B.; Graf, F.; Shih, W. M. *Nature* **2009**, *459*, 414.

[14] Seeman, N. C.; Lukeman, P. S. *Rep. Prog. Phys.* **2004**, *68*, 237.

[15] Dietz, H.; Douglas, S. M.; Shih, W. M. *Science* **2009**, *325*, 725–730.

[16] Sharma, R.; Schreck, J. S.; Romano, F.; Louis, A. A.; Doye, J. P. K. *ACS Nano* **2017**, *11*, 12426–12435.

[17] Marras, A. E.; Zhou, L.; Su, H.-J.; Castro, C. E. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 713–718.

[18] Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Lind-Thomsen, A.; Mamdouh, W.; Gothelf, K. V.; Besenbacher, F.; Kjems, J. *ACS nano* **2008**, *2*, 1213–1218.

[19] Stopar, A.; Coral, L.; Di Giacomo, S.; Adedeji, A. F.; Castronovo, M. *Nucleic Acids Res.* **2017**, *46*, 995–1006.

[20] Ding, B.; Deng, Z.; Yan, H.; Cabrini, S.; Zuckermann, R. N.; Bokor, J. *J. Am. Chem. Soc.* **2010**, *132*, 3248–3249.

[21] Zhang, Q.; Jiang, Q.; Li, N.; Dai, L.; Liu, Q.; Song, L.; Wang, J.; Li, Y.; Tian, J.; Ding, B.; Du, Y. *ACS nano* **2014**, *8*, 6633–6643.

[22] Gopinath, A.; Miyazono, E.; Faraon, A.; Rothemund, P. W. *Nature* **2016**, *535*, 401.

[23] Ouldridge, T. E.; Louis, A. A.; Doye, J. P. *Phys. Rev. Lett.* **2010**, *104*, 178101.

[24] Šulc, P.; Romano, F.; Ouldridge, T. E.; Rovigatti, L.; Doye, J. P. K.; Louis, A. A. *J. Chem. Phys* **2012**, *137*, 135101.

[25] Suma, A.; Micheletti, C. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2991–E2997.

[26] Coronel, L.; Suma, A.; Micheletti, C. *Nucleic Acids Res.* **2018**, *46*, 7533–7541.

[27] Kabsch, W. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography* **1976**, *32*, 922–923.

[28] Ke, Y.; Douglas, S. M.; Liu, M.; Sharma, J.; Cheng, A.; Leung, A.; Liu, Y.; Shih, W. M.; Yan, H. *J. Am. Chem. Soc.* **2009**, *131*, 15903–15908.

[29] García, A. E. *Phys. Rev. Lett.* **1992**, *68*, 2696–2699.

[30] Hess, B.; van der Spoel, D.; Lindahl, E.; Smith, J. C.; Shirts, M. R.; Bjelkmar, P.; Larsson, P.; Kasson, P. M.; Schulz, R.; Apostolov, R.; Pronk, S.; Pll, S. *Bioinformatics* **2013**, *29*, 845–854.

[31] Horton, J. R.; Zhang, X.; Maunus, R.; Yang, Z.; Wilson, G. G.; Roberts, R. J.; Cheng, X. *Nucleic Acids Res.* **2006**, *34*, 939–948.

[32] Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. *J. Chem. Phys* **2011**, *134*, 085101.

[33] Snodin, B. E. K.; Schreck, J. S.; Romano, F.; Louis, A. A.; Doye, J. P. K. *Nucleic Acids Res.* **2019**, *47*, 1585–1597.

[34] Engel, M. C.; Smith, D. M.; Jobst, M. A.; Sajfutdinow, M.; Liedl, T.; Romano, F.; Rovigatti, L.; Louis, A. A.; Doye, J. P. K. *ACS Nano* **2018**, *12*, 6734–6747.

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

[36] Shi, Z.; Castro, C. E.; Arya, G. *ACS Nano* **2017**, *11*, 4617–4630.

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

[38] Plimpton, S. *J. Comput. Phys.* **1995**, *117*, 1–19.

[39] Henrich, O.; Fosado, Y. A. G.; Curk, T.; Ouldridge, T. E. *Eur. Phys. J. E* **2018**, *41*, 57.