Polyhedral Nets Fold Cooperatively at High Temperature

Paul M. Dodd[a], **Pablo F. Damasceno**[b], **and Sharon C. Glotzer**[a,b,c,d]

[a]Chemical Engineering Department, University of Michigan, Ann Arbor, MI 48109, USA; [b]Applied Physics Program, University of Michigan, Ann Arbor, MI 48109, USA; [c]Dep. of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, USA; [d]Biointerfaces Institute, University of Michigan, Ann Arbor, MI 48109, USA

This manuscript was compiled on December 14, 2017

Low-dimensional objects such as molecular strands, ladders and sheets have intrinsic features that affect their propensity to fold into three-dimensional objects(1). Understanding this relationship remains a challenge for *de novo* design of functional structures, from protein machines (2) to self-folding robots. Using molecular dynamics simulations, we investigate the re-folding of the 24 possible two-dimensional unfoldings ("nets") of the simplest Platonic shapes – tetrahedron, cube, and octahedron. We show how net geometry and topology correlate with thermodynamic folding propensity by enumerating all possible folding pathways and their fluxes. We show there is a crossover temperature below which nets fold via non-native contacts (bonds that must break before the net can fold completely) and above which they fold via native contacts (bonds also present in the folded structure). Above the crossover temperature folding balances entropy reduction via elimination of internal degrees of freedom with potential energy gain via local edge bonding – even though the folding process occurs out of equilibrium. We further show that this balance is universal and cooperative. Exploiting this universality, we devised a numerical method to efficiently compute all high-temperature folding pathways for any net, allowing us to predict high-propensity nets for the remaining two Platonic solids – the dodecahedron and icosahedron – from among their combined 86,760 nets. Our results provide a general heuristic for the design of two-dimensional objects capable of stochastically folding into target 3-dimensional geometries, and suggest a universal relation between local geometric characteristics and global folding propensity in stochastic systems.

self-folding | origami | polyhedra nets | ... ?

In the 16th century, the Dutch artist Albrecht Dürer investigated which 2D cuts of non-overlapping, edge-joined polygons could be folded into Platonic and Archimedean polyhedra(3, 4). Dürer cuts were later called "nets" and are still a subject of study in several fields(3–5). Here we create nets from polyhedra via a process known as edge unfolding. In edge unfolding, one cuts along a set of prechosen edges, called the cutting tree, of a polyhedron (e.g. a cube) to create a single, contiguous flat 2D sheet of connected (e.g. square) faces; this sheet is the net. For the Platonic shapes, whose nets were enumerated by Buekenhout (6), this process can be repeated exhaustively until all distinct nets are discovered (see Methods for details). Here we are interested in the self-folding of these nets.

Self-folding origami adds a modern twist to the ancient art of origami. It provides a process to achieve complex geometries from simpler, low-dimensional objects. A type of self assembly, self-folding (or simply folding) has applications ranging from drug delivery(7) to robotics(8). Recent studies have leveraged physical forces to achieve controllable folding, from light(9), pH(10), and capillary forces(11), to cellular traction(12) and thermal expansion(13) . In natural systems, the canonical example of folding occurs for proteins. There, a string of amino acids, thermodynamically, navigate from a denatured (unfolded) state to a natured (folded) one. However, even after many decades of study, a universal relationship between molecular sequence and folded state – which could provide crucial insight into the causes and potential treatments of many diseases – remain out of reach(1, 14, 15). In larger scale artificial systems, this connection between local (geometry) and global (folding) structure can be investigated with greater control. In the macroscopic folding of kirigami sheets, for example, the effect of different cut patterns on the material's stress-strain behavior has been elucidated (1, 14, 15). Also for kirigami, the "inverse design problem" of finding cuts leading to the folding of a particular target structure has been recently solved(16). For sub-millimeter sized capsules formed from self-folding polyhedra, Gracias et al.(11, 17, 18) showed that nets fold with higher yield when they are compact.

Nevertheless, connections between such deterministic, irreversible foldings and those occurring stochastically and reversibly on nanometer to micron scales remains elusive. In addition to potentially giving insights into the thermodynamics of folding in naturally occurring systems, these connections could also provide a route for the fabrication of anisotropic Brownian shells, paving the way for the self-assembly of complex crystals from nano and colloidal building blocks(19–22) capable of encapsulating cargo. In this work, we study the thermodynamic foldability of 2D nets of all five Platonic solids in an effort to understand how 3D objects can be formed through the use of stochastic forces and origami-like folding processes.

Despite being the simplest and most symmetric 3-dimensional polytopes, the family of Platonic shapes suffices to demonstrate the rapid expansion of design space as the shape increases in complexity: a tetrahedron has two possible net representations, cubes and octahedra have 11 nets each, and both dodecahedra and icosahedra each have 43,380 distinct nets. We list all 86,784 nets for the Platonic solids in an online database (23), identified as follows. Beginning with a target Platonic shape (Fig. 1), we construct a graph whose vertices and edges correspond to those in the polyhedron. This mapping of the shape to a graph facilitates the exhaustive search of all distinct nets

Significance Statement

Authors must submit a 120-word maximum statement about the significance of their research paper written at a level understandable to an undergraduate educated scientist outside their field of speciality. The primary goal of the Significance Statement is to explain the relevance of the work in broad context to a broad readership. The Significance Statement appears in the paper itself and is required for all research papers.

All authors designed and planned research. P.M.D wrote and performed numerical simulations. All authors analyzed the data and wrote the paper.

The authors declare no conflict of interest.

[2]To whom correspondence should be addressed. E-mail: sglotzer@umich.edu

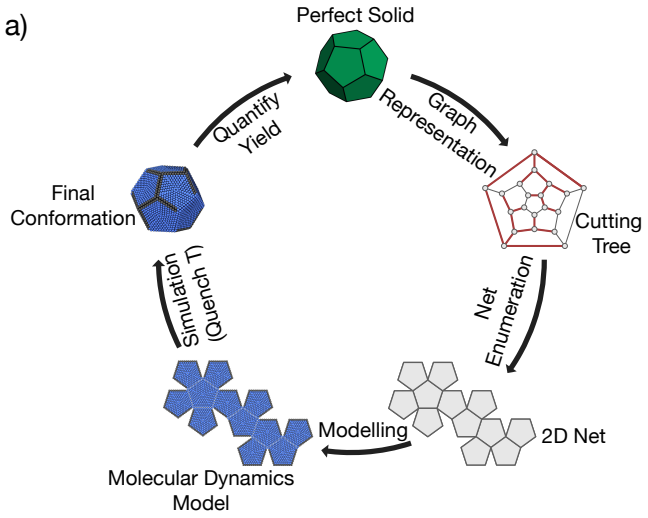

Fig. 1. Workflow used in this study. Starting with the target 3D shape, the graph representation is used to enumerate all the different ways to cut the solid. The vertices in the graph correspond to the vertices of the polyhedron and the edges in the graph correspond to the edges of the polyhedron. The edges marked in red represent the edges that are cut on the shape. Using the information in the cutting graph we obtain the 2-dimensional net of the polyhedron. Each face is modeled as a union of spheres held rigidly together. The faces are tethered together using harmonic springs. The blue spheres interact via a purely repulsive WCA potential and the gray spheres interact via the Lennard-Jones potential, which includes both repulsion and attraction. The blue and gray spheres interact repulsively. The simulations are initialized at high temperature and then the temperature is quenched to low temperature. Once the final temperature is reached the final configuration is compared to the target shape and the folding probability (yield) is calculated. **also include something about the networks here!!**

by allowing spanning tree enumeration(3). We model each net as a sheet of rigid polygons connected to adjacent polygons on the net via harmonic springs and use Langevin molecular dynamics (MD) simulations to simulate the influence of thermal fluctuations on a single net suspended in implicit solvent. We assign non-specific attractive (sticky) interactions between all edges of a net (details in Methods), so that the folded state – that is, the net re-folded into the original polyhedron – is also the ground state (zero temperature) configuration. Because the conformational entropy of any unfolded net is higher than that of the folded net, the ground state is also the minimum free energy configuration.

The non-specificity of the edge-edge interactions means that, as in protein folding, both native and non-native contacts between edges are possible. Consequently, when the system temperature is rapidly decreased (quenched) from high temperature, where the unfolded net is the preferred equilibrium structure, to low temperature, where the folded net is the preferred equilibrium structure, kinetic traps from mis-folds are possible, and some nets may fail to fold. Here we ask the following question: which nets – among the many thousands of candidates – offer the highest folding propensity (yield) for a particular folded polyhedron?

1. Results.

The tetrahedron has two distinct nets – triangular and linear (Fig. 2a). Out of 125 simulations for each of two quench rates (see Methods), we found that all triangular nets folded into the target tetrahedron. In contrast, only 68 of 125 linear nets successfully folded for the fastest cooling rate; the rest failed due to mis-folds. Similar simulations for

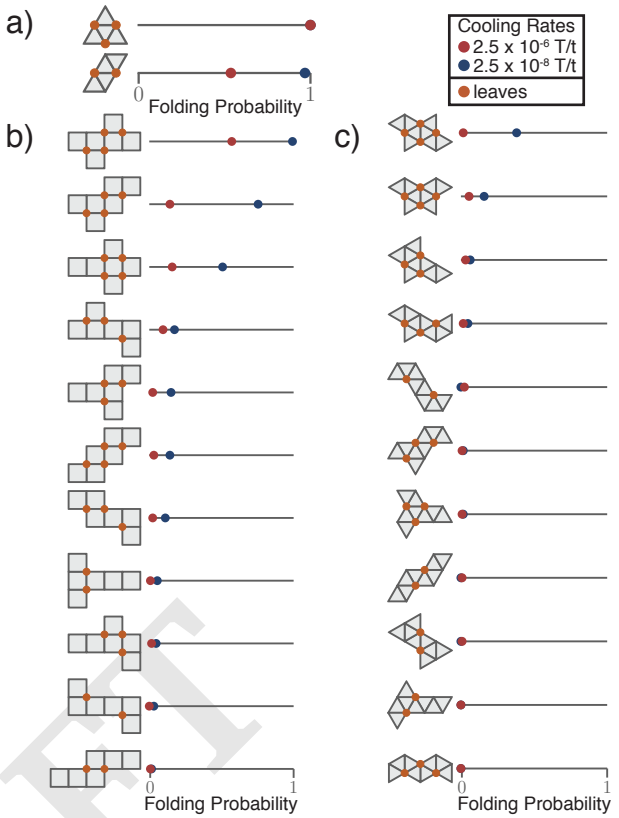

Fig. 2. Folding yield defined by the fraction of simulations that reached the folded state for the tetrahedron, cube, octahedron, and 11 of the 43,380 dodecahedron nets. The folding yield for the two tetrahedron nets. The uppermost net folds 100% of the time because there are no kinetic traps. The lower most net folds in a range between 54%-98%. **Cube:** The folding yield for all 11 nets of the cube ordered from highest yield (top) to lowest (bottom). The folding yield is dependent on the geometry of the net. Generally the more compact nets fold better and nets that have a high number of leaves on their cutting tree and a smaller diameter usually folds best.

each of the 11 nets of both the cube and octahedron revealed only a few nets able to reliably fold into their target shape. For the cube (Fig. 2b), it is clear that the number of leaves alone (i.e. the vertices corresponding to leaves on the cutting tree are shown in green) suffice to warrant high folding propensity, although some correlation exists.

For the octahedron nets (Fig. 2c), we found that folding commonly failed, regardless of quench rate. Instead, many nets partially folded into a concave, boat-like polyhedron with the same number of edge-edge contacts as the octahedron. This competing structure, which is less symmetric than the octahedron, also has higher rotational entropy, resulting in a lower free energy than the octahedron. However, under rapid quench rates, once sufficient edge-edge contacts form to create the boat structure, it is difficult for enough of them to unbind so that the folded state can be sampled. A similar competition between degenerate structures of equivalent potential energy was reported for clusters of six attractive spherical colloids(24), where symmetry breaking leads to a higher rotational entropy for the less symmetric conformation, resulting in lower free energy. **Describe a few of the differences here.**

Figure 2 shows that, despite having the same ground-state energy, not all nets of a polyhedron are created equal. Indeed, even nets differing only by the location of a single face can have folding propensities reduced from nearly 100% to approximately 20%. What causes one shape to fold nearly perfectly every time while a slightly different one

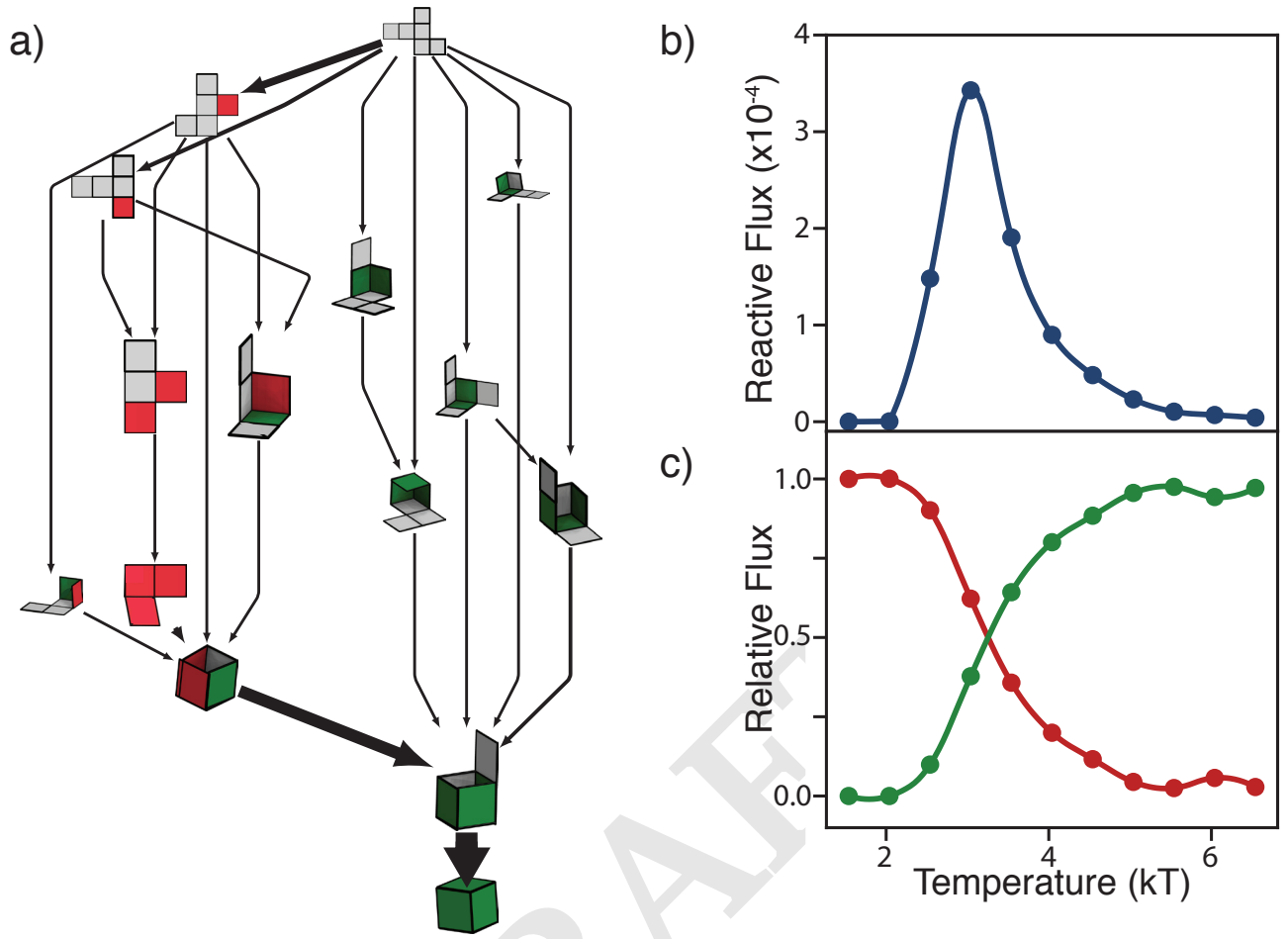

Fig. 3. (a) Folding pathways for one of the cubic nets. The network can be separated into two regions: native contacts only(left) and non-native containing (right). The pathways that contain only native contacts follow a "sequential" pathway where one face folds at a time. The non-native pathways are observed to behave in two different ways i) they can bring 2 edges that need to bind close together and later the non-native bonds break leaving the native contact ii) the polyhedron folds as if the non-native contact was not there and later this non-native contact breaks and the folding can complete to the final target structure. (b) The total probability flux of all the pathways. (c) The relative amount of flux going through pathways that use the native-exclusive (green curve) pathways and the non-native inclusive pathways (red curve).

fails to do so almost as frequently?

To answer this question, we computed the folding pathways using Markov state models (MSM)(25, 26). Two representative folding networks created using this method are shown in Fig. 3. In the case of one of the cube nets (Fig. 3a), there are two dominant pathways along which the net can successfully fold: one pathway uses only intermediate states having only native contacts (shown in green); the other main pathway uses non-native contacts (shown in red). We observe that non-native contacts play either an "active" or a "passive" role in the folding process. We say that non-native contacts are active when a native contact is brought together by the non-native contact. In general, this happens when a non-local bond needs to be formed to stabilize the structure until the native contact is formed. The "passive" type happens when a non-native contact is formed and then the net folds using native contacts, as if the non-native contact did not exist. In either case, at the end the non-native contact eventually breaks and the net achieves the final target structure.

We observe that the relative flux through the folding pathways depends on temperature (Figs 3b, 3d). For the tetrahedron, cube, and octahedron, at low temperatures the nets tend to fold through pathways that utilize non-native contacts. As the temperature increases we observed a transition to pathways that only use native contacts. We observe a correlation between both the total flux and the number of high temperature pathways of a given net with the folding yield.

Representative pathways calculated at fixed high temperature for the cube and octahedron nets are shown in Fig. 4. We see that the pathways in the high temperature limit are *cooperative*(27): nets fold using local bonds first, which in turn bring non-local contacts close together to be folded next. This suggests that as each net folds, it freezes out the fewest possible number of degrees of freedom, thereby maximizing the conformational entropy along the folding pathways.

To quantify this trade-off between maximization of the degrees of freedom and the formation of native contacts at high temperature, we plot all of the pathways (both native and non-native) as a function of N – the number of internal degrees of freedom – and Q – the number of native contacts. The upper right corner of the plot is a geometrically forbidden region arising from the fact that the net can not gain a bond without losing at least one degree of freedom.

Remarkably, all the nets follow a folding pathway that achieves a narrow balance between freezing degrees of freedom and gain of energy. In practical terms, high temperature folding happens locally such that, at each step of the process, the system strives to maximize its conformational entropy. From this observation, we can hypothesize the following mechanism for the folding of general nets at high

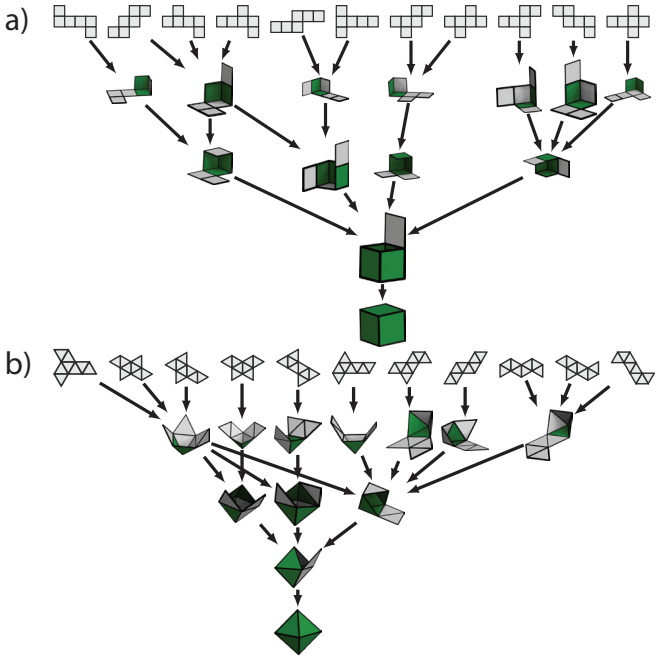

Fig. 4. Combined dominant pathways (a) the cube (b) the octahedron taken at high temperature. In both (a) and (b) we observe that the dominant pathways are sequential, generally speaking one face folding at a time. All of the observed pathways include only native contacts. If we project ALL the pathways of ALL the nets onto the two order parameters native contacts and number of internal degrees of freedom, we find that each net will trade their degrees of freedom in approximately the same way. This is demonstrated for the tetrahedron in (d) the cube in (e) and octahedron in (f).

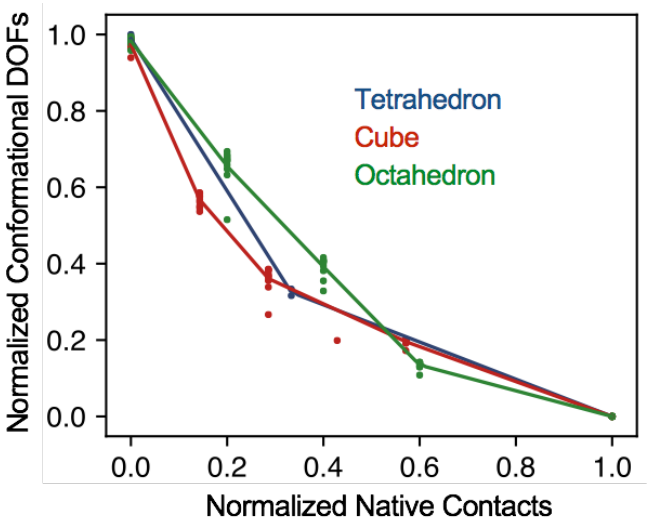

Fig. 5. Combined dominant pathways (a) the tetrahedron (b) the cube (c) the octahedron taken at a temperature $T > T^*$. In both a-c we observe that the dominant pathways are sequential, generally speaking one face folding at a time. All of the observed pathways include only native contacts. If we project ALL the pathways of ALL the nets onto the two order parameters native contacts and number of internal degrees of freedom, we find that each net will trade their degrees of freedom in approximately the same way. This is demonstrated for the tetrahedron in (d) the cube in (e) and octahedron in (f). The size of each data point corresponds to the calculated likelihood that any net will fold along a path that has any intermediate with those order values of N and Q, calculated by weighting averages of N while holding Q constant. Lines are drawn to guide the eye.

temperature. Folding should primarily happen: i) via nearby (local) connections – favoring nets with higher numbers of leaves; ii) along one of the optimal trade-off paths – favoring nets with high degeneracy in the number of such optimal paths.

In figure 6, we confirm that these two principles are obeyed for the cube and octahedron nets. With this information in hand, we then used these principles to predict, which nets among the 86,760 possible for the dodecahedron and icosahedron have lowest and highest folding yield (see SI for specific nets). The corresponding folding propensities are plotted in the rightmost columns of Figure 6. They demonstrate that these principles are good predictors of folding yield and can be used for the design or selection of nets based purely on *a priori* computable quantities.

2. Discussion and Conclusion

The observed preference for native contact pathways at high temperature is not unique to polyhedral nets. Several small proteins have been observed in simulation to fold via native-only contacts when close to their melting temperature. At lower temperature (or high hydrophobicity) the pathways shift to a hydrophobic collapse, in which non-native contacts are found to form followed by further rearrangements leading to the native state(28). Similarly at high temperature patchy colloids assemble icosahedra via monomeric addition whereas at low temperature particles first aggregate into large clusters, forming undesired bonds that later rearrange towards the correct assembly(29). Finally, systems of colloidal sticky spheres prefer to form the same concave, boat-like conformation that we observe to fold from octahedron nets(24).

Our toy model therefore draws connections between the macroscopic irreversible folding of polyhedra(11) and the folding of amino acids(**?**), assembly of patchy(29) and colloidal particles(24). The identified trade-off that dictates high temperature folding supplies guiding principles for the assembly of 3-dimensional complex geometries from potentially simpler-to-fabricate 2-dimensional nets. In fact, we demonstrated that judicious pathway engineering, whether via the creation of specific interactions designed to prevent traps, or whether the selection of nets with certain characteristics. We expect these results to impact future experiments on nanoscopic foldings such as from even smaller graphene sheets(15), graphene oxide layers (14), or DNA-origami polyhedral nets(30).

Materials and Methods

Enumeration of Polyhedral Nets. We create nets from polyhedra via a process known as edge-unfolding. In edge-unfolding, one cuts along a set of pre-chosen edges, called the cutting tree, of a polyhedron (e.g. a cube) to create a single, contiguous flat 2D sheet of connected (square) faces: a net. We enumerate all of the nets of each polyhedron by first inputing the skeleton-1 graph of the polyhedron. Then random weights were assigned to each edge on the interval [0,1]. The minimal spanning tree was found using Kruskal's algorithm. We then converted the spanning tree to a net and added it to the database if it did not already exist. We ran this loop for many iterations until we found all the nets for each shape.

Quantifying yield. To quantify the folding yield we ran 125 simulations starting from a high temperature and quench the temperature to near zero ($T^* = 0.1$). We then defined the yield as the fraction of simulations that folded. We defined the folding propensity as the average number of native contacts averaged over all of the quenching simulations we ran. If all nets folded perfectly in all runs, the folding propensity is one. We linearly quenched 125 systems from $0.1 \leq T \leq T_m + 2.5$, where T_m is the folding temperature defined as the maximum melting temperature among all nets for a given target shape. The T_m for each shape is listed in Table 1 of the SI. We used two different cooling rates to investigate their influence in the folding yield: $2.5\times10^{-6}T/t$ and $2.5\times10^{-8}T/t$.

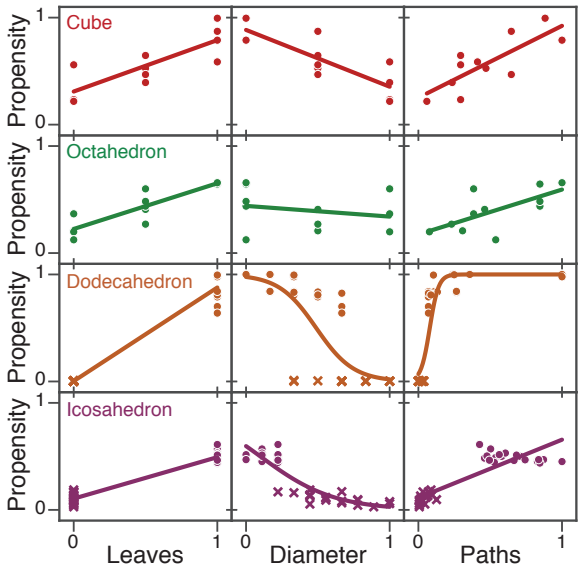

Fig. 6. (a) The average fraction of native contacts observed over 125 simulations at the slowest cooling rate for each net. Cubic nets are shown in red and octahedron nets are shown in blue. (will add the other data) There is a strong correlation between the average fraction of native contacts and the folding probability. The average fraction of native contacts helps distinguish poor folders. (b) The average fraction of native contacts vs the number of high temperature pathways shown in Fig. **??**a-c. Folding strongly correlates with the number of pathways geometrically accessible to the net. (c) The average fraction for the tetrahedron, cube, octahedron, and dodecahedron. Folding positively correlates with the number of leaves.

Markov State Models. Markov State Models (MSM)(31) have been used to study protein folding(25, 26) and virus capsid assembly(32) and can provide a detailed view into the dynamics and thermodynamics of the assembly landscape. Each simulation snapshot needs to be classified as a discrete state. In principal the configuration of the net is completely specified by the dihedral angles of each hinge and is therefore a good set of collective variables. We differentiate between folding 'up' and folding 'down' by keeping track of the dihedral angle on the interval $[0, 2\pi]$, by defining a local coordinate system on the corner of each face of the net. As the simulation is running we also compute the energy between each pair of free edges (edges not part of a hinge) on the net, if the potential energy between two edges is less than $E_{bond} = -5\epsilon$, then the edges are considered to be bonded, and a list of bonded edges is recorded along with the the list of dihedral angles. While these measures are enough to specify each state it is often desirable to account for the symmetry and blank of each intermediate. To do this we turn the molecular dynamics model into a graph, G, and by looking at the graph automorphisms we can determine the different symmetries to compare intermediates. Two intermediates are then clustered using DBSCAN with a variation of the Manhattan metric, $d(i,j) = \min_{\alpha \in \mathrm{Aut}(G_{\mathrm{net}})} \max |a_{\alpha(i),k} - a_{j,k}|$, which is a way to compare the dihedral angles of different states taking into account the symmetry of the net. The symmetry of each intermediate was analyzed and if the symmetry could lead to a relabeling of the vertices of the net then the clusters were merged. To build the MSM we ran 125 independent NVT simulations for 12.5×10^6 steps. The results became seeds for new independent simulations for a total of 1875 (14×125) trajectories (or, equivalently, 2.3×10^{10} time steps) for each net and each temperature. Bonded edges and dihedral angles for each hinge were computed every 10 time steps and combined to define a state in the MSM described above. The lag time was found by the standard protocol of identifying the time at which the eigenvalues of the transition probability matrix become constant (Figure SIXX). We used Transition Path Theory(26, 33, 34) to determine the reactive flux, $f_{ij} = q_j^+ \pi_i P_{ij} q_i^-$, along the dominant pathways. The dominant paths were computed using the 'bottle neck' algorithm(26, 33).

Folding Parameters. The number of native contacts, Q, were calculated by counting the number of edges that were bonded to the correct corresponding edge according to the criteria described above. The non-native contacts were calculated similarly. In general the number of degrees of freedom can be difficult to calculated because it can be difficult to deduce which constraints are redundant. In general one can use the pebble game(35), however if the linkage is not generic or has point group symmetries the pebble game will underestimate the number of degrees of freedom(**?**). First we applied the pebble game to each intermediate, then for closed loop motifs we applied an extensor analysis to determine the number of degrees of freedom(36), finally for intermediates that had high degrees of symmetry the intermediates were checked by hand.

ACKNOWLEDGMENTS. This research was supported in part through computational resources and services supported by Advanced Research Computing at the University of Michigan, Ann Arbor. This work was supported as part of the Center for Bio-Inspired Energy Science, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award # DE-SC0000989. Research supported by the National Science Foundation, Emerging Frontiers in Research and Innovation Award # EFRI-1240264.

1. Ken A Dill, S B Ozkan, T R Weikl, J D Chodera, and V A Voelz. The protein folding problem: when will it be solved? *Current Opinion in Structural Biology*, 17(3):342–6, 2007. ISSN 0959-440X. .
2. Po-Ssu Huang, Scott E Boyken, and David Baker. The coming of age of de novo protein design. *Nature*, 537(7620):320–327, 2016.
3. Erik D Demaine and Joseph O'Rourke. *Geometric Folding Algorithms: Linkages, Origami, Polyhedra*. Cambridge University Press, New York, NY, USA, 2007. ISBN 0521857570.
4. Wolfram Schlickenrieder. *Nets of Polyhedra*. PhD thesis, Berlin, Technische Universitat, 1997.
5. G. C. Shephard. Convex polytopes with convex nets. *Mathematical Proceedings of the Cambridge Philosophical Society*, 78(October):389, 1975. ISSN 0305-0041. .
6. Francis Buekenhout and Monique Parker. The number of nets of the regular convex polytopes in dimension? 4. *Discrete mathematics*, 186(1-3):69–94, 1998.
7. Rohan Fernandes and DH Gracias. Self-folding polymeric containers for encapsulation and delivery of drugs. *Advanced drug delivery reviews*, 64(14):1–11, nov 2012. ISSN 1872-8294. .
8. S. Felton, M. Tolley, E. Demaine, D. Rus, and R. Wood. A method for building self-folding machines. *Science*, 345(6197):644–646, 2014. ISSN 0036-8075. .
9. Ying Liu, Julie K. Boyles, Jan Genzer, and Michael D. Dickey. Self-folding of polymer sheets using local light absorption. *Soft Matter*, 8(6):1764, 2012. ISSN 1744-683X. .
10. Tae Soup Shim, Shin-Hyun Kim, Chul-Joon Heo, Hwan Chul Jeon, and Seung-Man Yang. Controlled Origami Folding of Hydrogel Bilayers with Sustained Reversibility for Robust Microcarriers. *Angewandte Chemie*, 124(6):1449–1452, feb 2012. ISSN 00448249. .
11. Shivendra Pandey and Margaret Ewing. Algorithmic design of self-folding polyhedra. *Proceedings of the . . .*, 108(50):19885–90, dec 2011. ISSN 1091-6490. .
12. Kaori Kuribayashi-Shigetomi, Hiroaki Onoe, and Shoji Takeuchi. Cell origami: self-folding of three-dimensional cell-laden microstructures driven by cell traction force. *PloS one*, 7(12): e51085, jan 2012. ISSN 1932-6203. .
13. VB Shenoy and DH Gracias. Self-folding thin-film materials: From nanopolyhedra to graphene origami. *MRS bulletin*, 37(09):847–854, sep 2012. ISSN 0883-7694. .
14. Terry C Shyu, Pablo F Damasceno, Paul M Dodd, Aaron Lamoureux, Lizhi Xu, Matthew Shlian, Max Shtein, Sharon C Glotzer, and Nicholas A Kotov. A kirigami approach to engineering elasticity in nanocomposites through patterned defects. *Nat Mater*, advance on, jun 2015. ISSN 1476-4660.
15. Melina K. Blees, Arthur W. Barnard, Peter a. Rose, Samantha P. Roberts, Kathryn L. McGill, Pinshane Y. Huang, Alexander R. Ruyack, Joshua W. Kevek, Bryce Kobrin, David a. Muller, and Paul L. McEuen. Graphene kirigami. *Nature*, 2015. ISSN 0028-0836. .
16. Daniel M. Sussman, Yigil Cho, Toen Castle, Xingting Gong, Euiyeon Jung, Shu Yang, and Randall D. Kamien. Algorithmic lattice kirigami: A route to pluripotent materials. *Proceedings of the National Academy of Sciences of the United States of America*, 112(24):7449–7453, 2015. ISSN 0027-8424. .
17. Anum Azam, Timothy G Leong, Aasiyeh M Zarafshar, and David H Gracias. Compactness determines the success of cube and octahedron self-assembly. *PloS one*, 4(2):e4451, jan 2009. ISSN 1932-6203. .
18. Ryan Kaplan, Joseph Klobušický, and Shivendra Pandey. Building Polyhedra by Self-Assembly: Theory and Experiment. *Artificial Life*, 31:1–31, 2014. .
19. Sharon C Glotzer and Michael J Solomon. Anisotropy of building blocks and their assembly into complex structures. *Nature materials*, 6(8):557–562, 2007.
20. Umang Agarwal and Fernando A Escobedo. Mesophase behaviour of polyhedral particles. *Nature materials*, 10(3):230–235, 2011.
21. Joost de Graaf and Liberato Manna. A roadmap for the assembly of polyhedral particles. *Science*, 337(6093):417–418, 2012.
22. Pablo F Damasceno, Michael Engel, and Sharon C Glotzer. Predictive self-assembly of polyhedra into complex structures. *Science (New York, N.Y.)*, 337(6093):453–7, jul 2012. ISSN 1095-9203. .
23.
24. Guangnan Meng, Natalie Arkus, Michael P. Brenner, and Vinothan N. Manoharan. The free-energy landscape of clusters of attractive hard spheres. *Science*, 327(5965):560–563, 2010. ISSN 0036-8075. .

25. Gregory R Bowman and Vijay S Pande. Protein folded states are kinetic hubs. *Proceedings of the National Academy of Sciences*, 107(38):16749–16749, aug 2010. ISSN 0027-8424. .
26. Frank Noé, Christof Schütte, Eric Vanden-Eijnden, Lothar Reich, and Thomas R Weikl. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45): 19011–6, nov 2009. ISSN 1091-6490. .
27. Ken A Dill, Klaus M Fiebig, and Hue Sun Chan. Cooperativity in protein-folding kinetics. *Proceedings of the National Academy of Sciences*, 90(5):1942–1946, 1993.
28. Nicholas D. Socci, José Nelson Onuchic, and Peter G. Wolynes. Protein folding mechanisms and the multidimensional folding funnel. *Proteins: Structure, Function and Genetics*, 32(2): 136–158, 1998. ISSN 08873585. .
29. Andrew W Long and Andrew L Ferguson. Nonlinear Machine Learning of Patchy Colloid Self-Assembly Pathways and Mechanisms. *The journal of physical chemistry. B*, apr 2014. ISSN 1520-5207. .
30. Fei Zhang, Shuoxing Jiang, Siyu Wu, Yulin Li, Chengde Mao, Yan Liu, and Hao Yan. Complex wireframe dna origami nanostructures with multi-arm junction vertices. *Nat Nano*, 10(9):779–784, 09 2015.
31. William C. Swope, Jed W. Pitera, and Frank Suits. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory †. *The Journal of Physical Chemistry B*, 108(21): 6571–6581, may 2004. ISSN 1520-6106. .
32. Matthew R. Perkett and Michael F. Hagan. Using markov state models to study self-assembly. *The Journal of Chemical Physics*, 140(21):214101, 2014. . URL https://doi.org/10.1063/1.4878494.
33. Philipp Metzner, C Schütte, and E Vanden-Eijnden. Transition path theory for Markov jump processes. *Multiscale Modeling & Simulation*, 7(3):1192–1219, 2009.
34. Weinan E and Eric Vanden-Eijnden. Transition-path theory and path-finding algorithms for the study of rare events. *Annual review of physical chemistry*, 61:391–420, jan 2010. ISSN 1545-1593. .
35. Audrey Lee and Ileana Streinu. Pebble game algorithms and sparse graphs. *Discrete Mathematics*, 308(8):1425–1437, 2008. ISSN 0012365X. .
36. Tiong-Seng Tay. Rigidity of multi-graphs. I. Linking rigid bodies in n-space. *Journal of Combinatorial Theory, Series B*, 36(1):95–112, 1984. ISSN 00958956. .