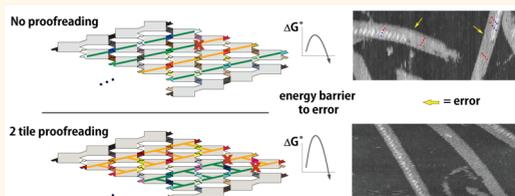


Increasing Redundancy Exponentially Reduces Error Rates during Algorithmic Self-Assembly

Rebecca Schulman,^{*,†,‡} Christina Wright,^{*,||} and Erik Winfree[§]

[†]Computation and Neural Systems, California Institute of Technology, Pasadena, California 91125, United States, [‡]Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States, and [§]Computer Science, Computation and Neural Systems, and Bioengineering, California Institute of Technology, Pasadena, California 91125, United States. [‡]Present address: Chemical and Biomolecular Engineering and Computer Science, Johns Hopkins University. ^{||}Present address: Google, Inc.

ABSTRACT While biology demonstrates that molecules can reliably transfer information and compute, design principles for implementing complex molecular computations *in vitro* are still being developed. In electronic computers, large-scale computation is made possible by redundancy, which allows errors to be detected and corrected. Increasing the amount of redundancy can exponentially reduce errors. Here, we use algorithmic self-assembly, a generalization of crystal growth in which the self-assembly process executes a program for growing an object, to examine experimentally whether redundancy can analogously reduce the rate at which errors occur during molecular self-assembly. We designed DNA double-crossover molecules to algorithmically self-assemble ribbon crystals that repeatedly copy a short bitstring, and we measured the error rate when each bit is encoded by 1 molecule, or redundantly encoded by 2, 3, or 4 molecules. Under our experimental conditions, each additional level of redundancy decreases the bitwise error rate by a factor of roughly 3, with the 4-redundant encoding yielding an error rate less than 0.1%. While theory and simulation predict that larger improvements in error rates are possible, our results already suggest that by using sufficient redundancy it may be possible to algorithmically self-assemble micrometer-sized objects with programmable, nanometer-scale features.



KEYWORDS: DNA nanotechnology · algorithmic self-assembly · error correction · crystal growth

Reliable digital computation and communication have changed electronics; biology demonstrates similarly that reliable complex molecular information transfer and computation could be equally transformative. Biological organisms reliably self-assemble large-scale objects such as the ribosome,¹ execute long signal-transduction cascades like those in animal development,² dynamically control cytoskeletal organization^{3,4} and tightly control metabolism.⁵ To engineer systems with similar capacities, it is necessary to develop design principles for chemical systems that can reliably process and transmit information.

In self-assembly, information encoded within molecules directs assembly. If each molecule is only used once in a final product, as in protein folding or DNA origami,⁶ self-assembly does not require recurring information transfer. However, using each species only once is not practical for the

assembly of micron-scale (or larger) objects from molecules. Biology generally reuses components to assemble such complex objects, as in the self-assembly of networks of actin,⁷ microtubules,⁸ viruses⁹ or in biomineralization.¹⁰ In these cases, the molecules encode a pathway of assembly, and reliable information transfer along the pathway is necessary to produce a correctly formed final product. Here we investigate how information can be reliably transferred during a designed self-assembly reaction.

Algorithmic self-assembly, a generalization of crystal growth, has been proposed as a general method for information-guided synthesis of supramolecular objects.¹¹ In algorithmic self-assembly, a set of molecular tiles containing four binding sites with particular affinities executes a “program” for the assembly of an object. This mechanism is surprisingly powerful; abstractly, the assembly of such tiles (Wang tiles) into a lattice can simulate universal computation,¹² and in

* Address correspondence to rschulm3@jhu.edu.

Received for review December 31, 2014 and accepted April 29, 2015.

Published online 10.1021/nn507493s

© XXXX American Chemical Society

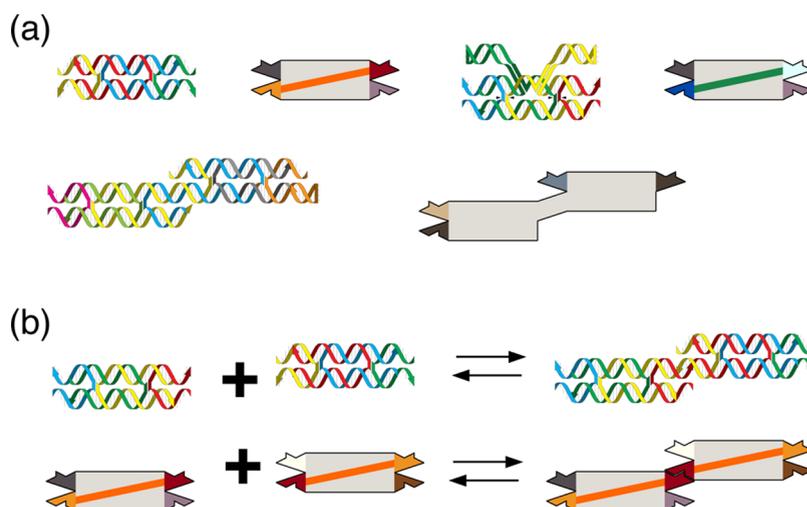


Figure 1. DNA tiles and tile hybridization. (a) The three kinds of DNA tiles (based on DAO-E double-crossover molecules) used in this work. Each DNA tile consists of 4–6 synthetic DNA strands that self-assemble into the shown structures as dictated by Watson–Crick complementarity of strand sequences. The upper right DNA tile contains two hairpins (which protrude into and out of the plane of the diagram) that provide contrast during microscopy. In composite diagrams, the double-stranded “core” is represented as a rectangle and the single-stranded “sticky ends” are shown as claws. Claws with the same color and complementary shapes represent sticky ends with complementary sequences. The colored stripe indicates the logic of tile interactions and information flow, as described in Figure 2. (b) Tiles bind by hybridization of their sticky ends. Noncomplementary sticky ends tend not to hybridize.

principle algorithmic self-assembly can be used to efficiently assemble arbitrary shapes.^{13–15}

DNA multicrossover molecules^{16–22} with binding sites consisting of single-stranded DNA segments have been used as tiles for algorithmic self-assembly (Figure 1). Algorithmic self-assembly of DNA tiles has been demonstrated in one-dimension^{23,24} as well as in two-dimensions, where it has been used to produce complex, aperiodic patterns.^{25–28} While these demonstrations indicate that the principles of algorithmic self-assembly are sound, the measured rate of assembly errors were generally too high for construction of large, perfect structures, because for most assembly targets of interest, even a single assembly error can result in a malformed final product.^{25,29} High error rates therefore limit the applicability of algorithmic self-assembly.

Efforts to decrease errors can focus either on the physical conditions of assembly (such as temperature, molecular design, or tile concentrations) and/or on logical properties of the tile set design, which determine what kinds of violations of assembly rules produce errors in the final structure. As an example of the latter, tile sets that propagate information only in a single direction (such as copying binary patterns from layer to layer) tend to have lower error rates, ranging from 0.01 to 2% per tile.^{26,27,30} In contrast, for tile sets that propagate and process information in two dimensions during most assembly steps, error rates have ranged from 1 to 10% per tile.^{25,27,28} This difference raises the question of whether some tile sets are logically more error-prone, and whether there are general techniques for designing tile sets that create desired shapes and patterns with lower error rates.

In electronic circuit design, redundancy can reduce error rates during computation. Because increasing the complexity of an electronic circuit by a linear factor can decrease the error rate exponentially,³¹ low error rates do not require a large increase in circuit size even when every logic gate is error-prone. “Proofreading” has been postulated as a technique to redundantly transfer information during algorithmic self-assembly.³² Analysis and simulations suggest that proofreading should also exponentially reduce the self-assembly error rate while increasing the size of the product by just a linear factor.^{32,33} Proofreading tile sets require a quadratic increase in the number of tile types in the general case of two-dimensional information transfer, but just a linear increase for one-dimensional information transfer. Here we test these predictions experimentally by measuring self-assembly error rates during a fundamental type of information transfer, sequence copying. We use four types of DNA tile ribbons³⁴ that copy a sequence of tile types as they self-assemble with four levels of redundancy (Figure 2). (In the following, we refer to a sequence of tile types along the cross-section of a ribbon as a “molecular sequence” or simply a “sequence.”) We compare the error rates for each type of ribbon to determine how redundancy affects assembly error rates.

Error-free algorithmic self-assembly consists of a series of steps in which a tile binds to a growing crystal by at least two matching sticky ends. Under slightly supersaturated conditions, the attachment of a tile by two sticky ends is energetically favorable, but the attachment of a tile to a crystal by fewer than two sticky ends is unfavorable.^{29,35} In practice, however,

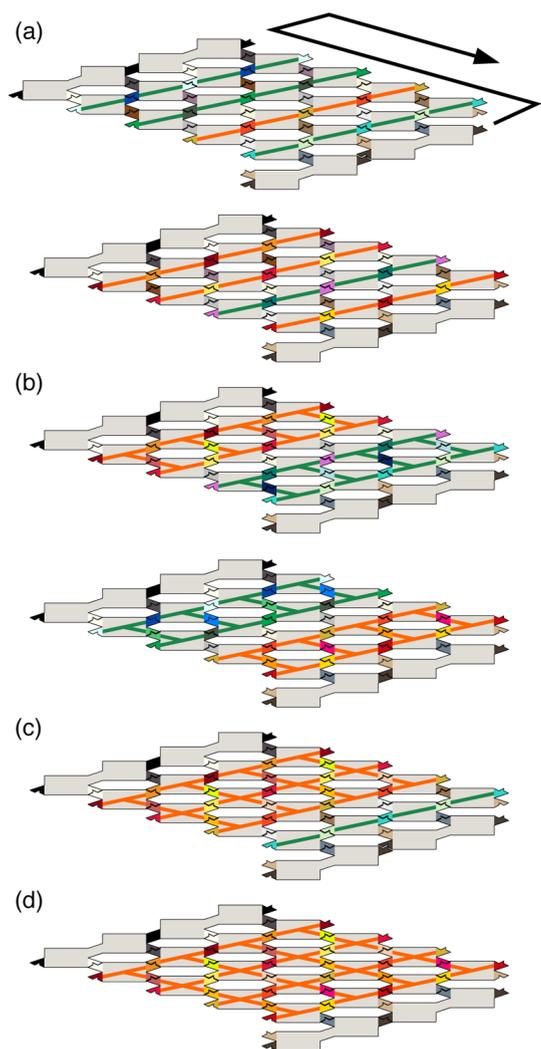


Figure 2. Tile sets that copy information with different levels of redundancy. The 4 tile sets used in experiments and how they fit together. Yellow and green lines (representing “0” and “1”, respectively) connect tiles that copy the same bit. (a) The *1-redundant* tile set (shown assembled into two ribbons) copies a 4-bit sequence with no redundancy. The arrow shows the preferred zig-zag growth path on one side of the ribbon. (Growth can also occur in a corresponding zig-zag path on the other side of the ribbon.) (b) The *2-redundant* tile set copies a 2-bit sequence with both bits encoded 2-redundantly. (c) The *3-redundant* tile set copies a 2-bit sequence with one bit encoded 3-redundantly, the other encoded 1-redundantly. Tiles for a complementary ribbon (“1–0” rather than “0–1”) are shown in the Supporting Information. (d) The *4-redundant* tile set copies a 1-bit sequence encoded 4-redundantly. Tiles for a complementary ribbon (“1” rather than “0”) are not shown. Each tile set can copy all possible binary sequences of its length. DNA sequences are given in the Supporting Information.

transient unfavorable tile attachments occur often, and if a second tile attaches to an unfavorably attached tile before it falls off, the mismatched tile can be “locked in,” resulting in an assembly mismatch error (Figure 3a).

In a proofreading tile set, each bit of information is represented by a block of tiles that redundantly encode the same input and output information as the

sticky ends of the single tile on the perimeter of the block. Binding domains on the interior of the block determine how the tiles within a block fit together and are unique to that block. Proofreading tile sets prevent assembly errors by interrupting the lock-in process. When an incorrect tile attaches, further mismatches must occur within the tile's block in order for the incorrect tile to become fixed in place (Figure 3b). Because sequences of multiple, co-localized errors are relatively rare, assembly stalls after an error, allowing the incorrect tile to fall off and correct assembly to proceed.³⁶ The block size determines the number of mismatches that must occur before lock-in. Each additional required mismatch should, in theory, exponentially reduce the error rate.³³ The proofreading mechanism can be applied to any algorithmic self-assembly process: a tile set without proofreading can be transformed into a proofreading tile set by replacing each tile by a block of tiles. The pattern assembled by the original tile set is then assembled more robustly by the corresponding proofreading tile set, except that the pattern is now larger.

There is already some experimental evidence that proofreading tile sets work. The first experimental implementation of a tile set that copied binary bit sequences from layer to layer²⁶ did not use proofreading and exhibited an error rate of at least 2% per bit copy event. Later implementations of bit-copying tile sets^{27,30} used 2-redundant proofreading tile sets and achieved error rates of 0.26% and 0.02% per bit, respectively. However, it is difficult to draw firm conclusions about the effect of proofreading on assembly error rates from these studies because each study used different methods to nucleate crystal growth and different growth conditions.

Here we experimentally test the prediction that each layer of additional redundancy in a proofreading tile set can exponentially reduce the error rate during algorithmic self-assembly in a controlled setting. We find that there is a reduction in error rate with increasing redundancy but the amount of this reduction is smaller than predicted by quantitative simulations. This difference may be the result of either nonidealities in our experiments or may suggest a need to further refine *in silico* models of algorithmic self-assembly.

RESULTS AND DISCUSSION

Experimental Design. To understand how redundant encoding of a molecular sequence affects the error rate as it is copied during algorithmic self-assembly, we chose to use a well-characterized DNA tile-based structure, the zig-zag ribbon.³⁴ The set of tiles that make up a zig-zag ribbon crystal are designed to control both the ribbon's structure and its growth pathway. Ribbons have a fixed number of tile rows; growth extends the number of ribbon columns. The ribbons we consider here have six rows, with the tiles in

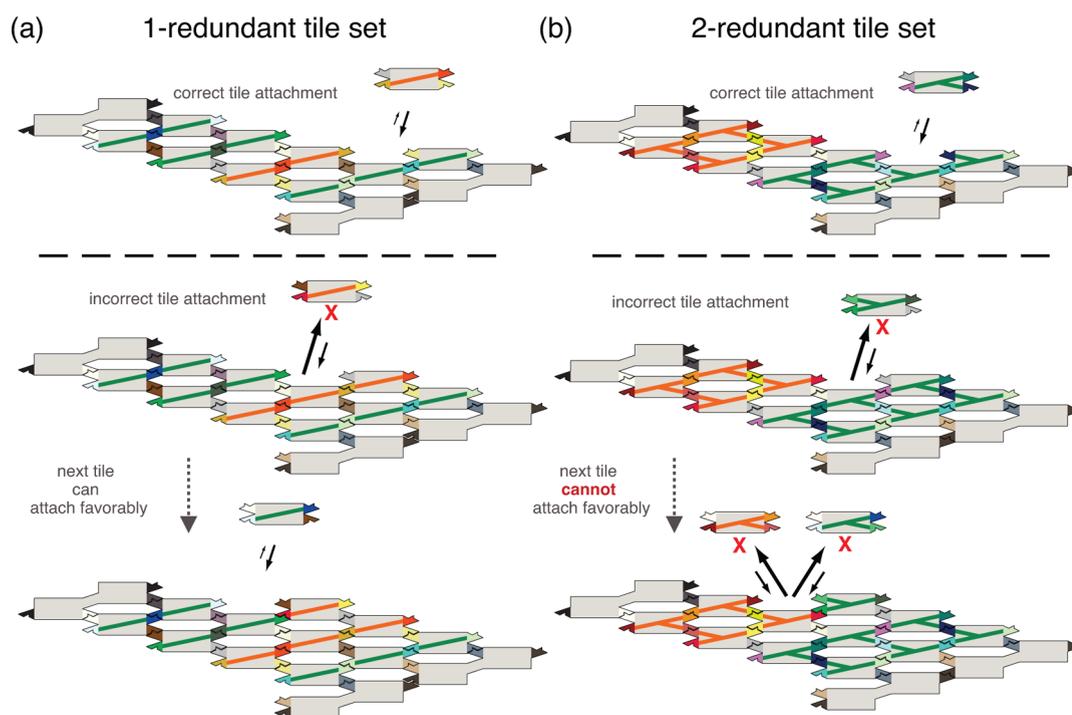


Figure 3. Proofreading uses redundancy to prevent errors during self-assembly. (a) Tile set growth without proofreading. Tiles can favorably attach to a crystal only if they form two new bonds (top). Tiles that form fewer than two bonds with the ribbon (mismatched sticky ends do not count as a bond) attach unfavorably and generally fall off quickly (center). However, if a second tile attaches before such a tile falls off (bottom), the first tile is locked in, potentially producing an error. (b) Proofreading tile set growth. Proofreading reduces the error rate by interrupting the lock-in process in (a). When a mismatched tile joins the crystal (middle), no tile matches more than sticky end at the new growth site. At least one additional unfavorable attachment must therefore occur before the mismatched tile can be locked in (bottom).

the middle four rows encoding a bit sequence that is copied during assembly. The top left and bottom right domains of these tiles are unique to their rows—they encode which row the tile fits into. The top right and bottom left domains encode either a “0” or “1” bit that is copied during assembly.

Under slightly supersaturated conditions, *i.e.*, just below the melting temperature, the thermodynamically favorable growth sites on a crystal are sites where tiles can attach to a ribbon by at least two binding domains. Zig-zag ribbons are designed such that there is exactly one growth site on each end of a crystal. A tile attachment at a growth site creates a new growth site at either the row above or below the newly attached tile. The designed zig-zag ribbon growth pathway zigs up the ribbon rows, then zags back down. A set of *double tiles* fit in the top and bottom rows. These tiles end a series of tile attachments in one direction (from bottom to top or top to bottom) and create a new binding site to initiate a new series of tile attachments in the opposite direction (Figure 2a). The use of two alternating tile types in each of the middle rows enforces the staggered placement of double tiles in the top and bottom rows, ensuring that zig-zag growth can continue. Information transfer occurs during ribbon growth because tiles must attach to a ribbon both by a binding site that encodes the correct row information *and* a second binding site that matches the logical

(“0” or “1”) value of the tile in the previous column (Figures 2 and 3, top).

To measure how the rates of assembly error decrease as the amount of redundancy increases, we designed four distinct zig-zag tile sets that employ either no redundancy (the 1-redundant tile set) or increasing amounts of redundancy (the 2-, 3- and 4-redundant tile sets) (Figure 2). Because each of the zig-zag ribbons has the same width, the redundant zig-zag tile sets copy sequences containing fewer distinct bits than the nonredundant tile set.

For any level of redundancy, the highest fidelity of copying is expected at low supersaturation, where assembly is reversible and where only attachment of tiles by two matching bonds is energetically favorable.²⁹ While ribbons can grow under these conditions, they do not nucleate readily. To reduce the barrier to ribbon nucleation at low supersaturation, we designed a ribbon “seed” structure, a self-assembled DNA structure with a facet analogous to a ribbon facet that is stable above the ribbons’ melting temperature (Figure 4a, see Supporting Information for design). The sequences of the binding sites on a seed’s facet determine the sequence copied by the ribbon that grows from it. The seed is designed so that no unfavorable tile attachments are required for ribbon growth: tiles can first assemble a cone of tiles in the center of the facet *via* exclusively favorable attachments.

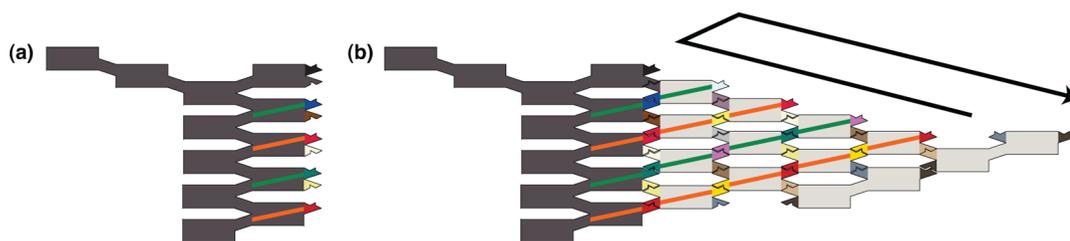


Figure 4. Nucleation of a ribbon from a crystal seed. (a) A composite illustration of a crystal seed. In this example, the binding sites propagate the illustrated green-yellow-green-yellow sequence; these binding sites can be modified by design to change the seeded sequence. (b) Initial error-free growth off the seed produces a V-shaped assembly of tiles. While assembly order is nondeterministic, all possible assembly orders that do not involve unfavorable assembly steps correctly copy the sequence. After enough layers of tiles have accumulated, a double tile can attach by two binding domains to the bottom edge. The attachment of a second adjacent double tile allows zig-zag growth (black arrow; also cf. Figure 2a) to commence.

Double tiles can then bind favorably to the assembly and initiate zig-zag growth (Figure 4b).

Simulation. We used simulations of tile assembly to determine how an established biophysical model of algorithmic self-assembly, the kinetic Tile Assembly Model (kTAM),²⁹ predicts that the level of redundancy should affect the error rate. The kTAM uses a single attachment rate constant, k_f , for all tiles, and a ΔG° of tile binding parametrized by the ΔH_{se}° and ΔS_{se}° for a single sticky-end bond. k_f was set to 10^6 /M/s, within a factor of 3 of the measured rate of oligonucleotide binding,³⁷ the measured attachment rate of DAO-O tiles (which are similar to the DAO-E molecules used here) during nanotube growth,³⁸ and the measured rate for dimerization of DAE-E tiles.³⁹ The hybridization rate presumably has some dependence on context (e.g., sticky-end sequence and orientation of the attaching tile) but for simplicity, and because explicit measurements in a similar DX tile system failed to find a statistically significant difference,⁴⁰ we neglect any such effects here. ΔH_{se}° and ΔS_{se}° were taken to be half the measured values of ΔH° and ΔS° for the attachment of a zig-zag ribbon tile to a ribbon by two sticky ends.³⁴ We thus assumed ideal cooperativity of tile binding sites. Further, we ignored any sequence dependence effects and any potential contribution of mismatched sticky end binding.⁴¹ Stochastic kinetics⁴² were used to simulate each tile binding and unbinding event. Attachment or detachment of blocks of tiles to or from each other, while possible in solution, were not modeled. The simulation is not intended to be quantitatively accurate, but rather is meant to examine tile assembly semiquantitatively, as the general characterization of phenomena relevant to proof-reading is likely to be robust over a wide range of parameter values.

The kTAM model is governed by two parameters that represent physical conditions: the **sticky end** free energy, G_{se} , which measures the interaction strength between two tiles joined by a single sticky end, and the **monomer concentration** free energy, G_{mc} , which measures the amount of entropy lost when a tile joins an assembly. Specifically, the rate of tile detachment

by b sticky ends is $r_{r,b} = k_f e^{b(\Delta H_{se}^\circ - T\Delta S_{se}^\circ)/RT - \Delta S_{init}/R} \stackrel{def}{=} \hat{k}_f e^{-bG_{se}}$ where k_f is the forward rate constant, $\Delta S_{init} = -6$ cal/M/K is the initiation entropy,⁴³ $\hat{k}_f = k_f e^{-\Delta S_{init}/R}$, T is absolute temperature, and $R \approx 2$ cal/M/K is the universal gas constant. The rate of tile attachment at a given site on a crystal is $r_f = k_f [t] \stackrel{def}{=} \hat{k}_f e^{-G_{mc}}$, where $[t]$ is the free tile concentration and $G_{mc} = -\ln[t] - \Delta S_{init}/R$. Thus, the dimensionless free energy for tile attachment by b binding domains is given by $\Delta G = \ln(r_{r,b}/r_f) = G_{mc} - bG_{se}$.

When $2G_{se} > G_{mc} > G_{se}$, tiles attach favorably only when they match at least two sticky ends on a growing ribbon. For a given G_{mc} , assembly occurs most accurately just below the melting temperature, when $G_{mc} \approx 2G_{se}$. For a fixed concentration, assembly becomes less accurate as G_{se} (and thus supersaturation) increases. The ratio $\tau \stackrel{def}{=} (G_{mc}/G_{se})$ is conventionally used as a measure of supersaturation for algorithmic self-assembly reactions.^{29,32,33}

This model is simple enough that it is possible to analytically estimate the assembly error rate for a given G_{mc} and G_{se} . In a 1-redundant tile set, a single mismatched tile can be locked in by the next tile, which can attach by two matching binding domains. The probability that a mismatched tile is locked in before it can fall off (as in Figure 3a) can be approximated as

$$\frac{r_f}{r_f + r_{r,1}} = \frac{e^{-G_{mc}}}{e^{-G_{mc}} + e^{-G_{se}}} \quad (1)$$

To obtain the rate at which mismatched tiles are locked in, we must multiply this probability by the rate at which mismatched tiles attach. That rate is the number of tile types that can bind by a single bond at this location (here, just 1) times the rate at which each such tile attaches (here, $k_f [t]$). The bitwise error rate for copying is then approximated as the ratio of this mismatch lock-in rate to the rate of correct growth (here, also $k_f [t]$), yielding again the term in eq 1. In this simple analysis, we ignore the probability that the tile that attached by two binding domains will subsequently fall off.

With a tile set that has $n > 1$ levels of redundancy, an incorrect attachment must be followed by another

incorrect attachment which also tends to fall off (Figure 3b). For such a tile set, the probability that an initial mismatched attachment will be locked in now depends on a sequence of n unfavorable single-bond attachments occurring before a favorable two-bond attachment can finally lock in the first error. Since the rate of attachment r_f is not dependent on the number of bonds formed, this probability is, roughly,

$$\left(\frac{e^{-G_{mc}}}{e^{-G_{mc}} + e^{-G_{se}}} \right)^n \quad (2)$$

This expression is approximate, as it neglects the probability that tiles will fall off and reattach during the lock-in process, or that, as above, the tile that attaches by two matching binding domains will fall off. But when τ is just below 2, eq 2 implies that each layer of redundancy is expected to reduce the error rate by the multiplicative factor

$$\frac{e^{-\tau G_{se}}}{e^{-\tau G_{se}} + e^{-G_{se}}} \approx e^{G_{se}} \quad (3)$$

We simulated the error rate during seeded ribbon growth for $G_{mc} = 13$, which corresponds to a tile concentration of about $46 \mu\text{M}$. While this concentration is much higher than the tile concentrations used in our experiments, we found that at lower simulated tile concentrations, accurately measuring the simulated error rates was computationally intractable. For near-optimal values of τ , the simulations predict a decrease in error rate by a multiplicative factor with each additional layer of redundancy, and for lower values of τ (*i.e.*, higher levels of supersaturation), the error rate increases markedly, which is qualitatively, though not quantitatively, consistent with eq 2 (see Figure 5a). For the actual experimental conditions, where tile concentrations are lower, simulations would predict even smaller error rates, and commensurate decreases in error rate with redundancy—but we would expect the qualitative trends to be the same.

Because the error rate depends strongly on the level of supersaturation, we anticipate a significant difference in error rates within crystals nucleated by a seed (which should grow at low supersaturation) and those that spontaneously nucleate (presumably at higher levels of supersaturation). We thus used a second set of simulations to estimate the degree of supersaturation at the point during the anneal where growth from seeds or spontaneous nucleation of new crystals would commence. In these simulations, it was computationally tractable to use a value of G_{mc} set so that $[t] = 50 \text{ nM}$, the same tile concentration used in experiments. In these simulations, the temperature was reduced at the same rate as in our experiments, because the degree of supersaturation experienced by growing crystals depends on the relative rates of nucleation, growth, and the annealing speed. For each

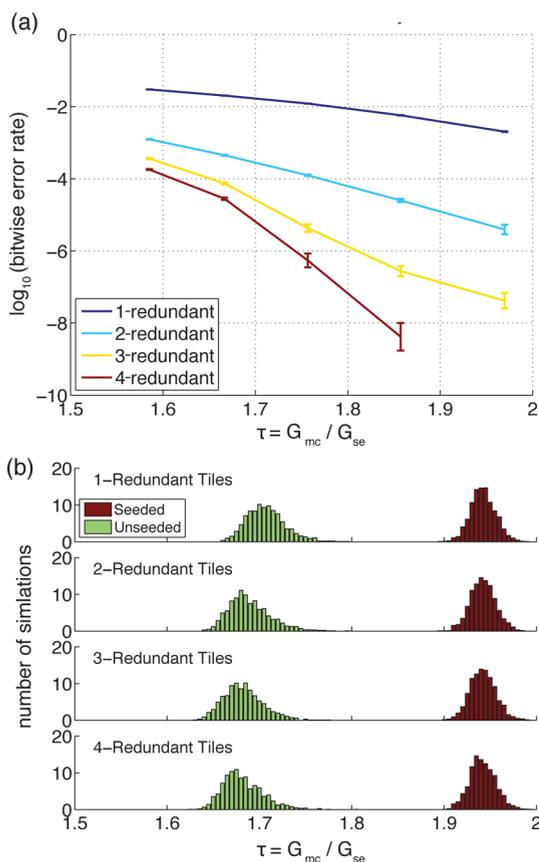


Figure 5. Simulated assembly error rates and nucleation temperatures. (a) Simulated error rates during constant temperature ribbon assembly from a seed, as a function of τ , a measure of supersaturation. (See text; $\tau = 2$ is the melting temperature of the ribbons.) Error rate measurements were made over the growth of a 500 column ribbon. To maintain a constant level of supersaturation, the free tile concentration was held constant at $46 \mu\text{M}$ over the course of the simulation. Error bars indicate 2 standard deviations. (b) Histogram of τ values at which crystals reach 80 tiles during kinetic simulations of annealing ribbon tiles. In the simulations, a reaction volume containing 1 seed, sized such that the seed concentration was 2 nM with 50 nM of each tile was annealed from 40 to 20 °C with the temperature decreasing 1 °C/h, the same protocol used in experiments. Free tile concentration was held constant over the course of the simulation. Each histogram comprises at least 100 simulations.

temperature, the value of G_{se} was set using ΔH_{se}° and ΔS_{se}° . A crystal was considered nucleated when it consisted of at least 80 tiles. These simulations predict (Figure 5b) that seeded crystals grow and quickly exceed the target size as soon as growth of any kind is favorable (and even slightly before, due to fluctuations). In contrast, spontaneously nucleated crystals do not form until the solution becomes fairly supersaturated and predicted error rates are much higher.

Tile Design. The zig-zag ribbons were assembled from DAO-E double crossover molecules (Figure 1a) with 5 base-pair sticky ends. The binding domains encoded by the sticky ends for the four tile sets are shown in Figure 2. The sequences of the double-stranded regions of the tiles were designed as reported

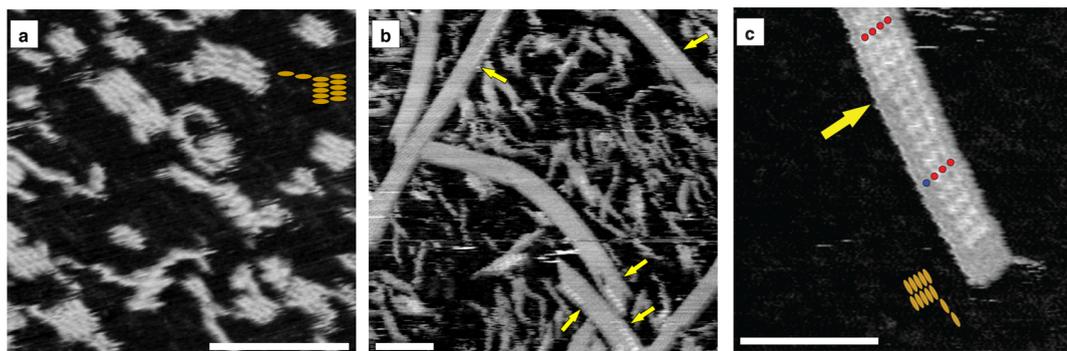


Figure 6. AFM images of seeds and ribbons. All scale bars are 100 nm. The brown illustration shows the expected seed structure. Yellow arrows indicate locations of assembly errors. (a) Crystal seeds. Seeds formed with less than 8% synthesis yield, but with a synthesis quality of roughly 40%. (See text and Supporting Information Figure S1 for details.) (b) The 1-redundant ribbons grown without seeds. The image contains both fully formed ribbons and smaller fragments in the background. The “1” tiles appear brighter because they protrude higher off the surface due to attached hairpin structures. (c) A seed-nucleated 3-redundant ribbon. Blue and red dots label the “1” and “0” elements of a sequence, respectively.

previously³⁴ using sequence symmetry minimization^{44,45} to prevent spurious interactions. By design, matching sticky ends have similar hybridization energies but all sticky ends encoding “0” and “1” in the same row have a cross-hybridization ΔG° of 0 (or weaker) according to the nearest neighbor model of DNA hybridization.⁴⁶ To differentiate “0” from “1” tiles during atomic force microscopy (AFM) imaging, two hairpins were added to the middle of one of the helices of the “1” tiles perpendicular to the plane of the lattice, which increased the height of the “1” tiles.¹⁷

The crystal seed structure is a DAO-E DNA crystal lattice¹⁷ woven such that no strand can detach from the structure without breaking at least 16 base pairs (see Supporting Information for design). The melting temperature of a similar seed structure at the seed concentration used during growth (2 nM), has been measured to be 62 °C, well above the melting temperature of the ribbons (about 34 °C at 50 nM).³⁴ To template the growth of a particular sequence, the binding domains along the edge of the seed were set to match those of the desired ribbon facet. Previous experimental work suggests that seeds lower, but do not eliminate, the nucleation barrier to the desired sequence.³⁴

Tile and Seed Assembly. To test that the tiles formed properly, we assembled individual tiles by annealing their component strands from 90 to 20 °C at 1 °C/min. (Buffer solution and electrophoresis and AFM protocols are as described in ref.34 and in the Methods section.) Polyacrylamide gel electrophoresis showed that each tile assembled into a single product with at least 80% yield. We think that this is an underestimate of the yield, but in any case, algorithmic crystal growth can be quite robust to impurities of malformed tiles.⁴⁷ AFM images of crystal seeds annealed at the same rate show that seeds form as designed, although the synthesis yield (the fraction of DNA material that was incorporated into well-formed seeds) was estimated to be less

than roughly 8% (Figure 6a). This synthesis yield is lower than that of other potential seed structures such as DNA origami,^{27,48} and suggests that a variety of partial seed structures could lower the barrier to nucleation for a wider variety of sequences during the annealing process. However, for our purposes the exact concentration of seeds (as measured by the synthesis yield) is not critical; our control over the sequence being nucleated depends rather on the fraction of well-formed seeds out of all DNA structures likely to stimulate ribbon nucleation, which would include only the large partial seed structures. This fraction, which we call the synthesis quality, was estimated to be roughly 40% and deemed to be adequate. (See Supporting Information Figure S1 for details.)

To test that 6-tile-wide ribbons formed and could propagate arbitrary sequences, we first annealed 50 nM of each 1-redundant tile strand from 90 to 40 °C at 1 °C/min and from 40 to 20 °C at 1 °C/h in the absence of seeds. Under these conditions, tiles are expected to form during the first part of the anneal and ribbon growth becomes favorable slightly below 40 °C.^{34,49} As the temperature decreases, spontaneous nucleation of ribbons containing arbitrary sequences becomes more likely. AFM imaging revealed that each ribbon copied a pattern with occasional errors (Figure 6b), and that each of the 10 distinguishable patterns formed. (The orientation of asymmetric patterns such as 1000 can not be determined.) In analogous tests of the 2-, 3- and 4-redundant tile sets, all 2-, 3- and 4-redundant patterns were also observed, indicating that all tiles correctly performed their logical function.

Ideally, assemblies with all possible patterns would nucleate and grow at the same rate. However, in previous experiments, segregation of “0” and “1” tiles into mostly “0” and mostly “1” regions has been observed,⁵⁰ possibly caused by a slight difference in the melting temperature of “0” and “1” tile lattices.

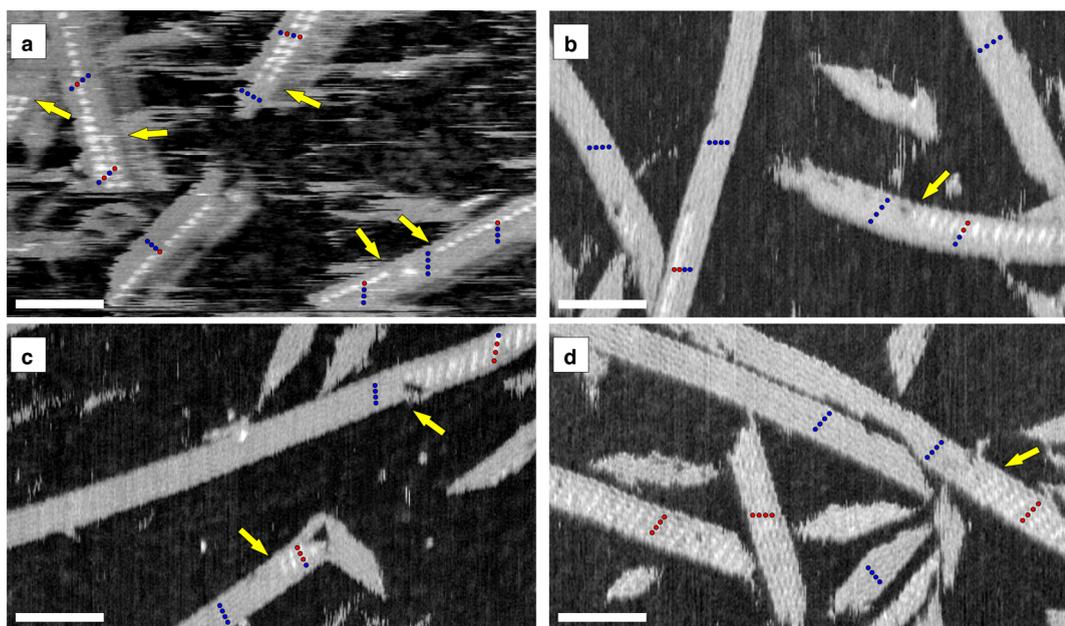


Figure 7. Ribbons copying sequences. All scale bars are 100 nm; yellow arrows point out locations where sequences are incorrectly copied. Blue and red dots label the “1” and “0” elements of the sequence, respectively. (a–d) 1-, 2-, 3-, and 4-redundant ribbons, grown from seeds as described in the text.

This difference may be due to the hairpins on the “1” tiles, which may alter the lattice structure by introducing twist or distortion that could weaken binding rates.^{25,27,49}

When the two tile types are annealed together, this difference could cause ribbons with all “0” tiles to form first, producing an excess of these ribbons, and preventing the “0” tiles from being used to form ribbons encoding other patterns. The over-representation of ribbons with mostly 0s would also cause “0” tiles to be used up faster than “1” tiles during growth. The resulting imbalance in concentrations of “0” and “1” tiles means that the ON rates for these two tile types would be different and that as a result, the mismatch error rates for different sequences would be different. The degree of these differences would also change over the course of the reaction as the two tile types are depleted. Thus, segregation of tile types into ribbons of mostly “0” or “1” tiles makes it hard to quantitatively measure how redundancy, rather the sequence being copied or when during the reaction the copying process takes place, affects error rates.

To measure sequence-dependent nucleation of unseeded ribbons, we counted the number of times each 1-redundant pattern was observed in 25 nonoverlapping 1 μm square images taken at random locations on the surface. Consistent with our expectations, ribbons containing all or almost all 0s or 1s were more common than expected by random chance, while those with a mixture of 0s or 1s were comparatively rare (Figure 8, green bars).

If most crystals grow from seeds that nucleate 0101 and 1010 patterns, and if 0101 and 1010 ribbons grow

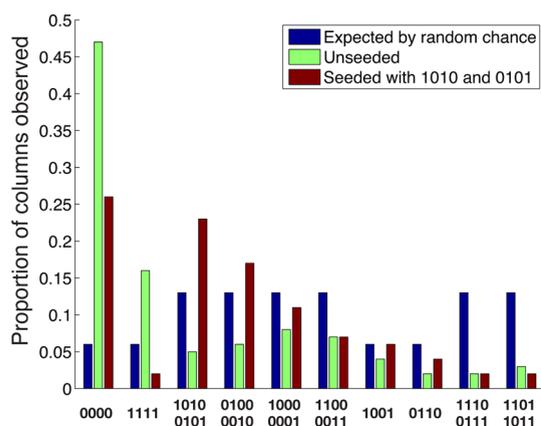


Figure 8. Proportions of observed 1-redundant patterns. Data are based on 25 seeded and 25 unseeded images. Each set of images contained, in total, approximately 100 ribbons.

at the same rate, then “0” and “1” tiles would be expected to be depleted at the same rate, so that concentrations of the two tile types would remain balanced. We therefore repeated the growth of 1-redundant ribbons using seeds for the 0101 and 1010 patterns with the goal of achieving this balance. We added 2 nM each of preformed crystal seeds for the 0101 and 1010 patterns when the anneal reached 50 $^{\circ}\text{C}$, which is above the melting temperature of ribbons but below the melting temperature of the seeds. (See Figure 6c for an AFM image of a seeded 3-redundant ribbon.) More than 20% of the resulting ribbons displayed the seeded pattern, 10 times more than without seeds, but the 0000 pattern still occurred most frequently (Figure 8, brown bars). From this we

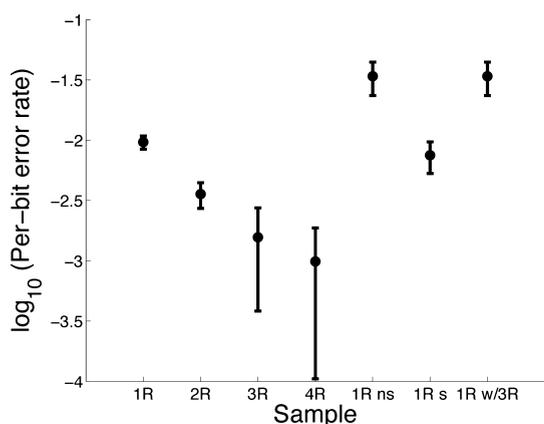


Figure 9. Per-bit error rates during ribbon growth. 1R = 1-redundant, 2R = 2-redundant, 3R = 3-redundant, 4R = 4-redundant. The graphs include all crystals grown with seeds, although since not all crystals grew from seeds, these quantities include both seeded and unseeded growth. The last three bars show error rates of 1-redundant growth under different conditions. “1R ns” is the error rate within 1-redundant crystals grown without seeds. “1R s” is the error rate of 1-redundant crystals grown with seeds that are seeded with very high probability because they begin with the seeded sequence, 0101 or 1010, that otherwise rarely nucleates. “1R w/3R” is the measured error rate for the fourth row of the 3-redundant ribbons, which copied 1 bit nonredundantly. For each level of redundancy, the number of errors are tabulated using 25 AFM images of ribbons $1\ \mu\text{m}$ on each side, about 5000–10 000 ribbon rows.

conclude that ribbons nucleated and grew at a variety of temperatures during the anneal, the seeded pattern largely at lower supersaturation and the unseed pattern largely at higher supersaturations, making it difficult to precisely compare our simulation results with our experiments. However, the use of 0101 and 1010 seeds did reduce the problem of uneven depletion of “0” and “1” tiles and made it possible to study the propagation of a wider variety of sequences.

Error Rates. To determine the rate at which assembly errors occurred in the 1-redundant tile set, we counted the number of bits that were correctly and incorrectly copied by ribbons annealed with seeds (Methods, Figure 9, bar 1). Crystals with the seeded pattern should nucleate first, and therefore should grow under less supersaturated conditions, where the error rate is lower, than other crystals. Given the average level of supersaturation when a ribbon starts to grow that we observed in our simulations, we would expect that the error rate for seeded crystals would be approximately 10-fold lower than the rate for unseeded crystals. In our experiments, the error rate difference between unseeded crystals and mostly seeded crystals was approximately 4.5, consistent with this prediction (Figure 9, points 5–6).

We characterized the copying error rates of the 2-, 3-, and 4-redundant tile sets using the same protocol. The crystal seeds nucleated the patterns 01 and 10 for the 2- and 3-redundant tile sets, and the patterns 0 and 1 for the 4-redundant tile set (Figure 7b–d). With each

increase in the level of redundancy, the error rate decreased by a factor of about 3, with the error rate for the 4-redundant tile set being less than 0.1% per bit copied (Figure 9, bars 1–4). These results qualitatively match our predictions, although the observed decrease in error rates with each additional degree of redundancy was much smaller than the 100-fold drop in error rates predicted by our simulations. This difference is even more striking given that these simulations overestimated the expected error rates, since they considered growth at a higher tile concentration than the concentration used in experiments. We also observed, surprisingly, that the error rate in the fourth row of the 3-redundant ribbons (which was encoded 1-redundantly) was about three times higher than the error rate in copying the bits in the 1-redundant tile set.

DISCUSSION

In our experiments, the observed error rates while copying information in zig-zag ribbon crystals qualitatively agreed with the predictions of our approximate theoretical analysis and our simulations: with each level of redundancy, we see an approximately multiplicative reduction in the error rate. The biggest reduction in error rates came with introducing the first level of redundancy, also in qualitative agreement with the simulations.

Together, these results suggest that redundancy can play an important role in ensuring that self-assembly proceeds with few defects, and that it is possible to use kinetic models to design energy landscapes for self-assembly pathways that systematically reduce errors. The approach considered here, *i.e.*, using redundancy, cooperativity, and proofreading to improve the accuracy of a self-assembly process, contrasts with the approach of using strand displacement to make each step of self-assembly robust.^{51,52} While the strand displacement approach is simple and results in low error rates, it places limits on the kinds of structures that can be assembled and so far has proved difficult to scale.

A remaining discrepancy in our work is that simulations and theory predict a much lower absolute error rate than was observed experimentally. Similar discrepancies between simulations and experimental results have been described in previous algorithmic self-assembly studies.^{25,34,49} Possible reasons for these discrepancies include differences in interaction energies between tiles with different sticky end sequences or with and without hairpins, limited cooperativity in the binding of tiles to multiple sticky ends simultaneously, malformed tiles, poor nucleation and nonideal assembly temperatures, or differences between tile concentrations and other tile depletion effects. More detailed simulations have shown that some of these effects can strongly affect the predicted error rate,^{41,47,53} but

systematic experiments to identify which effects are most important in practice have not been performed.

The higher-than-predicted error rates in the experiments reported here can be partially attributed to the imperfect performance of the seeds and the presumed consequence that assembly occurred at a relatively high level of supersaturation. It is therefore possible that we might observe much lower error rates using larger, more sturdy DNA origami seeds^{27,28,30,48} that appear to nucleate assemblies closer to the melting temperature. Another factor that may have affected our results is ribbon joining, which has been observed at high enough rates that some ribbons we observed can be presumed to have been formed by end-to-end joining reactions.^{30,34,54} Joining of ribbons could produce structures with mismatches that we would have counted as errors during AFM image analysis. We therefore expect that our measured error rates may overestimate the true rate of errors during growth by monomer addition. Reliable nucleation of ribbons from a seed would also reduce end-to-end joining and therefore allow a more accurate measurement of ribbon growth rates.

Despite the limitations to our experiments and modeling, the continued reduction in error rates with

increasing redundancy also suggests that it should be possible, without any improvement in experimental technique, to assemble ribbons with virtually no errors simply by using enough redundancy and by using similar techniques that prevent spurious DNA tile nucleation^{34,55} and facet growth.^{33,49} While large amounts of redundancy increase the feature size of the product, it may be possible to mitigate this increase by using either tile sets which introduce redundancy without increasing size,^{56,57} or by using smaller tiles.^{52,58}

Consider, for example, the 16-tile wide “Sierpinski ribbon” of Fujibayashi *et al.*,²⁸ which used a 1-redundant tile set and obtained a 1.4% error rate, comparable to the 1-redundant tile set in this work. The Sierpinski pattern can alternatively be produced using a 4-redundant tile set without increasing the pattern scale.^{57,59} If a similar reduction in error rates similar were to be observed, each 4×4 proofreading block would have a 0.14% chance of error, implying that the expected ribbon length prior to the first error would be more than 700 tiles or roughly $9 \mu\text{m}$. While this prediction remains to be tested experimentally, it illustrates the potential for the self-assembly of complex micrometer-scale molecular structures.

METHODS

Crystal Growth. For each experiment, strands were mixed at 50 nM per strand in TAE buffer with added Mg^{2+} (40 mM Tris-acetate, 1 mM EDTA, 12.5 mM magnesium acetate). Samples grown without seeds were annealed from 90 to 40 °C at 0.1 °C/6 s, during which time it was expected that tiles assembled from individual strands, and then from 40 to 20 °C at 0.1 °C/6 min, during which time crystals putatively nucleated and grew. For samples containing a crystal seed, the seed strands were first annealed separately at 50 nM per strand from 90 to 20 °C at 0.1 °C/6 s. Tile samples (without seeds) were then annealed from 90 to 50 °C, and held at 50 °C for 1 h, during which time the previously annealed seeds were added to a final concentration of 2 nM. After holding at 50 °C for 1 h, the samples were cooled to 40 °C at 0.1 °C/6 s and then from 40 to 20 °C at 0.1 °C/6 min. Samples were imaged by AFM soon after an anneal was completed.

Sample Imaging. AFM imaging was performed using a Digital Instruments Multimode AFM (with a Nanoscope IIIa controller) in tapping mode under the TAE/ Mg^{2+} buffer described above. After annealing, 2 μL of each sample was deposited on a freshly cleaved mica sheet with an added 40 μL layer of TAE/ Mg^{2+} buffer. To avoid sampling bias, images taken for statistical measurement of error rate were taken at random locations in a self-avoiding random walk along the surface.

Yield Determination for Crystal Seeds. Synthesis yield is here defined as the fraction of DNA material that is incorporated into well-formed seeds. We estimated the synthesis yield from the AFM image of Supporting Information Figure S1 as follows. The total DNA material was measured in pixels occupied by applying a height threshold to the image. The material incorporated into well-formed seeds was measured, similarly in pixels, by identifying structures that were five tiles high. The ratio, which was roughly 8%, is an estimated upper bound on the synthesis yield. We expect that it is a loose upper bound, because in our experience smaller DNA structures do not adhere to the mica as well as larger DNA structures, resulting in a bias for larger structures. It is difficult to assess the extent of

the bias because it appears to vary with mica samples and is sensitive to the Na^+ , Mg^{2+} , and Ni^{2+} salt concentrations.

Synthesis quality is here defined as the fraction of well-formed seeds among all full-sized (5 tiles in height) and reasonably large partially assembled (3 or 4 tiles in height) structures. Identification of said structures is also shown in Supporting Information Figure S1, where the synthesis quality is estimated to be roughly 40%. We believe this estimate to be fairly robust to the mica stickiness bias as well as to alternative interpretations of the image.

Error Rate Determination for Ribbon Growth. Approximately 50–75 images for each level of redundancy were used to determine error rates in assembly. In each image, the total number of mismatch errors, along with the total length of observed ribbons were tabulated. An error during assembly copying was defined as a change in a single bit from 0 to 1 or *vice versa* such that each value is copied for at least 3 columns—since changes to a redundantly encoded bit might involve several individual defects spread over several columns, and because imaging artifacts appear locally without changing the propagated pattern, structures unique to an individual column were not counted. Simultaneous changes in multiple bits, as in Figure 7, were counted as multiple errors.

The total number of rows copied was estimated from the total length using previously measured tile sizes of approximately 13 by 6 nm.^{17,34} The error rate was calculated by dividing the total number of errors by the total number of bits copied. The confidence interval is given by 2 standard deviations of this measurement. Ribbon fragments that were not full width or other artifacts were not included in error rate determination, but were often seen in AFM images.

Simulations. Simulations used *xgrow*, which simulates tile assembly using the kTAM model²⁹ and used the parameters described in the main text.

For simulations that measured error rates, growth of ribbons began from a structure with the same arrangement of sticky ends as the seed structure. Each simulation proceeded for 2000 simulated seconds. The maximum size the structure could grow

to was determined by the simulated lattice size, approximately 256 rows. For $\tau \ll 2$, assemblies reached this maximum in the allotted time, but samples with $\tau \approx 2$, assemblies did not. Error rates were calculated by dividing the number of bits copied after the initial nucleation stage (see Figure 4b, main paper) by the number of mismatch errors seen. To avoid counting tiles that may have been added transiently in error but were not locked in, the last 4, 8, 12, and 16 tiles added for 1-, 2- 3- and 4-redundant tile sets, respectively, were removed before tabulating the error rate. The confidence intervals for error rate measurements were two times the standard deviation, given by $\sigma = (\epsilon)^{1/2}/(m^{1/2})$, where m is the total number of bits copied and ϵ is the measured error rate.

To determine the average temperature at nucleation of both seeded and unseeded structures, simulations followed the same linear annealing schedule used for experiments, a 1 °C drop each hour. The strength of a single sticky-end bond, G_{se} , at each simulated temperature was determined using the formula $G_{se} = -(\Delta H_{se}^{\circ} - T\Delta S_{se}^{\circ})/RT$, where $\Delta H_{se}^{\circ} = -51.2$ kcal/mol and $\Delta S_{se}^{\circ} = -150$ cal/M/K are exactly half the previously measured energies for attachment to a ribbon by two sticky end bonds.³⁴ Each sample used $G_{mc} = 19.8$ which corresponds roughly to a tile concentration of 50 nM and boundary tile concentration of 100 nM. Simulations that measured average nucleation temperatures from a seeded structure used the same seeded structure as the simulations that measured error rate. To fairly compare the nucleation time of seeded with unseeded samples, all single tile attachment and detachment reactions occurred within a volume of solution that would be occupied by 1 seed (about 83 femtoliters) were tracked, while keeping free tile concentrations constant. Nucleation in both cases was defined as when an assembly reached 80 tiles.

The `xgrow` program and simulation scripts used in this study are available from the authors by request.

Conflict of Interest: The authors declare no competing financial interest.

Acknowledgment. The authors wish to thank Paul Rothmund, Robert Barish and Ho-Lin Chen for valuable advice and discussion. This work was supported by NSF Grants CCF-0432193, CCF-0523761, and CCF-0832824 (The Molecular Programming Project), NASA Astrobiology NNG06GA50G, FENA Theme 2, a Caltech SURF award and the Miller Institute for Basic Science.

Supporting Information Available: Additional controls, strand design information and all sequences. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/nn507493s.

REFERENCES AND NOTES

- Racine, M. F.; Senger, B.; Saveanu, C.; Fasiolob, F. Ribosome Assembly in Eukaryotes. *Gene* **2003**, *313*, 17–42.
- Davidson, E. H. *Genomic Regulatory Systems*; Academic Press: San Diego, CA, 2001.
- Geiger, B.; Spatz, J. P.; Bershadsky, A. D. Environmental Sensing through Focal Adhesions. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 21–33.
- Affolter, M.; Weijer, C. J. Signaling to Cytoskeletal Dynamics during Chemotaxis. *Dev. Cell* **2005**, *9*, 19–34.
- Fell, D. A. Metabolic Control Analysis: A Survey of Its Theoretical and Experimental Development. *Biochem. J.* **1992**, *286*, 313–330.
- Rothmund, P. W. K. Folding DNA To Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440*, 297–302.
- Pollard, T. D.; Cooper, J. A. Actin and Actin-Binding Proteins: A Critical Evaluation of Mechanisms and Functions. *Annu. Rev. Biochem.* **1986**, *55*, 987–1035.
- Nogales, E. Structural Insights into Microtubule Function. *Annu. Rev. Biochem.* **2000**, *30*, 397–420.
- Rossmann, M. G.; Johnson, J. E. Icosahedral RNA Virus Structure. *Annu. Rev. Biochem.* **1989**, *58*, 533–573.
- Mann, S. Molecular Recognition in Biomineralization. *Nature* **1988**, *332*, 119–124.

- Winfree, E. *On the Computational Power of DNA Annealing and Ligation*; DNA Based Computers: Providence, RI, 1996; pp 199–221.
- Wang, H. *An Unsolvability Problem on Dominoes*; Harvard Computation Laboratory Technical Report BL-30 (II-15), 1962.
- Rothmund, P. W. K.; Winfree, E. *The Program-Size Complexity of Self-Assembled Squares*; Symposium on Theory of Computing (STOC), New York, 2000; pp 459–468.
- Adleman, L. M.; Cheng, Q.; Goel, A.; Huang, M.-D. *Running Time and Program Size for Self-Assembled Squares*; Symposium on Theory of Computing (STOC), New York, 2001; pp 740–748.
- Soloveichik, D.; Winfree, E. Complexity of Self-Assembled Shapes. *SICOMP* **2007**, *36*, 1544–1569.
- Fu, T.-J.; Seeman, N. C. DNA Double-Crossover Molecules. *Biochemistry* **1993**, *32*, 3211–3220.
- Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. Design and Self-Assembly of Two-Dimensional DNA Crystals. *Nature* **1998**, *394*, 539–544.
- LaBean, T. H.; Yan, H.; Kopatsch, J.; Liu, F.; Winfree, E.; Reif, J. H.; Seeman, N. C. Construction, Analysis, Ligation, and Self-Assembly of DNA Triple Crossover Complexes. *J. Am. Chem. Soc.* **2000**, *122*, 1848–1860.
- Mao, C.; Sun, W.; Seeman, N. C. Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy. *J. Am. Chem. Soc.* **1999**, *121*, 5437–5443.
- Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H. DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires. *Science* **2003**, *301*, 1882–1884.
- He, Y.; Chen, Y.; Liu, H.; Ribbe, A. E.; Mao, C. Self-Assembly of Hexagonal DNA Two-Dimensional (2D) Arrays. *J. Am. Chem. Soc.* **2005**, *127*, 12202–12203.
- Chworos, A.; Severcan, I.; Koyfman, A. Y.; Weinkam, P.; Oroudjev, E.; Hansma, H. G.; Jaeger, L. Building Programmable Jigsaw Puzzles with RNA. *Science* **2004**, *306*, 2068–2072.
- Adleman, L. M. Molecular Computation of Solutions to Combinatorial Problems. *Science* **1994**, *266*, 1021–1024.
- Mao, C.; LaBean, T. H.; Reif, J. H.; Seeman, N. C. Logical Computation using Algorithmic Self-Assembly of DNA Triple-Crossover Molecules. *Nature* **2000**, *407*, 493–496.
- Rothmund, P. W. K.; Papadakis, N.; Winfree, E. Algorithmic Self-Assembly of DNA Sierpinski Triangles. *PLoS Biol.* **2004**, *2*, 424–436.
- Barish, R. D.; Rothmund, P. W. K.; Winfree, E. Two Computational Primitives for Algorithmic Self-Assembly: Copying and Counting. *Nano Lett.* **2005**, *5*, 2586–2592.
- Barish, R. D.; Schulman, R.; Rothmund, P. W. K.; Winfree, E. An Information-Bearing Seed for Nucleating Algorithmic Self-Assembly. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6054–6059.
- Fujibayashi, K.; Hariadi, R.; Park, S. H.; Winfree, E.; Murata, S. Toward Reliable Algorithmic Self-Assembly of DNA Tiles: A Fixed-Width Cellular Automaton Pattern. *Nano Lett.* **2008**, *8*, 3554–3560.
- Winfree, E. *Simulations of Computing by Self-Assembly*; Caltech Technical Report CS-TR:1998.22, 1998
- Schulman, R.; Yurke, B.; Winfree, E. Robust Self-Replication of Combinatorial Information by Crystal Growth and Scission. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 6405–6410.
- von Neumann, J. *Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components*; Automata Studies: Princeton, NJ, 1956; pp 43–98.
- Winfree, E.; Bekbolatov, R. *Proofreading Tile Sets: Error-Correction for Algorithmic Self-Assembly*; DNA Computing 9, LNCS 2943; Berlin Heidelberg, 2004; pp 126–144.
- Chen, H.-L.; Goel, A. *Error Free Self-Assembly using Error Prone Tiles*; DNA Computing 10, LNCS 3384; Berlin Heidelberg, 2005; pp 62–75.
- Schulman, R.; Winfree, E. Synthesis of Crystals with a Programmable Kinetic Barrier to Nucleation. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15236–15241.
- Winfree, E.; Yang, X.; Seeman, N. C. *Universal Computation via Self-Assembly of DNA: Some Theory and Experiments*;

- DNA Based Computers II: Providence, RI, 1998; pp 191–213.
36. Park, S.-H.; Rothmund, P. W. K.; Winfree, E. Proofreading for Error Correction in Two-Dimensional DNA Crystals that Copy Information. In preparation.
 37. Wetmur, J. G. DNA Probes: Applications of the Principles of Nucleic Acid Hybridization. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 227–259.
 38. Hariadi, R. F.; Yurke, B.; Winfree, E. Thermodynamics and Kinetics of DNA Nanotube Polymerization from Single-Filament Measurements. *Chem. Sci.* **2015**, *6*, 2252–2267.
 39. Jiang, S.; Yan, H.; Liu, Y. Kinetics of DNA Tile Dimerization. *ACS Nano* **2014**, *8*, 5826–5832.
 40. Evans, C. G.; Hariadi, R. F.; Winfree, E. Direct Atomic Force Microscopy Observation of DNA Tile Crystal Growth at the Single-Molecule Level. *J. Am. Chem. Soc.* **2012**, *134*, 10485–10492.
 41. Evans, C. G.; Winfree, E. *DNA Sticky End Design and Assignment for Robust Algorithmic Self-Assembly*; DNA Computing and Molecular Programming, LNCS 8141, 2013; pp 61–75.
 42. Gillespie, D. T. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *J. Comput. Phys.* **1976**, *22*, 403–434.
 43. SantaLucia, J., Jr.; Allawi, H. T.; Seneviratne, P. A. Improved Nearest-Neighbor Parameters for Predicting DNA Duplex Stability. *Biochemistry* **1996**, *35*, 3555–3562.
 44. Seeman, N. C. *De Novo* Design of Sequences for Nucleic Acid Structural Engineering. *J. Biomol. Struct. Dyn.* **1990**, *8*, 573–581.
 45. Dirks, R. M.; Lin, M.; Winfree, E.; Pierce, N. A. Paradigms for Computational Nucleic Acid Design. *Nucleic Acids Res.* **2004**, *32*, 1392.
 46. SantaLucia, J., Jr. A Unified View of Polymer, Dumbbell, and Oligonucleotide DNA Nearest-Neighbor Thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460–1465.
 47. Meng, Y.; Kashyap, N. The Effect of Malformed Tiles on Tile Assemblies within the Kinetic Tile Assembly Model. *Nat. Comput.* **2011**, *10*, 357–373.
 48. Mohammed, A. M.; Schulman, R. Directing Self-Assembly of DNA Nanotubes Using Programmable Seeds. *Nano Lett.* **2013**, *13*, 4006–4013.
 49. Chen, H.-L.; Schulman, R.; Goel, A.; Winfree, E. Reducing Facet Nucleation during Algorithmic Self-Assembly. *Nano Lett.* **2007**, *7*, 2912–2919.
 50. Rothmund, P. W. K. Theory and Experiments in Algorithmic Self-Assembly, Ph.D. Thesis, University of Southern California, Department of Computer Science, 2001.
 51. Dirks, R. M.; Pierce, N. A. Triggered Amplification by Hybridization Chain Reaction. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15275–15278.
 52. Yin, P.; Choi, H. M. T.; Calvert, C. R.; Pierce, N. A. Programming Biomolecular Self-Assembly Pathways. *Nature* **2008**, *451*, 318–322.
 53. Fujibayashi, K.; Murata, S. Precise Simulation Model for DNA Tile Self-Assembly. *IEEE Trans. Nanotechnol.* **2009**, *8*, 361–368.
 54. Ekani-Nkodo, A.; Kumar, A.; Fygenson, D. K. Joining and Scission in the Self-Assembly of Nanotubes from DNA Tiles. *Phys. Rev. Lett.* **2004**, *93*, 268301.
 55. Schulman, R.; Winfree, E. Programmable Control of Nucleation for Algorithmic Self-Assembly. *SIAM J. Comput.* **2009**, *39*, 1581–1616.
 56. Fujibayashi, K.; Zhang, D. Y.; Winfree, E.; Murata, S. Error Suppression Mechanisms for DNA Tile Self-Assembly and Their Simulation. *Nat. Comput.* **2009**, *8*, 589–612.
 57. Reif, J. H.; Sahu, S.; Yin, P. *Compact Error-Resilient Computational DNA Tiling Assemblies*; DNA Computing 10, LNCS 3384. Berlin Heidelberg, 2005; pp 293–307.
 58. Wei, B.; Dai, M.; Yin, P. Complex Shapes Self-Assembled from Single-Stranded DNA Tiles. *Nature* **2012**, *485*, 623–626.
 59. Soloveichik, D.; Winfree, E. *Complexity of Compact Proofreading for Self-Assembled Patterns*; DNA Computing 11, LNCS 3892, Berlin Heidelberg, 2005; pp 305–324.