CHEMISTRY

Prime factorization via localized tile assembly in a DNA origami framework

Yinan Zhang[1,2]†, Xiaoyao Yin[3]†, Chengjun Cui[1]†, Kun He[3], Fei Wang[1], Jie Chao[4], Tao Li[3], Xiaolei Zuo[5], Ailing Li[3], Lihua Wang[6,7], Na Wang[3], Xiaochen Bo[8]*, Chunhai Fan[1]*

Modern cybersecurity built on public-key cryptosystems like Rivest-Shamir-Adleman is compromised upon finding solutions to the prime factorization. Nevertheless, solving the prime factorization problem, given a large *N*, remains computationally challenging. Here, we design DNA origami frameworks (DOFs) to direct localized assembly of double-crossover (DX) tiles for solving prime factorization with a model consisting of the computing, decision-making, and reporting motifs. The model implementation is based on the sequential assembly of different DX tiles in the DOF cavity that carries overhangs encoding the prime and composite integers. The primes are multiplied and then verified with the composite, and the result is visualized under atomic force microscopy via the presence (success) or absence (failure) of biotin-streptavidin labels on the reporting DX tile. The factorization of semiprimes 6 and 15 is realized with this DOF-based demonstration. Given the potential of massively parallel processing ability of DNA, this strategy opens an avenue to solve complex mathematical puzzles like prime factoring with molecular computing.

INTRODUCTION

Public-key algorithms are fundamental security elements in modern cryptography, forming the basis of multiple internet Standards including Transport Layer Security (*1*), Secure/Multipurpose Internet Mail Extensions (*2*), and GNU Privacy Guard (*3*). In particular, the prime factorization–based Rivest-Shamir-Adleman algorithm (*4*) that provides both secrecy and digital signature is of paramount interest. However, no general approach has been proven to solve the factorization problem efficiently for a large composite number (especially a semiprime), rendering it in class nondeterministic polynomial time that could be impractical to search exact solutions with state-of-the-art electronic computers (*5*). Hence, alternative approaches that rely on emerging technologies [e.g., Shor's algorithm (*6*)] have aroused increasing attention for their potentials in solving prime factorization in polynomial time.

Molecular computing that exploits individual molecules to perform computations holds the promise to revolutionize the computational paradigm with the miniatured size of molecular devices and scalable parallel processing abilities (*7*–*9*). In particular, DNA computing that converts information into customed conformations shows intense computing capability for intrinsically specific base pairing and sequence programmability (*10*–*12*). Over the past decades, DNA computers have been used in the implementation of certain algorithms and dynamic logic circuits based on hybridization reactions (*13*–*16*). Meanwhile, various theoretical models have been proposed that intend to tackle difficult mathematical problems with DNA computing (*17*–*19*). A tile assembly model was given to compute the sum and product of two numbers using $\Theta(1)$ distinct tiles in $\Theta(n)$ steps, providing preliminary thoughts for prime factorization with DNA tiles (*20*).

Inspired by the segmentation of cellular compartments that processes information effectively in vivo (*21*), the idea of spatial localization is introduced to DNA computing in which preorganized, instead of diffusible, components compute in a prescribed manner on DNA origami (*22*, *23*). Assembled from the folding of a long scaffold strand by hundreds of short staple strands, DNA origami confers high programmability on the manipulation of matter with nanometer precision (*24*–*30*). By sterically favoring the expected reactions and hindering the interferences, these synthetic DNA devices require less complicated system design, thus exhibiting distinct advantages in cascading biomolecular reactions to realize algorithms for sophisticated computing models (*31*). Here, we develop a DNA origami framework (DOF)–based strategy to perform prime factorization by the algorithmic self-assembly of DNA tiles localized in a defined cavity of DOF (*32*, *33*). By combining the reported multiplication motif (*20*) with a decision-making motif that compares the product of prime inputs with the original integer and a reporting motif that gives the result, we establish a prime factoring model that is realizable through the assembly of double-crossover (DX) tiles (*34*). As a proof of concept, we demonstrate the factorization of two semiprimes 6 and 15 on DOF, respectively, using atomic force microscopy (AFM) imaging to visualize biotin-streptavidin labels on the reporting DX tile.

[1]School of Chemistry and Chemical Engineering, Frontiers Science Center for Transformative Molecules, and National Center for Translational Medicine, Shanghai Jiao Tong University, Shanghai 200240, China. [2]School of Chemical Science and Engineering, Tongji University, Shanghai 200092, China. [3]State Key Laboratory of Proteomics, National Center of Biomedical Analysis, Beijing 100850, China. [4]Key Laboratory for Organic Electronics & Information Displays (KLOEID), Institute of Advanced Materials (IAM) and School of Materials Science and Engineering, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China. [5]Institute of Molecular Medicine, Shanghai Key Laboratory for Nucleic Acid Chemistry, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China. [6]Bioimaging Center, Shanghai Synchrotron Radiation Facility, Zhangjiang Laboratory, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China. [7]Shanghai Key Laboratory of Green Chemistry and Chemical Processes, School of Chemistry and Molecular Engineering, East China Normal University, 500 Dongchuan Road, Shanghai 200127, China. [8]Institute of Health Service and Transfusion Medicine, Beijing 100850, China.
*Corresponding author. Email: boxc@bmi.ac.cn (X.B.); fanchunhai@sjtu.edu.cn (C.F.)
†These authors contributed equally to this work.

RESULTS

Tile assembly model for semiprime factorization

To factorize an integer, one may compare the product of prime numbers with the composite number or keep dividing the number by prime factors until the quotient equals 1. Given that division may generate decimals that are difficult to be represented by DNA, we herein used the former algorithm to solve prime factorization, which takes two prime numbers as the input, computes their product, and verifies with the original semiprime. Specifically, the factorization model contains three distinct motifs for multiplication, decision-making, and reporting, respectively. The multiplication motif accepts two binary prime inputs and calculates the product by performing left-shift and adding operations (*20*). To put it simply, if $b = \sum_i b_i 2^i$, for $b_i \in \{0,1\}$ that is the ith bit of b, the product of a and b is given by $ab = \sum_i b_i a 2^i$. Multiplying by two in binary is essentially a left-shift of a; hence, calculating the product of a and b equals summing up the products of a and appropriate powers of 2. Then, the decision-making motif compares the outputs of the multiplication with the original semiprime bit by bit and gives "equal" or "unequal" information to the reporting motif. Last, the reporting motif converts the information to a readable signal to demonstrate the result.

Figure 1A illustrates the logic circuits of our model that are composed of a 2-bit binary multiplier (in purple and cyan), a bit-by-bit comparator (in green), and a digital-to-analog converter (in red), respectively (fig. S1). In the 2-bit carry multiplier motif, each unit initially takes a bit a_i from A, a bit b_i from B, a bit s_{in} representing the positional cumulated sum, and a carry bit c_{in}. Then, the purple unit computes $s_{out} + c_{out} = a_i^* b_i + s_{in} + c_{in}$, where s_{out} and c_{out} are the positional cumulated sum and the carry bit passed to the next unit, respectively (table S1). The cyan unit delivers the positional cumulated sum and carry bit in orthogonal directions, working as a "transporter" that aids in computing with the purple unit. From Z_0 to Z_3 are the bits of the result of multiplication, i.e., the s_{out} from upstream multiplier units. The bit-by-bit comparator units of the decision-making motif take three values as the input, i.e., the ith bit of the multiplication result, the ith bit of the original semiprime and the comparison result of the $(i-1)$th comparator unit. Different logic gates are cascaded to implement the algorithm of each comparator unit, which gives an output to represent the running result of comparison so far (table S2). The decision-making motif only judges the factorization as successful when all comparator units at every bit consistently give positive outputs. The digital-to-analog converter reports the result of prime factorization by accepting the cumulated output from the decision-making motif as the input and converting it to a readable signal. Figure 1B presents the abstract tile assembly model to mimic the logic circuits (figs. S2 and S3 and Supplementary Text). Here, we take the factorization of 6 into the product of 2 and 3 as an example, in which different recognition domains are denoted in a special color. The binary primes 10, 011, and composite integer 110 are placed at the corners of the framework, with each digit encoded into a recognition domain that directs the stepwise attachment of screened tiles to the cavity. Individual tiles function as the multiplier (computing), comparator (decision-making), or convert units (reporting) shown in Fig. 1A, respectively. Only three comparator tiles are used here since the number 6 occupies three bits. Representative tiles of different functions are displayed. The inputs from the ongoing multiplication are on the top left and right of computing tiles, and the outputs of the tiles are given at the bottom sides that subsequently act as the inputs of the next tiles. The appropriate bit of $2^i a$ and that of the running sum are given at the center of each computing tile separately from top to bottom to present the multiplication process in real time. Decision-making tiles accept inputs from three sides, the top left indicates the running result of comparison so far, the top right is the output from the multiplication motif, and the bottom left accepts a bit from the original integer for comparison (given at the center of the tile). Their output is placed at the bottom right to give the immediate decision, in which a check or cross mark is used to represent the positive or negative result, respectively. The last decision-making tile presents the cumulated comparison result to the reporting tile that carries a recognizable signal. The success reporting tile (smiling face) indicates the correct solution, whereas a failure reporting tile (crying face) means that the answer is wrong. Other than the input from the decision-making motif, the reporting tile has two other inputs that aid in robust binding. Strong and weak bonds, denoted by large and small rectangles separately, are coengineered in the recognition domains to guarantee that a tile is added only when inputs from all sides are matched.

In DNA implementation of this model, we used the DNA origami and DX tile to represent the framework and abstract tile, respectively (Fig. 1C). Each DX tile carries four sticky ends that correspond to four sides of the abstract tile, and the DOF carries the sticky ends encoded from the recognition domains of the framework to direct the algorithmic assembly of the DX tiles that is localized in the prescribed DOF cavity. To mimic the model in Fig. 1B, a computing DX tile must attach to at least two sticky ends simultaneously to be added, and decision-making or reporting one is added only when three sticky ends are matched. Hence, other than using the 6–nucleotide (nt) overhang (long arrow) for the computing tiles, we used 3-nt overhangs (short arrow) in the decision-making or reporting tile that attaches to the bottom of the cavity for representing the original integer and the comparison information (primarily four characters including "1," "0," check, and cross marks). A three-thymine spacer, denoted by a light purple short segment, is inserted in the overhangs at the right ends of the computing tiles or both ends of the decision-making/reporting tile to accommodate the internal strain of tile assembly to the size of the cavity. The result is given by the presence (success) or absence (failure) of biotin labeling on the reporting tile in the DOF cavity, identified in AFM imaging by introducing streptavidin markers. We demonstrate the assembly process of screened tiles to the cavity to perform factorization of 6 as 2×3 using the model (Fig. 1D). All the tiles involved are given at the bottom. Since 6 has only two prime factors 2 and 3, the computing tiles that calculates 2×2 and 2×3 and associated decision-making and reporting tiles are presented for screening (fig. S4). The screened tiles at each step from (i) to (iv) are shown below the framework. For better presentation of the assembly details and accommodating to DNA binding characteristics, we distinguish the tiles that carry the same information but in complementary recognition domains by denoting them in different colors. The computing tile C1 fills in the top corner at first (i) and presents additional strong bonds that enable C2 and C4 in the second row to attach (ii). Likewise, C2 and C4 provide extra strong bonds to bind the decision-making tile D1, computing tiles C5 and C7 in the third row (iii). Upon addition, D1 presents a weak bond, while C5 and C3 expose strong bonds, enabling the addition of D2 and C6 (iv).

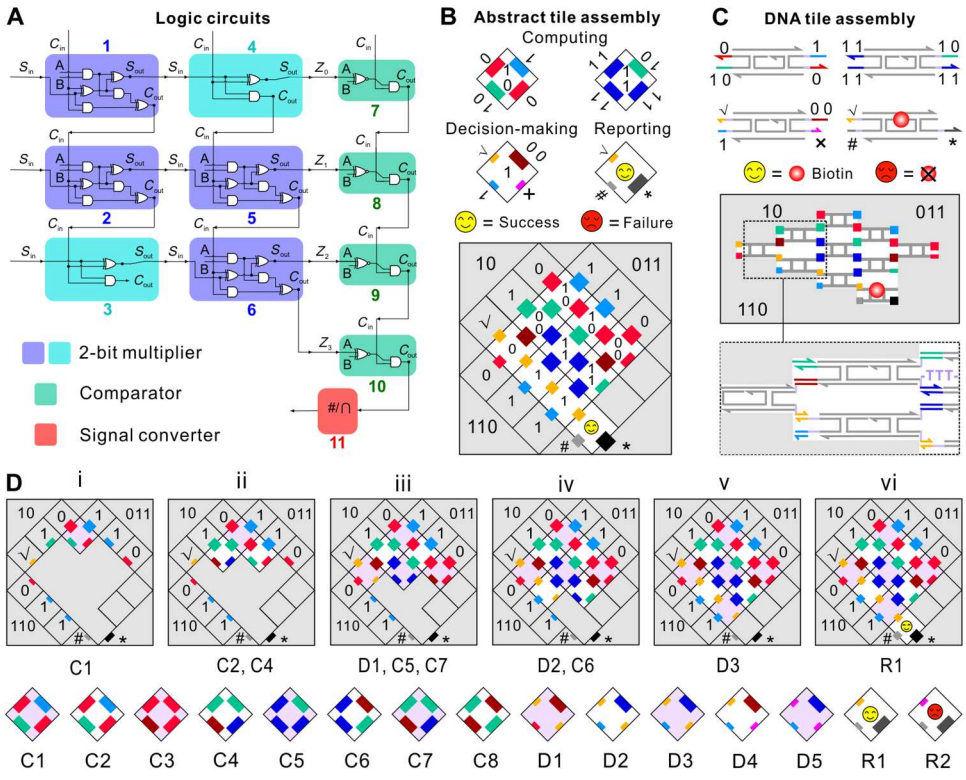

Fig. 1. The abstract model of prime factorization with the localized tile assembly and its DNA representation. (**A**) Logic circuits of this model that are composed of a 2-bit binary multiplier (purple and cyan), a bit-by-bit comparator (green), and a digital-to-analog converter (red). The multiplier computes the product of inputs that are subjected to the comparator to decide if it matches the original integer. The converter gives an indicating signal to report the result. (**B**) Abstract tile assembly model to implement the algorithm. Factoring 6 as the product of 2 and 3 is given herein as an example. Representative tiles of different functions are displayed, with each recognition domain denoted by a special color. The large and small rectangles represent different bond strengths between the recognition domains, respectively. (**C**) DNA mimic of this abstract tile assembly model. Each sticky end of the DNA tile corresponds to a side of the tile model. The positive or negative result is given by the presence or absence of biotin labeling on the reporting tile, respectively. A three-thymine spacer, colored in light purple, is added to the overhangs placed at the right ends of computing tiles or both ends of decision-making/reporting tiles to accommodate the cavity of DNA origami. (**D**) The assembly process of screened tiles in the framework that carries 6 and prime inputs 2 and 3. All the tiles associated are listed at the bottom. The screened tiles at each step from (i) to (vi) are given below the framework. The tiles carrying the same information but in complementary domains are colored differently for better presenting the assembly details and accommodating to DNA binding characteristics.

Subsequently, D2 and C6 provide a strong and weak bond, respectively, to attach D3 (v), which indicates the completion of the computing and decision-making process. Last, D3 exposes a weak bond to attach the reporting tile R1 (vi) that carries a success marker to the framework.

The DOFs exist in the solution on the nanomole scale per liter. Each DOF can function as a computer to perform the factorization independently at the same time, laying the foundation of the parallel processing abilities of this molecular paradigm. Apart from that, for a single DOF, the DNA tiles are added to the cavity in batches (Fig. 1D). Some tiles can attach to the DOF simultaneously in the assembly. This also contributes to the parallel processing abilities of the DOF strategy.

Programming the model with DNA

The sequences of the DX tiles were designed using a program written in Python (see Materials and Methods for details). Each DX tile has a length of 41 base pairs (bp) that is approximately four helical turns. The native polyacrylamide gel electrophoresis (PAGE) was used to verify the assembly of the DX tiles (Fig. 2A and fig. S5). By comparing the shifting mobility of assemblies containing different component strands to that of the ladder marker, we found that the hybridization of component strands was effective and the DX tile (taking C4 as an example) was properly synthesized with approximate mobility to the 100-bp band. Notably, the assembly containing either a + c or a + d strands shifts more slowly than that containing a + e strands despite their similar molecular weights, which is ascribed to their discrepancy in secondary structures. The DOF structure was designed using caDNAno (*35*), in which the gap distance in each DNA bundle was deliberately tailored to form a cavity for the tile assembly (fig. S6). As displayed in the AFM image (Fig. 2B), the DOF structure is well synthesized with the cavity unambiguously observed. A molecular dynamics (MD) simulation was conducted using oxDNA (*36*) to evaluate the effect of the cavity on the integral mechanics of the DOF (Fig. 2C). From the mean structure and the deviation analysis, the root mean square fluctuation (RMSF) of the overhang-carrying bundles along the cavity was estimated at ~5 nm, which certifies the DOF as a rigid flat template to direct the tile assembly.

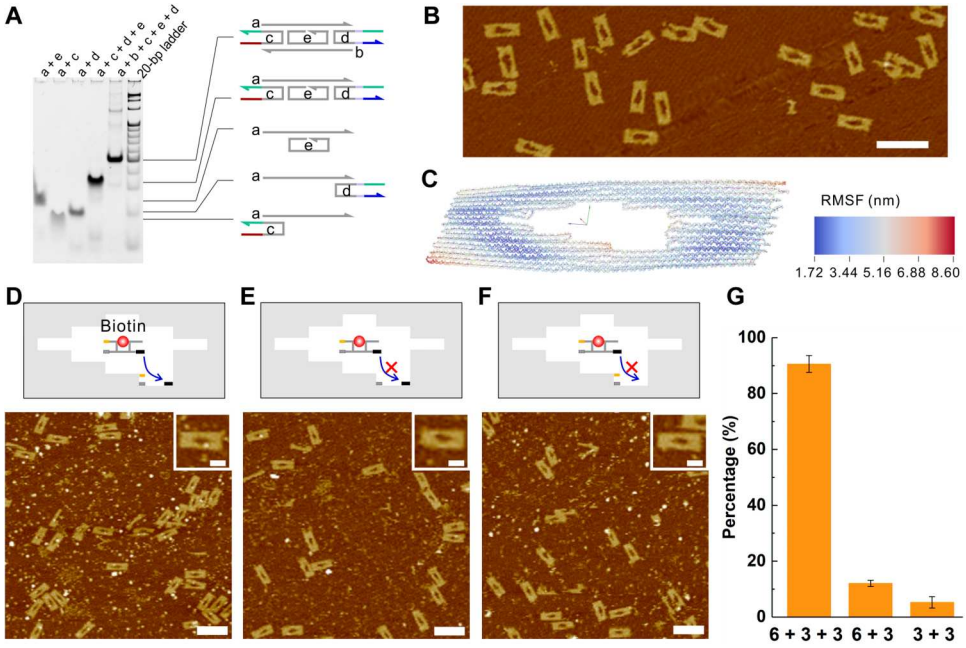

Fig. 2. Assembly of the DX tiles in the DOFs that carry different prime inputs for the factorization of 6. (**A**) Nondenaturing gel analysis of the formation of the DX tile. (**B**) Representative AFM image of the DOF. (**C**) MD simulation frame showing the mean structure and deviation of the DOF. Representative AFM images showing the attachment of the DX tile to the DOF by 6-nt + 3-nt + 3-nt (**D**), 6-nt + 3-nt (**E**), and two 3-nt overhangs (**F**). The DX tile used, R1, carries biotin labeling that can conjugate streptavidin markers for AFM imaging. (**G**) Statistics of streptavidin markers in the DOF cavity with different combinations of DNA overhangs. Error bars depicts the standard deviation. Scale bars, 200 nm (insets, 50 nm).

Then, we investigated the attachment of the DX tile to the DOF by different combinations of sticky ends to verify the strengths of DNA bonds for fulfilling the abstract model in Fig. 1B. Since two 6-nt overhangs have been proven to be a robust linkage in the formation of DNA tile arrays (*37*), we focused on the feasibility of adding DX tiles using the combinations of 6-nt + 3-nt + 3-nt, 6-nt + 3-nt, and 3-nt + 3-nt overhangs, respectively. A "success" reporting tile R1 with biotin labeling was attached to the DOFs carrying the three combinations of overhangs separately, and then streptavidin markers were added to identify if the tile was attached in AFM imaging (Fig. 2, D to F). We statistical calculated the distribution of streptavidin markers in the DOF cavity and found that 90.6 ± 3.0% of the DOFs carried a streptavidin marker at the expected position when the first combination of overhangs was used, which is close to the reported efficiency of biotin-streptavidin conjugation (Fig. 2G) (*38*, *39*). However, only 12.0 ± 1.2% and 5.2 ± 2.0% of the DOFs carried a streptavidin marker in the cavity when the 6-nt + 3-nt and 3-nt + 3-nt overhangs were used for tile attachment, respectively. It is therefore concluded that the DX tile can be attached to the DOF by cooperatively binding to a 6-nt and two 3-nt overhangs, whereas neither the 6-nt + 3-nt nor two 3-nt overhangs can fasten a DX tile to the cavity effectively, especially given the possible nonspecific absorption of streptavidin to the DOF. Streptavidin markers attached to the positions outside the DOF cavity were not included in the statistics. This lays the foundation that the decision-making and reporting tiles are added only when the inputs from three sides are all matched.

Factoring the semiprime 6 using the DOF strategy

Having physically mimicked the abstract model via localized DNA assembly, we demonstrate the prime factorization of 6 using the DOF strategy. As discussed in Fig. 1D, possible solutions of decomposing 6 include 2 × 2 (Fig. 3A) and 2 × 3 (Fig. 1B). To ensure that the product of the prime numbers equals the original number perfectly, only when all the three decision-making tiles that output a check mark (i.e., D1, D2, and D3) are added is the success reporting tile R1 attached. Otherwise, the failure reporting tile R2 is added, or the assembly is halted before reaching the end, both giving a negative signal. In the case of factoring 6, D1 occupies the first position of decision-making tiles, unless the assembly is terminated aforehand, while the addition of the following tiles is dependent on the product of prime inputs. Since the second digit of the product of 2 and 2 (100 in binary), i.e., 0, differs from that in 6 (110 in binary), a decision-making tile D4 takes place of D2 to output a cross mark (Fig. 3A). D5 that also has a cross output is then added and transfers the decision to a failure reporting tile R2.

We statistically analyzed the number of streptavidin markers in the cavity of DOFs (given by n) of different solutions in AFM imaging (Fig. 3B). For factorization of 6 as 2 × 2, only 4.6 ± 1.7% of the DOFs carry a streptavidin marker in the cavity, while 93.8 ± 3.1% of the DOFs have no streptavidin. Notably, more than one streptavidin marker could appear on the DOF mainly due to nonspecific absorption. In contrast, there are around 58.5 ± 4.9% of the DOFs that have a streptavidin marker in the cavity for factorization of 6 as 2 × 3. Given that the conjugation efficiency of biotin and streptavidin reaches ~92% (*39*), we estimated that up to ~63.6% of the DOFs carried the R1 tile of biotin labeling, revealing factoring 6 as 2 × 3 as the correct solution. The absence of

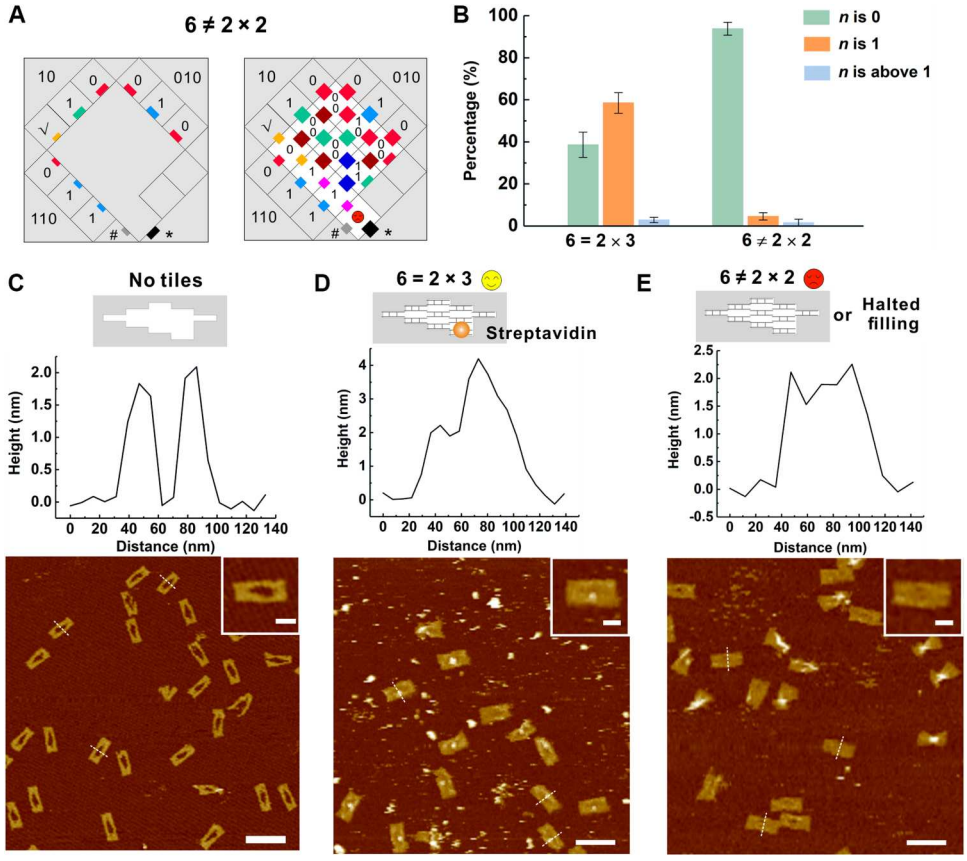

Fig. 3. Assembly of the DX tiles in the DOFs that carry different prime inputs for the factorization of 6. (**A**) Factoring 6 as 2 × 2 using the abstract model. Before (left) and after (right) the addition of the tiles. Because of the addition of decision-making tiles that output a cross mark, the failure reporting tile is attached at last and gives a negative result. (**B**) The distribution of streptavidin markers in the cavity of DOFs of different solutions for factoring 6. The number of streptavidin markers is denoted by *n*. Error bars depicts the SD. (**C**) Representative AFM image of the DOFs before the addition of DX tiles. Streptavidin markers were observed on the DOFs carrying primes 2 and 3 (**D**) but not on the DOFs carrying primes 2 and 2 (**E**) after the assembly of DX tiles. The height plot is averaged from the plots of height versus distance of the three DOFs denoted in white dotted lines. Scale bars, 200 nm (insets, 50 nm).

the R1 tile could be attributed to the false addition of R2 or disrupted tile assembly or scratched by the AFM tip.

Note that the DX tiles are added to the DOF in sequence (Fig. 1D). To be specific, a tile can be added only when its former tiles have all been added. Once a tile is missing from the DOF, the following tiles fail to be attached, and the factorization process for this solution is terminated. In this case, no reporting DX tile will occur on the DOF cavity. In light of this design, a final DOF displaying a streptavidin marker (on R1 tile) must contain all the tiles (for the right solution) in the cavity.

We constructed algorithmic structures using the DOF and an incomplete tile set for the right solution to validate that. The tiles D1, D2, D3, and a single D1 were omitted to investigate the effect of missing tiles on the DX tile assembly, respectively (fig. S7). The AFM images show that even the absence of a single D1 leads to unsuccessful filling of the DOF cavity, with less than 2% of DOFs carrying the streptavidin marker. Given the fact that missing tiles in the DOF lead to an unfilled cavity, we conclude that the presence of streptavidin marker on the DOF cavity would necessitate a complete tile set for the right solution added to the cavity.

The cavity appeared identically dark to the substrate in the AFM images before any DX tile was added (Fig. 3C). An average height plot of three representative DOFs denoted in white dotted lines is given to display the height details of the unfilled cavity featured by an abrupt drop in height from ~2 nm to nearly 0 nm. After the assembly of DX tiles, the cavity was filled and streptavidin markers were added to display the reporting result. Representative AFM images are given for the DOFs with prime inputs 2 × 3 (Fig. 3D and fig. S8) and 2 × 2 (Fig. 3E and fig. S9) after the addition of streptavidin, respectively. As shown in Fig. 3D, a ~2-nm protrusion appeared from the DOF substrate in the average height plot, corresponding to the streptavidin marker in the AFM image. The cavity with streptavidin remained slightly darker than the DNA bundles for some DOFs, supposedly ascribed to the segmented filling of DX tiles in the cavity that results in a lower DNA density. In addition, unlike the DOF extended on the mica, the geometrical flexibility of the DX tile array to adapt to the cavity could induce more influence of the AFM tip on the DX tiles by scratching the cavity in imaging. Besides, excess free DX tiles could also interfere with the AFM tip to influence the imaging quality and occupy the mica substrate to lower the contrast of DOF against the mica. To reduce the absorption of free DX tiles to the mica, 1× TAE buffer [40 mM tris, 20 mM acetic acid, and 2 mM EDTA (pH 8.0)] containing 12.5 mM Mg^{2+} and 100 mM Na^{+} was prepared as the imaging buffer

and used to rinse the mica gently before AFM imaging (*40*). In addition, we performed gel purification of the structures to recover DOFs of two solutions for factoring 6 (fig. S10). The DOFs with a streptavidin marker took up ~61% of the whole for the right solution, which is close to that for unpurified structures.

For the DOF with the inputs 2 and 2 (Fig. 3E), the cavity can be filled up after the assembly of DX tiles as well, featuring a ~0.5-nm depression between the substrate in the average height plot of the three representative DOFs. Nevertheless, there are few streptavidin markers on the DOFs due to the attachment of a biotin-free R2 tile instead of R1 or the absence of the reporting tile, indicating that the solution is incorrect. Note that the cavity in Fig. 3E appears slightly brighter than that in Fig. 3D, possibly because of the protruded height of streptavidin that affected the visual contrast in Fig. 3D. Alternatively, the streptavidin induced the geometrical changes of DX tiles, rendering them more likely to be scratched by the AFM tip.

The DOFs and DX tiles were incubated at 30°C following an anneal from 45°C for the tile assembly. Inspired by the earlier work that synthesized DNA crystals at constant temperatures (*40*, *41*), we considered the length of the sticky ends and initially set the range of incubation temperature from 26° to 34°C. Besides, an anneal from 45°C to the incubation temperature was added at the start to make the assembly more robust. We found that when incubated at 26°C, ~42.9% of the DOFs carried a streptavidin marker at the cavity (fig. S11). When the assembly temperature came to 34°C, only ~38.4% of the DOFs were with a streptavidin marker at the cavity. The decrease in the efficiency could largely be due to the influence of the elevated temperature to the 3-nt-short overhangs of the DOFs. Both the temperatures were inferior to 30°C in the efficiency of tile assembly. Hence, we chose to incubate the DOFs and DX tiles at 30°C following an anneal from 45°C for the tile assembly.

Factoring the semiprime 15 using the DOF strategy

Having established the feasibility of this DOF strategy, we sought to factor a larger semiprime 15. Since the smallest prime number is 2, only prime inputs smaller than 15 divided by 2 (i.e., 2, 3, 5, and 7) were considered for the factorization. We randomly chose three pairs of prime numbers 3 and 5, 2 and 3, and 2 and 7 and verified each pair using the DOF method to find the right solution (fig. S12). The logic circuits for the factorization (Fig. 4A) resemble that for factoring 6, except that more multiplication and comparator units are used because of the increasing number of bits for displaying 15 in binary. In the DNA implementation of factoring 15 (Fig. 4B), representative DX tiles are displayed to specify the correspondence between the binary numbers (or marks) and colors. Likewise, the light purple segment incorporated to the right ends of computing tiles or both ends of decision-making/reporting tiles is a three-thymine spacer to reduce the internal strain of the DNA assemblies. In addition, the verification result is revealed by the presence or absence of biotin labeling in the reporting DX tile. At several corners of the cavity, we placed 41-bp DNA duplexes (in black) joint to the DNA origami by staple crossovers to carry specific overhangs (Fig. 4B and fig. S13), thus allowing a more compact filling of the DX tiles without threading through the interhelical distances between adjacent tiles (Fig. 1C, inset). Consequently, the DX tiles could fill the cavity in a continuous rather than segmented manner. Because of the increased number of involved DX tiles that is beyond the scope of a DNA origami monomer, we combined two monomers that carried specific sticky ends to construct a larger DOF (Fig. 4C and fig. S14). The long or short thick lines are used separately to denote the 6- or 3-nt overhangs that are specially colored to represent the integer 15 and the inputs for directing the localized assembly, displayed here include the primes 3 and 5. The MD simulation gives the mean structure and the deviation of a DNA origami monomer (Fig. 4D). The concave interiors of the DOF that carry the overhangs have an RMSF of approximately 5 nm, close to that in the factorization of 6. The monomer shows a featured concave side in AFM imaging (Fig. 4E), and two monomers can make a DOF with an expected cavity at the center. Five linking strands were used to fix the DOF by hybridizing a 14-nt domain at every other bundle of each DNA origami monomer (fig. S14). The yield of the assembly of DOFs was calculated at ~69.6% by AFM imaging (Fig. 4F and fig. S15). As presented in the average height plot of the three DOFs (denoted by a white dotted line), the cavity is displayed as an abrupt drop of the height from more than 1.5 nm to nearly 0 nm.

For the factorization of 15 as 2×7 (Fig. 5A) or 2×3 (Fig. 5B), the reporting tile R2 is attached to the last position and reveals a negative signal following the addition of a decision-making tile with a cross output. Statistics of the streptavidin markers in the cavity of DOFs of different solutions in AFM imaging reveal that $51.8 \pm 3.2\%$, $12.0 \pm 4.7\%$, and $11.1 \pm 5.3\%$ of the DOFs carry a streptavidin in the cavity for factorization of 15 as 3×5, 2×7, and 2×3, respectively, suggesting the prime inputs 3 and 5 as the correct solution (Fig. 5C). Given the incompleteness of biotin-streptavidin conjugation, the percentage of DOFs with a biotinylated R1 tile can be up to ~56.3% for factoring 15 as 3×5. It is worth noting that there is a drop in the fraction of DOFs with an R1 tile in the correct factorization of 15, compared to that for factoring 6 as 2×3. We attribute that to the increased complexity of the model that requires the cooperative assembly of more DX tiles in the DNA implementation, which renders the factorization more susceptible to errors for weakened sequence orthogonality.

Representative AFM images of the DOFs that factor 15 as 3×5 are given in Fig. 5D and fig. S16. The streptavidin marker on the DOF induces a ~2-nm protrusion from the substrate as shown in the averaged height plot (Fig. 5D). The cavity of the DOF still appears darker than the DNA bundles, supposedly owing to the tile array with increased distortion that is more prone to be scratched by the tip. Besides, the DOF cavity is not thoroughly filled at the bottom right (Fig. 4B), with a margin remaining dark that could appear "defective" in AFM images. We omitted D1, D2, and a single D1 to verify the effect of missing tiles from the DOF on the assembly separately (fig. S17). Likewise, few streptavidin markers were observed on the DOF in the AFM images.

In the AFM images for the DOFs of the other solutions (Fig. 5, E and F, and fig. S18), the DOF cavity was filled with DX tiles but few streptavidin markers were observed in the cavity. Likewise, we found that the cavity seemed darker than the DNA bundles, featured by the depression of ~0.5 nm in the averaged height plot of the three representative DOFs that have a filled cavity.

Last, to verify if the DX tiles create unwanted connections with each other without the DOF, all the DX tiles (total concentration: ~0.6 and ~0.8 μM for tiles in factoring 6 and 15, respectively) were annealed and incubated without the DOF. In AFM imaging without rinsing the mica, a considerable amount of DX tiles was displayed (fig. S19), some tiles supposedly forming assembly. However, when diluted to one sixth the initial concentration, few

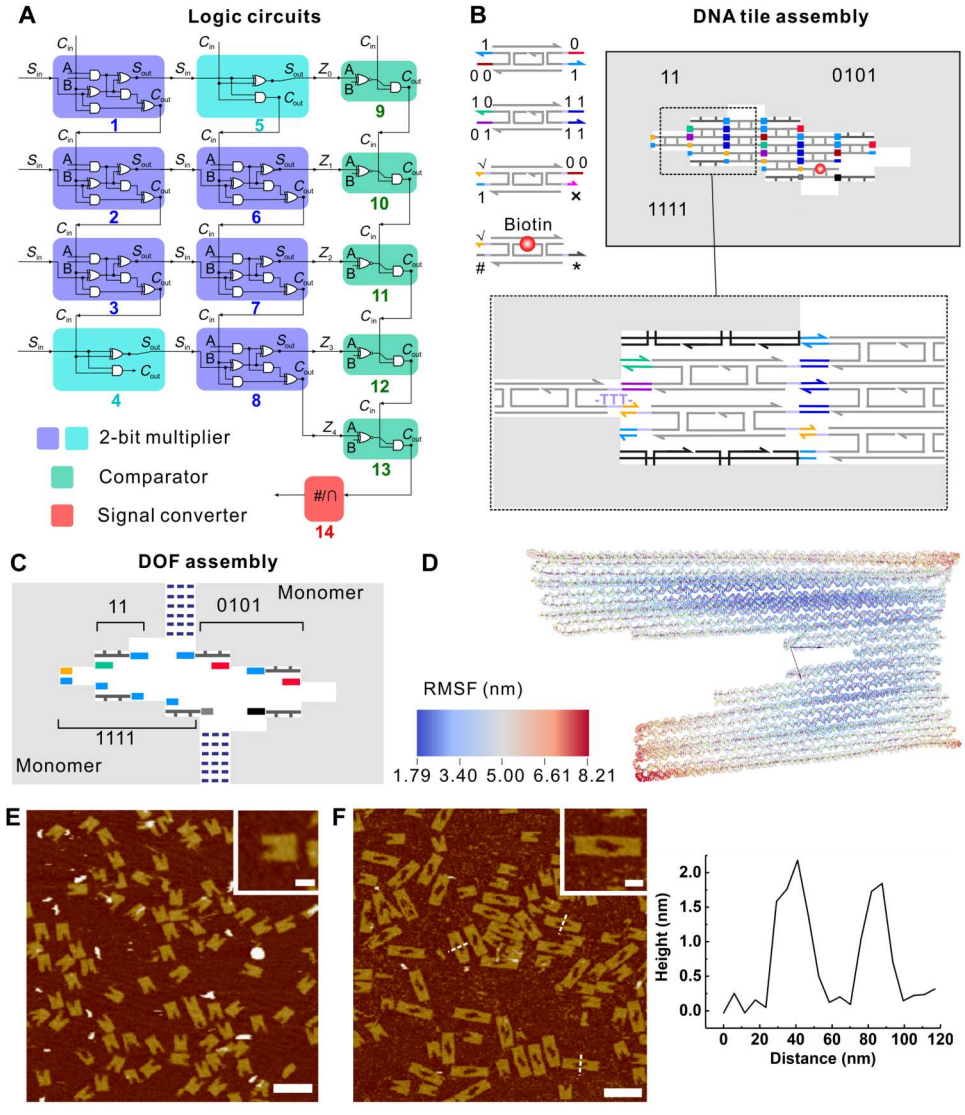

Fig. 4. Prime factorization of 15 using the assembly of DX tiles in DOF. (**A**) Logic circuits to perform the prime factorization. Since the number 15 occupies only four bits, four comparator tiles will suffice to realize the factorization. (**B**) DNA implementation of the factorization of 15. Representative DX tiles are given to display the correspondence between the binary numbers (or marks) and colors. In addition, a three-thymine spacer (light purple) is incorporated to the right ends of computing tiles or both ends of decision-making/reporting tiles. DNA duplexes (in black) joint to the DOF by staple crossovers are placed at several corners of the cavity to carry specific overhangs. (**C**) The DOF for the factorization of 15. Because of the limited size of a single DNA scaffold, we fabricated the DOF by joining two DNA origami monomers that carried specific overhangs. The long or short thick lines separately represent the 6-or 3-nt overhangs that are colored to specify the correspondence between the primes and the original integer. (**D**) MD simulation frame showing the mean structure and deviation of the DNA origami monomer. Representative AFM images of the DNA origami monomer (**E**) and the DOF (**F**). The height plot is averaged from the plots of height versus distance of the three representative DOFs denoted in white dotted lines. Scale bars, 200 nm (insets, 50 nm).

assembly-like structures were observed. Further evidence was provided by agarose gel images (fig. S20). A single DX tile (C1 for 6) of both concentrations was used as the reference. Both the high- and low-concentration DX tiles for 6 had a single band with the same mobility as that of a single C1, suggesting that there were few unwanted connections between different tiles, even at a high concentration. The low-concentration DX tiles for 15 had the same mobility as that of a single C1. Nonetheless, a faint lagging band that could be related to tile assembly appeared for the tiles at high concentration. Overall, we deduce that at the high concentration, the tile assembly has a chance to occur (especially tiles for 15) but cannot be overwhelming.

DISCUSSION

We have demonstrated a molecular programming paradigm that performs the factorization of semiprimes through the localized assembly of DX tiles in a prescribed cavity of the DOF. The abstract model of factorization consists of a multiplication core, a decision-making motif, and a reporting motif, in which the arithmetic operations are practically implemented by the stepwise addition of the

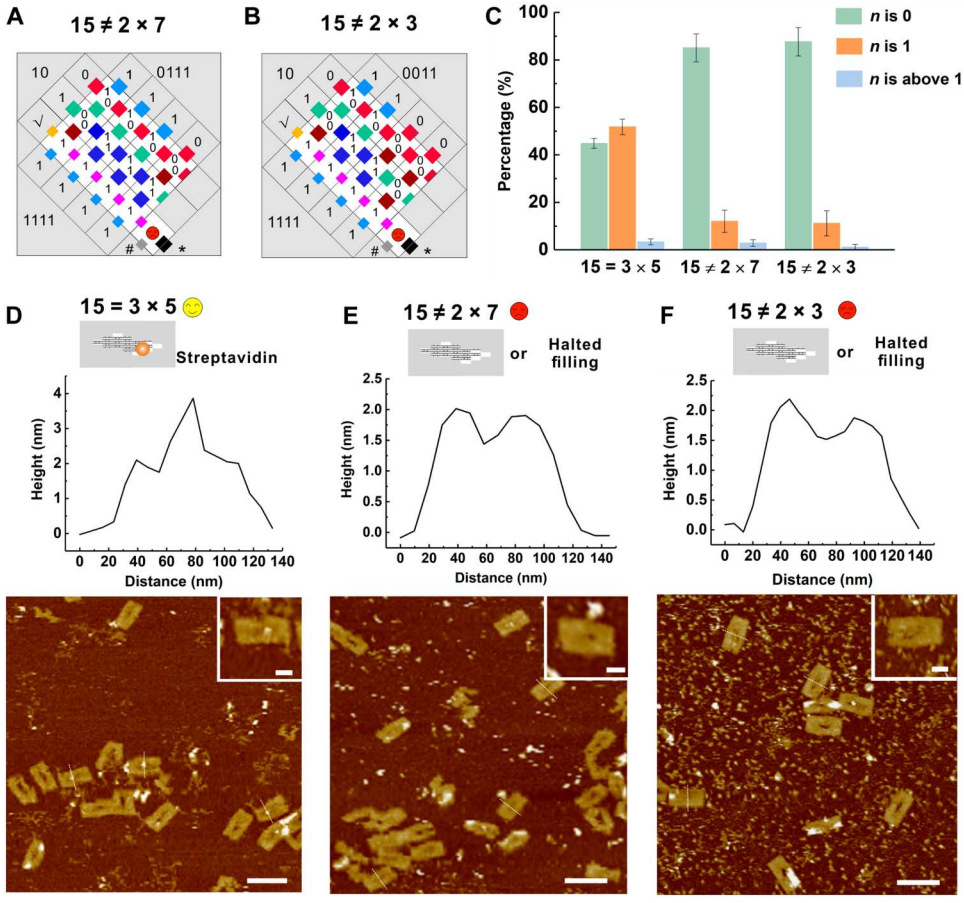

Fig. 5. Assembly of the DX tiles in the DOFs carrying different factors for prime factorization of the integer 15. Factoring 15 as 2 × 7 (**A**) or 2 × 3 (**B**) using the abstract model. The result is reported as negative for the addition of decision-making tiles that output a cross mark. (**C**) The distribution of streptavidin markers in the cavity of DOFs of different solutions for factoring 15. The number of streptavidin markers is denoted by *n*. Error bars depicts the SD. Representative AFM images of the DOFs carrying primes 3 and 5 (**D**), 2 and 7 (**E**), and 2 and 3 (**F**) after the streptavidin markers were added. The height plot is averaged from the plots of height versus distance of the three DOFs denoted in white dotted lines. Scale bars, 200 nm (insets, 50 nm).

DX tiles to the DOF. The binary primes and composite integer are encoded as DNA overhangs spatially organized along the cavity of the DOF to direct the assembly, and the DX tile has sticky ends encoding the algorithmic information of the input or output for the intertile or DOF-tile recognition. Localization of the DNA components on the DOF interface reduces interferences from diffusible strands and favors increased complexity of the molecular reaction networks. Through the modular assembly of the DNA overhangs that represent the binary digits, we explored the factorization of semiprimes 6 and 15 into different pairs of primes respectively and picked out the right solutions.

We developed the DNA machine to realize the prime factorization that is fundamental to the Rivest-Shamir-Adleman public cryptography by molecular computing. This strategy is built on a mathematical model that is substantially more complex than that of the earlier work of Winfree and Rothemund on programmable DNA logic (*42*–*44*).

First, the model we used is superior in realizing more complex functions. In the earlier work of Winfree and Rothemund, exemplified by the self-assembled Sierpinski triangle (*42*, *45*) and the DNA algorithms for copying and counting (*43*, *46*), commonly no more than two logic operations, XOR and AND, were demonstrated in computing. By comparison, in the model presented for the prime factorization, we introduced the functions including multiplication and bit-by-bit comparison that require AND, OR, XOR, and NOT logic operations to collaboratively perform complex calculations, thus finalizing the factorization.

Second, to fulfill the custom functions, the model we used comprises more kinds of tiles, featuring more complexity in physical design. In some earlier work of Winfree and Rothemund, four different DNA tiles are typically used to perform the logic operations (XOR and AND) through tile self-assembly (*43*). In later work that used DNA origami seeds for nucleating algorithmic tile assembly (*46*), up to 18 different tiles were used in the assembly. In addition, we have noted that ~22 DNA tiles have been involved to construct algorithmic structures (*47*). By comparison, the abstract model we gave for prime factorization can be associated with up to 46 abstract tiles. Given the complementarity of DNA domains, the number of corresponding DNA tiles could double. In the factorization of 6 and 15, the DOFs picked up the right solutions from 15 and 21 tiles, respectively.

In particular, to implement the decision-making motifs, we set up three inputs in the DNA tiles and engineered the bond strengths to guarantee that a tile is added only when all three inputs were matched. This design differs from the earlier work in which only two inputs are used for each tile. Notably, it allows the bit-by-bit comparison of the product of primes with the original integer, which is a critical function of the model. We calculate that the overall error rate per tile is approximately 4% in the prime factorization, sufficing for picking the right solutions for small composite numbers. Next work would include refining the physical design and experimental conditions to allow for factoring larger numbers.

Because of the limited size of the scaffold DNA available (*48*), it remains difficult to build giant DOF structures that allow the incorporation of more DX tiles to represent digits for a larger semiprime. Synthetic advances have been made that enable mass production of scaffold DNA at a low cost, whereas the efficient synthesis of extra-long single-strand DNA still poses a challenge that practically restricts the size of the DOF (*49*, *50*). Alternatively, higher-order structures self-assembled from a set of DNA origami monomers (*51*), similar to the one presented in Fig. 4C, can be used to provide more space for the addition of DX tiles. The DOF assembly should retain mechanical rigidity instead of being floppy aggregates to restrain the positions of DNA overhangs. DNA single-stranded tiles (SSTs) present a modular approach for assembling DNA nanoarchitectures surpassing the limit of available scaffold DNA (*52*). Using the SST method, we could make larger frameworks to allow more DNA tiles to represent the prime numbers. Nevertheless, some shortcomings of the SST need to be addressed. The yield of SST structures is relatively low, especially for large ones. In addition, individual strands are required to be stoichiometrically equal, which could render the preparation process rather laborious and time-consuming.

Alternative approaches to circumventing this problem of scaffold DNA include using smaller DNA blocks to denote the binary digits for the primes. The DX tile has a length of about four helical turns (41 bp) that is ~13.9 nm. Reducing the tile length, e.g., using a shorter e strand, could help incorporate more tiles into the DOF. Nevertheless, there is rather limited room for improvement for implementation with DX tiles. The next attempts also include using single-stranded DNA tiles instead to carry information.

Another notable problem is that the efficiency of the tile assembly is diminished by the added complexity of DX tiles, thus restraining the potential of the DOF method to explore the factorization of large semiprimes that requires the cooperation of more tiles. The number of possible solutions increases as the composite number N gets larger. As a result, the proportion of DNA tiles that constitute the correct solution inevitably shrinks, and the tiles of the wrong answers bring more disturbances to the factoring. This could pull down the efficiency of the DOF method. By sorting possible solutions into different groups and conducting parallel tests, it is feasible, although troublesome, to find the right solution for a large semiprime using this method. Marking different primes on the DOFs and performing one-tube tests for each group [i.e., identifying various patterns simultaneously in AFM imaging (*39*)] favor the acceleration of the search process for the right solution.

The factoring process typically took more than 4 hours to give the result from the tile-filled DOF, including the time for characterization and analytical statistics of the streptavidin sample. We used normal AFM to characterize the samples. Each scan typically took from 3 to 6 min. In addition, the preparation of the DNA origami and tiles, as well as the addition of DNA tiles to the DNA origami, took more than 12 hours. The whole process can be greatly accelerated by the high-speed AFM, automated sample processing, and image analysis systems. Although the time of DOF-based factoring is still long compared to that of the silicon-based approach for small numbers, this method could show merits for factorization of large numbers as a molecular route of parallel processing.

The DOF method demonstrates the potentials of the intelligent molecule DNA to handle complicated mathematical problems. Despite the existing tackles on the robust factorization of large semiprimes, this strategy opens avenues to boost molecular computing by combing the information-processing potential of DNA and its ability in precisely fabricating custom structures that can locate the computing components.

MATERIALS AND METHODS

Materials

M13mp18 viral DNA was purchased from Tilibit Nanosystems and used without further purification. The staple strands with PAGE purification were purchased from Jie Li Biology. Biotinylated DNA strands with high-performance liquid chromatography purification were purchased from Sangon Biotech. Streptavidin was purchased from Sigma-Aldrich. Freeze 'N Squeeze columns were purchased from Bio-Rad. All other chemicals were purchased from Sinopharm. Water was purified with a Millipore Milli-Q Integral water purification system (resistivity = 18.2 megohm·cm).

DNA sequence design

DNA sequences of the DX tiles were designed using a custom program written in Python. To determine the appropriate sequences of the sticky ends, the program first took the sequence length as the input, then automatically screened all the possible sequences with the desired length, and lastly outputted the chosen sequences as a text file. The basic constraints for screening are as follows:

1) The sequences should be of a certain length with A, T, C, and G only.

2) The proportion of (G + C) should be within 40 to 60% to make a relatively narrow annealing temperature.

3) No consecutive nucleotide repeats of longer than 3 are permitted.

4) No repeated sequences of longer than 5 are permitted.

5) The Hamming distance of any two sequences should be more than 40% of the sequence length to ensure recognition orthogonality.

6) The Hamming distance between a sequence and its reverse-complement counterpart should be less than 5 nt to avoid the formation of hairpin structures.

Note that to guarantee the addition of tiles to the framework in the correct direction, the recognition domains flanking the cavity are deliberately devised to differ from the others. For instance, the 0 domain at the top right of tile C3 bound to the cavity is different from the 0 domain at the left of this tile, thus forbidding the rotation of this tile when added (Fig. 1B). Besides, the bottom left of decision-making tiles is encoded as a 3-nt DNA overhang that differs from the 6-nt overhang for other sides.

DX tile assembly

The strands composing each DX tile were mixed at an equal molar ratio in 1× TAE buffer with 12.5 mM Mg^{2+}. The mixture was then subjected to an anneal from 85°C to 4°C at a rate of −5°C/min and held at 4°C finally. To characterize the efficiency of the assembly, the products were analyzed by 6% native PAGE using 1× TAE buffer with 12.5 mM Mg^{2+} in a gel box immersed in an ice bath. The gel was stained with GelRed (Biotium) and imaged with a G:BOX Chemi system (Syngene) after running at 100 V for about 1 hour.

DNA origami folding

The DOF was designed using caDNAno. The staple strands were mixed with the scaffold strand at a molar ratio of 10:1 in 1× TAE buffer with 12.5 mM Mg^{2+}. Subsequently, the mixture was slowly annealed from 85° to 4°C in 2 hours using a Veriti Thermal Cycler (Applied Biosystems). Folded DNA origami was purified with 100-kDa (molecular weight cutoff) centrifugal filters (Amicon) three times at a speed of 3000*g* to remove excess staple strands.

Assembly of DNA origami dimers

For assembly of the DOF composed of two DNA origami monomers, each monomer was mixed at equal amounts in 1× TAE buffer with 12.5 mM Mg^{2+}. Linking strands and the duplex-forming strands that carry the DNA overhangs were then added to the monomers at a molar ratio of 50:50:1. The mixture was slowly annealed from 45° to 30°C at a rate of −0.1°C/min and held at 4°C finally. Free strands were removed through ultrafiltration.

DX tile assembly in the DOF

The DOFs at a concentration of 2 nM were mixed with DX tiles of 40 nM each in 1× TAE buffer with 12.5 mM Mg^{2+}. The mixture was annealed from 45° to 30°C at a rate of −0.1°C/min and then incubated at 30°C for 12 hours. Afterward, streptavidin markers were added at a molar ratio of 10:1 to the biotin labels on the DX tiles. The mixture was incubated for 2 hours at room temperature before characterizing with the AFM.

AFM imaging

A droplet (~2 μl) was deposited on a freshly cleaved mica surface and left to absorb for 3 min. After that, 40 μl of 1× TAE buffer with 12.5 mM Mg^{2+} was added to the liquid cell, and a ScanAsyst Fluid tip (Bruker) was used to scan the sample in a PeakForce-tapping mode on a Multimode VIII AFM (Bruker). A minimum force was maintained in imaging to reduce the effect of scratching of streptavidin by the AFM tip that could lead to false-negative results. For statistical analysis of the streptavidin marker in the DOF cavity, three repeated measurements were taken with ~1.5 nM sample in a view of ~2 μm by 2 μm or 3 μm by 3 μm. When excess free tiles absorbed on the mica compromised the imaging quality, 1× TAE buffer containing 12.5 mM Mg^{2+} and 100 mM Na^{+} was prepared as the imaging buffer and used to rinse the mica gently before AFM imaging.

Gel purification

A 0.5% agarose gel analysis (running buffer: 1× TAE buffer with 12.5 mM Mg^{2+}, 5 V/cm) at an ice bath was performed to purify the structures from the samples. The concentration of DOFs should be less than 5 nM since a high concentration of DOFs tends to aggregate in tile assembly. The target band was excised from the gel and chopped into pieces. The product was then extracted with Freeze 'N Squeeze columns.

Supplementary Materials

This PDF file includes:

Supplementary Text
Tables S1 and S2
Figs. S1 to S20
DNA Sequences

REFERENCES AND NOTES

1. N. R. Potlapally, S. Ravi, A. Raghunathan, N. K. Jha, A study of the energy consumption characteristics of cryptographic algorithms and security protocols. *IEEE Trans. Mob. Comput.* **5**, 128–143 (2006).
2. G. Kelly, B. McKenzie, Security, privacy, and confidentiality issues on the internet. *J. Med. Internet Res.* **4**, e12 (2002).
3. P. Q. Nguyen, Can we trust cryptographic software? Cryptographic flaws in GNU privacy guard v1.2.3. *Lect. Notes Comput. Sci.* **3027**, 555–570 (2004).
4. R. Steinfeld, Y. L. Zheng, A signcryption scheme based on integer factorization. *Lect. Notes Comput. Sci.* **1975**, 308–322 (2000).
5. P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM. J. Comput.* **26**, 1484–1509 (1997).
6. M. N. Leuenberger, D. Loss, Quantum computing in molecular magnets. *Nature* **410**, 789–793 (2001).
7. L. M. Adleman, Molecular computation of solutions to combinatorial problems. *Science* **266**, 1021–1024 (1994).
8. Y. Benenson, B. Gil, U. Ben-Dor, R. Adar, E. Shapiro, An autonomous molecular computer for logical control of gene expression. *Nature* **429**, 423–429 (2004).
9. M. N. Stojanovic, D. Stefanovic, S. Rudchenko, Exercises in molecular computing. *Acc. Chem. Res.* **47**, 1845–1852 (2014).
10. D. Woods, D. Doty, C. Myhrvold, J. Hui, F. Zhou, P. Yin, E. Winfree, Diverse and robust molecular algorithms using reprogrammable DNA self-assembly. *Nature* **567**, 366–372 (2019).
11. L. Qian, E. Winfree, Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**, 1196–1201 (2011).
12. R. S. Braich, N. Chelyapov, C. Johnson, P. W. Rothemund, L. Adleman, Solution of a 20-variable 3-SAT problem on a DNA computer. *Science* **296**, 499–502 (2002).
13. E. Winfree, Algorithmic self-assembly of DNA: Theoretical motivations and 2D assembly experiments. *J. Biomol. Struct. Dyn.* **17** (Suppl. 1), 263–270 (2000).
14. G. Seelig, D. Soloveichik, D. Y. Zhang, E. Winfree, Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–1588 (2006).
15. W. Li, Y. Yang, H. Yan, Y. Liu, Three-input majority logic gate and multiple input logic circuit based on DNA strand displacement. *Nano Lett.* **13**, 2980–2988 (2013).
16. T. Song, A. Eshra, S. Shah, H. Bui, D. Fu, M. Yang, R. Mokhtar, J. Reif, Fast and compact DNA logic circuits based on single-stranded gates using strand-displacing polymerase. *Nat. Nanotechnol.* **14**, 1075–1081 (2019).
17. S. Roweis, E. Winfree, R. Burgoyne, N. V. Chelyapov, M. F. Goodman, P. W. Rothemund, L. M. Adleman, A sticker-based model for DNA computation. *J. Comput. Biol.* **5**, 615–629 (1998).
18. P. Yin, A. J. Turberfield, S. Sahu, J. H. Reif, Design of an autonomous DNA nanomechanical device capable of universal computation and universal translational motion. *Lect. Notes Comput. Sci.* **3384**, 426–444 (2005).
19. Z. Wang, Y. Zhang, W. Zhou, H. Liu, Solving traveling salesman problem in the Adleman–Lipton model. *Appl. Math Comput.* **219**, 2267–2270 (2012).
20. Y. Brun, Arithmetic computation in the tile assembly model: Addition and multiplication. *Theor. Comput. Sci.* **378**, 17–31 (2007).
21. C. M. Agapakis, P. M. Boyle, P. A. Silver, Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nat. Chem. Biol.* **8**, 527–535 (2012).
22. G. Chatterjee, N. Dalchau, R. A. Muscat, A. Phillips, G. Seelig, A spatially localized architecture for fast and modular DNA computing. *Nat. Nanotechnol.* **12**, 920–927 (2017).
23. H. Bui, S. Shah, R. Mokhtar, T. Song, S. Garg, J. Reif, Localized DNA hybridization chain reactions on DNA origami. *ACS Nano* **12**, 1146–1155 (2018).

24. A. J. Thubagere, W. Li, R. F. Johnson, Z. Chen, S. Doroudi, Y. L. Lee, G. Izatt, S. Wittman, N. Srinivas, D. Woods, E. Winfree, L. Qian, A cargo-sorting DNA robot. *Science* **357**, eaan6558 (2017).
25. K. Lund, A. J. Manzo, N. Dabby, N. Michelotti, A. Johnson-Buck, J. Nangreave, S. Taylor, R. Pei, M. N. Stojanovic, N. G. Walter, E. Winfree, H. Yan, Molecular robots guided by prescriptive landscapes. *Nature* **465**, 206–210 (2010).
26. H. Gu, J. Chao, S. J. Xiao, N. C. Seeman, A proximity-based programmable DNA nanoscale assembly line. *Nature* **465**, 202–205 (2010).
27. P. W. K. Rothemund, Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
28. S. Dey, C. Fan, K. V. Gothelf, J. Li, C. Lin, L. Liu, N. Liu, M. A. D. Nijenhuis, B. Saccà, F. C. Simmel, H. Yan, P. Zhan, DNA origami. *Nat. Rev. Methods Primers* **1**, 13 (2021).
29. Y. Zhang, Z. B. Qu, C. Jiang, Y. Liu, R. Pradeep Narayanan, D. Williams, X. Zuo, L. Wang, H. Yan, H. Liu, C. Fan, Prescribing silver chirality with DNA origami. *J. Am. Chem. Soc.* **143**, 8639–8646 (2021).
30. J. Shen, W. Sun, D. Liu, T. Schaus, P. Yin, Three-dimensional nanolithography guided by DNA modular epitaxy. *Nat. Mater.* **20**, 683–690 (2021).
31. J. Chao, J. Wang, F. Wang, X. Ouyang, E. Kopperger, H. Liu, Q. Li, J. Shi, L. Wang, J. Hu, L. Wang, W. Huang, F. C. Simmel, C. Fan, Solving mazes with single-molecule DNA navigators. *Nat. Mater.* **18**, 273–279 (2019).
32. N. C. Seeman, DNA in a material world. *Nature* **421**, 427–431 (2003).
33. S. X. Jiang, N. Pal, F. Hong, N. E. Fahmi, H. Y. Hu, M. Vrbanac, H. Yan, N. G. Walter, Y. Liu, Regulating DNA self-assembly dynamics with controlled nucleation. *ACS Nano* **15**, 5384–5396 (2021).
34. P. Sa-Ardyen, A. V. Vologodskii, N. C. Seeman, The flexibility of DNA double crossover molecules. *Biophys. J.* **84**, 3829–3837 (2003).
35. S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, W. M. Shih, Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res.* **37**, 5001–5006 (2009).
36. E. Poppleton, R. Romero, A. Mallya, L. Rovigatti, P. Sulc, OxDNA.org: A public webserver for coarse-grained simulations of DNA and RNA nanostructures. *Nucleic Acids Res.* **49**, W491–W498 (2021).
37. E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, Design and self-assembly of two-dimensional DNA crystals. *Nature* **394**, 539–544 (1998).
38. N. Y. Wong, H. Xing, L. H. Tan, Y. Lu, Nano-encrypted Morse code: A versatile approach to programmable and reversible nanoscale assembly and disassembly. *J. Am. Chem. Soc.* **135**, 2931–2934 (2013).
39. Y. Zhang, Q. Li, X. Liu, C. Fan, H. Liu, L. Wang, Prescribing DNA origami patterns via scaffold decoration. *Small* **16**, e2000793 (2020).
40. R. Schulman, B. Yurke, E. Winfree, Robust self-replication of combinatorial information via crystal growth and scission. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6405–6410 (2012).
41. R. Schulman, E. Winfree, Synthesis of crystals with a programmable kinetic barrier to nucleation. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15236–15241 (2007).
42. P. W. Rothemund, N. Papadakis, E. Winfree, Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol.* **2**, e424 (2004).
43. R. D. Barish, P. W. Rothemund, E. Winfree, Two computational primitives for algorithmic self-assembly: Copying and counting. *Nano Lett.* **5**, 2586–2592 (2005).
44. D. Y. Zhang, R. F. Hariadi, H. M. T. Choi, E. Winfree, Integrating DNA strand-displacement circuitry with DNA tile self-assembly. *Nat. Commun.* **4**, 1965 (2013).
45. K. Fujibayashi, R. Hariadi, S. H. Park, E. Winfree, S. Murata, Toward reliable algorithmic self-assembly of DNA tiles: A fixed-width cellular automaton pattern. *Nano Lett.* **8**, 1791–1797 (2008).
46. R. D. Barish, R. Schulman, P. W. Rothemund, E. Winfree, An information-bearing seed for nucleating algorithmic self-assembly. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6054–6059 (2009).
47. C. G. Evans, "Crystals that count! Physical principles and experimental investigations of DNA tile self-assembly," thesis, Caltech (2014).
48. G. Tikhomirov, P. Petersen, L. Qian, Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature* **552**, 67–71 (2017).
49. F. Praetorius, B. Kick, K. L. Behler, M. N. Honemann, D. Weuster-Botz, H. Dietz, Biotechnological mass production of DNA origami. *Nature* **552**, 84–87 (2017).
50. Y. Jia, L. Chen, J. Liu, W. Li, H. Gu, DNA-catalyzed efficient production of single-stranded DNA nanostructures. *Chem* **7**, 959–981 (2021).
51. G. Yao, F. Zhang, F. Wang, T. Peng, H. Liu, E. Poppleton, P. Sulc, S. Jiang, L. Liu, C. Gong, X. Jing, X. Liu, L. Wang, Y. Liu, C. Fan, H. Yan, Meta-DNA structures. *Nat. Chem.* **12**, 1067–1075 (2020).
52. W. Wang, S. Chen, B. An, K. Huang, T. Bai, M. Xu, G. Bellot, Y. Ke, Y. Xiang, B. Wei, Complex wireframe DNA nanostructures from simple building blocks. *Nat. Commun.* **10**, 1067 (2019).

Acknowledgments

Funding: This work is supported by the National Key R&D Program of China 2018YFA0902600, NSFC 21991134, and NSFC T2188102. **Author contributions:** Y.Z., X.Y., X.B., and C.F. conceived the project. Y.Z., X.Y., C.C., and J.C. performed the experiments. Y.Z., X.Y., and X.Z. designed the DNA sequences on the algorithm assembly. K.H., T.L., F.W., and N.W. made valuable suggestions on optimizing the experimental conditions. Y.Z., X.Y., A.L., L.W., and C.F. wrote the manuscript. X.Z. and C.F. supervised the research. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The Python codes that support the findings of this study are available at https://doi.org/10.5281/zenodo.7554909.

Submitted 16 November 2022
Accepted 27 February 2023
Published 31 March 2023
10.1126/sciadv.adf8263