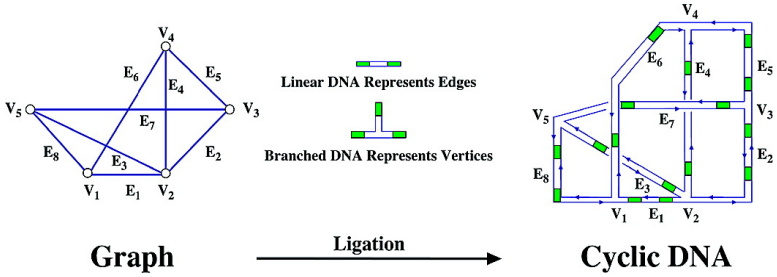
Article

Self-Assembly of Irregular Graphs Whose Edges Are DNA Helix Axes

Phiset Sa-Ardyen, Natasha Jonoska, and Nadrian C. Seeman

J. Am. Chem. Soc., **2004**, 126 (21), 6648-6657• DOI: 10.1021/ja049953d • Publication Date (Web): 07 May 2004



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 05/07/2004

Self-Assembly of Irregular Graphs Whose Edges Are DNA Helix Axes

Phiset Sa-Ardyen,[†,§] Natasha Jonoska,[*,‡] and Nadrian C. Seeman[*,†]

Contribution from the Department of Chemistry, New York University, New York, New York 10003, and Department of Mathematics, University of South Florida, Tampa, Florida 33620

Received January 4, 2004; E-mail: ned.seeman@nyu.edu; jonoska@math.usf.edu

Abstract: A variety of computational models have been introduced recently that are based on the properties of DNA. In particular, branched junction molecules and graphlike DNA structures have been proposed as computational devices, although such models have yet to be confirmed experimentally. DNA branched junction molecules have been used previously to form graph-like three-dimensional DNA structures, such as a cube and a truncated octahedron, but these DNA constructs represent regular graphs, where the connectivities of all of the vertexes are the same. Here, we demonstrate the construction of an irregular DNA graph structure by a single step of self-assembly. A graph made of five vertexes and eight edges was chosen for this experiment. DNA branched junction molecules represent the vertexes, and duplex molecules represent the edges; in contrast to previous work, specific edge molecules are included as components. We demonstrate that the product is a closed cyclic single-stranded molecule that corresponds to a double cover of the graph and that the DNA double helix axes represent the designed graph. The correct assembly of the target molecule has been demonstrated unambiguously by restriction analysis.

Introduction

Structural DNA nanotechnology uses reciprocal exchange between DNA double helices or hairpins to produce branched DNA motifs. These branched motifs can be combined via sticky-ended cohesion to produce specific graphlike structures.[1,2] The power of sticky-ended cohesion is that it leads to predictable adhesion between components whose product has a known structure.[3] DNA stick-polyhedra, such as a cube[4] and a truncated octahedron,[5] have been constructed from simple branched junctions; the vertexes correspond to the branch points of branched junctions, and the edges are DNA double helices that connect these vertexes. The cube was built by combining squares in solution, and the truncated octahedron entailed the use of a solid support methodology;[6] both syntheses contained a number of steps. These Platonic and Archimedean solids are regular graphs, all of whose vertexes have the same degree of connectivity, three. From the computational point of view, it is important to be able to construct irregular graphs, molecules whose vertexes have varying degrees of connectivity. From the perspective of computation, it is most desirable to construct these graphs by a single step of self-assembly, because, in this way, the power of parallel computation using DNA can be maximized and the number of computational steps will not depend on the size of the graph.

It has been proposed that through the use of three-dimensional graph structures achieved by DNA self-assembly, it is possible to solve a number of NP-complete problems with a constant number of steps. Algorithms for solving the Hamiltonian Path problem,[7] the three-vertex colorability problem,[8] and the three-SAT problem[9] have been described. In contrast to many algorithms proposed for a large class of NP-complete problems where DNA molecules are treated as linear strings, the general idea here is to encode a problem in branched DNA molecules that represent the vertexes of a graph encoding a solution. A graph corresponding to a solution to the problem may then be obtained, through the assembly of a set of vertex and edge building blocks, which represent the variable inputs and the rules they must obey to generate the output. It can be shown that the whole graph can be constructed from DNA, if and only if a solution to the problem exists. Consequently, to establish that a solution exists for a particular problem, one need only verify that DNA graphs corresponding to the solution have actually resulted from the assembly. The details of the solution can be analyzed by conventional DNA analytical techniques, such as sequencing or restriction analysis. Encoding rules and actual graphs vary according to the nature of the problem. Nevertheless,

[†] New York University.
[‡] University of South Florida.
[§] Current Address: Department of Biochemistry, Faculty of Science, Chulalongkorn University, Phyathai Road, Patumwan, Bangkok Thailand 10330.

(1) Seeman, N. C. *J. Theor. Biol.* **1982**, *99*, 237−247.
(2) Seeman, N. C. *Nature* **2003**, *421*, 427−433.
(3) Qiu, H.; Dewan, J. C.; Seeman, N. C. *J. Mol. Biol.* **1997**, *267*, 881−898.
(4) Chen, J.; Seeman, N. C. *Nature* **1991**, *350*, 631−633.
(5) Zhang, Y.; Seeman, N. C. *J. Am. Chem. Soc.* **1994**, *116*, 1661−1669.
(6) Zhang, Y.; Seeman, N. C. *J. Am. Chem. Soc.* **1992**, *114*, 2656−2663.
(7) Jonoska, N.; Karl, S. A.; Saito, M. In *DNA Computers III*; Rubin, H., Wood, D., Eds.; American Mathematical Society: Providence, RI, 1999; pp 123−135.
(8) Jonoska, N.; Karl, S. A.; Saito, M. *Biosystems* **1999**, *52*, 143−153.
(9) Jonoska, N.; Sa-Ardyen, P.; Seeman, N. C. *J. Genetic Programming And Evolvable Machines* **2003**, *4*, 123−137.

10.1021/ja049953d CCC: $27.50

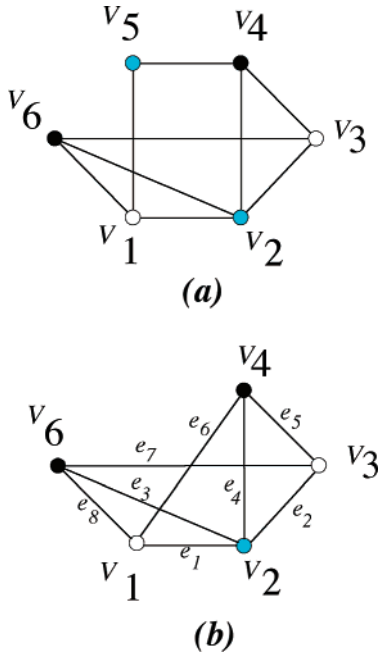


Figure 1. A Graph with three-coloring and its prototype for solving a three-vertex colorability problem. (a) A graph is three-colorable if it is possible to assign one color to each vertex such that no two adjacent vertexes (those that share the same edge) are colored with the same color. A vertex that has degree k is constructed using a k-armed branched DNA molecule. In this example, one two-armed branched molecule, four three-armed branched molecules, and one four-armed branched molecule are needed. (b) The same graph was chosen for the construction. Since the vertex V5 in panel a has degree 2, for the experiment a double helical DNA is used to represent the vertex V5 and the edges connecting V5 with V1 and V4. The target graph to be made consists of five vertexes and eight edges.

since the presence of a solution is predicated on the construction-cum-proof concept, the number of experimental steps required to obtain the solution for the problem need not increase as the problem size increases. Thus, in principle, the method scales very favorably with increasing input size of the problem.

To illustrate the approach, we have prototyped an example of the algorithm for solving the three-vertex colorability. A graph is said to be three-vertex colorable if three colors can be assigned to its vertexes so that no two adjacent vertexes (connected by an edge) have the same color. An example of a three-vertex colorable graph is shown is Figure 1a. Such a colorability requirement may be translated from the realm of graph theory to DNA nanotechnology by taking the following steps in designing the building blocks:

(1) For each vertex building block, all the arms connecting this vertex to any given edge must bear identical colors. A vertex of degree k is represented by a k-armed DNA branched junction molecule. The end of each arm is terminated by a single-stranded extension ("sticky end") that serves two functions: (a) to specify the connection between the vertex and an incident edge and (b) to encode for the color. For each vertex, three different vertex building blocks, corresponding to the three possible colors, are used for the self-assembly. Since the way any given vertex is connected to a certain edge follows the graph in question, the sticky section corresponding to this part is the same for all the building blocks of a vertex; the sticky part that specifies the vertex color differs from one type of building block to another.

(2) For each edge that connects two vertexes, a conventional double helical DNA molecule is used. Both ends of the DNA

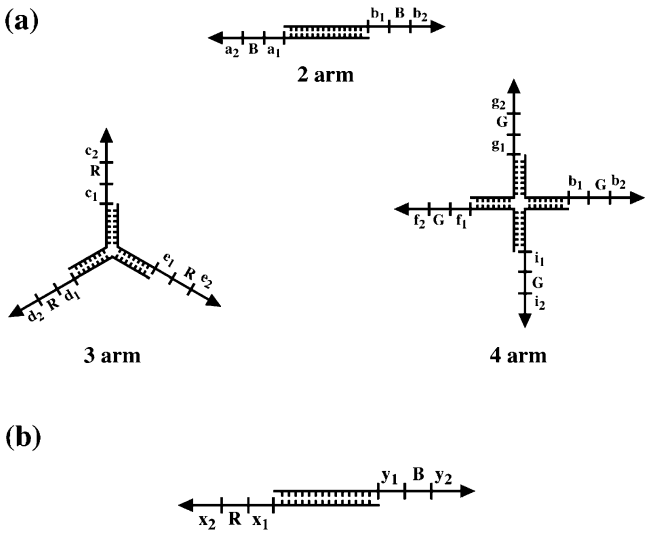


Figure 2. Schematic drawings of vertex and edge building blocks in self-assembling a three-colorable graph. (a) A single-stranded DNA strand is represented by a line terminated by an arrowhead, which indicates the 3′ end. Three different multiarm junction molecules are depicted that represent vertexes of degree 2, 3, and 4. Double helical DNA can be regarded as a two-arm branched junction in this context. The 3′ ends of the junction molecules are terminated by single-stranded extensions or sticky ends. For use in a computation, these extensions will consist of three parts. The first and the third parts, for example, x_1 and x_2 where $x \in \{a,...,i\}$, are specific encodings for the edge that is represented by the given arm of the molecule. The middle part of the encoding is the same for all arms of the vertex molecule and represents the color of the vertex. For each vertex, three molecules are needed, each corresponding to one of the three possible colors of the vertex, indicated here by three alphabets: R, B, and G. (b) The edges are regular double-stranded DNA molecules. The 3′ ends of an edge molecule are terminated by the single-stranded extensions that are complementary to the corresponding arms of the vertex that is incident to the edge. Hence, the first and third part of the encodings at the 3′ end are $\sim x_1$ and $\sim x_2$ which are Watson−Crick complementary to x_1 and x_2 ($x \in \{a,...,i\}$) of the corresponding arm of the incident vertex molecule. The middle part of the encoding segment is complementary to the color of the incident vertex, but here, the color sequence at one 3′ end is different from the color sequence at the other 3′ end. For each edge, six double-stranded molecules are needed, each representing a pair of distinct colors at the endpoints of the edge molecule.

are terminated by a single-stranded extension, as in the vertex. Similar to its counterpart in the vertexes, a section of this extension encodes the pairing rule, which is represented by a sequence that is complementary to the corresponding arm of the vertex incident to the edge. The remaining part of the extension is complementary to the color code of the incident vertex. The key is to insist that the two ends of any given edge bear different colors. For each edge, six different double helical DNA molecules are required, corresponding to the number of differently ended edges one can make with three colors. An example of vertex and edge building blocks is shown in Figure 2.

(3) To form the graph, all edge molecules and all vertex molecules are combined and their compatible ends are allowed to cohere. Following this self-assembly step, the nicks are sealed in a reaction catalyzed by T4 DNA ligase.

To solve the colorability problem, all one must do is determine whether the DNA graph has formed. The graph molecules will be topologically closed, with all the nicks sealed by ligase, so that two-dimensional denaturing gel electrophoresis (e.g., ref 10) can be employed to remove all partially formed graphs, which will correspond to linear molecules. Depending

on the conditions used, it is possible to obtain graphs of the expected size, thus excluding potential contamination by graphs having higher molecular weight. The latter must be removed because they can erroneously indicate that a solution exists where none does.

The approach described above provides a good starting point for assessing the potential of computing with DNA graphs. Based on the scheme described for the instance of three-vertex colorable graph, we set out to explore the feasibility of using the self-assembly of linear and branched DNA molecules to construct a graph with comparable complexity, as shown in Figure 2. We have not performed the calculation described above, but we have demonstrated its feasibility by building a "monochrome" graph, where all of the vertexes and all of the edges are of the same color, so that they do not have to compete to assemble into a graph structure represented by a cyclic molecule.

Following self-assembly and ligation, the target product was a single-stranded circle or knot, indicating that the target was successfully made. We have applied restriction enzyme digestion systematically to this product. The restriction analysis was used to establish the order of the components. This method has demonstrated that the graph obtained conforms to the design of the correct solution.

Materials and Methods

Design, Synthesis, and Purification of DNA: The sequences for each of the strands were designed by use of the program *SEQUIN*.[11] The sticky ends were designed by *DNA Sequence Generator*.[12] Special care was taken so that no sticky-end sequence appears in the whole final product more than once. In the design of the four-arm junction molecule that represents V2, the first eight-base sequence of the arms was taken from the well-characterized J1 molecule.[13] DNA strands containing biotin-dT (Glen Research) units or additional hairpins were synthesized on an Applied Biosystems 380B automatic DNA synthesizer, removed from the support, and deprotected, using routine phosphoramidite procedures.[14] All other strands were purchased from Integrated DNA Technologies. DNA strands were purified by denaturing polyacrylamide gel electrophoresis containing 8−12% acrylamide (19:1 acrylamide/bisacrylamide). Bands corresponding to DNA strands of expected size were excised from denaturing gels stained with ethidium bromide. DNA was eluted in a solution containing 500 mM ammonium acetate, 10 mM magnesium acetate, and 1 mM EDTA at 4 °C overnight. The eluates were then extracted with butanol to remove ethidium, and the DNA was recovered by ethanol precipitation. The amount of DNA was estimated by OD_{260}.

Electrophoresis: See the Supporting Information.

Phosphorylation: See the Supporting Information.

Assembly and Ligation: The following annealed and phosphorylated subunits V1, V2, V4, E3, E4, E6, and E8 were first mixed together, and the reaction mixture was brought to 40 °C for 20 min. The mixture was then allowed to cool slowly to 37 °C, at which point subunits V3, V5, and E7 were added, and incubation at 37 °C proceeded for another 20 min. The remaining subunits E1, E2, and E5 were added after the mixture had cooled further to 20 °C. 270 units of T4 DNA ligase (NEB) were added directly to the mixture, and the ligation reaction was allowed to proceed at 16 °C for up to 48 h. The reaction mixture was then extracted by phenol/chloroform, precipitated by ethanol, and resuspended in 60 μL of 1× ligase buffer (NEB). Another round of ligation was performed by incubating the solution with another 10−45 units of fresh ligase (NEB) at 16 °C for another 16 h. When further removal of unincorporated radioactive ATP was desirable, the reaction mixture was filtered through a G-25 micro-spin column (Amersham Pharmacia Biotech) prior to subsequent purification.

Two-Dimensional Denaturing Polyacrylamide Gel Electrophoresis: These gels were Fischer-Lerman[14] 100% denaturing gels. The first dimension contained 4% acrylamide (37.5:1 acrylamide/bisacrylamide), and the second dimension contained 5.5% acrylamide (37.5:1 acrylamide/bisacrylamide); both were run at 55 °C. After electrophoresis, the gel was exposed to a Kodak X-OMAT AR film (Fisher Scientific) for up to 16 h to obtain the autoradiogram. Bands corresponding to topologically closed molecules appeared on a second arc above the one corresponding to linear species. Normally two well-defined bands were visible on the second arc. The gel area containing these bands was excised and mixed with a solution containing 500 mM ammonium acetate, 10 mM magnesium acetate, and 1 mM EDTA to elute the target samples at 4 °C overnight. The purified samples were then recovered by precipitation from ethanol.

Characterization with Topoisomerase I: About 5 fmol each of two distinct DNA graphs that were purified independently from two distinct locations along the upper arc of the two-dimensional gel were incubated with 5−15 units of topoisomerase I. A 50-μL reaction buffer contained 50 mM Tris·HCl, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.5 mM DTT, and 30 μg/mL BSA. Incubation proceeded at 37 °C overnight. The reaction was stopped by phenol/chloroform extraction. The samples were recovered by ethanol precipitation and run along with untreated control materials (kept at room temperature and at 37 °C) on a Fischer-Lerman 100% denaturing gel containing 6% acrylamide (37.5:1 acrylamide/bisacrylamide) at 55 °C. The gel was dried onto Whatman 3MM paper and scanned using a Storm 860 Gel and Blot Imaging System.

Restriction Endonuclease Mapping: There are four different restriction endonuclease sites: Pst I, EcoR I, Xho I, and Pvu II, designed to be located at each of the four hairpins on the four long edges of the graph. All restriction enzymes were purchased and supplied with corresponding reaction buffers from NEB. Digestions were run in buffers recommended by the supplier. To perform a series of digestions, the buffers had to be prioritized. EcoR I buffer received first priority; in the absence of EcoR I, Pst I buffer was used. When neither EcoR I nor Pst I was present, the recommended Xho I buffer was used.

A. Preliminary Characterization with Restriction Endonucleases: About 10 fmol of purified cyclic graphs were incubated with 20−40 units of the appropriate restriction enzyme in a 50-μL reaction buffer recommended for each enzyme at 37 °C for 1 h. A 30-μL aliquot of the reaction mixture was then withdrawn and set aside for exonuclease digestion, where 100 units of exonuclease III (USB) and 15 units of exonuclease I (USB) were added directly and incubation proceeded for 1 h. All reactions were stopped by phenol/chloroform extraction, followed by ethanol precipitation. Recovered samples were run on a Fischer-Lerman 100% denaturing gel containing 5.5% acrylamide (37.5:1 acrylamide/bisacrylamide) at 55 °C. After electrophoresis, the gel was exposed to a Kodak X-OMAT AR film (Fisher Scientific) for up to 16 h to obtain the autoradiogram.

B. Preliminary Double Digestion to Generate Restriction Fragments: 40 fmol of purified cyclic graph molecules were incubated with 10 units each of the two enzymes in a 50 μL reaction volume at 37 °C for 2 h. The reaction was stopped by ethanol precipitation, and the sample was dried and purified on a Fischer-Lerman 100% denaturing gel containing 5.5% acrylamide (37.5:1 acrylamide/bisacrylamide). The autoradiogram of the gel was used to locate and excise the bands corresponding to desired fragments. A gel slab containing long DNA fragments were further chopped into smaller pieces, mixed with 500 μL of elution buffer (500 mM ammonium acetate, 10 mM magnesium

(10) Wang, H.; Di Gate, R. J.; Seeman, N. C. *Proc. Nat. Acad. Sci. U.S.A.* **1996**, *93*, 9477−9482.
(11) Seeman, N. C. *J. Biomol. Struct. Dyn.* **1990**, *8*, 573−581.
(12) Feldkamp, U.; Saghafi, S.; Rauhe, H. In *DNA Computers VII*; Jonoska, N., Seeman, N. C., Eds.; *Lec. Notes Comput. Sci.* 2340; Springer: Berlin, 2002; pp 23−32.
(13) Seeman, N. C.; Kallenbach, N. R. *Biophys. J.* **1983**, *44*, 201−209.
(14) Caruthers, M. H. *Science* **1985**, *230*, 281−285.

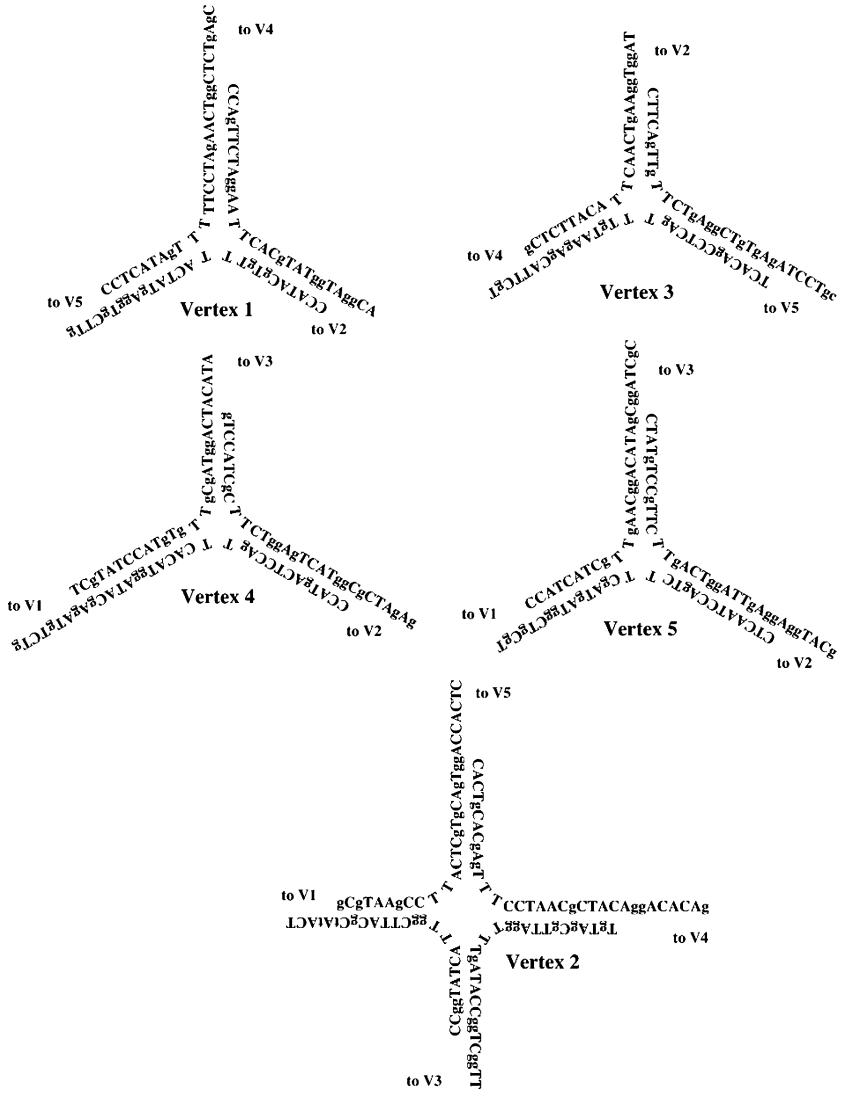


Figure 3. Sequences of the vertexes. The five vertexes, V1, V2, V3, V4, and V5, are shown. A strand of DNA is represented by a string of nucleotide sequences: A, T, g, and C ("g" is used to make it easier to distinguish from "C"). The 5′ end of a strand is located at the terminus of the single-stranded extension of an arm. Each component strand of a vertex can be uniquely identified by associating the name of the vertex where such a strand is located, with the name of two adjacent vertexes. For instance, three strands that make up V1 are V1-4-5, V1-5-2, and V1-2-4. For a vertex, the names of neighboring vertexes are indicated at the termini of each arm. Note that bulges of two T's were added to the branch point region to increase flexibility.

acetate, and 1 mM EDTA), and distributed equally into four 200-μL PCR tubes. Elution was performed in a thermo-cycler programmed to heat alternately at 60 °C for 5 min and at 37 °C for 20 min for 20−30 cycles. Elution of short fragments was done in 1.7-mL microcentrifuge tubes at 4 °C overnight.

C. Single- and Double-Digestion of Restriction Fragments: For single-enzyme mapping, 2 fmol of purified restriction fragments were incubated with 1 unit of enzyme in 25 μL of reaction buffer. For the double-enzyme mapping, 6 fmol of purified restriction fragments and 2 units of each of the two enzymes were incubated in a 50-μL reaction buffer. In both cases, incubation proceeded at 37 °C for 2 h. After digestion, when no further treatment was required, the samples were dried for subsequent analysis on a Fischer-Lerman 100% denaturing gel containing 6% acrylamide (37.5:1 acrylamide/bisacrylamide). The gel was dried onto Whatman 3MM paper and scanned using a Storm 860 Gel and Blot Imaging System.

D. Streptavidin Bead Purification: An additional step was needed for the doubly digested products, to remove undigested and partially digested materials. This was achieved by streptavidin-coated magnetic beads (Promega) which bind specifically to the loops of the restrictable hairpins, where two biotin-modified dT nucleotides are present in each loop. The beads were supplied as 1 mg/mL magnetic particles in a phosphate buffer. Based on the binding capacity suggested by the supplier, 0.75−1.25 nmol/mg, each microliter of bead solution is capable of binding 0.75−1.25 pmol of singly biotinylated DNA substrates. A 30-μL aliquot of a 50-μL reaction was withdrawn after a 2-h double digestion; it was treated with 50 μL of the bead solution. Following rinsing and equilibration with the corresponding digestion buffer, the beads were incubated with the doubly restricted DNA fragments at room temperature for 10 min. The physical separation of the beads from the supernatant was done by setting the sample tube on a magnet stand for 3 min to immobilize the beads along the side of the tube. The supernatant that contained only hairpin-free DNA fragments was then carefully withdrawn and dried for subsequent analysis.

E. Buffer Exchange for EcoRI-Treated Samples: It was found that Triton X-100, the detergent component present in the EcoRI reaction buffer decreased the quality of the subsequent denaturing gel analysis, by causing severe smears in all the lanes treated with the buffer. This problem was overcome by exchanging regular EcoRI buffer with a modified EcoRI buffer where Triton X-100 was omitted. The EcoRI treated samples in the conventional buffer were first diluted with the modified buffer to make a total volume of 500 μL and then spun through Microcon-3 columns (Millipore) at 14 000× g until the final

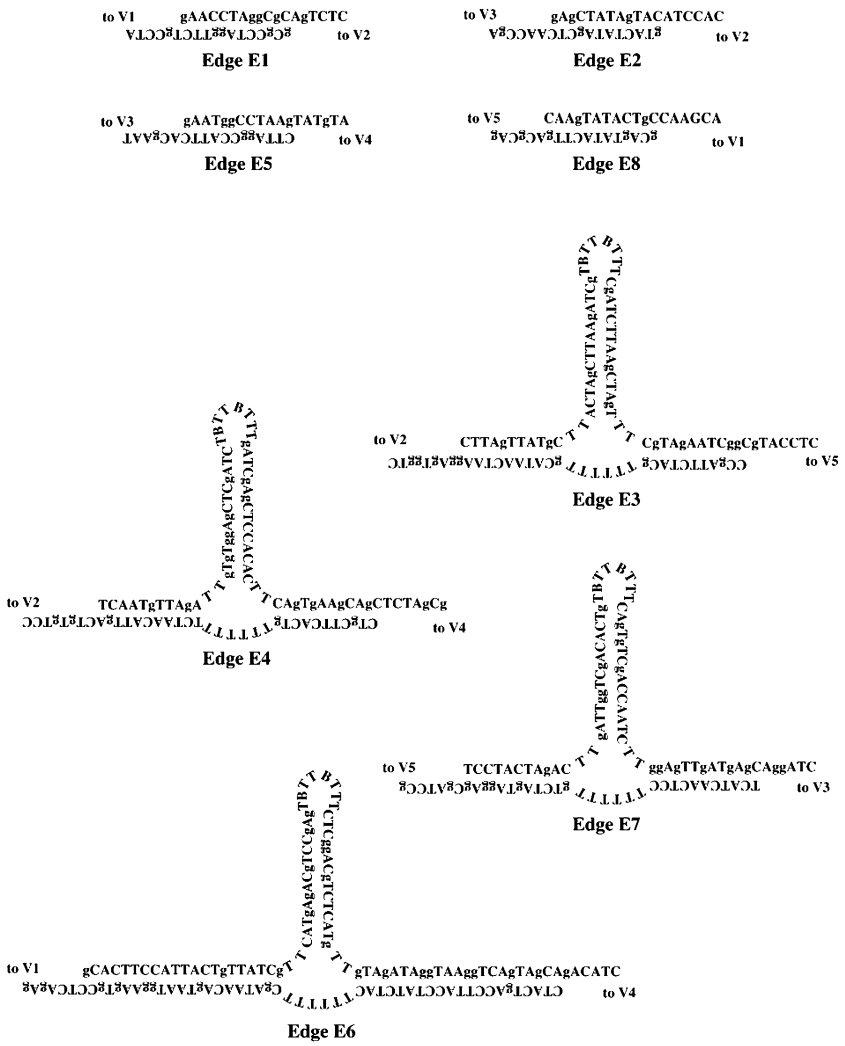


Figure 4. Sequences of the Edges. The four short edges, E1, E2, E5, and E8, as well as the long edges, E3, E4, E6, and E7, are shown. A component strand of an edge can be identified in the same way as the strands forming a vertex. For an edge, the names of the two connecting vertexes are shown at each end of the molecule. Note the bulges of two T's (on the hairpin strands) and six T's (on the opposite strands) flanking the branch point where a hairpin is fused to the middle part of each long arm for increased flexibility. In this drawing, recognition sites for restriction enzymes located on each hairpin are not shown explicitly.

volume was reduced to about 10 μL. Two rounds of dilution and spinning were usually sufficient to complete the buffer exchange.

Results

Design of the Graph. The target graph to be constructed in this study is the same as the one used as the example in the algorithms proposed by Jonoska et al. to solve the Hamiltonian path problem[7] and the three-vertex colorability problem.[8] Both graphs are shown in Figure 1. Since the vertex V5 in part a has degree two, V5 and the edges connecting V5 with V1 and V4 are replaced by one edge as depicted in part b. The graph contains five vertexes and eight edges. Three-arm junctions were used to make V1, V3, V4, and V5. The vertex V2, which has degree four, is represented by a four-arm branched junction molecule. To increase flexibility around the branch point at each vertex, tracts of dT's ranging from 2 to 6 nucleotides were placed on each strand where it flanks a junction.

A linear duplex DNA molecule was used to represent an edge connecting two vertexes. The distance between any two vertexes was chosen to be an integral multiple of helical repeat to avoid torsional stress insofar as possible, while maximizing the flexibility of the connection to allow successful assembly of the graph. The lengths of the edges were modeled using physical models.[16] The range of lengths is from 4 helical turns or 42 base pairs for the short edges (E1, E2, E5, and E8) to 6 helical turns for the medium ones (E3, E4, E7) and 8 helical turns for E6. At the midsection of each long edge, a biotinylated hairpin (15-nucleotide stem) containing a restriction site has been fused[17] to the double helix to form a three-arm junction. As employed previously,[18] the addition of these hairpins does not affect the fundamental topology of the graph, but the feature can be useful both for synthesis and for analysis of the product. Here, the additional hairpins serve three purposes: First its extra dT's at the junction are expected to confer added flexibility to the edge. Second the hairpin provides a restriction site for subsequent characterization. Finally, biotin-modified bases in the loops allow removal of incomplete digestion products when the sample is treated with streptavidin magnetic beads. The sequences of the DNA strands used for constructing the vertexes are shown in Figure 3, and those for each edge are shown in Figure 4.

(15) Fischer, S. G.; Lerman, L. S. *Cell* **1979**, *16*, 191−200.
(16) Seeman, N. C. *J. Biomol. Struct. Dyn.* **1988**, *5*, 997−1004.
(17) Seeman, N. C. *NanoLett.* **2001**, *1*, 22−26.
(18) Mao, C.; Sun, W.; Seeman, N. C. *Nature* **1997**, *386*, 137−138.

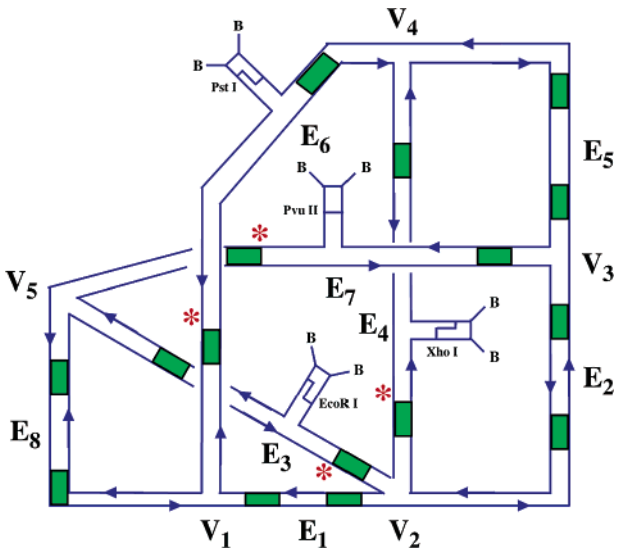


Figure 5. Schematic drawing of the prototype DNA graph. One possible structure of the target DNA graph to be constructed is shown. Small half-arrowheads indicate 3′ ends of strands. The shaded green areas represent the position of the sticky ends. Red asterisks indicate the positions of labeling with ^{32}P. All vertexes and edges are indicated. Restriction sites and the loci of biotin groups are indicated. Note that the final graph that forms results in one single-stranded circle or knot.

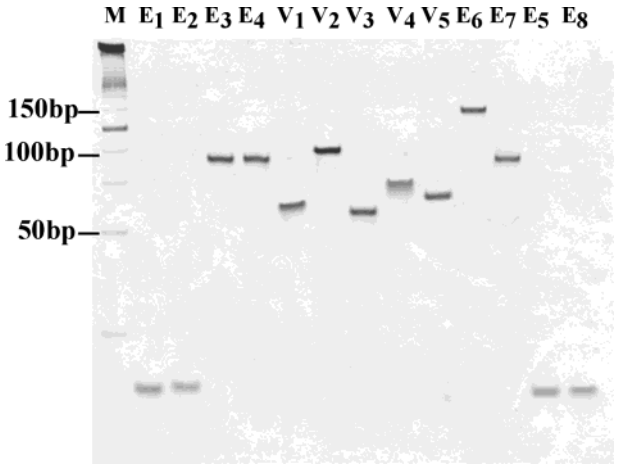


Figure 6. Nondenaturing gel characterization of the self-assembled vertex and edge subunits. This is a 12% nondenaturing gel that was run at 30 °C. Lane 1 contains linear markers. Lane 2 to 14 contain the annealed vertexes and edges. The names of the species characterized are indicated on top of each lane. Note that vertexes and edges that contain branched junctions will migrate with an apparent mobility much lower than would be expected if the same number of nucleotide pairs were in linear duplex molecules.

Although the ultimate computation would likely require more information, for this prototypical assembly sticky ends of 6−8 bases are used to direct the cohesion of vertexes and edges to their designated partners. The final graph is designed to be a continuous single-stranded circle or knot when all the strands present in the building blocks are ligated to covalency. The final graph is illustrated schematically in Figure 5.

Assembly of Individual Vertexes and Edges. All the vertexes and edges were first assembled individually from their component strands by cooling slowly from 90 °C to the desired temperature. The formation of each vertex and edge was then characterized on a nondenaturing gel. The characterization is shown in Figure 6. Discrete bands were obtained for most of the assembled vertexes and edges in a gel run at 22 °C, indicating that the molecules formed well (not shown). In case of V1 and V4, the bands were less discrete, showing slightly different degrees of smearing. The smearing found for both vertexes was significantly reduced in Figure 6 where the characterization temperature was raised to 30 °C, while at 40 °C the smearing completely disappeared (not shown).

Self-Assembly of the Graph. The strategy used to construct the graph entailed the following steps: (1) assemble individual vertexes and edges; (2) phosphorylate 5′ ends of every strand to facilitate covalent-joining via ligation, where a set of strands in certain vertexes were chosen for radioactive labeling to aid purification and subsequent characterizations; (3) ligate all the nicks in the whole graph that resulted when all vertexes and edges were assembled.

There are two approaches for self-assembly of the graph. The self-assembly may be restricted so that certain assembly reactions can proceed only after others have been completed (stepwise or serial self-assembly), e.g., ref 6. Alternatively, self-assembly reactions may be limited by no such restrictions (free self-assembly). Serial self-assembly can decrease the possibilities of ′cross-talk′ among components, but it violates the spirit of the one-pot reaction that is desirable for computational ends. Given the self-association noted for V4, we have adopted a modified free self-assembly protocol here. We allowed the sticky ends with a higher propensity for self-association to be assembled with their designated partners, thereby grouping subunits for the annealing protocol indicated in materials and methods. Nevertheless, we did try the one-pot reaction and obtained cyclic molecules of the same length from it.

As noted above, the product is designed to have the topology of a circle or knot. Evidence that the cyclic molecules were obtained was derived by characterizing the ligation products on two-dimensional denaturing gels. It is key that the complex secondary structure of the graph molecules be completely denatured in all analyses of the graph, either in full-length or in parts, for unambiguous comparison of the bands of interest with linear standard markers. A typical characterization result is shown in the inset to Figure 7. The products can be seen on two different arcs, one arc (virtually linear) corresponding to linear molecules and a second arc corresponding to cyclic molecules. The most prominent band in the linear arc corresponds to a linear molecule whose length is very close to that expected for a linear (nicked) version of the graph.

There are two prominent bands in the second arc, corresponding to cyclic molecules. Each band, designated G1 and G2, was recovered independently from the gel. By subjecting these species to moderate heating, it is possible to induce random single nicking that converts them into linear strands of the same length. By comparing the length of the linearized product with standard linear markers, we obtained strong evidence that both G1 and G2 are cyclic ligation products of the expected size of 1084 nucleotides, equivalent to the length of the linear species on the inset gel. There are all together 32 sites in the whole assembly that require ligation reaction to seal the nicks. The estimated yield for a typical ligation reaction was obtained by quantitating the amount of cyclic products and all remaining incomplete linear ligation products. By using the protocol described in methods, the ligation yield of about 1% or lower was obtained for the entire graph molecule; assuming 1% total yield, this corresponds to ligation step yields of about 87%.

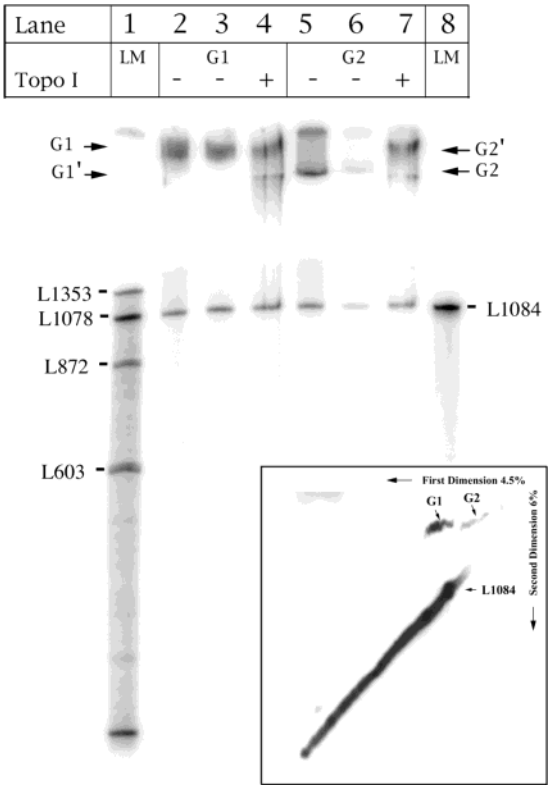


Figure 7. Topoisomerase I characterization of DNA graph molecules isolated from a two-dimensional denaturing gel. Two of the topologically closed ligation species G1 and G2 were purified separately from two distinct off-diagonal bands on a two-dimensional Fischer-Lerman 100% denaturing gel (inset) and treated with *E. coli* topoisomerase. Lane 1 contains linear markers. Lanes 2−4 contain G1, and lanes 5−7 contain G2. Lanes 2 and 3 contain untreated controls that were stored at −20 °C (lane 2) or incubated in buffer at 37 °C overnight without topoisomerase (lane 3). Lane 4 contains the material treated with topoisomerase I. Incubating materials in lane 2 or 3 with topoisomerase I followed by phenol extraction resulted in an extra band characteristic of the untreated materials of the other species as seen in lane 4. Similarly, lane 7 shows that topoisomerase I treatment of materials in lane 5 or 6 produced an extra band that ran with the same mobility as the untreated band in lanes 2 and 3. Thus, G1 and G2 are likely topoisomers of each other. The upper band in lane 5 corresponds to the well.

Topoisomers. Although it will have no impact on the final answer, the topological state of the product is always of concern when producing knotted DNA constructs.[19] *E. coli* topoisomerase I was used to analyze the two major cyclic graph molecules obtained from the two-dimensional denaturing gels, bands G1 and G2 (Figure 7). *E. coli* topoisomerase I catalyzes single-strand passage reactions that require no extra energy source, so that the products of the reaction approach the equilibrium distribution. When G1 was treated, a new band with the mobility of G2 was obtained; similarly, a new band with the mobility of G1 was obtained when G2 was treated. Thus, the two species are likely to be topoisomers of each other. In both experiments, the final ratio between the G1 and G2 is similar, with G1 in greater abundance, suggesting that G1 is the favored topoisomer. As noted in the preceding section, a band corresponding to linearized breakdown products can be seen in all the lanes that contain purified cyclic graphs.

Restriction Analysis. The proof that the final product has been assembled as designed utilizes the restriction approach employed to demonstrate a previous DNA covalent assembly,

(19) Du, S. M.; Stollar, B. D.; Seeman, N. C. *J. Am. Chem. Soc.* **1995**, *117*, 1194−1200.

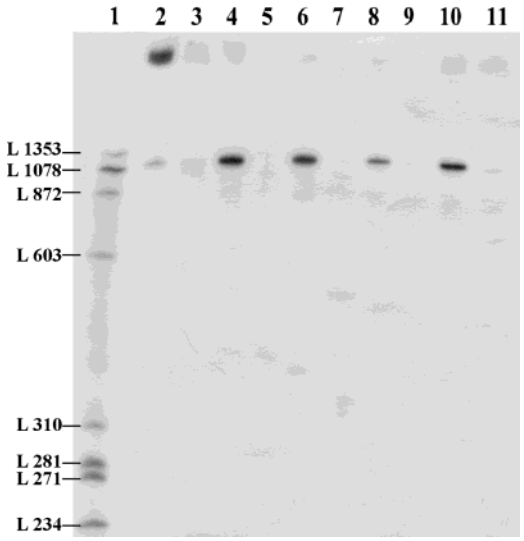


Figure 8. Individual cleavages of the graph molecule by four restriction enzymes. This is an autoradiogram of a denaturing gel showing characterization of the product by restricting it individually by each of the four enzymes. The first section (lanes 2−3) contains the control molecules, with untreated materials shown in lane 2 and the exonuclease III treated materials shown in lane 3. Overdigestion by exonuclease III leading to the significant decrease in the amount of the original products is seen in lane 3. In the following sections, the results of restricting the product with the following restriction enzymes are shown: PvuII (lane 4); EcoRI (lane 6); XhoI (lane 8); and PstI (lane 10). Lanes 5, 7, 9, and 11 contain the products of digesting aliquots of the restricted materials with exonuclease III. It can be seen that singly restricting the product gives rise to a linear strand of expected length.

Borromean rings:[18] This method entails the insertion of specific restriction sites on hairpins whose presence does not affect the topology of the molecular graph. The digestion products from the correct assembly are well-defined linear molecules, corresponding to a series of bands of particular lengths. The covalent sequence of the target molecule is a single cyclic molecule that traces the path of the graph twice. The double coverage of the graph makes it impossible to sequence the molecule by methods involving the polymerase chain reaction, because it is not possible to denature it well enough for the polymerase to get through it. As noted above, four of the edges have been constructed by using "T-shaped" molecules (Figure 4); these molecules, edges E3, E4, E6, and E7, are three-arm branched junctions that contain a hairpin on the external arm that is not an edge of the graph. We have placed four different restriction sites on these external arms. These sites, EcoR I, Xho I, Pst I, and Pvu II, respectively, enable us to cut the graph at specific sites and demonstrate the proper sequence of the molecule. Note that digestion at a single site serves only to remove a small piece of the hairpin, linearizing the molecule. Digestion of the four single sites individually produces a band of approximately the same length, differing only by the site of the restriction site within the hairpin (Figure 8).

To demonstrate the correct order of assembly, we have used two adjacent restriction enzymes to cleave the cyclic single strand. This produces a fragment lacking any further restriction sites and a second fragment that contains both of the remaining sites. This second fragment is cleaved by the two remaining enzymes, and the products are sized by electrophoresis on a Fischer-Lerman 100% denaturing gel.[15] The protocol is illustrated at the bottom of each panel of Figure 9. This analysis has been repeated for all four successive enzyme pairs. The resulting products in all cases confirm that the molecule

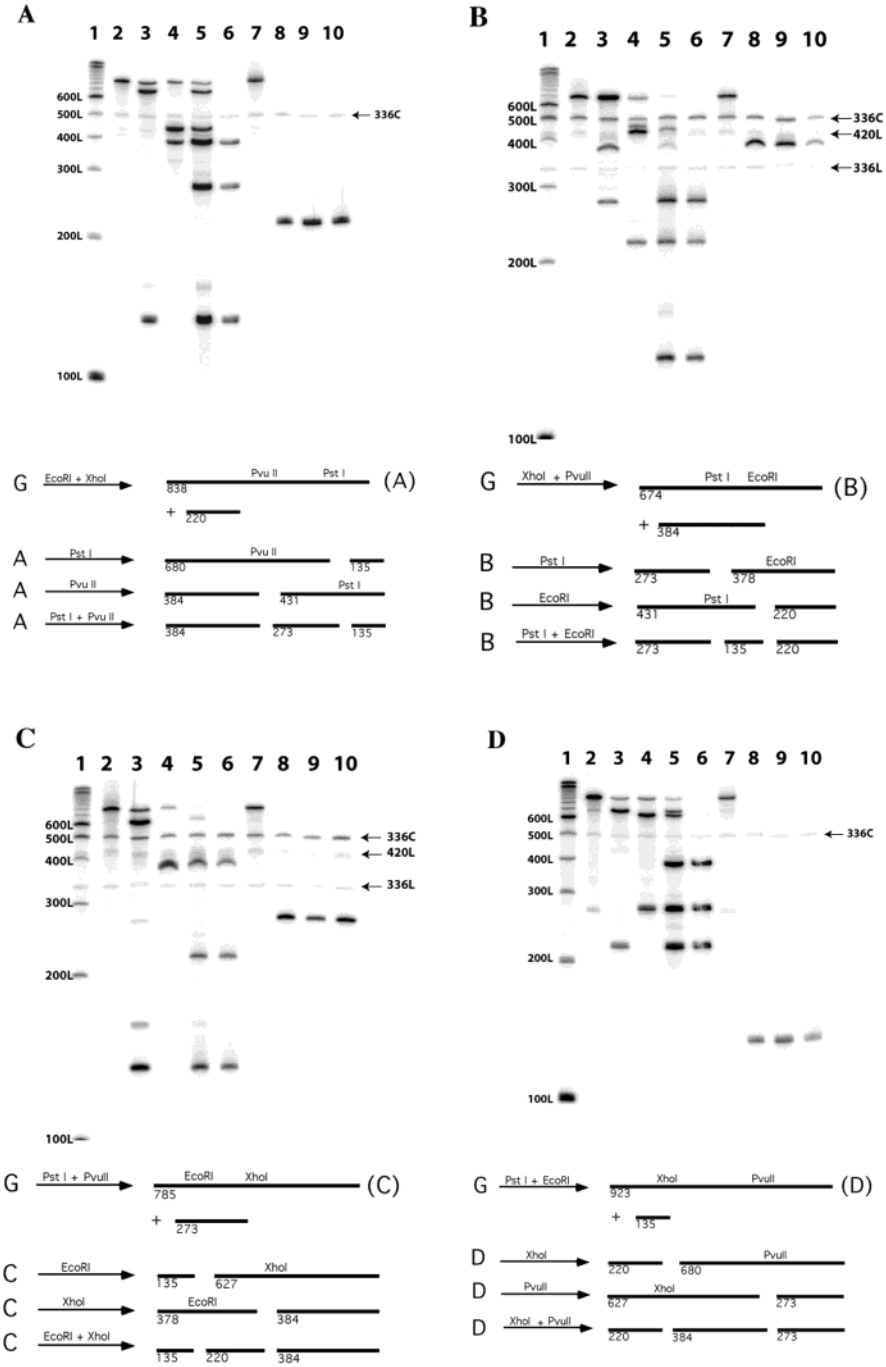


Figure 9. Restriction mapping analysis. These are the autoradiograms of Fischer-Lerman 100% denaturing gels that show the restriction mapping analysis. Shown beneath each gel is a diagram outlining the steps for obtaining the restriction fragments, as well as the following restriction steps. The initial graph is indicated by "G", and the long fragments containing further restriction sites are indicated by the letter of the figure panel. The expected size of each species is indicated, as is the location of each restriction site along the fragment. (a) The result of the analysis using the enzymes Pst I and Pvu II. Lane 1 contains a ladder of linear 100 nt markers. The first section (lane 2−7) contains the restriction fragments obtained from simultaneously digesting the graph with two enzymes: EcoRI and XhoI. The expected size of this fragment is 838 nt. The second section (lane 8−10) contains the shorter restriction products after the double-digestion by the same set of enzymes. The length of this fragment is expected to be 220 nt. Both 838 nt and 220 nt were first denaturing gel purified prior to the restriction analysis shown in this gel. Lanes 2 and 8 contain untreated controls. Lane 3 contains the material digested with PstI. Lane 4 contains the products of single digestion by PvuII. Lanes 4, 5, and 9 contain the products of treating with both enzymes simultaneously. Lane 6 contain the doubly digested products treated with streptavidin beads. Materials treated with EcoRI and XhoI as controls to show that no sites for these enzymes are present are shown in lanes 7 and 10. A cyclic 336 nt was added to each lane to indicate lack of endonuclease contamination. (b) Similar analysis using the enzymes PstI and EcoRI on the fragments (674 nt for the long and 384 nt for the short ones) generated by first double digesting the graph with XhoI and PvuII. There are three types of internal markers in this case: cyclic 336 nt, its linear counterpart and linear 420L; they are indicated in the same way as in (a). (c) Similar analysis using the enzymes EcoRI and XhoI on the fragments (785 nt for the long and 273 nt for the short ones) generated by first double digesting the graph with PstI and PvuII. (d) Similar analysis using the two enzymes XhoI and PvuII on the fragments (923 nt for the long and 135 nt for the short ones) generated by first double digesting the graph with PstI and EcoRI.

assembled is the target molecule. Biotinylation of the hairpins enabled the use of streptavidin beads to remove partially digested molecules.

The data are shown in Figure 9. Panel D illustrates the analysis of the material following cleavage of the cyclic graph by Pst I and EcoR I. These two enzymes leave a 135 nt fragment

and a second fragment containing 923 nt, including the Xho I and Pvu II sites. Cleavage of the 923 nt fragment by Xho I should leave a 220 nt fragment and a 680 nt fragment containing the Pvu II site. This digestion is shown in lane 3. Likewise, digestion of the 923 nt fragment by Pvu II should leave a 223 fragment and a 627 fragment containing the Xho I site. This digestion is shown in lane 4. The double digestion by Xho I and Pvu II is shown in lane 5. Lane 6 shows the results of treating the material in lane 5 with streptavidin beads. Only the expected double digest products are seen, along with added cyclic markers to demonstrate the lack of random endonuclease activity. Lane 7 contains a control experiment indicating no further Pst I and EcoR I sites in the 923 nt fragment; lane 8 contains the 135 nt fragment, and lane 9 and lane 10 are controls indicating that there are no further restriction sites in the 135-mer fragment.

Figure 9A−C are similar to Figure 9D. All of the expected bands are there. There is an extraneous band in lane 3 of Figure 9C, of length ∼270 nt, and another extraneous band in lane 3 of Figure 9D, this one of length ∼430 nt. These extraneous bands constitute about 2% of the material present and may well indicate the inherent error rate of the assembly under the conditions used here.

Discussion

Novel Design Features. This construction differs from previous DNA nanoconstructions in a number of ways. With a well-defined structural target, such as a polyhedron, a device, or an array component, we usually try to maximize the rigidity of the target. By contrast, here we are interested only in maximizing the chances that the assembly of the irregular graph will not be inhibited by an overly stiff edge preventing two sticky ends from cohering. Consequently, we have maximized the flexibility of the molecule, so as to increase the ability of all sticky ends to find each other.

A second unusual feature of the design here is that explicit edges have been included in the design. In all previous constructions, edges were built out from the vertexes so that they made contact with other vertexes, and the edges were formed implicitly, without having particular molecules designated to make them. This feature was needed here, because of the nature of the assembly that we are prototyping, but it should be pointed out that we have effectively doubled the number of points where ligation is needed, relative to the previous types of synthesis.

The third way in which this assembly differs from previous assemblies is in its "single-pot" character. For nanotechnological purposes, successive steps of assembly (e.g., ref 5) provide greater amounts of control. However, the goal here is to get a single cyclic molecule from all the components in the reaction. It is remarkable that the yield was as high as it was, considering the number of ligation steps that needed to occur successfully. The limits on the computational potential of this approach are closely related with the detectability of a species corresponding to a solution. The intensity of the G1 band in the Figure 7 inset suggests that at this scale a band at least 1 order of magnitude weaker could be detected, indicating the presence of a solution. Larger scale assemblies using components of very high specific activities might be another way to increase the range of this technique.

The Molecule as the Solution to a Computational Problem. We have demonstrated that a molecular graph whose edges are represented by DNA helix axes can be made by self-assembly and ligation of branched DNA molecules and linear duplex DNA. This is a prototype system where only unique vertexes and edges were designed to form a graph and no algorithm was implemented for computation. For a practical system where computation can be executed by self-assembly, additional variation in the sticky ends of both the input edges and vertexes must be present in the reaction mixture. However, the self-assembly step would not be different, and once completed, the same approach indicated here can be applied to read out a unique solution. The presence of multiple solutions would likely require the separation of the molecules representing them before implementing the type of analysis used here.

The advantage of using a three-dimensional graph structure in computation is that if a solution to the problem exists, its presence is established directly by the presence of a properly sized cyclic molecule. Thus the same approach used in this study can be applied, in principle, to obtain the information about the existence of one or more solutions. Ultimately, the actual information regarding the nature of the solutions will likely have to be read out by direct sequencing of at least parts of the graph where the values are encoded. This might be feasible if graphs are designed to contain features that aid subsequent amplification by polymerase chain reaction. The intrinsic double cover structure of the graph built here prevents the use of PCR. For example, building edges with incomplete complementarity might prove helpful in this regard.

Topoisomers. Multiple topoisomers were obtained for the graphs made. This was revealed by topoisomerase analysis. When assembling a flexible cyclic molecule, it is possible that the product will be knotted. With enough flexibility in the molecule, it is likely that multiple knots will be produced with the same sequence. We have determined that at least two topoisomeric knots have been produced in the ligation procedure. The two knots have been linearized to produce the same strand. We have demonstrated previously that multiple knotted topoisomers can be produced from the same strand[19] and that multiple catenanes can be produced from the same design.[20] It is most likely that the topoisomers resulted from strand flexibility, rather from different folding of the graph backbone. The topoisomerase used produces single-strand passage events, and only two topoisomers were detected. Backbone topoisomers would require double-strand passage events, and one would expect an intermediate topoisomer in this system if backbones were being passed through each other one strand at a time. The presence of multiple topoisomers is, of course, only a preparative artifact, and it has no bearing on the validity of the solution.

Errors. Fidelity of the assembly of subunits into the graph depends heavily on the specificity of the sticky ends. We noted the smearing of the V4 band in nondenaturing gel analysis (indicating self-association) and took ad hoc measures to ensure that it did not interfere with the correct assembly. As noted above, this is a prototype assembly system where no computation was intended. For a real computation, such ad hoc measures are inappropriate, and a higher degree of fidelity will be

(20) Qi, J.; Li, X.; Yang, X.; Seeman, N. C. *J. Am. Chem. Soc.* **1996**, *118*, 6121−6130.

required: the correct subunits should have to compete only with incorrect and not partially correct subunits[21] in assembling the graph.

We have noted that one band, corresponding to about 2% of the material in the lane, is seen in Figure 9C and D. This is material for which we cannot account by any reasonable process, and we assume that it represents erroneous ligation that nevertheless leads to cyclic molecules of the appropriate size. We estimate that this material represents the inherent error rate of the system, at least under the conditions used in these experiments.

Another potential source of error for this approach when applied to NP-complete problems is the self-assembly byproducts known in graph theory as covering graphs. It has been shown that whenever a graph is constructed by DNA molecules representing vertexes and edges, a covering space of the graph is possible.[7] Since the vertex and edge building blocks connect locally and possibly build larger pieces globally, it is possible that the covering graphs, which can be regarded here as the result of multimerization of the original graph, will be obtained. In the case of the three-vertex colorability problem, covering graphs can lead to false positives, since it is possible that an original graph is not three-colorable but a covering graph is. Therefore, care must be taken in isolating the graphs that have been constructed, such that no covering graphs are misidentified and inadvertently included in the subsequent steps. Fortunately a combination of denaturing gel purification and heat-induced linearization can be employed to identify reliably the correct size of a target graph that is a single-stranded circle or knot or to reveal multilinked circles which result from the covering graph formation.

(21) Mao, C.; LaBean, T.; Reif, J. H.; Seeman, N. C. *Nature* **2000**, *407*, 493−496.

Acknowledgment. This research has been supported by Grants GM-29554 from the National Institute of General Medical Sciences; N00014-98-1-0093 from the Office of Naval Research; DMI-0210844, EIA-0086015, EIA-0074808, DMR-01138790, and CTS-0103002 from the National Science Foundation; and F30602-01-2-0561 from DARPA/AFSOR.

Supporting Information Available: The Supporting Information contains the experimental details of denaturing and nondenaturing gel electrophoretic procedures. It also describes the protocols used for the radioactive and nonradioactive phosphorylation of DNA strands. This material is available free of charge via the Internet at http://pubs.acs.org.

JA049953D