



(51) International Patent Classification:

B82B 3/00 (2006.01) *C12Q 1/68* (2006.01)
B82Y 40/00 (2011.01) *C12N 15/10* (2006.01)
C07H 1/00 (2006.01) *C12N 15/11* (2006.01)
C07H 21/04 (2006.01)

(21) International Application Number:

PCT/US2017/026959

(22) International Filing Date:

11 April 2017 (11.04.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/321,146 11 April 2016 (11.04.2016) US

(71) Applicant: **PRESIDENT AND FELLOWS OF HARVARD COLLEGE** [US/US]; 17 Quincy Street, Cambridge, MA 02138 (US).

(72) Inventors: **GOPALKRISHNAN, Nikhil**; 30 Jamaicaaway, Apt. 7, Jamaica Plain, MA 02130 (US). **SCHAUS, Thomas, E.**; 90 Mount Vernon, Arlington, MA 02476 (US). **YIN, Peng**; 51 Winthrop Road, Brookline, MA 02445 (US).

(74) Agent: **DIPIETRANTONIO, Heather, J.**; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: SEQUENCE DESIGN FOR EFFICIENT ASSEMBLY OF NUCLEIC ACID STRUCTURES

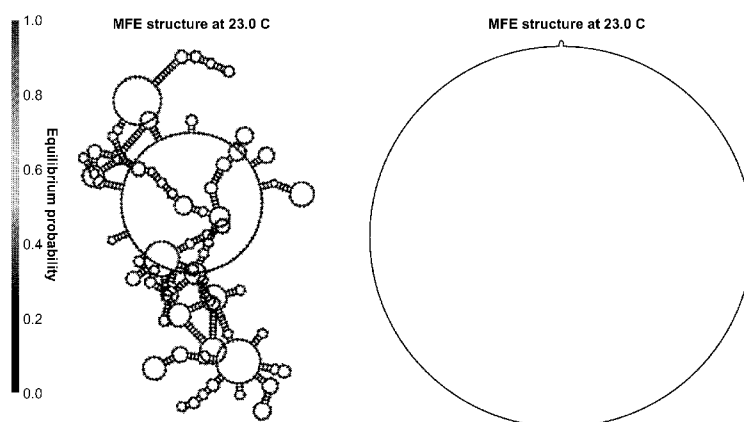


FIG. 1

(57) Abstract: The present disclosure provides, in some aspects, methods and compositions for producing nucleic acid nanostructures having little to no kinetic barriers to self-assembly.

SEQUENCE DESIGN FOR EFFICIENT ASSEMBLY OF NUCLEIC ACID STRUCTURES

RELATED APPLICATION

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application number 62/321,146, filed April 11, 2016, which is incorporated by reference herein in its entirety.

FEDERALLY SPONSORED RESEARCH

This invention was made with government support under Contract No. EB018659 awarded by National Institutes of Health, under Grant Nos. 1540214 and 1317291 awarded by National Science Foundation, and under Contract Nos. N00014-13-1-0593 and N00014-14-1-0610 awarded by United States Department of Defense Office of Naval Research. The government has certain rights in the invention.

BACKGROUND

Deoxyribonucleic acid (DNA) origami is the nanoscale folding of DNA to create non-arbitrary two-dimensional and three-dimensional shapes at the nanoscale. The specificity of the interactions between complementary base pairs make DNA a useful construction material, through design of its base sequences. The conventional DNA origami process involves the folding of a long single strand of viral DNA (often referred to as a “scaffold” strand) aided by multiple shorter “staple” strands. These shorter staple strands bind the longer scaffold strand in various places, resulting in various shapes. Thus, the sequence of a staple strand is determined by the sequence of the scaffold strand to which it binds. DNA nanostructures have been widely utilized in material science, synthetic biology, biological imaging, structural determination, and drug delivery.

SUMMARY

Provided herein, in some embodiments, are compositions and methods that permit efficient isothermal self-assembly of non-arbitrary nucleic acid (*e.g.*, DNA) nanostructures at a wide range of temperatures, including at physiological temperature and room temperature. Self-assembly of nanostructures, as provided herein, relies, at least in part, on minimizing or eliminating kinetic barriers that often hinder the nucleic acid self-assembly process by introducing secondary structure, intramolecular nucleic acid interactions (*e.g.*, hybridization

between subsequences within the same scaffold strand) and unintended intermolecular nucleic acid interactions (*e.g.*, hybridization between staple strands).

A contributing kinetic barrier to nucleic acid nanostructure self-assembly (*e.g.*, at room temperature) includes (i) secondary structure of the nucleic acid scaffold strand and nucleic acid staple strands, (ii) intermolecular staple-staple interactions and (iii) unintended intermolecular staple-scaffold interactions. Secondary structure occurs when two complementary or near-complementary regions within a nucleic acid strand hybridize, even transiently, to each other to form loops, duplexes, hairpin structures, three-way structures, four-way structures, or any other structure. Intermolecular interactions occur when two complementary or near-complementary regions across two or more nucleic acid strands hybridize, even transiently, to each other to form loops, duplexes, hairpin structures, three-way structures, four-way structures, or any other structure. The present disclosure demonstrates that secondary structure and undesired/unintended intermolecular interactions may be minimized or eliminated by rationally designing the nucleotide sequence of the strands to minimize or eliminate one of the four nucleotides (*e.g.*, A, T, C or G) and to minimize or eliminate repeat subsequences. A nucleic acid strand having a 3-letter code favors a linear arrangement, in part because there is one less binding partner facilitating intramolecular binding (*e.g.*, without a G, C cannot pair with another nucleotide within the strand). Two or more nucleic acid strands having a 3-letter code are also less likely to have intermolecular interactions because there is one less binding partner facilitating intermolecular binding. Minimizing repeated subsequences helps to prevent unintended intermolecular interactions by increasing the likelihood that complementary subsequences occur only where intended and are not present elsewhere in the scaffold strand.

Thus, nucleic acid “scaffold” and “staple” strand sequences of the present disclosure have primarily three different types of nucleotides (*e.g.*, A, T and C, or A, T and G), thus minimizing secondary structure formation and intermolecular staple-staple interactions. Also, nucleic acid “scaffold” and “staple” strand sequences of the present disclosure are rationally designed to minimize sequence repeats, thus minimizing unintended scaffold-staple interactions. A nucleic acid “scaffold” strand is typically a strand of nucleotides having a length of greater than 100 nucleotides. By contrast a nucleic acid “staple strand” is typically a strand of nucleotides having a length of 100 nucleotides or less. It should be understood that the lengths of scaffold strands and complementary staple strands may vary, as long as the scaffold strand is longer than the staple strands. Thus, a scaffold strand may be shorter than 100 nucleotides, provided the scaffold strand is longer than complementary staple strands.

Likewise, the staple strands may be longer than 100 nucleotides, provided the staple strands are shorter than a complementary scaffold strand (a strand that contains subsequences that are complementary to the staple strands).

In some embodiments, the GC content of nucleic acid scaffold and staple strands is minimized to reduce the strength of secondary structure and intermolecular staple-staple base pairing. G and C nucleotides form a 3-hydrogen bond when paired together, whereas A and T nucleotides form a weaker 2-hydrogen bond. Reducing or eliminating G-C binding weakens the strength of undesired interactions (*e.g.*, via base pairing). Thus, nucleic acid scaffold strands, in some embodiments, have a nucleotide sequence (*e.g.*, full-length sequence) that comprises less than 10% (*e.g.*, less than 9, 8, 7, 6, 5, 4, 3, 2 or 1%) of one of C, G, T or A nucleotides. For example, at least 90% of the nucleotides of a nucleic acid scaffold strand may be comprised of a combination of A, T and C, while less than 10% (*e.g.*, less than 5%) of the nucleotides of the nucleic acid scaffold strand may be comprised of G. As another example, at least 90% of the nucleotides of a nucleic acid scaffold strand may be comprised of a combination of A, T and G, while less than 10% (*e.g.*, less than 5%) of the nucleotides of the nucleic acid scaffold strand may be comprised of C.

Moreover, nucleic acid scaffold and staple sequences have minimal sequence repeats. This design also minimizes or eliminates unintended intermolecular scaffold-staple base pairing and, thus, kinetic traps.

Fig. 1 shows a comparison of secondary structures formed from a 1000-nucleotide scaffold strand having a conventional 4-letter (A, T, C, G) code (left) and a 1000-nucleotide scaffold strand having a 3-letter (*e.g.*, A, T, C) code (right). The scaffold strand having a 3-letter code exhibits little to no secondary structure.

Fig. 2 shows a comparison of secondary structure for corresponding sets of staple strands. The staple strands corresponding to a 4-letter (A, T, C, G) code are themselves 4-letter sequences (A, T, C, G) and exhibit significant secondary structure (as signified by a more negative free energy). The staple strands corresponding to a 3-letter code (A, T, C) are themselves 3-letter sequences (A, T, G) and exhibit little to no secondary structure (as signified by less negative free energy).

Fig. 3 shows a comparison of staple-staple interactions for corresponding sets of staple strands. Some of the staple strands corresponding to a 4-letter (A, T, C, G) code have significant undesired staple-staple interactions (signified by a low percentage of the staple strands as monomers in solution). The staple strands corresponding to a 3-letter code (A, T, C) are themselves 3-letter sequences (A, T, G) and exhibit little to no undesired staple-staple

interactions (signified by a high (close to 100%) percentage of the staple strands as monomers in solution).

Also provided herein are engineered nucleic acid scaffold strands having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides.

In some embodiments, a nucleic acid scaffold strand comprises greater than 90% (*e.g.*, 95% or 98%) of A, T/U and C nucleotides and less than 10% (*e.g.*, 5% or 2%) of G nucleotides. In some embodiments, a nucleic acid scaffold strand comprises greater than 90% of A, T/U and G nucleotides and less than 10% of C nucleotides. In some embodiments, the nucleic acid scaffold strand comprises a nucleotide sequence comprising 1-5% of one of C, G, T or A nucleotides. In some embodiments, the nucleic acid scaffold strand consists of: A, T and C; A, T and G; A, G and C; or T; G and C.

In some embodiments, the nucleic acid scaffold strand comprises nucleotide subsequences (contiguous sequences shorter than the full-length sequence), each of which is present only once (is not repeated) in the nucleic acid scaffold strand. In some embodiments, the nucleic acid scaffold strand comprises nucleotide subsequences, wherein any subsequence is repeated, if at all, only 1, 2 or 3 times.

In some embodiments, the nucleic acid scaffold strand has a length of at least 500 nucleotides. For example, a nucleic acid scaffold strand may have a length of at least 1000 nucleotides. In some embodiments, the nucleic acid scaffold strand has a length of at least 100-100,000 nucleotides. In some embodiments, the nucleic acid scaffold strand (and the nucleic acid nanostructure) is DNA. In some embodiments, the nucleic acid scaffold strand (and the nucleic acid nanostructure) is RNA or a combination of DNA and RNA.

In some embodiments, the nucleic acid nanostructure comprises a plurality of oligonucleotides (referred to herein as “staple” strands). For example, a nucleic acid nanostructure may include 10-2000 oligonucleotides (*e.g.*, DNA oligonucleotides). In some embodiments, the oligonucleotides have a length of 5 to 10 nucleotides, or a length of 3 to 30 nucleotides.

In some embodiments, a nucleic acid scaffold strand has a C-content or G-content of 20%-60% (*e.g.*, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55% or 60%) for each subsequence having a length of at least 10 nucleotides (*e.g.*, at least 15, at least 20, at least 30, at least 35, at least 40, at least 45 or at least 50 nucleotides).

In some embodiments, the nucleic acid nanostructure is two-dimensional or three-dimensional. In some embodiments, the nucleic acid nanostructure has a diameter of 10-500 nm.

Other aspects of the present disclosure provide a kit comprising (a) a nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides; and (b) a plurality of oligonucleotides, wherein each oligonucleotide is at least partially complementary to and at least partially bound to the nucleic acid strand. In some embodiments, the kit further comprising buffer supplemented with monovalent cations and/or divalent cations.

Still other aspects of the present disclosure provide a method of producing a nucleic acid nanostructure, comprising incubating, in a reaction vessel, (a) a nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides, and (b) a plurality of oligonucleotides, wherein each oligonucleotide is at least partially complementary to the nucleic acid strand, under conditions that result in binding of the oligonucleotides to the nucleic acid strand to form a nucleic acid nanostructure. In some embodiments, the method comprises conditions including incubating a nucleic acid strand and a plurality of oligonucleotides in a buffer at a temperature of 4-70 °C for at least 5 minutes.

BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 shows a random 1000 base four letter scaffold sequence (left) with significant secondary structure, while a three letter scaffold sequence (right) shows almost no secondary structure under the same conditions. The free energies of the two secondary structures are -165.53 kcal/mol and -0.77 kcal/mol, respectively. The scaffolds were analyzed with NUPACK, an online tool for analyzing nucleic acid thermodynamics.

Fig. 2 shows a comparison of secondary structure for corresponding sets of staple strands. The staple strands corresponding to a 4-letter (A, T, C, G) code are themselves 4-letter sequences (A, T, C, G) and exhibit significant secondary structure (as signified by a more negative free energy). The staple strands corresponding to a 3-letter code (A, T, C) are themselves 3-letter sequences (A, T, G) and exhibit little to no secondary structure (as signified by less negative free energy).

Fig. 3 shows a comparison of staple-staple interactions for corresponding sets of staple strands. Some of the staple strands corresponding to a 4-letter (A, T, C, G) code have significant undesired staple-staple interactions (signified by a low percentage of the staple

strands as monomers in solution). The staple strands corresponding to a 3-letter code (A, T, C) are themselves 3-letter sequences (A, T, G) and exhibit little to no undesired staple-staple interactions (signified by a high (close to 100%) percentage of the staple strands as monomers in solution).

Fig. 4 shows DNA origami assembled from a 1000 base long DeBruijn-derived scaffold sequence and a corresponding staple set. The scaffold and staple strands were mixed together and incubated at room temperature for five minutes in 1X TE buffer supplemented with 10mM Mg^{++} . Image size: 1 μ m x 1 μ m.

DESCRIPTION

Nucleic acid nanostructure self-assembly using, for example, a conventional origami approach (Rothemund, *Nature* 440 (7082): 297–302), typically requires annealing of nucleic acid strands at high temperature (*e.g.*, >80 °C), which are not always compatible with temperature-sensitive components or contexts.

Provided herein are compositions and methods for self-assembly of nucleic acid (*e.g.*, DNA) nanostructures at a wide range of temperatures, including at physiological temperature and room temperature). Self-assembly of nanostructures, as provided herein, relies, in part, on minimizing or eliminating kinetic barriers such as undesired secondary structure, undesired intermolecular nucleic acid interactions (*e.g.*, staple-staple interactions) and unintended intermolecular nucleic acid interactions (*e.g.* staple-scaffold interactions) that often hinder the self-assembly process. Nucleic acid sequences of the present disclosure are engineered to minimize or eliminate secondary structure and undesired/unintended intermolecular base pairing, thereby facilitating self-assembly. More specifically, nucleic acid sequences disclosed herein have primarily three different types of nucleotides (*e.g.*, A, T and C, or A, T and G) and have minimal sequence repeats.

Nucleic acid strands

Some aspects of the present disclosure provide engineered nucleic acid strands (*e.g.*, scaffold strands and staple strands) that self-assemble in a pre-determined matter to form a nucleic acid nanostructure.

Kinetic barriers to nucleic acid self-assembly are minimized or eliminated by engineering a nucleic acid strand in a way that minimizes or eliminates secondary structure formation and undesired/unintended intermolecular interactions. The present disclosure includes varying the percentage of C, G, T or A nucleotides in an engineered nucleic acid

strand to minimize or eliminate kinetic barriers to self-assembly. The percentage of C, G, T or A nucleotides with reference to a nucleic acid scaffold strand refers to the percentage of that nucleotide in the entirety) full length of the strand. Percent nucleotide composition may also be defined in the context of subsequences. Thus, particular subsequences (sequences shorter than the full length sequence) may contain a particular percentage of one or more nucleotide types (*e.g.*, A particular percentage of A, T or C/G, or a particular percentage of a combination of A, T and C/G.) “C/G” refers to C or G.

In some embodiments, the engineered nucleic acid strand has a length of at least 100 nucleotides and has a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides. For example, a nucleic acid scaffold strand may be primarily composed of A, T and C nucleotides and contain only a few G nucleotides. Thus, in some embodiments, the engineered nucleic acid strand comprises greater than 90% (*e.g.*, greater than 95%) of A, T/U and C nucleotides and less than 10% (*e.g.*, less than 5%) of G nucleotides. In some embodiments, the engineered nucleic acid strand comprises greater than 90% (*e.g.*, greater than 95%) of A, T/U and G nucleotides and less than 10% (*e.g.*, less than 5%) of C nucleotides. In some embodiments, the engineered nucleic acid strand comprises greater than 90% (*e.g.*, greater than 95%) of A, C and G nucleotides and less than 10% (*e.g.*, less than 5%) of T nucleotides. In some embodiments, the engineered nucleic acid strand comprises greater than 90% (*e.g.*, greater than 95%) of T, C and G nucleotides and less than 10% (*e.g.*, less than 5%) of A nucleotides.

The engineered nucleic acid strand percentages are not limited to less than 10% of one of C, G, T, or A nucleotides. In some embodiments, the engineered nucleic acid strand comprises less than 5% of one of C, G, T or A nucleotides. In some embodiments, the engineered nucleic acid strand comprises 1-5% (*e.g.*, 1%, 2%, 3%, 4% or 5%) of one of C, G, T or A nucleotides.

The present disclosure includes minimizing subsequence repeats in the engineered nucleic acid strand sequence to minimize or eliminate kinetic barriers to self-assembly. Despite reduced sequence diversity of a nucleic acid strand comprising three different nucleotides compared to a nucleic acid strand comprising four different nucleotides, many distinct nucleotide sequences may be generated. For example, a nucleotide sequence having a length of 8 and consisting of 3 different nucleotides yields ~ 6500 ($3^8 = \sim 6500$) distinct nucleotide sequences that may occur in the nucleic acid strand. A nucleotide sequence having a length of 8 and consisting of 4 different nucleotides yields ~ 65000 ($4^8 = \sim 65000$) distinct nucleotide sequences that may occur in the nucleic acid strand.

A “subsequence” refers to a discrete, contiguous sequence of nucleotides or nucleotide base pairs, depending on whether the domain is unpaired (single-stranded nucleotides) or paired (double-stranded nucleotide base pairs), respectively. One subsequence is “complementary to” another subsequence if one subsequence contains nucleotides that base pair (hybridize/bind through Watson-Crick nucleotide base pairing) with nucleotides of the other subsequence such that the two subsequences form a paired (double-stranded) or partially-paired molecular species/structure. In some embodiments, a subsequence in a first nucleic acid strand is described as “complementary to” a subsequence in a second nucleic acid strand such that, by pairing up complementary regions, the nucleic acid strand becomes double-stranded. In some embodiments, a subsequence in the nucleic acid strand is described as “complementary to” an oligonucleotide such that, by pairing up complementary regions, the DNA assembles into a predefined nanostructure. Complementary subsequences need not be perfectly (100%) complementary to form a paired structure, although perfect complementarity is provided, in some embodiments.

In some embodiments, a nucleic acid strand sequence comprises nucleotide subsequences, each of which is present only once (or up to twice or up to three times) in a nucleic acid scaffold strand. In some embodiments, subsequences have a length of 3 to 50 nucleotides. For example, subsequences may have a length of 3-40, 3-30, 3-20, 3-10, 5-50, 5-40, 5-30, 5-20, or 5-10 nucleotides.

In some embodiments, a nucleic acid strand comprises a nucleotide subsequence having a length of 3 to 30 nucleotides. In some embodiments, the engineered nucleic acid strand comprises a nucleotide subsequence having a length of 3 to 35 nucleotides, 3 to 40 nucleotides, 3 to 45 nucleotides, 3 to 50 nucleotides, 3 to 100 nucleotides, or 3 to 200 nucleotides. In some embodiments, the engineered nucleic acid strand comprises a nucleotide subsequence having a length of 3 to 25 nucleotides, 3 to 20 nucleotides, 3 to 15 nucleotides, 3 to 10 nucleotides, or 3 to 5 nucleotides. In some embodiments, the engineered nucleic acid strand comprises a nucleotide subsequence having a length of 4 to 11 nucleotides, 4 to 10 nucleotides, 4 to 9 nucleotides, 4 to 8 nucleotides, 4 to 7 nucleotides, 4 to 6 nucleotides, or 4 to 5 nucleotides. In some embodiments, the engineered nucleic acid strand comprises a nucleotide subsequence having a length of 11 to 10 nucleotides, 11 to 9 nucleotides, 11 to 8 nucleotides, 11 to 7 nucleotides, 11 to 6 nucleotides, or 11 to 5 nucleotides. In some embodiments, the engineered nucleic acid strand comprises a nucleotide subsequence having a length of 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides.

In some embodiments, a nucleic acid strand comprises one or more nucleotide subsequences. In some embodiments, the engineered nucleic acid strand comprises 1-10 subsequences, 1-100 subsequences, 1-500 subsequences, or 1-1000 subsequences. In some embodiments, the engineered nucleic acid strand comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, or more nucleic acid subsequences.

In some embodiments, the nucleic acid strand has a GC-content of 20% to 60% for each subsequence having a length of at least 10, 20, 30, 40 or 50 nucleotides. In some embodiments, the nucleic acid subsequence having a length of at least 10, 20, 30, 40 or 50 nucleotides has a GC-content of 20% to 60%, 20% to 50%, 20% to 40%, or 20% to 30%. In some embodiments, the nucleic acid subsequence having a length of at least 10, 20, 30, 40 or 50 nucleotides has a GC-content of 25% to 60%, 30% to 60%, 40% to 60%, or 50% to 60%. In some embodiments, the nucleic acid subsequence having a length of at least 10, 20, 30, 40 or 50 nucleotides has a GC-content of 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, or 60%.

In some embodiments, a nucleic acid strand has a length of at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1500, at least 2000, at least 3000, at least 4000, or at least 5000 nucleotides. In some embodiments, a nucleic acid strand may have a length of 100-100000, 100-50000, 100-25000, 100-10000, 100-5000, 100-4000, 100-3000, 100-2000, 100-1500, 100-1000, 100-500, or 100-250 nucleotides. In some embodiments, a nucleic acid strand is shorter than 100, shorter than 90, shorter than 80, shorter than 70, shorter than 60, or shorter than 50 nucleotides. In some embodiments, a nucleic acid strand has a length of 100, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500 or 5000 nucleotides. For example, a nucleic acid strand having a length of 1000 nucleotides is shown in Fig. 1.

A nucleic acid strand may comprise DNA (*e.g.*, genomic DNA, cDNA or a combination of genomic DNA and cDNA), RNA or a hybrid molecule, for example, where the nucleic acid contains any combination of deoxyribonucleotides and ribonucleotides (*e.g.*, artificial or natural), and any combination of two or more bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine, hypoxanthine, isocytosine and isoguanine. In some embodiments, the engineered nucleic acid strand is a DNA strand.

It should be understood that the nucleic acid sequences of the present disclosure are engineered to minimize or eliminate kinetic barriers to self-assembly. Thus, the nucleic acid strand may be referred to as an “engineered nucleic acid strand” or an “engineered scaffold strand.” An engineered nucleic acid strand is a contiguous sequence of nucleic acids

comprising less than 10% of one of C, G, T or A nucleotides and nucleotide sequences having minimal (*e.g.*, less than 4, less than 3, less than 2) repetitive subsequences.

Nucleic acid sequences of the present disclosure may include backbone modifications, base modifications, and/or sugar modifications. Non-limiting examples of these modifications are provided herein.

In some embodiments, a nucleic acid strand of the present disclosure may comprise a backbone other than a phosphodiester backbone. For example, an engineered nucleic acid strand, in some embodiments, may comprise phosphoramidate, phosphorothioate, phosphorodithioate, O-methylphosphoroamidite linkages, peptide nucleic acids or a combination of any two or more of the foregoing linkages. An engineered nucleic acid strand may be single-stranded (ss) or double-stranded (ds), as specified, or an engineered nucleic acid strand may contain portions of both single-stranded and double-stranded sequence. In some embodiments, an engineered nucleic acid strand contains portions of triple-stranded sequence, or other non-Watson-Crick base pairing such as G-quartets, G-quadruplexes, and i-motifs.

In some embodiments, a nucleic acid strand may comprise a sugar modification. For example, a β -D-ribose unit or a β -D-2'-deoxyribose unit can be replaced by a modified sugar unit, wherein the modified sugar unit is for example selected from β -D-ribose, α -D-2'-deoxyribose, L-2'-deoxyribose, 2'-F-2'-deoxyribose, arabinose, 2'-F-arabinose, 2'-O-(C₁-C₆)alkyl-ribose, preferably 2'-O-(C₁-C₆) alkyl-ribose is 2'-O-methylribose, 2'-O(C₂-C₆)alkenyl-ribose, 2'-[O-(C₁-C₆)alkyl-O(C₁-C₆)alkyl]-ribose, 2'-NH₂-2'-deoxyribose, 13-D-xylo-furanose, α -arabinofuranose, 2,4-dideoxy-3-D-erythro-hexo-pyranose, and carbocyclic and/or open-chain sugar analogs and/or bicyclosugar analogs.

In some embodiments, a nucleic acid strand may comprise a base modifications. Modified bases include modified cytosines (such as 5-substituted cytosines (*e.g.*, 5-methyl-cytosine, 5-fluoro-cytosine, 5-chloro-cytosine, 5-bromo-cytosine, 5-iodo-cytosine, 5-hydroxy-cytosine, 5-hydroxymethyl-cytosine, 5-difluoromethyl-cytosine, and unsubstituted or substituted 5-alkynyl-cytosine), 6-substituted cytosines, N4-substituted cytosines (*e.g.*, N4-ethyl-cytosine), 5-aza-cytosine, 2-mercapto-cytosine, isocytosine, pseudo-isocytosine, cytosine analogs with condensed ring systems (*e.g.*, N,N'- propylene cytosine or phenoxazine), and uracil and its derivatives (*e.g.*, 5-fluoro-uracil, 5-bromo-uracil, 5-bromovinyl-uracil, 4-thio-uracil, 5-hydroxy-uracil, 5-propynyl-uracil), modified guanines such as 7-deazaguanine, 7-deaza-7-substituted guanine (such as 7-deaza-7-(C₂-C₆)alkynylguanine), 7-deaza-8-substituted guanine, hypoxanthine, N2-substituted guanines

(*e.g.* N2-methyl-guanine), 5-amino-3-methyl-3H,6H-thiazolo[4,5-d]pyrimidine-2,7-dione, 2,6- diaminopurine, 2-aminopurine, purine, indole, adenine, substituted adenines (*e.g.* N6-methyl-adenine, 8-oxo-adenine) 8-substituted guanine (*e.g.* 8-hydroxyguanine and 8-bromoguanine), and 6-thioguanine. The nucleic acids may comprise universal bases (*e.g.* 3-nitropyrrole, P-base, 4-methyl-indole, 5-nitro-indole, and K-base) and/or aromatic ring systems (*e.g.* fluorobenzene, difluorobenzene, benzimidazole or dichloro-benzimidazole, 1-methyl-1H-[1,2,4]triazole-3-carboxylic acid amide).

Nucleic acid nanostructure

Nucleic acid nanostructures, as provided herein, are two-dimensional or three-dimensional structures that self-assemble into a predesigned shape by hybridization of a plurality of shorter staple oligonucleotides to a longer scaffold nucleic acid strand. In order to self-assemble, nuclei acid nanostructures must overcome kinetic barriers to assembly including intramolecular base pairing and secondary structure formation. The fundamental principle for nanostructure assembly, as provided herein, is engineering nucleic acid sequences in a way that minimizes or eliminates secondary structure formation and undesired/unintended intermolecular interactions, thereby minimizing or eliminating kinetic barriers to assembly. More specifically, nucleic acid sequences disclosed herein have primarily three different types of nucleotides (*e.g.*, A, T and C, or A, T and G) and have minimal subsequence repeats.

Nucleic acid “nanostructures” produced in accordance with the present disclosure are typically nanometer-scale structures (*e.g.*, having lengths of 1 to 1000 nanometers). In some embodiments, the term “nanostructure” herein may refer to micrometer-scale structures (*e.g.*, assembled from more than one nanometer-scale or micrometer-scale structure). In some embodiments, a nanostructure has a diameter of greater than 500 nm or greater than 1000 nm. In some embodiments, a nanostructure has a length scale of 1 micrometer to 2 micrometers. In some embodiments, a nanostructure has a length scale of 10 to 500 nm, 10 to 450 nm, 10 to 400 nm, 10 to 350 nm, 10 to 300 nm, 10 to 250 nm, 10 to 200 nm, 10 to 150 nm, 10 to 100 nm, 10 to 50 nm, or 10 to 25 nm. In some embodiments, the nanostructure has a diameter of 500 to 450 nm, 500 to 400 nm, 500 to 350 nm, 500 to 300 nm, 500 to 250 nm, 500 to 200 nm, 500 to 150 nm, 500 to 100 nm, 500 to 50 nm, or 500 to 25nm. In some embodiments, the nanostructure has a diameter of 10, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, or 500 nm.

A plurality of oligonucleotides anneals to the nucleic acid strand through nucleotide base pairing to fold the nanostructure into an arbitrary shape. In some embodiments, the plurality of oligonucleotides is a plurality of DNA oligonucleotides. In some embodiments, the plurality of oligonucleotides comprises 10-2000 oligonucleotides, 10-1500 oligonucleotides, 10-1000 oligonucleotides, 10-500 oligonucleotides, 10-250 oligonucleotides, 10-100 oligonucleotides, 10-50 oligonucleotides or 10-25 oligonucleotides. In some embodiments, the plurality of oligonucleotides comprises 2000-1500 oligonucleotides, 2000-1000 oligonucleotides, 2000-500 oligonucleotides, 2000-250 oligonucleotides, 2000-100 oligonucleotides, 2000-50 oligonucleotides, or 2000-25 oligonucleotides. In some embodiments, the plurality of oligonucleotides comprises 500-600 oligonucleotides, 600-700 oligonucleotides, 700-800 oligonucleotides, 800-900 oligonucleotides, 1000-1100 oligonucleotides, 1100-1200 oligonucleotides, 1200-1300 oligonucleotides, 1300-1400 oligonucleotides, or 1400-1500 oligonucleotides. In some embodiments, the plurality of oligonucleotides comprises 10, 25, 50, 100, 250, 500, 1000, 1500, or 2000 oligonucleotides.

The nucleic acid nanostructures produced in accordance with the present disclosure may take any shape or form. For example, a rectangular DNA nanostructure comprising a nucleic acid strand and a plurality of oligonucleotides was assembled as shown in Fig. 4. Using the methods disclosed herein, it is possible to predesign and assemble nanostructures comprising various shapes, forms, and sizes.

In some embodiments, the nucleic acid nanostructure comprises a nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides; and a plurality of oligonucleotides, wherein each oligonucleotide is at least partially complementary to and at least partially bound to the nucleic acid strand. Nucleic acid strand embodiments of a have been disclosed herein.

It should be also understood that the nucleic acid sequences may be designed manually or by computer means based on the teachings provided herein. In some embodiments, the nucleic acid sequences are DeBruijn derived nucleic acid sequences.

The nucleic acid structures may be isolated and/or purified. Isolation, as used herein, refers to the physical separation of the desired entity (*e.g.*, nucleic acid structures, etc.) from the environment in which it was generated. The isolation may be partial or complete.

Isolation of the nucleic acid structure may be carried out by running a hybridization reaction mixture on a gel and isolating nucleic acid structures that migrate at a particular molecular weight and are thereby distinguished from the nucleic acid substrates and the

spurious products of the hybridization reaction. As another example, isolation of nucleic acid structures may be carried out using a buoyant density gradient, sedimentation gradient centrifugation, or through filtration means.

Nucleic acid nanostructure assembly

Provided herein, in some aspects, are methods of minimizing or eliminating kinetic barriers through sequence design, thereby providing for a means of rapid assembly of nucleic acid nanostructures at room temperature.

In some embodiments, a method of producing a nucleic acid nanostructure comprises incubating, in a reaction vessel, (a) a nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides; and (b) a plurality of oligonucleotides, wherein each oligonucleotide is at least partially complementary to and at least partially bound to the nucleic acid strand, under conditions that result in binding of the oligonucleotides to the nucleic acid strand to form a nucleic acid nanostructure.

In some embodiments, the conditions include incubating nucleic acid strand and plurality of oligonucleotides for at least 5 minutes, at least 10 minutes, at least 15 minutes, at least 20 minutes, at least 25 minutes, at least 30 minutes, at least 35 minutes, at least 40 minutes, at least 45 minutes, at least 50 minutes, at least 55 minutes, or at least 60 minutes. In some embodiments, the conditions include incubating nucleic acid strand and plurality of oligonucleotides for more than 1 hour, more than 2 hours, more than 4 hours, more than 8 hours, more than 24 hours, more than 48 hours, or more than 72 hours.

In some embodiments, the conditions include incubating nucleic acid strand and plurality of oligonucleotides at room temperature. In some embodiments, the conditions include incubating nucleic acid strand and plurality of oligonucleotides isothermally at a temperature ranging from 4°C to 10°C, 4°C to 20°C, 4°C to 30°C, 4°C to 40°C, 4°C to 50°C, 4°C to 60°C, or 4°C to 70°C. In some embodiments, the conditions include incubating nucleic acid strand and plurality of oligonucleotides isothermally at 20°C, 21°C, 22°C, 23°C, 24°C, 25°C, 26°C, 27°C, 28°C, 29°C, 30°C.

In some embodiments, a nucleic acid strand and complementary oligonucleotides are combined in a reaction vessel such as, but not limited to a tube, a well, a vial, and the like. In some embodiments, a reaction vessel is incubated in any constant temperature environment. In some embodiments, a reaction vessel is incubated in a PCR machine.

In an exemplary method, a rectangular nanostructure comprising a nucleic acid strand having a length of 1000 nucleotides and a plurality of oligonucleotides were incubated in

buffer containing 1X TE and 10 mM Mg^{2+} and assembled in only 5 minutes at room temperature (Fig. 2).

Applications and Uses

It should be understood that while the nucleic acid nanostructures themselves (the base structure) typically consist of a engineered nucleic acid strand and a plurality of oligonucleotides folded into a desired shape, the nanostructures may comprise agents or other molecules that are added to or attached to the folded nanostructure, as described elsewhere herein. For example, in some embodiments, nanostructures of the present disclosure may be decorated with therapeutic moieties (*e.g.*, antibodies) that are more stable at room temperature and, thus the therapeutic moieties maintain stability during rapid room temperature assembly which lacks exposure to high annealing temperatures.

The nanostructures of the present disclosure may be used as therapeutics, or therapeutic delivery vehicles, for a multitude of biomedical applications, including cancer therapy and immunotherapy. Nanostructures can be used, for example, to organize various agents (*e.g.*, proteins, such as antibodies) with controlled spatial distance. In another example, nanostructures can be used to encapsulate therapeutic agents (*e.g.*, antibodies) and deploy them specifically in the vicinity of a tissue in vivo (*e.g.*, a tumor) using, for example, targeting peptides (*e.g.*, tumor targeting peptides).

The nanostructures may contain any agent that is intended for use in vivo and/or in vitro, in a biological or non-biological application. For example, an agent may be any atom, molecule, or compound that can be used to provide benefit to a subject (including without limitation prophylactic or therapeutic benefit) or that can be used for diagnosis and/or detection (for example, imaging) in vivo, or that may be used for effect in an in vitro setting (for example, a tissue or organ culture, a clean-up process, and the like). The agents may be without limitation therapeutic agents and diagnostic agents. Any agent may be incorporated in a disclosed nanostructure provided that it can be assembled and/or loaded into the nanostructure.

The agent may be without limitation a chemical compound including a small molecule, a protein, a polypeptide, a peptide, a nucleic acid, a virus-like particle, a steroid, a proteoglycan, a lipid, a carbohydrate, and analogs, derivatives, mixtures, fusions, combinations or conjugates thereof. The agent may be a prodrug that is metabolized and thus converted in vivo to its active (and/or stable) form. More than one agent or more than one type of agent may be assembled and/or loaded in the nanostructure.

A variety of agents that are currently used for therapeutic or diagnostic purposes include without limitation imaging agents, immunomodulatory agents such as immunostimulatory agents and immunoinhibitory agents (*e.g.*, cyclosporine), antigens, adjuvants, cytokines, chemokines, anti-cancer agents, anti-infective agents, nucleic acids, antibodies or fragments thereof, fusion proteins such as cytokine-antibody fusion proteins, Fc-fusion proteins, analgesics, opioids, enzyme inhibitors, neurotoxins, hypnotics, anti-histamines, lubricants, tranquilizers, anti-convulsants, muscle relaxants, anti-Parkinson agents, anti-spasmodics, muscle contractants including channel blockers, miotics and anti-cholinergics, anti-glaucoma compounds, modulators of cell-extracellular matrix interactions including cell growth inhibitors and anti-adhesion molecules, vasodilating agents, inhibitors of DNA, RNA or protein synthesis, anti-hypertensives, anti-pyretics, steroidal and non-steroidal anti-inflammatory agents, anti-angiogenic factors, anti-secretory factors, anticoagulants and/or antithrombotic agents, local anesthetics, ophthalmics, prostaglandins, targeting agents, neurotransmitters, proteins, cell response modifiers, and vaccines.

In some embodiments, an agent is a diagnostic agent such as an imaging agent. As used herein, an imaging agent is an agent that emits signal directly or indirectly thereby allowing its detection *in vivo*. Imaging agents such as contrast agents and radioactive agents can be detected using medical imaging techniques such as nuclear medicine scans and magnetic resonance imaging (MRI). Imaging agents for magnetic resonance imaging (MRI) include Gd(DOTA), iron oxide or gold nanoparticles; imaging agents for nuclear medicine include ²⁰¹Tl, gamma-emitting radionuclide ^{99m}Tc; imaging agents for positron-emission tomography (PET) include positron-emitting isotopes, (18)F-fluorodeoxyglucose ((18)FDG), (18)F-fluoride, copper-64, gadoamide, and radioisotopes of Pb(II) such as ²⁰³Pb, and ¹¹¹In; imaging agents for *in vivo* fluorescence imaging such as fluorescent dyes or dye-conjugated nanoparticles.

Nucleic acid nanostructure kit

In one aspect, the present disclosure provides a kit for assembling, at room temperature, a nanostructure comprising a nucleic acid strand having three different nucleotides and a plurality of oligonucleotides.

A kit may contain one or more components related to assembly of a nanostructure at room temperature. In some embodiments, a kit comprises a nucleic acid “scaffold” strand as described herein. In some embodiments, a kit comprises a plurality of oligonucleotides as presently disclosed, wherein each oligonucleotide is at least partially complementary to and

at least capable of partially binding to the nucleic acid strand. In some embodiments, a kit component comprises instructions for assembling a nanostructure as described herein. In some embodiments, the instructions relate to composition of one or more kit components and/or methods of use of nanostructures.

The components of the kit may be provided in any suitable form. In some embodiments, components are provided as liquid solutions or as dried powders. In some embodiments, components provided as a dry powder may be reconstituted by the addition of a suitable buffer, which may also be provided. In some embodiments, the buffer comprises magnesium, sodium, monovalent cations, and/or divalent cations. For example, a nanostructure was assembled in buffer comprising 1X TE buffer supplemented with 10 mM Mg^{2+} . In some embodiments, components provided as liquid solutions may be concentrated or ready to use.

EXAMPLE

Eliminating Kinetic Barriers through a Sequence Design Allows Rapid Assembly of DNA Origami at Room Temperature

DNA origami [3] is a robust and programmable technique to make DNA nanostructures. A key use of DNA origami structures has been as nanoscale breadboards that organize components like small proteins, nanoparticles, small molecules and other DNA strands. Typical DNA origami assembly requires that the scaffold strand be annealed with staple strands in the appropriate buffer. The high temperatures ($> 80^\circ\text{C}$) used in this process are not always compatible with temperature-sensitive components or contexts. In response, there have been attempts to isothermally fold DNA. Studies have shown that DNA origami can assemble isothermally if held at temperatures around 60°C [4, 5], or at room temperature by using denaturing agents [1, 2]. However, rapid assembly of DNA origami at room temperature without the use of denaturing agents, which can be toxic or have other unintended consequences, has not been achieved.

A rectangular DNA origami was assembled from a 1000 base long DeBruijn-derived scaffold sequence and a corresponding staple set. The origami assembled isothermally in as little as five minutes at temperatures ranging from 25°C to 50°C .

Traditional DNA origami designs are faced with prohibitive kinetic barriers to assembly at room temperature. Heating, or using denaturing agents, helps overcome these barriers. An alternate approach was taken. Computational sequence design was used to

eliminate prohibitive kinetic barriers and thus assemble DNA origami at room temperature in as little as five minutes.

The key contributions to kinetic barriers are: secondary structure of the scaffold, secondary structure of staple strands, and undesired staple-staple interactions. These issues were circumvented by the use of a three letter code. The scaffold sequence was chosen from the three letter alphabet {A, T, C} and the staple sequences from the complementary alphabet {A, T, G}. Mfold and NUPACK analysis of the resulting sequences showed dramatic reduction in predicted secondary structure (Fig. 1) and unintended staple-staple interactions through this strategy. A potential drawback of a three letter code is reduced sequence diversity. A four letter code affords $4^8 \approx 65000$ distinct 8-mers while a three letter code affords only $3^8 \approx 6500$. Thus, a poorly chosen three letter scaffold sequence could have poor sequence diversity that results in unintended staple-scaffold interactions, resulting in kinetic barriers to assembly. Such “poor choices” were minimized by using DeBruijn sequences to maximize sequence diversity and prevent unintended staple-scaffold interactions. A DeBruijn sequence S of order n over the three letter alphabet {A, T, C} is a sequence where every possible subsequence of length n appears exactly once. DeBruijn sequences can be generated computationally in time linear in the length of the sequence. DeBruijn sequences were adapted to filter out degenerate subsequences and to ensure that the C content was evenly distributed across the sequence. The resulting sequence is considered to be DeBruijn-derived and is selected as the scaffold sequence (Fig. 2).

The following example illustrates derivation of DeBruijn sequences. Given an alphabet of size k (i.e. there are k letters in the alphabet; for example, {A, T, C} is a 3-letter alphabet) there are a possible k^n distinct words of length n. For example, the alphabet {A, T, C} has $3^2 = 9$ distinct words of length 2: AA, AT, AC, TA, TT, TC, CA, CT, CC.

The number of subsequences is calculated from the sequence length (l). Consider a sequence (S) of length (l), wherein S has $l - n + 1$ subsequences of length n. For example, the sequence ATTTATACCTATTA (SEQ ID NO: 1) has length 14 and has $14 - 3 + 1 = 12$ subsequences of length 3, namely ATT, TTT, TTA, TAT, ATA, TAC, ACC, CCT, CTA, TAC, ATT, TTA which are obtained by taking a window of length 3 and sliding it along the sequence. Note that not all subsequences are necessarily unique. For instance, the subsequence ATT and TTA both occur twice, and we say they are repeated.

Sequence repeats are minimized according to the following: a k-ary DeBruijn sequence of order n is a sequence over a k letter alphabet where every subsequence of size n

occurs exactly once. There are k^n such subsequences and hence the length of the DeBruijn sequence is $l = k^n + n - 1$.

Given a sequence of length l over a k letter alphabet, we minimize sequence repeats by the following procedure:

- Choose the smallest positive integer n that satisfies $k^n + n - 1 \leq l$
- Compute a k -ary DeBruijn sequence of order n by taking an Eulerian tour of a DeBruijn graph (this is a standard algorithm)
- Choose any subsequence of length l of this DeBruijn sequence. This is our target sequence. Note that our target sequence has the property that all subsequences of length n or larger occur at most once. That is, no subsequences of length n or larger are repeated.

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.” The term “or” is generally employed in its sense including “and/or” unless the content clearly dictates otherwise.

It should be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

In the claims, as well as in the specification above, all transitional phrases such as “comprising,” “including,” “carrying,” “having,” “containing,” “involving,” “holding,” “composed of,” and the like are to be understood to be open-ended (including but not limited to). Only the transitional phrases “consisting of” and “consisting essentially of” shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03. It should also be understood that all open-ended transitional phrases may be substituted with closed or semi-closed transitional phrases. Thus, the term “comprising” may be substituted with “consisting of” or “consisting essentially of.”

References

- [1] Ralf Jungmann, Tim Liedl, Thomas L. Sobey, William Shih, and Friedrich C. Simmel. Isothermal assembly of DNA origami structures using denaturing agents. *Journal of the American Chemical Society*, 130(31):10062–10063, 2008. PMID: 18613687.
- [2] Andreas Kapielski, Anne Schneider, Andrea Csaki, and Wolfgang Fritzsche. Isothermal DNA origami folding: avoiding denaturing conditions for one-pot, hybrid-component annealing. *Nanoscale*, 7:2102–2106, 2015.
- [3] Paul Rothmund. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature*, 440:297–302, 2006.
- [4] Jean-Philippe J. Sobczak, Thomas G. Martin, Thomas Gerling, and Hendrik Dietz. Rapid folding of DNA into nanoscale shapes at constant temperature. *Science*, 338(6113):1458–1461, 2012.
- [5] Jie Song, Zhao Zhang, Shuai Zhang, Lei Liu, Qiang Li, Erqing Xie, Kurt Vesterager Gothelf, Flemming Besenbacher, and Mingdong Dong. Isothermal hybridization kinetics of DNA assembly of two-dimensional DNA origami. *Small*, 9(17):2954–2959, 2013.

What is claimed is:

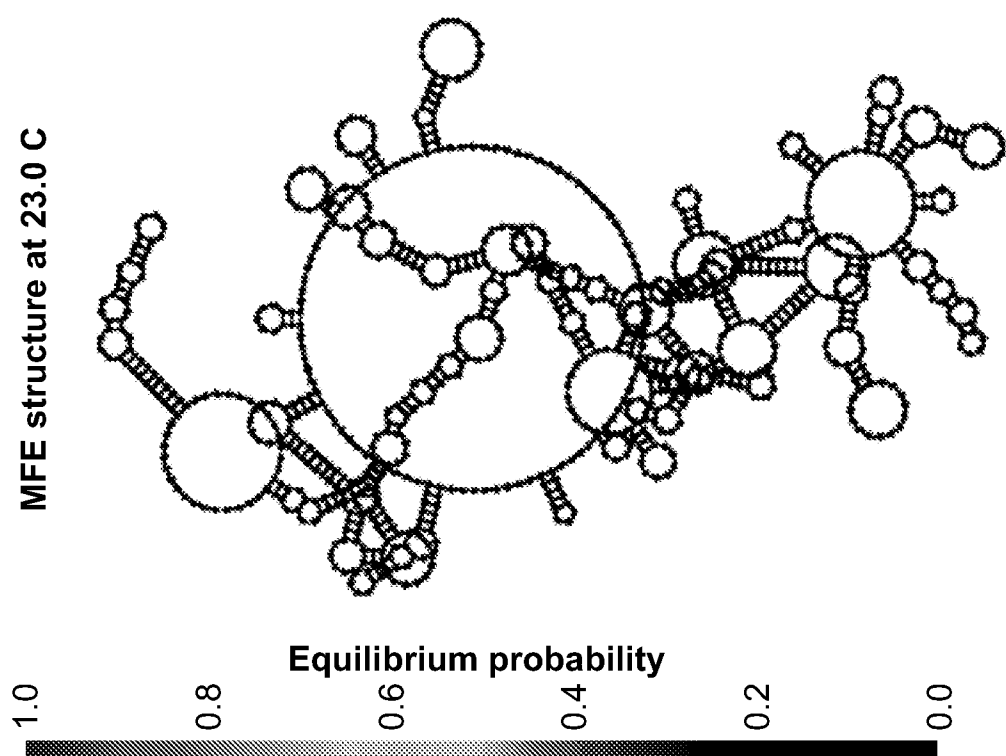
CLAIMS

1. A nucleic acid nanostructure comprising
 - (a) a nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides; and
 - (b) a plurality of oligonucleotides, wherein each oligonucleotide is at least partially complementary to and at least partially bound to the nucleic acid strand.
2. The nucleic acid nanostructure of claim 1, wherein the nucleic acid strand comprises greater than 90% of A, T/U and C nucleotides and less than 10% of G nucleotides.
3. The nucleic acid nanostructure of claim 1, wherein the nucleic acid strand comprises greater than 90% of A, T/U and G nucleotides and less than 10% of C nucleotides.
4. The nucleic acid nanostructure of any one of claims 1-3, wherein the nucleotide sequence comprises less than 5% of one of C, G, T or A nucleotides.
5. The nucleic acid nanostructure of any one of claims 1-4, wherein the nucleotide sequence comprises 1-5% of one of C, G, T or A nucleotides.
6. The nucleic acid nanostructure of any one of claims 1-5, wherein that the nucleotide sequence comprises nucleotide subsequences, each of which is present only once in the nucleic acid strand.
7. The nucleic acid nanostructure of any one of claims 1-6, wherein the nucleic acid strand has a length of at least 500 nucleotides.
8. The nucleic acid nanostructure of any one of claims 1-7, wherein the nucleic acid strand is a DNA strand.
9. The nucleic acid nanostructure of any one of claims 1-8, wherein the plurality of oligonucleotides is a plurality of DNA oligonucleotides.

10. The nucleic acid nanostructure of any one of claims 1-9, wherein the nucleotide subsequences have a length of 5 to 10 nucleotides.
11. The nucleic acid nanostructure of any one of claims 1-10, wherein the nucleic acid strand has a length of at least 1000 nucleotides.
12. The nucleic acid nanostructure of any one of claims 1-10, wherein the nucleic acid strand has a length of 100-100,000 nucleotides.
13. The nucleic acid nanostructure of any one of claims 1-12, wherein the plurality of oligonucleotides comprises 10-2000 oligonucleotides.
14. The nucleic acid nanostructure of any one of claims 6-13, wherein the nucleic acid strand has a C- or G-content of 20%-60% for each subsequences having a length of at least 30 nucleotides.
15. The nucleic acid nanostructure of any one of claims 1-14, wherein the nucleic acid nanostructure is three-dimensional.
16. The nucleic acid nanostructure of claim 15, wherein the nanostructure has a diameter of 10-500 nm.
17. An engineered nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides.
18. The engineered nucleic acid strand of claim 17, wherein the nucleic acid strand comprises greater than 90% of A, T/U and C nucleotides and less than 10% of G nucleotides.
19. The engineered nucleic acid strand of claim 17, wherein the nucleic acid strand comprises greater than 90% of A, T/U and G nucleotides and less than 10% of C nucleotides.
20. The engineered nucleic acid strand of any one of claims 17-19, wherein the nucleic acid strand comprises less than 5% of one of C, G, T or A nucleotides.

21. The engineered nucleic acid strand of any one of claims 17-20, wherein the nucleic acid strand comprises less than 1-5% of one of C, G, T or A nucleotides.
22. The engineered nucleic acid strand of any one of claims 17-21, wherein that the nucleotide strand sequence comprises nucleotide subsequences, each of which is present only once in the nucleic acid strand.
23. The engineered nucleic acid strand of any one of claims 17-22, wherein the nucleic acid strand has a length of at least 500 nucleotides.
24. The engineered nucleic acid strand of any one of claims 17-22, wherein the nucleic acid strand is a DNA strand.
25. The engineered nucleic acid strand of any one of claims 17-24, wherein the subsequences have a length of 5 to 10 nucleotides.
26. The engineered nucleic acid strand of any one of claims 17-25, wherein the nucleic acid strand has a length of at least 1000 nucleotides.
27. The engineered nucleic acid strand of any one of claims 17-26, wherein the nucleic acid strand has a length of 100-100,000 nucleotides.
28. The engineered nucleic acid strand of any one of claims 17-27, wherein the nucleic acid strand has a C- or G-content of 20%-60% for each subsequences having a length of at least 30 nucleotides.
29. A kit comprising
 - (a) a nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides; and
 - (b) a plurality of oligonucleotides, wherein each oligonucleotide is at least partially complementary to and at least partially bound to the nucleic acid strand.
30. The kit of claim 29 further comprising buffer, monovalent cations and/or divalent cations.

31. A method of producing a nucleic acid nanostructure, comprising incubating, in a reaction vessel,
- (a) a nucleic acid strand having a length of at least 100 nucleotides and having a nucleotide sequence that comprises less than 10% of one of C, G, T or A nucleotides, and
 - (b) a plurality of oligonucleotides, wherein each oligonucleotide is at least partially complementary to the nucleic acid strand,
- under conditions that result in binding of the oligonucleotides to the nucleic acid strand to form a nucleic acid nanostructure.
32. The method of claim 31, wherein the conditions include incubating nucleic acid strand and plurality of oligonucleotides in buffer at a temperature of 4-70 °C for at least 5 minutes.



1/4

FIG. 1

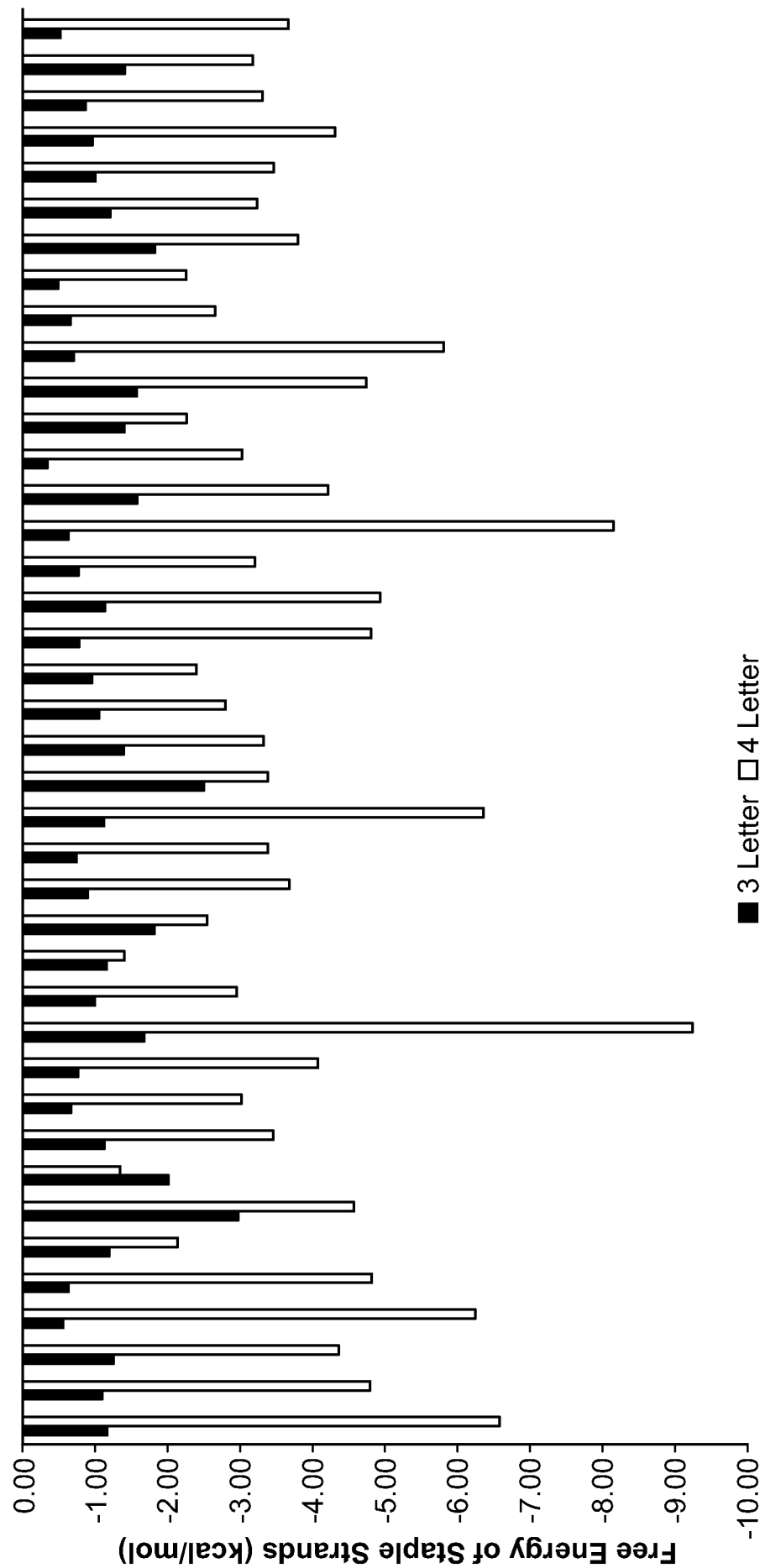


FIG. 2

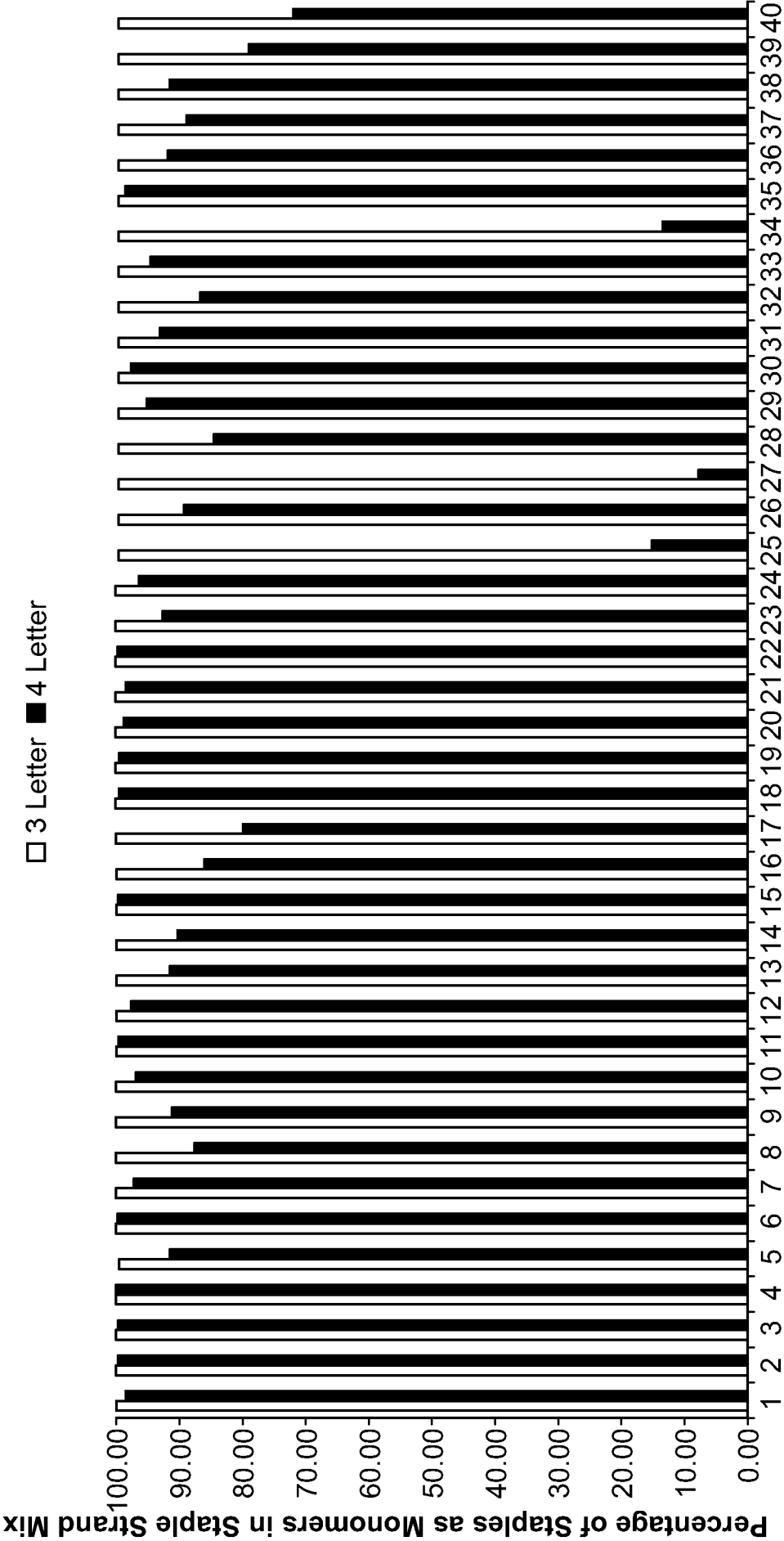
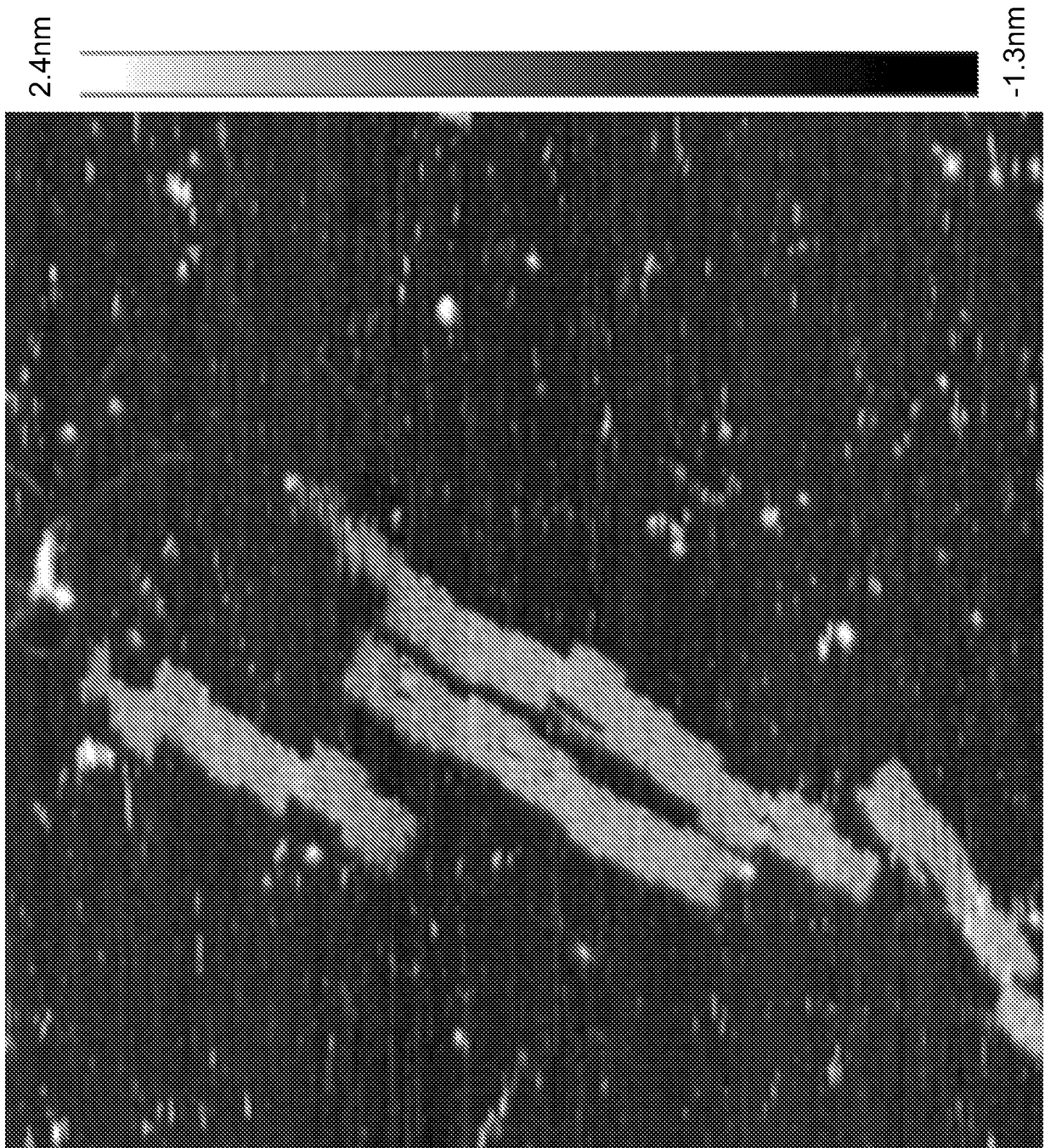


FIG. 3

FIG. 4



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US17/26959

A. CLASSIFICATION OF SUBJECT MATTER

IPC - B82B 3/00; B82Y 40/00; C07H 1/00, 21/04; C12Q 1/68; C12N 15/10, 15/11 (2017.01)

CPC - B82B 3/00; B82Y 5/00, 15/00, 40/00; C07H 21/04; C12N 15/10, 15/11; G01N 21/6486; C12Q 1/68, 1/6874

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	(TSUTSUMI, T et al.) GENESUS: A two-step sequence design program for DNA nanostructure self-assembly. BioTechniques. April 2014, Vol. 56, No. 4; pages 180-185; page 180, 1st column, 1st paragraph; page 182, 1st column, 4th paragraph; page 184, 2nd column, 2nd paragraph; DOI: 10.2144/000114157	1-4, 17-20, 29-32
A	(MYHRVOLD, C et al.) Isothermal Self-Assembly of Complex DNA Structures under Diverse and Biocompatible Conditions. Nano Letters. 11 September 2013, Vol. 13, No. 9; pages 4242-4248; Figure 4; page 4242, 1st column, 1st paragraph; page 4245, 2nd column, 1st paragraph; page 4247, 1st column 1st paragraph; DOI: 10.1021/nl4019512	1-4, 17-20, 29-32
A	(EICKERT, G) Using Modular Preformed DNA Origami Building Blocks to Fold Dynamic 3D Structures. Thesis [online]. 2014 [Retrieved on 11 June 2017]. Retrieved from the Internet: <URL: https://etd.ohiolink.edu/etd.send_file?accession=osu1397742084&disposition=inline>; page 9, 2nd paragraph; page 16, 2nd paragraph; page 25, 2nd paragraph	1-4, 17-20, 29-32
A	US 2014/0303320 A1 (SCHAUS, TE et al.) 9 October 2014; entire document	1-4, 17-20, 29-32

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

22 June 2017 (22.06.2017)

Date of mailing of the international search report

24 JUL 2017

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US17/26959

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 5-16, 21-28
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.