

# US Patent & Trademark Office

## Patent Public Search | Text View

---

United States Patent Application Publication

20250257366

Kind Code

A1

Publication Date

August 14, 2025

Inventor(s)

Kretzmann; Jessica et al.

---

### **DNA ORIGAMI ENCODING FOR GENE EXPRESSION AND CONTRANSFECTION**

---

#### **Abstract**

The present invention relates to a nucleic acid nanostructure comprising at least one scaffold strand and a plurality of staple strands, wherein said nanostructure, preferably said at least one scaffold strand, comprises at least one nucleic acid sequence encoding a gene. The present invention further relates to a composition comprising a nucleic acid nanostructure, and to a collection of nucleic acid sequences or collection of plasmids encoding a nucleic acid nanostructure. Furthermore, the present invention relates to a nucleic acid nanostructure or composition comprising a nucleic acid nanostructure for use in medicine; preferably for use in a method of preventing, treating and/or diagnosing a disease or disorder. The present invention also relates to a method of expressing a gene from a nucleic acid nanostructure, and to a use of a nanostructure or of a composition for gene expression.

---

**Inventors:** Kretzmann; Jessica (Manning, AU), Dietz; Hendrik (Haar, DE)

**Applicant:** Technische Universität München (München, DE)

**Family ID:** 84370200

**Appl. No.:** 19/193742

**Filed:** April 29, 2025

#### **Foreign Application Priority Data**

EP 22207474.2

Nov. 15, 2022

#### **Related U.S. Application Data**

parent WO continuation PCT/EP2023/081932 20231115 PENDING child US 19193742

---

#### **Publication Classification**

**Int. Cl.: C12N15/85 (20060101); C12N15/87 (20060101)**

**U.S. Cl.:**

**CPC C12N15/85 (20130101); C12N15/87 (20130101); C12N2310/531 (20130101); C12N2830/42 (20130101); C12N2830/48 (20130101); C12N2830/50 (20130101)**

---

## **Background/Summary**

### **FIELD OF THE INVENTION**

[0001] The present invention relates to a nucleic acid nanostructure comprising at least one scaffold strand and a plurality of staple strands, wherein said nanostructure, preferably said at least one scaffold strand, comprises at least one nucleic acid sequence encoding a gene. The present invention further relates to a composition comprising a nucleic acid nanostructure, and to a collection of nucleic acid sequences or collection of plasmids encoding a nucleic acid nanostructure. Furthermore, the present invention relates to a nucleic acid nanostructure or composition comprising a nucleic acid nanostructure for use in medicine; preferably for use in a method of preventing, treating and/or diagnosing a disease or disorder. The present invention also relates to a method of expressing a gene from a nucleic acid nanostructure, and to a use of a nanostructure or of a composition for gene expression.

### **BACKGROUND OF THE INVENTION**

[0002] Delivering and expressing genes faces a series of obstacles, including how to package, target, and release the nucleic acid to be delivered. These obstacles become further apparent in the delivery of multi-component systems, where precise and tunable amounts of several genes need to be delivered, such as in the case include CRISPR-based technology and construction of gene circuits, for example. Multiplexing to achieve genome or epigenome editing, transcriptional modulation, and/or construction of gene circuits, offers vast potential for tailoring genetic networks for therapeutic (re)programming, bio-production, and basic research. However, despite rapid progress in these fields, simultaneous delivery and expression of multiple genes remains challenging, particularly when looking to deliver in vivo.

[0003] Artificial structures formed from nucleic acids, such as DNA origami, have been discussed as having enormous potential for the field of biotechnology. DNA origami enables long single-stranded DNA to be packaged into a compact structure with an unparalleled level of structural programmability and homology, spatial addressability, and biocompatibility. Further, multi-component assemblies of different DNA origami structures can be realized, with structures composed of 220 monomers and reaching over 1 GDa in size. However, while there has been significant progress in the use of DNA origami for applications such as drug delivery, sensing and imaging, there has been limited development of DNA origami for gene therapy. To date, DNA origami has only been utilized as hybrids, with either RNA or proteins, for gene therapeutic studies.

[0004] There remains the need for tools for efficiently expressing genes, such as mammalian genes. Particularly, there remains the need for assemblies encoding two or more genes in controlled stoichiometric ratios. Furthermore, there remains the need for simultaneous delivery and expression of one or more genes, e.g. multiple genes. Moreover, there remains the need for tools that efficiently deliver and allow to express genes in vivo.

### **SUMMARY OF THE INVENTION**

[0005] In the following, the elements of the invention will be described. These elements are listed with specific embodiments, however, it should be understood that they may be combined in any manner and in any number to create additional embodiments. The variously described examples

and preferred embodiments should not be construed to limit the present invention to only the explicitly described embodiments. This description should be understood to support and encompass embodiments which combine two or more of the explicitly described embodiments or which combine the one or more of the explicitly described embodiments with any number of the disclosed and/or preferred elements. Furthermore, any permutations and combinations of all described elements in this application should be considered disclosed by the description of the present application unless the context indicates otherwise.

[0006] In a first aspect, the present invention relates to a nucleic acid nanostructure comprising at least one scaffold strand and a plurality of staple strands, wherein said nanostructure, preferably said at least one scaffold strand, comprises at least one nucleic acid sequence encoding a gene.

[0007] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand, comprises a first nucleic acid sequence encoding a gene and a second nucleic acid sequence encoding a gene; wherein, optionally, said nucleic acid nanostructure, preferably said at least one scaffold strand, comprises a plurality of nucleic acid sequences encoding a gene.

[0008] In one embodiment, said nucleic acid nanostructure comprises a first subunit and a second subunit; wherein, preferably, said first subunit and said second subunit each comprise a nucleic acid sequence encoding a gene.

[0009] In one embodiment, said nanostructure comprises an enhancer staple strand having a length in a range of from about 60 to about 250 nucleic acid bases, preferably from about 80 to about 220 nucleic acid bases, more preferably from about 90 to about 200 nucleic acid bases; [0010] wherein, optionally, said enhancer staple strand is configured such that it binds to said scaffold strand at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene.

[0011] In one embodiment, said nanostructure comprises an enhancer staple strand having a length in a range of from about 60 to about 250 nucleic acid bases, preferably from about 80 to about 220 nucleic acid bases, more preferably from about 90 to about 200 nucleic acid bases; [0012] wherein, optionally, said enhancer staple strand comprises a nucleic acid sequence which is complementary to a nucleic acid sequence of said scaffold strand, wherein said nucleic acid sequence of said scaffold strand is located at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene.

[0013] In one embodiment, said nanostructure, preferably said scaffold strand and/or at least one staple strand of said plurality of staple strands, comprises a nucleic acid sequence configured to form a loop structure, preferably an inverted-terminal repeat nucleic acid sequence configured to form a hairpin. In one embodiment, said nanostructure, preferably said scaffold strand and/or at least one staple strand of said plurality of staple strands, comprises a loop structure, preferably a hairpin.

[0014] In one embodiment, said nucleic acid sequence configured to form a loop structure is configured such that a loop is formed at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene.

[0015] In one embodiment, said nanostructure, preferably said scaffold strand and/or at least one staple strand of said plurality of staple strands, comprises a loop structure, preferably a hairpin, at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene. In one embodiment, said nanostructure, preferably said at least one scaffold strand, comprises at least one nuclear targeting sequence, preferably a DNA nuclear targeting sequence, more preferably a simian virus 40 DNA nuclear targeting sequence; wherein, optionally, said nanostructure, preferably said at least one scaffold strand, comprises a plurality of nuclear targeting sequences.

[0016] In one embodiment, said nanostructure, preferably said at least one scaffold strand, comprises a promoter such as a CMV promoter, a terminator, a polyadenylation signal sequence, an intron, a kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element.

[0017] In one embodiment, said nanostructure has an aspect ratio in the range of from about 1:1 to about 1000:1, preferably 1.5:1 to about 20:1, more preferably from about 2:1 to about 15:1.

[0018] In one embodiment, said scaffold strand comprises said at least one nucleic acid sequence encoding a gene, a promoter, and a terminator; optionally further comprises a nucleic acid sequence configured to form a loop structure, an intron, a DNA nuclear targeting sequence, a polyadenylation signal sequence, a Kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element.

[0019] In one embodiment, said scaffold strand comprises said at least one nucleic acid sequence encoding a gene, a promoter, and a terminator; wherein said scaffold strand optionally further comprises a loop structure, an intron, a DNA nuclear targeting sequence, a polyadenylation signal sequence, a Kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element.

[0020] In one embodiment, the nucleic acid sequence encoding a gene is a nucleic acid sequence encoding a eukaryotic gene, preferably a nucleic acid sequence encoding a mammalian gene, more preferably a nucleic acid sequence encoding a human gene.

[0021] In a further aspect, the present invention relates to a composition, preferably pharmaceutical composition, comprising a nucleic acid nanostructure as defined herein.

[0022] In a further aspect, the present invention relates to a collection of nucleic acid sequences or collection of plasmids encoding a nucleic acid nanostructure as defined herein.

[0023] In a further aspect, the present invention relates to a nucleic acid nanostructure, as defined herein, or composition, as defined herein, for use in medicine; preferably for use in a method of preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder; optionally for use in gene therapy and/or immunotherapy.

[0024] In one embodiment, the nucleic acid nanostructure, as defined herein, or composition, as defined herein, is for use in gene therapy and/or immunotherapy.

[0025] In a further aspect, the present invention relates to a method of expressing a gene from a nucleic acid nanostructure, preferably a nucleic acid nanostructure as defined herein, comprising

[0026] i) providing a nucleic acid nanostructure comprising at least one nucleic acid sequence encoding a gene, preferably a nucleic acid nanostructure as defined herein; [0027] ii) delivering said nucleic acid nanostructure provided in step i) to a cell; wherein, preferably, said delivering comprises transfecting or transforming said cell; [0028] iii) allowing said cell to express said gene;

[0029] wherein, optionally, said providing in step i) comprises providing a plasmid, preferably a phagemid, or a collection of plasmids, preferably a collection of phagemids, wherein said plasmid or collection of plasmids encodes said nucleic acid nanostructure, and preparing said nucleic acid nanostructure using said plasmid or collection of plasmids, preferably by using bacteriophages.

[0030] In a further aspect, the present invention relates to a use of a nanostructure, as defined herein, or of a composition, as defined herein, for gene expression, preferably for in vitro gene expression.

[0031] In a further aspect, the present invention relates to a method of preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder, comprising administering a nanostructure, as defined herein, or a composition, as defined herein, to a patient in need thereof.

[0032] In one embodiment, said method of preventing, treating, and/or diagnosing a disease or disorder is a method of gene therapy and/or immunotherapy.

[0033] In one embodiment, said administering comprises administering an effective amount of a nanostructure, as defined herein, and/or a composition, as defined herein, to a patient in need thereof.

[0034] In a further aspect, the present invention relates to a use of a nanostructure, as defined herein, or of a composition, as defined herein, for the manufacture of a medicament, e.g. a medicament for preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder; optionally for gene therapy and/or immunotherapy.

[0035] In one embodiment, the nanostructure, as defined herein, or the composition, as defined

herein, is used for the manufacture of a medicament for gene therapy and/or immunotherapy.

## DETAILED DESCRIPTION

[0036] It is an aim of the invention to provide tools for efficiently expressing genes, such as eukaryotic genes e.g. mammalian genes. Particularly, it is an aim of the invention to provide nanostructures encoding two or more genes in controlled stoichiometric ratios. Furthermore, it is an aim of the invention to simultaneously deliver and express one or more genes, e.g. multiple genes. Moreover, it is an aim of the invention to efficiently deliver and express genes in vitro, ex vivo, and in vivo.

[0037] The inventors have successfully expressed genes from encoded DNA origami. Particularly, the inventors synthesized a library of custom ssDNA scaffolds for mammalian gene expression e.g. using bacteriophage-based production. The inventors found that genes were readily expressed from nucleic acid nanostructure, regardless of gene position in nanostructure, or nanostructure shape. The inventors found that gene expression can be even further enhanced with nucleic acid sequences configured to form a loop structure, e.g. adeno-associated virus-inspired inverted-terminal repeat (ITR) hairpin sequences, either upstream of the expression cassette, or flanking either side the expression cassette with the loop structure, e.g. ITR hairpin, featured on the staples. Overall, the inventors demonstrate highly efficient gene expression encoded within the nucleic acid nanostructure of the invention, particularly DNA origami structures of the invention.

[0038] The inventors have successfully delivered and expressed genes from encoded ssDNA scaffolds and customized DNA origami objects, as shown in FIG. 1a. The inventors found that nucleic acid nanostructures of the invention readily unfold within the intracellular environment and efficiently express genes from the ssDNA scaffold strand, and that gene expression can be further optimized with targeted staple design. The inventors further produced and tested a library of ssDNA scaffolds optimizing for increased gene expression, high-yield scaffold production and purity, and origami folding quality. Overall, the inventors find that, surprisingly, inclusion of at least one loop structure such as ITR secondary DNA structure, preferably upstream of the expression cassette, and/or the inclusion of one or more nuclear targeting sequences such as DTS sequences, e.g. three SV40 DTS sequences, enables a highly efficient and robust gene expression without compromising scaffold production or origami folding quality. Finally, the inventors demonstrate efficient and controlled assembly of gene-encoded origami structures in stoichiometric ratios. These “plug and play” architected origami's enabled successful co-delivery and expression of an array of genes with unprecedented control.

[0039] Advantageously, the nanostructures of the invention exploit both the ability to encode for genetic information, together with the unique design possibilities of DNA origami structures. Thus, advantageously, the nanostructures of the invention allow to encode and express an array of genes in controlled stoichiometries.

[0040] The present invention relates to a nucleic acid nanostructure comprising at least one scaffold strand and a plurality of staple strands, wherein said nanostructure, preferably said at least one scaffold strand, comprises at least one nucleic acid sequence encoding a gene. Advantageously, one or more genes can be efficiently expressed using a nanostructure of the invention.

[0041] The term “nanostructure”, as used herein, relates to a nucleic acid nanostructure, preferably a DNA origami structure which is composed of one or more DNA origami subunits. In one embodiment, the nanostructure comprises or consists of a DNA origami structure. Readily available nucleic acid nanostructure techniques e.g. DNA origami techniques, which involve comparatively less complex procedures to assemble nanostructures than standard nanomanufacturing techniques, may be used to manufacture the nanostructure. In one embodiment, the nanostructure is at least partially manufactured using DNA origami techniques. Owing to the self-assembly of DNA origami structures and the readily available software for designing the corresponding scaffold and staple strands, this is a comparatively less complex manufacturing process compared to standard nanomanufacturing techniques. In one embodiment, a nanostructure of the invention is a DNA

origami structure. In one embodiment, the nanostructure of the invention has a maximum length smaller than 1000 nm, e.g. in the range of from about 10 nm to about 150 nm, preferably about 20 nm to about 100 nm.

[0042] In one embodiment, a nucleic acid nanostructure of the invention, e.g. a DNA origami structure, comprises at least one scaffold strand and a plurality of staple strands, e.g. single-stranded oligonucleotide staple strands. The term “staple strand”, as used in this invention, shall refer to a single-stranded oligonucleotide molecule, which is at least partially complementary to a scaffold strand. In one embodiment, when referring to a “staple”, a staple strand of a plurality of staple strands is meant. In general, staple strands can be used to introduce, e.g., coupling sites into a DNA origami structure and/or a DNA origami subunit. The term “plurality of staple strands”, as used herein, relates to a plurality, e.g. at least three, staple strands. For example, a plurality of staple strands may relate to at least 3, 4, 5, 6, 7, 8, 9, 10, or more staple strands. In one embodiment, staple strands, particularly the staple strands of the plurality of staple strands, have a length of from 20 to 100 nucleic acid bases. In one embodiment, the staple strands of the plurality of staple strands have a length of from 20 to 80 nucleic acid bases, and the enhancer staple strand(s) has/have a length of from 90 to 250 nucleic acid bases. In one embodiment, the enhancer staple strand(s) differs from the staple strands of the plurality of staple strands in that it is at least 5, preferably at least 10 nucleic acid bases longer than each of the staple strands of the plurality of staple strands.

[0043] The term “scaffold strand” is a nucleic acid strand, preferably DNA strand, such as a single-stranded nucleic acid strand e.g. single-stranded polynucleotide strand. In one embodiment, a scaffold strand makes up and/or traverses the main part of a DNA origami structure and/or a DNA origami subunit. In one embodiment, a scaffold strand has a length of from 100 to 20000 nucleic acid bases, preferably from 120 to 15000 nucleic acid bases, more preferably from 260 to 11000 nucleic acid bases, e.g. from 1000 to 11000 nucleic acid bases. In one embodiment, said scaffold strand is a circular or linear scaffold strand. In one embodiment, said scaffold strand is a circular ssDNA scaffold strand. In one embodiment, the scaffold strand has a length of from 260 to 20000 nucleic acid bases, the enhancer staple strand(s) has/have a length of from 90 to 250 nucleic acid bases, and the staple strands of the plurality of staple strands have a length of from 20 to 80 nucleic acid bases.

[0044] The nanostructure, e.g. DNA origami structure, may comprise at least one scaffold strand, i.e. single-stranded polynucleotide scaffold DNA with a known sequence. The DNA origami structure may further comprise a plurality of single-stranded oligonucleotide staple strands, wherein each staple strand may be at least partially complementary to at least one scaffold strand. Further, each of the staple strands may be configured to bind to the at least one scaffold strand, wherein the at least one scaffold strand may be folded and/or arranged such that the desired nanostructure may be formed. The term “strand”, as used herein, relates to a nucleic acid strand, e.g. a DNA and/or RNA strand, preferably a DNA strand. A three-dimensional nanostructure may be realized using DNA origami, i.e. by combining scaffold strands and staple strands to form the required portions and the overall device. Such designs may for example be performed using software such as caDNAno. That is, a nanostructure comprising multiple portions may in some embodiments be made out of one scaffold strand, whereas in other embodiments portions of a nanostructure may be constructed utilizing a plurality of scaffold strands.

[0045] In one embodiment, the shape of a nanostructure may be any shape, for example, a brick, a rod, a triangular shape, a round shape, a cuboid, i.e. a rectangular shape, a star shape, or any other shape. The nanostructure of this invention can be of any length. In a preferred embodiment, the nanostructure comprises a maximum length, and in a particularly preferred embodiment, the maximum length is smaller than 1000 nm, preferably smaller than 500 nm, such as around 100 nm, or smaller. In one embodiment, the terms “nanostructure”, “nano-object”, and “nucleic acid nanostructure” are used interchangeably.

[0046] In one embodiment, each of the staple strands is configured to bind to at least one of the at

least one scaffold strand in at least one, preferably two or more distinct places. In one embodiment, a nucleic acid nanostructure of the invention, e.g. a DNA origami, comprises a scaffold strand and one or more staple strands. In one embodiment, a nucleic acid nanostructure of the invention comprises  $\leq 100$  staple strands. In one embodiment, the nucleic acid nanostructure of the invention comprises  $\geq 10$  DNA strands, preferably  $\geq 15$  DNA strands, e.g. at least one scaffold strand and at least nine staple strands.

[0047] The terms “DNA origami structure”, “DNA origami” and “DNA origami objects”, as used herein, relate to a nanostructure that comprises DNA as a building material to make nanoscale shapes. Preparing and/or providing a DNA origami involves folding of one or more scaffold DNA strands into a defined shape using a plurality of rationally designed staple DNA strands, e.g. by self-assembly. A scaffold strand is typically longer than a staple strand. The nucleic acid sequences of the staple strands are designed such that the staple strands hybridize to defined portions of the scaffold strands and, due to the hybridization, a particular shape of the nanostructure is obtained.

[0048] In one embodiment, the term “nucleic acid”, as used herein, relates to a nucleotide sequence, such as ribonucleic acid or deoxyribonucleic acid. In a preferred embodiment, the nucleic acid nanostructure is a DNA nanostructure. In a preferred embodiment, the nucleic acid nanostructure of the invention is a DNA origami structure. In a preferred embodiment, the nucleic acid nanostructure comprises or consists of a DNA origami structure. In one embodiment, the nucleic acid nanostructure is provided in the form of a DNA origami structure. An advantage of a nucleic acid nanostructure such as a DNA origami structure is that nucleic acid nanostructures such as DNA origami structures may comprise a plurality of genes, particularly in a desired stoichiometry, due to their rational design. Advantageously, with nucleic acid nanostructures such as DNA origami structures, the number of genes and their expression can be precisely controlled. Furthermore, the arrangement of the genes and further nucleic acid sequences of interest can be precisely controlled. A further advantage of nucleic acid nanostructures such as DNA origami structures is that they can be stabilized against nucleases. In one embodiment, the nanostructure is configured such that it is stabilized against nucleases. A further advantage of DNA origami structures is that the assembly, e.g. self-assembly, and purification of the structure(s) are more robust and simple compared to non-DNA origami nucleic acid nanostructures such as DNA tetrahedrons or RNA assemblies. All embodiments described herein with respect to “a/the/said nucleic acid nanostructure” or “a/the/said nucleic acid nanostructure of the invention” are meant to be understood as also relating to the nanostructure(s) comprised by the composition of the invention, to the nanostructure used in accordance with the invention, and to the nanostructure(s) provided in any method of the invention.

[0049] The term “nucleic acid sequence encoding a gene”, as used herein, relates to a nucleic acid sequence encoding any gene of interest, e.g. a gene involved in a pathological pathway, a gene suitable for a vaccination, and/or a CRISPR-based gene. In one embodiment, the gene is selected from prokaryotic genes, viral genes, and eukaryotic genes. In one embodiment, the gene is selected from prokaryotic genes such as CRISPR-based genes and eukaryotic genes such as human genes. In a preferred embodiment, the gene is a eukaryotic gene, preferably a mammalian gene, e.g. a human gene. An advantage of the nucleic acid nanostructure of the invention is that it allows to express mammalian genes, e.g. different mammalian genes with defined stoichiometries. A further advantage is that the nanostructure allows to express prokaryotic genes, e.g. for CRISPR-based gene editing, and viral genes, e.g. for DNA- or RNA-based vaccinations. In a preferred embodiment, the nucleic acid sequence encoding a gene encodes a eukaryotic gene, preferably mammalian gene, more preferably a human gene. In one embodiment, the nucleic acid nanostructure of the invention, preferably said at least one scaffold strand, comprises a first nucleic acid sequence encoding a gene and a second nucleic acid sequence encoding a gene; wherein, optionally, said nucleic acid nanostructure, preferably said at least one scaffold strand, comprises a plurality of nucleic acid sequences encoding a gene. Advantageously, the nanostructure of the

invention comprising a first nucleic acid sequence encoding a gene and a second nucleic acid sequence encoding a gene, e.g. comprising a plurality of nucleic acid sequences encoding a gene, allows to efficiently express genes of interest in a defined stoichiometry. In one embodiment, the first nucleic acid sequence encoding a gene and the second nucleic acid sequence encoding a gene each encode a eukaryotic gene, preferably a mammalian gene, more preferably a human gene. The nanostructure of the invention advantageously allows to express an array of genes in controlled stoichiometries. In one embodiment, the first nucleic acid sequence encoding a gene and the second nucleic acid sequence encoding a gene may encode the same gene or different genes. In one embodiment, the nanostructure of the invention comprises one or several copies of a gene of interest, e.g. one or more nucleic acid sequences encoding a gene, such as one or more first nucleic acid sequence(s) encoding a gene and one or more second nucleic acid sequence(s) encoding a gene.

[0050] In one embodiment, the nucleic acid nanostructure comprises a plurality of nucleic acid sequences encoding a gene. In one embodiment, each nucleic acid sequence encoding a gene of said plurality of nucleic acid sequences encoding a gene encodes for a different gene. In one embodiment, each nucleic acid sequence encoding a gene of said plurality of nucleic acid sequences encoding a gene encodes for a gene different from the genes encoded by the other nucleic acid sequences encoding a gene of said plurality of nucleic acid sequences encoding a gene. Thus, the plurality of nucleic acid sequences encoding a gene may comprise a plurality of encoded genes. In one embodiment, the nucleic acid sequence(s) encoding a gene is(are) positioned at any site of the nucleic acid nanostructure, preferably is(are) positioned at any site of the scaffold strand.

[0051] In one embodiment, said nucleic acid nanostructure comprises a first subunit and a second subunit; wherein, preferably, said first subunit and said second subunit each comprise a nucleic acid sequence encoding a gene. The advantage of a nucleic acid nanostructure comprising a first subunit and a second subunit, e.g. a nucleic acid nanostructure comprising a plurality of subunits, is that each of the subunits may comprise one or more nucleic acid sequences encoding a gene, and thus the stoichiometry of multiple nucleic acid sequences encoding a gene can be rationally designed. In one embodiment, the nucleic acid nanostructure comprises a first subunit and a second subunit, optionally a plurality of subunits, which are stacked. In one embodiment, the nucleic acid nanostructure comprises two or more stacked subunits comprising a first subunit and the second subunit. In one embodiment, the first subunit and the second subunit each comprises a nucleic acid sequence encoding a gene, wherein the nucleic acid sequence encoding a gene comprised by the first subunit is the same or is different from the nucleic acid sequence encoding a gene comprised by the second subunit. In one embodiment, the first subunit and the second subunit encode the same or different genes. In one embodiment, the first subunit comprises a first nucleic acid sequence encoding gene and the second subunit comprises a second nucleic acid sequence encoding a gene. In one embodiment, the first subunit and the second subunit, optionally each subunit of a plurality of subunits, each comprise a scaffold strand and a plurality of staple strands. In one embodiment, the nucleic acid nanostructure comprises a first subunit and a second subunit, wherein each subunit comprises a scaffold strand comprising a nucleic acid sequence encoding a gene and wherein each subunit comprises a plurality of staple strands, optionally wherein at least one subunit comprises an enhancer staple strand. The subunits of the nucleic acid nanostructure may be connected by any means, e.g. by shape complementarity, by nucleotide base-stacking interactions, by nucleic acid mediated interaction such as by base pairing, and/or by covalent binding such as disulfide bridges. Advantageously, an assembly of multiple DNA origami subunits containing multiple genes can be achieved in controlled stoichiometric ratios (FIG. 6) to enable the simultaneous delivery of multiple components. This is of particular importance in areas such as CRISPR-based technologies, for gene/base editing or epigenetic modulation. Further, production of virus-like particles (VLPs), lentiviruses and adeno-associated viruses, etc. additionally requires delivery of multiple components in controlled ratios.



[0052] In one embodiment, said nanostructure comprises an enhancer staple strand having a length in a range of from about 60 to about 250 nucleic acid bases, preferably from about 80 to about 220 nucleic acid bases, more preferably from about 90 to about 200 nucleic acid bases, e.g. about 154 bases. In one embodiment, said enhancer staple strand is configured such that it binds to said scaffold strand at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene. The inventors have found that an enhancer staple strand stabilizes the nanostructure and increases gene expression from the nanostructure. Furthermore, the inventors have found that, surprisingly, an efficient expression of genes such as mammalian genes can be achieved with a nanostructure comprising an enhancer staple strand. In one embodiment, the term "enhancer staple strand" relates to a staple strand having a length of at least 60 nucleic acid bases, preferably of at least 80 nucleic acid bases, more preferably of at least 90 nucleic acid bases, e.g. having a length in a range of from about 90 to about 250 nucleic acid bases. In one embodiment, the terms "enhancer staple strand", "staple strand having a length of at least 60 nucleic acid bases", "continuous staple strand", and "stabilizing staple strand" are used interchangeably. In one embodiment, the enhancer staple strand binds to said scaffold strand such that one or multiple continuous double-helical domains with the scaffold are formed which comprise at least 10 nucleic acid base pairings, preferably at least 15 nucleic acid base pairings, more preferably at least 20 nucleic acid base pairings, even more preferably at least 30 nucleic acid base pairings, even more preferably at least 80, 85, or 90 nucleic acid base pairings. In one embodiment, enhancer staple strands provide a long region of dsDNA to help with recognition and binding of the polymerase and/or other helper proteins needed for gene expression.

[0053] The staple strands of the plurality of staple strands typically cross from one region of a scaffold strand over to other regions of the scaffold strand, thereby creating a folding pattern. The staple strands of the plurality of staple strands typically cross multiple regions of the scaffold strand. In contrast thereto, the enhancer staple strand performs less or none of these typical crosses, such that it remains continuous along one region or two regions of the scaffold strand. In one embodiment, the enhancer staple strand binds to said scaffold strand along one region or two regions of the scaffold strand. In one embodiment, the enhancer staple strand binds to said scaffold strand along one region or two regions of consecutive nucleic acids of the nucleic acid sequence of the scaffold strand. In one embodiment, the enhancer staple strand comprises or consists of one or two parts, wherein, if the enhancer staple strand comprises or consists of one part, at least 90% of the nucleic acid bases of said one part bind to a region of said scaffold strand, and wherein, if the enhancer staple strand comprises or consists of two parts, at least 90% of the nucleic acid bases of the first part of said two parts bind to a first region of said scaffold strand and at least 90% of the nucleic acid bases of the second part of said two parts bind to a second region of said scaffold strand. In one embodiment, the enhancer staple strand is arranged such that it is or at least a part of it is substantially parallel to a longitudinal extension of said nanostructure. In one embodiment, the term "enhancer" in the expression "enhancer staple strand" means that the enhancer staple strand binds to at least 10, preferably at least 15, more preferably at least 20, even more preferably at least 30, even more preferably at least 80, 85, or 90 consecutive nucleic acids of a nucleic acid sequence of said scaffold strand; wherein, preferably, the assembling of the nanostructure and the gene expression therefrom are enhanced. In one embodiment, said scaffold strand comprises said nucleic acid sequence encoding a gene, and said enhancer staple strand binds to at least 10, preferably at least 15, more preferably at least 20, even more preferably at least 30, even more preferably at least 80, 85, or 90 consecutive nucleic acids of a nucleic acid sequence of said scaffold strand at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene. In one embodiment, said enhancer staple strand binds to at least 10, preferably at least 15, more preferably at least 20, even more preferably at least 30, even more preferably at least 80, 85, or 90 consecutive nucleic acids of a nucleic acid sequence of said scaffold strand at a 5' end of said nucleic acid sequence encoding a gene and said enhancer staple strand binds to at least 10, preferably at least 15, more preferably at

least 20, even more preferably at least 30, even more preferably at least 80, 85, or 90 consecutive nucleic acids of a nucleic acid sequence of said scaffold strand at a 3' end of said nucleic acid sequence encoding a gene. For example, the enhancer staple strand may have a first part, e.g. a first half, and second part, e.g. a second half, and said first part, e.g. first half, binds to at least 10, preferably at least 15, more preferably at least 20, even more preferably at least 30, even more preferably at least 80, 85, or 90 consecutive nucleic acids of a nucleic acid sequence of said scaffold strand at a 5' end of said nucleic acid sequence encoding a gene and said second part, e.g. a second half, binds to at least 10, preferably at least 15, more preferably at least 20, even more preferably at least 30, even more preferably at least 80, 85, or 90 consecutive nucleic acids of a nucleic acid sequence of said scaffold strand at a 3' end of said nucleic acid sequence encoding a gene. The inventors have surprisingly found that a enhancer staple strand placed flanking either side of the expression cassette leads to a significant increase in transfection efficiency and gene expression. In one embodiment, the scaffold strand comprises said nucleic acid sequence encoding a gene and a polyadenylation signal sequence, wherein, optionally, an enhancer staple strand binds to said scaffold strand at a 5' end of said nucleic acid sequence encoding a gene and at a 3' end of said polyadenylation signal sequence. In one embodiment, the enhancer staple strand is configured such that it binds to said scaffold strand such that a circular structure, e.g. a circular scaffold strand or a circular expression cassette, is formed.

[0054] In one embodiment, said nanostructure, preferably said scaffold strand and/or at least one staple strand of said plurality of staple strands, comprises a nucleic acid sequence configured to form a loop structure, preferably an inverted-terminal repeat nucleic acid sequence configured to form a hairpin. For example, a loop structure is formed when two regions of nucleic acid strands, e.g. two regions of the same strand, typically complementary in nucleotide sequence when read in opposite directions, base-pair to form a double helix that ends in an unpaired loop. For example, an inverted-terminal repeat nucleic acid sequence may be a single stranded sequence of nucleotides followed downstream by its reverse complement. The intervening sequence of nucleotides between an initial sequence and the reverse complement thereof can be any length including zero.

Advantageously, a nanostructure comprising a nucleic acid sequence configured to form a loop structure, e.g. an inverted-terminal repeat nucleic acid sequence, such as an adeno-associated virus-inspired inverted-terminal repeat (ITR) hairpin sequence, shows highly effective gene expression. In one embodiment, the nanostructure, e.g. said scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a nucleic acid sequence configured to form a loop structure, preferably an inverted-terminal repeat nucleic acid sequence such as an adeno-associated virus-inspired inverted-terminal repeat hairpin sequence. In one embodiment, said nucleic acid sequence configured to form a loop structure is configured such that a loop is formed at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene. In one embodiment, said scaffold strand comprises a nucleic acid sequence configured to form a loop structure, preferably upstream of said nucleic acid sequence encoding a gene, more preferably upstream of an expression cassette comprising said nucleic acid sequence encoding a gene. In one embodiment, a staple strand, e.g. a staple strand of a plurality of staple strands, comprises a nucleic acid sequence configured to form a loop structure, preferably a staple strand flanking either side of said nucleic acid sequence encoding a gene, more preferably flanking either side of an expression cassette comprising said nucleic acid sequence encoding a gene. In one embodiment, the nucleic acid sequence configured to form a loop structure comprises or consists of a sequence as defined in any of SEQ ID NO. 13-14. In one embodiment, said nanostructure comprises a nucleic acid sequence encoding a loop structure, preferably an inverted-terminal repeat nucleic acid sequence encoding a hairpin. In one embodiment, said nanostructure comprises a loop structure, preferably a hairpin. In one embodiment, the term "nucleic acid sequence encoding a loop structure", as used herein, relates to a nucleic acid sequence configured to form a loop structure. In one embodiment, the nucleic acid sequence configured to form a loop structure is configured such that the loop is

formed by the nucleic acid sequence configured to form a loop structure alone, or by the nucleic acid sequence configured to form a loop structure and a further nucleic acid sequence e.g. a further nucleic acid sequence configured to form a loop structure. In one embodiment, the nucleic acid nanostructure comprises a first nucleic acid sequence configured to form a loop structure and a second nucleic acid sequence configured to form a loop structure. In one embodiment, the first and second nucleic acid sequences configured to form a loop structure each form a loop structure and/or jointly form a loop structure. The inventors have surprisingly found that an enhancement of gene expression efficiency can be achieved by including nucleic acid sequences encoding a loop structure such as adeno-associated virus-inspired inverted-terminal repeat (ITR) hairpin sequences, for example either upstream of the expression cassette or flanking either side of the expression cassette with the loop structure featured on the staple strands. In one embodiment, said nucleic acid sequence configured to form a loop structure is located upstream or downstream of the nucleic acid sequence encoding a gene, particularly upstream or downstream of an expression cassette. In one embodiment, said loop structure is formed upstream or downstream of the nucleic acid sequence encoding a gene, particularly upstream or downstream of an expression cassette. In one embodiment, said nucleic acid sequence configured to form a loop structure is located on one or two staple strands binding to said scaffold strand upstream or downstream of the nucleic acid sequence encoding a gene, particularly upstream or downstream of an expression cassette.

[0055] In one embodiment, said nanostructure, preferably said at least one scaffold strand, comprises at least one nuclear targeting sequence, preferably a DNA nuclear targeting sequence, more preferably a simian virus 40 DNA nuclear targeting sequence; wherein, optionally, said nanostructure, preferably said at least one scaffold strand, comprises a plurality of nuclear targeting sequences. In one embodiment, the nuclear targeting sequence comprises or consists of a sequence as defined in SEQ ID NO. 21. DNA nuclear targeting sequences (DTSs) are consensus motifs recognized by transcription factors, and can be used to shuttle DNA from the cytosol through nuclear pores to the nucleus. For instance, the simian virus 40 DNA nuclear targeting sequence (SV40 DTS) can be used as a DTS since it is recognized by various TFs. Advantageously, a nanostructure comprising one or more DTS sequences, such as three SV40 DTS sequences, enables a highly efficient and robust gene expression. In one embodiment, the nucleic acid nanostructure comprises one, two, or three nuclear targeting sequences, preferably DNA nuclear targeting sequences. The inventors have surprisingly found that the transfection efficiency can be further increased with inclusion of nuclear targeting sequences, e.g. DNA nuclear targeting sequences such as the SV40 (simian vacuolating virus 40) DTS, within the nanostructure, preferably the scaffold strand. The inventors have found a maximum effect with an inclusion of 1-3 nuclear targeting sequence repeats such as 1-3 SV40 sequence repeats, for both dividing and non-dividing (arrested) cells (FIG. 5).

[0056] In one embodiment, said nanostructure, preferably said at least one scaffold strand, comprises a promoter such as a CMV promoter, a terminator, a polyadenylation signal sequence, an intron, a kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element. Promoters are typically sequences of DNA to which proteins bind to initiate transcription of a single RNA transcript from the DNA downstream of the promoter, e.g. bacterial or eukaryotic promoters such as mammalian promoters. Terminators, particularly transcription terminators, are typically a section of nucleic acid sequence that marks the end of a gene or operon during transcription, e.g. mammalian terminators such as SV40, hGH, BGH, and rbGlob terminators. In one embodiment, the terminator is a SV40 terminator, hGH terminator, BGH terminator, or rbGlob terminator. For example, the terminator mediates transcriptional termination by providing signals in the newly synthesized transcript RNA that trigger processes which release the transcript RNA from the transcriptional complex. Polyadenylation is the addition of a poly(A) tail to an RNA transcript, typically a messenger RNA (mRNA). A poly(A) tail enhances the nuclear export, translation and stability of mRNA. Polyadenylation signal sequences typically comprise a consensus sequence for

the addition of a poly(A) tail (polyadenylation) and/or a terminator sequence. In one embodiment, the polyadenylation signal sequence comprises an AAUAAA-motif. In one embodiment, the polyadenylation signal sequence comprises or consists of a sequence as defined in SEQ ID NO. 19. The kozak sequence may be a nucleic acid motif that functions as the protein translation initiation site in the transcript, e.g. an eukaryotic mRNA transcript. In one embodiment, the kozak sequence comprises or consists of a sequence as defined in SEQ ID NO. 16. The woodchuck hepatitis virus posttranscriptional regulatory element may be a DNA sequence that, when transcribed, creates a tertiary structure enhancing expression. In one embodiment, the woodchuck hepatitis virus posttranscriptional regulatory element comprises or consists of a sequence as defined in SEQ ID NO. 17.

[0057] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a nucleic acid sequence configured to form a loop structure, preferably an inverted-terminal repeat nucleic acid sequence such as an adeno-associated virus-inspired inverted-terminal repeat hairpin sequence, and a promoter such as a CMV promoter, a terminator, a polyadenylation signal sequence, an intron, a kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0058] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) an inverted-terminal repeat nucleic acid sequence such as an adeno-associated virus-inspired inverted-terminal repeat hairpin sequence, and a promoter, such as a CMV promoter, a terminator, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, a kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0059] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) an adeno-associated virus-inspired inverted-terminal repeat hairpin sequence, and a promoter, such as a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, a kozak sequence, and a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0060] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) an adeno-associated virus-inspired inverted-terminal repeat hairpin sequence, and a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, a kozak sequence, and a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0061] In one embodiment, said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand, wherein said nucleic acid nanostructure comprise(s) an adeno-associated virus-inspired inverted-terminal repeat hairpin sequence, and a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, a kozak sequence, and a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE).

[0062] In one embodiment, said nucleic acid nanostructure is a ssDNA scaffold strand, wherein said nucleic acid nanostructure comprise(s) an adeno-associated virus-inspired inverted-terminal repeat hairpin sequence, a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, a kozak sequence, and a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE).

[0063] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a promoter, such as a CMV promoter, a terminator, a polyadenylation signal sequence, an intron, a kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), and one or more nuclear targeting sequences, such as one or more DTS sequences, for example one or more SV40 DTS sequences, optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0064] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a promoter, such as a CMV promoter, a terminator, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, a kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), and one or more nuclear targeting sequences, such as one or more DTS sequences, for example one or more SV40 DTS sequences, optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0065] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a promoter, such as a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, and one or more nuclear targeting sequences, such as one or more DTS sequences, for example one or more SV40 DTS sequences, optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0066] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a promoter, such as a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, and one or more SV40 DTS sequences, optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0067] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a promoter, such as a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, and one or three SV40 DTS sequences, optionally wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0068] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a promoter, such as a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, and one or three SV40 DTS sequences, wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0069] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, and one or three SV40 DTS sequences, wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand.

[0070] In one embodiment, said nucleic acid nanostructure, preferably said at least one scaffold strand and/or at least one staple strand of said plurality of staple strands of said nanostructure, comprise(s) a CMV promoter, a polyadenylation signal sequence, at least one nucleic acid sequence encoding a gene, and three SV40 DTS sequences, wherein said at least one scaffold strand of said nucleic acid nanostructure is a ssDNA scaffold strand, and the shape of the nanostructure is a triangular shape.

[0071] According to this invention, the term “ssDNA” shall refer to single stranded DNA.

Accordingly, the term “ssDNA scaffold” shall refer to a single stranded DNA scaffold, and the term “ssDNA scaffold strand” shall refer to a single stranded DNA scaffold strand.

[0072] In one embodiment, said nanostructure has an aspect ratio in the range of from about 1:1 to about 1000:1, preferably 1.5:1 to about 20:1, more preferably from about 2:1 to about 15:1. In one embodiment, the aspect ratio is the ratio of the sizes of the nanostructure in different dimensions, e.g., a ratio of the longitudinal extension to a transverse extension. In one embodiment, the nanostructure has a maximum longitudinal extension along said longitudinal axis which is larger than a maximum transverse extension along said transverse axis. For example, a nanostructure having an aspect ratio of 20:1 may have a longitudinal extension of 20 nm and a transverse extension of 1 nm.

[0073] In one embodiment, said scaffold strand comprises said at least one nucleic acid sequence encoding a gene, a promoter, and a terminator; optionally further comprises a nucleic acid sequence configured to form a loop structure, an intron, a DNA nuclear targeting sequence, a polyadenylation signal sequence, a Kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element. In one embodiment, said nanostructure, preferably, said scaffold strand, comprises an expression cassette comprising said at least one nucleic acid sequence encoding a gene, a promoter, and a terminator; optionally further comprising a nucleic acid sequence configured to form a loop structure, an intron, a DNA nuclear targeting sequence, a polyadenylation signal sequence, a Kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element.

[0074] In a further aspect, the present invention relates to a composition, preferably a pharmaceutical composition, comprising a nucleic acid nanostructure as defined herein. In one embodiment, the composition, preferably pharmaceutical composition, comprises a pharmaceutically acceptable excipient. The composition, preferably pharmaceutical composition, of the invention shall be formulated to be compatible with its intended route of administration. In a particularly preferred embodiment, examples of routes of administration of the pharmaceutical and/or nanostructure of this invention include intravenous, oral, intranasal, intrathecal, intra-arterial, intradermal, subcutaneous, transdermal (topical), intracerebroventricular, intraparenchymal, intratumoral, transmucosal, rectal, vaginal, bronchial, parenteral administration, and any other clinically/medically accepted method for administration of a pharmaceutical and/or a compound.

[0075] In a further aspect, the present invention relates to a collection of nucleic acid sequences or collection of plasmids encoding a nucleic acid nanostructure as defined herein. In one embodiment, the collection of nucleic acid sequences comprises or consists of one or more nucleic acid sequences encoding a nucleic acid nanostructure as defined herein. In one embodiment, the collection of nucleic acid sequences comprises a scaffold strand and a plurality of staple strands. In one embodiment, the collection of plasmids comprises or consists of one or more plasmids encoding a nucleic acid nanostructure as defined herein. In one embodiment, the plasmid is a phagemid. In one embodiment, the collection of plasmids is a collection of phagemids.

[0076] In a further aspect, the present invention relates to a nucleic acid nanostructure, as defined herein, or composition, as defined herein, for use in medicine. In one embodiment, the nucleic acid nanostructure, as defined herein, or composition, as defined herein, is for use in a method of preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder. In one embodiment, the nucleic acid nanostructure, as defined herein, or composition, as defined herein, is for use in gene therapy and/or immunotherapy.

Advantageously, the nanostructure of the invention allows to efficiently express genes, such as mammalian genes, which is highly useful for the prevention, treatment, and diagnosis of genetic diseases or disorders and of immunological diseases or disorders. For example, the nanostructures of the invention can be used for gene therapy and immunotherapy by administering the nanostructures to a patient in need thereof, and expressing genes from said nanostructure in said patient, e.g. genes deficient in said patient and/or genes involved in a pathology. In one embodiment, said gene therapy and/or immunotherapy comprises or consists of a vaccination,

preferably a vaccination using said nanostructure or composition.

[0077] In a further aspect, the present invention relates to a method of expressing a gene from a nucleic acid nanostructure, preferably a nucleic acid nanostructure as defined herein, comprising [0078] i) providing a nucleic acid nanostructure comprising at least one nucleic acid sequence encoding a gene, preferably a nucleic acid nanostructure as defined herein; [0079] ii) delivering said nucleic acid nanostructure provided in step i) to a cell; wherein, preferably, said delivering comprises transfecting or transforming said cell; [0080] iii) allowing said cell to express said gene; [0081] wherein, optionally, said providing in step i) comprises providing a plasmid, preferably a phagemid, or a collection of plasmids, preferably a collection of phagemids, wherein said plasmid or collection of plasmids encodes said nucleic acid nanostructure, and preparing said nucleic acid nanostructure using said plasmid or collection of plasmids, preferably by using bacteriophages. In one embodiment, said collection of plasmids is a collection of plasmids as defined herein.

[0082] In one embodiment, said method of expressing a gene is an in vitro or ex vivo method. In one embodiment, delivering said nanostructure to said cell comprises contacting said nanostructure with said cell, optionally further comprising electroporation, lipofection, endocytosis such as chemically induced endocytosis and/or receptor-mediated endocytosis, phagocytosis, membrane fusion, heat shock, calcium phosphate, liposomes, nanoparticles, biolistics, microinjection, sonoporation, photoporation, magnetofection, and/or hydroporation. In one embodiment, said delivering in step ii) is performed using electroporation, lipofection, endocytosis such as chemically induced endocytosis and/or receptor-mediated endocytosis, phagocytosis, membrane fusion, heat shock, calcium phosphate, liposomes, nanoparticles, biolistics, microinjection, sonoporation, photoporation, magnetofection, and/or hydroporation. For example, chemically induced endocytosis may comprise peptide-targeted endocytosis, protein-targeted endocytosis, polysaccharide-targeted endocytosis, carbohydrate-targeted endocytosis, lipid-targeted endocytosis, and/or aptamer-targeted endocytosis.

[0083] In one embodiment, said allowing said cell to express said gene in step iii) comprises cultivating said cell(s) in a cell culture, preferably at a temperature in a range of from about 25° C. to about 40° C., preferably of from about 30° C. to about 38° C., more preferably at about 37° C., e.g. for about 1 h to about 72 h. In one embodiment, said allowing said cell(s) to express said gene in step iii) comprises providing suitable growth conditions for said cell. In one embodiment, said cell is a eukaryotic or prokaryotic cell, e.g. a mammalian cell, a fungal cell, a yeast cell, or a bacterial cell. In one embodiment, said cell is a mammalian cell.

[0084] In one embodiment, in the context of a method of the invention, the term “providing a nucleic acid nanostructure” comprises providing an assembled nucleic acid nanostructure and/or providing a building material for a nucleic acid nanostructure, such as a scaffold strand and one or more staple strands. In one embodiment, a method of preparing a nanostructure comprises a step of allowing self-assembly of a nanostructure and purifying said self-assembled nanostructure. For example, providing a nucleic acid nanostructure may comprise a step of allowing self-assembly and subsequent purification. In one embodiment, allowing self-assembly comprises mixing at least one scaffold strand and one or more staple strands, optionally further comprises adjusting the ionic strength, e.g. by adding about 10 mM to about 20 mM MgCl<sub>2</sub>, and/or further comprises using a temperature protocol running a sequence of temperatures. In one embodiment, said purifying comprises removing the remaining excess of staple strands, e.g. by a precipitation such as PEG-precipitation, filtration, and/or liquid chromatography. In one embodiment, self-assembly and/or a step of allowing self-assembly comprises a step of denaturation and a step of cooling. In one embodiment, a step of denaturation is performed at a temperature of from 50° C. to 80° C., preferably 60° C. to 70° C., e.g. about 65° C., for a period of from 1 minute to 45 minutes, preferably 10 to 20 minutes, e.g. about 15 minutes. In one embodiment, a step of cooling is performed at a temperature of from 0° C. to 70° C., preferably 20° C. to 60° C., e.g., about 50° C. to 58° C. In a preferred embodiment, a step of cooling is performed as a gradual cooling, preferably

at a temperature from about 58° C. to about 50° C. with a decrease of 1° C. per hour. A person skilled in the art understands that the protocol for self-assembly depends on the design and/or nucleic acid sequence of the nanostructure, and that the protocol can be adjusted in line with known protocols for preparing nanostructures.

[0085] In one embodiment, when referring to a method, the method is an in vivo, ex vivo, in vitro, or in situ method, e.g. an in vitro method. In one embodiment, when referring to a use, the use is an in vivo, ex vivo, in vitro, or in situ use, e.g. an in vitro use. In one embodiment, the nanostructure of the invention is for use in vivo or in vitro, preferably in vivo.

[0086] In a further aspect, the present invention relates to a use of a nanostructure, as defined herein, or of a composition, as defined herein, for gene expression, preferably for in vitro gene expression. For example, the nanostructure of the invention can be used to efficiently express molecules of interest, for example proteins of interest, in cell culture, e.g. for large-scale production of therapeutic proteins such as mammalian enzymes or antibodies.

[0087] In a further aspect, the present invention relates to a method of preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder, comprising administering a nanostructure, as defined herein, or a composition, as defined herein, to a patient in need thereof. In one embodiment, said method of preventing, treating, and/or diagnosing a disease or disorder is a method of gene therapy and/or immunotherapy. In one embodiment, said administering comprises administering an effective amount of a nanostructure, as defined herein, and/or a composition, as defined herein, to a patient in need thereof. The term “patient”, as used herein, may relate to a human or an animal. The term “effective amount”, as used herein, relates to an amount sufficient to evoke a desired effect, e.g. a sufficient labeling for in vivo imaging.

[0088] In a further aspect, the present invention relates to a use of a nanostructure, as defined herein, or of a composition, as defined herein, for the manufacture of a medicament, e.g. a medicament for preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder; optionally for gene therapy and/or immunotherapy. For example, the nanostructure may comprise genes which are deficient in a patient.

[0089] The terms “of the [present] invention”, “in accordance with the invention”, “according to the invention” and the like, as used herein are intended to refer to all aspects and embodiments of the invention described and/or claimed herein.

[0090] As used herein, the term “comprising” is to be construed as encompassing both “including” and “consisting of”, both meanings being specifically intended, and hence individually disclosed embodiments in accordance with the present invention. Where used herein, “and/or” is to be taken as specific disclosure of each of the two specified features or components with or without the other. For example, “A and/or B” is to be taken as specific disclosure of each of (i) A, (ii) B and (iii) A and B, just as if each is set out individually herein. In the context of the present invention, the terms “about” and “approximately” denote an interval of accuracy that the person skilled in the art will understand to still ensure the technical effect of the feature in question. The term typically indicates deviation from the indicated numerical value by  $\pm 20\%$ ,  $\pm 15\%$ ,  $\pm 10\%$ , and for example  $\pm 5\%$ . As will be appreciated by the person of ordinary skill, the specific such deviation for a numerical value for a given technical effect will depend on the nature of the technical effect. For example, a natural or biological technical effect may generally have a larger such deviation than one for a man-made or engineering technical effect. Where an indefinite or definite article is used when referring to a singular noun, e.g. “a”, “an” or “the”, this includes a plural of that noun unless something else is specifically stated.

---

## Description



## BRIEF DESCRIPTION OF THE FIGURES

[0091] The present invention is now further described by reference to the following figures.

[0092] All methods mentioned in the figure descriptions below were carried out as described in detail in the examples.

[0093] FIG. 1 shows folding and expressing genes from origami structures, and the impact of gene position and origami aspect ratio on gene expression. a, Schematics of the overall workflow: ssDNA is produced from plasmid DNA via phagemids (i), and then folded into 20 helix bundle (20HB) DNA origami objects (ii). Objects were delivered to cells, and gene expression from the origami structure was assessed by detection of positive fluorescence read-out (iii). CMV promoter sequence is shown in blue, gene encoding for enhanced green fluorescent protein (EGFP) in green, and the polyA in purple. b, Cylinder models and negative-staining transmission electron micrographs of the 20HB, 12HB and 32 HB are given in the upper and lower panels, respectively (scale bar 100 nm, insets 20 nm), (HB: helix bundles). Coloring demonstrates the positions of the scaffold features, for example 20HB-ext demonstrates the CMV (blue), EGFP (green) and polyA (purple) encoding sequences present along exterior helices, while 20HB-int presents the sequences encoding for EGFP and polyA within interior helices. c, Transfection efficiency in HEK293T cells (sc: scaffold; st: staples). d, Schematics explaining internal crosslinking via UV irradiation. UV-welded structures 20HB-ext-W and 20HB-int-W, EGFP expression was silenced. e, Transfection efficiency in HEK293T cells by electroporation seen for 20HB(-ext), 32HB and 12HB structures. Data collected in c and e were quantified using flow cytometry and are presented as mean±standard deviation (s.d.) for n=3 biologically independent experiments. Individual data points are overlaid. Statistical analysis was performed using one-way ANOVA with Tukey's multiple comparison (\*p≤0.05, ns p>0.05).

[0094] FIG. 2 shows the optimization of gene expression through alternative staple design and scaffold orientation. Scaffold routing and schematic of unfolded scaffold for 20HB-LPv2 and 20HB-Circ designs, a and b respectively. 20HB-LPv2 incorporates two continuous 154-mer staples (enhancer staples; pink), and 20HB-Circ design has been routed so that the 200-mer staple (pink), which acts as a splint to bring together the 5' start of the CMV, and the 3'. c, TEM micrographs of 20HB-LPv2 and 20HB-Circ, scale bar 100 nm. d, Delivery into HEK293T cells via electroporation revealed statistically significant increased transfection efficiency for samples 20HB-LP, 20HB-LPv2 and 20HB-Circ, when compared to the standard 20HB (20HB-LP comprising two 93-mer continuous staples (enhancer staples)). e, Scaffold used until now has encoded for the 'coding strand', where the expression cassette is present in the 5' to 3' direction ("sc\_EGFP1", upper panel). Scaffold encoding for the reverse complementary sequence of the expression cassette, thus the "template strand" ("sc\_EGFP2", lower panel) was designed and produced. f, HEK293T cells transfected with either sc\_EGFP2 scaffold+staple mixture, or scaffold only, demonstrated significantly higher transfection efficiency than that of the sc\_EGFP1 counterpart. No significant difference in the transfection efficiency from 20HB structures was observed. g, EGFP mean fluorescence intensity (MFI) for 20HB, 20HB-LPv2 and 20HB-Circ structures folded with either the coding or the template strand as the scaffold. Structures 20HB and 20HB-Circ demonstrated no significant difference in MFI, while a significant difference was observed for structure 20HB-LPv2. Data collected in d, f and g were quantified using flow cytometry and are presented as mean±s.d. for n=3 biologically independent experiments, individual data points are overlaid. Statistical analysis in d was performed using one-way ANOVA with Tukey's multiple comparison, while statistical analysis for f and g was performed using Student's t-tests (\*p≤0.05, \*\*p≤0.01, \*\*\*p≤0.001, \*\*\*\*p≤0.0001, ns p>0.05). Advantageously, an efficient gene expression is achieved with a nanostructure comprising the nucleic acid sequence encoding a gene in the scaffold strand and with a nanostructure comprising the nucleic acid sequence encoding a gene in the plurality of staple strands.

[0095] FIG. 3 shows enhancing gene expression through alternative scaffold sequences. a, Scaffold designs where sc\_EGFP1 represents the initial scaffold design, and sc\_EGFP3/4/5/6 include additional sequence features such as ITRs (light pink), or ITR binding domains (ITR\*), kozak sequence (black), and WPRE (dark pink). b and c, Comparison of transfection efficiency as determined by EGFP+ cells (b) and mean fluorescent intensity of EGFP+ cells (c). Data collected in a and b were quantified using flow cytometry and are presented as mean±s.d. for n=3 biologically independent experiments, individual data points are overlaid. Statistical analysis was performed using one-way ANOVA with Tukey's multiple comparison (\*p≤0.05, \*\*p≤0.01, \*\*\*p≤0.001, \*\*\*\*p≤0.0001, ns p>0.05).

[0096] FIG. 4 shows enhancing gene expression through scaffold sequence design. a, 20HB design for the sc\_EGFP5 structure included an external single stranded loop to allow the ITR sequence to self-anneal and form the hairpin structure. 20HB design for the sc\_EGFP6 scaffold included two external loops to expose the ITR binding domain, enabling the ITR hairpin staples to anneal. b, c Transfection efficiency and MFI seen for 20HB structures folded with sc\_EGFP5 and sc\_EGFP6. Data collected in b and c were quantified using flow cytometry and are presented as mean±s.d. for n=3 biologically independent experiments, individual data points are overlaid. Statistical analysis in b and c was performed using one-way ANOVA with Tukey's multiple comparison, (\*p≤0.05, \*\*\*p≤0.001, \*\*\*\*p≤0.0001, ns p>0.05). d, Representative epifluorescence microscopy images showing EGFP expression from cells transfected with DNA origami objects folded with sc\_EGFP5 and sc\_EGFP6 relative to sc\_EGFP1. Images in the bottom row have been purposely contrast enhanced to reveal EGFP positive cells that have poor EGFP intensity in the sc\_EGFP1 sample. The images are representative of one of n=3 biologically independent experiments; similar results were observed each time. Scale bar 100 μm.

[0097] FIG. 5 shows structure transfection summary and optimization. a, Comparison of transfection efficiency (%) and EGFP MFI (A.U.) across all structures investigated, grouped by scaffold. sc\_EGFP1 20HB-ext was used as an internal control in all experiments, and EGFP MFI is represented as fold change compared to this sample. Three clusters are highlighted: 1, structures with high folding quality but low overall gene expression (grey); 2, structures with low folding quality and medium expression levels (orange); 3, structures with medium to high folding quality and high expression levels (yellow). b, Representative epifluorescent microscopy images showing the expression of EGFP by successfully transfected HEK293T cells after optimization of electroporation settings. For each of the conditions, eGFP expression (green), cell (phase contrast) and the overlay are given. The images are representative of one of n=2 biologically independent experiments; similar results were observed each time. Scale bar 100 μm.

[0098] FIG. 6 shows plasmid designs with varying numbers of SV40 DTS sequences included (0×, 1×, 3× and 6×SV40 repeats) for production of custom scaffolds.

[0099] FIG. 7 shows production and characterization of custom scaffolds and corresponding DNA origami structures. a, Agarose gel demonstrating all custom scaffolds produced, and the corresponding purified DNA origami structures. b, Representative negative stain TEM images showing the 20HB DNA origami structures for each of the custom scaffolds produced. Scale bar 100 nm.

[0100] FIG. 8 shows that cell cycle arrest diminishes gene delivery efficiency. Representative transmission image and corresponding fluorescence image of dividing a and arrested b HEK293T cells 24 h after electroporation with the 20HB-mCh without any SV40 sequences. mCherry signal is shown in red, nuclei in blue, scale bar 100 μm. c, Flow cytometry histogram plot demonstrating cell cycle populations of actively dividing and chemically arrested HEK293T cells. (d) Quantification of mCherry+ cells (%) in dividing and chemically arrested HEK293T populations 24 h after electroporation with the 20HB-mCh.

[0101] FIG. 9 shows the effect of including SV40DTS sequences in DNA origami structures on the gene expression in dividing and arrested HEK293T cells after electroporation. a, Transfection

efficiency (bar graph) and mean fluorescence intensity (square symbols) 24 h after electroporation with DNA origami structures containing one or several DTS compared to the control structure, with no DTS sequence, in normally dividing HEK293T cells. b, Transfection efficiency (bar graph) and mean fluorescent intensity (square symbols) 24 h after electroporation with DNA origami structures containing one or several DTS compared to the control structure, with no DTS sequence, in arrested HEK293T cells. The data were quantified using flow cytometry and are presented as mean values $\pm$ s.d. for n=3 biologically independent experiments. One-way ANOVA was performed to test statistically significant differences in gene expression compared to the control. For normally dividing cells (a), the inclusion of one copy of the SV40 DTS resulted in a statistically significant increase in transfection efficiency (\* $p\leq 0.05$ ) and mean fluorescence intensity (\*\* $p\leq 0.01$ ) compared to the control without DTS sequence. For arrested cells (b), the inclusion of three copies of the SV40 DTS resulted in a statistically significant increase in transfection efficiency (\*\* $p\leq 0.01$ ) and mean fluorescence intensity (\*\*\*\* $p\leq 0.0001$ ) compared to the control without DTS sequence. The inclusion of one copy did not result in a statistically significant increased transfection efficiency but did result in a statistically significant increase in mean gene expression (\*\*\*\* $p\leq 0.0001$ ) for arrested cells.

[0102] FIG. 10 shows delivery of multimeric origami assemblies enables codelivery of genes in defined ratios. a, Cylindrical models of DNA origami objects for programmed assembly via shape-complementary protrusions and recesses. Schematic demonstrates unique interaction patterns to build dimer (i), trimer (i and ii) and tetramer (i, ii, iii and iv) higher-order assemblies. b, Representative comparison tomogram slice through dimer, trimer and tetramer structures, scale bar 100 nm, taken from a sample with mixed assembly products. c, Schematic demonstrating passivated overhangs for inhibiting assembly, and assembly assisted via complementary 5 nt sticky ends. d, Cotransfection (mCherry+/EGFP+) efficiency in HEK293T cells after delivery of mCherry and EGFP as individual monomers (Pass.), or as a dimer connected through 5 nt or 8 nt sticky ends. e, Cotransfection (mCherry+/EGFP+) efficiency in HEK293T cells with assembly multimeric DNA origami structures including mCherry and EGFP-encoded monomers in ratios of 1:1, 1:2 and 1:3 mCherry:EGFP. EGFP MFI (A.U.) is given on the right y-axis. Data collected in d and e were quantified using flow cytometry and are presented as mean $\pm$ s.d. for n=3 biologically independent experiments, individual data points are overlaid. Statistical analysis in d and e was performed using one-way ANOVA with Tukey's multiple comparison, (\* $p\leq 0.05$ , \*\* $p\leq 0.01$ , ns  $p>0.05$ ). f, Top: schematic design of mCherry and EGFP monomer blocks for no assembly (passivated), or assembly into dimer, trimer or tetramer structures in the ratio of 1:1, 1:2 and 1:3 mCherry:EGFP, from left to right. Bottom: representative epifluorescent microscopy images demonstrated the expression of mCherry (red), EGFP (green), or coexpression (yellow) by successfully transfected HEK293T cells. Cell nuclei are given in blue, scale bar 100  $\mu$ m. The images are representative of one of n=3 biologically independent experiments; similar results were observed each time.

[0103] FIG. 11 shows DNA origami triangles for gene expression. a, Scaffold design with three SV40 DTS sequences for production of custom scaffolds, wherein the phagemid encoding mCherry has 8064 bases. b, Representative negative stain TEM images showing the mCherry-encoding DNA origami structures (triangular structures) for the custom scaffolds produced. Scale bar: 50 nm. c, Representative epifluorescence microscopy images showing mCherry expression from HEK293T cells transfected with DNA origami objects, 48 hours after electroporation of mCherry-encoding origami triangles. Scale bar: 50  $\mu$ m.

[0104] In the following, reference is made to the examples, which are given to illustrate, not to limit the present invention.

## EXAMPLES

### Example 1: Materials and Methods

#### Scaffold Production.

[0105] The design and cloning methods of our customized scaffolds are given in detail in Examples

2 and 6. In brief, gene fragments from EGFP-containing plasmids (Addgene plasmids #13031 and #105530, with and without ITR sequences respectively) were assembled with a fragment for phage origin of replication bacterial resistance (Addgene plasmid #126854) using either Golden Gate or digestion ligation cloning. Plasmids were verified using restriction digests and DNA sequencing (Eurofins genomics, Ebersberg Germany). Exact primer sequences and methods can be found in Examples 2 and 6, and sequences of the custom scaffolds can be found in SEQ ID NOS 1-11. Exemplary sequences of staple strands can be found in SEQ ID NOs 76-1507.

[0106] Production of the ssDNA custom scaffolds was performed as previously described.<sup>sup.1,2</sup> Briefly, chemically competent DH5 $\alpha$  *E. coli* cells were cotransformed with the plasmid of interest, and a helper plasmid (Addgene plasmid #120346). Single colonies were picked and grown for ~10 h in a 5 mL pre-culture (2 $\times$ YT, 30  $\mu$ g/mL kanamycin, 30  $\mu$ g/mL carbenicillin) before being transferred to 750 mL of 2 $\times$ YT (30  $\mu$ g/mL kanamycin, 30  $\mu$ g/mL carbenicillin, 5 mM MgCl<sub>2</sub>.sub.2) in Ultra Yield flasks (Thomson). Cells then were grown in a shaking incubator at 37° C. overnight. Bacteria were pelleted by centrifugation (45 min, 4500 g, and the supernatant collected. Phagemid particles were precipitated from the supernatant with addition of polyethylene glycol 8000 (PEG-8000, final concentration 3% w/w) and NaCl (final concentration 0.5 M) and incubated with stirring for 1 h at rt, before being collected by centrifugation (45 min, 4500 g, 4° C.). The pellet was resuspended in 4 mL of 1 $\times$  TE buffer (10 mM Tris, 1 mM EDTA, pH 8) and centrifuged again (15 min, 16000 g, 4° C.) to remove residual bacterial components. The ssDNA scaffold was then extracted via phagemid lysis and purified via ethanol precipitation.

DNA Origami Design, Folding and Purification.

[0107] All origami objects were folded in standardized ‘folding buffers’ containing x mM MgCl<sub>2</sub>.sub.2 in addition to 5 mM Tris base, 1 mM EDTA and 5 mM NaCl, pH 8 (FoBx). All reactions were subjected to thermal annealing ramps in Tetrad (Bio-Rad) thermal cycling devices. Exact folding conditions for each structure is given in Tables 2 and 3. Staple strands were purchased from Integrated DNA Technologies, as exemplified in Example 6, and used with standard desalting unless stated otherwise. Origami objects were purified by either PEG precipitation, or gel purification, as previously described.<sup>sup.3,4</sup>

[0108] Assembly of multi-component DNA origami structures. To assemble the origami subunits to form the dimer, trimer and tetramer samples, monomers were mixed in molar ratios in 1  $\times$  FoB5 buffer and incubated for 48 h at 37° C. Passivated samples were treated identically.

UV Welding.

[0109] UV weldable samples were designed with additional thymine bases located at all potential staple crossover position, and UV-crosslinked as described previously [5] with UV light (310 nm, 2 h) using Asahi Spectra Xenon Light source (300 W, MAX-303) with a high transmission bandpass filter centered around 310 nm (XAQA310, Asahi Spectra). Samples were in FoB10 buffer at the time of UV-crosslinking.

PAGE Purification of Ultramers.

[0110] Long staple oligomers (93-mers, 154-mers and 200-mers) were purchased from IDT as ultramers and purified in-house via denaturing urea polyacrylamide gel electrophoresis (Urea-PAGE). Bands corresponding to the correct MW were cut away and crushed prior to the addition of 1 $\times$  TEN buffer (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 8.00). Pure ultramers were recovered via EtOH precipitation, redissolved in MilliQ H<sub>2</sub>O, and stored at 4° C.

Gel Electrophoresis.

[0111] For characterization of PCR products and plasmids, 1% agarose gels containing 0.5 $\times$  TBE buffer (22.25 mM tris base, 22.25 mM boric acid, 0.5 mM EDTA) were used. Gel electrophoresis was performed with an identical buffer solution for 1 h at a voltage of 110 V. To characterize assembled origami and scaffolds, the inventors used 2% agarose gels containing 0.5 $\times$  TBE buffer and 5.5 mM MgCl<sub>2</sub>.sub.2. Gel electrophoresis was performed with an identical buffer solution for 1-2 h at a voltage of 90 V, gels were placed in a water bath for cooling. All gels were imaged using a

Typhoon FLA 9500 laser scanner (GE Healthcare) with a pixel size of 50  $\mu\text{m}$ /pixel.

#### Negative Staining TEM.

[0112] Samples were incubated on glow-discharged copper TEM grids (FCF400-CU, Electron Microscopy Sciences), for 30-60 s. Grids were then stained for 30 s (2% aqueous uranyl formate, 25 mM NaOH). Imaging was performed at magnifications of 21,000-42,000 $\times$ . Data was acquired with SerialEM software, using a FEI Tecnai T12 microscope (120 kV, Tietz TEMCAM-F416 camera). Images were processed using ImageJ..sup.5 TEM micrographs were high-pass filtered to remove long-range staining gradients and the contrast was auto-leveled using Adobe Photoshop CS5.

[0113] The tilt series were performed from  $-50^\circ$  to  $+50^\circ$  and micrographs were acquired in  $2^\circ$  increments, the tomogram was then generated using a filtered back-projection, processed with Etomo (IMOD) to acquire tomograms..sup.6 The Gaussian-Filter used a cutoff between 0.25 and 0.5, and a fall-off of 0.035.

#### Cell Culture.

[0114] HEK293T cells (DSMZ) were cultured routinely in Dulbecco's modified Eagle's medium (DMEM, Gibco, cat. no. 31966047), supplemented with 10% heat-inactivated fetal bovine serum (FBS, Sigma-Aldrich, cat. No. F9665). Cells were grown in a humidified incubator at  $37^\circ\text{C}$ . with 5%  $\text{CO}_2$ .sub.2.

#### Cell Cycle Arrest.

[0115] HEK293T cells were arrested for 24 h prior to electroporation using arrest media (DMEM supplemented 10% FBS and 5 ng/ $\mu\text{L}$  Aphidicolin, Sigma-Aldrich, cat. no. A0781, dissolved in dimethyl sulfoxide, DMSO, Sigma-Aldrich). Cells were kept in arresting media for the entire time of the experiment.

#### Electroporation.

[0116] Electroporation experiments were carried out according to the Manufacturer's protocol (Neon<sup>TM</sup> transfection protocol, ThermoFisher). Briefly, HEK293T cells were washed with phosphate buffered saline solution (PBS) and collected using TrypLE. Cells were pelleted via centrifugation (5 min, 300 g), resuspended in PBS and counted. Cells were centrifuged again (5 min, 300 g), and then resuspended in Buffer R (Neon<sup>TM</sup> Transfection System) at a concentration of  $5 \times 10^6$  cells/mL. Mixtures for each condition were prepared so that each electroporation event contained 0.5  $\mu\text{g}$  total DNA, and the volume was supplemented to a total of 1  $\mu\text{L}$  with  $1 \times$  FOB5 buffer (folding buffer, 1 mM Tris, 1 mM EDTA, 5 mM NaCl, 5 mM  $\text{MgCl}_2$ .sub.2), which was mixed with 9  $\mu\text{L}$  of the cell suspension. Electroporation occurred in the 10  $\mu\text{L}$  transfection tips, with two pulses at pulse voltage of 1150 V and width of 20 ms. After electroporation, cells were immediately transferred to a 48 well plate which had been pre-prepared with a poly-L-lysine coating, and 240  $\mu\text{L}$  of complete DMEM growth media or arresting media.

[0117] After 48 h, samples were imaged using the EVOS<sup>TM</sup> M7000 Imaging System, and the transfection efficiency was quantified via flow cytometry. For arrest experiments, the cells were analyzed 24 h after electroporation to avoid extensive cell death. Briefly, samples were acquired using Attune Nxt Flow Cytometer and software (Thermo Fisher). In total, 20,000 single cell events, gated on side scatter area versus height, were recorded for analysis. EGFP was excited with a 488 nm laser, and emission was measured with a 530/30 nm bandpass filter. mCherry was excited with a 561 nm laser and emission was measured with a 620/15 nm bandpass filter. Untreated cells, and cells electroporated with buffer only, were used as negative controls. Cells electroporated with the corresponding EGFP plasmid was used as a positive control. Cell cycle arrest was confirmed by cell cycle analysis via flow cytometry. The cells were stained with FxCycle<sup>TM</sup> Far Red Stain (Invitrogen, Thermo Fisher Scientific) according to Manufacturer's protocol. The dye was excited with a 638 nm laser and emission was measured with a 670/14 nm bandpass filter. Data was analyzed post-acquisition using FlowJo software (v10.7.1).

#### Statistics and Reproducibility.

[0118] Statistical analyses were performed with GraphPad Prism (GraphPad Software Inc. v9). The data is illustrated as the mean±standard deviation, and the individual data points representing biological replicates are shown. The specific analysis performed is detailed in the corresponding figure caption. For all tests, differences were considered significant at  $p \leq 0.05$  (\*),  $p \leq 0.01$  (\*\*),  $p \leq 0.001$  (\*\*\*),  $p \leq 0.0001$  (\*\*\*\*).

## Example 2: Further Materials and Methods

### Scaffold Cloning.

[0119] Plasmids encoding for custom scaffolds were created via standard cloning techniques. All plasmids, with the exception of sc\_EGFP3, were created via Golden gate assembly using either Esp3I (NEB cat. no. R0734), or BsaI-HF®v2 (NEB cat. no. R3733), together with T4 DNA ligase (NEB cat. no. M0202). For each plasmid, appropriate cut sites were introduced with PCR, and the assembly was conducted as per manufacturer's protocol.

[0120] The exception, sc\_EGFP3, was assembled via digestion ligation, and compatible enzyme cut sites were added to the fragment of interest by means of PCR. Enzymes EcoNI and PacI (NEB cat. no. R0521 and R0547 respectively) were used to digest fragments of interest, and then ligated using NEB T4 DNA ligase, as above.

[0121] Touchdown PCR with primers (see Table 1) for all constructs was performed. In all cases, PCR products were confirmed by agarose gel electrophoresis (AGE), bands were excised, and fragments were extracted (Qiagen QIAquick Gel Extraction Kit) as per manufacturer's protocol.

TABLE-US-00001 TABLE 1 Primer sequences used for the construction of plasmids for custom scaffold production. Plasmid Template Primer Sequence 5' - 3'

Plasmid	Template	Primer Sequence (5' - 3')
sc_EGFP1	Addgene	FWD CTGGATGGTCTCCgtgacattaagcgcgggcggtg #126854
	plasmid	FWD CTGGATGGTCTCCaatgagttagcaaaaggccagca
	Addgene	plasmid FWD CTGGATGGTCTCCcattgatatacgcttgacattga #13031
	REV	CTGGATGGTCTCCtcacattccgcctcagaagccat
sc_EGFP2	Addgene	plasmid FWD CACTGACGTCTCTgtgacattaagcgcgggcggtg #126854
	REV	CACTGACGTCTCTaatgagttagcaaaaggccagca
	Addgene	plasmid FWD CACTGACGTCTCTcattcgctcagaagccataga #13031
	REV	CACTGACGTCTCTtcacagatatacgcttgacattga
sc_EGFP3	Addgene	plasmid FWD TTAATTAACattaagcgcgggcggtgt #105530
	REV	CCTAATTAAGGgtgagcaaaaggccagcaaa
sc_EGFP4	sc_EGFP3	FWD CACTGACGTCTCTcattatgctctaggaagatcgga
	REV	CGTGATCGTCTCTaatgtaagggtgagcaaaaggcca
sc_EGFP5	sc_EGFP3	FWD CACTGACGTCTCTcattgccttaattaacattaagcgc
	REV	CACTGACGTCTCTaatgcgcatgctacttatctacg
sc_EGFP6	sc_EGFP3	FWD AGATGGCGTCTCCtaggggcctcagttagcgagcg
	REV	AGATGGCGTCTCCccttaattaacattaagcgcg
	Addgene	plasmid FWD AGATGGCGTCTCCctaatTAAGGgtgagcaaaagg
	sc_mCherry5	sc_EGFP5 FWD TGGCTACGTCTCGgtgattgatccaatcaacctctg
	REV	TGGCTACGTCTCGaatgcctggacacctgtggaga
	Addgene	plasmid FWD TGGCTACGTCTCGcattattcgccaccatggtgag #127813
	REV	TGGCTACGTCTCGtccccgctcactgtacagct
sc_mCherry	Addgene	plasmid FWD CAAGGTGGTCTCCgtgacattaagcgcgggcggtg #126854
	REV	CAAGGTGGTCTCCaatgagttagcaaaaggccagca
	Addgene	plasmid FWD CAAGGTGGTCTCGcattcgcatgtacgggccaga #128744
	REV	CAAGGTGGTCTCGtcacagagccccagctggttctt
	sc_mCherry_	plasmid FWD TTTCCGGGTCTCGgtgacattaagcgcgggcggtg 1xSV40 #126854
	REV	TTTCCGGGTCTCGaatgagttagcaaaaggccagca
	sc_mCherry	FWD TTTCCGGGTCTCCcattgtacgggccagatatag
	REV	TTTCCGGGTCTCCATctctagactcgagcgggcc
	sc_mCherry	FWD TTTCCGGGTCTCCGcttaaacccgctgatcgc
	REV	TTTCCGGGTCTCCtcacaggttcttccgcctcaga
	sc_mCherry_	sc_mCherry_ FWD TTGTGGGGTCTCGgcctcgactgtgccttctag 3xSV40_ 1xSV40
	REV	TTGTGGGGTCTCGgagcgccgctcactgtacagc
	intermediate	sc_mCherry_ FWD

TTGTGGGGTCTCGgctcagcttagagatccg 1xSV40 REV TTGTGGGGGTCTCGtttggttgactaattgag  
sc\_mCherry\_ FWD TTGTGGGGGTCTCGcaaagctctagagatccggtgtgg 1xSV40 REV  
TTGTGGGGGTCTCGgatgggttaaagctttggttgct sc\_mCherry\_ FWD  
TTGTGGGGGTCTCGcatccggtgtggaaagtcc 1xSV40 REV  
TTGTGGGGGTCTCGaggctgatcagcgggttaa sc\_mCherry\_ sc\_mCherry FWD  
TTCGAGGGTCTCCttgtgacattaagcgcggc 3xSV40 REV TTCGAGGGTCTCCcttaccggccctctaga  
sc\_mCherry FWD TTCGAGGGTCTCGtaaggagggccgttaaacc REV  
TTCGAGGGTCTCGagccatagagcccaccgcat sc\_mCherry FWD  
TTCGAGGGTCTCGggctcgtttcttgctgtcc REV TTCGAGGGTCTCGcatccccagtttagtagtgg  
sc\_mCherry\_ FWD CAAGAGGGTCTCCatgcggccgctcagcttag 3xSV40 REV  
CAAGAGGGTCTCCcgaggctgatcagcgggt intermediate Addgene plasmid FWD  
AACACCGGTCTCGgactacaacaaggcaaggct #128744 REV  
AACACCGGTCTCGacaaagcagcgcaaaacgcct sc\_mCherry\_ sc\_mCherry\_ FWD  
GTACACGGTCTCGgactatacgcttgacattgattat 6xSV40 3xSV40 REV  
GTACACGGTCTCGcaatgagtgagcaaaaggcc sc\_mCherry\_ FWD  
GTACACGGTCTCGttggggatgcggccgctcga 3xSV40 REV  
GTACACGGTCTCGagtcgaggctgatcagcgg

#### DNA Origami Folding.

[0122] Nucleic acid nanostructures, for example CS3\_EGFP3, CS3\_EGFP1, CS3\_EGFP4, CS3\_EGFP4A, CS3\_EGFP4B, CS3EGFP5, were successfully produced. The nucleic acid nanostructures were folded using the conditions defined in tables 2 and 3.

TABLE-US-00002 TABLE 2 Folding conditions summary for each of the structures and scaffolds used. 10× v(scaffold, v(staples, Folding 100 nM,) 100 μM), Scaffold Structure Program Buffer μL\* μL\* sc\_EGFP1 20HB-ext 3 FoB10 10 8 20HB-int 1 FoB10 10 8 20HB-ext-W 1 FoB15 10 8 20HB-int-W 1 FoB15 10 8 32HB 1 FoB20 10 8 12HB 2 FoB10 10 8 20HB LS 3 FoB15 10 8 20HB LP 1 FoB20 10 8 20HB LPv2 1 FoB20 10 8 20HB Circ 3 FoB10 10 8 sc\_EGFP2 20HB 1 FoB10 10 8 20HB LPv2 3 FoB15 10 8 20HB Circ 3 FoB10 10 8 sc\_EGFP3 20HB 5 FoB20 2.5 .sup. 8.sup.‡ sc\_EGFP4 20HB 2 FoB15 10 8 sc\_EGFP5 20HB 3 FoB15 4 14 20HB loop 5 FoB10 4 8 20HB loop LPv2 4 FoB15 4 8 16HB\_2 7 FoB12.5 10 8 16HB\_3 7 FoB25 10 8 16HB\_4 7 FoB7.5 10 8 sc\_EGFP6 20HB 2 FoB15 4 14 20HB LPv2 6 FoB10 4 8 sc\_mCherry5 16HB\_1 1 FoB10 10 8 sc\_mCherry 20HB-mCh 1 FoB10 4 8 sc\_mCherry\_1xSV40 20HB-1xSV40 1 FoB15 10 8 sc\_mCherry\_3xSV40 20HB-3xSV40 2 FoB25 4 14 sc\_mCherry\_6xSV40 20HB-6xSV40 8 FoB15 10 8 \*For 20 μL folding reaction, where necessary volume supplemented to 20 μL total with ddH.sub.2O. .sup.‡Staples used here were at 500 μM concentration.

TABLE-US-00003 TABLE 3 Folding programs used for the folding reactions. 1. Denaturation time 3. Storage Program 30 s 15 min 2. Temperature ramp temperature 1 — 65° C. 60-44° C., at 1° C./1 h 20° C. 2 — 65° C. 60-44° C., at 1° C./2 h 20° C. 3 70° C. 65° C. 60-35° C., at 1° C./1 h 20° C. 4 70° C. 65° C. 60-35° C., at 1° C./2 h 20° C. 5 70° C. 65° C. 60-35° C., at 1° C./4 h 20° C. 6 70° C. 65° C. 60-30° C., at 1° C./1 h 20° C. 7 — 65° C. 60-44° C., at 1° C./4 h 20° C. 8 — 65° C. 60-25° C., at 1° C./2 h 20° C.

#### Example 3: Analysis of Structure Integrity

[0123] The structure integrity of the nucleic acid nanostructure of the invention was analyzed. For example, 20 HB structure integrity after electroporation was analyzed. Structural integrity of the 20 HB was maintained when diluted in RPMI 1640 media. Furthermore, 20 HBs were stable in both the electroporation buffer (EB, buffer R in kit) and after electroporation using the Neon™ transfection system. The nucleic acid nanostructure of the invention shows an advantageous structural integrity.

#### Example 4: Results

Genes are Readily Expressed From DNA Origami Independent of Gene Position or Origami Shape.

[0124] The inventors first investigation was to determine basic parameters of origami design that

mammalian cells will express. To do this, the inventors created a customized circular ssDNA scaffold which encoded for an enhanced green fluorescent protein (EGFP) in the 5' to 3' direction (coding strand), as given in FIG. 1a. Thus, cells which successfully express EGFP from the nucleic acid nanostructure can be monitored via fluorescence detection. The inventors use two observables in this study: the fraction of cells showing green fluorescence (termed transfection efficiency), and the fluorescence intensity per cell which the inventors use as a proxy for expression efficiency. The inventors used electroporation as the method to deliver the origami directly to the cell in order to circumnavigate issues such as cellular uptake and endosomal escape, which could confound data interpretation, and rather focus on parameters directly affecting expression. The inventors used electroporation via the Neon™ transfection system which did not compromise (and prevented aggregation of) the DNA origami structures.

[0125] The custom EGFP scaffold (sc\_EGFP1) expressed in high yield and purity via phagemid production, and the scaffold folded efficiently into the designed target objects (FIG. 1b). To address whether the spatial position of the gene within the DNA origami object affects expression, the inventors designed two 20-helix bundles (20HBs) variants where the EGFP gene was positioned either on the exterior (20HB-ext) or in the interior (20HB-int) of the multi-layer DNA origami (FIG. 1b, first two panels). Gene expression occurred from both 20HB variants in human embryonic kidney 293T (HEK293T) cells after electroporation (FIG. 1c). The inventors found no statistically significant difference in either the transfection or expression efficiency from the two objects.

[0126] The aspect ratio of DNA origami has previously been reported to influence cellular uptake.<sup>sup.7,8</sup> To elucidate whether aspect ratio influences expression, the inventors designed a ~114 nm long twelve-helix bundle (12HB), a ~69 nm long 20HB-ext and a ~42 nm long 32HB with the EGFP gene and recognition sequences presented in all cases on the exterior of the bundles (FIG. 1b). These objects have aspect ratios of ~15, 5 and 2 for the 12HB, 20HB and 32HB, respectively. The EGFP and associated genes are presented in the 12HB as long continuous regions with minimal scaffold crossovers, while the 32HB presents them within the shortest continuous regions. When delivered to HEK293T cells the inventors found no statistically significant difference in transfection and expression efficiency from the 20HB and the 32HB samples (FIG. 1e). A small decrease in transfection efficiency was seen from the 12HB sample ( $p \leq 0.05$ ) relative to 20HB and 32HB. However, the cell density after electroporation was also lower for the 12HB object.

[0127] Hence, for both transfection and expression efficiency it did not matter how the gene of interest was packaged among our panel of test DNA origami. This observation suggested that unfolding of the DNA origami occurs prior to gene expression. The inventors tested this hypothesis with EGFP encoding objects that cannot unfold. To this end the inventors included extra thymidine residues in the staple strands for 20HB-ext and 20HB-int to enable internal crosslinking via UV point welding.<sup>sup.9</sup> The objects were then internally stabilized by several hundred UV-induced cyclobutane pyrimidine dimer bonds between staple strands and at crossovers, which topologically prevents strand dissociation (FIG. 1d). When the inventors delivered the UV point welded 20HB variants to the HEK293T cells, the inventors found almost complete suppression of the EGFP signal (FIG. 1c). While exposure to UV radiation can also have an inhibitory effect on gene expression from plasmids.<sup>sup.10,11</sup> the pronounced, near-complete inhibition of gene expression from the covalently crosslinked DNA origami supports the inventors hypothesis that DNA origami must unfold prior for gene expression.

Targeted Design Changes in Promoter Region Enhance Gene Expression.

[0128] The inventors observed that electroporating a premixed, non-annealed cocktail of ssDNA scaffold and staple strands, which do not form structured objects, resulted in slightly higher transfection efficiency compared to administering the ssDNA scaffold alone (FIG. 1c, 1e). The inventors hypothesized that partial association occurs between the staple and scaffold strands,



resulting in double-stranded DNA regions around the promoter regions that enhance gene expression. The inventors thus tested whether simply increasing the average staple length in a DNA origami object would lead to enhanced expression, which was not the case, suggesting that a more targeted design is required. We redesigned the 20HB object to incorporate long continuous staple segments (enhancer staples) with no crossovers in the promoter region, resulting in a structure with continuous 93-mer and 154-mer staples (enhancer staples) flanking the expression region and at the 5' start region of the CMV promoter, and at the 3' end of the polyA sequence (20HB-LP and 20HB-LPv2, respectively). A schematic of 20HB-LPv2 design and staple localization is given in FIG. 2a. Inspired by the partially double stranded hepatitis B genome.<sup>sup.12</sup>, the inventors also prepared a design in which a 200-mer staple acts as a splint between the 5' start of the CMV, and the 3' end of the polyA to form a partially double stranded circular structure when unfolded (20HB-Circ) (FIG. 2b). All designs folded readily into defined 20HBs, as seen by direct imaging by TEM (FIG. 2c). Delivery of these objects into cells resulted in up to 50% enhancement of gene expression efficiency for objects 20HB-LP, 20HB-LPv2 and 20HB-Circ when compared to the standard 20HB staple routing (FIG. 2d).

[0129] Next, the inventors determined whether the orientation of the target gene on the scaffold impacts gene expression. As the scaffold is ssDNA, delivery of the coding strand requires synthesis of the complementary sequence (template strand) prior to transcription. The inventors created a 'template strand' scaffold with the reverse complementary gene sequences (sc\_EGFP2, FIG. 2e). Delivering these scaffolds with and without staples demonstrated significantly increased transfection efficiency for the template strand compared to the coding strand (FIG. 2f). However, when folded into the 20HB DNA origami object, the difference in transfection efficiency disappeared (FIG. 2f, right). The overall transfection efficiency of the 20HBs thus does not depend on having either the coding or template strand as the scaffold. Yet the inventors observed a minor trend of increased mean fluorescent intensity (MFI) of EGFP in EGFP-positive cells for all objects with the template-strand scaffold relative to those where the coding strand was used as scaffold (FIG. 2g).

#### Inclusion of Scaffold Sequence Features Boosts Gene Expression.

[0130] To further enhance the gene expression, the inventors included additional features in the scaffold sequence based on ssDNA AAV2 expression cassettes (FIG. 3a). The inventors placed a kozak sequence, which functions as a protein translation site,<sup>sup.13</sup> upstream of EGFP; a chimeric intron, and a woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) downstream of EGFP (before the polyA). The WPRE is thought to improve mRNA stability and protein yield.<sup>sup.14</sup> Additionally, the inventors included inverted terminal repeats (ITRs) flanking the expression cassette. ITRs are palindromic sequences which form a T-shaped hairpin, and are used by adeno-associated viruses as an origin of replication for their ssDNA genome, in addition to other functions.<sup>sup.15,16</sup> The inventors attempted producing a synthetic scaffold ssDNA containing all of these features (sc\_EGFP3, FIG. 3a). The scaffold was produced at low yield and low quality, which the inventors attributed to the repetitive ITR structures. To improve scaffold yield and quality we produced a further series of scaffolds that included only a single ITR downstream or upstream of the expression cassette (sc\_EGFP4 and sc\_EGFP5, respectively). Additionally, the inventors created a scaffold which includes partial sequences of both ITRs, but where the ITR hairpin would be provided by a complementary staple oligonucleotide during DNA origami folding (sc\_EGFP6). 20HBs with standard staple designs were produced for all these scaffold variants.

[0131] The inventors observed a trend of increased transfection efficiency and enhanced gene expression in the scaffold-only controls in all cases relative to the original sc\_EGFP1 scaffold (FIG. 3b). The 20HB samples folded from the "enhanced" scaffolds demonstrated a similar range of transfection efficiencies relative to those observed for 20HB folded from the sc\_EGFP1 scaffold but the MFI was significantly increased in the positive cells for 20HBs folded from the "enhanced"

scaffolds sc\_EGFP3/4/5/6 relative to the original sc\_EGFP1 (FIG. 3b), meaning that the additional features in the scaffolds enhanced intracellular gene expression.

[0132] In the inventors designs discussed thus far, the ITR sequences were masked within double-helical DNA domains in the object, to become available only once the object denatures within the cell. The inventors hypothesized that the gene expression from the 20HB may be further improved by positioning the ITR sequence motif such that it can already assemble into its hairpin secondary structure during folding of the object (FIG. 4a, design 20HB-ex). Additionally, the inventors included a continuous 154-mer staple (enhancer staple) at the 5' region of the promoter for design 20HB-exLP, encoded with both sc\_EGFP5 and 6 scaffolds. Indeed, delivery of these designs demonstrated up to 2-fold transfection enhancement (FIG. 4b) and up 6-fold and 9-fold increased expression efficiency, respectively, as measured by MFI (FIG. 4c, d) compared to the original 20HB design using the sc\_EGFP1 scaffold.

[0133] The inventors plotted the transfection efficiency achieved against the fold-change EGFP MFI (A.U.) for all designs, relative to the sc\_EGFP1 20HB-ext as an internal control (FIG. 5a). The samples separate into three clusters. Cluster 1 includes objects built from sc\_EGFP1/2 scaffolds, which have a high quality of folding and high transfection efficiencies, but featured low overall gene expression. Cluster 2 includes objects based on sc\_EGFP3/4/5/6 scaffolds which had low quality of folding and low transfection efficiencies but enhanced gene expression. Cluster 3 then had objects with improved folding quality, as is the case with 20HB-exLP for sc\_EGFP5 and 6 scaffolds, and improved transfection efficiency and improved gene expression. Finally, the inventors further optimized the transfection efficiency by titrating the amount of material administered and by varying the electroporation conditions, resulting in even higher transfection efficiencies (~80%) and MFIs (FIG. 5b).

#### Encoding Active Nuclear Import Into DNA Origami

[0134] The inventors designed and investigated DNA origami objects inherently encoded with instructions for active nuclear import within mammalian cells using a mCherry gene expression cassette to enable easy fluorescent read-out to assess successful nuclear access (FIG. 6). The inventors designed DNA origami scaffolds to encode for mCherry expression and DNA nuclear targeting sequences (DTS) within the origami structures. The custom ssDNA scaffold was designed to include a CMV promoter, mCherry reporter gene, and a polyA signal, to be encoded in the 5' to 3' direction (coding strand). The inventors chose to use the 72 bp Simian virus 40 (SV40) DTS.sup.17, and incorporated either 0×, 1×, 3× or 6×SV40 repeats into the scaffold design (FIG. 6). The inventors designed a 20-helix bundle (20HB) and orientated the scaffold so that the gene features and DTS sequences are on the exterior helices of the objects.

[0135] Four DNA origami structures, corresponding to each of the custom scaffolds, were folded and purified. The 20HB-mCh and 20HB-1×SV40 folded at high yield with a clear leading band and without major structure impurities, while 20HB-3×SV40 and 20HB-6×SV40 displayed some structural impurities (FIG. 7). The 20HB-3×SV40 structure demonstrated low level presence of higher order bands. The higher order impurities were even more prominent within the 20HB-6×SV40 object and so this object was further gel purified. The inventors attribute the difficulty in folding to the increasing number of repetitive DTS sequences in the scaffold impeding the folding.

[0136] The inventors were interested in both dividing cells, and cells arrested at the G1/S phases in the cell cycle to avoid passive nuclear uptake during mitosis. The inventors tested whether the cell cycle arrest leads to the expected decrease in gene expression, as it should inhibit occurrence of passive nuclear transport. For this purpose, the inventors transfected both dividing and arrested cells with the 20HB-mCh via electroporation. Cells were analyzed qualitatively by fluorescence microscopy and quantified by flow cytometry. The inventors observed a statistically significant decrease in the percentage of mCherry+ cells after electroporation in the arrested cell population (FIG. 8). Next, the 20HB variants, 20HB-mCh, 20HB-1×SV40, 20HB-3×SV40 and 20HB-6×SV40 were tested in both in dividing and chemically arrested HEK293T. Cells were quantitatively

assessed via flow cytometry for proportion of mCherry+ cells (%) and MFI (A.U.) to signify gene expression levels, and values were compared to the control, 20HB-mCh (FIG. 9). In dividing cells, the inventors observed a slight increase in both the percentage of mCherry+ cells and the MFI for both 20HB-1×SV40 and 20HB-3×SV40. Inclusion of the SV40 DTS sequences had a greater effect in arrested cells, where the percentage of mCherry+ cells increased for the 20HB-1×SV40 (~1.4-fold), and even further for the 20HB-3×SV40 (~1.8-fold), when compared to 20HB-mCh. The MFI showed a similar trend with a 3-fold increase for the 20HB-1×SV40 and 4.5-fold increase for 20HB-3×SV40 in chemically arrested cells. Interestingly, 20HB-6×SV40 demonstrated consistently lower proportions of mCherry+ cells and MFI of mCherry expression. The inventors attribute this observation to the lower folding quality of the sample.

#### Multiplexed Gene Assemblies for Cotransfection.

[0137] The inventors designed DNA origami objects encoding for either mCherry or EGFP expression to enable assembly and delivery in stoichiometric ratios of 1:1, 1:2 and 1:3 mCherry to EGFP (FIG. 10a, b, using scaffolds sc\_mCherry5 and sc\_EGFP5). The individual gene blocks were programmed to interact with each other via shape-complementary docking sites. sup.18 lined with sequence-complementary sticky-ends which were either five or eight base pairs long (referred to as 5 nt or 8 nt sticky ends, respectively). The inventors also made control objects that had deactivated docking sites passivated with five thymidine long single-strand overhangs (FIG. 10c). Gene assemblies of either dimer, trimer or tetramers formed as designed, as we saw by negative-staining TEM tomography and AGE (FIG. 10b). The cotransfection efficiency for a 1:1 stoichiometric mixture of non-connected mCherry and EGFP monomers was ~5.4 +/-1.4%. By contrast, the inventors observed ~17.5 +/-2.9% cotransfection efficiency when using a pre-assembled dimer object that included both mCherry and EGFP as expressible genes (FIG. 10d, f). The near-four-fold increase in cotransfection relative to when delivering the genes in separate objects indicates that the delivery and expression of both components is now linked, and no longer occurs at random.

[0138] Finally, the inventors delivered the multimeric origami objects in the form of a dimer, trimer and tetramer with the ratios of 1:1, 1:2 and 1:3 mCherry:EGFP. The molar concentration of the multimeric origami objects was conserved across samples and the total cotransfection efficiency thus remained comparable (FIG. 10e, black bars). However, the expression level of EGFP was directly proportional to the number of monomers present within the object (FIG. 10e (green bars)). Direct imaging of cells using two-channel fluorescence microscopy agreed with the observations made in flow cytometry (FIG. 10f). Therefore, the inventors succeeded delivering and expressing genes in a designed, stoichiometric ratio simply by “clicking” the genes together in a higher-order DNA origami assembly.

#### Example 5: Discussion

[0139] Here the inventors have investigated gene expression from encoded DNA origami structures. The inventors present scaffold and structural design features allowing for highly efficient gene expression. The nanostructures of the invention are highly effective tools for therapeutic gene delivery applications. In addition to the electroporation described above, the nanostructures of the invention can be delivered using techniques other than electroporation, e.g. techniques that enable even further optimization of transfection efficiency and gene expression. Instead or in addition to electroporation, inclusion of chemical moieties, aptamers, peptides or antibodies on the origami surface can be used for targeted delivery and gene expression. The nanostructures of the invention are highly valuable for therapeutic applications, for example comprising designing scaffolds with the inclusion of sequences encoding for expression of therapeutic proteins, or gene-editing technology such as CRISPR-Cas allows for therapeutic gene delivery and vaccines. Furthermore, the inventions provides a valuable tool for probing the intracellular, or in vivo fate, of DNA nanotechnology which has, to date, proved to be difficult due to the attachment of tracking molecules being largely on the staples, rather than the scaffold.

#### Example 6: Exemplary Sequences

## SEQ ID NOS 1-11: Scaffold Sequences

[0140] Sequences of scaffolds sc\_EGFP1 (SEQ ID NO 1); sc\_EGFP2 (SEQ ID NO 2); sc\_EGFP3 (SEQ ID NO 3); sc\_EGFP4 (SEQ ID NO 4); sc\_EGFP5 (SEQ ID NO 5); sc\_EGFP6 (SEQ ID NO 6), sc\_mCherry5 (SEQ ID NO 7); sc\_mCherry (SEQ ID NO 8); sc\_mCherry\_1×SV40 (SEQ ID NO 9); sc\_mCherry\_3×SV40 (SEQ ID NO 10); and sc\_mCherry\_6×SV40 (SEQ ID NO 11).

## SEQ ID NOS 12-21: Specific Features

[0141] Sequences of CMV promoter/enhancer (SEQ ID NO 12), 5' ITR (SEQ ID NO 13), 3' ITR (SEQ ID NO 14), Chimeric Intron (SEQ ID NO 15), Kozac (SEQ ID NO 16), WPRE (SEQ ID NO 17), EGFP (SEQ ID NO 18), bGH polyA (SEQ ID NO 19), mCherry (SEQ ID NO 20), and SV40 (SEQ ID NO 21).

SEQ ID NOS 22-75: Primer sequences (see also Table 1).

SEQ ID NOS 76-1507: Exemplary staple strand sequences

SEQ ID NOS 76-177: Staple sequences for sc\_EGFP1; 20HB-ext.

SEQ ID NOS 178-279: Staple sequences for sc\_EGFP1; 20HB-ext-W.

SEQ ID NOS 280-381: Staple sequences for sc\_EGFP1; 20HB-int.

SEQ ID NOS 382-483: Staple sequences for sc\_EGFP1; 20HB-int-W.

SEQ ID NOS 484-584: Staple sequences for sc\_EGFP1; 32HB.

SEQ ID NOS 585-692: Staple sequences for sc\_EGFP1; 12HB.

SEQ ID NOS 693-766: Staple sequences for sc\_EGFP1; 20HB-LS.

SEQ ID NOS 767-863: Staple sequences for sc\_EGFP1; 20HB-LP.

SEQ ID NOS 864-957: Staple sequences for sc\_EGFP1; 20HB-LPv2.

SEQ ID NOS 958-1055: Staple sequences for sc\_EGFP1; 20HB-Circ.

SEQ ID NOS 1056-1157: Staple sequences for sc\_EGFP2; 20HB.

SEQ ID NOS 1158-1279: Staple sequences for sc\_EGFP3; 20HB.

SEQ ID NOS 1280-1392: Staple sequences for sc\_EGFP4; 20HB.

SEQ ID NOS 1393-1507: Staple sequences for sc\_EGFP5; 20HB-exLP.

## REFERENCES

[0142] 1. Engelhardt, F. A. S. et al. Custom-Size, Functional, and Durable DNA Origami with Design-Specific Scaffolds. *ACS Nano* 13, 5015-5027 (2019). [0143] 2. Praetorius, F. et al. Biotechnological mass production of DNA origami. *Nature* 552, 84-87 (2017). [0144] 3. Stahl, E., Martin, T. G., Praetorius, F. & Dietz, H. Facile and Scalable Preparation of Pure and Dense DNA Origami Solutions. *Angewandte Chemie* 126, 12949-12954 (2014). [0145] 4. Wagenbauer, K. F. et al. How We Make DNA Origami. *ChemBioChem* 18, 1873-1885 (2017). [0146] 5. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676-682 (2012). [0147] 6. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *Journal of Structural Biology* 116, 71-76 (1996). [0148] 7. Bastings, M. M. C. et al. Modulation of the Cellular Uptake of DNA Origami through Control over Mass and Shape. *Nano Lett.* 18, 3557-3564 (2018). [0149] 8. Wang, P. et al. Visualization of the Cellular Uptake and Trafficking of DNA Origami Nanostructures in Cancer Cells. *J. Am. Chem. Soc.* 140, 2478-2484 (2018). [0150] 9. Gerling, T., Kube, M., Kick, B. & Dietz, H. Sequence-programmable covalent bonding of designed DNA assemblies. *Science Advances* (2018). [0151] 10. Mitchell, D. L., Vaughan, J. E. & Nairn, R. S. Inhibition of transient gene expression in Chinese hamster ovary cells by cyclobutane dimers and (6-4) photoproducts in transfected ultraviolet-irradiated plasmid DNA. *Plasmid* 21, 21-30 (1989). [0152] 11. Jiang, Y., Ke, C., Mieczkowski, P. A. & Marszalek, P. E. Detecting Ultraviolet Damage in Single DNA Molecules by Atomic Force Microscopy. *Biophys J* 93, 1758-1767 (2007). [0153] 12. Wei, L. & Ploss, A. Hepatitis B virus cccDNA is formed through distinct repair processes of each strand. *Nat Commun* 12, 1591 (2021). [0154] 13. Acevedo, J. M., Hoermann, B., Schlimbach, T. & Teleman, A. A. Changes in global translation elongation or initiation rates shape the proteome via the Kozak sequence. *Sci Rep* 8, 4018 (2018). [0155] 14. Brun, S., Faucon-Biguier, N. & Mallet, J. Optimization of transgene

expression at the posttranscriptional level in neural cells: implications for gene therapy. *Molecular Therapy* 7, 782-789 (2003). [0156] 15. Ping, H., Liu, X., Zhu, D., Li, T. & Zhang, C. Construction and Gene Expression Analysis of a Single-Stranded DNA Minivector Based on an Inverted Terminal Repeat of Adeno-Associated Virus. *Mol Biotechnol* 57, 382-390 (2015). [0157] 16. Cao, L., During, M. & Xiao, W. Replication competent helper functions for recombinant AAV vector generation. *Gene Ther* 9, 1199-1206 (2002). [0158] 17. Dean, D. A., Dean, B. S., Muller, S. & Smith, L. C. Sequence Requirements for Plasmid Nuclear Import. *Experimental Cell Research* 253, 713-722 (1999). [0159] 18. Gerling, T., Wagenbauer, K. F., Neuner, A. M. & Dietz, H. Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components. *Science* 347, 1446-1452 (2015).

[0160] The features of the present invention disclosed in the specification, the claims, and/or in the accompanying figures may, both separately and in any combination thereof, be material for realizing the invention in various forms thereof.

## Claims

1. A nucleic acid nanostructure for gene expression in a eukaryotic cell comprising at least one single-stranded DNA (ssDNA) scaffold strand and a plurality of staple strands, wherein said scaffold strand comprises said at least one nucleic acid sequence encoding a gene, wherein said nucleic acid encoding a gene is part of an expression cassette comprising a promoter, a terminator, and a polyadenylation signal sequence; and wherein said gene is expressed from the scaffold of the nucleic acid nanostructure.
2. The nucleic acid nanostructure according to claim 1, wherein said nucleic acid nanostructure, preferably said at least one scaffold strand, comprises a second nucleic acid sequence encoding a gene; wherein, optionally, said nucleic acid nanostructure, preferably said at least one scaffold strand, comprises a plurality of nucleic acid sequences encoding a gene.
3. The nucleic acid nanostructure according to claim 1, wherein said nucleic acid nanostructure comprises a first subunit and a second subunit; wherein, preferably, said first subunit and said second subunit each comprise a nucleic acid sequence encoding a gene.
4. The nucleic acid nanostructure according to claim 1, wherein said nanostructure comprises an enhancer staple strand having a length in a range of from about 60 to about 250 nucleic acid bases, preferably from about 80 to about 220 nucleic acid bases, more preferably from about 90 to about 200 nucleic acid bases; wherein, optionally, said enhancer staple strand is configured such that it binds to said scaffold strand at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene.
5. The nucleic acid nanostructure according to claim 1, wherein said nanostructure, preferably said scaffold strand and/or at least one staple strand of said plurality of staple strands, comprises a nucleic acid sequence configured to form a loop structure, preferably an inverted-terminal repeat nucleic acid sequence configured to form a hairpin.
6. The nucleic acid nanostructure according to claim 5, wherein said nucleic acid sequence configured to form a loop structure is configured such that a loop is formed at a 5' end and/or 3' end of said nucleic acid sequence encoding a gene.
7. The nucleic acid nanostructure according to claim 1, wherein said nanostructure, preferably said at least one scaffold strand, comprises at least one nuclear targeting sequence, preferably a DNA nuclear targeting sequence, more preferably a simian virus 40 DNA nuclear targeting sequence; wherein, optionally, said nanostructure, preferably said at least one scaffold strand, comprises a plurality of nuclear targeting sequences.
8. The nucleic acid nanostructure according to claim 1, wherein said nanostructure, preferably said at least one scaffold strand, comprises an intron, a kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element.
9. The nucleic acid nanostructure according to claim 1, wherein said nanostructure has an aspect

ratio in the range of from about 1:1 to about 1000:1, preferably 1.5:1 to about 20:1, more preferably from about 2:1 to about 15:1.

**10.** The nucleic acid nanostructure according to claim 1, wherein said scaffold strand further comprises a nucleic acid sequence configured to form a loop structure, an intron, a DNA nuclear targeting sequence, a Kozak sequence, and/or a woodchuck hepatitis virus posttranscriptional regulatory element.

**11.** A composition, preferably pharmaceutical composition, comprising a nucleic acid nanostructure according to claim 1.

**12.** A collection of nucleic acid sequences or collection of plasmids encoding a nucleic acid nanostructure according to claim 1.

**13.** A method of preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder; optionally for use in gene therapy and/or immunotherapy wherein said method comprises administering a nucleic acid composition according to claim 1.

**14.** A method of expressing a gene from a nucleic acid nanostructure of claim 1, comprising i) providing a nucleic acid nanostructure of claim 1; ii) delivering said nucleic acid nanostructure provided in step i) to a cell; wherein, preferably, said delivering comprises transfecting or transforming said cell; iii) allowing said cell to express said gene; wherein, optionally, said providing in step i) comprises providing a plasmid, preferably a phagemid, or a collection of plasmids, preferably a collection of phagemids, wherein said plasmid or collection of plasmids encodes said nucleic acid nanostructure, and preparing said nucleic acid nanostructure using said plasmid or collection of plasmids, preferably by using bacteriophages.

**15.** Use of a nanostructure according to claim 1 for gene expression, preferably for in vitro gene expression.

**16.** A method of preventing, treating, and/or diagnosing a disease or disorder, preferably a genetic and/or immunological disease or disorder, optionally for use in gene therapy and/or immunotherapy, wherein said method comprises administering a composition according to claim 11.

---