Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

Click chemistry for the modification of oligonucleotides and their applications

Stefano Croce

Aus

Rome, Italy

2020

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung von 28. November 2011 von Herrn Prof. Dr. Thomas Carell betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde selbstständig, ohne unerlaubte Hilfe erarbeitet.

München, den ……20.01.2021……..

Stefano Croce

Dissertation eingereicht am: ………27.01.2021………..

1 Gutachter   Prof. Dr. Thomas Carell

2 Gutachter   Dr. Pavel Kielkowski

Mündliche Prüfung am: …………26.02.2021………….

To my family

Acknowledgments

I am extremely grateful to P.D. Dr. Thomas Frischmuth for his kindness, priceless advices and support. As well for giving me the opportunity to work at baseclick.

I can´t express my gratitude to Prof. Dr. Thomas Carell, who kindly welcomed me as an external student, included me in his group and opened the doors of his laboratories and activities.

I would like to thank Dr. Antonio Manetto, for transmitting me his passion for science and for selecting me for this PhD student position at baseclick.

A special thanks goes also to Dr. Sascha Serdjukow, with whom I shared the office, I will treasure all the discussions we had, scientific and not.

I couldn't be happier to work with great collagues at baseclick. It is always a pleasure to spend the day with all of them.

It was a great pleasure to be part of the Carell group at the LMU, in here I did't find only collegues but also friendship that I hope will last.

I want to thank Dr. Georgia Menounou, who has always supported me with love during the period of the thesis. I will always remember our wonderful time together.

Finallay I need to thank my family, my mother Anna Rita, my sister Roberta and my brother Marco for their vital support.

List of Publications

Publications

- Ilenia Manuguerra[#], **Stefano Croce[#]**, Afaf H. El-Sagheer, Abhichart Krissanaprasit, Tom Brown, Kurt V. Gothelf and Antonio Manetto*, Gene Assembly via One-Pot Chemical Ligation of DNA Promoted by DNA nanostructure*, ChemComm, **2018**, 54, 4529-4532.
  ([#]Co-first author)

- Sofia Raniolo, **Stefano Croce**, Rasmus P. Thompsen, Anders H. Okholm, Valeria Unida, Federico Iacovelli, Antonio Manetto, Jørgen Kjems, Alessandro Desideri, Silvia Biocca, *Cellular uptake of covalent and non-covalent DNA nanostructures with different size and geometry*, Nanoscale, **2019**, 11, 10808-10818

- **Stefano Croce**, Sascha Serdjukow, Thomas Carell and Thomas Frischmuth, *Chemo-Enzymatic Preparation of Functional Click Labelled Messenger RNA*, *ChemBioChem*, **2020**, 21, 1641-1646

- Nada Raddaoui[#], **Stefano Croce[#]**, Florian Geiger, Alexander Borodavka, Leonhard Möckl, Samuele Stazzoni, Bastien Viverge, Christoph Bräuchle, Thomas Frischmuth, Hanna Engelke, and Thomas Carell, *Super-sensitive multi-fluorophore RNA-FISH using click chemistry*, ChemBioChem, **2020,** 21, 2214– 2218
  ([#]Co-first author)

Conferences

- COST Action CM1201 inter-WG meeting, Grenoble, France, April **2016**
  Small talk: *Ultra-stable DNA nanostructure for gene detection*

- Functional DNA nanotechnology workshop (FDN), Rome, Italy, June **2016**
  Poster presentation: *DNA-chainmail nanostructures: enhanced stability and applications*

- Symposium on Chemistry of Nucleic Acid Components (SCNAC), Cesky Krumlov, Czech Republic, June **2017**.
  Poster presentation: *New EdU Cell Proliferation Assays*

- The Complex Life of RNA, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany, October **2018**.
  Poster presentation: Novel mRNA bioconjugation system for mRNA therapy

- Annual meeting of the RNA society, Krakow, Poland, June **2019**
  Poster presentation: *Novel mRNA bioconjugation system for mRNA therapy*

- International mRNA health Conference, Berlin, Germany, November **2019**
  Talk: *Chemo-Enzymatic Methods for Modification of IVT mRNA*

Poster presentation: *Chemo-Enzymatic Methods for Modification of IVT mRNA*

Patent Applications

- WO 2019/121803 A1: *Click-modified mRNA.* Priority applications (3) filed, pending PCT phase

Acknowledgements

Contents

1. Summary

This PhD thesis reports the published work done in the laboratories of baselick GmbH and the Ludwig Maximilians Universität (LMU). baseclick GmbH was founded by Prof. Dr. Thomas Carell (LMU) in collaboration with the chemical company BASF in 2008. Main focus of baseclick is the usage of click chemistry for modification of nucleic acids. In the described research work click chemistry is applied to oligonucleotides and in particular to DNA nanostructures, to mRNA to be then used in drug development and to produce highly labeled probes for fluorescent *in situ* hybridization.

Continuing the previous work done at baseclick[1], at first click chemistry was applied on DNA nanostructures. In this field DNA, is not used as a carrier of genetic information but as material for the production of structures with different size, geometry and shape[2–8]. The main concept behind the technology is based on Watson and Crick base pairing interactions, which bring portions of a ssDNA to hybridize with a complementary sequence, usually of another ssDNA strand, to form a rigid dsDNA helix. This is used in the rational design of the sequences to form complex structures with nanometric precision[9,10].

DNA nanostructures and click chemistry were used to find an alternative to the current state of the art methods for gene synthesis *in vitro*. To date enzymatic synthesis of long DNA fragment is the method of choice since solid phase synthesis is not suitable for very long sequences[11,12]. Enzymatic synthesis approaches are based on the activity of either DNA polymerases or DNA ligases reactions[13,14], but those methods suffer from some limitations: *e.g.* in the case of ligases the final gene is assembled by overlapping of strands then ligated to form a longer fragment, but it starts to be insufficient when a big number of strands need to be ligated together. With DNA polymerases the final product is formed by different cycles of the enzyme in a multiple step assembly, with the limitation due to the mispriming and formation of secondary structures which then lead to errors.

Therefore in here, in collaboration with the group of Prof. Dr. Tom Brown from University of Oxford, a method was developed where, with the help of the DNA origami technique, ligase activity is replaced by chemical ligation, in this case click chemistry. 14 oligonucleotides were designed and synthesized with a 5'-terminal azide and 3'-terminal alkyne. The oligos where then preorganized in a DNA nanostructure to bring the alkyne and azide in close proximity and, most importantly, in a selective order. In this way after the click reaction occurs, the full length of the defined sequence is established, with a bio-compatible triazole linkage replacing the phosphate bond at the point of the oligo connections.

In a second project click chemistry was used to stabilize a DNA nanostructure, in this case composed of 24 different interlocked oligonucleotides[15], and at the same time achieve selective labeling with biotin molecules in a one pot reaction. This work was done in collaboration with the group of Prof. Dr. Silvia Biocca of University of Rome Tor Vergata. DNA nanostructures thanks to their properties such as bio-compatibility, non-toxicity and bio-degradability have been used for different applications: *e.g.* drug delivery,

nanocontainer, cellular biosensor and *in vivo* imaging[16–19]. Anyhow, crucial for these applications is the understanding of how different DNA nanostructure enter the mammalian cells[3]. For this reason, in this work five different topological configurations and functionalizations, with size varying from 8 to 80 nm and shape from tetrahedral, octahedral, cylindrical, square box and rectangular, have been investigated for their ability to interact with the scavenger receptor LOX-1, which overexpression has been associated with tumor development in many cancer cells.

Inspired by the big success that mRNA therapy[20] has in the last decade, methods to enable modification of very long RNA oligonucleotides, such as mRNA, were established using click chemistry. *In vitro* transcribed (IVT) mRNA, consisting in mRNA produced by RNA polymerases from a DNA template, is nowadays considered to be a valid candidate for a novel class of drugs. It was already demonstrated to be efficient in several diseases including: vaccination, protein replacement and cancer therapy[21–25]. Indeed nowadays mRNA is playing a central role in vaccination programs against SARS-Cov-2 pandemic. The main idea behind the mRNA therapy is to provide IVT mRNA to the patients to help them developing their own cure. For example in vaccination, mRNA coding for a specific viral antigen is used to produce an immune response leading to the immunogenicity.

Besides stability issues deriving from using RNA molecule as drugs, another problem arises from the cellular delivery of such molecules[26–28]. Indeed, delivery to a specific target is still an unsolved problem. To date the mRNA is delivered to patients using lipid nanoparticle (LNPs) to act as carrier and at the same time as a barrier from the extracellular environment[29–31]. Recent studies demonstrate however that the usage of LNP is not the ideal method to deliver mRNA. It has been proven to be less efficient *in vivo* than what was observed *in vitro*[32]. Especially in living organisms the main destination is the liver, which is often not the final target for the therapy. Also driven by the recent FDA approval of the first siRNA drug (GIVLAARI[TM]), where the siRNA has been chemically modified using N-acetylgalactosamine molecules (GalNac) to enable efficient targeted delivery[33], it is described here a chemoenzymatic approach based on the incorporation of modified nucleotides bearing an alkyne or azide moiety (such as 5-ethynyl-UTP or 3'-azido-dd-ATP), that can then be labelled post-transcriptionally, using click chemistry. This method allows for example the incorporation of specific modifications inside the mRNA that would not be accepted by the RNA polymerases, e.g. targeting-molecules for specific delivery, or fluorescent dyes for tracking, thus potentially improving the biochemical properties of the mRNA.

Click chemistry was also used to improve the current methods for the preparation of probes used in fluorescent *in situ* hybridization (FISH). FISH is a cytogenetic analysis that allows the detection and the spatial localization of specific nucleotide sequences[34–38] in tissues or cells. The fluorescent probes, consisting of ssDNA, were designed to hybridize only to those parts of the target DNA/RNA with a high degree of complementarity. Then by utilizing fluorescent microscopy it was possible to localize where the fluorescent probes are hybridized. This technique is largely used for diagnosis of genetic abnormalities such as gene

fusion, aneuploidy, loss of chromosomal regions, detection of oncogenes and diagnosis of viral infections and to date it can also detect other targets such RNA (mRNA) in cells and tissue samples[38,39].

The probes can be prepared in various ways, such as nick translation, random priming, PCR, end labelling or NHS-ester chemistry. Most of the probes, especially for RNA detection, are composed of ssDNA which are approximately 20-25 nt long, conjugated to a fluorophore *via* coupling of an amino group introduced at the 3' end and an activated ester of the fluorophore[40].

Finally a method for preparation of mRNA FISH probes based on click chemistry is described, where each probe contains three fluorophore instead of a single one, thus giving an increment of fluorescent yields and allowing microscopy analysis without the need of special deconvolution software. Furthermore this allows the detection using fluorescent activated cell sorting (FACS) devices. Enabling FACS analysis is of outmost importance especially for clinical studies, where up to now the detection of specific mRNA or chromosomes sequences is still done manually by clinicians, analyzing all the samples through visual inspection.

2. Introduction

2.1. DNA

The deoxyribonucleic acid (DNA) is the central molecule involved in the storage of genetic information used in the growth, development, functioning and reproduction in the majority of living organisms including many viruses[41]. The structure of the DNA was first solved by J.D. Watson and F. H. C. Crick in 1953 helped by the X-ray diffraction analysis done by M. Wilkins and R. Franklin[42]. It is composed of repetitive elements called nucleotides. Each nucleotide is a unit formed by a phosphate residue, a pentose sugar (the 2-deoxyribose) and a planar aromatic heterocyclic base. The DNA is composed of four bases divided as monocyclic pyridines: Cytosine (C) and thymine (T), and bicyclic purines: adenine (A) and guanine (G). The nucleotides are linked together by a covalent bond, the phosphodiester bond, between the sugar of one nucleotide to the phosphate of the next, forming the so called polynucleotides[43]. In a single strand DNA filament (ssDNA) the hydroxyl group at the C3' is called 3'-end, and the phosphate group at the C5' is called 5'-end, thus giving orientation to the DNA strand (5'→3'). In a double stranded DNA filament (dsDNA) the two strands of the DNA run in opposite orientation. The bases of the opposite strands are bound together *via* hydrogen bonds according to the base pairing rules where thymine forms two hydrogen bonds with adenine and guanine forms three hydrogen bonds with cytosine, as shown in Figure 1. This is giving a stabilizing effect on the double helix, thus meaning that increasing the G and C content increases the stabilizing effect. As the hydrogen bond is not a covalent bond they can be relatively easy broken and re-joined, an important aspect that allows, alongside with some biological activities such as replication, also other applications, e.g. fluorescent *in situ* hybridization (FISH) and DNA origami[10,38].
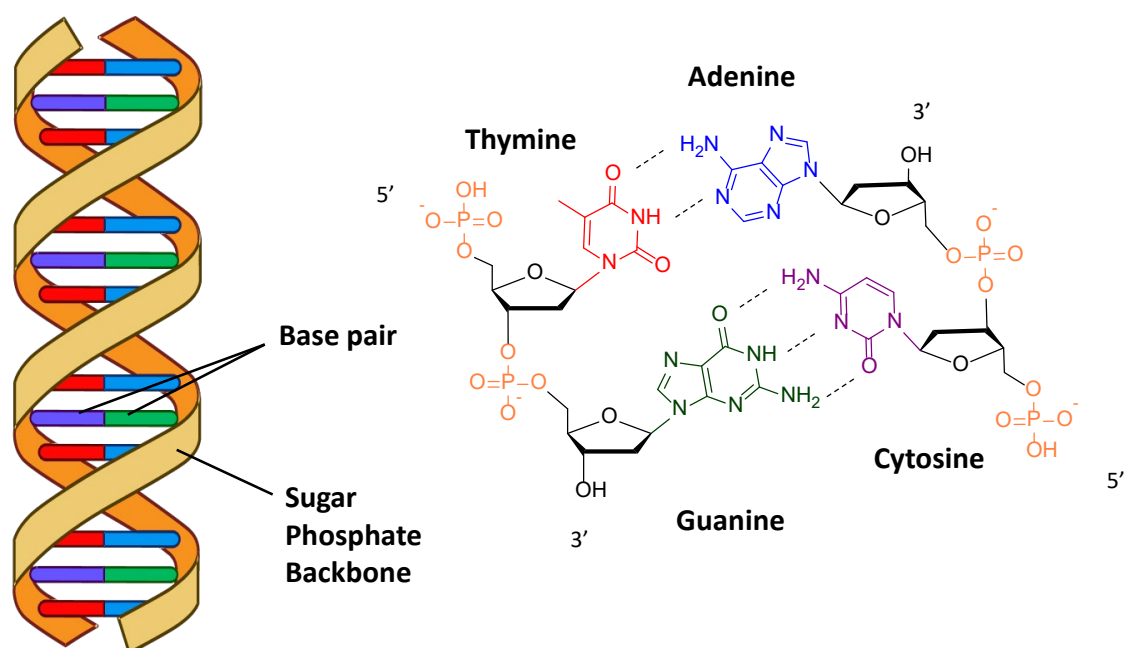

Figure 1: Base pair interaction within a dsDNA.

Hydrogen bonds are not the only stabilizing effect on dsDNA, another stabilizing effect is due to the so called base stacking between adjacent bases[41]. This as well is a noncovalent interaction which gives a major contribution to duplex stability, as it is more prevalent in dsDNA than ssDNA. Base stacking depends on the aromaticity of the bases and is a hydrophobic and electrostatic force. Like for the hydrogen bond, different sequences can influence the degree of stabilization, some combinations give a more stable interaction than others. The salt concentration can influence the base stacking effect as well, high salt concentration, e.g. $MgCl_2$, can mask the destabilizing effect of the repulsive negative charge on the phosphodiester bonds.

DNA is known in three different conformations, A-DNA, B-DNA and Z-DNA, depending among others on hydration levels, DNA sequence, amount of supercoiling and chemical modified bases. However the most common conformation in cells, and used in the projects described below, is the B-DNA[44]. In this conformation the helix is right handed with 10.5 bases per turn of helix, with a distance between adjacent base of 0.34 nm and a helical diameter of 2 nm. The knowledge of the structural properties of DNA is of fundamental importance in applications such as DNA nanotechnology.

2.2. DNA synthesis

Artificial DNA synthesis represents one of the most important advances in molecular biology. The deep understanding of the chemical structure and its replication mechanism allowed scientists to develop different methods to synthetized DNA *in vitro*.

One of the most important strategy is represented by the PCR, developed by K.Mullis[45] in 1985. With the discovery of thermostable DNA polymerases, it was possible to amplify fragments of DNA exponentially during different thermal cycles, generating a high copy number of the original DNA used as template.

Another important approach is represented by chemical synthesis of short ssDNA fragments. By the usage of phosphoramidite chemistry on solid-phase synthesis it is nowadays possible to produce high quality DNA oligonucleotide with defined sequences of up to 100-120 nt[46]. The solid supports are insoluble materials where the oligonucleotide is coupled during synthesis. Different types of solid support have been tested and so far, controlled pore glass (CPG) and polystyrene turned out to be the most used ones. The synthesis is based on the usage of nucleoside 3'-(2-cyanoethyl-N,N-diisopropylphosphoramidite), phosphoramidites as building blocks, added one by one per synthetic cycle in a sequence specific manner, from 3' to 5' orientation. The phosphoramidite synthesis cycle consists of a series of steps summarized in Figure 2. Using a solid support, with the first nucleoside already on it, the cycle starts with a detritylation reaction in which the 5'-DMTr protecting group (4,4'-dimethoxytrityl) on the first nucleoside is removed. Following detritylation, the nucleoside is ready to react with the next base, after activation with tetrazole (or a derivate). The activated phosphoramidite then reacts with the 5'-hydroxyl group of the nucleoside bound to the support, generating a phosphite triester linkage. Even assuming more than 99% yield during the coupling step, there will be a few unreacted 5'-hydroxyl groups on the support bound nucleotide. For this reason, a capping step is performed

to prevent such unreacted groups to react with the incoming phosphoramidite. At this point the unstable trivalent phosphite triester is oxidized to the corresponding pentavalent phosphotriester, Figure 2. The cycle is repeated until the desired sequence of the oligonucleotide is complete, then the oligonucleotide is removed from the solid support by ammonium hydroxide treatment.
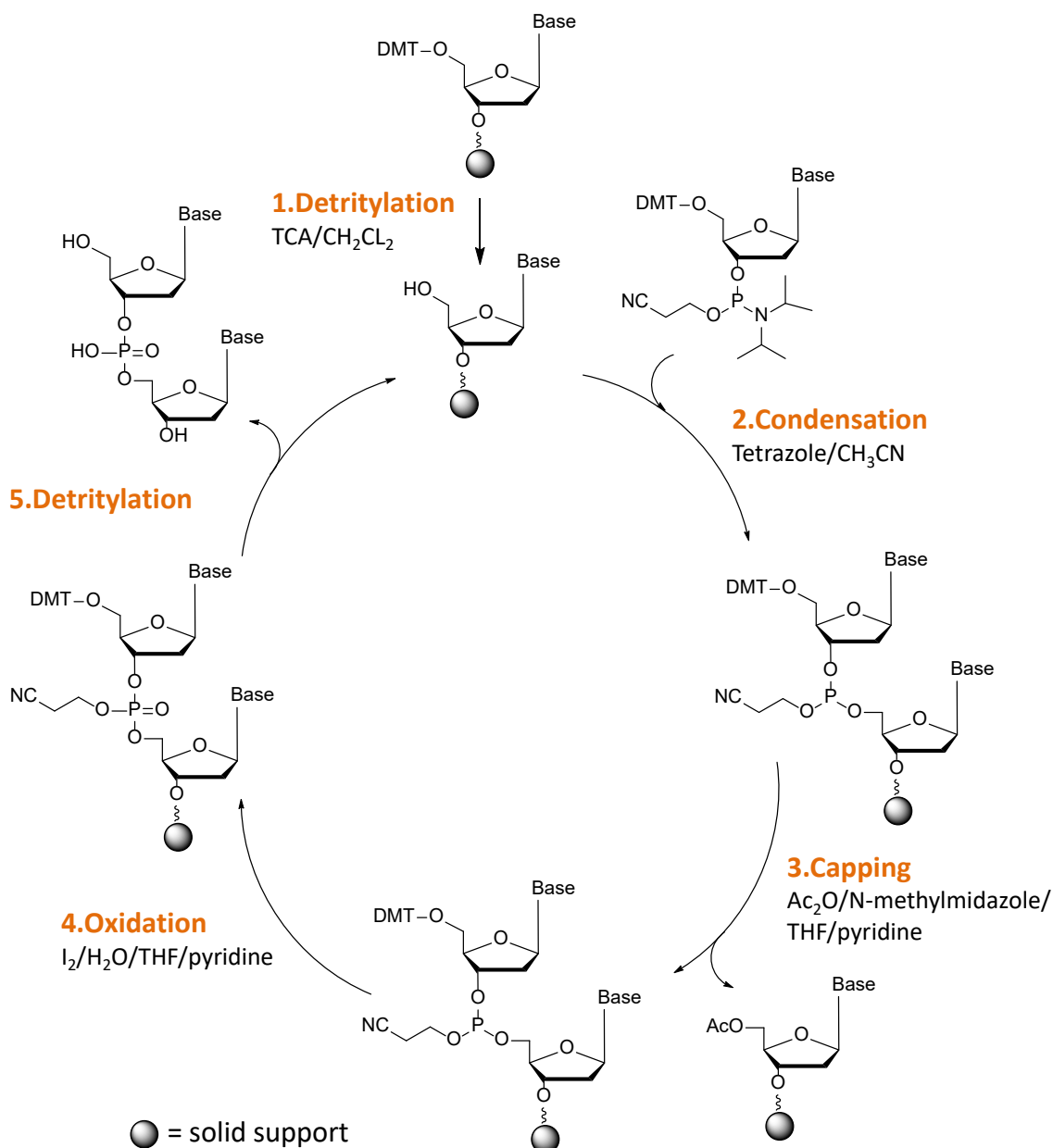
Figure 2: Phosphoramidite synthesis cycle for oligonucleotide synthesis.

2.3. RNA

Some organisms, e.g. retrovirus and flavivirus, use ribonucleic acid (RNA) instead of DNA as depositary of the genetic information[41]. In all other organism, RNA acts in different biological roles like coding, regulation and expression of genes like for example messenger RNA (mRNA), described below, and the transfer RNA (tRNA) that serve as physical link between mRNA and protein sequence during translation. Similar to DNA, RNA is composed of repetitive units of nucleotides as well, but unlike DNA it is found mostly as single strand folded

to secondary structures. However, some secondary structures, due to the complementary sequence within the same strand, are forming regions of dsRNA like for example tRNAs. Another characteristic that differs from DNA is the sugar forming the sugar-phosphate backbone that in this case consists of a ribose instead of a deoxyribose. The presence of the hydroxyl group in 2' position makes the RNA less stable, enhancing the hydrolysis of the adjacent phosphodiester bond. In RNA the complementary base to adenine is uracil (U) instead thymine, Figure 3, forming always two hydrogen bonds.
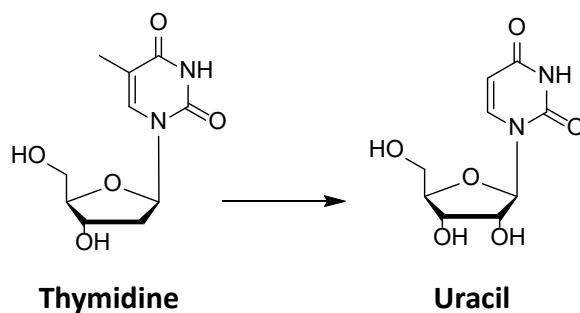

Thymidine   **Uracil**

Figure 3: In RNA thymidine is replaced by uracil, forming anyway always 2 hydrogen bonds in a double strand.

2.3.1. mRNA

mRNA is the molecule involved in the transfer of the genetic information from DNA to ribosomes, the site in the cell where proteins are synthetized in a process called translation[41]. Eukaryotic mRNA is composed of different regions: A 5' cap (or m^{7}G cap), a 5' untranslated region (5'-UTR), an open reading frame (ORF) of the gene, a 3' untranslated region (3'-UTR) and a poly adenine tail[47,48]. The 5' cap consists of a 7-methylguanosine residue linked through a 5'-5' triphosphate bond before the first transcribed nucleotide. The presence of the 5' cap enhances the stability of the mRNA itself and is important for its recognition by the ribosome complex for the start of translation. UTRs are sequences not translated, located before the starting codon (5'-UTR) and after the stop codon (3'-UTR), those regions are often folded into secondary structures and have been assigned to different functions including stability, localization, ribosome binding and translation efficiency by interaction with different proteins. The poly(A) tail is a long repetition of adenine nucleotides (often more than 100 nt) added to the 3' end. It promotes the export from the nucleus and enhance the protection from exonuclease degradation. Furthermore it was observed that poly(A) binding protein (PABP) can interact with the protein complex at the 5' cap and form a loop capable to speed up and improve the protein production by the ribosomes[41]. In the last decade, mRNA has attracted attention as a valid candidate for a new class of therapeutics, such as vaccination, cancer immunotherapy and protein replacement[20,49–51]. Therefore part of this thesis, chapter 5, describes methods for the modification of such mRNA in order to enhance its biological properties.

2.4. Click Chemistry

The term click chemistry, introduced for the first time by Finn and Sharpless in 2001[52], integrates a series of chemical reactions which have common criteria. The so called "click" reactions should be modular, wide in scope, resulting in very high yields, stereospecific and produce only non-toxic byproducts that can be easily removed from the solution *via* non-chromatographic methods. In addition, the reactions could be conducted in aqueous solution and in the presence of oxygen. Typically, the products of a click reaction require simple isolation methods and they are stable under physiological conditions.

The Cu(I)-catalyzed alkyne/azide cycloaddition (CuAAC) is the most representative and widely used click reaction[53]. The alkyne-azide cycloaddition was originally studied by Huisgen in the 1960s at the university of Munich (LMU), where an azide and an alkyne react at high temperature (about 100°C) and with long reactions times, generating a mixture of 1,4 and 1,5 regioisomers, thus limiting the usage of the reaction for practical applications, Figure 4.
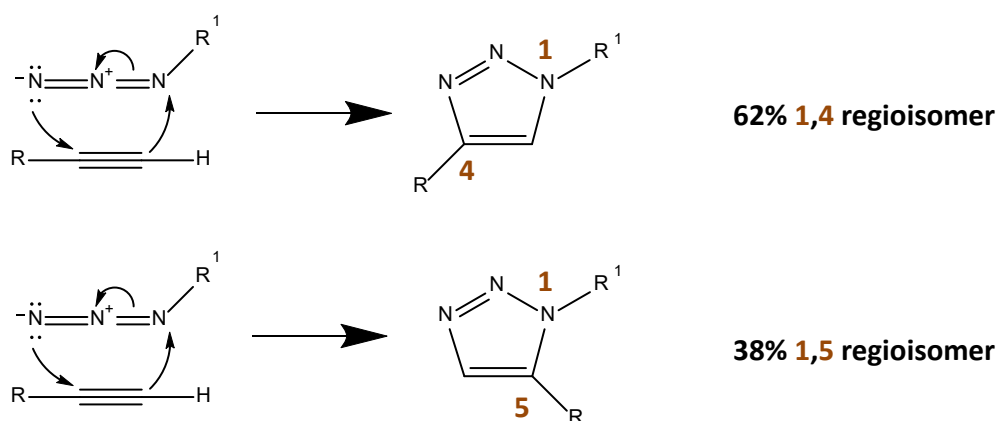

62% 1,4 regioisomer

38% 1,5 regioisomer

Figure 4. Huisgen cycloaddition.

2.4.1. Cu-Catalyzed Azide Alkyne Cycloaddition

In 2002, two different research groups independently improved the regioselectivity and the kinetic of the Huisgen reaction, discovering that Cu(I) catalysis leads to only one regioisomer, the 1,4-triazole. Furthermore the reaction rate is increased, can be carried at ambient conditions and can tolerate a broad range of functional groups[54,55]. Despite the large use of the CuAAC, the mechanism remained difficult to establish because of the multiple equilibria between different reactive intermediates occurring during the reaction. While the classical Huisgen 1,3-dipolar cycloaddition is a one-step process, the CuAAC is considered to be a step-wise process involving copper within the intermediates. At the first step, there is the formation of Cu(I)-alkyne complex. This complex lower the pKa of the alkyne which helps the deprotonation and results in copper acetylide coordinating with the alkyne. In the next step, the azide binds to the copper forming an unconventional Cu(III) metallocycle. This intermediate undergoes then ring contraction to give a copper triazolyl derivative, which upon protonolysis gives the 1,2,3-triazole product[56], Figure 5. Usually, the

8

formation of active Cu(I) species can be generated by Cu(I) or Cu(II) salts by using sodium ascorbate as the reducing agent.
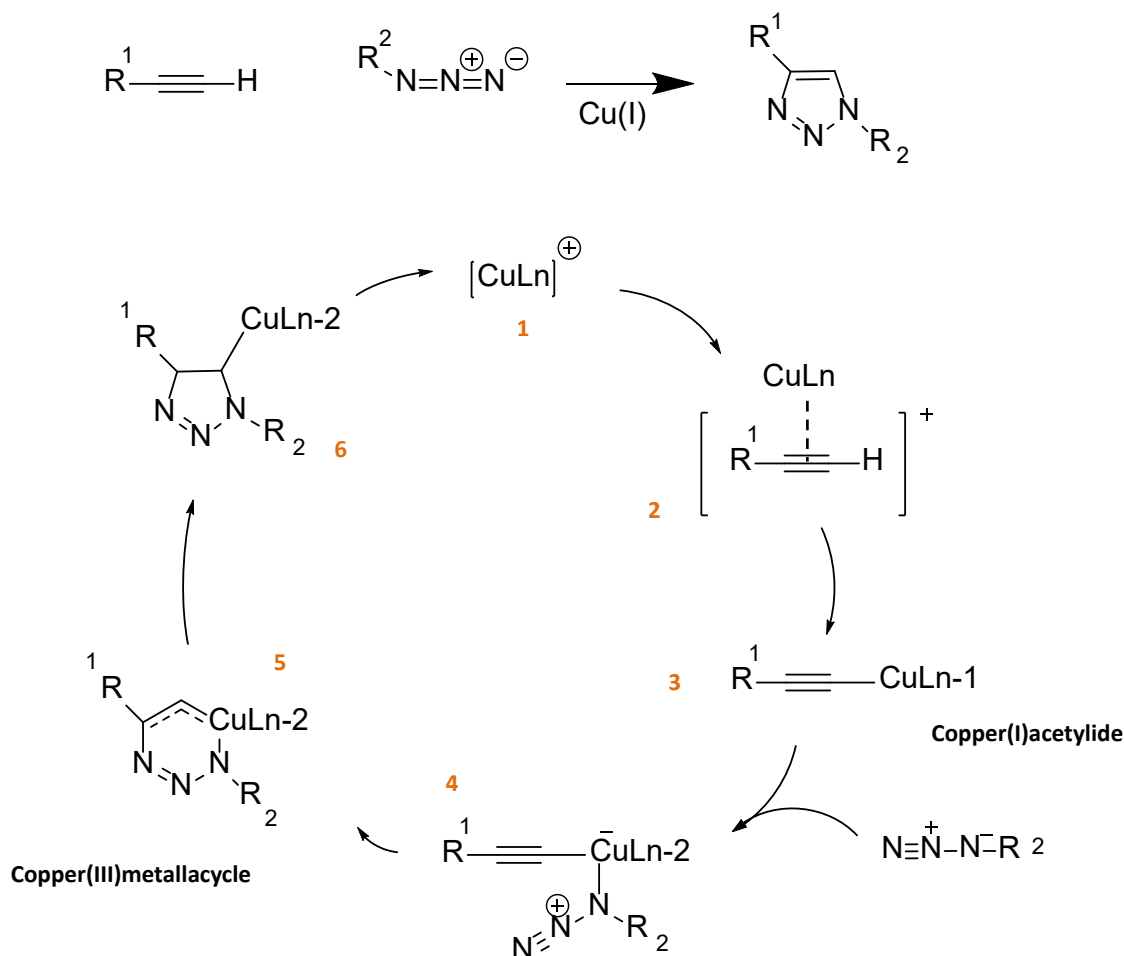
Figure 5. Proposal mechanism for the copper-catalyzed azide-alkyne cycloaddition (CuAAC).

An important finding that contributed to the wide usage of the CuAAC, in particular for biological applications, is that by protecting the Cu(I) species from oxidation with nitrogen-based ligands, the reaction can be further accelerated, increasing the reaction rate up to 10.000 fold. Wu *et al.* introduce the tis (triazolylmethyl) amine (BTTES)[57], that coordinates copper during the reaction, leading to increasing the reaction rates in living systems without apparent toxicity. Other examples of ligands are THPTA: Tris((1-hydroxy-propyl-1H-1,2,3-triazol-4-yl)methyl)amine, TBTA: Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, BTTAA: 2-(4-((bis((1-(tert-butyl)-1H-1,2,3-triazol-4-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)acetic acid[58].

2.4.2. Strain Promoted Azide Alkyne Cyclo Addition

In this variant of the classical CuAAC, without the necessity of any catalyst, a stable triazole product is formed by the reaction of strained cycloalkynes, most often a cyclooctyne, with an organic azide in what takes the

name of strain-promoted azide-alkyne cycloaddition (SPAAC). It is a powerful and versatile chemical reaction with broad academic and commercial applications[59,60].

The stability of a cycloalkyne is directly correlated with the C-C C-C bond angle, which, in a ring form cannot have the ideal 180°C angle for the sp^2-hybridized carbon. An explanation for the spontaneous reaction of the cycloalkyne with the azide lies in the favorable enthalpic release of the ring strain, changing from a strained ring to a fused ring system with more favorable angle for the sp^2-hybridized carbon of the resulting triazole. The usage of the CuAAC can be compromised in the context of biological and especially *in vivo* applications, due to the toxicity of the copper(I) species, for this reason selective chemical reactions such SPAAC, that does not require the Cu catalysis and are orthogonal to the diverse functionality of biological systems have become important in the field of chemical biology[60,61].
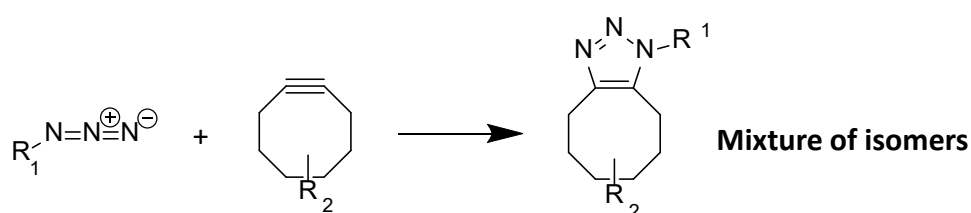

Figure 6: Strain promoted azide alkyne cycloaddition (SPAAC).

2.4.3. Click chemistry on nucleic acids

In the recent years, click chemistry has become a powerful tool in many applications, including drug discovery, cell biology, peptide chemistry and in particular oligonucleotide chemistry [62–65]. The usage of click chemistry on nucleic acid was thoroughly examined by Prof. Dr. Thomas Carell at the LMU Munich, which has led to the founding of the spin-off company baseclick GmbH. The studies from the Carell group[66,67] established the CuAAC as highly efficient method for post synthetic labelling of nucleic acid. Generally, alkyne groups were introduced in the 5-position of the pyrimidines nucleobase and position 7 of purines. Alkyne phosphoramidites were synthetized and used in solid phase synthesis of DNA oligonucleotide. Two alkyne linker were studied in combination with different labels varying in density such as coumarin, fluorescein and sugar azide. Furthermore, the modified deoxynucleoside triphosphate, uridine and cytosine analog, were accepted as substrates in a PCR[67]. In this study more than 800 alkyne groups were introduced in a 2000 bp PCR product, showing quantitative labeling rate with at least 95% reacted with the azido sugar.

Nowadays click chemistry is an extraordinary powerful and versatile reaction used routinely in the field of nucleic acid chemistry, as shown also in this thesis.

3. Aim of the thesis

I joined the baseclick team thanks to a European grant (Horizon 2020 Innovative Training Network, ITN) with the name of ClickGene. As suggested by the name the project was unified by the usage of click chemistry (Click) on nucleic acid (Gene), therefore the aim was to apply click chemistry on oligonucleotides. Herein it is described as a powerful tool that allows the conjugation of two (bio)molecules in different applications: e.g. alternative methods for gene synthesis *in vitro*, labeling and stabilization of DNA nanostructures, IVT mRNA modification, and mRNA detection *via* FISH.

Considering that part of this thesis work was carried out in a company we also aimed to the commercialization of products, endeed it has to be noted that part of the work resulted in commercially available products, such as improved version of the EdU cell proliferation assay and kits for the labeling of RNA molecules using click chemistry.

4. Gene Assembly via One-Pot chemical Ligation of DNA Promoted by DNA Nanostructure

4.1. Prologue

A high quality solid phase synthesis of nucleic acids depends on different factors: on the sequence length, nucleotide composition and purification system[46]. Therefore, syntheses below 100 nucleotides are preferred to achieve reliable oligonucleotides in the required amount using solid phase methods. In order to overcome these limitations and to achieve a production method for very long oligonucleotides or gene fragments, various approaches based on joining multiple oligonucleotides have been developed. Currently two main strategies are considered for gene synthesis from oligonucleotides *in vitro*[68–70], both based on enzymatic methods: the first is based on DNA ligase (Figure 7) and the second on DNA polymerase (Figure 8). In the ligation methods, which represent the earliest example of synthetic gene synthesis, the double-stranded DNA is formed by complementary overlapping strands, subsequently joined by the ligase to produce larger fragments. The ligase requires the 5' ends to be phosphorylated, so the set of oligonucleotides used must be phosphorylated before. An advancement of this technology was the discovery of thermostable DNA ligases and the development of ligase chain reaction, very similar to the polymerase chain reaction, where the gene fragments are produced by several cycles of ligation[13]. This method has become very common and largely used, however once the ligation is finished, the gene needs to be amplified by standard PCR.
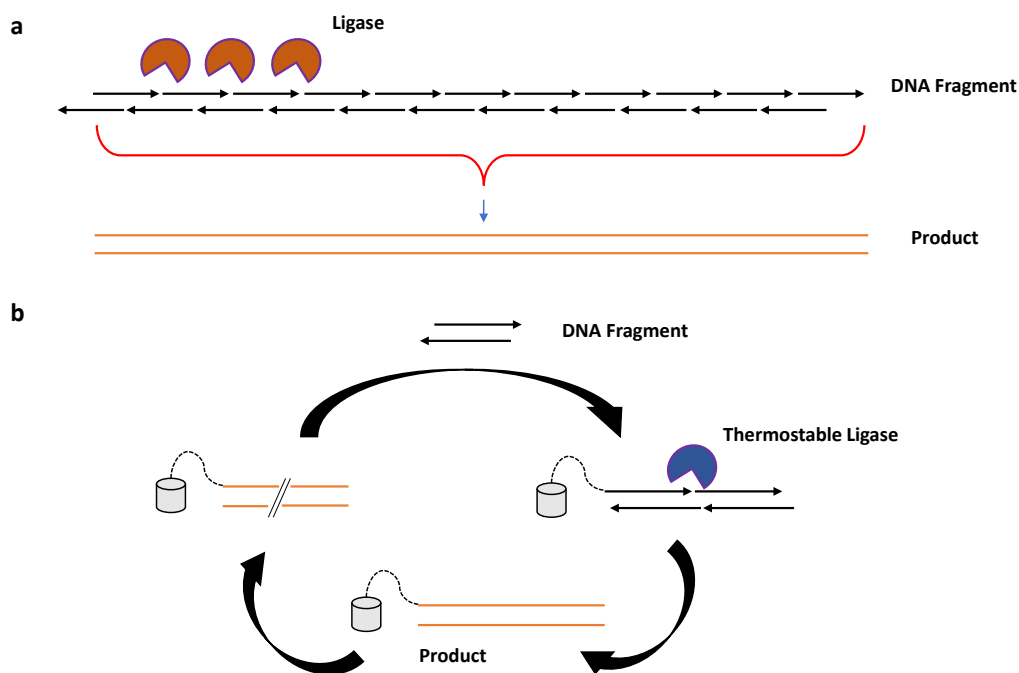

Figure 7. Ligase based methods. **a)** The ligase is used to ligate-assemble overlapping oligonucleotides into a single strand. **b)** A thermostable ligase is used in a stepwise synthesis where the growing DNA sequence is anchored to a solid support and each cycle new pair of oligonucleotides are added.

In the second method, based on DNA polymerase cycling assembly (PCA), the desire gene is produced as multiple step assembly. A solution containing all required oligonucleotides and a DNA polymerase, is subjected to a similar cycling as in the PCR. During each cycle, pairs of oligos anneal and extend, forming a duplex including the sequence between the ends of the two olignucleotides overlapping. Bigger duplex will be formed and after several cycles a full-length fragment will exist (most often as a set PCR products with various lengths). As the ligation technology, the PCA is based on a series of overlapping oligonucleotides, however there is no need that the oligos form the entire double-stranded product but there can be spaces between oligos on the same strand, since during the first PCR assembly reaction, overlapping oligos will anneal and extend using each other as template to form longer DNA fragments. Same as the other method, the PCA does not result in exponential growth of the DNA, so a final PCR reaction is needed to amplify the full lenght product[14]. Although these methods – along with all documented variations - give access to a large variety of DNA fragments, there are some limitations due to mispriming during the assembly, formation of secondary structures or to repetitive sequences that can hinder the polymerase activity. Moreover, in the fast evolving genomic, epigenomic and DNA-nanotechnology fields, the demand for functional gene fragments are not yet matched and alternative approaches are strongly needed. Thus, in our work we propose an alternative to the enzymatic methods for the production of gene fragments, using chemical ligation.
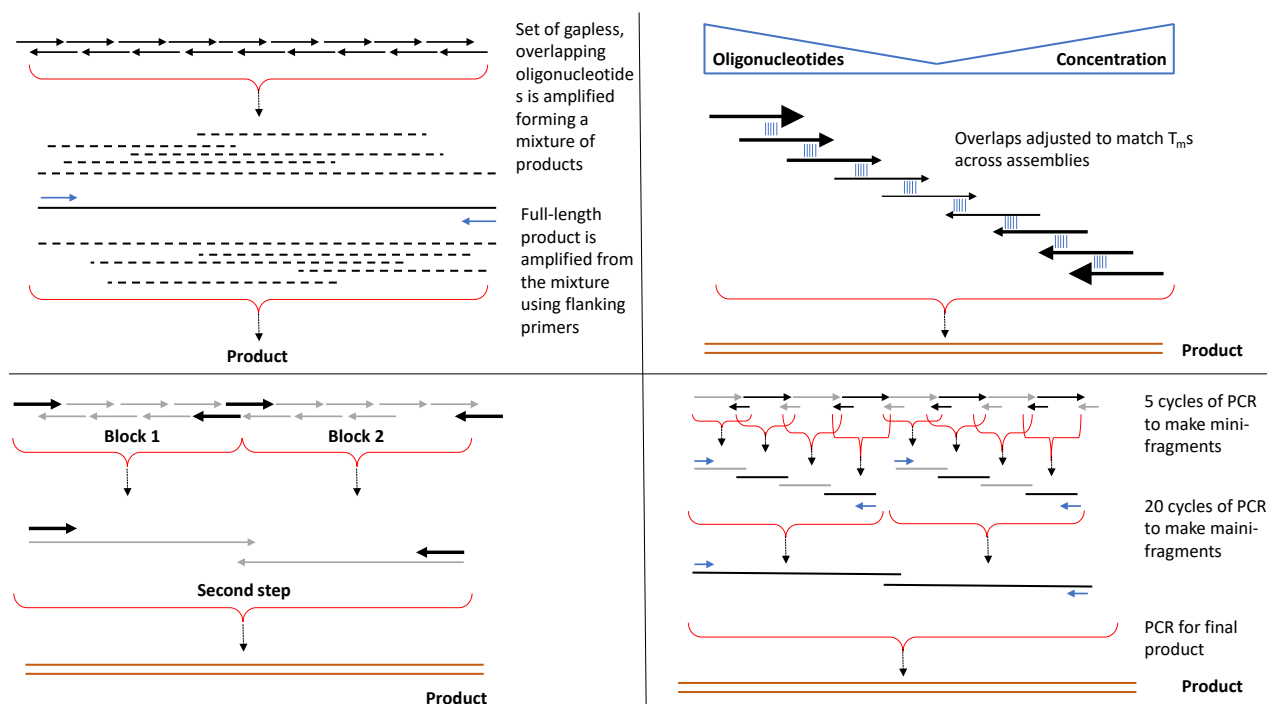
Figure 8. Top left, A series of overlapping oligonucleotides amplify one another to produce a mixture of product, eventually including the full-length gene. Then this mixture is used for a final PCR reaction using primers to selectively amplify only the full length product. Bottom left, in this case the product is produced by at first assembling ca 500bp blocks of overlapping oligonucleotides. Higher concentration of the flanking primers

(expressed here by the thickness of the arrows) facilitate the assembly of the block. Then the blocks are assembled into a full-length gene with a final PCR. Top right, a gradient concentration of the primers resulting in selective pressure for amplification of eventually the full-length gene. Bottom right, this method is based on the formation of mini-fragments by briefly extending primers pair. These pairs then are used to build larger fragments, then assembled in the final product.

4.2. Abstract of the publication " Gene Assembly via One-Pot Chemical Ligation of DNA Promoted by DNA nanostructure"

Current gene synthesis methods are driven by enzymatic reactions. Here we report the one-pot synthesis of a chemically-ligated gene from 14 oligonucleotides. The chemical ligation benefits from the highly efficient click chemistry approach templated by DNA nanostructures, and produces modified DNA that is compatible with polymerase enzymes.

4.3. Author contribution

The author planned the experiments and carried out the folding reactions, the chemical and enzymatic ligations, the PCR experiments and helped in the preparation of the manuscript.

4.4. Associated publication

Gene Assembly via One-Pot Chemical Ligation of DNA Promoted by DNA nanostructure

Ilenia Manuguerra[#], **Stefano Croce[#]**, Afaf H. El-Sagheer, Abhichart Krissanaprasit, Tom Brown, Kurt V. Gothelf and Antonio Manetto*, *ChemComm*, **2018**, 54, 4529-4532. DOI: 10.1039/c8cc00738a

COMMUNICATION

Check for updates

Gene assembly *via* one-pot chemical ligation of DNA promoted by DNA nanostructures†

Ilenia Manuguerra, [ID]‡[a] Stefano Croce, [ID]‡[bc] Afaf H. El-Sagheer, [ID][de]
Abhichart Krissanaprasit,§[a] Tom Brown, [ID][d] Kurt V. Gothelf [ID][a] and
Antonio Manetto [ID]*[f]

Current gene synthesis methods are driven by enzymatic reactions. Here we report the one-pot synthesis of a chemically-ligated gene from 14 oligonucleotides. The chemical ligation benefits from the highly efficient click chemistry approach templated by DNA nanostructures, and produces modified DNA that is compatible with polymerase enzymes.

The DNA nanotechnology and synthetic biology fields rely on synthetic oligonucleotides that are assembled to form nanostructures and artificial genetic systems. Synthesis of high quality oligonucleotides by solid phase synthesis[1] depends among other factors on the sequence length, nucleotide composition and the purification system used. Yields exceeding 99% are not rare for each coupling step, although even the most efficient synthesis setup cannot reach 100% coupling efficiency. Therefore, the overall percentage yield of oligonucleotides strongly depends on their length. For instance, the synthesis of a 200-mer, where each cycle has an efficiency of incorporation of 99%, yields 13% of full-length product without taking in account further purification steps. Therefore, synthesis of oligonucleotides that are shorter than 100 nts is preferred in order to achieve reliable yields.

[a] *Interdisciplinary Nanoscience Center (iNANO), Gustav Wieds Vej 14 and Department of Chemistry, Aarhus University, Langelandsgade 140, 8000 Aarhus, Denmark*
[b] *Center for Integrated Protein Science, Department of Chemistry, Ludwig-Maximilians-Universität (LMU), Butenandtstrasse 5-13, 81377 Munich, Germany*
[c] *Baseclick GmbH, Floriansbogen 2-4, 82061 Neuried Munich, Germany*
[d] *Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK*
[e] *Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, 43721 Suez, Egypt*
[f] *Metabion GmbH, Semmelweisstrasse 3, 82152 Planegg Munich, Germany. E-mail: a.manetto@metabion.com*

† Electronic supplementary information (ESI) available: Experimental procedures, supplementary figures, sequencing data, oligonucleotides sequences. See DOI: 10.1039/c8cc00738a
‡ Both authors contributed equally to this work.
§ Present address: Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695, USA.

Furthermore, the acidic reagents used for the de-tritylation step can lead to the formation of abasic sites and cleavage of the biopolymer, further decreasing the yield of full-length oligonucleotides.[2] Various approaches based on joining multiple short oligonucleotides have been developed to overcome these limitations with the goal of assembling long synthetic DNA strands (genes and gene fragments).[3–5] Currently two main strategies are used, both based on procedures that involve the use of enzymes. The first utilises DNA ligase whereas the second relies on the activity of DNA polymerases.[3,4] In the ligation method, which represents the earliest example of synthetic gene synthesis, the double-stranded DNA is assembled from complementary overlapping strands subsequently joined by the ligase to produce longer fragments, which requires 5′-phosphorylated oligonucleotides. This method becomes inefficient when large numbers of oligonucleotides are ligated.[3] In the second method, called DNA polymerase cycling assembly (PCA), which is based on the activity of DNA polymerase enzymes, the desired gene fragment is produced in a multiple step assembly.[4] Although these methods give access to a large variety of DNA fragments, there are some limitations due to mispriming, formation of secondary structures, and mistakes that occur when assembling repetitive sequences that hinder the polymerase activity. Therefore, in the rapidly evolving genomic, epigenomic and DNA nanotechnology fields the demand for functional gene fragments are not yet fully met, and alternative approaches are urgently needed.

To address this issue non-enzymatic ligation methods have been explored. Brown and collaborators reported a biocompatible chemical linkage that is read by polymerases both *in vitro* and *in vivo* in *E. coli* and human cells.[5–9] The authors previously proved the efficiency of copper-mediated azide/alkyne cycloaddition (CuAAC),[10,11] the classic click chemistry reaction,[12] as a ligation method for oligonucleotides. A major advantage of this technique is the high reaction efficiency in aqueous buffer, which makes CuAAC a suitable reaction for conjugation of biomolecules.[13]

Following this work, recently, Kukwikila *et al.* demonstrated enzyme-free, click-mediated gene assembly starting from 10 functionalized oligonucleotides that overlap to create a small

335 bp gene. The assembled gene is functional both *in vitro* and *in vivo*, confirming the biocompatibility of the triazole linkage.[14] The assembly approach used by Kukwikila *et al.* is based on splint oligonucleotides, but it is known that when increasing the number of strands to create a long gene, the complexity of the assembly increases significantly, often leading to failure in the synthesis of full-length sequences.[15] An alternative method to chemically ligate DNA strands is based on formation of phosphoramidate linkages. In this case, 3′-amino-modified oligonucleotides react with 5′-phosphorylated partner strands in templated reactions. This method has been recently used for gene synthesis[16] and also to circularize DNA nanostructures.[17] The field of DNA nanotechnology has produced a large number of sophisticated 2D and 3D DNA nanostructures that have been applied in many research fields. It has been proven that DNA origami is a robust assembly method where a long single-stranded DNA scaffold is folded with the help of short DNA strands called staples.[18] DNA origami has been widely used for organization of bio/nanomaterials at nanoscale precision, however, the use of DNA origami for gene assembly has not been reported. One of the great features of this technique is that the designed structure has the most stable conformation among all possible.[19] Therefore it provides control over position and stoichiometry of each strand involved in the assembly. In this manuscript we explore these features and adapt them to gene synthesis. For this purpose, we tested a derivative of the DNA origami technique, where the scaffold is fragmented into ∼60 nt long strands (gene oligonucleotides, GOs) and it folds with the help of staple strands.

In this manuscript we describe a one-pot click assembly procedure inspired by established self-assembly techniques from the DNA nanotechnology field, affirming that genes can be obtained by chemical ligation of several short DNA strands. The work paves the way to the synthesis of long DNA fragments and genes by combining the geometrical precision achieved with DNA nanostructures and the highly efficient click chemistry-mediated ligation. In this work we demonstrate the synthesis of a 762 bp gene encoding the enhanced green fluorescent protein (EGFP) from 14 functionalized oligonucleotides. The system employs a DNA nanostructure – a 6-helix bundle (6HB)[20] – as vehicle for assembling single stranded DNA bearing triazole linkages that are converted into double strands after PCR amplification. Using this technique, we assembled a 6HB where all GOs (3′ alkyne, 5′ azide-modified) are brought in close proximity, ordered in a predesigned fashion with an equimolar stoichiometry and ligated through click chemistry. The resulting product is then amplified by PCR to convert the triazole linkages in a canonical phosphodiester backbone (Fig. 1). The design was executed using the caDNAno software package[21] using the following principles: (1) the 762 nt long gene runs through the nanostructure forming the single-stranded scaffold of a 6HB of ∼40 nm in length. (2) The gene scaffold is fragmented into strands of ∼60 nt to assure reliable chemical synthesis of the double functionalized GOs. (3) The staples are designed to allow the structure to fold in a hierarchical order. The gene was divided into 14 GOs; 12 internal GOs bearing a 5′-terminal
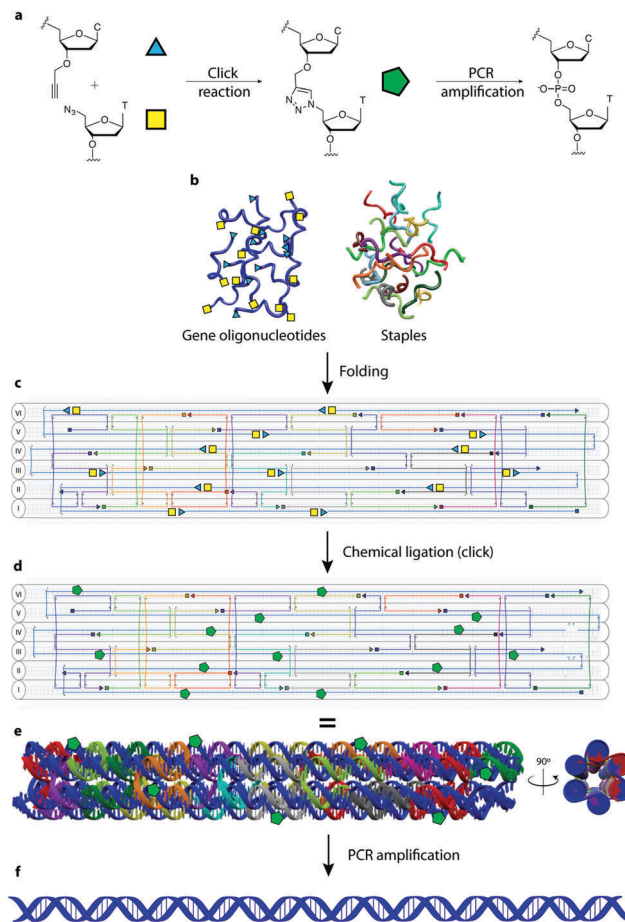

Fig. 1 Gene assembly process. (a) Chemical ligation mechanism. Coloured shapes represent the molecule on their left. (b) Gene oligonucleotides (blue, GOs) – bearing chemical modifications at the 3′ and 5′-ends – and staples (colours other than blue) are folded, forming the 6HB construct (c) caDNAno blueprint: helices are marked with roman numerals. (d) Click chemistry of the GOs in the 6HB forms a long linear scaffold (in blue). (e) 3D views of the designed DNA nanostructure with green pentagons to highlight click points (triazoles). (f) The ligated DNA strand in the nanostructure is used as a template for PCR amplification forming a product with a canonical phosphate backbone. In figure c and d, the tips of the small arrows represent the 3′-end of DNA and the small squares represent the 5′-end.

azide-modified thymidine and a 3′-terminal alkyne-modified deoxycytidine and two terminal GOs mono-functionalized as 5′-azide and 3′-alkyne respectively.

To test the folding of the 6HB nanostructure, unmodified GOs were initially used. We found that the 6HB presented here folds in presence of 20 mM MgCl$_2$ with formation of two species (Fig. 2a). The sample was analyzed by AFM and, as expected, the species were found to be monomers and dimers of the designed 6HB, with an average length of 43 ± 4.5 nm for the monomers and 82 ± 3.4 nm for the dimers (Fig. 2e). The length of monomers from AFM results agrees with theoretical calculation (42 nm). Dimer formation is probably due to stacking interactions between terminal base pairs of two different 6HB. The fact that only dimers, but not trimers or larger assemblies are formed, indicates that only one end of the 6HB tends to participate in base stacking. We speculated, and later confirmed, that the
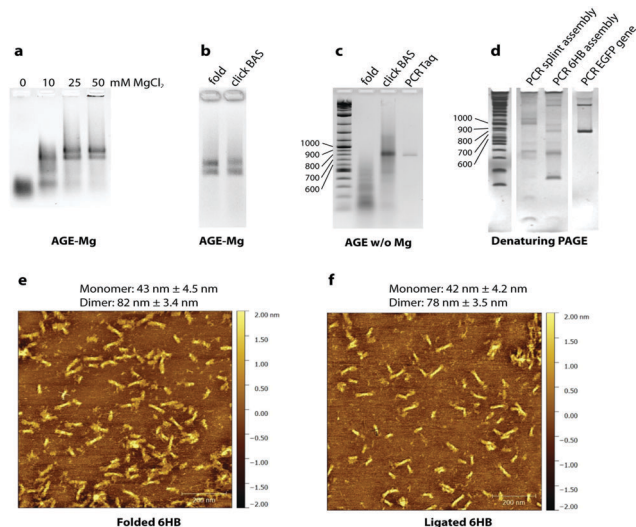
Fig. 2 Assembly experiments. (a) Salt screening test (mM MgCl$_2$): the structure formed at concentrations above 10 mM. Two species products (monomer and dimer) were found. (b) The product of the click reaction with the heterogeneous catalyst contained in the vial "reactor M" (BAS) runs like the folded sample. (c) The 6HB after click reaction is stable in absence of Mg ions. The PCR with Taq polymerase shows a product of the correct length. (d) Comparison between splint assembly without a nanostructure and assembly in the 6HB. The latter shows a product of the same length as the positive control (last lane, PCR on the EGFP gene). (e) AFM of the folded sample: monomers of ∼43 nm and dimers of ∼82 nm are formed. (f) AFM of the ligated sample: monomers of ∼42 nm and dimers of ∼78 nm are shown.

presence of the two species does not interfere with the gene assembly.

We then carried out chemical ligation of the EGFP gene, folding the 6HB using the 14 modified GOs and the complementary staples. Ligation through click chemistry – assisted by the close proximity of the GOs pre-organized in the nanostructure – was conducted using a protocol from Cassinelli et al.[22] The gel containing magnesium ions (AGE-Mg) in Fig. 2b shows that the structure retains its conformation after the click reaction. In contrast, in an AGE without magnesium ions (Fig. 2c), the structure prior to the click reaction unfolds, whereas the ligated structure ("click BAS" in Fig. 2c) entirely retains its conformation. Chemically-ligated 6HB constructs were examined by AFM and they were shown to have a similar length to the 6HB having an unligated scaffold (42 ± 4.2 nm for the monomer and 78 ± 3.5 nm for the dimer) (Fig. 2f). However, PAGE analyses in denaturing conditions of the EGFP gene assembly was not very informative: many bands formed after the click reaction without any particular predominance (Fig. S1, ESI†). We speculated that the band corresponding to the full-length gene might be the closest to the well of the denaturing gel, but since we observed that triazole-containing oligonucleotides run slower than their unmodified counterpart, we could not reliably refer to the ladder in assessing their size.

The crude chemical ligation mixture was then used as template for PCR amplification. Primers were designed in order to amplify only the full length gene. Amplification of the full-length

EGFP gene was successful when both Taq polymerase (low fidelity) and Baseclick polymerase (high fidelity) were employed (Fig. 2c). To assess the accuracy of the gene assembly method, PCR products were cloned and sequenced. In both cases, 5% of the screened clones resulted in 100% identity with the designed gene sequence (Tables S1–S3, ESI†). This is an encouraging result if we consider that one of the polymerase tested in the PCR step is Taq polymerase, classified as relatively low-fidelity due to its error-rate of 2.3×10^{-5} (vs. 9.5×10^{-7} of a high fidelity polymerase).[23] At this point we calculated the error rate of the polymerase in our gene assembly method to understand if the system is prone to mutations, or whether the triazole groups interfere with the correct incorporation of bases during PCR. An estimation of the error rate of the system was obtained by comparing our results to published data for the fidelity of Taq polymerase, which is reported to incorporate 1 error every 700–1700 bp depending on the source of the mutation data.[24] In our system Taq polymerase incorporated 1 error every 254 bp. This may indicate that the high concentration of metal ions present in the crude ligation mixture used as template for PCR may interfere with the activity of DNA polymerases, this can be addressed by exchanging the ion rich buffer with water. However, we cannot exclude that some of these mutations may be produced during the chemical synthesis of the starting oligonucleotides. Sequencing results show that mutations are homogeneously localized along the assembled gene fragment and that the triazole backbone is correctly replicated as previously shown.[5]

Finally, the method was compared to splint-assisted ligation in the absence of a nanoconstruct to prove the utility of the DNA nanostructure in assembling multiple gene oligonucleotides in equimolar ratio. The 14 GOs were assembled using 13 complementary splint oligonucleotides and chemically ligated with the same procedure used for the 6HB nanostructure. The ligation product was used as template for a PCR reaction where we employed KOD XL DNA polymerase, which is expected to efficiently read through the triazole linkage. Fig. 2d shows the PCR products of the splint-mediated assembly, the 6HB assembly and a positive control (PCR of the EGFP gene). PCR of the splint assembly did not produce full-length EGFP gene, but artifacts of higher and lower molecular weight, while PCR of the 6HB assembly showed a product of the same length as the control. However, the PCR products obtained using KOD XL polymerase were not as homogeneous as the ones employing Taq polymerase or Baseclick polymerase.

In conclusion, we have developed a system for gene fragment assembly by chemical ligation promoted by a DNA nanostructure, where gene fragments are part of the scaffold that runs inside the nanostructure. These are assembled in a predefined fashion, so that 3′-alkyne and 5′-azide are in close proximity, forming a 6HB nanostructure. The use of the nanostructure proved to be an efficient method to achieve an equimolar ratio of oligonucleotides, which is otherwise difficult when several oligonucleotides have to be ligated together. With this technique we were able to assemble 14 gene oligonucleotides to create a 762 nt long DNA strand, that after PCR is converted into a canonical double-stranded gene, encoding for EGFP. The method proved to be more efficient

than the equivalent ligation performed using splint oligonucleotides in the absence of the nanostructure. The chemical ligation method based on the CuAAC reaction is fast and efficient and can be carried out in a variety of biologically compatible buffers. Interestingly this gene is twice the size of the only one previously synthesized by CuAAC-mediated ligation.[15] This method should provide a general route to the synthesis of larger genes as well as long DNA strands for use in DNA nanotechnology and synthetic biology for the construction of complex nanostructures and synthetic organisms. We further envision the use of this technology to chemically assemble genes decorated with modifications such as epigenetic bases, fluorophores or haptens, which could have important applications in the fields of DNA nanotechnology and synthetic biology.

Conflicts of interest

The authors declare no conflicts of interest.

References

1 M. H. Caruthers, *et al.*, Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method, *Methods Enzymol.*, 1987, **154**, 287–313.
2 S. Kosuri and G. M. Church, Large-scale de novo DNA synthesis: technologies and applications, *Nat. Methods*, 2014, **11**, 499–507.
3 F. Chalmers and K. Curnow, Scaling Up the Ligase Chain Reaction-Based Approach to Gene Synthesis, *Biotechniques*, 2001, **30**, 249–252.
4 W. P. C. Stemmer, A. Crameri, D. H. Kim, T. M. Brennan and H. L. Heyneker, Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides, *Gene*, 1995, **164**, 49–53.
5 A. H. El-Sagheer, A. P. Sanzone, R. Gao, A. Tavassoli and T. Brown, Biocompatible artificial DNA linker that is read through by DNA polymerases and is functional in Escherichia coli, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 11338–11343.
6 C. N. Birts, *et al.*, Transcription of click-linked DNA in human cells, *Angew. Chem., Int. Ed.*, 2014, **53**, 2362–2365.
7 A. H. El-Sagheer and T. Brown, A triazole linkage that mimics the DNA phosphodiester group in living systems, *Q. Rev. Biophys.*, 2015, **48**, 429–436.
8 A. H. El-Sagheer and T. Brown, Efficient RNA synthesis by *in vitro* transcription of a triazole-modified DNA template, *Chem. Commun.*, 2011, **47**, 12057–12058.
9 A. H. El-Sagheer and T. Brown, Click nucleic acid ligation: Applications in biology and nanotechnology, *Acc. Chem. Res.*, 2012, **45**, 1258–1267.
10 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, A stepwise huisgen cycloaddition process: Copper(I)-catalyzed regio-selective 'ligation' of azides and terminal alkynes, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599.
11 C. W. Tornøe, C. Christensen and M. Meldal, Peptidotriazoles on solid phase: [1,2,3]-Triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides, *J. Org. Chem.*, 2002, **67**, 3057–3064.
12 H. C. Kolb, M. G. Finn and K. B. Sharpless, Click Chemistry: Diverse Chemical Function from a Few Good Reactions, *Angew. Chem., Int. Ed.*, 2001, **40**, 2004–2021.
13 J. Gierlich, G. a. Burley, P. M. E. Gramlich, D. M. Hammond and T. Carell, Click chemistry as a reliable method for the high-density functionalisation of alkyne-modified oligodeoxyribonucleotides, *Org. Lett.*, 2006, **8**, 3639–3642.
14 M. Kukwikila, N. Gale, A. H. El-Sagheer, T. Brown and A. Tavassoli, Assembly of a biocompatible triazole-linked gene by one-pot click-DNA ligation, *Nat. Chem.*, 2017, **9**, 1089–1098.
15 R. A. Hughes, A. E. Miklos and A. D. Ellington, Gene synthesis: Methods and applications, *Methods Enzymol.*, 2011, **498**, 277–309.
16 A. H. El-Sagheer and T. Brown, Single tube gene synthesis by phosphoramidate chemical ligation, *Chem. Commun.*, 2017, **53**, 10700–10702.
17 M. Kalinowski, *et al.*, Phosphoramidate Ligation of Oligonucleotides in Nanoscale Structures, *ChemBioChem*, 2016, **17**, 1150–1155.
18 P. W. K. Rothemund, Folding DNA to create nanoscale shapes and patterns, *Nature*, 2006, **440**, 297–302.
19 K. E. Dunn, *et al.*, Guiding the folding pathway of DNA origami, *Nature*, 2015, **525**, 82–86.
20 S. M. Douglas, *et al.*, Rapid prototyping of 3D DNA-origami shapes with caDNAno, *Nucleic Acids Res.*, 2009, **37**, 5001–5006.
21 S. M. Douglas, *et al.*, Self-assembly of DNA into nanoscale three-dimensional shapes, *Nature*, 2009, **459**, 414–418.
22 V. Cassinelli, *et al.*, One-Step Formation of 'Chain-Armor'-Stabilized DNA Nanostructures, *Angew. Chem., Int. Ed.*, 2015, **54**, 7795–7798.
23 Thermo Scientific. PCR fidelity calculator. https://www.thermofisher.com/dk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/pcr-fidelity-calculator.html.
24 P. McInerney, P. Adams and M. Z. Hadi, Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase, *Mol. Biol. Int.*, 2014, **2014**, 287430.

Supplementary Information

Gene Fragment Assembly via One-Pot Chemical Ligation of DNA Promoted by Nanostructures

Ilenia Manuguerra[#1], Stefano Croce[#2,5], Afaf H. El-Sagheer[3,4], Abhichart Krissanaprasit[1], Tom Brown[3], Kurt V. Gothelf[1] and Antonio Manetto[*6]

[1]Interdisciplinary Nanoscience Center (iNANO), Gustav Wieds Vej 14 and Department of Chemistry, Langelandsgade 140, Aarhus University, 8000 Aarhus (Denmark);

[2]Center for Integrated Protein science, Department of Chemistry, Ludwig-Maximilians-Universität (LMU), Butenandtstrasse 5-13, 81377 Munich (Germany);

[3]Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, (United Kingdom);

[4]Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, 43721 Suez (Egypt);

[5]baseclick GmbH, Floriansbogen 2-4, 82061 Neuried Munich (Germany);

[6]Metabion GmbH, Semmelweisstrasse 3, 82152 Planegg Munich (Germany).

[#]Both authors contributed equally to this work.

[*]E-mail: a.manetto@metabion.com

Experimental procedures

Design of the 6HB

The design of the 6HB was supported by caDNAno software. After introducing the sequence of the scaffold the staples were added automatically by the software and then manually modified to achieve a hierarchical folding.

In Supplementary Figure 3 a high-resolution blueprint with gene oligonucleotides and staples is shown.

Folding of the 6HB

Samples were made mixing staples and in-house synthetized GOs in ratio 1:1 to a final concentration of 500 nM/each oligo in 1X TE buffer with 20 mM $MgCl_2$. The sample was folded in a thermocycler using the following program: from 95 ºC to 80 ºC with a ramp of 1º C/min, from 80 ºC to 40 ºC with a ramp of 0.03ºC/min, from 40 ºC to 23 ºC with a ramp of 0.1 ºC/min and final hold at 8 ºC.

Click reactions on the 6HB

- **Click reaction (CuAAC) using an heterogeneous catalyst**

A volume of 15 µl of THPTA 0.1 M was added to the "reactor M", a vial containing the heterogeneous catalyst (baseclick), then 20 µl of folding reaction was added to the "reactor". The sample was incubated at 32ºC with gentle shaking (200 rpm) for 5h.

- **Click reaction (CuAAC) using $CuSO_4$**

The Baseclick EdU kit (reaction buffer, catalyst solution and reducing agent/buffer additive) was used for the experiments with $CuSO_4$ as source of Cu(I). The indications of the producer were used for the ligation assay using the click reaction. In this case, 40 µl of folding reaction were used in the assays.

- **Click reaction (CuAAC) using CuBr**

For the experiments with CuBr as source of Cu(I): 1 mg of CuBr was weighted under inert atmosphere and dissolved in 70 µl of "click solution". A total of 5 mg of the ligand TBTA were dissolved in 94 µl click solution and the two solutions were combined in ratio 1:2 (TBTA:CuBr). A volume of 20 µl of the mix were mixed with 4 µl of the folding reaction and incubated at 32ºC with gentle shaking.

AGE/AGE-Mg

AGE/AGE-Mg were prepared dissolving agarose (Ultra-pure, Thermo Scientific) to achieve a 1% gel in 0.5X TBE buffer (and 11 mM $MgCl_2$ final concentration for AGE-Mg, 1X TAE was used when the bands were extracted from gel). The gel was casted and left solidify at RT for 30 min.

PCR reactions

All the PCR reactions were tested using the following primers (Metabion):

FW EGFP 1-14   TATCACTATCGACGGTA

REV EGFP1-14   ACTTACAGCTTTACTTG

A volume of 1 µl of click reaction was used as template for PCR. The incubation with Taq Polymerase (NEB) and KOD XL (Millipore) proceeded as follows: 94 ºC for 3 min, 80ºC for 30 sec (add polymerase); 94 ºC for 45 sec, 30 ºC for 30 sec, 72 ºC for 12 min, repeat for 4 times; 94 ºC for 45 sec, 46 ºC for 30 sec, 72 ºC for 72 sec, repeat 10 times; 72 ºC for 10 min. Incubation with Baseclick polymerase proceeded as follows: 98 ºC for 90 sec; 98 ºC for 10 sec, 58 ºC for 20 sec, 72 ºC for 15 sec, repeat 20 times; 72 ºC for 8 min.

Sequencing

PCR reactions were cloned into a plasmid using the TOPO PCR cloning kit by Thermo Fisher Scientific and following manufacturer instructions. Plasmids from clones were extracted with Gene Elute HP Miniprep by Sigma-Aldrich and sent to sequencing to GATC Biotech.

AFM imaging

A volume of 2 µl sample was deposited on freshly cleaved mica and incubated for 1 min. Subsequently, 400 µl of imaging buffer (1× TAE-Mg2+ containing 5 mM NiCl2, 1× TAE-Mg-Ni) were added into the cell. Nanostructures were visualized by tapping mode AFM (Agilent AFM series 5100) using silicon nitride cantilevers (Olympus). All recorded AFM images were processed and analyzed by Gwyddion software.

Supplementary Figures

Denaturing PAGE   Denaturing PAGE
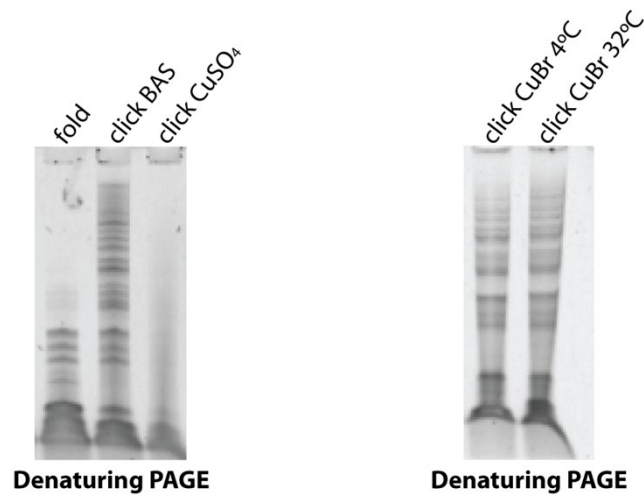
Supplementary Figure 1. Denaturing PAGE of the 6HB after folding and after ligation with different catalysts. Many bands are formed, but the complete EGFP gene is not detectable using PAGE, probably due to the low yield of the long sequence. To understand whether a different catalyst could give higher yields in the formation of the "ligated gene", CuAAC reaction using $CuSO_4$ as source of Cu(I) was tested. The sample appears like a light smear, probably because the reagent interferes with either the electrophoresis or the staining of DNA. CuBr was also tested as source of Cu(I), but the results are similar to the ones obtained with the heterogeneous catalyst.
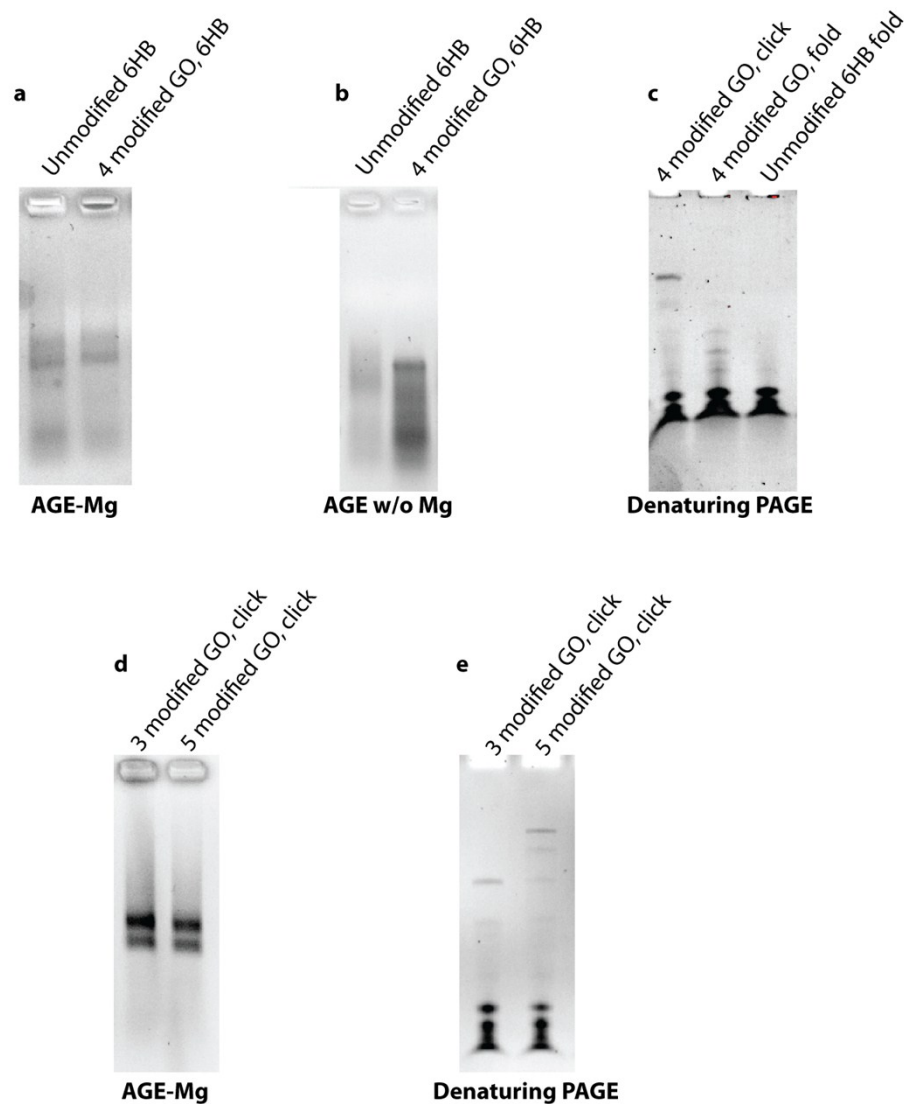
Supplementary Figure 2. Gene assembly in the 6HB of (a, b, c) 4, (d, e) 3 and 5 modified gene oligonucleotides. (a) Folding and click reaction on the 6HB using 4 modified gene oligonucleotides. The nanostructure forms using both a set of unmodified gene oligonucleotides or a set containing only 4 modified gene oligonucleotides. (b) Only the ligated structure partially retains its conformation in an AGE without Mg^{2+} ions. (c) Denaturing PAGE of the folded and ligated structures with 4 modified gene oligonucleotides or fully unmodified. Only after the click reaction is a band of high molecular weight corresponding to 4 ligated fragments visible (first lane). (d) Folding and click reaction on the 6HB using 3 and 5 modified gene oligonucleotides. (e) Denaturing PAGE of the ligation product in (d). Bands corresponding to 3 and 5 ligated gene oligonucleotides are visible on gel.
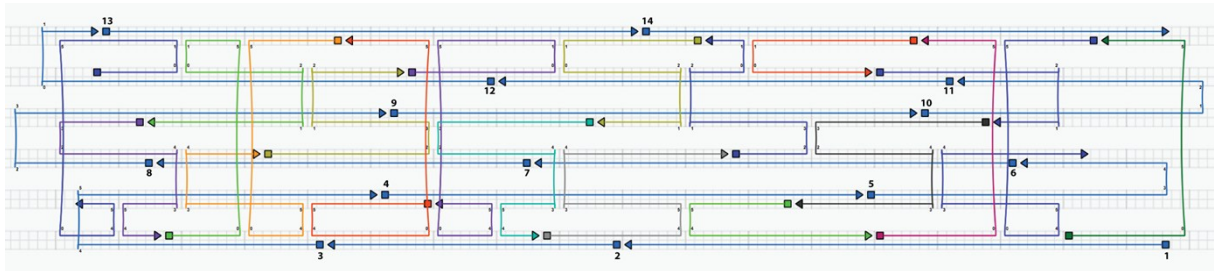
Supplementary Figure 3. Blueprint of the 6HB used in this manuscript. GOs are in blue, staples are represented in other colors. Arrow heads indicate the 3'end. Numbers on the 5'-end of the blue strands indicate the name of the GOs (i.e. 1 = EGFP1, 12 = EGFP12).

Supplementary Text

Sequencing result analysis of PCR products using Taq polymerase (medium-low fidelity)

Table S1: number of mutations (divided into substitutions, deletions and insertions) for each sequenced clone containing ~762 bp insert.

Clone Name	Substitutions	Deletions	Insertions
A02	2 (9 bp after click point)	1 (13 bp before click point)	1
B01	2 (12 bp, 9bp before click point)	1 (9 bp before click point)	1 (8 bp before click point)
D02	3 (16 bp before click point)	0	0
D03	**0**	**0**	**0**
H03	2 (13 bp before click point)	0	0
H04	1	0	0
H06	**0**	**0**	**0**

48 clones were sent for sequencing.

38 clones produced sequencing result.

33 clones had the PCR products as insert.

7 clones have a 762 bp insert.

2 clones (D03 and H06) have 100% identity with the designed sequence.

~5% of the clones that produced sequencing results showed 100% identity with the designed sequence.

Sequencing result analysis of PCR products using Baseclick polymerase (high fidelity)

Table S2: number of mutations (divided into substitutions, deletions and insertions) for each sequenced clone containing ~762 bp insert.

20 clones were sequenced

Clone name	Deletions	Substitutions	Insertions
EGFP_3_1	0	0	0
EGFP_3_2	3	1	0
EGFP_3_3	0	1	0
EGFP_3_4	0	2 + 2 c	1
EGFP_3_5	0	1	0
EGFP_3_6	0	1	0
EGFP_5_1	1	2	0
EGFP_5_2	2 c, 15 c (part of GO5)	2	0
EGFP_5_3	1	1	0
EGFP_5_4	60 c (whole GO10)+ 1	0	0
EGFP_5_5	0	2	0
EGFP_5_6	2 c	1	0
EGFP_5_7	3	0	0
EGFP_8_1	1 + 3 c	0	1
EGFP_8_2	1	0	0
EGFP_8_3	4	1	0
EGFP_8_4	0	1	0
EGFP_8_5	4	1	0
EGFP_8_6	2	2	0
EGFP_8_7	7	0	0
Tot (considering long deletions as 1)	37	21	2

(GO10: gene oligonucleotides EGFP10; 2 c: 2 consecutive mutations)

Table S3: Description of mutations close to click points

Clone name	Sequence surrounding CT click point (-: deletions)	Gene oligonucleotides involved
EGFP_5_2	C-CCTACG	3-4
	-CGCTACC	5-6
EGFP_5_4	GGAGTAC(-15 bp-)AACTTCAAGATCC	9-11 (10 is missing)
EGFP_5_5	CGACTACAACA	9-10
EGFP_5_6	C--CTGGAGTT	13-14
EGFP_3_2	CAACTAC-ACA	9-10
	TA-CTGAGCA	12-13

The click region of gene oligonucleotides 9-10 appears to be prone to mutations. We speculate that GO10 might not have enough complementary bases at the level of the staples (see Supplementary Figure 3). This might explain why gene oligonucleotides 10 (represented in red box, Fig Sx) is missing in one of the clones sequenced.

Oligonucleotides

All reagents for chemical synthesis were purchased from Sigma-Aldrich. Unmodified oligonucleotides were acquired from Metabion International AG, while 3'-alkyne and 5'-azide modified gene oligonucleotides were synthesized in-house using a previously reported procedure.[6]

Sequences

GENE OLIGONUCLEOTIDES (GOs)		
Name	Sequence (5' to 3' end)	Length (nt)
EGFP1	TCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGC	61
EGFP2	TGTTCACCGGGGTGGTGCCCATCCTGGTCGAGC	33
EGFP3	TGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC	62
EGFP4	TACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCC	54

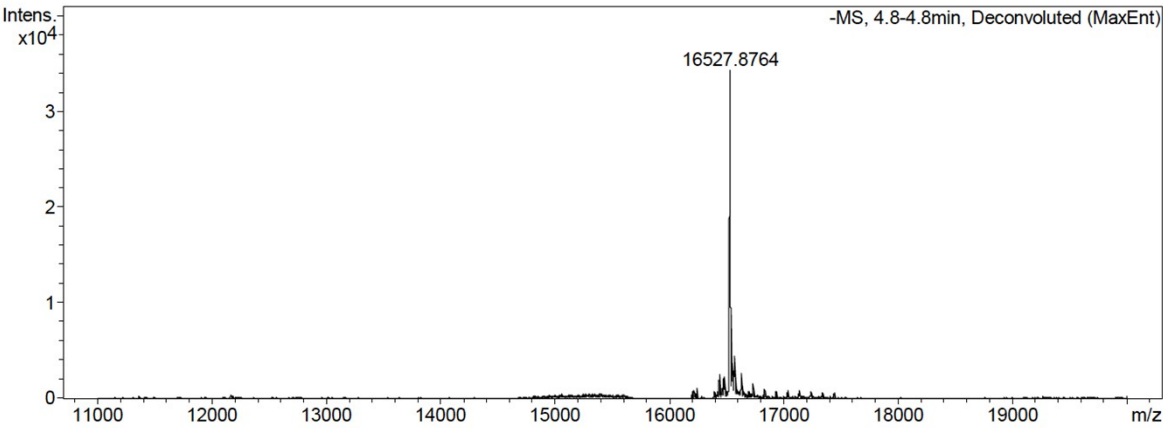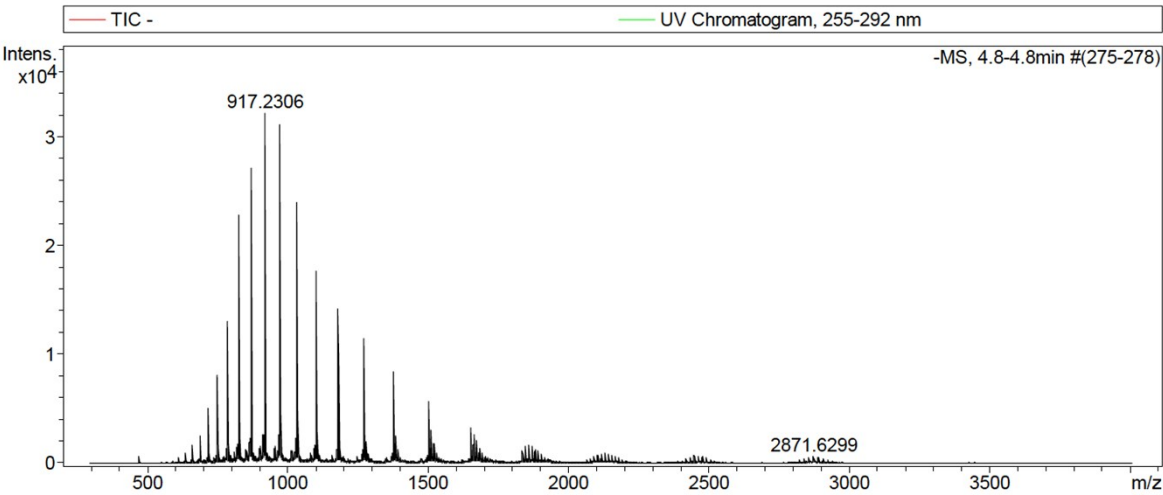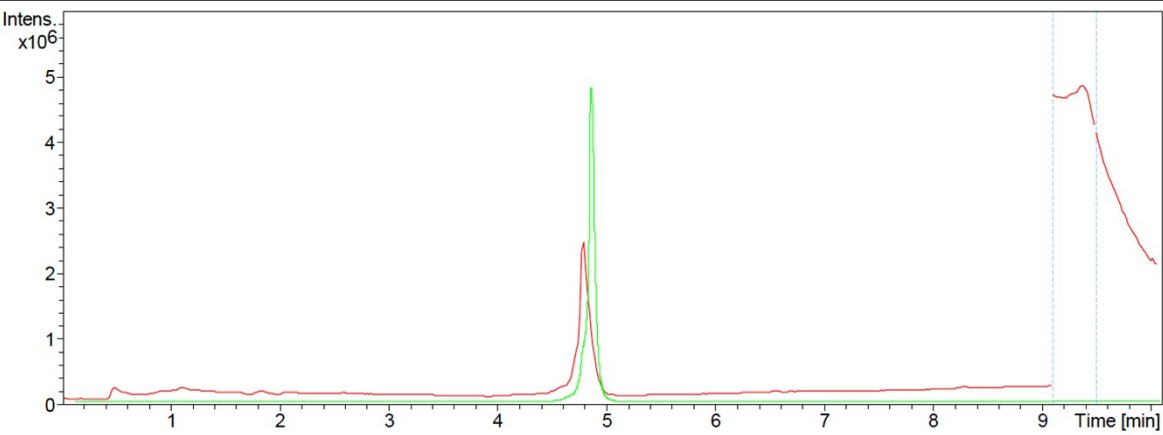| EGFP5	TGGCCCACCCTCGTGACCACCCTGACCTACGGTGTACAGT GCTTCAGCCGC	51
EGFP6	TACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCG CCATGCCCGAAGGC	54
EGFP7	TACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCA AC	42
EGFP8	TACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC CTGGTGAACCGCATCGAGC	58
EGFP9	TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG GCACAAGCTGGAGTACAAC	59
EGFP10	TACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA GAACGGCATCAAGGTGAAC	60
EGFP11	TTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGC TCGCCGACCAC	51
EGFP12	TACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGC TGCCCGACAACCACTACC	58
EGFP13	TGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAA GCGCGATCACATGGTCCTGC	60
EGFP14	TGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA CGAGCTGTACAAGTAAAGC	59
STAPLE STRANDS		
Name	Sequence (5' to 3' end)	Length (nt)
I6M0	GCCGTAGGTGGCATCAGCTCGACCAGGATCGTTGGGGT	38
I6M1	AAGAAGTCGCTTGTGCCCCAGGAGCCGTCCTCACG	35
I6M2	CTCCTTGAAGTCGATGCCTCCTGGACGTAGCCTTCAGGGC ACCC	44
I6M3	GGGCAGCAGCGGGTGCTCAGGTAGTTAACTTCGCTG	36
I6M4	GCAGCTTGCCGGGCCCTTGCTCACCATGGTGGC	33
I6M5	GATCTTGAAGTTCACCTTGATCGTTGTGG	29
I6M6	AACTCCAGCAGGACCAGCGAGCTGCACGCTTGTTGCCGTC	40
I6M7	CGGTGAACAGCTCCTCTGGTGCAGATGAACTTCGGGCATG GCGGACTTG	49
I6M8	CTTTGCTCAGGCGCCGTCCGCCCTCGCCCTCGGCCGTCGT C	41
I6M9	CTGTTGTAGTTGTACTCCAGTGCTGCTTCATGTGGTGGGCC AGGGCACGG	50
I6M10	GACCGGTGGATCCCTCCATGCCG	23
I6M11	TGTACAGCTCGGGGCCCGGGGTGGTCACGAGGGTCGGGG TAGCGGCTGA	49
I6M12	CGTTTACGTGCGGACTACGGGGCCGTCGCCGGTTCACCAG GGTGTCGC	48
I6M13	CCTCGAACTTCGGGTCTTGTAGTTCCGGACAGTGGC	36
I6M14	TCTGCTGGTAGTGGTCGTGTGATCGCGCTTCTGGGCACTC	45

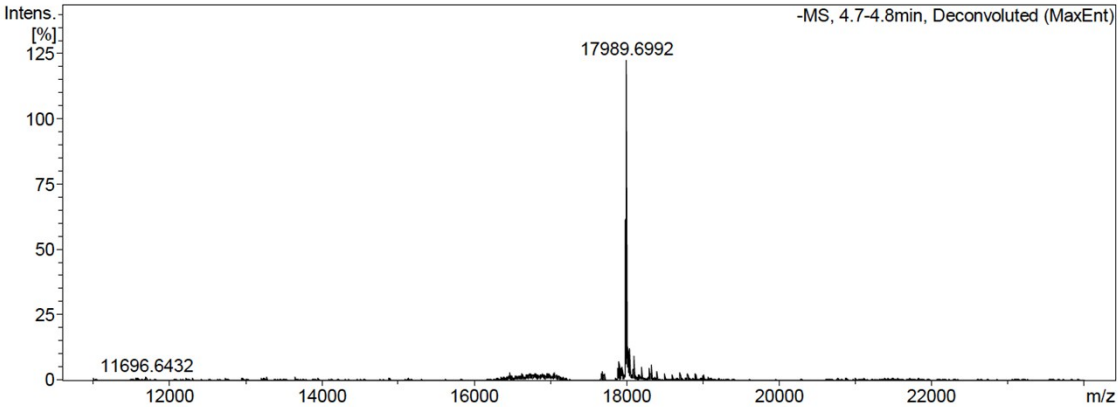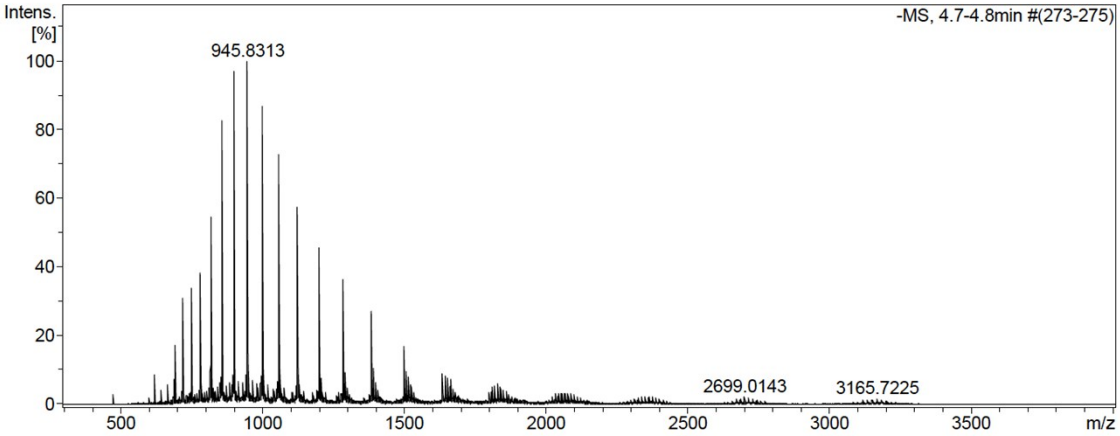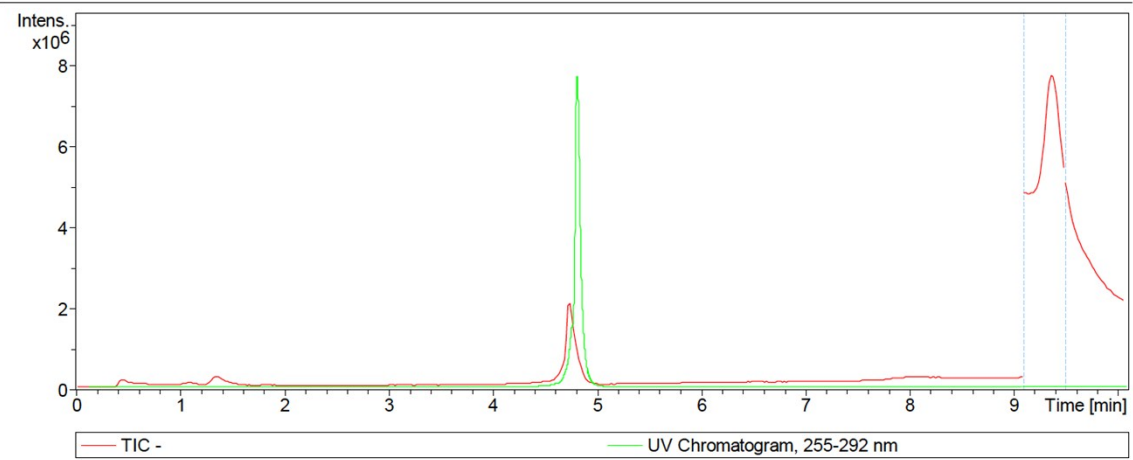|	AGCTT	
I6M15	AGAGTGATCCCGGCGGCGGTCGATGTTGTGGCG	33
I6M16	CGGTACCGTCGATTTTGCTTTACT	24
I6M17	CTTGAAGAAGATGGTGCGCCTTCAGCTCGATGCGATGGGG GTGT	44

EGFPn modified sequences were synthesized with 3'-alkyne and 5'-azide.

Typical LC/MS analyses are shown in the following chromatograms and deconvoluted mass spectra.
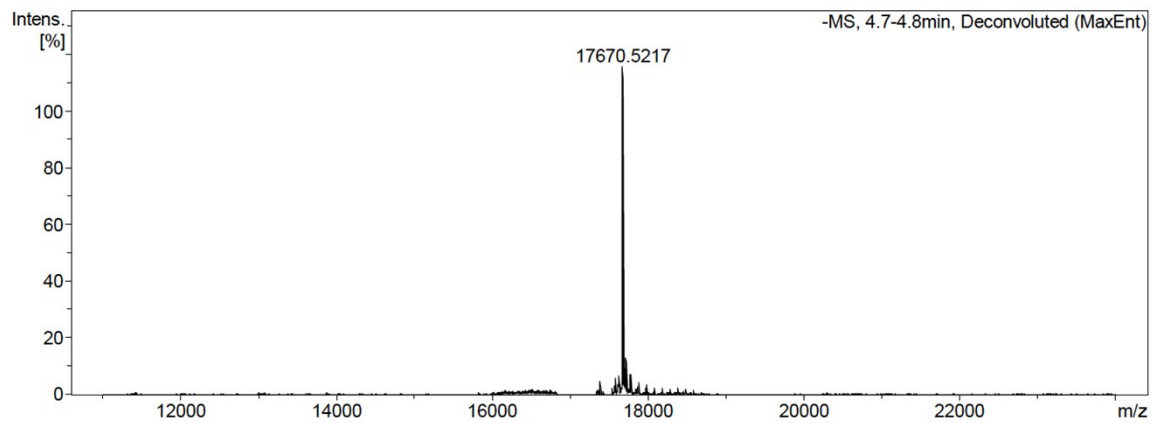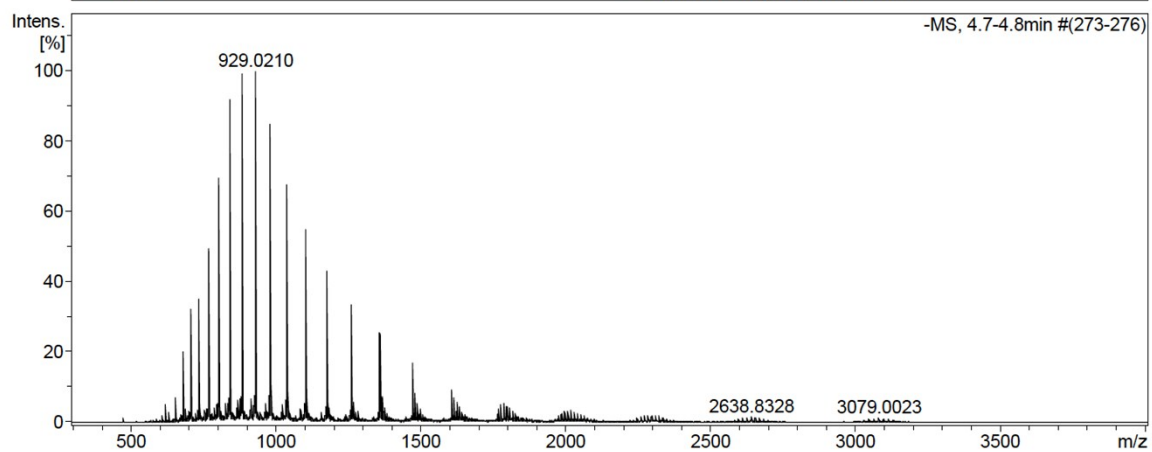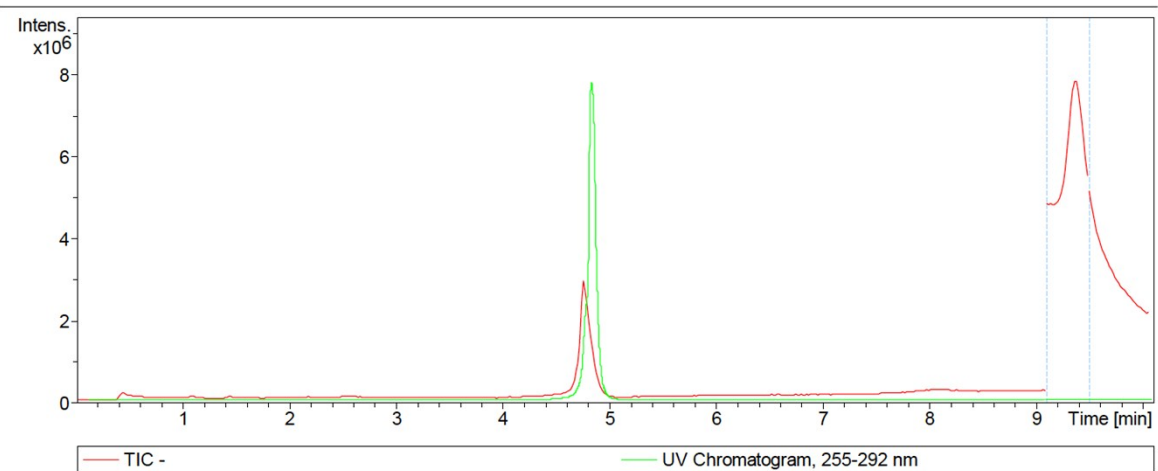
LC/MS EGFP6

LC/MS EGFP8

-MS, 4.7-4.8min #(273-275)

945.8313

2699.0143   3165.7225

-MS, 4.7-4.8min, Deconvoluted (MaxEnt)

17989.6992

11696.6432

LC/MS EGFP12

5. Cellular uptake of covalent and non-covalent DNA nanostructures with different size and geometry

5.1. Prologue

In recent years, DNA has been used as building block for the production of nanostructures with different biomedical applications, such as drug delivery and cellular biosensors, which has led to a rapid growth of the DNA nanotechnology area[6,18,71,72]. The need for spatial precision and organization at the nanometer level for biomedical applications is an important aspect to consider for the introduction of exogenous particles into living systems, and this can potentially be achieved by employing DNA nanostructures. Indeed, they can be easily and precisely modified with a variety of different molecules such as receptor ligands for targeting purposes[3]. To date, a plethora of DNA nanostructures based on Watson and Crick base pairing has been prepared and studied. Different methods have been developed for the production of the so-called DNA nanostructures, Figure 9. The technique of DNA origami was the first one to be developed and demonstrated in 2006, with the pioneering work of P. Rothemund[6]. The method is based on the usage of a very long ssDNA, usually from bacterial or viral origin, called "scaffold DNA", folded by the help of short ssDNA oligos, called "staples", designed to bring different parts of the scaffold together. Thanks to the robustness of the methods, it has widely and often been used as main technique for DNA nanotechnology preparation.
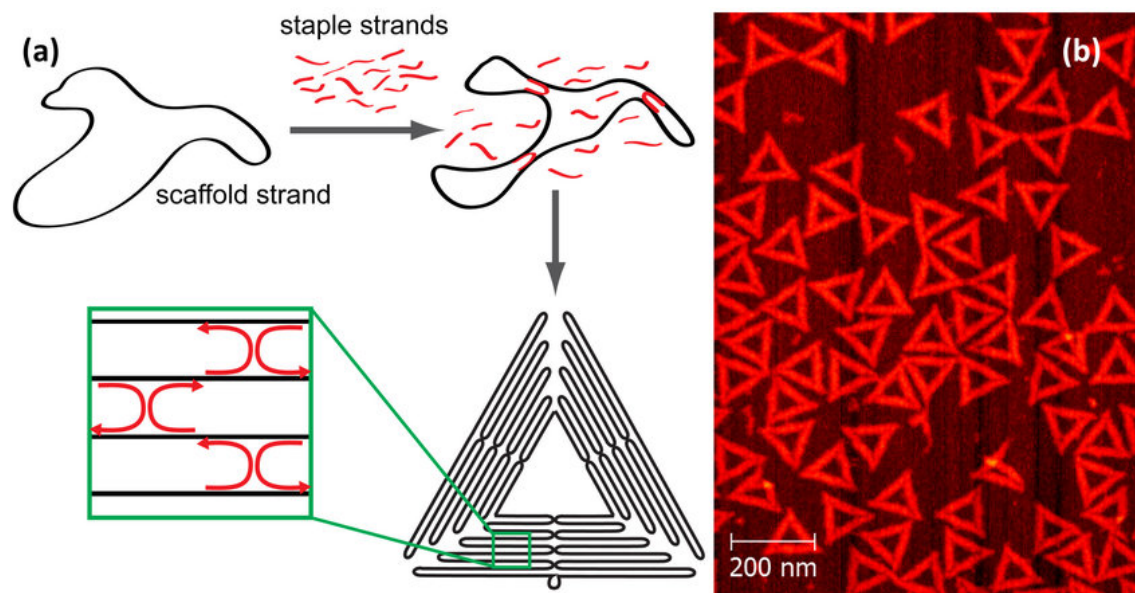
Figure 9: Left, schematic representation of the DNA origami technique, a long ssDNA (scaffold) is folded through the usage of short ssDNA (staples) in a precise geometrical structure. Right AFM imagine of the DNA nanostructure. I.Bald *et al.* Molecules, 2014, 19, 13803-13823.

Since then, other methods have been developed as well, such as the so called "tile assembly"[73] where shorter oligonucleotides, called "tiles" are used to self-assemble into bigger structures, where a short ssDNA region

hybridizes with another ssDNA region of another tile, thus avoiding the usage of the very long ssDNA, Figure 10.
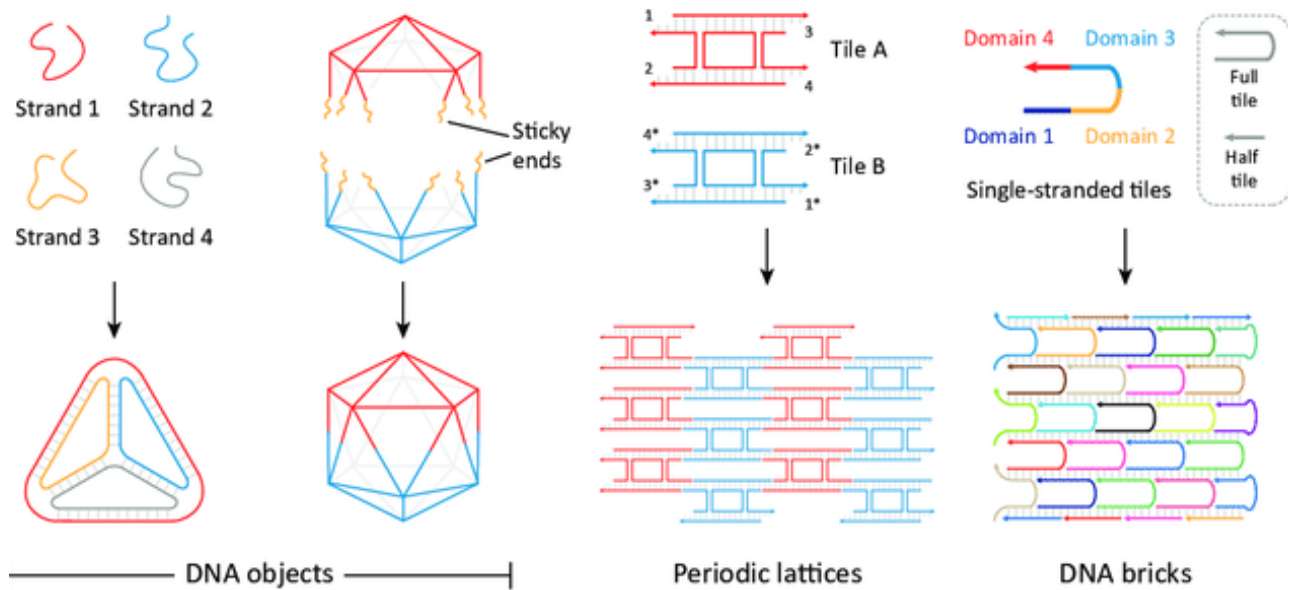
Figure 10: Different tile assembly method, based on short ssDNA, for the production of DNA nanostructures. M.B.Ra *et al*., Trends in Biochemical Sciences, 2018, 17,1512

To be considered a viable tool to build biomedically relevant carriers, the material should maintain long term stability in various physiological environments as well as represent a suitable shuttle to the target and/or subcellular destination, thus DNA nanostructures have already been used as delivery vehicle for different molecules, including anticancer drugs.

Furthermore, signal transduction is an important aspect where DNA nanostructures may have potential advantages, due to their intrinsic bio-chemical property to "communicate" with biological systems and therefore be an excellent connection for exchange of information. Driven by these findings, we thought to employ different DNA nanostructures and evaluate their response in cellular uptake experiments.

5.2. Abstract of the publication "Cellular uptake of covalent and non-covalent DNA nanostructures with different size and geometry"

DNA nanostructures with different size and shape, assembled through either covalent or non covalent bonds, namely tetrahedral and octahedral nanocages, rod-shaped chainmails, square box and rectangular DNA origami structures, were compared for their stability in serum, cell surface binding, internalization efficiency, and intracellular degradation rate. For cell internalization a specific cell system, highly expressing the scavenger receptor LOX-1, was used. The results indicate that LOX-1 binds and internalizes a broad family of DNA structures of different size that, however, have a different fate and lifetime inside the cells. Covalently-linked tetrahedra, octahedra or chainmails are intact inside cells for up to 18 hours whilst the same DNA nanostructures without covalent bonds along with square box and rectangular origami are rapidly degraded. These data suggest that noncovalently linked structures may be useful for a fast drug release whilst the covalently-linked structures can be appropriate vehicles for slow releasing molecules.

5.3. Author contribution

The author carried out the labeling and characterization of the oligo containing biotin molecule. The synthesis of the DNA nanostructure, the so called chainmail, it´s chemical ligation *via* click chemistry and helped in the revision of the manuscript.

5.4. Associated Publication

Cellular uptake of covalent and non-covalent DNA nanostructures with different size and geometry

Sofia Raniolo, **Stefano Croce**, Rasmus P. Thompsen, Anders H. Okholm, Valeria Unida, Federico Iacovelli,

Antonio Manetto, Jørgen Kjems, Alessandro Desideri, Silvia Biocca, *Nanoscale*, **2019**, 11, 10808-10818.

Check for updates

Cellular uptake of covalent and non-covalent DNA nanostructures with different sizes and geometries†

Sofia Raniolo, [ID] [a,d] Stefano Croce,[b] Rasmus P. Thomsen,[c] Anders H. Okholm,‡[c] Valeria Unida,[d] Federico Iacovelli, [ID] [d] Antonio Manetto, [ID] [e] Jørgen Kjems,[c] Alessandro Desideri [ID] [d] and Silvia Biocca [ID] *[a]

DNA nanostructures with different sizes and shapes, assembled through either covalent or non-covalent bonds, namely tetrahedral and octahedral nanocages, rod-shaped chainmails, square box and rectangular DNA origami structures, were compared for their stability in serum, cell surface binding, internalization efficiency, and intracellular degradation rate. For cell internalization a specific cell system, highly expressing the scavenger receptor LOX-1 was used. The results indicate that LOX-1 binds and internalizes a broad family of DNA structures of different sizes that, however, have a different fate and lifetime inside the cells. Covalently linked tetrahedra, octahedra or chainmails are intact inside cells for up to 18 hours whilst the same DNA nanostructures without covalent bonds along with square box and rectangular origami are rapidly degraded. These data suggest that non-covalently linked structures may be useful for fast drug release whilst the covalently-linked structures could be appropriate vehicles for slow release of molecules.

Introduction

In recent years, the nanotechnology field has grown quickly and an impressive variety of materials have been used to build nanoparticles of different sizes and shapes. In particular, DNA has been extensively investigated for designing programmable, self-assembling nanostructures for various biomedical applications, such as drug delivery, cellular biosensors and *in vivo* imaging, due to its intrinsic biocompatible, nontoxic and stable properties.[1,2] Crucial for these applications is the ability of DNA-based nanostructures (DNS) to enter mammalian cells. Studies addressing different topological configurations and functionalization of DNS that may improve selective targeting and interaction between DNS and cell membranes may offer insights for their optimization.

[a]Department of Systems Medicine, University of Rome Tor Vergata, Via Montpellier 1, 00133 Roma, Italy. E-mail: biocca@med.uniroma2.it
[b]Baseclick GmbH, Floriansbogen 2-4, 82061 Neuried, Germany
[c]Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Aarhus, Denmark
[d]Department of Biology, University of Rome 'Tor Vergata', Via della Ricerca Scientifica 1, 00133 Roma, Italy
[e]Metabion, Gesellschaft für angewandte Biotechnologie mbH, 82152 Planegg, Germany
†Electronic supplementary information (ESI) available. See DOI: 10.1039/c9nr02006c
‡Present address: Arla Innovation Centre, Arla Foods, Agro Food Park 19, 8200, Aarhus N, Denmark.

For entering into cells, short single- or double-stranded DNA pieces need cationic transfection reagents that help to overcome the electrostatic repulsive forces between their phosphate backbone and the cell membrane.[3] Notably, DNA assembled into three-dimensional structures, such as nanocages, chainmails and origami, can enter a wide variety of cell types without the aid of transfection agents.[4–10] Selective targeting can be achieved through the functionalization of DNS with specific ligands such as folate, transferrin or aptamers, leading to recognition of specific receptors and improved entry into specific cell types.[11–17] While the intracellular fate and localization of DNS is strongly dependent on the receptor mediating their uptake,[18,19] the importance of DNS size and shape in cellular uptake has been taken into consideration just recently.[20,21]

The main responsible targets for the internalization of DNA-based nanostructures into cells are scavenger receptors, having a broad ligand-binding specificity.[20,22–24] These receptors constitute a large family of cell-surface receptors that typically bind multiple ligands, including a wide variety of polyanionic molecules.[25] Oligonucleotide-decorated gold nanoparticles, wrapped by a dense monolayer of DNA on their surface, have been shown to interact with scavenger receptors of class A (SR-A) and class B (SR-B)[22] and, more recently, we demonstrated that pristine octahedral DNA nanocages specifically interact with LOX-1, a scavenger receptor of class E (SR-E1).[24] Pristine octahedral nanocages are very efficiently internalized in vesicular structures, which traffic through lyso-

somes, by a LOX-1-mediated uptake. The amount of internalized cages in LOX-1-expressing cells is 30 times higher than in cells not expressing this receptor.[24]

LOX-1 is the major receptor for oxidized low-density lipoprotein (ox-LDL) in endothelial cells.[26] It binds its natural substrate ox-LDL through electrostatic interactions due to the presence of a basic spine formed by positively charged amino acids.[27–30] LOX-1 is a multivalent scavenger receptor responsible for the internalization, beyond ox-LDL, of a variety of different macromolecules, including activated platelets, apoptotic bodies, bacteria, and advanced glycation end products (AGEs).[31] Its expression is up-regulated in many pathological conditions affecting the cardiovascular system and it has been associated with atherosclerosis, obesity and inflammation.[32] Recent studies have demonstrated that LOX-1 is up-regulated in many cancer types including bladder, cervix, mammary gland, lung, and colorectal, suggesting its involvement in tumor development.[33–35] During the past years, our group, with a combined simulative and experimental approach, has characterized the molecular mechanism of LOX-1 binding activity[36,37] and set up a model cell system based on the ectopic expression of the human isoform of the receptor.[38] This system allows a very high expression level of LOX-1 receptors on the cell membrane, mimicking the *in vivo* characteristics of up-regulation typical of pathological states.[35] Taking advantage of this cell system, we have previously characterized the LOX-1-mediated cellular uptake of octahedral nanocages.[24] Here we use the same cellular system for investigating either the influence of size and shape along with the presence of covalent or non-covalent bonds on the DNS receptor-mediated cellular uptake. We assembled tetrahedral and octahedral nanocages, rod-shaped chainmails, square box and rectangular DNA origami structures and compared them for their stability in serum, cell membrane binding, internalization efficiency and intracellular degradation rate.

Materials and methods

Preparation of DNS

Tetrahedral and octahedral nanocages. All oligonucleotides used for the assembly of non-covalent and covalent biotinylated-DNA tetrahedral (TD) and octahedral (OD) nanocages were HPLC purified and purchased from Sigma-Aldrich. Sequences of oligonucleotides are reported in the ESI S1, Table S1† and in Vindigni *et al.*,[24] respectively. DNA nanocages were prepared as previously described.[24]

Chainmails. All azide/alkyne-oligonucleotides used for the assembly of non-covalent and covalent biotinylated chainmails (CM) were HPLC purified and purchased from Metabion, Germany. Sequences of the oligonucleotides are reported in the ESI S2, Table S2.† DNA chainmails were folded by cooling a buffer solution (1 × TE, 20 mM MgCl$_2$) containing azide/alkyne-oligonucleotides from 80 °C to 65 °C in steps of 1 °C min^{-1} followed by a slower ramp from 65 °C to 25 °C in steps of 0.5 °C every 12 min, as previously described.[39] Namely, five

bct-ODNs (body functional click tiles) and three hct-ODNs (head functional click tile) were mixed with 16 ct-ODNs (click tile). Thus, through a single one-pot CuAAC three different kinds of click reactions are simultaneously accomplished. Specifically, 24 strands undergo catenation, five positions are labeled with an external azide and three "head" alkyne residues are conjugated with a strand carrying an additional functionality. A graphical representation of the components of the chainmail participating in the folding procedure and the subsequent one-pot click reaction on the structure is shown in the ESI S3, Fig. S1.†

Square box (SBO) and rectangular (RO) origami. DNA origami structures were assembled in one-pot reactions containing single-stranded M13 DNA, unmodified and biotin-modified staple strands in TAEM buffer (Tris-acetate pH 8 40 mM, EDTA 1 mM and Mg(OAc)$_2$ 12.5 mM) using an overnight thermal annealing ramp from 90 °C to 10 °C. Excess staple strands were removed by 100k Amicon filters as described by the manufacturer (Merck Millipore).[40]

TEM studies

The structural integrity of DNA nanostructures was verified using negative-stain transmission electron microscopy (TEM). The TEM samples were prepared by dropping 5 μL of 0.1–0.3 μg μl^{-1} DNS solution in TAE (TD, OD and CM) or TAEM (RO and SBO) onto a glow-discharged thin carbon-coated CF400-CU-UL grids (EMS, Hatfield, PA, USA) for 1.5 min. The grids were quickly blotted and immediately stained using a negative stain procedure of 2 consecutive rounds of addition of 2% uranyl formate solution for 5 and 20 s, respectively (5 s + 10 s for the rectangular origami), each followed by a blotting step. After the second stain-blotting step, samples were left to air dry prior to imaging.

Imaging was performed with a Tecnai G2 Spirit electron microscope (FEI Company) operated at 120 kV using a Lab6 electron gun. Image acquisition was done using a bottom-mounted TVIPS CMOS 4k camera (TEM-cam-F416) at magnifications of ×42.000–×52.000.

DNS biotinylation

Biotinylated TD, OD and CM were obtained using oligonucleotides modified with biotin through copper(ı)-catalysed azide/alkyne cycloaddition (CuAAC, click reaction), as described in the ESI, S3.† Biotinylated RO and SBO were assembled using biotin-modified staple strands, generated reacting the terminal transferase enzyme with biotinylated dUTPs.[41]

Cell cultures and transfection

COS-7 cells, a monkey kidney fibroblast-like cell line, were grown in DMEM (Dulbecco's Modified Eagle's Medium; Biowest) supplemented with 10% fetal bovine serum (FBS, Gibco), ʟ-glutamine 1 mM (Sigma Aldrich), sodium pyruvate 1 mM (Biowest) and 100 U ml^{-1} penicillin–streptomycin (Euroclone). For the expression of LOX-1 receptors, human LOX-1 was subcloned into pEF/V5-His vectors (Invitrogen) and used for transiently transfecting COS cells, as previously described.[38]

Stability of DNS

Biotinylated DNS were incubated in 10% FBS at 37 °C for different times. To prevent the denaturation of origami structures, RO and SBO stability experiments were performed by supplementing 10% FBS with 6.25 mM $MgCl_2$. After incubation, each DNS sample was treated with proteinase K (100 μg mL^{-1}) for 1 h at 37 °C and protein digestion was stopped by adding PMSF to a final concentration of 5 mM. Samples were mixed with loading buffer (Tris-Cl 500 mM pH 6.8; glycerol 20%; SDS 4%; Bromophenol Blue 0.02%), run on 1% agarose gel in TBE (Tris-Cl 89 mM pH 8, boric acid 89 mM, EDTA 2 mM), supplemented with 6.25 mM $MgCl_2$ for non-covalent DNS, and blotted.

Purification of DNS from cell lysates and blotting

Cells were plated in 48 wells per plate at a density of 3×10^4 cells per well, transiently transfected with LOX-1-V5 and, after 24 hours, incubated with 6 μg mL^{-1} of biotinylated DNS for different times at 37 °C. DNSs were purified from cell lysates and treated with proteinase K (100 μg mL^{-1}) for 1 h at 37 °C, as previously described.[24] Samples were run on 1% agarose gel in TBE, transferred to a positively charged nylon membrane (Zeta Probe, BioRad) for 30 min at 23 V (SemiDry Transfer Cell, BioRad) and treated as previously described.[24] Biotin detection was carried out using streptavidin–HRP (Abcam) and visualized by enhanced chemiluminescence (ECL EXTEND, Euroclone). For image processing and densitometric analysis, photographic films were digitized by scanning. Bands were analyzed using ImageJ software. Values were calculated using GraphPad Prism and expressed as a mean ± SEM.

Confocal analysis

COS cells were seeded on poly-L-lysine-coated glass coverslides, transfected with LOX-1-V5 and incubated with biotinylated DNS. Binding experiments at 4 °C were performed as previously described.[42] Uptake experiments were performed by incubating cells with DNS in DMEM 10% FBS for different times at 37 °C. Cells were then washed in Phosphate-Buffered Saline (PBS), fixed with 4% paraformaldehyde and permeabilized with Tris-HCl 0.1 M pH 7.6/Triton 0,1% for 4 min. Co-localization experiments were performed as previously described.[24] Biotinylated DNS were detected with streptavidin-FITC (Jackson ImmunoResearch); the LOX-1 receptor was visualized using Mab anti-V5 (Invitrogen) and lysosomes were detected using rabbit monoclonal anti-LAMP-1 antibody (Cell Signaling Technology). Rhodamine Red-X-conjugated AffiniPure donkey anti-mouse IgG and Red-X-conjugated AffiniPure donkey anti-rabbit IgG (Jackson ImmunoResearch) were used as secondary antibodies. Nuclei were stained with DAPI (Invitrogen). At the end of incubation, coverslips were washed in PBS and mounted with ProLong Gold Antifade Mountant (Thermofisher). Images were obtained with a laser confocal fluorescence microscope Olympus FV1000 at 60× magnification.

Results

Design, assembly, and characterization of DNS

Different covalently linked DNA structures (DNS) and non-covalent structures (ncDNS) were designed and assembled (Fig. 1A). Tetrahedra (TD) and octahedra (OD) are composed of four and eight oligonucleotides, respectively (ESI, S1† and Vindigni et al.[24]). For covalent TD and OD, DNA ligase is added after their assembly, in order to link the strands through a 5′–3′ covalent bond. Rod-shaped chainmail structures (CM) are assembled using 5′-azide, 3′-alkyne-modified oligonucleotides reported in the ESI, S2.† For covalent CM, 5′-azide, 3′-alkyne-modified oligonucleotides are covalently connected by chemical ligation to form rings and interlocked DNA single strands (ESI, S3†). Square box (SBO) and rectangu-
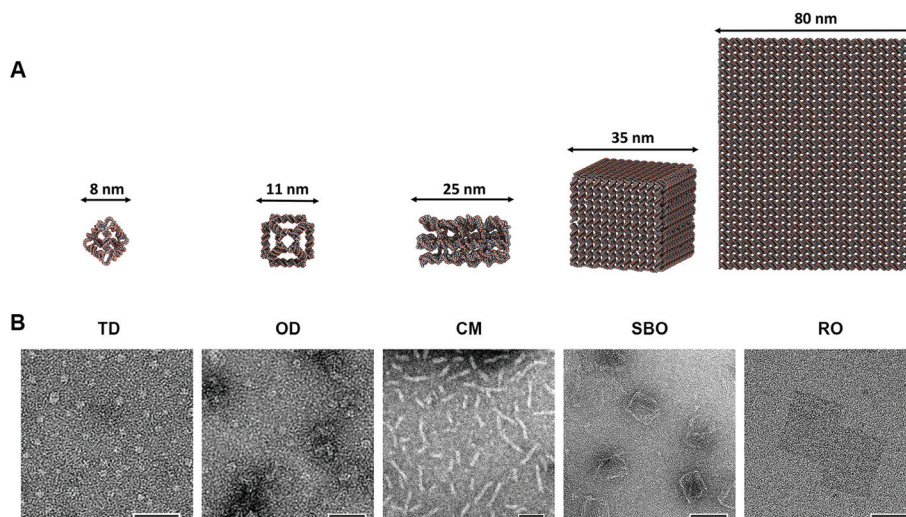

Fig. 1 (A) Computer models of DNS used in this study. From left to right: Tetrahedral (TD) and octahedral (OD) nanocages, rod-shaped chainmail (CM), square box (SBO) and rectangular (RO) origami. (B) TEM images of each of the DNS used in the study. Scale bars: 50 nm.

lar (RO) origami are non-covalently linked structures, assembled from a 7249-nucleotide-long single-stranded M13 DNA (the scaffold strand), folded into a defined shape by hundreds of shorter staple strands using thermal annealing in a one-pot reaction. The structural integrity of each DNS shape was verified by negative-stain transmission electron microscopy (TEM) (Fig. 1B). All constructed DNS were free from aggregation and were predominantly found in the monomeric state.

Biotin groups were added to the DNS. In detail, TD and OD have one biotin molecule bound per cage, CM contains 5 and RO and SBO structures have 32 biotin molecules for each structure.

Stability of DNS in fetal bovine serum

As an important prerequisite for the use of DNS in biomedical applications, we investigated the *in vitro* stability in biological fluids at the physiological temperature of 37 °C and compared their resistance to nuclease digestion. After incubation with 10% fetal bovine serum (FBS) for different time intervals, non-covalent and covalent DNSs were run in 1% agarose gel and blotted to examine their structural integrity (Fig. 2 and ESI S4†). The presence of biotin molecules inserted in the structures allows the detection of DNS through a streptavidin–HRP–biotin reaction in DNA blots.

As shown in Fig. 2, non-covalent ncTD, ncOD and ncCM are stable for at least 4 h in 10% FBS at 37 °C and then the ncDNS are degraded as a function of time (Fig. 2A, lanes 1–5; 2B, lanes 1–4 and 2C, lanes 1–4). The relative intensity of each band, quantified by densitometric analysis using ImageJ software, was normalized to the intensity of the band corresponding to the input and reported in the graph. The half-lives of ncTD, ncOD and ncCM in FBS are 25 ± 2.4, 15 ± 2.7 and 17 ± 1.6 h, respectively, calculated by the relative intensity of each band visualized through a DNA blot. Instead, covalent TD, OD and CM are stable up to 24 h in 10% FBS at 37 °C before gradual degradation (Fig. 2A, lanes 6–11; 2B, lanes 6–10; 2C, lanes 5–10), indicating higher resistance than non-covalent DNS. In particular, the presence of covalent bonds in TD, OD and CM extends the half-life in 10% FBS to 48 ± 2.5, 27 ± 1.9 and 66 ± 2.1 h, respectively.

Regarding origami structures, the non-covalently-linked ROs are fully stable in 10% FBS at 37 °C for 4 h (Fig. 2D, lanes 1–3). Between 4 and 24 h, a lower molecular band of about 1 kb is visible, suggesting a partial disassembly (Fig. 2D, lanes 4). At 48 h, no band with the same mobility as the input is detectable, while the intensity of the lower-molecular-weight band increases, indicating that the original RO structures are almost fully degraded (Fig. 2D, lane 5). It is worth noting that the intensity of the 2 kb origami band increases when the lower molecular weight band appears (Fig. 2D, lane 4). This may be explained by an increased exposure of the biotin groups upon partial unfolding of the origami structure. SBO structures show similar behavior in 10% FBS at 37 °C (ESI, S4†).

Receptor-mediated entry of DNS into LOX-1-expressing cells

Binding and uptake of the DNS in mammalian cells were monitored in COS fibroblasts transiently transfected with a plasmid encoding for the human full-length LOX-1 receptor, containing a V5 tag at the C terminus (LOX-1-V5), as previously described.[24] This cell system allows a very high expression level of functional LOX-1 receptors and a high transfection efficiency (about 40% of total cell population). Since not-transfected COS cells do not express endogenous LOX-1, the study of the specific interaction between LOX-1 receptors and DNS can be studied within the same cell population.

Twenty-four hours after transfection with the LOX-1 receptor gene, COS cells were incubated with the biotin-labeled DNS for 1 h at 4 °C or different times at 37 °C and stained with anti-V5 antibodies and streptavidin–FITC, for visualizing LOX-1 receptors and biotinylated DNS, respectively. DAPI was used for localizing the nuclei. DNS binding to the cells at 4 °C was visualized by confocal analysis as intense dot-like green fluorescence in association with the outer surface of the plasma membrane (Fig. 3A, B, and 4, panels a and b). The finding that an almost identical dot-like fluorescence is obtained when cells are stained with Mab anti-V5 IgG, which selectively visualizes transfected LOX-1 receptors (ESI, S5†), is consistent with an interaction between DNS and LOX-1 receptors.

As already demonstrated for octahedral DNA nanocages,[19,24] DNS and LOX-1 receptors cannot be visualized simultaneously in a binding experiment at 4 °C, since DNS and anti-V5 antibodies compete for the same binding site on LOX-1 receptors, when situated on the cell surface. However, the co-detection of intracellular LOX-1 receptors and internalized DNS in the same cell is possible since newly synthetized LOX-1 is present in an intracellular pool in the endoplasmic reticulum and Golgi apparatus.[43] To enable their co-detection, DNS were incubated with cells for different times at 37 °C and confocal microscopy used for visualizing internalized DNS and intracellular LOX-1 receptors. Fig. 3 shows representative images in which LOX-1-expressing cells (red fluorescence, Fig. 3A and B, panels c) exhibit internalized TD (green fluorescence, Fig. 3A, panels d and e) and CM (Fig. 3B, panels d and e) after 1 h incubation at 37 °C. Notably, both TD and CM, notwithstanding the different sizes and shapes, are efficiently internalized in LOX-1-expressing cells. TD and CM nanostructures appear as fluorescent dots in the cytoplasm, larger than those seen in the binding experiments (compare panels a and d of Fig. 3A and B), confirming that TD and CM are distributed in endocytic vesicles that tend to fuse together to form larger endosomes, as shown for octahedral cages in Vindigni *et al.*[24]

As shown for TD and CM, also RO nanostructures bind to LOX-1 receptors at the cell surface and are distributed in small green fluorescent dots (Fig. 4, panels a and b). For the co-detection of RO and LOX-1 in the same cells we incubated for different time intervals at 37 °C RO structures with LOX-1 expressing cells and analyzed cells with confocal microscopy.
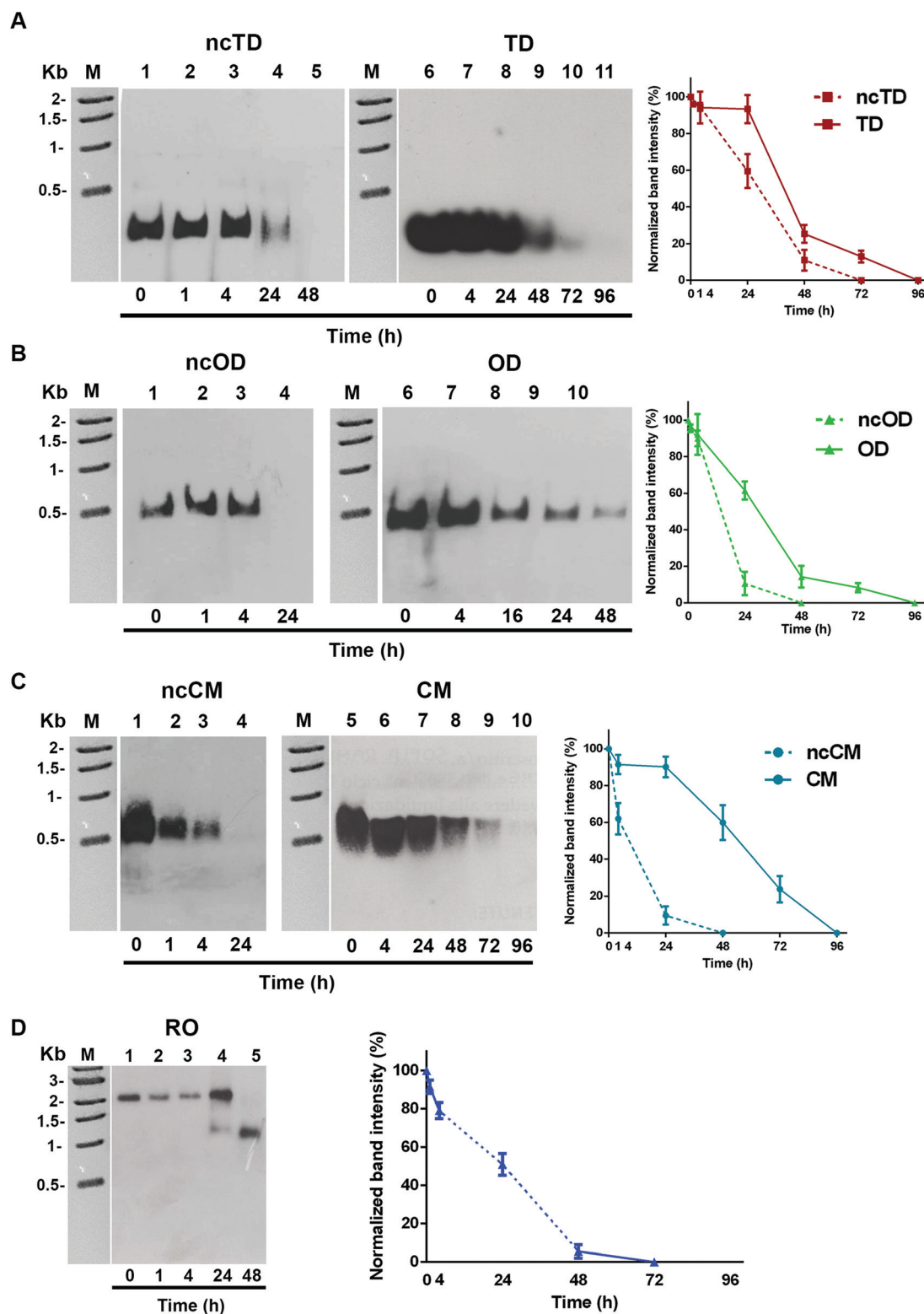
Fig. 2 *In vitro* stability of biotinylated non-covalent (ncDNS) and covalent DNS. DNA blot analysis of (A) ncTD and TD (B) ncOD and OD (C) ncCM and CM and (D) RO structures incubated with 10% FBS for different time intervals, as indicated. Purified DNS before incubation with serum proteins (time 0) are shown in lane 1 of each panel. The graphs show the densitometric analysis, performed by using ImageJ software, of three different experiments. The relative intensity of each band was normalized to the intensity of the band corresponding to each time 0 (lane 1) and reported in the graph. In (D) the dotted line represents the normalized band intensity of the RO origami obtained summing the intensity of the two visible bands and calculating the percentage of the upper band compared to the total. Values are expressed as an average ± SEM.

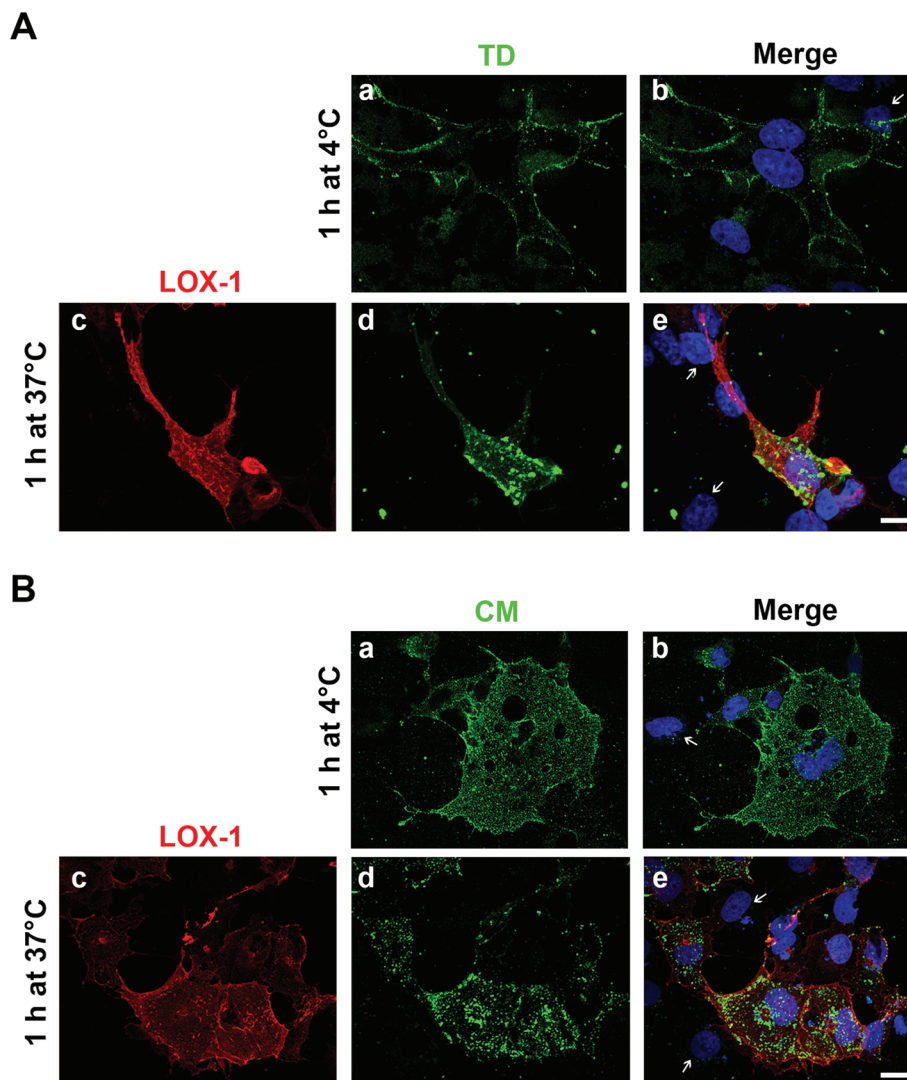**Fig. 3** Co-detection of LOX-1 receptors with TD and CM in COS transfected cells. Confocal analysis of double fluorescence of LOX-1-V5-expressing COS cells incubated with biotinylated TD (A) and CM (B) for 1 h at 4 °C (panels a and b) and 1 h at 37 °C (panels c–e). Intracellular LOX-1 receptors (panel c) were visualized using mouse monoclonal anti-V5 antibody (red), and biotinylated DNS were detected by using streptavidin-FITC (green). The nuclei were stained with DAPI. White arrows in the merge images indicate not-transfected cells. Scale bar: 20 μm.

After 1 h at 37 °C, RO structures appear in many small green fluorescent dots inside LOX-1-positive cells (red fluorescence) (Fig. 4, panels c and d). At 3 h of incubation at 37 °C, RO structures are detectable inside cells in larger fluorescent dots (Fig. 4, panel g), as already observed for TD and CM after 1 h (Fig. 3A and B, panels d). The merged images (Fig. 4, panels e and h) clearly demonstrate that RO structures are internalized only in LOX-1-positive cells (yellow and orange color in merged images). Co-detection experiments were repeated for SBO structures (ESI S6†). Fig. S6† shows representative confocal images in which LOX-1-expressing cells have unequivocally internalized SBO structures.

Notably, no binding or internalization of different biotinylated DNS was detected in not-transfected cells (ESI, S7†). Since the scavenger LOX-1 receptor interacts with negatively charged molecules, competition assays were performed to test

the specificity of the recognition for DNS (ESI, S8†). We show that a single-stranded oligonucleotide with a random sequence, at a concentration 10 times higher than TD, does not impair the TD binding to LOX-1 receptors (Fig. S6,† panel b). As a control, non-biotinylated assembled TD structures compete for the binding of biotinylated TD and lead to complete loss of binding, visualized by streptavidin–FITC (Fig. S6,† panel c).

Intracellular stability of DNS

Since confocal analysis does not allow verifying whether DNA nanostructures are partially degraded or maintain their structural integrity inside cells, we evaluated the integrity of internalized DNS by DNA blot after DNS purification from cells. In detail, structural stability during the internalization process was monitored by purifying biotinylated DNS after incubation
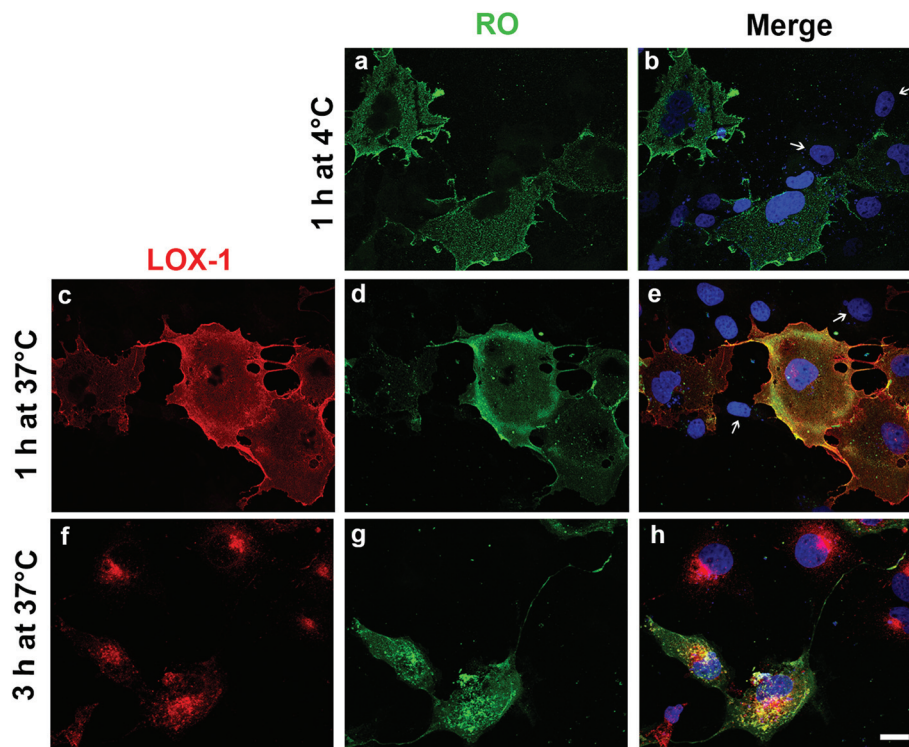
Fig. 4 Co-detection of LOX-1 receptor with RO in COS transfected cells. Confocal analysis of double fluorescence of LOX-1-V5-expressing COS cells incubated with biotinylated RO for 1 h at 4 °C (panels a and b), 1 h at 37 °C (panels c–e) and 3 h at 37 °C (panels f–h). Intracellular LOX-1 receptors (panels c and f) were visualized using mouse monoclonal anti-V5 antibody (red), and biotinylated origami were detected by using streptavidin-FITC (green). The nuclei were stained with DAPI. White arrows in the merge images indicate not-transfected cells. Scale bar: 20 μm.

with LOX-1-transfected COS cells at 37 °C for different time intervals. Samples were analyzed in 1% agarose gels, transferred onto Zeta-Probe membranes and visualized by using streptavidin–HRP (Fig. 5 and 6 and ESI S9–S11†). Covalently linked TD and CM, purified from cells after 1, 3 and 18 h of incubation (Fig. 5A and B), have electrophoretic mobility comparable to the input (lane 4 in Fig. 5A and lane 1 in Fig. 5B), indicating that they remain intact inside cells for at least 18 h. The densitometric analysis allowed us to calculate the amount, expressed in ng per 10^6 cells, of intact TD and CM internalized

by LOX-1-expressing COS cells at different times (Fig. 4C). After 1 h incubation, we detect 140 ± 33 ng per 10^6 cells and 261 ± 69 ng per 10^6 cells internalized TD and CM, respectively, increasing to 489 ± 107 ng per 10^6 cells and 541 ± 104 ng per 10^6 cells at 3 h, respectively. After 18 h of incubation, the amount of intact structures inside cells decreases to 348 ± 65 ng per 10^6 cells for TD and 406 ± 74 ng per 10^6 cells for CM, suggesting that DNSs undergo intracellular degradation associated with the LOX-1 pathway. The amount of intact internalized OD in LOX-1-expressing COS cells, calculated by setting up identical
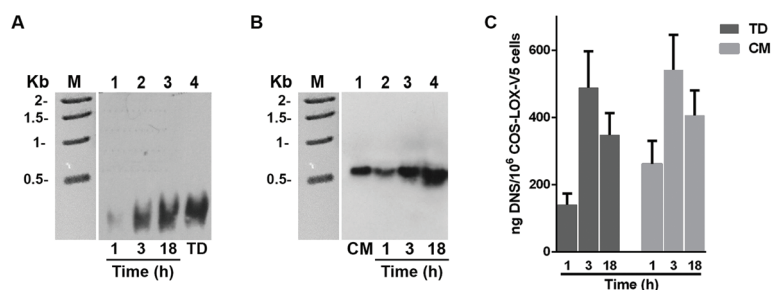
Fig. 5 LOX-1-mediated uptake of covalent TD and CM in cells. Representative DNA blots of cell extracts obtained from LOX-1-transfected COS cells incubated with 6 μg ml^{-1} of biotinylated TD (A) and CM (B) for 1, 3 and 18 h at 37 °C. DNS were detected with streptavidin-HRP. (C) Histogram shows the amount of internalized TD and CM detected in cell extracts of LOX-1-transfected cells at different times, as indicated and expressed in ng per 106 cells. For image processing and densitometric analysis, photographic films were digitized by scanning. Bands were analyzed by using ImageJ software. Values were normalized for the total number of LOX-1-transfected cells and expressed as a mean ± SEM.
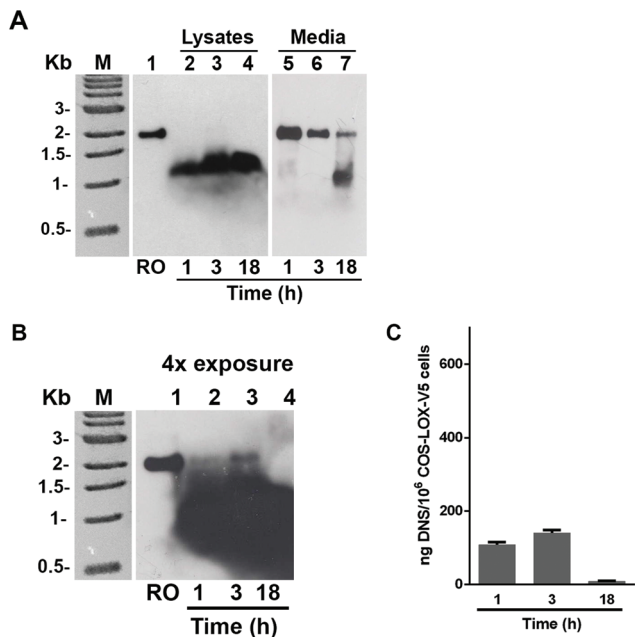
Fig. 6 LOX-1-mediated uptake of RO origami in cells. (A) Representative DNA blot of cell lysates and conditioned media obtained from LOX-1-transfected cells incubated with 6 μg ml⁻¹ of biotinylated RO for 1, 3 and 18 h at 37 °C. DNS were detected with streptavidin-HRP. (B) 4× exposure of DNA blot shown in A. (C) Histogram shows the amount of internalized RO detected in cell extracts of LOX-1-transfected cells at different times, as indicated and expressed in ng per 106 cells. Bands were analyzed by using ImageJ software. Values were normalized for the total number of LOX-1-transfected cells and expressed as a mean ± SEM.

experiments is 95 ± 15 and 219 ± 9 ng per 10^6 cells after 1 and 3 h, respectively, and 246 ± 25 ng per 10^6 cells after 18 h, values of the same order of magnitude of that obtained with TD and CM.[19] It is worth noting that not-transfected COS cells, after 1 h of incubation at 37 °C, internalize very low amounts of TD and CM, namely 4.6 ± 0.7 ng per 10^6 cells and 8.4 ± 0.9 ng per 10^6 cells, respectively (ESI, S9†).

Non-covalently linked TD, OD and CM structures are less stable than their covalently linked forms, as shown in Fig. 2. The lower stability of ncOD becomes more evident when we evaluated their stability inside COS-LOX-1 cells (ESI, S10†). Fig. S8A† shows that ncODs are barely detectable in cell lysates after 3 h incubation and no intact structures are found after 18 h of incubation.

Fig. 6 shows a representative DNA blot of RO structures extracted after incubation with cells for 1, 3 and 18 h. After 1 h incubation, the origami undergoes a degradation process that leads to the formation of a smaller band at 1 kb. The 2 kb band, corresponding to the intact structure before incubation with cells (input, Fig. 6A, lane 1), becomes barely visible after 1 h (Fig. 6A, lanes 2, 3 and 4). As shown in Fig. 6B, a 4× exposure of the blot is necessary to reveal the intact 2 kb band. The smaller 1 kb band accumulates in cells in a time-dependent manner. The densitometric analysis allowed us to evaluate the amount of intact RO (2 kb band) inside cells expressed

in ng per 10^6 cells at different time points. After 1 and 3 h of incubation (Fig. 6C), we calculate 65 ± 13 ng per 10^6 cells and 117 ± 11 ng per 10^6 cells and only a tiny amount of intact RO was detectable after 18 h incubation. Similar results were obtained for SBO structures after internalization in cells (ESI, S11†). Notably, the analysis of the RO structures in the conditional medium at the same time points (Fig. 6A, lanes 5, 6 and 7), shows that the structures are intact outside the cells and, only after 18 h incubation, the band at 1 kb becomes evident. This result indicates that RO structures initiate a degradation process only after cell internalization.

To study whether endocyted RO structures follow a degradation pathway where they are trafficked toward lysosomes, we incubated LOX-1-transfected COS cells with biotinylated RO at different time intervals. By performing a double immunofluorescence analysis using streptavidin–FITC and antibodies against the lysosomal-associated membrane protein (LAMP-1), we could track biotinylated-DNA structures and lysosomes. Fig. 7 shows representative confocal images of cells incubated for 1 and 3 h with 4 μg mL⁻¹ of RO. After 1 h, RO structures (green fluorescence, panel a) appear in small vesicles in the cytoplasm of LOX-1-expressing cells and only very few of them co-localize to lysosomes (Fig. 7, panels b and c). After 3 h of incubation at 37 °C, vesicles become larger and some of them appear to fuse with lysosomes (see the merged image in panel f). The correlation of the intensity values of the green and red pixels was evaluated by the Pearson's correlation coefficient (PCC).[44] In our analysis, the PCC values for RO and LAMP-1 marker is very low (0.18) after 1 h of incubation and increases to 0.38 after 3 h of incubation, indicating increased co-localization. This finding indicates that the internalized RO structures finally reach lysosomes after extended incubation times.

Discussion

DNA nanostructures of different sizes (from 8 to 80 nm) and shapes (tetrahedral, octahedral, cylindrical, square box and rectangular) are efficiently recognized by the scavenger receptor LOX-1 and internalized through a receptor-mediated mechanism. The recognition process occurs independently of the covalent or non-covalent nature of DNS that strongly perturbs their stability. Presence of covalent bonds make these structures more stable in biological liquids compared to the non-covalent isoforms (ncTD, ncOD and ncCM in Fig. 2A–C). In line with this observation, the half-life in serum of square box and rectangular origami structures is similar to that obtained with non-covalent DNS (Fig. 2D).

The LOX-1 receptor-mediated endocytosis of the natural substrate ox-LDL leads, after 1 hour incubation, to the formation of visible vesicles within cells that typically increase in volume over time, indicative of a vesicle fusion phenomenon.[45] Despite the efficient recognition of all DNS by cell surface-exposed LOX-1 receptors, we find differences in the kinetics of intracellular entry of the covalent tetrahedra and chainmails
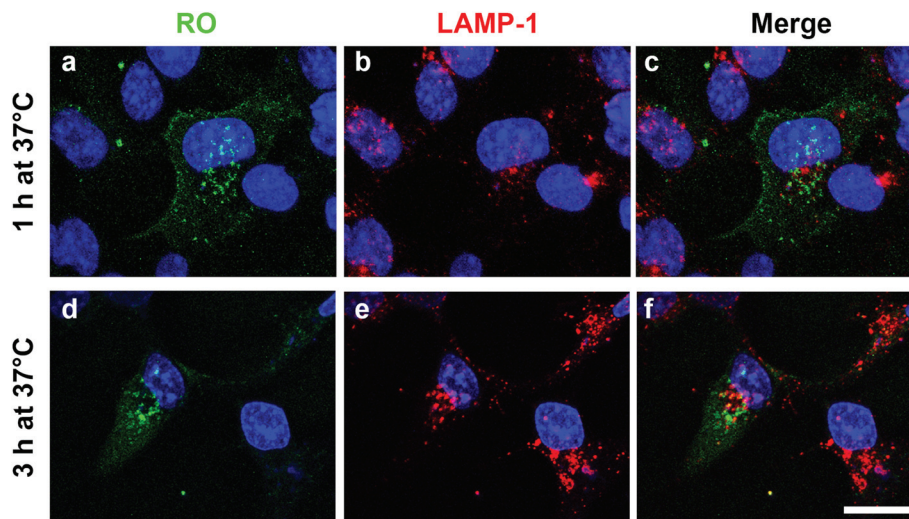
Fig. 7 Co-localization analysis of RO structures with the lysosomal marker LAMP-1. Double immunofluorescence of LOX-1-V5-expressing COS cells incubated with RO at 37 °C for 1 h (panels a–c) and 3 h (panels d–f). Biotinylated ROs were detected using streptavidin–FITC (green, panels a and d) and lysosomes were visualized using anti-LAMP-1 antibody (red, panels b and e). The nuclei were stained with DAPI. Merged images are shown in panels c and f. Co-localization analysis was performed using Imaris software. Scale bar: 20 μm.

when compared to the much larger square box and rectangular origami. Tetrahedral cages and chainmail, after incubation with LOX-1 expressing cells for 1 h at 37 °C, are found in small dots that tend to fuse in larger vesicles within a few hours, indicating that they traffic through lysosomes for destruction, as we have also demonstrated in the case of octahedral DNA cages.[19,24] When cells are incubated with RO or SBO structures, most of the green fluorescence is diffused into the cytoplasm in very small dots after 1 h at 37 °C (Fig. 4 and ESI, S6†). Co-localization analysis with the lysosomal marker LAMP-1 shows that most of these dispersed fluorescent dots are not co-localizing with lysosomes and, only after 3 h, vesicles become larger and some of them appear to fuse with lysosomes, indicating that they traffic towards lysosomes as TD, OD and CM, although with an apparent different kinetics. The difference in intracellular distribution of RO and SBO compared to TD and CM can be explained at least with two putative events, which are not mutually exclusive: (i) a rapid disassembly and release of the partially broken origami from vesicles to the cytosol or (ii) a different trafficking kinetics of origami structures compared to TD and CM, influenced by size and compactness.

It is well known that non-covalent DNS can be sensitive to divalent cation depletion.[46] Although high Mg^{2+} concentration is required for the assembly of all DNS described here, the stabilities of ncTD, ncOD and ncCM are not sensitive to post-assembly alteration in Mg^{2+} concentrations or to its depletion, while RO and SBO structures need at least 6 mM Mg^{2+} for maintaining their structural integrity in solution.[40] DNA origami susceptibility to divalent cations depletion provide a supporting argument for the first hypothesis. Notwithstanding that the intracellular Mg^{2+} concentration ranges from 5 to 20 mM, 95% of it is associated with proteins, phospholipids,

nucleic acids, chromatin and nucleotides and only a small percentage (1–5%) is present in its ionized form.[47] Thus, the low available Mg^{2+} concentration (ranging from 0.25 to 1 mM) may induce a rapid disassembly of the origami structures inside endocytic vesicles after entry in cells and release of partially broken origami in the cytoplasm (Fig. 4 and 6).

The nanoparticle visualization system used in this study, based on the recognition of the biotinylated structures with streptavidin, allows verifying the state of integrity of the particles inside cells, as a function of time. We find that the covalently linked DNA structures display higher stability in cells than the non-covalently linked ones. Covalently linked TD, OD, and CM remain intact inside cells at least for 18 h. On the other hand, origami structures, already after 1 h, appear to partially disassemble and degrade in a smaller structure of 1 kb band, still present after 24 hours inside cells. At present, the shape of this new structure is not clear. It may represent a more compact origami inside cells, but we need a more in-depth analysis of it to prove this. It is worth noting that adding covalent bonds to origami structures results in enhanced resistance against nuclease activity and in a longer half-life in 10% FBS at 37 °C, as recently reported.[48]

In conclusion, the scavenger receptor LOX-1 very efficiently binds and internalizes a broad family of DNA nanostructures, indicating that they can be used for efficient and selective delivery of useful cargo to the cells. However, the DNS studied in this work have a different fate and lifetime inside cells. RO and SBO structures are rapidly degraded whilst covalently linked tetrahedra, octahedra or rod-shaped chainmails are more stable. LOX-1 expression is not only markedly increased in many pathological conditions affecting the cardiovascular system but its up-regulation has been associated with tumor development in many cancer cells, including bladder, cervix,

mammary gland, lung, and colorectal, indicating that it is an interesting target molecule in different diseases.[33,34] Depending on the desired effect, different DNS may be preferred, such as, for example, origami structures for rapid burst drug release and covalent DNS for slower release.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

References

1 V. Linko, A. Ora and M. A. Kostiainen, *Trends Biotechnol.*, 2015, **33**, 586–594.

2 N. Xie, S. Liu, X. Yang, X. He, J. Huang and K. Wang, *Analyst*, 2017, **142**, 3322–3332.

3 J. J. Fakhoury, T. G. Edwardson, J. W. Conway, T. Trinh, F. Khan, M. Barłóg, H. S. Bazzi and H. F. Sleiman, *Nanoscale*, 2015, **7**, 20625–20634.

4 J. W. Keum and H. Bermudez, *Chem. Commun.*, 2009, **45**, 7036–7038.

5 A. S. Walsh, H. Yin, C. M. Erben, M. J. Wood and A. J. Turberfield, *ACS Nano*, 2011, **5**, 5427–5432.

6 L. Liang, J. Li, Q. Li, Q. Huang, J. Shi, H. Yan and C. Fan, *Angew. Chem., Int. Ed.*, 2014, **53**, 7745–7750.

7 J. Li, H. Pei, B. Zhu, L. Liang, M. Wei, Y. He, N. Chen, D. Li, Q. Huang and C. Fan, *ACS Nano*, 2011, **5**, 8783–8789.

8 Q. Jiang, C. Song, J. Nangreave, X. Liu, L. Lin, D. Qiu, Z. G. Wang, G. Zou, X. Liang, H. Yan and B. Ding, *J. Am. Chem. Soc.*, 2012, **134**, 13396–13403.

9 Y. X. Zhao, A. Shaw, X. Zeng, E. Benson, A. M. Nyström and B. Högberg, *ACS Nano*, 2012, **6**, 8684–8691.

10 V. Kumar, S. Bayda, M. Hadla, I. Caligiuri, C. Russo Spena, S. Palazzolo, S. Kempter, G. Corona, G. Toffoli and F. Rizzolio, *J. Cell. Physiol.*, 2016, **1**, 106–110.

11 M. Chang, C. S. Yang and D. Huang, *ACS Nano*, 2011, **5**, 6156–6163.

12 T. Watanabe, K. Hirano, A. Takahashi, K. Yamaguchi, M. Beppu, H. Fujiki and M. Suganuma, *Biol. Pharm. Bull.*, 2010, **33**, 796–803.

13 M. Koutsioumpa and E. Papadimitriou, *Recent Pat. Anti-Cancer Drug Discovery*, 2014, **9**, 137–152.

14 H. Lee, A. K. Lytton-Jean, Y. Chen, K. T. Love, A. I. Park, E. D. Karagiannis, A. Sehgal, W. Querbes, C. S. Zurenko, M. Jayaraman, C. G. Peng, K. Charisse, A. Borodovsky, M. Manoharan, J. S. Donahoe, J. Truelove, M. Nahrendorf, R. Langer and D. G. Anderson, *Nat. Nanotechnol.*, 2012, **7**, 389–393.

15 Z. Xia, P. Wang, X. Liu, T. Liu, Y. Yan, J. Yan, J. Zhong, G. Sun and D. He, *Biochemistry*, 2016, **55**, 1326–1331.

16 D. H. Schaffert, A. H. Okholm, R. S. Sørensen, J. S. Nielsen, T. Tørring, C. B. Rosen, A. L. Kodal, M. R. Mortensen, K. V. Gothelf and J. Kjems, *Small*, 2016, **12**, 2634–2640.

17 S. Raniolo, G. Vindigni, A. Ottaviani, V. Unida, F. Iacovelli, A. Manetto, M. Figini, L. Stella, A. Desideri and S. Biocca, *Nanomedicine*, 2018, **14**, 1181–1190.

18 D. Bhatia, S. Arumugam, M. Nasilowski, H. Joshi, C. Wunder, V. Chambon, V. Prakash, C. Grazon, B. Nadal, P. K. Maiti, L. Johannes, B. Dubertret and Y. Krishnan, *Nat. Nanotechnol.*, 2016, **11**, 1112–1119.

19 S. Raniolo, G. Vindigni, V. Unida, A. Ottaviani, E. Romano, A. Desideri and S. Biocca, *Nanoscale*, 2018, **10**, 12078–12086.

20 P. Wang, M. A. Rahman, Z. Zhao, K. Weiss, C. Zhang, Z. Chen, S. J. Hurwitz, Z. G. Chen, D. M. Shin and Y. Ke, *J. Am. Chem. Soc.*, 2018, **140**, 2478–2484.

21 M. M. C. Bastings, F. M. Anastassacos, N. Ponnuswamy, F. G. Leifer, G. Cuneo, C. Lin, D. E. Ingber, J. H. Ryu and W. M. Shih, *Nano Lett.*, 2018, **18**, 3557–3564.

22 P. C. Patel, D. A. Giljohann, W. L. Daniel, D. Zheng, A. E. Prigodich and C. A. Mirkin, *Bioconjugate Chem.*, 2010, **21**, 2250–2256.

23 C. H. Choi, L. Hao, S. P. Narayan, E. Auyeung and C. A. Mirkin, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 7625–7630.

24 G. Vindigni, S. Raniolo, A. Ottaviani, M. Falconi, O. Franch, B. R. Knudsen, A. Desideri and S. Biocca, *ACS Nano*, 2016, **10**, 5971–5979.

25 M. R. PrabhuDas, C. L. Baldwin, P. L. Bollyky, D. M. E. Bowdish, K. Drickamer, M. Febbraio, J. Herz, L. Kobzik, M. Krieger, J. Loike, B. McVicker, T. K. Means, S. K. Moestrup, S. R. Post, T. Sawamura, S. Silverstein, R. C. Speth, J. C. Telfer, G. M. Thiele, X. Y. Wang, S. D. Wright and J. El Khoury, *J. Immunol.*, 2017, **198**, 3775–3789.

26 T. Sawamura, N. Kume, T. Aoyama, H. Moriwaki, H. Hoshikawa, Y. Aiba, T. Tanaka, S. Miwa, Y. Katsura, T. Kita and T. Masaki, *Nature*, 1997, **386**, 73–77.

27 I. Ohki, T. Ishigaki, T. Oyama, S. Matsunaga, Q. Xie, M. Ohnishi-Kameyama, T. Murata, D. Tsuchiya, S. Machida, K. Morikawa and S. Tate, *Structure*, 2005, **13**, 905–917.

28 H. Park, F. G. Adsit, J. C. Boyington and J. Biol, *Chem.*, 2005, **280**, 13593–13599.

29 M. Falconi, S. Biocca, G. Novelli and A. Desideri, *BMC Struct. Biol.*, 2007, **7**, 73.

30 S. Biocca, M. Falconi, I. Filesi, F. Baldini, L. Vecchione, R. Mango, F. Romeo, G. Federici, A. Desideri and G. Novelli, *PLoS One*, 2009, **4**, e4648.

31 R. Yoshimoto, Y. Fujita, A. Kakino, S. Iwamoto, T. Takaya and T. Sawamura, *Cardiovasc. Drugs Ther.*, 2011, **225**, 379–391.

32 S. Balzan and V. Lubrano, *Life Sci.*, 2018, **198**, 79–86.

33 A. Hirsch, D. Iliopoulos, A. Joshi, Y. Zhang, S. A. Jaeger, M. Bulyk, P. N. Tsichlis, X. Shirley Liu and K. A. Struhl, *Cancer Cell*, 2010, **13**, 348–361.

34 M. Murdocca, R. Mango, S. Pucci, S. Biocca, B. Testa, R. Capuano, R. Paolesse, M. Sanchez, A. Orlandi, C. di Natale, G. Novelli and F. Sangiuolo, *Oncotarget*, 2016, **7**, 14765–14780.

35 S. Raniolo, G. Vindigni and S. Biocca, *Biomed. Spectrosc. Imaging*, 2016, **5**, S87–S99.

36 M. Falconi, S. Ciccone, P. D'Arrigo, F. Viani, R. Sorge, G. Novelli, P. Patrizi, A. Desideri and S. Biocca, *Biochem. Biophys. Res. Commun.*, 2013, **438**, 340–345.

37 S. Biocca, F. Iacovelli, S. Matarazzo, G. Vindigni, F. Oteri, A. Desideri and M. Falconi, *Cell Cycle*, 2015, **14**, 1583–1595.

38 S. Biocca, I. Filesi, R. Mango, L. Maggiore, F. Baldini, L. Vecchione, A. Viola, G. Citro, G. Federici, F. Romeo and G. Novelli, *J. Mol. Cell. Cardiol.*, 2008, **44**, 561–570.

39 V. Cassinelli, B. Oberleitner, J. Sobotta, P. Nickels, G. Grossi, S. Kempter, T. Frischmuth, T. Liedl and A. Manetto, *Angew. Chem., Int. Ed.*, 2015, **54**, 7795–7798.

40 A. H. Okholm, J. S. Nielsen, M. Vinther, R. S. Sørensen, D. Schaffert and J. Kjems, *Methods*, 2014, **67**, 193–197.

41 R. S. Sørensen, A. H. Okholm, D. Schaffert, A. L. Kodal, K. V. Gothelf and J. Kjems, *ACS Nano*, 2013, **7**, 8098–8104.

42 A. Cardinale, I. Filesi, V. Vetrugno, M. Pocchiari, M. S. Sy and S. Biocca, *J. Biol. Chem.*, 2005, **280**, 685–694.

43 R. Mango, S. Biocca, F. del Vecchio, F. Clementi, F. Sangiuolo, F. Amati, A. Filareto, S. Grelli, P. Spitalieri, I. Filesi, C. Favalli, R. Lauro, J. L. Mehta, F. Romeo and G. Novelli, *Circ. Res.*, 2005, **97**, 152–158.

44 S. Bolte and F. P. Cordelières, *J. Microsc.*, 2006, **224**, 213–232.

45 S. Matarazzo, M. C. Quitadamo, R. Mango, S. Ciccone, G. Novelli and S. Biocca, *Mol. Pharmacol.*, 2012, **82**, 246–254.

46 J. Hahn, S. F. Wickham, W. M. Shih and S. D. Perrault, *ACS Nano*, 2014, **8**, 8765–8775.

47 A. M. P. Romani, *Arch. Biochem. Biophys.*, 2011, **512**, 1–23.

48 T. Gerling, M. Kube, B. Kick and H. Dietz, *Sci. Adv.*, 2018, **4**, eaau1157.

Electronic Supporting Information

(ESI)

Cellular uptake of covalent and non-covalent DNA nanostructures with different size and geometry

Sofia Raniolo[1, 4], Stefano Croce[2], Rasmus P. Thomsen[3], Anders H. Okholm[3*], Valeria Unida[4], Federico Iacovelli[4], Antonio Manetto[5], Jørgen Kjems[3], Alessandro Desideri[4], Silvia Biocca[1§],

[1]Department of Systems Medicine, University of Rome Tor Vergata, Via Montpellier 1, 00133, Rome, Italy.

[2]Baseclick GmbH, Floriansbogen 2-4, 82061 Neuried, Germany

[3]Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Aarhus , Denmark.

[4]Department of Biology, University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133, Rome, Italy.

[5]Metabion, Gesellschaft für angewandte Biotechnologie mbH, 82152 Planegg, Germany

*Present address: Arla Innovation Centre, Arla Foods amba, Agro Food Park 19, 8200, Aarhus N, Denmark

§Correspondence to: biocca@med.uniroma2.it

S1.   Sequences of oligonucleotides used for the assembly of non-covalent and covalent DNA tetrahedra

Unmodified oligonucleotides were HPLC purified and purchased from Sigma Aldrich. The sequences of the oligonucleotides are reported in Table S1. The 5' of each oligonucleotide is phosphorylated. TTTTT represents a short non-pairing spacer inserted within the strands as a DNA junction at each vertex of the assembled 3D structure. OL3$_{BIO}$ has a biotin tetra-ethylene-glycol molecule (BtndT) at the T represented in red.

Oligo	Sequences
1	5'-P-CCACGGCGCTTTTTTTGCTGCTCCTTCCTTATCCTTTTTTTCGTCGACCGCCCAATTACTTTTTTTCCACTAATT-3'
2	5'-P-ACTCTCCCGTTTTTTTGACGACGCGCATTCCAGCTTTTTTTGTAATTGGGCGGTCGACGTTTTTTTGGCAAACCC-3'
3$_{BIO}$	5'-P-CGCGTCGTCTTTTTTTGTAAGTATGATTGTAGGGTTTTTTGCGCCGTGGAATTAGTGGTTTTTTTGCTGGAATG -3'
4	5'-P-CATACTTACTTTTTTTCGGGAGAGTGGGTTTGCCTTTTTTTGGAGTAAGGAAGGAGCAGCTTTTTTTCCCTACAAT-3'

Table S1.   Sequences of the oligonucleotides used for the assembly of Bio-DNA tetrahedra. In red is depicted the biotinylated T in OL3$_{BIO}$.

S2.   Sequences of oligonucleotides used for the assembly of non-covalent and covalent DNA chainmails

5'-Azide, 3'-Alkyne oligonucleotides were HPLC purified and purchased from Metabion, Germany. The sequences of the oligonucleotides are reported in Table S2. After folding, the oligonucleotides form a 6-helix bundle in which the 5' and 3' click functionalities are pre-organized and ready to react efficiently in presence of Cu(I).

Oligo	Sequences
hct-ODN1	5'-ZAAAACGCTAAGCCACCTTXAGATCCAAXCAGATACTCGGT-3'
ct-ODN2	5'-ZGGTCGTGCGGACTGTCGAACACCAACGATGCCTGATAGAAGX-3'
ct-ODN3	5'-ZGCGTGGCAATAGCCATAAATTCATACATAACGGCGCCAGACX-3'
ct-ODN4	5'-ZTTTCAAGACCAGCACTTGTATGGCGTAGGGCGGGTTTAGCGX-3'
ct-ODN5	5'-ZGGATCTAAAGGACTTCTATCAAAGACGGGACGACTCCGGGAX-3'
ct-ODN6	5'-ZGGCATCGTTGGAGTCTGGCGCACGACTTCGATTTCGGATCCX-3'
bct-ODN1	5'-ZCGTTATGTATGACGCTAAACCTTGCAATGACTGAACTCGAAX-3'
ct-ODN8	5'-ZCGCCCTACGCCAAAAAAGATGGGAGCTX-3'
hct-ODN2	5'-ZCGTTAXGTATGACGCTAAACCTTGCAATGACTGAACTCGAAT-3'
bct-ODN2	5'-ZGTCCCGTCTTAGGATCCGAAAGCCAXAATATATCGAGACGGX-3'
ct-ODN11	5'-ZTCGAAGTCGTATTCGAGTTCAAATGTCTATGCGATGCAGCAX-3'
ct-ODN12	5'-ZGTCATTGCAAAAGCTCCCATCATTTAATGTCGTTTACAGTAX-3'
ct-ODN13	5'-ZGATCAGCAGCGACCGTCTCGACTGCAGAAATAGGACCCCCAX-3'
ct-ODN14	5'-ZTATATTATGGCATGCTGCATCTTCCTGGCATGGCTGAATTCX-3'
bct-ODN3	5'-ZGCATAGACATTATACXGTAAAACCTTACGTAACTTACAGCCX-3'
ct-ODN16	5'-ZCGACATTAAATAAAAAAGATGAGTATTX-3'
hct-ODN3	5'-ZAAAATGGGGGTCCTCGAGGCGAAACAAXCAGATACTCGGT-3'
bct-ODN4	5'-ZATTTCXGCAGAGAATTCAGCCTATTCACATAGGCGAAGGCTX-3'
ct-ODN19	5'-ZATGCCAGGAAAGGCTGTAAGTTGCATCATGGGGGTCCTCAAX-3'
ct-ODN19	5'-ZTACGTAAGGTAAATACTCATCCCTGAGTGATCCATGACCCTX-3'
ct-ODN21	5'-ZGTTTCGCCTCGAAGCCTTCGCCCGCACGACCTGGCTTAGCGX-3'
ct-ODN22	5'-ZCTATGTGAATAATTGAGGACCATTGCCACGCTGTTCGACAGX-3'
bct-ODN5	5'-ZCCCATGATGCAAAGGGTCATGGGTCTTGAAAAATTTAXGGCX-3'
ct-ODN24	5'-ZGATCACTCAGGAAAAAAATACAAGTGCX-3'

Table S2.   Sequences of the oligonucleotides used for the assembly of DNA chainmails.
Z: Azide; X: Alkyne

S3.   Click reaction

Starting from the original design previously described[1], eight click tail oligonucleotides (ODN) were replaced with 5 bct-ODNs having the same sequence with an extra alkyne moiety and three hct-ODNs having the same sequence with an extra strand carrying an optional modification. The click reaction was performed using 40 µl of the solution containing the folded nanostructures (0.5 µM) where it was added 6 mg of the heterogenous Cu catalyst, 10 µL of Cu ligand Tris (benzyltriazolylmethyl) amine (THPTA baseclick GmbH) 0,1 M and 1 µL of Biotin-Peg3-Azide (baseclick GmbH) 1 mM. The solution was then incubated at 32 °C for 5 hours gently mixing at 200 rpm. The mixture was then transferred in a fresh vial and the Cu catalyst discharged. The sample was analyzed without any further purification step. For cell experiments, the chainmails were purified *via* EtOH precipitation.
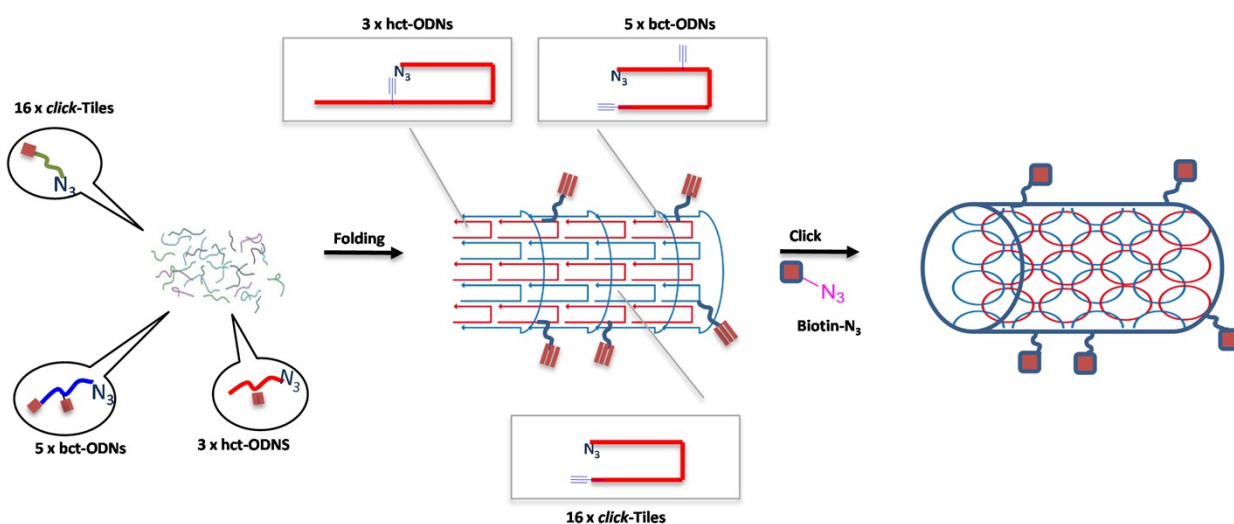

Figure S1.   Schematic representation of the DNA chainmail functionalization. After the folding reaction *via* one-pot click reaction, it is possible to close the click tails leading to chainmail formation and, at the same time, to labeling the structure with biotin. Nomenclature: bct-ODN = body (functional) click tile; hct-ODN = head (functional) click tile.

S4.   Stability of square box origami structures in 10%FBS

Square box origami (SBO) were incubated at 37°C with 10% fetal bovine serum (FBS) for different time intervals, as indicated. After incubation, SBO samples were run in 1% agarose gel to examine the structural integrity. The presence of 32 biotin molecules allows the detection of SBO through a streptavidin-HRP-biotin reaction in DNA blots.

As shown in Figure S4, SBO structures start to lose their integrity between 3 and 24 h. Notably, the SBO loss of structure and degradation causes an increase in the intensity of the signal of the band detected by the streptavidin-biotin reaction in DNA blot (lane 4) as observed for RO (Figure 2, lanes 4 and 5).
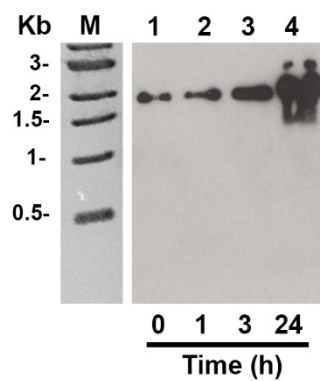

Figure S2.   Stability of biotinylated SBO in serum. DNA blot analysis of SBO incubated with 10% FBS. 20 ng of SBO before incubation with serum proteins (time 0) are shown in lane 1. Incubation times are indicated under the gel.

S5.   Membrane LOX-1 receptors visualized with Mab anti-V5 IgG

COS cells were plated on poly-L-lysine coated glass slides and then transfected with pEF/V5-LOX-1, a plasmid encoding for the full-length LOX-1 receptor containing a V5-tag at the C-terminus (LOX-1-V5). After 24 hours from transfection, cells were incubated with Mab anti-V5 for 1 h at 4°C for visualizing LOX-1-V5 receptors expressed at the cell surface (Figure S5). Cells were washed in PBS, fixed in 4% paraformaldehyde and neutralized with $NaBH_4$. Rhodamine Red-X-conjugated AffiniPure donkey anti-mouse IgG was used as secondary antibody and nuclei were stained with DAPI. In Figure S5 LOX-1 receptors are represented by red fluorescent intense dots, indicating that the receptors localize on the outer surface of the plasma membranes. It is worth noting that within transfected LOX-1 cells some cells (white arrows) do not express LOX-1 according to the transfection efficiency of 40% and do not show any red membrane fluorescence around the blue nuclei.
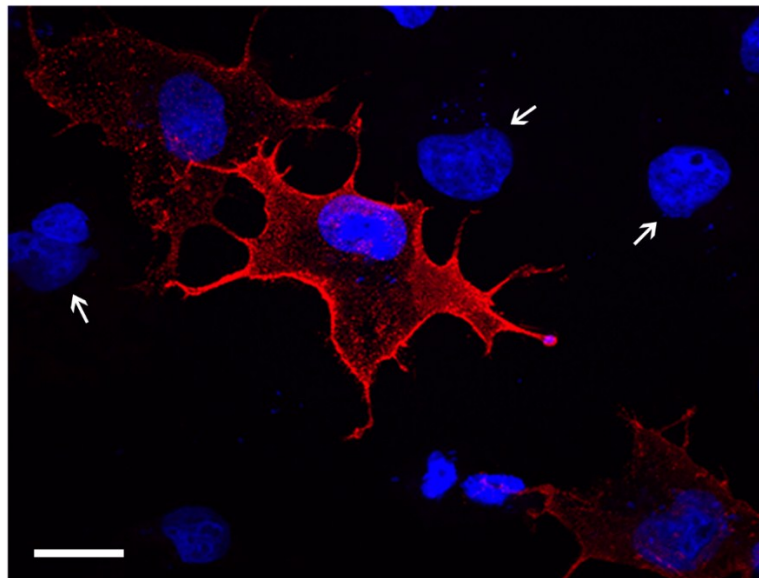

Figure S3.   Confocal analysis of LOX-1-expressing COS cells incubated with Mab anti-V5 for 1 h at 4°C. White arrows indicate some not-transfected cells. Images were obtained with a laser confocal fluorescence microscope Olympus FV1000 at 60X magnification and processed by IMARIS software. Scale bar: 20 μm.

S6. Co-detection of LOX-1 receptors and SBO structures in LOX-1 expressing COS cells.

For the binding experiment, COS cells were transfected with pEF/V5-LOX-1 plasmid and, after 24 h, incubated with biotin-labeled SBO origami (4 µg/mL) for 1 h at 4°C, fixed in 4% paraformaldehyde and neutralized with NaBH$_4$. Biotinylated SBO were detected with streptavidin-FITC (Jackson) and nuclei stained with DAPI (Invitrogen). In Figure S6, panels a and b, SBO origami are detected as green fluorescent intense dots, indicating that they localize on the plasma membrane surface.

For co-detection of LOX-1 receptors and SBO structures, LOX-1-transfected COS cells were incubated with biotinylated SBO (4 µg/mL) for 1 h at 37°C, fixed with 4% paraformaldehyde and permeabilized with Tris-HCl 0.1M pH 7.6/Triton 0,1%, for allowing the detection of both internalized structures and intracellular LOX-1 receptors. Cells were stained with anti-V5 antibodies and streptavidin−FITC, for visualizing LOX-1 receptors and biotinylated SBO respectively and with DAPI for localizing the nuclei. Figure S6 shows representative confocal images in which unequivocally LOX-1-expressing cells (red fluorescence, panel c) have internalized SBO structures (green fluorescence, panels d and e).
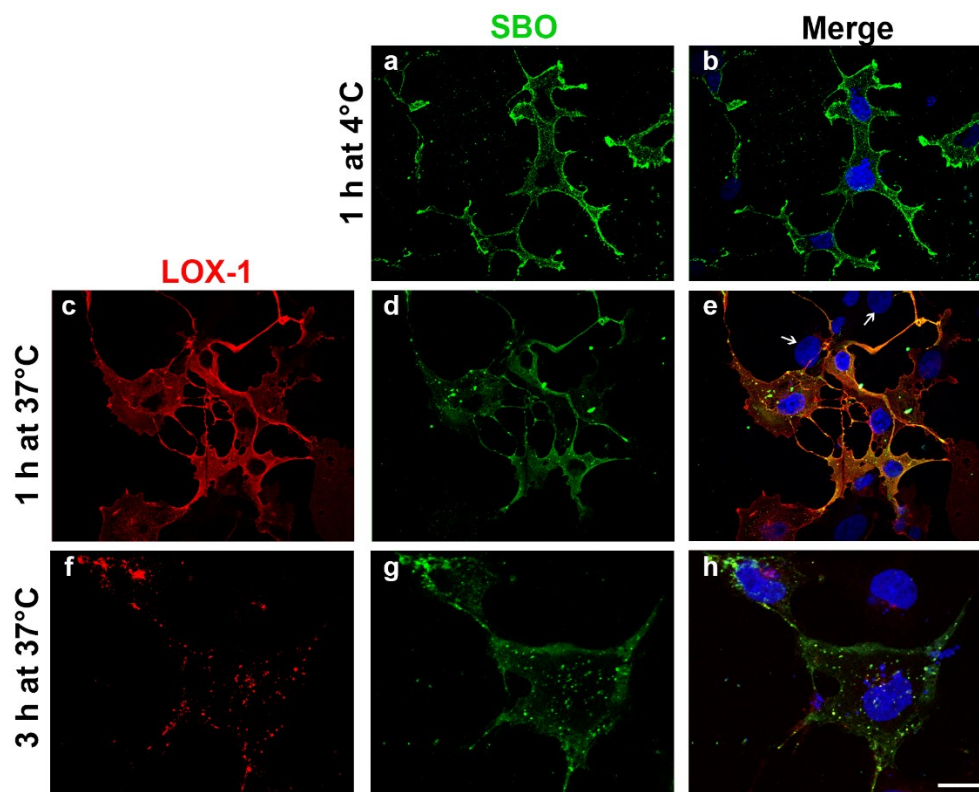
Figure S4. **Binding of SBO to LOX-1-expressing COS cells and co-detection of LOX-1 receptors and SBO structures.** Confocal analysis of LOX-1-expressing COS cells incubated with biotinylated SBO for 1 h at 4°C (panels a and b) and with biotinylated SBO for 1 h at 37°C (panels c, d, and e). SBO were detected by using streptavidin−FITC (green) and LOX-1 receptors were visualized using anti-V5 antibody (red). Nuclei were stained with DAPI (blue). White arrows show not-transfected cells. Images were obtained with a laser confocal fluorescence microscope Olympus FV1000 at 60X magnification and processed by IMARIS software. Scale bar: 20 µm.

S7. Confocal analysis of biotinylated DNS internalized in not-transfected COS cells compared to LOX-1 transfected cells.

In order to demonstrate a specific binding between DNS and LOX-1 receptors, we have performed uptake experiments on LOX-1 transfected (COS-LOX-1) or not-transfected (COS nt) COS cells. In detail, cells were incubated with biotinylated DNS (10 µg/mL for TD and CM and 4 µg/mL for RO) for 1 h at 37°C, fixed with 4% paraformaldehyde and permeabilized with Tris-HCl 0.1M pH 7.6/Triton 0,1%, for allowing the detection of the internalized structures. Cells were stained with streptavidin−FITC for visualizing biotinylated DNS and with DAPI for localizing the nuclei. Figure S7 shows representative confocal analysis in which LOX-1-expressing cells (panel a, c, and e) have internalized biotinylated DNS while no green fluorescence was detectable inside the COS nt cells (panels b, d, and f) for any of the DNS tested, indicating a LOX-1-dependent cellular uptake of DNS.
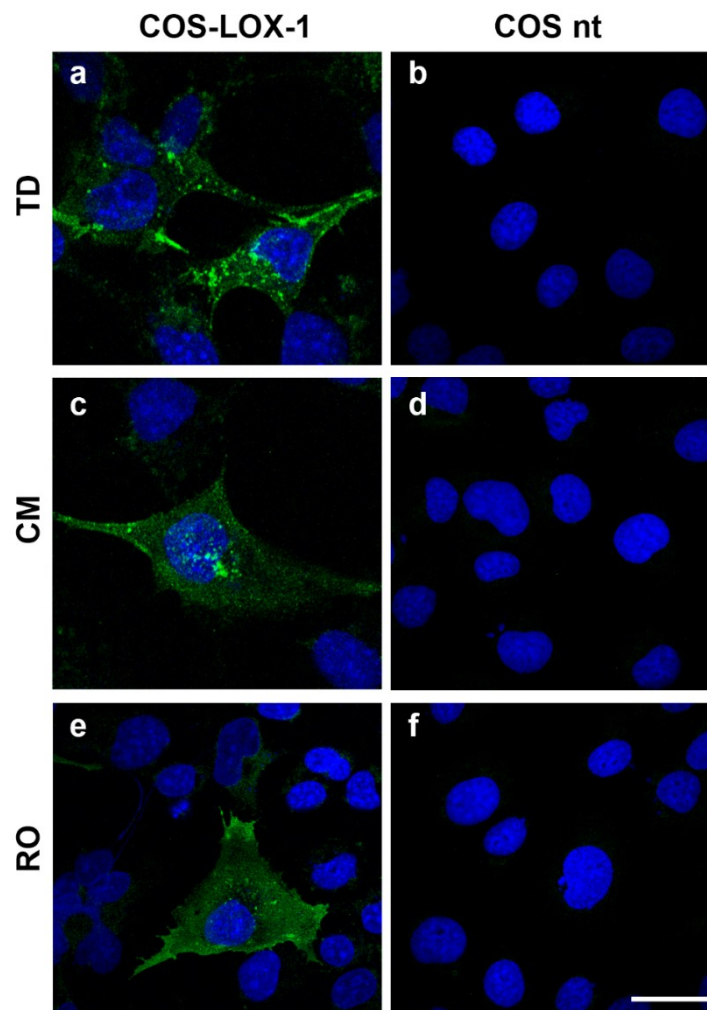

Figure S5. **Representative confocal images of DNS uptake in COS cells.** COS-LOX-1 (panels a, c, and e) and COS nt cells (panels b, d, and f) were incubated for 1 h at 37°C with biotinylated TD, CM, and RO. Biotinylated DNS were detected with streptavidin-FITC and nuclei were blue stained

with DAPI. Images were obtained with a laser confocal fluorescence microscope Olympus FV1000 at 60X magnification and processed by IMARIS software. Scale bar, 20 μm.

S8. Competition assays.

To study the specificity of LOX-1 recognition, two competition binding assays were performed, one in the presence of 10 times more concentrated single stranded oligonucleotide and one in the presence of non-biotinylated DNS. Figure S8 shows representative confocal images of the competition assays performed by incubating LOX-1 expressing cells with 144 nM biotinylated tetrahedral nanocages (Bio-TD) in the absence (panel a), in the presence of 1,4 μM 29 nt single-stranded oligonucleotide with random sequence (10x oligo, panel b) and in the presence of 1,4 μM of non-biotinylated (10x TD) (panel c). In detail, 24h after transfection, COS cells were simultaneously incubated with Bio-TD and 10x oligo or 10x TD for 1 h at 4°C. Cells were then fixed in 4% paraformaldehyde and neutralized with NaBH4. Biotinylated TD were detected with streptavidin-FITC (Jackson) and nuclei stained with DAPI (Invitrogen). As shown in Figure S8, panel b, the presence of 10x oligo does not impair Bio-TD binding to LOX-1 expressing cells. Of note, 10-fold excess of not-biotinylated fully assembled TD almost completely abolishes TD binding to LOX-1 (panel c), indicating the specificity of the binding.
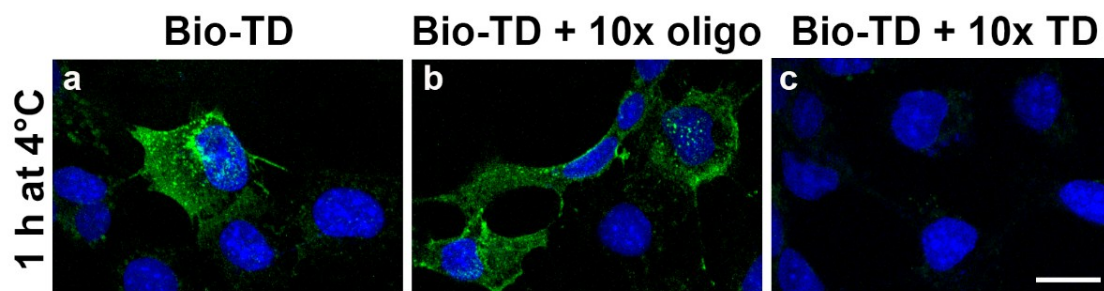
Figure S6. Competition assays. LOX-1 transfected COS cells were incubated with 144 nM Bio-TD (panel a) in the simultaneous presence of 10-fold excess of a single-stranded oligonucleotide (panel b) or non-biotinylated TD (panel c) for 1 h at 4°C. Biotinylated TD were detected with streptavidin-FITC and nuclei are blue stained with DAPI. Images were obtained with a laser confocal fluorescence microscope Olympus FV1000 at 60X magnification and processed by IMARIS software. Scale bar, 20 μm.

S9. TD and CM internalization in not-transfected COS cells compared to LOX-1 transfected cells.

Not expressing (-) and LOX-1 expressing (+) COS cells were incubated with 6 μg/mL of biotinylated TD and CM for 1 h at 37 °C. After incubation, cells were lysed, centrifuged and supernatants digested with proteinase K for removing the bound proteins that surround the surface of DNS. Samples were analyzed by DNA blot. Figure S9 shows the amount of tetrahedral (TD, panel A) and chainmail (CM, panel B) internalized by not-transfected (-) compared to LOX-1 transfected COS cells (+). The amount of DNS found in not-transfected cells was calculated by densitometric analysis. After 1 h of incubation at 37°C, not-transfected COS cells internalize a very low amount of TD and CM that have been quantified to be 4.6 ± 0.7 ng/10^{6cells} and 8.4 ± 0.9 ng/10^{6cells}, respectively. At the same time of incubation, LOX-1-expressing COS cells internalize 140 ± 33ng/10^{6cells} of TD and 261 ± 69 ng/10^{6cells} of CM, confirming that LOX-1-mediated uptake leads to an increase of about 30-fold in the amount of internalized DNS.
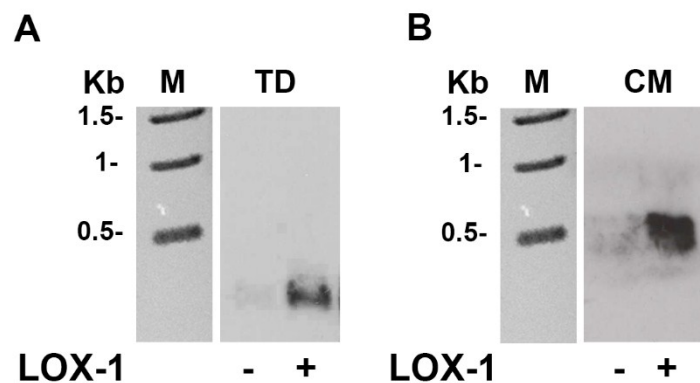

Figure S7. Representative DNA blots of cell extracts obtained from COS cells expressing (+) or not expressing (-) LOX-1 receptor incubated with TD (A) and CM (B). Biotinylated DNS were detected with streptavidin-HRP and visualized by enhanced chemiluminescence.

S10.   Stability of non-covalently-linked OD structures in cells.

Non covalently-linked DNS are more unstable in biological fluids. For evaluating their stability inside cells, we incubated biotinylated non covalent octahedral nanocages (ncOD) (6 µg/mL) with LOX-1-transfected COS cells for 3 and 18 h at 37 °C, we purified them from cell lysates and conditioned media, treated with proteinase K (100 µg/mL) for 1 h at 37°C and run on 1% agarose gel in TBE (Tris-Cl 89 mM pH 8, boric acid 89 mM, EDTA 2mM). A representative DNA blots is shown in Figure S10A. Lanes 1 and 4 show the electrophoretic mobility of ncOD before incubation with cells (time 0). After 3 h of incubation, intact ncOD structures are barely detectable in cells lysates (lane 2) and, after 18 h of incubation, only a smear is visible in cell lysates indicating that the nanostructures have been degraded (lane 3). Notably, the analysis of the ncOD in the conditional medium at the same time points (lanes 5 and 6), shows that after 18 h the ncOD structures are completely degraded, confirming their lower stability, when compared to the covalent DNA nanostructures.

For evaluating whether ncOD undergoes disassembly or degradation at the level of the plasma membrane, biotinylated ncOD (6 µg/mL) were incubated with LOX-1-transfected COS cells for 1 h at 4 °C, for avoiding cell internalization and analyzing the ncOD bound to the cell surface. Analysis of the purified ncOD by DNA blot indicates that, despite the low amount of membrane bound ncOD (Figure S10B, lane 2), only a single product is present in the gel with a mobility comparable to the input (lane 1), indicating that binding of ncOD to the receptor is not inducing disassembly or degradation of the structures.

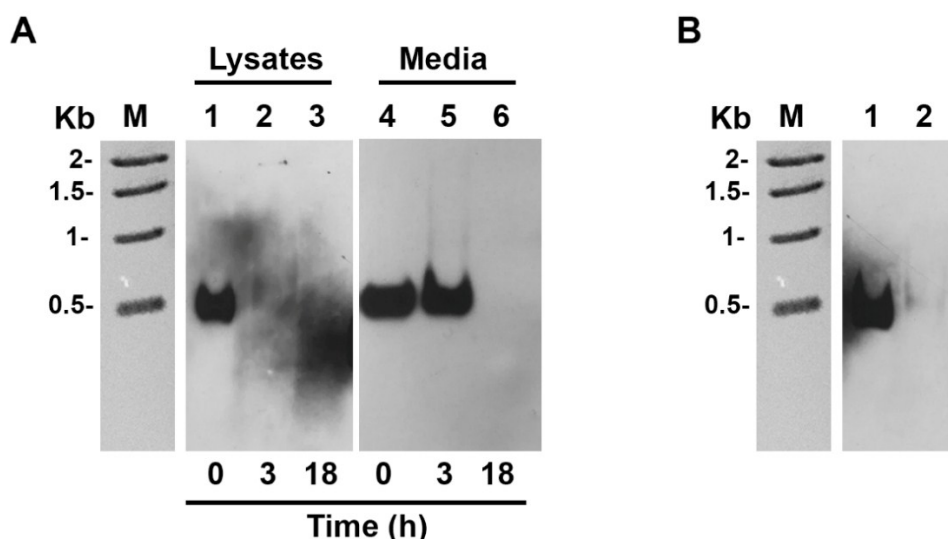**Figure S8.   LOX-1-mediated uptake of ncOD and their binding to cells.** (A) Representative DNA blots of lysates and conditioned media from LOX-1-transfected cells incubated with biotinylated ncOD for different times at 37 °C, as indicated. (B) Representative DNA blot of lysates of LOX-1-transfected cells incubated with biotinylated ncOD for 1 h at 4°C. NcOD were detected with streptavidin-HRP.

S11.   Stability of SBO structures in cells.

Biotinylated SBO (6 µg/mL) were incubated with LOX-1-transfected COS cells for 1, 3 and 18 h at 37 °C, treated with proteinase K (100 µg/mL) for 1 h at 37°C and run on 1% agarose gel in TBE (Tris-Cl 89 mM pH 8, boric acid 89 mM, EDTA 2mM), supplemented with 6.25 mM MgCl$_2$.

DNA blots were visualized by using streptavidin-HRP. Lane 4 in Figure S11 shows the electrophoretic mobility of SBO before incubation with cells (input). Purified SBO, after incubation with cells, run with a lower electrophoretic mobility compared to the input (Figure S11, compare lanes 1, 2 and lane 3 with lane 4) and a DNA product with a lower molecular weight, corresponding to 1 kb, is detected inside cells, as reported for RO (Figure 6, lanes 2, 3 and 4).
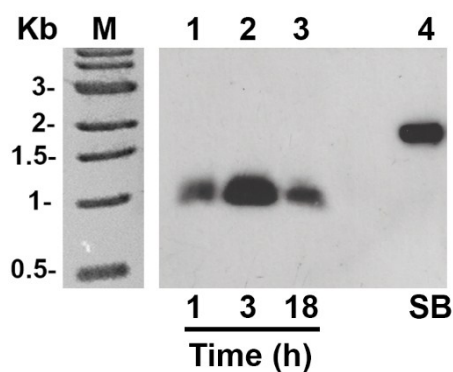

Figure S9.   LOX-1-mediated uptake of SBO origami in COS cells. Representative DNA blot of cell extracts derived from LOX-1-transfected COS cells incubated with biotinylated SBO for 1, 3 and 18 h at 37 °C.

References:

1. V. Cassinelli, B. Oberleitner, J. Sobotta, P. Nickels, G. Grossi, S. Kempter, T. Frischmuth, T. Liedl, A. Manetto. Angew Chem Int Ed Engl. 2015, **54**, 7795-8.

6. Chemo-Enzymatic Preparation of functional click labelled Messenger RNA

6.1. Prologue

Nowadays, *in vitro* transcribed (IVT) messenger RNA (mRNA) has attracted attention as a valid candidate in next generation-therapeutics and is considered an asset in treatment of several diseases[20]. This novel class of drugs has already demonstrated feasibility in several areas including vaccination[49,50], protein replacement[74] and cancer therapy[21]. In the last decades, research groups and especially several companies base their own research and business on the development of mRNA therapeutics, including *e.g.* ModeRNA (USA), CureVac, BioNTech (Germany) and AstraZeneca (UK). These companies have gathered tremendous attention and were able to raise important fundings for their next development steps[20], since this method is expected to play a key role in developing and facilitating a future approach to medicine.

The main idea behind the so-called "mRNA therapy" is to provide IVT mRNA as a carrier of genetic information, thus allowing the organism to develop its own cure. For example in vaccination, mRNA coding for a specific antigen is used to generate an immune response and lead to the formation of B and T-memory cells Figure 11. In the case of protein replacement therapy, mRNA produces the missing protein causing the disease. Several advantages have made mRNA more attractive compared to plasmid DNA. Among others, the usage of mRNA avoids insertional mutagenesis and potential oncogenesis considering that its activity does not require nuclear and genome integration. For this reason, it is also an ideal candidate for protein expression in non-dividing cells.
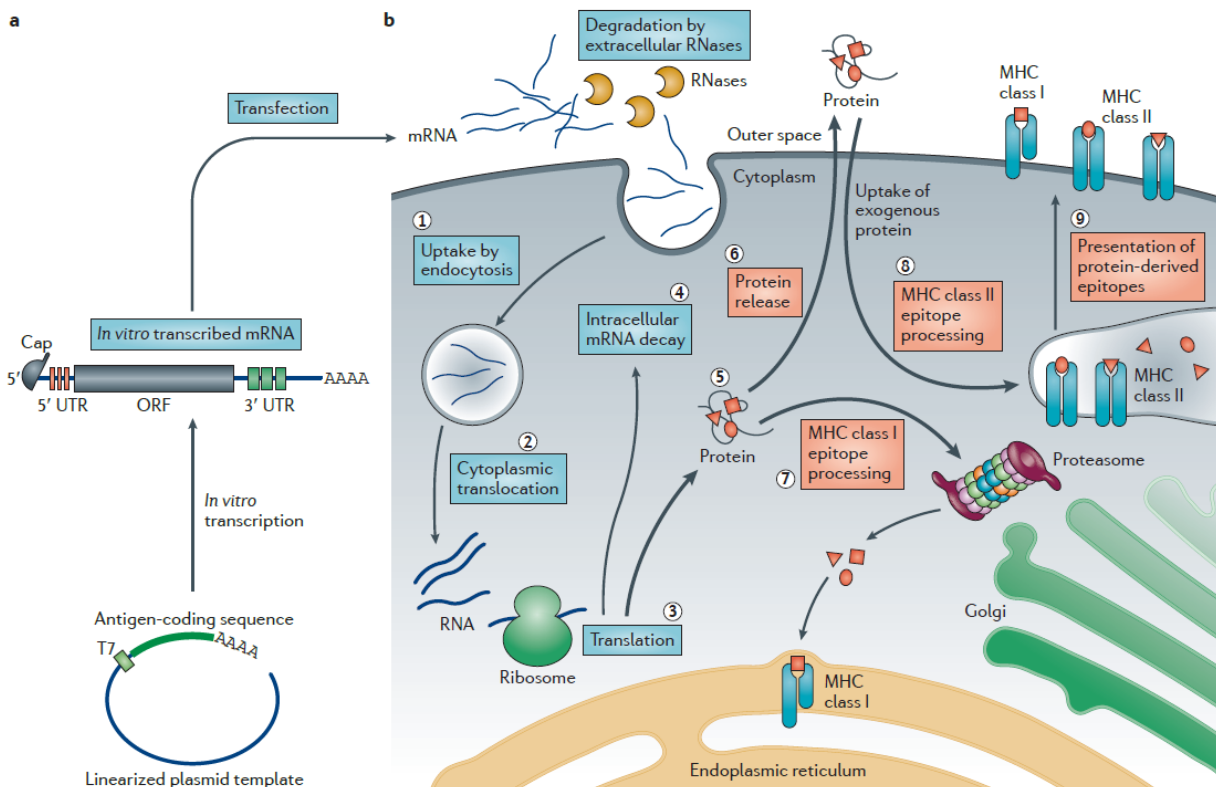
Figure 11: Schematic representation of the usage of mRNA as therapeutic agent, in this particular case in vaccination. U.Sahin *et al., Nature reviews, 2014, 13, 759-780*

Several groups started to work on chemically modified RNA to expand its chemical properties[75,76], but so far the biocompatibility of such modification has not been shown.

Despite the recent advances, there are still some challenges that need to be addressed using IVT mRNA. The bio-chemical properties of the RNA render the molecule sensitive towards physical and biochemical degradation either in intracellular and extracellular environments. Delivery to a specific target, tissue or organ, is still a major challenge for the usage in therapies. To date, mRNA is delivered by liposomal drug delivery systems[29] , that protect mRNA from degradation in extracellular environments. Anyhow, it was observed that lipid nanoparticles are mostly directed and degraded to the liver, that in most of the cases is not the final target for the therapy. Furthermore this approach has crucial variations between *in vitro* and *in vivo* protein expression[32]. Delivery with lipid nanoparticles containing mRNA has a better expression level in cell culture compared to administration in living organisms, while the opposite effect is observed when "naked" mRNA is transfected.

One important discovery to the advancement of mRNA technology was the incorporation of modified bases during transcription. The most prominent example is the usage of the N1-methyl-pseudouridine or pseudouridine. The replacement of natural uridine with pseudouridine during *in vitro* transcription, drastically reduces the immune response against external RNA molecules and at the same time increase its translation efficiency[77,78].

The main goal of this project was to develop the cutting-edge technology to enhance the biochemical properties of chemically modified mRNA in a cross-disciplinary academia-industry collaboration between the LMU München and baseclick GmbH. The technology is based on the incorporation of azide and/or alkyne nucleotides during the enzymatic synthesis of the IVT mRNA in order to achieve the modification of the resulting molecules post-transcriptionally, making possible to choose among a plethora of modifications using click chemistry.

6.2. Abstract of the publication "Chemo-Enzymatic Preparation of functional click labelled Messenger RNA"

Alkyne and or azide-modified nucleotides are enzymatically incorporated into mRNA and subsequently conjugated to fluorescent dyes using click chemistry. This allows visualization of the labelled mRNA inside of cells. mRNA coding for the enhanced green fluorescent protein (eGFP) was chosen as a model system and the successful expression of eGFP demonstrated that our modified mRNA is accepted by the translation machinery.

6.3. Author contribution

The author designed the experiments and analyzed the data. Performed all the mRNA preparations,the Hela cells experiments and wrote the manuscript.

6.4. Associated Publication

Chemo-Enzymatic Preparation of Functional Click Labelled Messenger RNA

Chemoenzymatic Preparation of Functional Click-Labeled Messenger RNA

Stefano Croce,[a, b] Sascha Serdjukow,[a] Thomas Carell,[b] and Thomas Frischmuth*[a]

Synthetic mRNAs are promising candidates for a new class of transformative drugs that provide genetic information for patients' cells to develop their own cure. One key advancement to develop so-called druggable mRNAs was the preparation of chemically modified mRNAs, by replacing standard bases with modified bases, such as uridine with pseudouridine, which can ameliorate the immunogenic profile and translation efficiency of the mRNA. Thus the introduction of modified nucleobases was the foundation for the clinical use of such mRNAs. Herein we describe modular and simple methods to chemoenzymatically modify mRNA. Alkyne- and/or azide-modified nucleotides are enzymatically incorporated into mRNA and subsequently conjugated to fluorescent dyes using click chemistry. This allows visualization of the labeled mRNA inside cells. mRNA coding for the enhanced green fluorescent protein (eGFP) was chosen as a model system and the successful expression of eGFP demonstrated that our modified mRNA is accepted by the translation machinery.

Introduction

Messenger RNA (mRNA) is the central molecule involved in transferring the genetic information from DNA into proteins. Eukaryotic mRNA is usually composed of five different parts: a 5′-CAP,[1–3] a 5′-untranslated region (5′-UTR), an open reading frame (ORF), a 3′-untranslated region (3′-UTR)[4, 5] and a poly(A) tail.[6] While the amino acid sequence is encoded in the open reading frame, the CAP and the poly(A) tail enhance cellular stability and translation efficiency.[6] The 5′-UTR and 3′-UTR are often folded in a secondary structure, which provides certain stability to the mRNA, thus influencing mRNA's half-life and turnover rate.[4, 5] To date synthetic mRNAs are prepared enzymatically by in vitro transcription from DNA templates,[7] as the chemical synthesis of long RNAs is still not feasible.

In vitro-transcribed (IVT) mRNA has recently become a promising candidate for the development of a new class of drugs,[8] which in principal should allow treatment of all diseases that require a gain-of-function such as enzyme replacement therapy, vaccination, immune response, etc. Due to its transient properties, mRNA has almost no chance for genome integration in contrast to DNA based gene-therapy approaches. At the same time high protein expression levels are often required. Therefore, the administered mRNA must be sufficiently stable to allow efficient translation in the target cells.

Preparation of chemically modified mRNA has been one key element to advance the mRNA technology toward drug development. By using various modified nucleotides, the half-life of the mRNA, the translation efficiency and its immunological profile are improved. For example, N1-methyl-pseudouridine results in more efficient translation and at the same time represses innate immune responses in comparison with the unmodified mRNA.[9]

Apart from incorporation of naturally occurring modified nucleotides, artificial modified nucleotides offer an interesting option to expand the (chemical) properties of the resulting mRNA. The seminal work from the groups of Hirao, Benner, and Romesberg has resulted in several artificial base pairs that can be replicated and even transcribed into RNA.[10–14] For example by using the artificial NaM:TPT3 base pair it was possible to achieve site-specific RNA labeling. A cyclopropene-modified TPT3 nucleotide is incorporated by T7 RNA polymerase during in vitro transcription opposite to the dNaM base in the DNA template and was subsequently labeled with dye conjugates through an inverse electron demand Diels–Alder reaction.[15] More recently, Walunj et al. described an RNA-labeling method which involves incorporation of 5-iodouridine triphosphate (IUTP) by in vitro transcription and subsequent labeling using a Suzuki–Miyaura cross-coupling reaction.[16] Modification of mRNA by these artificial nucleotides could profit from increased stability and functionality, but biocompatibility for cells and the translation machinery are yet unknown.

Herein we present four different methods to modify IVT mRNA via enzymatic incorporation of alkyne and/or azide-modified unnatural nucleotides and subsequent labeling using click chemistry. We were focused on a strategy to chemically modify IVT mRNA that could modulate properties such as stability, delivery and tracking but still allows translation into functional proteins. We believe that the establishment of an effec-

[a] S. Croce, Dr. S. Serdjukow, Dr. T. Frischmuth
     baseclick GmbH
     Floriansbogen 2–4, 82061 Neuried (bei München) (Germany)
     E-mail: t.frischmuth@baseclick.eu

[b] S. Croce, Prof. Dr. T. Carell
     Department of Chemistry, Ludwig-Maximilians-Universität München
     Butenandtstrasse 5–13, 81377 München (Germany)

     Supporting information and the ORCID identification numbers for the
     authors of this article can be found under https://doi.org/10.1002/
     cbic.201900718.

ChemBioChem

Full Papers
doi.org/10.1002/cbic.201900718

Chemistry Europe
European Chemical
Societies Publishing

tive chemo-enzymatic method for modification of still functional IVT mRNA will allow advancement of the current mRNA therapy development, as chemical modification has already been demonstrated to be useful for a short interfering (si)RNA drug,[17] which has been chemically coupled to the sugar *N*-acetylgalactosamine for better receptor-mediated uptake.[18]

Results and Discussion

eGFP mRNA model

To establish our mRNA labeling technology, an mRNA encoding for the enhanced green fluorescent protein (eGFP) was selected, because correct translation of the functional protein can be easily monitored directly via its green fluorescence properties. The desired mRNA constructs were prepared by in vitro transcriptions from a linearized DNA template (sequence in the Supporting Information) containing a T7 promotor, a 5′-UTR, the eGFP gene, a 3′-UTR composed of two repeats of the 3′-UTR of the β-globin and a 120-mer poly(A) tail. In addition, a 5′-CAP structure was co-transcriptionally introduced by T7 RNA polymerase using an anti-reverse capping analogue (ARCA). In this way, the transcription resulted in a stable mRNA including CAP structure and poly(A) tail, ready for efficient expression in eukaryotic cells. Chemoenzymatic approaches involving enzymatic incorporation of artificial nucleotides and subsequent labeling using click reactions[19,20] have performed well for long DNA, so it was straightforward for us to adapt them to mRNA.

Incorporation of EUTP during IVT

Our aim was to find a simple method to introduce modifications at multiple sequence positions of mRNA, which would first of all enable tracking of mRNA within the cell. Alkyne functional groups were introduced by incorporation of 5-ethynyl-uridine-trisphosphate (EUTP) using T7 RNA polymerase during in vitro transcription, Figure 1 A. By replacing all natural uridine triphosphates with EUTP an mRNA was produced containing 220 alkynes. It was observed that the T7 RNA polymerase incorporates alkyne-modified EUTP almost as well as natural UTP, as similar mRNA yields were obtained for both reactions (< 10% decrease relative to the natural nucleotides mix). In a first experiment TAMRA azide was chosen to serve as reporter label for analysis of the click labeling reaction's efficiency. Assisted by the THPTA ligand, the heterogeneous catalyst was pre-activated and in the presence of an excess of the dye azide the mixture was incubated at 45 °C for 30 minutes. After reaction completion, the dye-labeled mRNA was purified using a spin-column method and then analyzed by agarose gel electrophoresis. Fluorescence of TAMRA labeled mRNA was clearly recognized in direct comparison with the unclicked alkyne or unmodified mRNA, Figure 2 B and Figure S1 in the Supporting Information. Due to the modularity of the click reaction, preparation of several different dye-labeled mRNAs from the same transcription is possible. To allow optimal tracking of mRNA inside of cells some of the alkynylated mRNA was reacted with Eterneon Red azide. Its red fluorescence was ideal for simulta-
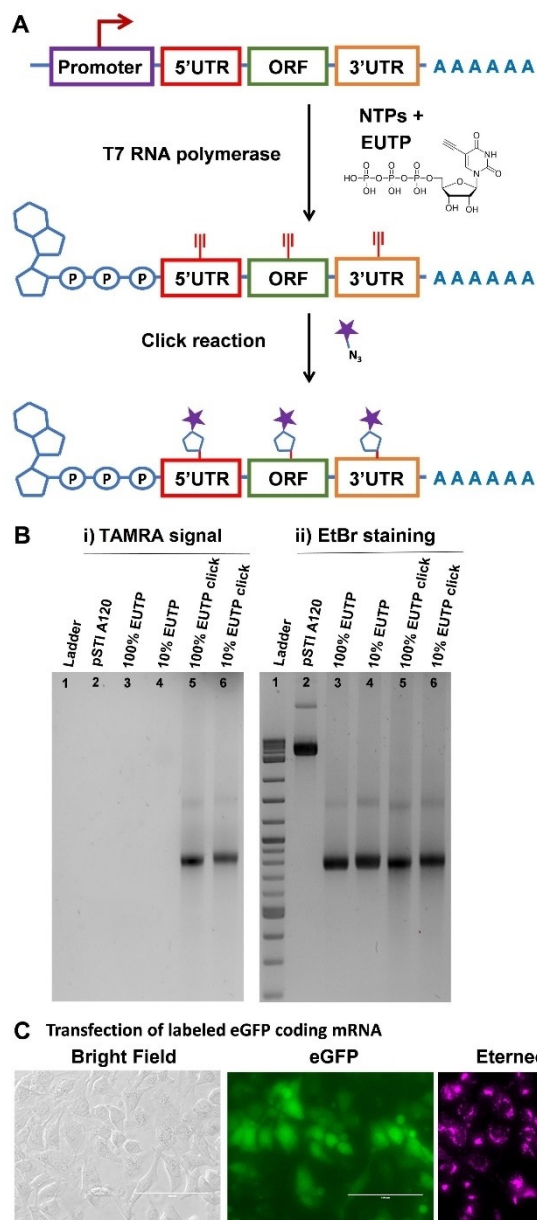
Figure 1. EUTP incorporation. A) Schematic representation of the EUTP incorporation into IVT mRNA. B) Unstained (i) and stained (ii) agarose gel containing linearized plasmid DNA template (lane 2), and its transcription product (lanes 3–6). The alkyne-containing mRNA (lanes 5–6) was click-labeled (lanes 5–6) with TAMRA-PEG3-azide, thus fluorescence without staining of the gel (left, lanes 5–6 B) proves successful click labeling. C) Microscopy images of HeLa cells transfected with eGFP mRNA labeled with Eterneon Red fluorescent dye in the 5′-UTR, ORF and 3′-UTR. Left to right: bright field, green fluorescence, and Cy5 channel. Scale bar: 100 μm.

neous detection of eGFP fluorescence and dye-labeled mRNA in HeLa cell culture experiments.

To stress and observe the limit of the labeling technology, at first an mRNA containing complete replacement of uridine by EUTP was studied. After 24 h of incubation, bright green fluorescence of eGFP was observed and no visible changes in cell morphology were detected (Figure S3). This is a good indication that the alkyne functional group at the nucleobase is well tolerated, even if it is in the coding region of the mRNA. We

ChemBioChem

Full Papers
doi.org/10.1002/cbic.201900718

Chemistry
Europe
European Chemical
Societies Publishing

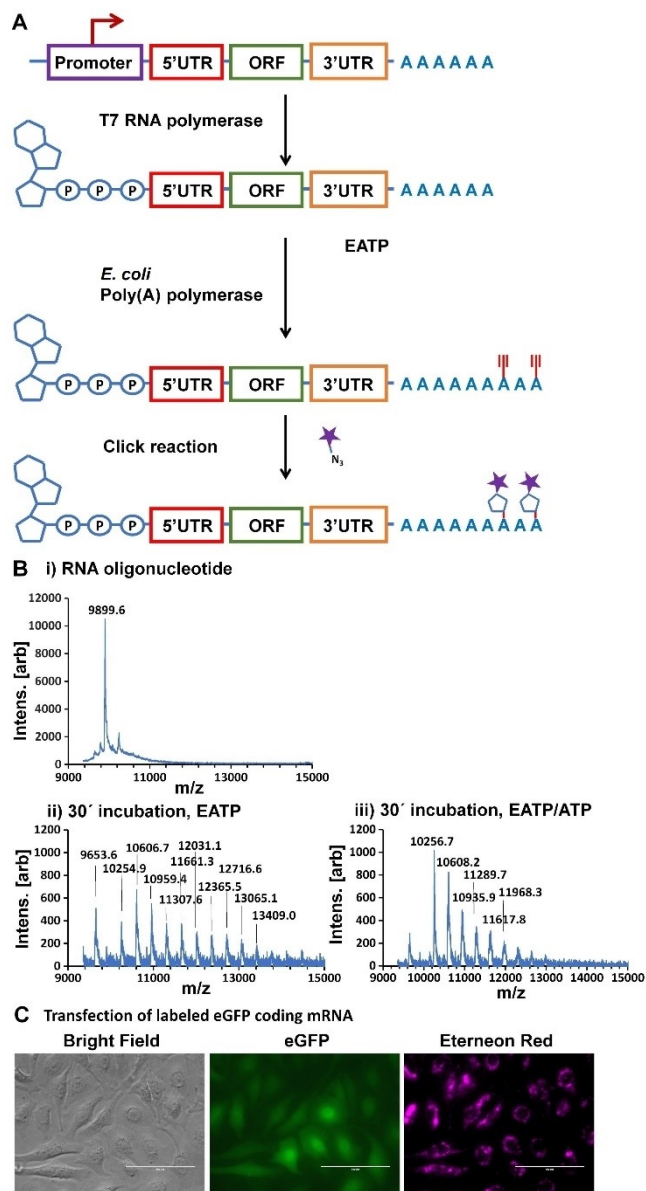**Figure 2.** EATP labeling of the poly(A) tail. A) Schematic representation of EATP incorporation into the poly(A) tail of IVT mRNA. B) MALDI-TOF results from *E. coli* poly(A) polymerase reactions after 30 min incubation using i) no nucleotide, ii) EATP and iii) EATP/ATP. C) Microscopy images of HeLa cells transfected with eGFP mRNA labeled with Eterneon Red fluorescent dye in poly(A) tail. Left to right: bright field, green fluorescence, and Cy5 channel. Scale bar: 100 μm.

unmodified mRNA expression in transfected cells, but it was still possible to visualize the presence of the mRNA due to its fluorescent dye labeling, Figure 1 C.

To prove that the observed eGFP fluorescence was not due to small amounts of unmodified mRNA (no alkyne, no label) present in our mRNA preparation with 10 % of EUTP replacement, a transfection experiment was designed where the amount of unlabeled mRNA was gradually decreased from 100 % (500 ng) to 15 % (75 ng; Table S2). It was observed that upon reaching 30 % (150 ng) of the unmodified mRNA amount no eGFP expression was detected using our microscopy setup (Figure S4). Considering the 220 possible EUTP incorporation positions and an enzymatic acceptance rate almost like a natural UTP (as described previously),[23] the probability to produce unmodified mRNA using 10 % EUTP and 90 % UTP during IVT is very low. Even assuming lower enzymatic acceptance rates, for example, 99 % UTP incorporation at each position, only 11 % of the total mRNAs would result label-free, which is proved to be insufficient for detect eGFP production in our set-up.

Modification of the poly(A) tail by EATP incorporation

Having been able to introduce alkyne modifications within the mRNA, we developed a second labeling method, which restricts labeling to the poly(A) tail of the mRNA. The concept here is based on incorporation of alkyne groups in the poly(A) tail with 7-ethynyl-7-deazaadenosine 5′-triphosphate (EATP) via a poly(A) polymerase reaction (Figure 2 A). To determine appropriate enzymatic labeling conditions, a short RNA served as a model system to allow more simple analytical studies and condition screens. A chemically synthetized 31-mer RNA was used to either incorporate natural ATP or EATP or a mixture of both by poly(A) polymerase from *Escherichia coli*. The resulting product was then analyzed by denaturing polyacrylamide gel electrophoresis and MALDI-TOF. When only EATP was used, the RNA was extended by a few nucleotides compared with about 100 nucleotides extension when the natural ATP was present (Figure S2). By MALDI-TOF the typical spectra of EATP incorporation is observed with a mass difference of about 350 Da between adjacent peaks (e.g., $10\,606.7-10\,254.9 = 351.8$), which corresponds well to the expected mass difference of 352.2 Da ($C_{13}H_{13}N_4O_6P$) upon EATP incorporation and release of pyrophosphate Figure 2 B ii. For incorporation of ATP/EATP mixtures the mass spectrum gets more complex. An extension by a few nucleotides was also found for an equimolar mixture of ATP and EATP, mass differences of 330 ($C_{10}H_{12}N_5O_6P$) and 350 Da can be observed, depending on incorporation of ATP or EATP, respectively Figure 2 B iii. Also, as the incorporation length increases, peak mixtures and broadening of the peaks are visible, as the probability for mixed (labeled and unlabeled incorporation) increases. MALDI-TOF measurements proved successful incorporation of the EATP nucleotide even when competing with the natural ATP for enzymatic incorporation. In the next step, the method was adapted for optimal EATP labeling in IVT mRNA. After EATP incorporation, the mRNA was labeled with Eterneon Red azide for tracking in HeLa cell culture experiments. The poly(A) labeling with EATP of the mRNA results in

speculate that the codon-codon interaction with the 5-EU is similar to natural modifications such as m5C and m6A in mRNAs.[21,22] Then, the alkyne functional groups were labeled with the red fluorescent dye azide. This time eGFP expression decreased relative to transfected cells with non-labeled control mRNA (Figure S3). This indicated that due to the high labeling density the translation at the ribosomes was reduced. Therefore we decided to partly label the mRNA just replacing 10 % of the uridine by EUTP during in vitro transcription. In this case even after the labeling reaction with the fluorescent dye, the green fluorescence eGFP expression intensity was similar to an

similar eGFP expression as the unmodified mRNA in transfected cells, Figure 2 C. In our case, the poly(A) labeled mRNA contained a 120-mer unmodified poly(A) tail, followed by a short part containing a mixture of labeled-alkyne and non-alkyne nucleotides. This way it is reasonable to speculate that the poly(A) binding proteins have sufficient binding motifs to enhance translation, despite the labels.

3′ END modification by AzddATP

To find methods that would allow introduction of azido modifications in the IVT mRNA several azide-modified nucleotides were studied for incorporation by T7 RNA polymerase and poly(A) polymerases. To our delight yeast poly(A) polymerase was found able to incorporate 3′-azido-2′,3′-dideoxyadenosine (AzddATP) quantitatively at the end of a short RNA model template in just 5–20 min as confirmed by MALDI-TOF measurements. A single new peak at m/z of 10 229.9 was observed for the sample after azide labeling, with the peak of unlabeled RNA oligonucleotide (at an m/z of 9898.9) being absent. A mass shift of about 331 Da was detected which is in accordance with a theoretical calculated mass difference of 338 Da (Az-ddATP added, $C_{10}H_{11}N_8O_4P$) taking the measurement error of 0.1 % ($= 10$ Da) into account (Figure 3 B). The lack of subsequent peaks (even after prolonged incubation, data not shown) proved single addition and termination of elongation by incorporation of 3′-azide-2′,3′-ddATP. The azide labeling conditions were transferred to IVT mRNA (Figure 3 A) and specific labeling under strain promoted azide–alkyne cycloaddition (SPAAC) conditions using the DBCO-sulfo-Cy3 resulted in a singly labeled mRNA.

This mRNA was then used for HeLa cells transfection experiments and similar to the EATP incorporation, the production of the eGFP was similar to an unmodified mRNA (Figure 3 C).

Dual modification of mRNA

With the ability to modify two different regions of the mRNA independently, we developed a method to introduce two different modifications on the same mRNA construct. This would allow in future approaches to label mRNAs with different chemical entities on one mRNA molecule, for example, ligands for targeted delivery of the mRNA with labels for tracking or stabilizing moieties. One approach would be combining our established labeling methods in a sequential manner (first internal labeling and click, followed by poly(A) tail labeling and a second click). In a second and more practical approach we initially incorporated alkyne modified nucleotides during in vitro transcription into the mRNA and in a second step azide modified nucleotides during poly(A)-addition. Combination of strain-promoted azide–alkyne cycloaddition (SPAAC) followed by Cu^I-catalyzed azide–alkyne cycloaddition (CuAAC) enables dual orthogonal labeling after enzymatic incorporation.[24]

Dual labeled mRNA was generated by combining incorporation of EUTP during in vitro transcription followed by site-specific terminal addition of AzddATP Figure 4 A. Such modified mRNA was initially labeled via reacting the modified mRNA
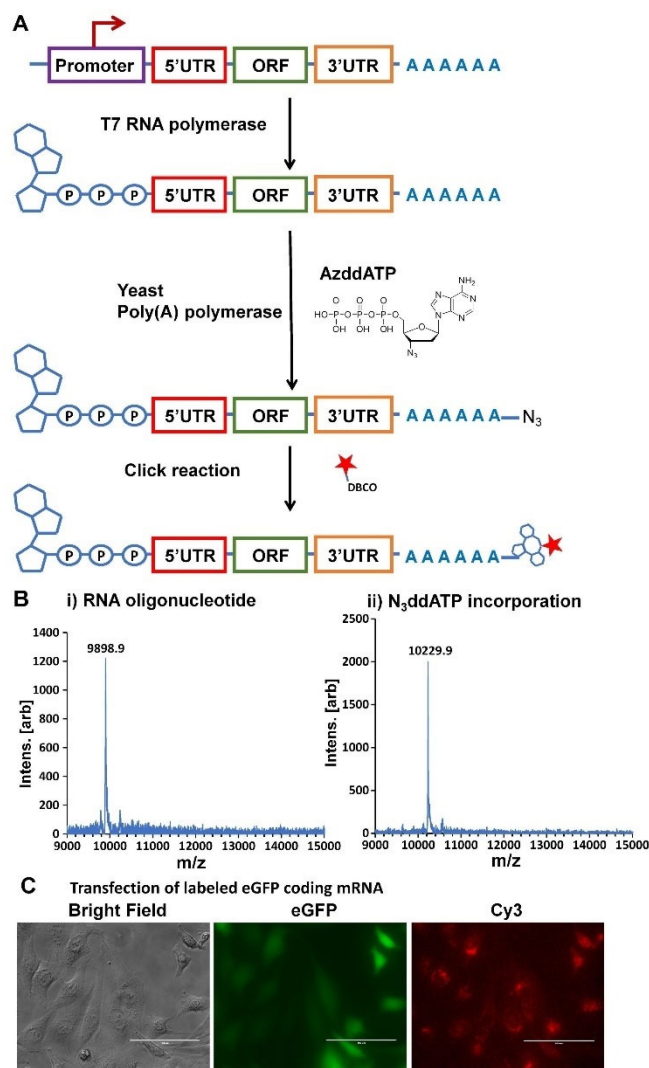

Figure 3. AzddATP labeling of the poly(A) tail. A) Schematic representation of the site-specific modification using AzddATP. B) MALDI-TOF analysis of model RNA oligonucleotide i) before and ii) after yeast poly(A) polymerase reaction using AzddATP. C) Microscopy images of HeLa cells transfected with eGFP mRNA labeled with Cy3 fluorescent dye in the poly(A) tail. Left to right: bright field, green fluorescence, and Cy5 channel. Scale bar: 100 μm.

with DBCO-Cy3 by SPAAC. Because the alkynes inside the mRNA are not able to react with azides without Cu^I catalysis, only specific reaction of the azide with the cyclooctyne modified dye occurred. Subsequently, the alkynes were clicked to Eterneon Red azide using the established CuAAC conditions.

The dual labeled mRNA was then transfected in HeLa cells. In such transfected cells we observed eGFP expression and detected fluorescence of both dyes, the Cy5 analogue (Eterneon Red) and Cy3 as well (Figure 4 B). This demonstrates that dual labeling of mRNA is feasible via click chemistry and that such modified mRNA is still accepted by the translation apparatus, producing a functional protein.
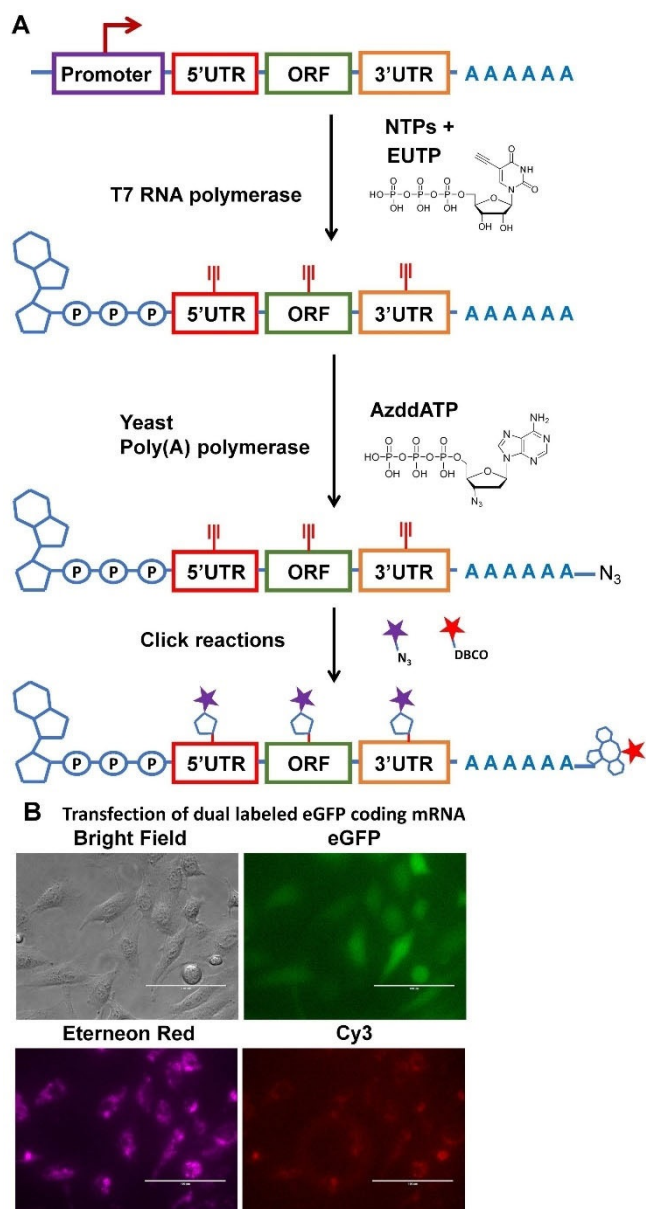
ChemBioChem

Full Papers
doi.org/10.1002/cbic.201900718

Chemistry
Europe
European Chemical
Societies Publishing

Figure 4. Dual-labeled mRNA preparation. A) Schematic representation of the dual-labeled mRNA preparation. B) Microscopy images of HeLa cells transfected with eGFP mRNA labeled with Eterneon Red and Cy3 fluorescent dyes. Left to right: bright field, green fluorescence, Cy5 and Cy3 channel. Scale bar: 100 μm.

Quantification of protein expression by FACS

So far, all experiments were analyzed using standard fluorescence microscopy imaging. To determine eGFP expression levels from modified mRNA, a fluorescence-activated cell sorting (FACS) method was established. HeLa cells were transfected with mRNA and incubated for 24 h, then they were fixed and analyzed by FACS using two fluorescent channels (green for eGFP fluorescence, red for dye-labeled mRNA detection) in correlation to the side scattering (SSC).

When transfected with unmodified eGFP mRNA almost all the cells with a defined gate P1 equal to 96.5% were expressing the eGFP (Figure 5B, red population) as compared with
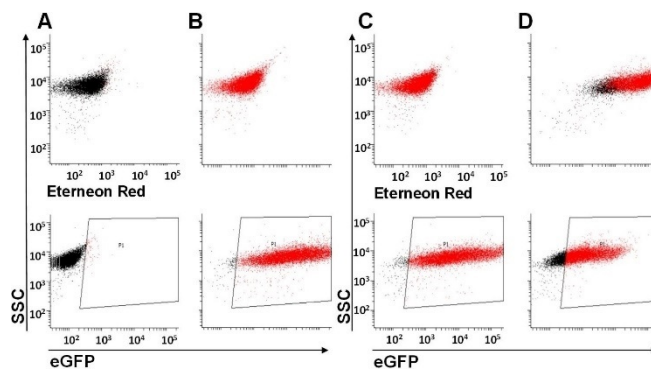

Figure 5. Dot plots of FACS experiments using mRNA-transfected cells after 24 h incubation. eGFP (bottom) and Eterneon Red (top) fluorescence were studied for A) untransfected cells, B) eGFP mRNA, C) alkyne eGFP mRNA, and D) Eterneon Red eGFP mRNA transfected cells.

0.5% of artefactual detection, when no mRNA was transfected (Figure 5A). Practically identical eGFP expression levels were detected for HeLa cells transfected with alkyne modified eGFP mRNA and a P1 value of 96.4% (Figure 5C). This quantification confirms that the alkyne groups do not change the final protein amount, which might indicate that they are well tolerated by the translation apparatus of transfected cells. For HeLa cells which were transfected with Eterneon Red eGFP labeled mRNA, a P1 population of 75% was observed (Figure 5D), meaning that even by attaching a sterically demanding dye molecule to the eGFP mRNA, the ribosomes are still able to translate the mRNA into a functional protein (75% compared with 95% of unmodified or alkyne modified mRNA). Furthermore, because the mRNA was labeled with the Eterneon Red dye it was possible to observe the relative mRNA amount taken up by the cell. Cells that had taken up more mRNA and thus contained more red fluorescent dye, were also expressing higher eGFP levels (Figure 5D), demonstrating how tracking of transfected mRNA labeled with fluorescent dyes can provide important information.

Conclusions

In summary, we have developed several methods to produce chemically modified mRNA in a modular manner by combining enzymatic incorporation of alkyne and azide modified nucleotides with subsequent labeling via click chemistry. Nucleotide analogues like 5-ethynyl uridine triphosphate, 7-ethynyl-7-deaza triphosphate or 3′-azido-2′,3′-dideoxy adenosine have been found to be well accepted by T7 RNA or poly(A) polymerases. The resulting mRNAs were readily functionalized here with dye molecules. This allows to modify mRNA with bulky chemical moieties which, when coupled directly to the nucleotide, during enzymatic incorporation would not be accepted. By varying the amount of nucleotide or choosing the right enzyme, it is possible to control the labeling density or even the position. Site selective terminal labeling of mRNA was achieved by incorporation of an azido nucleotide by yeast poly(A) polymerase and subsequent SPAAC.

ChemBioChem

Full Papers
doi.org/10.1002/cbic.201900718

**Chemistry
Europe**
European Chemical
Societies Publishing

Furthermore, by combining enzymatic labeling using T7 RNA polymerase and poly(A) polymerase and subsequent click labeling with SPAAC and CuAAC we were able to produce a dual modified mRNA molecule. The presented labeling technique could give access to mRNA based active drugs that can be chemically engineered to exhibit for example, better stability, targeting, tracking etc. or even combine such functions. It can be speculated that the modification within the sequence, such as the 5-EU, can alter the formation of secondary and tertiary structures that are important translation regulatory factors and in this way modulate the protein production and mRNA half-life.[25] This might have valuable applications for clinical drug development programs.

Transfection experiments in HeLa cells have shown that the ribosomal translation machinery can accept several bulky modifications inside the eGFP mRNA and still produce a functional protein. When quantification of the protein expression level was performed using FACS, it was observed that the expression level for alkyne labeled eGFP mRNA are identical to unmodified eGFP mRNA. Whereas expression of Eterneon Red labeled eGFP mRNA was slightly decreased, confirming the tolerance of the ribosome to the modified mRNA.

Furthermore, it must be pointed out that selecting uridine as a major modification target could have another benefit, beyond good T7 RNA polymerase acceptance during IVT production. We speculate that because uridine-rich sequences are known activators of several human RNA sensors,[26,27] which are involved in innate immune responses,[28] it is possible that by modifying uridine immunogenicity of the mRNA will be reduced.

Our experiments highlight once again the importance of click chemistry, which is not just a reaction, able to connect two molecules, but a sophisticated tool capable of expanding the chemical space of (bio)molecules, thus enabling new applications.

Acknowledgements

Conflict of Interest

S.C., S.S., and T.F. are co-inventors on a patent application related to this work.

[1] S. A. Martin, E. Paoletti, B. Moss, *J. Biol. Chem.* **1975**, *250*, 9322–9329.
[2] J. Stepinski, C. Waddell, R. Stolarski, E. Darzynkiewicz, R. E. Rhoads, *RNA* **2001**, *7*, 1486–1495.
[3] J. Jemielity, T. Fowler, J. Zuberek, J. Stepinski, M. Lewdorowicz, A. Niedzwiecka, R. Stolarski, E. Darzynkiewicz, R. E. Rhoads, *RNA* **2003**, *9*, 1108–1122.
[4] S. Holtkamp, S. Kreiter, A. Selmi, P. Simon, M. Koslowski, C. Huber, O. Türeci, U. Sahin, *Blood* **2006**, *108*, 4009–4018.
[5] J. Ross, T. D. Sullivan, *Blood* **1985**, *66*, 1149–1154.
[6] D. R. Gallie, *Genes Dev.* **1991**, *5*, 2108–2116.
[7] P. A. Krieg, D. A. Melton, *Nucleic Acids Res.* **1984**, *12*, 7057–7070.
[8] U. Sahin, K. Karikó, Ö. Türeci, *Nat. Rev. Drug Discovery* **2014**, *13*, 759–780.
[9] Y. V. Svitkin, Y. M. Cheng, T. Chakraborty, V. Presnyak, M. John, N. Sonenberg, *Nucleic Acids Res.* **2017**, *45*, 6023–6036.
[10] Y. J. Seo, S. Matsuda, F. E. Romesberg, *J. Am. Chem. Soc.* **2009**, *131*, 5046–5047.
[11] Y. J. Seo, D. A. Malyshev, T. Lavergne, P. Ordoukhanian, F. E. Romesberg, *J. Am. Chem. Soc.* **2011**, *133*, 19878–19888.
[12] T. Someya, A. Ando, M. Kimoto, I. Hirao, *Nucleic Acids Res.* **2015**, *43*, 6665–6676.
[13] T. Ishizuka, M. Kimoto, A. Sato, I. Hirao, *Chem. Commun.* **2012**, *48*, 10835–10837.
[14] N. A. Leal, H. Kim, S. Hoshika, M. Kim, M. A. Carrigan, S. A. Benner, *ACS Synth. Biol.* **2015**, *4*, 407–413.
[15] F. Eggert, S. Kath-Schorr, *Chem. Commun.* **2016**, *52*, 7284–7287.
[16] M. B. Walunj, A. A. Tanpure, S. G. Srivatsan, *Nucleic Acid Res.* **2018**, *46*, e65.
[17] D. Adams, A. Gonzalez-Duarte, W. D. O'Riordan, C.-C. Yang, M. Ueda, A. V. Kristen, I. Tournev, H. H. Schmidt, T. Coelho, J. L. Berk, et al., *N. Engl. J. Med.* **2018**, *379*, 11–21.
[18] M. Eisenstein, *Nature* **2019**, *574*, S4–S6.
[19] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.
[20] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599; *Angew. Chem.* **2002**, *114*, 2708–2711.
[21] J. Choi, K. Ieong, H. Demirci, J. Chen, A. Petrov, D. Dominissini, G. Rechavi, S. M. Soltis, J. D. Puglisi, M. Park, *Nat. Struct. Mol. Biol.* **2016**, *23*, 110–115.
[22] T. P. Hoernes, N. Clementi, K. Faserl, H. Glasner, M. D. Erlacher, K. Breuker, H. Lindner, H. Alexander, *Nucleic Acids Res.* **2016**, *44*, 852–862.
[23] N. Milisavljevič, P. Perlíková, R. Pohl, M. Hocek, *Org. Biomol. Chem.* **2018**, *16*, 5800–5807.
[24] R. R. Ramsubhag, G. B. Dudley, *Org. Biomol. Chem.* **2016**, *14*, 5028–5031.
[25] R. Y. Puah, H. Jia, M. Maraswami, D.-F. Kaixin Toh, R. Ero, L. Yang, K. M. Patil, A. A. Lerk Ong, M. S. Krishna, R. Sun, et al., *Biochemistry* **2018**, *57*, 149–159.
[26] S. S. Diebold, C. Massacrier, S. Akira, C. Paturel, Y. Morel, C. Reis e Sousa, *Eur. J. Immunol.* **2006**, *36*, 3256–3267.
[27] F. Heil, H. Hemmi, H. Hochrein, F. Ampenberger, C. Kirschning, S. Akira, G. Lipford, H. Wagner, S. Bauer, *Science* **2004**, *303*, 1526–1529.
[28] T. Saito, D. M. Owen, F. Jiang, J. Marcotrigiano, M. Gale, *Nature* **2008**, *454*, 523–527.

CHEMBIOCHEM

Supporting Information

Chemoenzymatic Preparation of Functional Click-Labeled Messenger RNA

Stefano Croce,[a, b] Sascha Serdjukow,[a] Thomas Carell,[b] and Thomas Frischmuth*[a]

Supporting Information

1. Chemicals, Materials and Methods

Materials and Chemicals

Chemicals and salts for buffer preparation (e.g. TRIS, etc.) were ordered from Carl Roth, Sigma Aldrich or VWR.

Anti-reverse cap analogue (ARCA), pseudouridine 5'-triphosphate (ΨUTP), EUTP, EATP, AzddATP and all unmodified nucleotides were provided by baseclick GmbH.

T7 RNA polymerase, yeast poly(A) polymerase and DNase I were purchased from Thermo Fischer Scientific, *E. coli* poly(A) polymerase and BspQI were purchased from NEB.

The fluorescent dye azides were provided by baseclick GmbH, including Eterneon Red azide, FAM azide and Tamra azide. Cy3 DBCO was purchased from Jena Bioscence.

RNase free water (Ultra pure distilled water, DNase & RNase free) was purchased from Invitrogen.

Qiaquick PCR purification and plasmid plus midi purification kits were ordered from Qiagen.

Methods

For **agarose gel electrophoresis** of plasmid, template DNA and RNA transcripts (4.7 kbp and 1200 mer respectively) 2 % agarose gels (10 x 15 cm) were prepared in 1x TAE buffer (20 mM TRIS, 10 mM acetic acid, 0.5 mM EDTA, pH 8.5). 1.5 g agarose (CARL ROTH GmbH, Roti®agarose) was dissolved in 75 g 1x TAE buffer in a 250 mL Erlenmeyer flask by alternating microwave heating and manual mixing. Loss of water during heating was controlled by weighing and replaced if necessary with dH_2O. The warm mixture was poured into a planned gel cast equipped with either a 15 or 20 sample comb. Samples were prepared with 20 % purple loading dye (NEB), 2-log DNA ladder (0.1-10 kbp, NEB, N3200) was prepared accordingly, usually 0.5 µL marker were used in 5 µL loading volume. Gels were run in 1x TAE buffer applying constant power (10 W, max. 500 V, max. 100 mA) for 60 min. Then, gels were incubated in a freshly prepared 1:10 000 ethidium bromide dilution for 15 min and then destained in dH_2O for 15 min. For visualization a Gel Doc EZ Imager (BIO RAD) was used.

For **polyacrylamide gel electrophoresis** of RNA (31 mer) 7 % polyacrylamide gels (8.6 x 6.7 cm) were prepared in 1x TBE buffer (0.1 M TRIS, 0.1 M boric acid, 2 mM EDTA, pH 8.3) and 7 M urea (Sigma Aldrich). 3.7 mL of polyacrylamide 30 % (Sigma Aldrich) and 6.7 g of urea were dissolved in a final volume of 16 mL of TBE 1x, 10 µL of tetramethylethyleneamine (TEMED Sigma Aldrich) and 100 µL of ammonium persulfate (APS) 10 % (VWR) freshly prepared were added for polymerization. The mixture was poured into a gel cast (BIO RAD) equipped with a 10 sample comb. Samples were prepared with 20 % Purple loading dye (NEB), low molecular weight ladder (25-766 bp, NEB, N3234) was prepared accordingly. The samples were heated 5 min at 90 °C before loading. Gels were run in TBE buffer 1x at 100 V (constant) for 60 min. Then, gels were incubated in a freshly prepared 1:10 000 ethidium bromide dilution for 15 min and then destained in dH_2O for 15 min. For visualization a Gel Doc EZ Imager (BIO RAD) was used.

Purifications of plasmid DNA was done using the plasmid plus midi kit according to manufacturer's instructions. **Purification of IVT mRNA** was done using the PCR purification kit following manufacturer's recommendations for PCR fragments with one exception: RNase free water was applied for elution.

HeLA cells were **cultured** at 37 °C, 5 % CO_2 and 95 % humidity in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10 % heat-inactivated fetal bovine serum (FBS, Gibco), 2 mM L-glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin, Sigma Aldrich. Cell experiments were performed in a µ-Slide plates (Ibidi®).

Transfection of mRNA was done using the mRNA transfection kit from baseclick GmbH following manufacturer's instructions, 0.5 µg of mRNA was applied to 25,000 HeLa cells (CLS GmbH) reaching confluence. The cells were incubated at 37 °C for 24 hours before analysis under the fluorescent microscope, GFP filter: (470/22 excitation; 510/42 emission), Cy5 filter (628/40 excitation; 692/40 emission) and Cy3 filter (531/40 nm excitation; 593/40 nm emission) were used for detection.

Matrix-assisted laser desorption/ionization-time-of-light (**MALDI-TOF**) mass spectra were recorded on a BRUKER AUTOFLEX II. For MALDI-TOF measurements analysis of short oligonucleotides the samples were dialyzed on a 0.025 µm VSWP filter (Merk Millipore) against ddH_2O for 2 h, then co-crystallized in 3-hydroxypicolinic matrix (HPA: 25 mg 3-hydroxypicolinic acid, 5 mg ammonium citrate, 5 µL 15-crown-5 in 0.5 mL H_2O/MeCN = 1:1).

Microscopy of mRNA transfected cells. HeLa cells were imaged on a µ-slide from Ibidi® using EVOS FL cell imaging system equipped with the GFP EVOS light cube (470/22 nm excitation; 510/42 nm emission), RFP EVOS light cube (531/40 nm excitation; 593/40 nm emission) and Cy5 EVOS light cube (628/40 nm excitation; 692/40 nm emission).

FACS. HeLa cells flow cytometer data were obtained using the BD FACS Canto II equipped with air cooled green (488 nm solid state; 20 mW laser output) and red (633 nm HeNe; 17 W output) lasers, and the data were analyzed with the FACS Diva software. The fluorescent detectors/filters relevant for this study were FITC (530 +/- 30 nm) and APC (660 +/- 30 nm), respectively. The cells were detached from the flask immediately before analysis and the samples were kept in buffer during acquisition, with a flow rate of 10 µL/min. FSC and SSC were used for detection of forward light scatter parameter and side light scatter parameter. The machine was cleaned before and after each measurement. The photomultiplier tube (PMT) voltage for each parameter was adjusted in order to have the cells displayed in the center of the investigating plot.

2. DNA Template Preparation

The plasmid containing the eGFP gene mRNA under control of a T7 promotor was a generous gift from the research group of Prof. Rädler (LMU München). About 100 ng plasmid were transformed into 5-alpha *Escherichia coli* cells (NEB) according to manufacturer's recommendations and then used to inoculate 20 mL LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) containing 50 µg/mL kanamycin as a selection marker. This was incubated overnight at 37 °C, 200 rpm before the cells were harvested by centrifugation (13,000 rpm 15 min, 4 °C). The plasmid was isolated using the plasmid plus midi purification kit (QIAGEN) according to manufacturer's instructions. Usually this yielded 50-60 µg of plasmid DNA, which was linearized using BspQ1 restriction enzyme (NEB) according to manufacturer's recommendations. This results in a linear DNA template that ends directly after a 120 mer poly A tail coding sequence. Purity of the DNA template was assessed by nanophotometer measurement (A_{260}/A_{280} ratio) and agarose gel electrophoresis.

3. In vitro Transcription Reactions

In a 50 µL final reaction volume 20 units of T7 RNA polymerase (THERMO FISHER), 1 µg of template DNA and several nucleotides (Supporting Table 1) were combined in transcription buffer (40 mM Tris-HCl, pH 7.9, 6 mM $MgCl_2$, 4 mM spermidine, 10 mM DTT).

In order to generate differently labeled mRNA molecules, different mixtures of nucleotides were generated and used for the mRNA synthesis as listed below as final concentrations (Supporting Table 1). One mixture was used to produce unmodified mRNA (mixture A), one was used for the EUTP modified mRNA (mixture B) and the other one for the dye labeled mRNA (mixture C).

Supporting Table 1. Final concentrations of nucleotide mixtures used for the production of (modified) mRNAs.

| Mixture	ARCA c [mM]	ATP c [mM]	CTP c [mM]	GTP c [mM]	UTP c [mM]	ψUTP c [mM]	EUTP c [mM]
A	4.00	1.50	1.25	1.00	1.25	1.25	-
B	4.00	1.50	1.25	1.00	-	1.25	1.25
C	4.00	1.50	1.25	1.00	0.63	0.63	0.63

Transcription reactions were incubated for 2 hours at 37 °C and then 2 units of DNAse I (THERMO FISHER) were added and incubated for 15 minutes at 37 °C to digest the DNA template and make its removal easier. The mRNA was purified by a spin column method according to manufacturer's instruction (QIAGEN). When click labeling was performed (e.g. for nucleotide mixture C), 2 µg of purified alkyne mRNA, 1 nmol e.g. Eterneon Red 645 Azide (baseclick GmbH), a single reactor pellet and 0.7 µL 10x Activator² (baseclick GmbH, Oligo² Click Kit) were combined in a total reaction volume of 7 µL. The reaction mixture was incubated at 45 °C for 30 min and then cleaned using a spin column method according to manufacturer's instruction (QIAGEN).

4. Gel Analysis of Plasmid and Modified mRNA

The plasmid and its linearized form were prepared according to section 2, the mRNA was prepared according to section 3 with minor exceptions and then analyzed by agarose gel electrophoresis. The nucleotide mixture used for these *in vitro* transcription experiments was 1.25 mM EUTP/GTP/CTP and 1.5 mM ATP and click reactions were performed using either FAM azide or TAMRA azide (baseclick GmbH). In order to find the best incubation duration for the CuAAC reaction conditions a time course with 5, 10, 20, 30 and 45 minutes at 45 °C was done and directly detected *via* the TAMRA label fluorescence (Supporting Figure 1 lane 7-11). Unlabeled oligonucleotides were detected after ethidium bromide staining (Supporting Figure 1, lane 1-4, 6).

It was found that after 45 min TAMRA labeling (Supporting Figure 1, lane 11) is most efficient, but a decrease in RNA amount (Supporting Figure 1, lane 7-11, EtBr) indicated incompatibility for our protocol with higher labeling rates. We established 30 min incubation as a compromise between labeling rate and RNA recovery. FAM-azide labeling (Supporting Figure 1, lane 5) was used as a second proof for successful mRNA labeling. In lane 6 the mRNA was clicked to

a 30 mer oligonucleotide azide (baseclick GmbH), not relevant for the purpose of this manuscript.
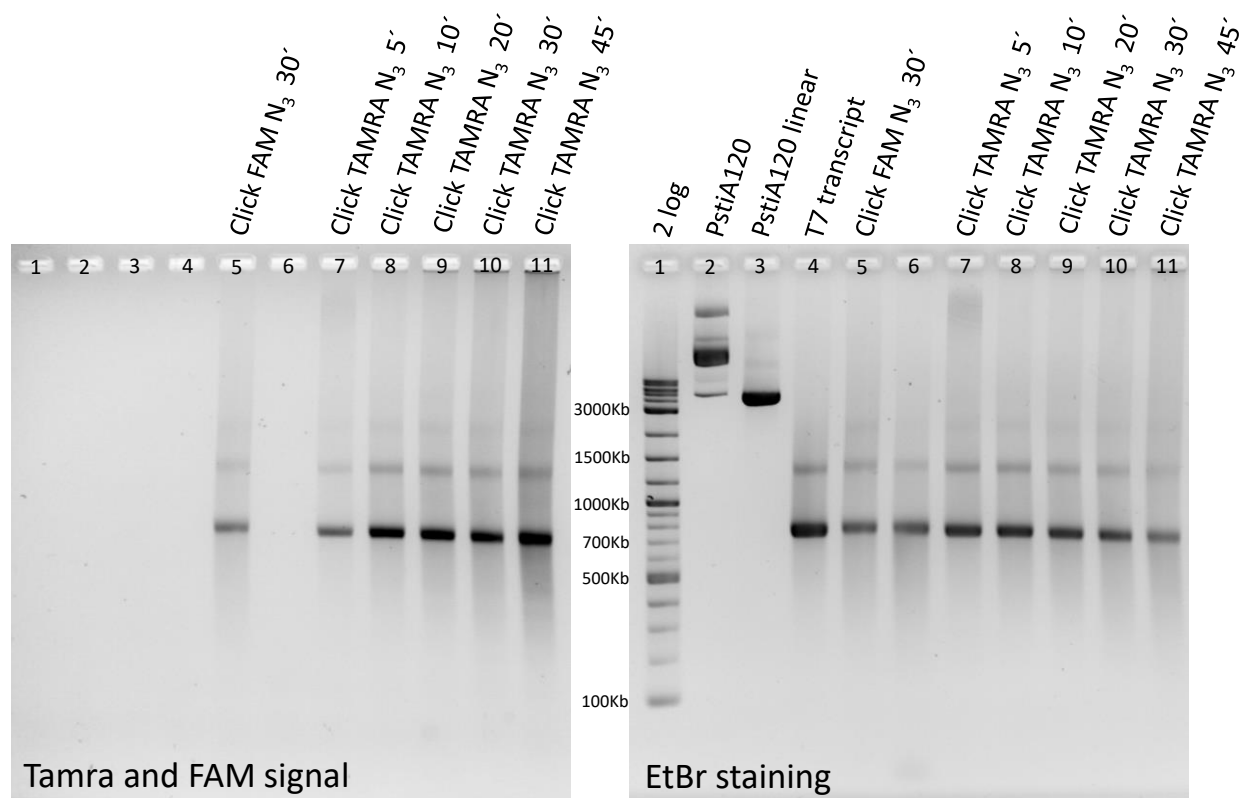
Supporting Figure 1. 2 % agarose gel of plasmid, linearized plasmid and mRNA from IVT and click reactions.

5. General Production of Alkyne-Labeled mRNA using EATP

mRNA was produced according to section 3 and Supporting Table 1 A. Then in a 20 µL reaction volume 5 units of *Escherichia coli* poly(A) polymerase, 6 µg of mRNA and nucleotides (EATP and ATP 2mM final concentration each) were combined in reaction buffer (250 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, pH 7.9). The mixtures were incubated for 30 minutes at 37 °C or for 1 hour at 37 °C. The mRNA was purified by a spin column method according to manufacturer's instruction (Qiagen). The click reaction was done according to section 3.

6. Model System Study for Poly(A) Labeling

In order to prove incorporation of the EATP within the poly(A) tail a short RNA oligonucleotide (31 mer, CUAGUGCAGUACAUGUAAUCGACCAGAUCAA) was used as template for the poly(A) polymerase reaction using final concentrations of:

A) 1 mM ATP

B) 1 mM EATP;

C) 0.5 mM ATP and 0.5 mM EATP.

In a 20 µL reaction volume 5 units of *Escherichia coli* poly(A) polymerase, 2 µg of RNA (31 mer) and nucleotides (final concentration of A-C) were combined in reaction buffer (250 mM

NaCl, 50 mM Tris-HCl, 10 mM MgCl$_2$, pH 7.9). The mixtures were incubated for 30 minutes at 37 °C or for 16 hours at 37 °C.

The results were analyzed by denaturing polyacrylamide gel electrophoresis. Compared to the template RNA oligonucleotide (Supporting Figure 2, Lane 2) a band or smear at higher molecular weight appeared for all samples, which were incubated in the presence of the poly(A) polymerase using different nucleotides and incubation durations (Supporting Figure 2, Lane 3-7). This indicated successful incorporation of ATP or its alkyne analog EATP. Within 30 min incubation the incorporation of ATP (Supporting Figure 2, Lane 3) was more efficient compared to EATP (Supporting Figure 2, Lane 4) or a mixture of EATP and ATP (Supporting Figure 2, Lane 5). By extending the incubation time for the incorporation of EATP to 16 h, the length of the poly-EA-addition was increased (Supporting Figure 2, Lane 6) in comparison to 30 min incubation (Supporting Figure 2, Lane 4). Interestingly, for the nucleotide mixture containing ATP and EATP no change was observed after 16 h (Supporting Figure 2, Lane 7) compared to 30 min.
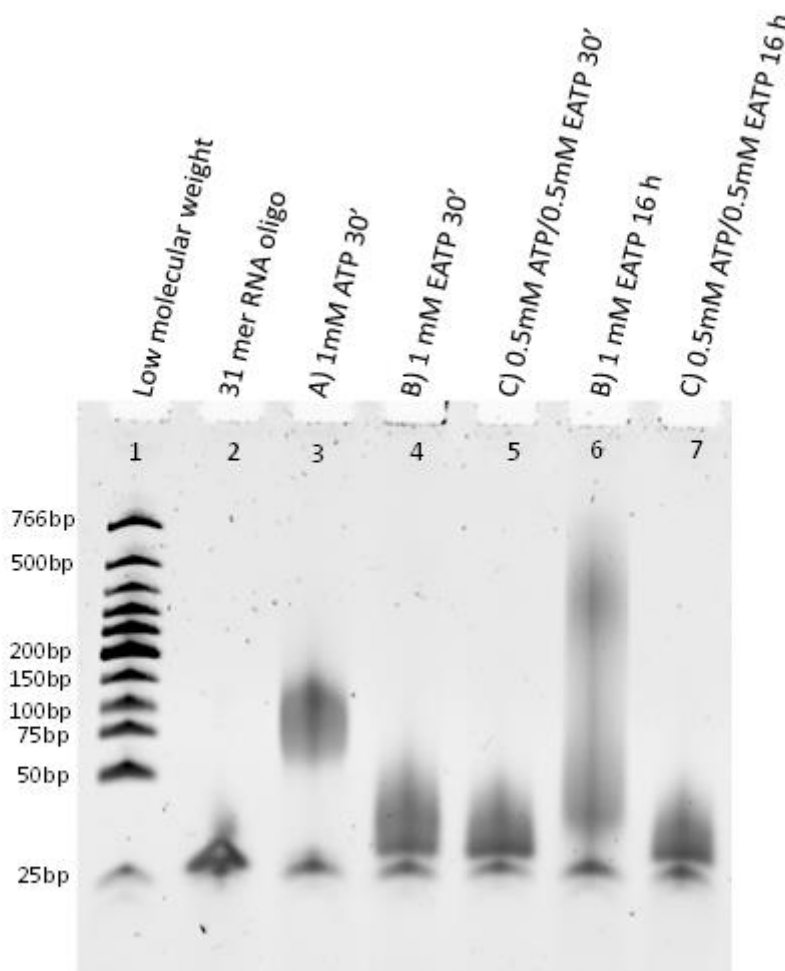
Supporting Figure 2. Stained 7 % denaturing polyacrylamide gel with samples from poly(A) addition experiments using ATP or EATP in different conditions.

7. Site-Specific Azide Labeling in mRNA

mRNA was produced according to section 3 and nucleotides as in Supporting Table 1 A. Then in a 25 µL reaction volume 600 units of *yeast* poly(A) polymerase, 6 µg of mRNA and 3´-Azido-

2′,3′-ddATP (0.5 mM final concentration) were combined in reaction buffer (20 mM Tris-HCl, pH 7.0, 0.6 mM MnCl$_2$, 20 µM EDTA, 200 µM DTT, 10 µg/mL acetylated BSA, 10 % glycerol). The mixture was incubated for 20 minutes at 37 °C. mRNA was purified by a spin column method according to manufacturer's instruction (Qiagen). When click labeling was performed, 5 µg of purified azide mRNA, 2 nmol DBCO-Sulfo-Cy3 (Jena Bioscence), were combined in a total reaction volume of 30 µL. The reaction mixture was incubated at room temperature overnight and then cleaned using a spin column method according to manufacturer's instruction (Qiagen).

## 8. Model System Study for Azide Labeling

In order to prove incorporation of the 3′-Azido-2′,3′-ddATP at the end of the poly(A) tail a short RNA oligonucleotide (31 mer, CUAGUGCAGUACAUGUAAUCGACCAGAUCAA) was used as model template for the poly(A) polymerase reaction using:
In 20 µL reaction volume 20 pmol of RNA oligonucleotide, 500 µM 3′-Azido-2′,3′-ddATP, 600 U of yeast poly(A) polymerase were mixed in buffer (25mM Tris-HCl (pH 7.0), 40 mM KCl, 0.5 mM MnCl$_2$, 50 µM EDTA, 0.5 mM DTT, 0.2 mg/mL BSA, 10 % glycerol) and were incubated for 5 minutes at 37 °C.

9. Sequence of the eGFP DNA Template and Map of the Plasmid

This is the DNA sequence (5' to 3' direction) necessary for the production of the mRNA that was used in this study. Sequence parts are highlighted in various colors: The T7 promotor in yellow, the ribosome binding site in violet, the start codon in green and the stop codon in red.

TAATACGACTCACTATAGGGCGAACTAGTAAGCAAGGAGGCGTGCAGATGGTGAGCAA
GGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGT
AAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAA
GCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTC
GTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGC
AGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTT
CTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACAC
CCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGC
AGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGT
GCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT
GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAG
CGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG
ACGAGCTGTACAAGTCCGGCCGGACTCAGATCTCGAGCTCAAGCTTCGAATTGATCCA
GATCTTAAGTAAGTAAGCTCGAGAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTT
CCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCT
GGATTCTGCCTAATAAAAAACATTTATTTTCATTGCTGCGTCGAGAGCTCGCTTTCTTGC
TGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATA
TTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCTG
CGTCGAGAGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAA
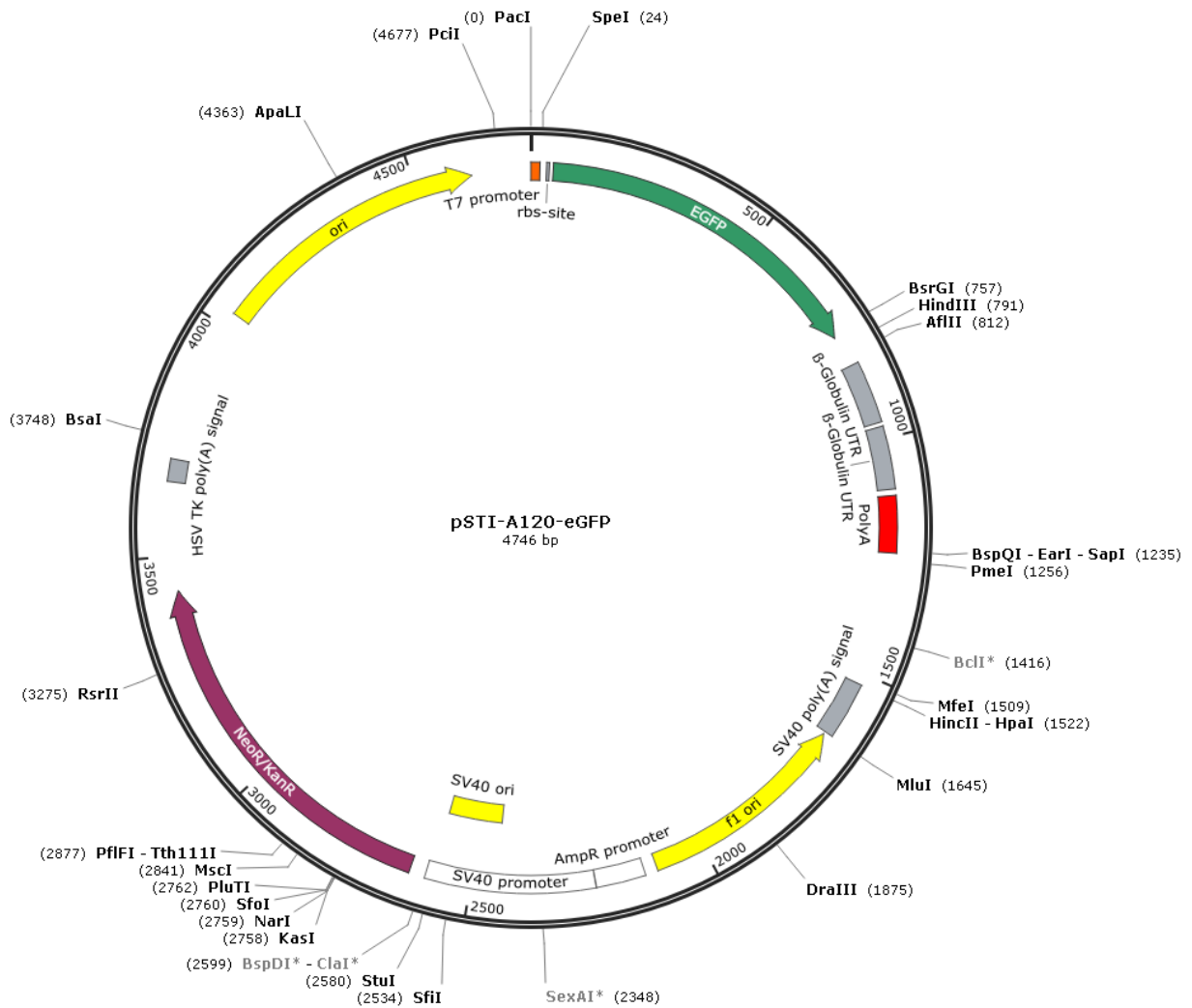
Supporting Figure 3: Map of the plasmid pSTI-A120 used for the production of the eGFP mRNA *via* T7 RNA polymerase reaction. It contains a T7 promotor (orange), a ribosome binding site (RBS, violet), the eGFP coding sequence (EGFP, green), two times the β-globin UTR (grey) and PolyA sequence (red) followed by a BspQ1 restriction site for convenient IVT from linearized plasmids.

10. Additional Cell Culture Experiments

100 % EUTP Replacement

In this experiment the mRNA was prepared according to section 3 and nucleotides as in Supporting Table 1 B. When labeling was required the click reaction was done according to section 3 using Eterneon Red Azide (baseclick). Cell culture, transfection and fluorescence microscopy were done according to general methods section 1.
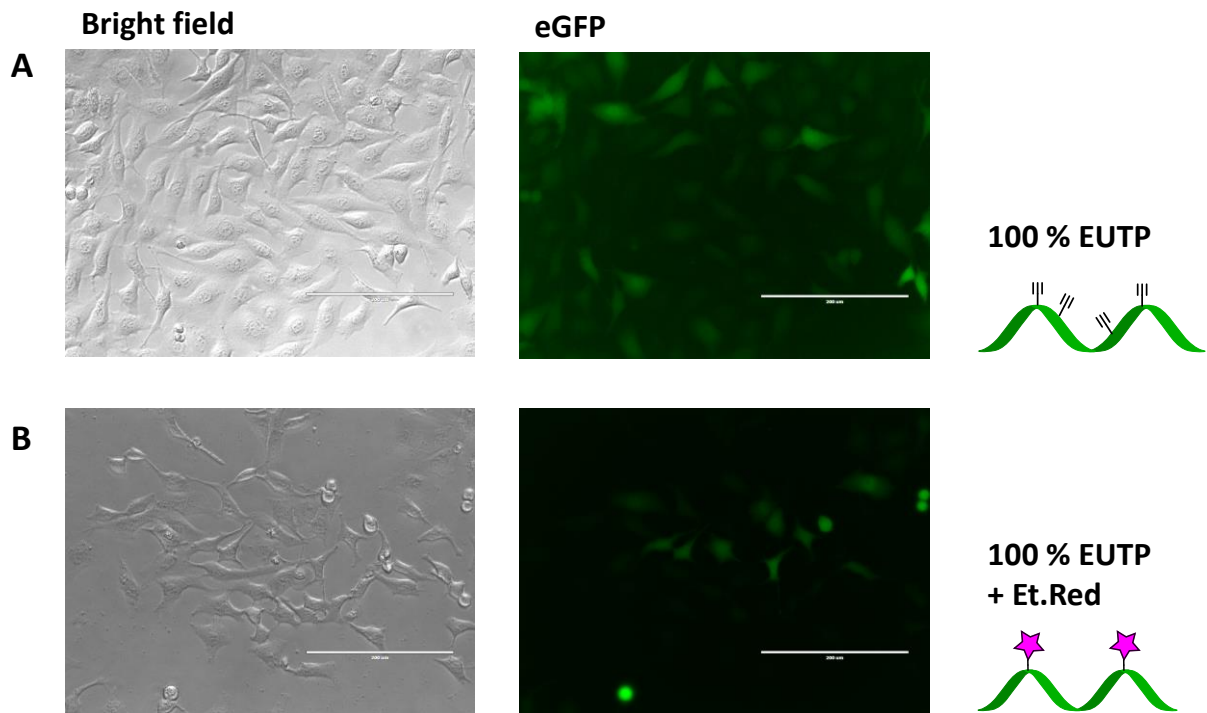
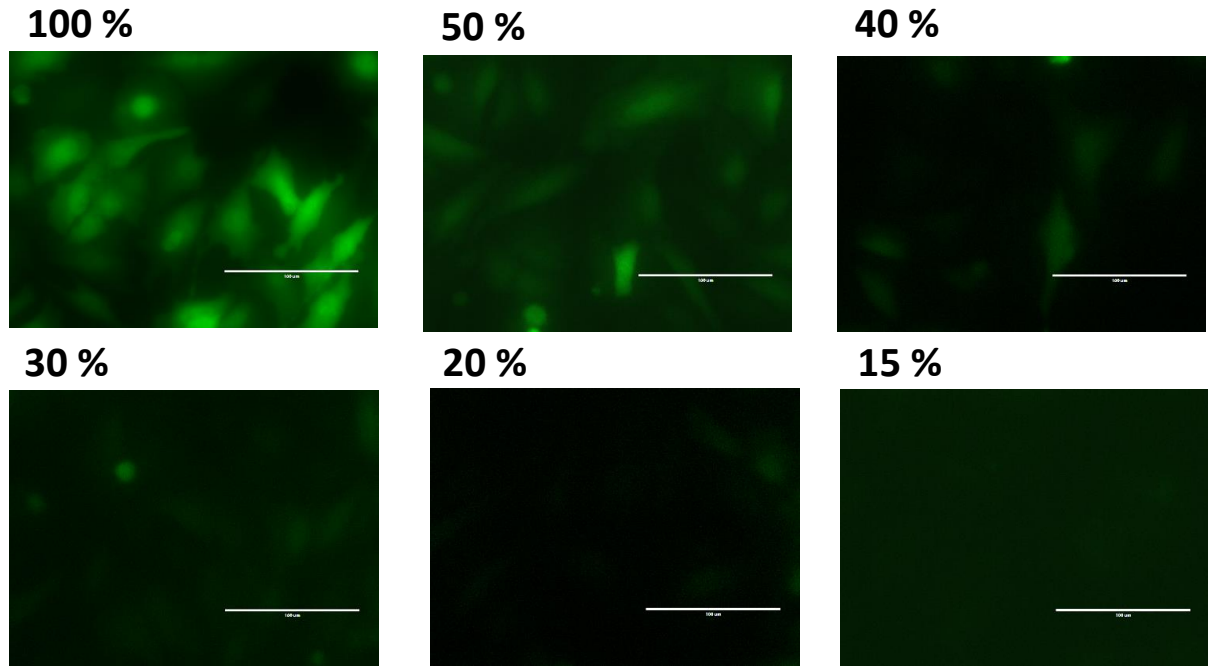Supporting Figure 4: Microscopy images of HeLa cells, which have been transfected with mRNAs containing A) no natural uridine (but 5-ethynyl uridine instead) and B) the mRNA as in A) after click to Eterneon Red-azide. In B) it is visible how by replacing all the natural uridine with an ethynyl derivative and click labelling the translation efficiency has been reduced, but the ribosomes are still able to produce active protein. Scale bar is 100 µm and both pictures were taken with 120 ms exposure time.

Decreasing Amount of mRNA

In order to find out whether the signal of the green fluorescent protein (eGFP) is due to the small fraction of mRNA that did not incorporate the EUTP, transfections using decreasing amounts of unmodified eGFP mRNA were performed. The amount of mRNA coding for the eGFP was gradually decreased according to Supporting Table 2 and the missing material was replaced with an inactive RNA. The unmodified eGFP mRNA for this study was produced according to section 3 and nucleotides as in Supporting Table 1 A.

Supporting Table 2: Decreasing amount of unmodified eGFP mRNA used for transfection in cell culture experiments.

	100 %	50 %	40 %	30 %	20 %	15 %
eGFP mRNA	500 ng	250 ng	200 ng	150 ng	100 ng	75 ng
Inactive RNA	0 ng	250 ng	300 ng	350 ng	400 ng	425 ng

Supporting Figure 5: Microscopy images showing the eGFP signal 24 h after transfection of different mRNA amounts compared to the standard amount (100 %). Scale bar is 100 µm and all pictures were taken with 60 ms exposure time.

As visible from Supporting Figure 5 by decreasing the amount of mRNA, the signal of the eGFP decreases at the same time. Reaching 30 % of the original mRNA amount the signal of the fluorescent protein is hardly visible. In this experiment we successfully proved that the active protein is not produced from the small fraction of mRNA that did not incorporate the EUTP.

11. SI References

[1]   J. Shen, A. Busch, T. S. Day, X. Meng, C. I. Yu, P. Dabrowska-schlepp, B. Fode, H. Niederkrüger, S. Forni, S. Chen, et al., **2016**, 293–303.

7. Super-sensitive multi-fluorophore RNA-FISH using click chemistry

7.1. Prologue

In situ hybridization is a cytogenetic technique that allows researchers to visualize and map genetic material in single cells, *e.g.* entire chromosomes, specific gene or portion of it[37,79] in animals and plants. Thus it can be used to diagnose a variety of chromosomal abnormalities (monosomy and trisomy), re-arrangements such as translocations, microdeletions and duplication by clinicians. The technique is based on probe consisting of ssDNA bearing a tag, designed and synthesized in order to hybridize with part of genomic DNA with a high degree of complementarity. Since its first appearance, that involved radioactive probes, (Gall & Pardue, 1969)[80], different variations have been developed with an enormous increase in sensitivity, nowadays using fluorescent tags.

Theoretically, fluorescent *in situ* hybridization (FISH) is a very simple assay, it is based on the ability of probes to hybridized with their complementary sequence of a chromosome previously fixed on a slide for microscopy. When the probes already contain fluorescent tags, after hybridization the position can be detected directly using a fluorescent microscope, Figure 12. When the probes do not contain a fluorescent tag, but haptens or biotin, a following step for fluorescent labelling is required, e.g using an antibody bearing a fluorophore and binding to the hapten.

Since the late 1990 (Femino et.al 1998)[81] FISH is not only restricted to DNA or chromosomal analysis but also to RNA detection such as expression of specific alleles, viral infections or mRNA distribution.
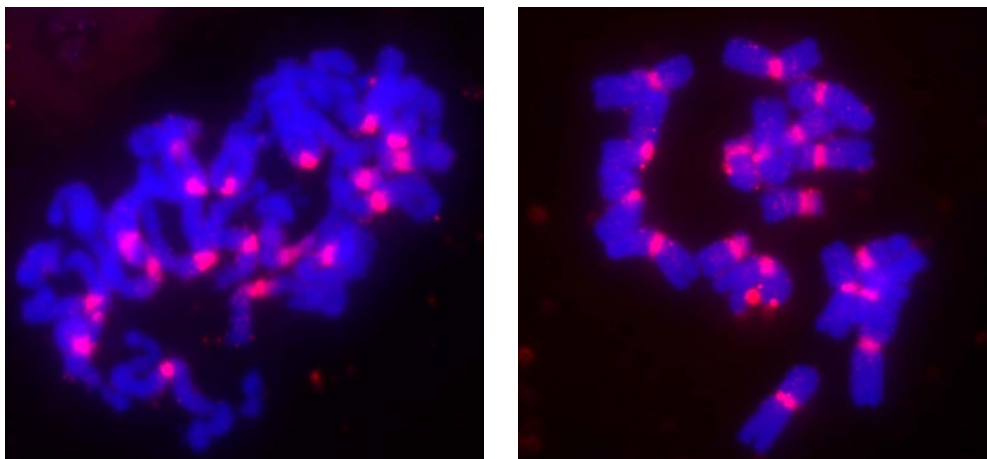
Figure 12: Mitotic chromosomes of rye (*Secale cereal*) hybridized with centromere-specific repeat Bilby fluorescent probes. M.Said *et al.*, The Plant Genome, 2018, 12, 1-13.

The fluorescent probes can be prepared with different methods such as nick translation, random priming, PCR and end labelling. By nick translation is intended a method for the preparation of the probes where deoxyribonuclease I (DNase I) creates "nicks" in the DNA, leaving a 3' hydroxyl group used then by DNA polymerases to introduce fluorescent nucleotides[36]. This result in a mix of labelled dsDNA that needs to be

denatured before the assay takes place. In random priming, small random oligonucleotides, usually between 6 and 9 nt long, are used to anneal to a denatured DNA fragment and then with the usage of the Klenow fragment polymerase, the oligonucleotides are extended, thereby incorporating fluorescent nucleotides, resulting in a uniformly labelled dsDNA. PCR is mostly used for the preparation of labelled versions of chromosomal regions, where the preparation of very long labelled probes is based on the ability of DNA polymerase to introduce labelled nucleotides during the polymerase cycles. End-labelling can be divided in 5' end labelling or 3' end labelling. The first one uses the activity of the T4 polynucleotide kinase to introduce phosphate groups which can then be conjugated to fluorescent labels. The second one uses the ability of the terminal deoxynucleotidyl transferase (TdT) to introduce labelled dideoxynucleotides triphosphates[82,83]. Nowdays also the use of NHS chemistry is one of the most preferred methods for the preparation of fluorescent probes, where an amino group at the 3' end introduced during solid phase synthesis of the oligonucleotide, then reacted with an NHS-ester on a fluorophore.

As obvious consequence, probe-labelling is a critical part of the total assay, that still needs to be optimized and ameliorated. Therefore we propose here a method for the labelling of FISH probes based on click chemistry. Click chemistry has proven to be a very efficient method that allows the production of probes bearing not only one label but several of them. The production of probes with an increased fluorescence can even allow the detection of the target by flow cytometry, as proven by this work where the expression of specific mRNA was detected in a mixed cells population, furthermore we could prove that the so prepared probes are even able to detect virus infections by the detection of viral mRNA, often folded in secondary structures and difficult to be detected by other traditional methods. In particular FACS detection can be of fundamental importance in clinical studies where the detection is still done using fluorescent microscopy by long and tedious work of the clinicians, thus speeding up the analysis and reducing their costs.

7.2. Abstract of the publication "Super-sensitive multi-fluorophore RNA-FISH using click chemistry"

The reliable detection of transcription events via the quantification of the corresponding mRNA is of paramount importance for the diagnostics of infections and diseases. The quantification and localization of the mRNA of a particular gene allows characterizing disease states more directly compared to an analysis on the DNA level. This is particularly needed for the early detection of virus infections as now required for Corona diagnostic. In situ mRNA analysis, however, is a formidable challenge and currently performed with sets of single-fluorophore containing oligonucleotide probes that hybridize to the mRNA at question. Often a large number of probe strands (often >30) are required to get a reliable signal. The more oligonucleotide probes are however used, the higher are potential off-target binding effects that create background noise. Here, we used click chemistry and alkyne modified DNA oligonucleotides for preparing multiple-fluorophore mRNA probes. We found that these multiple-dye probes allow reliable detection and localization of mRNA with only

a very small number (5-10) of probe strands. The new method is so sensitive that it even enabled the *in situ* detection of viral transcripts just 4 hours after infection.

7.3. Author contribution

The Author planned and executed the FACS experiments, prepared the click probes and did the HPLC analisys for characterization. Helped in the manuscript revision.

7.4. Associated Publication

Super-sensitive multi-fluorophore RNA-FISH using click chemistry

Nada Raddaoui,[#] **Stefano Croce,[#]** Florian Geiger, Alexander Borodavka, Leonhard Möckl, Samuele Stazzoni, Bastien Viverge, Christoph Bräuchle, Thomas Frischmuth, Hanna Engelke, and Thomas Carell*,

Supersensitive Multifluorophore RNA-FISH for Early Virus Detection and Flow-FISH by Using Click Chemistry

Nada Raddaoui+,[a] Stefano Croce+,[a, b] Florian Geiger,[a] Alexander Borodavka,[c, d] Leonhard Möckl,[a] Samuele Stazzoni,[a] Bastien Viverge,[a] Christoph Bräuchle,[a] Thomas Frischmuth,[b] Hanna Engelke,[a] and Thomas Carell*[a]

The reliable detection of transcription events through the quantification of the corresponding mRNA is of paramount importance for the diagnostics of infections and diseases. The quantification and localization analysis of the transcripts of a particular gene allows disease states to be characterized more directly compared to an analysis on the transcriptome wide level. This is particularly needed for the early detection of virus infections as now required for emergent viral diseases, e. g. Covid-19. *In situ* mRNA analysis, however, is a formidable challenge and currently performed with sets of single-fluorophore-containing oligonucleotide probes that hybridize to the mRNA in question. Often a large number of probe strands (> 30) are required to get a reliable signal. The more oligonucleotide probes are used, however, the higher the potential off-target binding effects that create background noise. Here, we used click chemistry and alkyne-modified DNA oligonucleotides to prepare multiple-fluorophore-containing probes. We found that these multiple-dye probes allow reliable detection and direct visualization of mRNA with only a very small number (5–10) of probe strands. The new method enabled the *in situ* detection of viral transcripts as early as 4 hours after infection.

Introduction

Gene expression varies significantly between individual cells and it is strongly altered in disease states. Viral infections for example lead to early transcription of virus-specific genes that could be exploited for an early diagnosis and characterization of the infection. In general, basically all malfunctioning processes in cells induce transcriptional changes.[1,2] These go in hand with altered levels of messenger RNAs (mRNAs). In extreme cases disease related mRNA may not be present at all in the normal state. In most cases however, the levels of specific mRNA will be changed, which requires reliable methods to quantify mRNA transcripts. Detection and quantification of a specific mRNA is thus highly desirable from a diagnostic point of view. Particularly informative are methods that allow the quantification of mRNA levels with spatial resolution. Currently, however, intracellular localization and quantification of mRNA faces a number of challenges that hinder routine use. The most common way to detect mRNA (or other RNAs) in cells is fluorescence *in situ* hybridization (RNA-FISH).[3] The method reveals localization patterns of individual RNA transcripts in cells or tissues and as such, RNA-FISH is the method of choice for quantitative single-cell transcriptomic studies.[4–8] The currently available technology behind RNA-FISH technologies is based on multiple (up to 50) individual anti-sense single-stranded (ss) DNA probes, which are approximately 22 nucleotides long. Each probe oligonucleotide carries a single fluorophore, which is typically introduced as its activated NHS ester to an amino group present at the 3'-end of the probe.[4,5] The pooled fluorescent ssDNA probes are finally added to fixed and permeabilized cells for hybridization with the target RNA. The large number of probe strands in such experiments is needed to create a sufficiently strong fluorescence signal. However, generally the larger the number of probe oligonucleotides that are used, the larger is often also off-target staining, which obscures the signal-to-noise ratio. A solution to the problem is deconvolution software that is able to increase the specific signal.[9] From a chemical point of view reduction of the number of probe strands is desirable and this has led to efforts to modify the probe oligonucleotides with, for example, LNA to increase binding. Importantly, mRNA analysis based on flow-cytometry is so far very challenging with contemporary RNA-FISH.

Here we report a small FISH-probe set for mRNA, where every probe contains three fluorophores instead of just one. These multichromophore probes were conveniently prepared

[a] N. Raddaoui,+ S. Croce,+ F. Geiger, L. Möckl, S. Stazzoni, B. Viverge,
 C. Bräuchle, H. Engelke, T. Carell
 Department of Chemistry
 Ludwig-Maximilians-Universität München
 Butenandtstraße 5–13, 81377 Munich (Germany)
 E-mail: Thomas.carell@lmu.de
[b] S. Croce,+ T. Frischmuth
 Baseclick GmbH
 Floriansbogen 2–4, 82061 Neuried (München) (Germany)
[c] A. Borodavka
 Astbury Centre for Structural Molecular Biology
 School of Molecular and Cellular Biology, University of Leeds
 Leeds, LS2 9JT (UK)
[d] A. Borodavka
 Department of Biochemistry
 University of Cambridge,
 Cambridge, CB2 1QW (UK)
[+] These authors contributed equally to this manuscript
 Supporting information for this article is available on the WWW under
 https://doi.org/10.1002/cbic.202000081

ChemBioChem

Full Papers
doi.org/10.1002/cbic.202000081

Chemistry Europe
European Chemical
Societies Publishing

using the Cu^I-catalyzed azide-alkyne click reaction.[10–19] In order to avoid stacking of the fluorophores on top of each other, which might induce self-quenching, we chose a fluorophores with two additional sulfonate groups, which provide two negative charges per fluorophore. This is supposed to minimize the interaction with the fluorophores with each other and with the negatively charged DNA. Indeed, with this design a small number of probe strands (5–10) was found to be sufficient for the visualization of RNA transcripts. The new probes design allowed not only transcript quantification and localization by microscopy, but it also enabled transcript analysis using flow-cytometry.

Results and Discussion

The new procedure based on click chemistry is illustrated in Figure 1. As a test-system for the mRNA-FISH we used a HEK293T cell line transfected with a plasmid containing the gene coding for the enhanced green fluorescent protein (eGFP). We synthesized ten DNA probe oligonucleotides targeting specific areas of the eGFP-mRNA transcript, with each one containing 22 nucleotides (Figure S1 in the Supporting Information). Into each of the ten DNA probe strands we inserted three C8-alkyne-dU building blocks at former dT positions (Figure 1 and Table S1) using phosphoramidites that we had developed previously.[20] The ten triple-alkyne DNA probes were finally purified by HPLC and individually subjected to a click reaction with Eterneon-Red 645 azide (cyanine-5 analogue). The so obtained three Eterneon containing DNA probes (10×3) were finally purified by a simple ethanol precipitation. Due to the high efficiency of the click reaction, no further HPLC purification of the probes was required. For the probes prepared for the

virus experiment (*vide infra*) we also confirmed the purity of the obtained probe strand by HPL chromatography (example shown in Figure S2). HPL chromatographic analysis of our 10×3 probe set proved good absorption and fluorescence properties (Figures S3 and S4).

In order to compare the obtained data with the state-of-the-art we performed in parallel studies with a commercially available RNA-FISH probe set. The provider suggested for the requested detection a set of 30 probe oligonucleotides (Table S2) each one carrying one Quasar 670 fluorophore. The purchased oligonucleotide probes are shown in Figure S1.

In order to investigate the ability to detect RNA, we first performed *in vitro* experiments with isolated total-RNA obtained from wild-type HEK293T and HEK293T-eGFP, stably expressing the eGFP gene. The data are depicted in Figure S5, we obtained clearly visible spots with the (10×3) triple modified probes. Importantly, the 10×3 set provided bright spots even without the use of the special deconvolution software. In order to exclude that the high spot density obtained with the new probes is caused by unspecific binding, we performed a negative control with total RNA isolated from HEK293T cells not expressing the eGFP-protein (control probes). Here, as expected far fewer spots were obtained, which rules out this possibility.

After these *in vitro* experiments, we next investigated the properties of the 10×3 probe set in fixed cells (Figure 2). To this end, the HEK293T cells were grown on 8-well μ-Slide (ibidi) and transfected with a plasmid DNA containing a gene coding for eGFP. The cells were fixed and permeabilized using the standard protocols (see Materials in the Supporting Information). We next added the mixture containing our 10×3 probe set and as a positive control, we also performed an experiment with the 30×1 set. Both probe sets were incubated over night at 37 °C. After washing, we analyzed the cells by fluorescence microscopy. The result of the study is shown in Figure 2a. In order to quantify the background fluorescence signal, we calculated the signal intensity obtained after *in situ* hybridization of the probe sets lacking the eGFP-locus. This background signals for the 10×3 and 30×1 experiment were
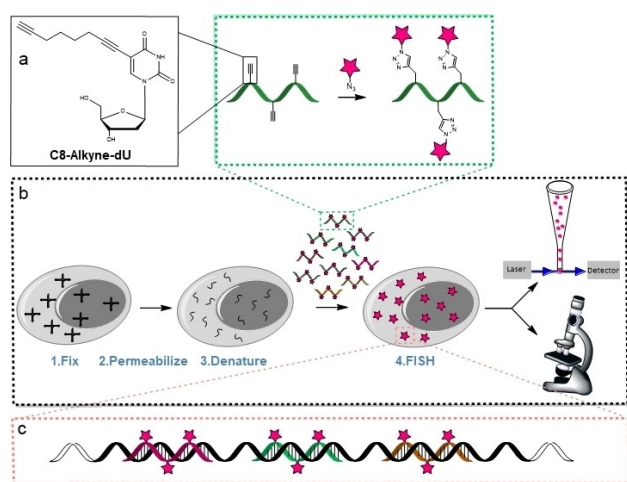
Figure 1. Depiction of RNA-FISH and schematic representation of the probe synthesis. a) Synthetic oligonucleotides with C8-alkyne-dU modifications in various positions were individually conjugated with a fluorescent dye azide. After reaction, the oligonucleotides were mixed to a probe set. b) The probe set was hybridized to the mRNA. After in situ hybridization, the mRNA molecules can be detected by flow-cytometry and/or microscopy. c) Depiction of the probes labeled with click chemistry hybridized to the target mRNA.

Figure 2. mRNA of eGFP-expressing HEK cells labeled with 10×3 and with 30×1 probe sets. a) FISH microscopy images of the 10×3 and 30×1 set of probes (scale bar: 10 μm) b) Comparison of the signal and background intensities. c) Signal and background obtained with CuAAC-labeled probes when using 3 and 10 oligos.

subtracted from 10×3 and 30×1 data sets obtained with the eGFP expressing cells (Figure 2b). As depicted in Figure 2a we saw for the 10×3 probe se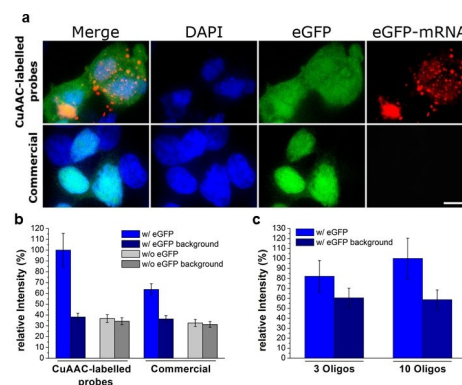t diffraction-limited spots were detected showing clearly the presence of the complementary mRNAs (Figure 2a, red channel). The signal-to-noise ratio was strongly increased. A small caveat is that we noted a slightly increased background signal with the 10×3 probe set (Figure 2b, dark gray bar) likely because the oligonucleotide containing three dyes are more hydrophobic, which may give slightly increased unspecific binding, which increases the background. This effect is obviously overcompensated by the strong increase of the fluorescence signal obtained from specific binding events.

We next experimented with different number of probe strands and dye loading and found that the most reproducible data were indeed obtained with ten oligonucleotide probe strands containing each three fluorophores. In our hands this probe design provided in all investigated cases the best signal-to-noise ratios (Figures 2c and S6).

We next investigated if we could further reduce the background signal when we performed the click reaction after *in situ* hybridization as shown in (Figure S7). For this study, we used a set of 10 probe strands with 3 alkyne units, hybridized them with the cellular mRNA and performed the CuAAC-reaction subsequently *in situ* with a TAMRA azide. While the signal-to-noise ratio indeed improved, we noted that we had to perform extensive washing in order to remove additional dye, which is typically used in large excess. This makes this procedure a little more accurate but cumbersome to perform.

We next investigated if the performance of the 10×3 probe design allows to detect mRNA even using flow-cytometry in a mixed cell population. For the experiment, we mixed HEK293T cells with and without eGFP-expression in a ratio 20:1 (95% HEK293T + 5% HEK293T-eGFP). Then, flow-cytometric measurements of the mixed cell populations were performed at $\lambda_{ex} = 488$ nm/$\lambda_{em} = 530$ nm for the detection of the eGFP protein. The correct ratio of the mixed cell population was nicely reproduced (Figure 3a). We then used the flow-FISH protocol described by Arrigucci et al.,[21] which involves trypsinization and resuspension of cells. The cells in suspension were permeabilized, fixed and subsequently incubated with the probe sets. Again, we performed the study with the new 10×3 set in comparison to a classical 30×1 design. When we measured at $\lambda_{ex} = 633$ nm/$\lambda_{em} = 660$ nm, which is suitable for both the Eterneon-Red 645 azide and the Quasar 670 dye, in the absence of hybridized probes, a single population was observed (Figure S8i). The upper plot of Figure 3b shows the hybridization experiment using the classical probe set at 0.05 ng/μL. Here, only a single population containing both GFP-positive and GFP-negative cells was detected at 660 nm. When the 10×3 probe set was used however at the same concentration (lower plot), the GFP-positive population nicely separated from the GFP-negative cells. When the separated population was gated (in blue), the exact proportion of cells expressing the eGFP was observed for the two different detection wavelengths. The same result with the classical 30×1 probe set was only obtained, when the concentration was
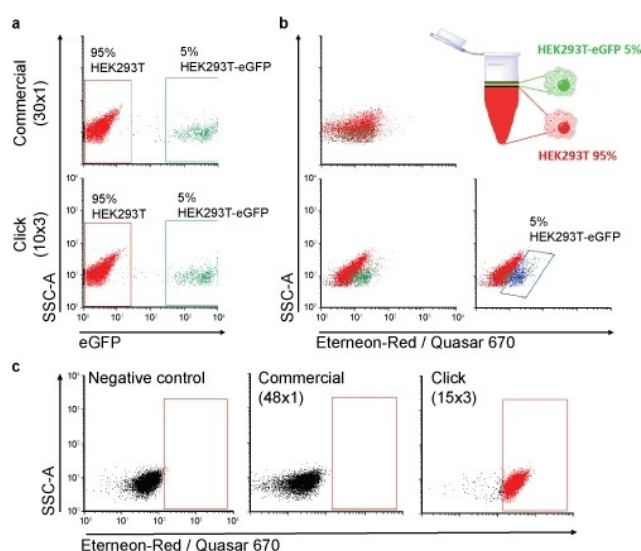

Figure 3. Flow cytometry analysis of mixed HEK293T and HEK293T expressing eGFP in a ratio 20:1. a) Mixing ratio determined on the basis of eGFP fluorescence. b) Mixing ratio determined by flow-FISH. At 0.05 ng/μL of probe, separation was only possible for the 10×3 set and establishes the same ratio when gated, in blue (5%). c) Flow-FISH for the detection of the endogenous ABL transcript.

increased fourfold to 0.2 ng DNA/μL (Figure S8ii). These flow-FISH data show again that the 10×3 probe allows to perform flow-FISH. Although an exact comparison between the 10×3 and the 30×1 probe sets is not possible, because of the differences in the fluorophore, the connectivity of the fluorophore to the probe and the quality of the strands, we believe that the data support the idea that flow-FISH is possible with our new probe design that has a limited number of probe oligonucleotides each one carrying three fluorophores.

In order to show that the new probe oligonucleotides are able to report transcription of a relevant endogenous gene by flow-cytometry, we next tested if the above method would be suitable for the detection of the ABL1-transcript. This time, we used a slightly larger 15×3 probe set. In order to again obtain comparative information we compared our design with a reported detection that used in this case 48 single labelled oligonucleotides (Semrau et al.).[5] The results are depicted in Figure 3c. While the 15×3 probe set with only 0.05 ng DNA/μL provides a clear shift in fluorescence compared to the negative control, the 48×1 probe set was under these conditions unable to provide a specific signal separation at this concentration.

We finally explored if the new (10×3) probe design enables imaging of RNA targets that are hard to image with conventional FISH probes. Such targets are characterized by extensive secondary structures that provides only few accessible sites for probe hybridization. We tested our probes by targeting a ~1 kb RNA viral transcript of the rotavirus A (RVA) gene segment-7. The idea was to test whether our set of probes detects these transcripts shortly after virus infection when the concentration of the transcript is expected to be very low. This specific target was chosen because of its extensive secondary structure that precludes hybridization of multiple probes, posing additional

ChemBioChem

Full Papers
doi.org/10.1002/cbic.202000081

Chemistry
Europe
European Chemical
Societies Publishing

challenges for conventional FISH.[22] We fixed rotavirus-infected cells, 4 hours post infection and incubated them with the (10× 3) DNA probe set targeting the RNA regions that were predicted to have less pronounced secondary structures.[23] To facilitate the detection of virus-infected cells, we took advantage of a stable cell line that expresses the rotavirus non-structural protein (NSP5)[24] fused to eGFP (see Methods in the Supporting Information).

After 4 hours post infection RVA, transcripts could be readily detected (Figure 4, top panel, red signal). No transcript specific signal was observed in mock-infected cells (Figure 4, bottom panel).

In parallel, we also analyzed eGFP-NSP5-marked foci in RVA-infected cells to identify cells at a more advanced stage of infection (Figure 4, top panel, green signal). As expected, these cells showed higher levels of the RNA transcript fully consistent with an increase of the amount of viral RNA transcripts over the course of the viral infection.

Importantly, the increased sensitivity of the new 10×3 detection approach uncovered a number of interesting insights. First, the data show a large variation in the amount of transcripts in different cells at the onset of infection. These variations likely reflect differences in the transcriptional activities of different rotavirus particles.[25] Furthermore, the transcripts do not seem to accumulate in a particular cellular compartment or in virus-induced organelles during early infection. These results thus show that the new 10×3 probe design is a significant step forward because it provides new biological insights.

Conclusions

We show here that by decreasing the number of probes and increasing the number of fluorophores on oligonucleotide probes using click chemistry provides fluorescent probe strands that allow efficient detection of RNA transcripts in live cells. The probes have a superb sensitivity and allow detection of transcripts that due to high secondary structure content provide only a limited number of binding sites. Self-quenching seems to play only a minor role, possibly because we used fluorophores that carry two negative charges each. The new probe design is
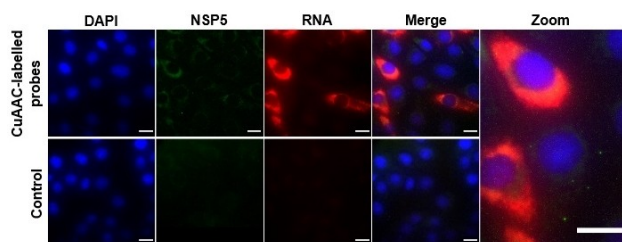
so sensitive that it even allows flow RNA-FISH to be established for demanding flow cytometry applications. These results pave the way for the detection of small highly structure RNA transcripts and transcripts with low abundance. A potential application could be the monitoring of leukemia therapy by flow-FISH to prevent dangerous relapse cases or the very early detection of virus infections now needed for the detection of the Corona virus to reduce the diagnostic gap and prevent uncontrolled propagation of the disease.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: click chemistry · fluorescence probes · mRNA detection · RNA-FISH · viral infection

Figure 4. Rotavirus gene segment-7 RNA transcripts imaged with 10×3 probes. Upper panel: RNA transcripts (red) in virus-infected cells stably expressing rotavirus non-structural protein NSP5 fused to eGFP (green) detected 4 hours post infection. Lower panel: Mock-infected MA104 cells stably expressing NSP5-eGFP. Scale bars: 20 μm.

[1] A. Raj, A. Van Oudenaarden, *Cell* **2008**, *135*, 216–226.
[2] A. Eldar, M. B. Elowitz, *Nature* **2010**, *467*, 167–173.
[3] S. Itzkovitz, A. van Oudenaarden, *Nat. Methods* **2011**, *8*, 12–19.
[4] A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, *Nat. Methods* **2008**, *5*, 877–879.
[5] A. Raj, S.Tyagi. Detection of individual endogenous RNA transcripts in situ using multiple singly labeled probes. In: Single molecule Tools: Chapter 17, Fluorescence based approaches, Part A; Walter, N. G. B. T.-M. in E., Ed.; Academic Press, **2010**, *472*, 365–368.
[6] S. Semrau, N. Crosetto, M. Bienko, M. Boni, P. Bernasconi, *Cell Rep.* **2011**, *6*, 18–23.
[7] T. Muramoto, D. Cannon, M. Gierliński, A. Corrigan, G. J. Barton, J. R. Chubb, *Proc. Mont. Acad. Sci.* **2012**, *109*, 7350–7355.
[8] A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, *Science* **1998**, *280*, 585–590.
[9] C. Larsson, I. Grundberg, O. Söderberg, M. Nilsson, *Nat. Methods* **2010**, *7*, 395–397.
[10] T. Trcek, T. Lionnet, H. Shroff, R. Lehmann, *Nat. Protoc.* **2017**, *12*, 1326–1348.
[11] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021; *Angew. Chem.* **2001**, *113*, 2056–2075.
[12] K. V. Gothelf, K. A. Jørgensen, *Chem. Rev.* **1998**, *98*, 863–910.
[13] S. Bräse, C. Gil, K. Knepper, V. Zimmermann, *Angew. Chem. Int. Ed.* **2005**, *44*, 5188–5240.
[14] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.
[15] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.
[16] K. Gutsmiedl, D. Fazio, T. Carell, *Chem. Eur. J.* **2010**, *16*, 6877–6883.
[17] J. Gierlich, G. A. Burley, P. M. E. Gramlich, D. M. Hammond, T. Carell, *Org. Lett.* **2006**, *8*, 3639–3642.

[18] P. M. E. Gramlich, S. Warncke, J. Gierlich, T. Carell, *Angew. Chem. Int. Ed.* **2008**, *47*, 8350–8358.

[19] P. M. E. Gramlich, C. T. Wirges, A. Manetto, T. Carell, *Angew. Chem. Int. Ed.* **2008**, *47*, 8350–8358.

[20] S. Hesse, A. Manetto, V. Cassinelli, J. Fuchs, L. Ma, N. Raddaoui, A. Houben, *Chromosom. Res.* **2016**, *24*, 299–307.

[21] J. Gierlich, G. A. Burley, P. M. E. Gramlich, D. M. Hammond, T. Carell, *Org. Lett.* **2006**, *8*, 3639–3642.

[22] R. Arrigucci, Y. Bushkin, F. Radford, K. Lakehal, P. Vir, R. Pine, D. Martin, J. Sugarman, Y. Zhao, G. S. Yap, *Nat. Protoc.* **2017**, *12*, 1245–1260.

[23] A. Borodavka, E. C. Dykeman, W. Schrimpf, D. C. Lamb, *elife* **2017**, *6*, e27453.

[24] F. Rodriguez, O. R. Burrone, C. Eichwald, *J. Gen. Virol.* **2004**, *85*, 625–634.

[25] E. N. Salgado, S. Upadhyayula, S. C. Harrison, *J. Virol.* **2017**, *91*, e00651–17.

[26] A. Borodavka, U. Desselberger, J. T. Patton, *Curr. Opin. Virol.* **2018**, *33*, 106–112.

ChemBioChem

Supporting Information

Supersensitive Multifluorophore RNA-FISH for Early Virus Detection and Flow-FISH by Using Click Chemistry

Nada Raddaoui[+], Stefano Croce[+], Florian Geiger, Alexander Borodavka, Leonhard Möckl, Samuele Stazzoni, Bastien Viverge, Christoph Bräuchle, Thomas Frischmuth, Hanna Engelke, and Thomas Carell*

Eternion Red 645   Quasar 670
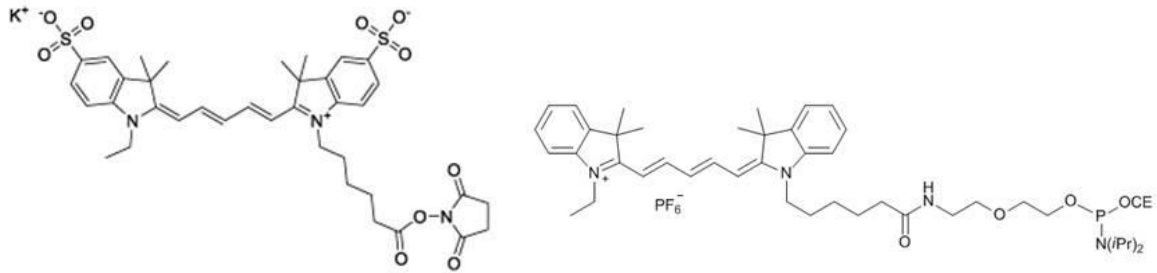
Probe design and sequences for eGFP mRNA

5'-ATGGTGAGCA AGGGCGAGGA GCTGTTCACC GGGGTGGTGC CCATCCTGGT CGAGCTGGAC GGCGACGTAA ACGGCCACAA GTTCAGCGTG TCCGGCGAGG GCGAGGGCGA TGC
3'-TACCACTCGT TCCCGCTCCT CGACAAGTGG CCCCACCACG GGTAGGACCA GCTCGACCTG CCGCTGCATT TGCCGGTGTT CAAGTCGCAC AGGCCGCTCC CGCTCCCGCT ACG

1 TCGT TCCCGCTCCT   CGACAAGT   **2** CTG CCGCTGCATT   TGCCGGTGT

CACCTAC GGCAAGCTGA CCCTGAAGTT CATCTGCACC ACCGGCAAGC TGCCCGTGCC CTGGCCCACC CTCGTGACCA CCCTGACCTA CGGCGTGCAG TGCTTCAGCC GCTACCCC
GTGGATGCCGTTCGACT GGGACTTCAA GTAGACGTGG TGGCCGTTCG ACGGGCACGG GACCGGGTGG GAGCACTGGT GGGACTGGAT GCCGCACGTC ACGAAGTCGGCGATGGGG

3 TTCAA GTAGACGTGG TGGCCGT

GA CCACATGAAG CAGCACGACT TCTTCAAGTC CGCCATGCCCGAAGGCTACG TCCAGGAGCG CACCATCTTC TTCAAGGACG ACGGCAACTA CAAGACCCGC GCCGAGGTGA AGTT
CT GGTGTACTTC GTCGTGCTGA AGAAGTTCAG GCGGTACGGG CTTCCGATGC AGGTCCTCGC GTGGTAGAAG AAGTTCCTGC TGCCGTTGAT GTTCTGGGCG CGGCTCCACT TCAA

4 AGAAGTTCAG GCGGTACGGGCT   **5** TTCCTGC TGCCGTTGAT GTTCT   **6** CTCCACT TCAA

CGAGGG CGACACCCTG GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG ACGGCAACAT CCTGGGGCAC AAGCTGGAGT ACAACTACAA CAGCCACAACGTCTATAT
GCTCCC GCTGTGGGACCACTTGGCGT AGCTCGACTT CCCGTAGCTG AAGTTCCTCC TGCCGTTGTA GGACCCCGTG TTCGACCTCA TGTTGATGTT GTCGGTGTTG CAGATATA

GCTCCC GCTGT   **7** CTG AAGTTCCTCC TGCCGTTGT

CA TGGCCGACAA GCAGAAGAAC GGCATCAAGG TGAACTTCAA GATCCGCCAC AACATCGAGG ACGGCAGCGT GCAGCTCGCC GACCACTACC AGCAGAACAC CCCCATCGGC
GT ACCGGCTGTT CGTCTTCTTG CCGTAGTTCC ACTTGAAGTT CTAGGCGGTG TTGTAGCTCC TGCCGTCGCA CGTCGAGCGG CTGGTGATGG TCGTCTTGTG GGGGTAGCCG

8 TGTAGCTCC TGCCGTCGCA CGT

GACGGCCCCG TGCTGCTGCC CGACAACCAC TACCTGAGCA CCCAGTCCGC CCTGAGC AAA GACCCCAACG AGAAGCGCGA TCACATGGTC CTGCTGGAGT TCGTGACCGC CGCC
CTGCCGGGGC ACGACGACGG GCTGTTGGTG ATGGACTCGT GGGTCAGGCG GGACTC GTTT CTGGGGTTGC TCTTCGCGCT AGTGTACCAG GACGACCTCA AGCACTGGCG GCGG

9 GACGG GCTGTTGGTG ATGGACT

GGGATC ACTCTCGGCA TGGACGAGCTGTACAAGTAA-3'
CCCTAG TGAGAGCCGT ACCTGCTCGA CATGTTCATT-5'

10 GAGCCGT ACCTGCTCGA CATGT

Sense sequence of eGFP-mRNA   Control anti-sense oligonucleotides   CuAAA-labelled antti-sense oligonucleotides

Figure S1. Sequence of the eGFP transcript in grey, together with the sequence of the antisense 10x3 probes in red. The comparison 30x1 probes are shown in green / light green.

Oligonucleotide synthesis

Into each of the DNA strands, three C8-alkyne-dU building blocks were inserted at former dT positions using phosphoramidites developed before[1].

Table S1. Sequences of the 10 oligonucleotide RNA FISH probes and the positions of the C8-alkyne-dU building blocks. The Oligonucleotides listed below were generated for eGFP mRNA using the Stellaris RNA FISH probe designer (https://www.biosearchtech.com/stellaris-designer). The positions of the modifications are highlighted in blue.

Number	Sequence 5'-> 3'
1	TGA ACA GCT CCT CGC CCT TGC T
2	TGT GGC CGT TTA CGT CGC CGT C
3	TGC CGG TGG TGC AGA TGA ACT T
4	TCG GGC ATG GCG GAC TTG AAG A
5	TCT TGT AGT TGC CGT CGT CCT T
6	TGT CGC CCT CGA ACT TCA CCT C
7	TGT TGC CGT CCT CCT TGA AGT C
8	TGC ACG CTG CCG TCC TCG ATG T
9	TCA GGT AGT GGT TGT CGG GCA G
10	TGT ACA GCT CGT CCA TGC CGA G

Click reaction

Each probe was prepared separately. The Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) was performed combining 5 nmole of DNA oligo and 45 nmol dye azide: TAMRA-azide (baseclick GmbH) or Eterneon Red 647-azide (baseclick GmbH) in 20 µL reaction volume in the presence of the catalyst CuBr 6.25 mM, tris-hydroxypropyltriazolylmethylamine (THPTA) 12.5 mM, and 2.5 % DMSO. The solution is then incubated for 1 h at 45 °C at 700 RPM protected from light and then cleaned by standard EtOH precipitation before the application in FISH experiments. Further purification was not needed.

HPLC example

For RP-HPLC analysis of labelled probes, a Waters system (Waters e2695 Separation Modul and Waters 2998 PDA) equipped with the XBridgeTM OST C18 column (2.5 µm, 4.6 mm x 50 mm) was used. Linear gradients of buffer A (0.1 M triethylammonium acetate, pH 7.5) to buffer B (0.1 M triethylammonium acetate in 80 % acetonitrile, pH 7.5) were run at a flow rate of 1.5 mL/min at 40 °C column temperature. For detection of oligonucleotide mixtures 260 nm absorbance was chosen.

Method:
Buffer A: 100 % to 70 % in 8 min, 70 % to 15 % in 2 min, 15 % to 0 % in 1 min.

Buffer B: 0 % to 30 % in 8 min, 30 % to 85 % in 2 min, 85 % to 100 % in 1 min

HPL chromatogram example of the oligo before (i) and after click reaction (ii)
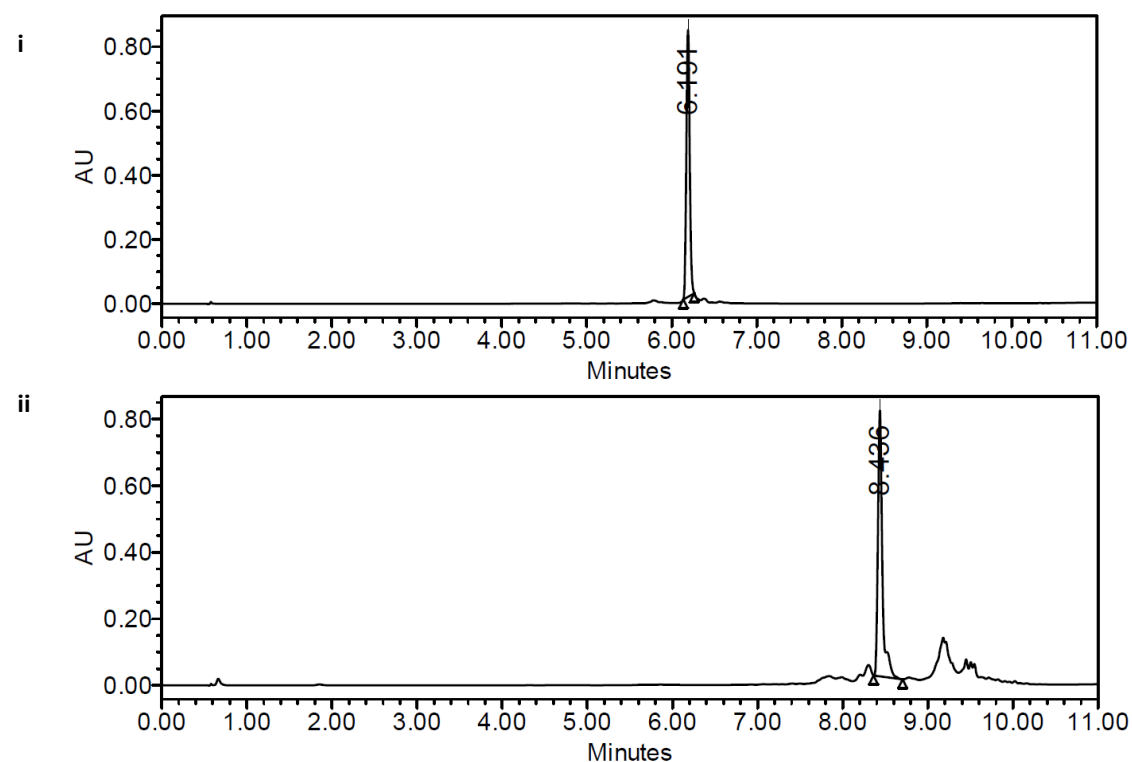

Figure S2. HPLC chromatogram of the probe before (i) and after the click reaction (ii)

Comparison of commercial set of probes and click probes

Analytical RP-HPLC was performed using a Macherey-Nagel Nucleodur 100-3 C18ec column on a Waters Alliance 2695 Separation Module coupled with a 2996 Photodiode Array Detector and a 2475 Multi wavelength Fluorescence Detector, using a flow of 0.5 mL/min. Linear gradients of buffer A (0.1 M triethylammonium acetate, pH 7.0) to buffer B (0.1 M triethylammonium acetate in 80 % acetonitrile, pH 7.0) were run at a flow rate of 0.5 mL/min. For detection of oligonucleotide mixture 260 nm absorbance was chosen, while 647 nm absorbance was used to detect the fluorophore. Method: 0-60 % B in 20 min, 60 to 100 % B in 1 min, 100 % B for 4 min

Table S2. Sequence of the commercial FISH probes, consisting of 30 oligonucleotides with one dye each (30x1)

Number	Sequence 5'-> 3'
1	GCT CCT CGC CCT TGC TCA CCA T
2	ATG GGC ACC ACC CCG GTG AAC A
3	GTC GCC GTC CAG CTC GAC CAG G
4	CGC TGA ACT TGT GGC CGT TTA C
5	TCG CCC TCG CCC TCG CCG GAC A
6	GGT CAG CTT GCC GTA GGT GGC A
7	CGG TGG TGC AGA TGA ACT TCA G
8	GGC CAG GGC ACG GGC AGC TTG C
9	GGT CAG GGT GGT CAC GAG GGT G
10	GGC TGA AGC ACT GCA CGC CGT A
11	TGC TTC ATG TGG TCG GGG TAG C
12	GGC GGA CTT GAA GAA GTC GTG C
13	CCT GGA CGT AGC CTT CGG GCA T
14	TCC TTG AAG AAG ATG GTG CGC T
15	GCG GGT CTT GTA GTT GCC GTC G
16	TCG ATG CGG TTC ACC AGG GTG T
17	CCT TGA AGT CGA TGC CCT TCA G
18	TGC CCC AGG ATG TTG CCG TCC T
19	GTT GTA GTT GTA CTC CAG CTT G
20	GCC ATG ATA TAG ACG TTG TGG C
21	GAT GCC GTT CTT CTG CTT GTC G
22	GGC GGA TCT TGA AGT TCA CCT T
23	ACG CTG CCG TCC TCG ATG TTG T
24	GCT GGT AGT GGT CGG CGA GCT G
25	CCG TCG CCG ATG GGG GTG TTC T
26	GGT TGT CGG GCA GCA GCA CGG G
27	GCG GAC TGG GTG CTC AGG TAG T
28	CTC GTT GGG GTC TTT GCT CAG G
29	GCA GGA CCA TGT GAT CGC GCT T
30	GCT CGT CCA TGC CGA GAG TGA T
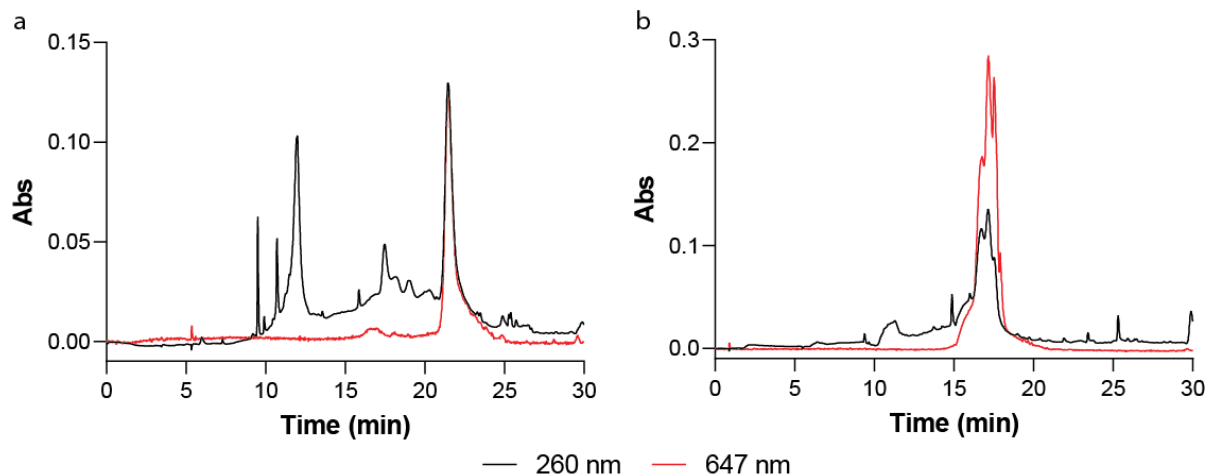
Figure S3. Comparison of (a) a commercial probe set (30x1 Quasar 670 dye) and (b) a clicked probe set (10x3 Eterneon-Red dye). Both sets were loaded in equal amounts, based on 260 nm absorption. Both dyes are very similar cyanine-5 fluorophores.
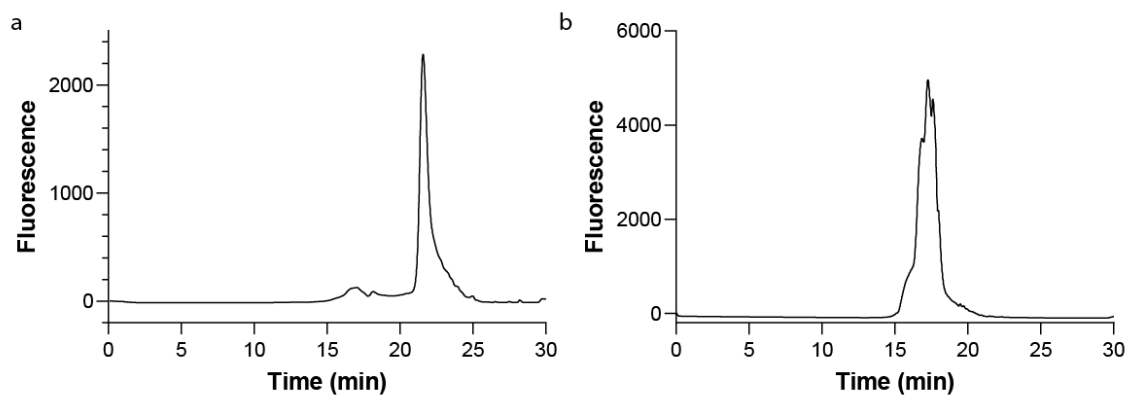
Figure S4. Fluorescence analysis: The fluorescence analysis of the two sets of probes confirm what described previously following the absorbance at 647 nm. Using the same amount of DNA, the commercial probes (a) show an overall lower fluorescence value than the click probes (b). Detection parameters: λ_{exc}: 647 nm, λ_{em}: 665 nm.

In-vitro-FISH with total RNA from eGFP- HEK293T cell line

Total RNA was extracted from eGFP-HEK239T cells using Total-RNA-kit, peqGOLD, catalogue number 732-2868 (VWR). FISH in total RNA was performed as reported in Semrau *et. al.* with minor changes. After DNase digestion, 0.5 µl of total RNA was spotted in the center of an RNase-free cover slips, which was attached to a microscope slide. The total RNA spotted on cover slips were dried for 20 min at 80 °C. After fixation of the spots for 5 min at RT, a washing step with 2x saline sodium citrate (SSC) followed. Either the commercial (30x1, labelled with Quasar 670) or the CuAAC-labelled probes (10x3, Eterneon-Red 647 azide) were added to the spotted RNA in hybridization buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0 in nuclease-free water, 10 % (v/v) formamide) and incubated for 5 min at 80 °C. The hybridization buffer including the FISH probes was washed with washing buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0, 10 % (v/v) formamide in RNase-free water). The samples were imaged in presence of 2xSSC.
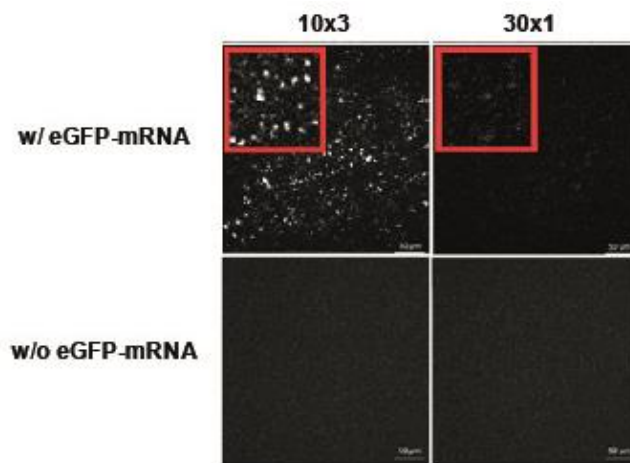
Figure S5. In vitro FISH experiment of total RNA extracted from eGFP-HEK239T cells. Total RNA from eGFP-HEK239T cells hybridized with the CuAAC-labelled probes (10x3) showed a very high signal of diffraction limited spots. In the control experiment, total RNA from eGFP-HEK239T cells was hybridized with the commercial probe set (30x1). Only a very week specific signal was detected. Scale bars: 50 µm

Titration for different numbers of triple labelled oligonucleotides

Since the oligonucleotides were labelled separately via click chemistry, a setting of different numbers of oligonucleotides to a probe set was possible. Four probe sets were prepared: 3x3, 5x3, 7x3 and 10x3 with 10 ng total amount of the oligonucleotides in each set. The sets were hybridized with fixed and permeabilized eGFP-HEK293T and with HEK293T for the negative control without eGFP expression.
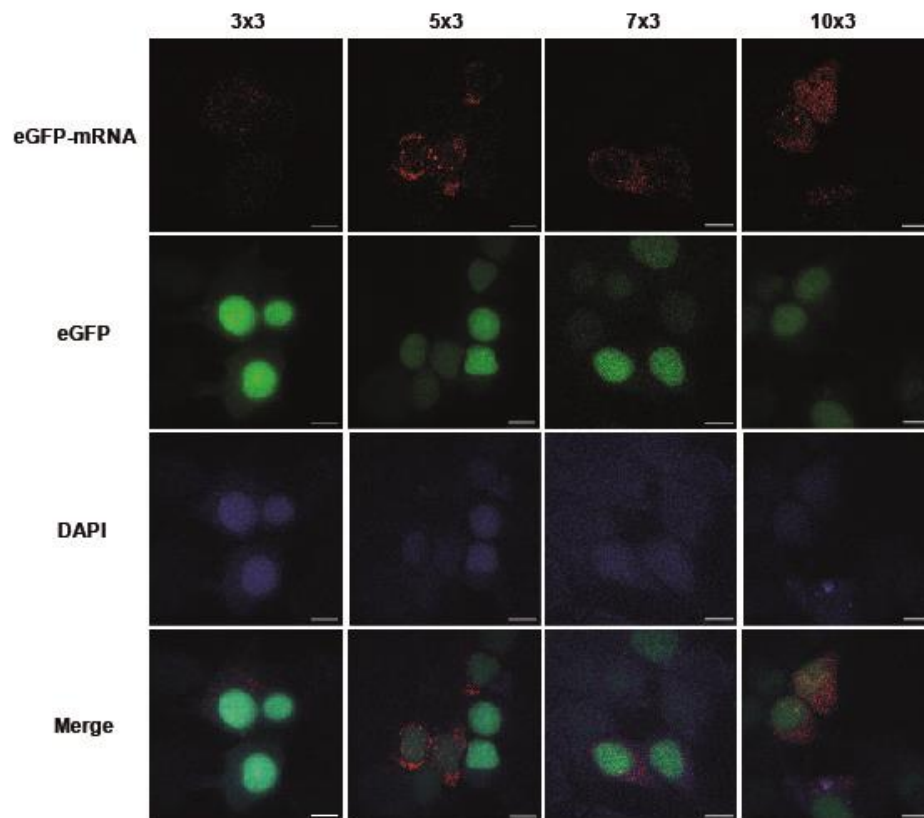
Figure S6. Differences in the signal intensities of eGFP-mRNA using different numbers of triple labelled probes (3x3, 5x3, 7x3 and 10x3) in eGFP-expressing HEK293T cells. Scale bar: 50 µm

Click post hybridization

eGFP-HEK293T were grown in 8 chamber µ-slides cell culture plates (ibidi) until 80 % cell density is reached. After removing the culture medium, cells were washed with 1x RNase-free phosphate saline buffer (PBS) and fixed 15 minutes at RT with the fixation buffer (4 % paraformaldehyde (v/v) in nuclease-free PBS. Cells were then washed twice with 3 % bovine serum albumin (BSA) in 1x PBS and permeabilized with 100 % EtOH for at least 1 h at 4 °C. After permeabilization and removing of the EtOH, cells were washed twice with 3 % BSA in 1x PBS. 5 pmol of each oligonucleotide in 150 µL hybridization buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0 in nuclease-free water, 10 % (v/v) formamide) was added to each chamber. The hybridization occurred for 3 h at 37 °C. The unbound oligonucleotides and the hybridization buffer were washed using 10 % washing buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0, 10 % (v/v) formamide in RNase-free water) followed by two washing steps with 3 % BSA in 1x PBS. 250 µL of click cocktail ($CuSO_4$ 50 µM, THPTA 250 µM, 7 equivalents 5-TAMRA-PEG3-azide to a final concentration of 25 µM and NaAscorbate to final concentration of 2.5 mM) was added to the cells, which were then incubated 30 minutes at RT protected from light. After hybridization, cells were washed three times either with 3 % BSA (w/v) or DMSO (5, 10, 20, 30, 40 and 50 % in 1x PBS (v/v)). Afterwards, 2x SSC was added to the cells and analysis via microscopy was performed. The signal to noise ratio is reduced when the concentration of DMSO is increased. Scale bar: 50 µm.
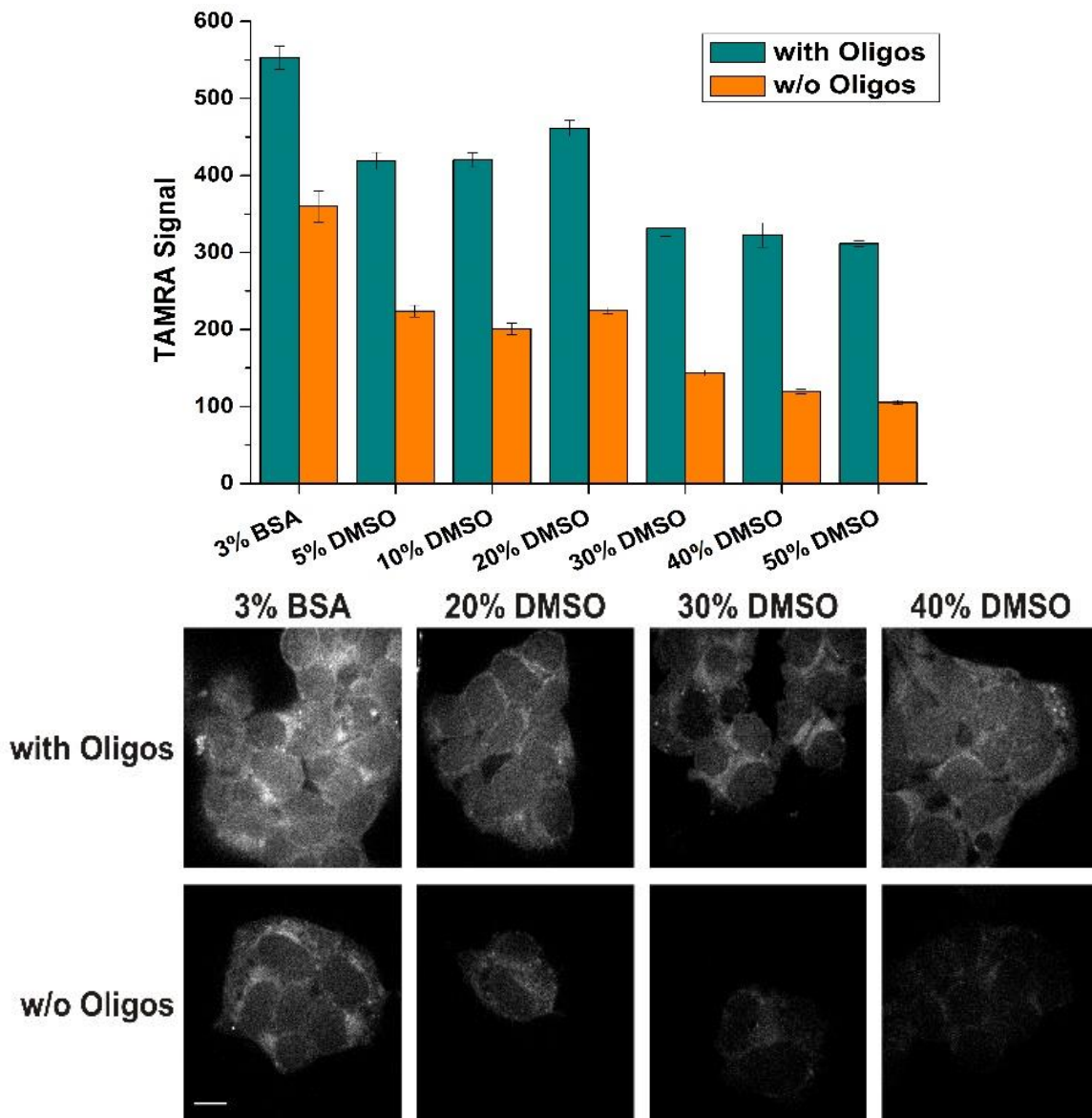

Figure S7. Click reaction of the FISH probes after hybridization.

Cell culture of human cell strains

HEK293T and HEK293T-eGFP cells were cultivated at 37 °C with CO_2-enriched (5 %) atmosphere. Dulbecco's modified Eagle medium (DMEM) supplemented with 10 % fetal bovine serum (FBS), 1 % penicillin and streptomycin, was used as growing medium. When reaching a confluence of 70 % to 80 %, the cells were passaged in a new culture flask. For staining experiments, 1.5×10^4 cells were seeded in each well of a µ-Slide 8 Well from ibidi (ibiTreat, #1.5 polymer coverslip, catalog No. 80826) and cultured for two days or until a density of 80 % is reached.

FISH protocol for HEK293T cells

HEK293T cells were fixed with 4 % (v/v) paraformaldehyde in nuclease-free phosphate saline buffer (PBS) for 10 min at room temperature. Samples were then washed twice with PBS, and the fixed cells were permeabilized with 70 % (v/v) ethanol (200 proof) in RNAse-free water and stored in ethanol at +4 °C for at least 12 hours prior to hybridization. Permeabilized cells were then re-hydrated for 5 min in a pre-hybridization buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0 in nuclease-free water, 10 % v/v formamide, supplemented with 2 mM vanadyl ribonucleoside complex). Re-hydrated samples were hybridized with an equimolar mixture of 10 RNA FISH DNA probes specific to bovine rotavirus strain RF gene segment 7 (62.5 nM final concentration, see Table S4) in a total volume of 200 µl of the hybridization buffer (Stellaris RNA FISH hybridization buffer, SMF-HB1-10, Biosearch Technologies, supplemented with 10 % v/v deionized formamide). After 4 hours of incubation at 37 °C in a humidified chamber, samples were briefly rinsed with the wash buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0, 10 % v/v formamide in nuclease-free water) after which a fresh aliquot of 300 µl of the wash buffer was applied to each well and incubated twice at 37 °C for 30 min. After washes, nuclei were briefly stained with 300 nM 4',6-diamidino-2-phenylindole (DAPI) solution in 300 mM NaCl, 30 mM trisodium citrate, pH 7.0) and the samples were finally rinsed with and stored in the same buffer without DAPI prior to imaging.

Flow cytometry

FISH experiments for flow cytometry analysis were performed as described previously by Arrigucci et al[2].

Table S3. Sequences of the BCR and ABL FISH probes labelled with click chemistry. The sequences were generated using the Stellaris RNA FISH probe designer (https://www.biosearchtech.com/stellaris-designer). The positions of the C8-Alkyne-dU building blocks are highlighted with blue

Number	Sequence 5'-> 3'
BCR#1	**T**AG CTC T**T**C TTT TCC **T**TG GC
BCR#2	A**T**C CGC **T**CG AAG T**T**G GAC CT
BCR#3	AAC **T**CG ACG T**T**C ACG **T**AG AA
BCR#4	**T**TT TTG CGC **T**CC ATC **T**GC AT
BCR#5	**T**TC AAC **T**CG GCG TCC **T**CG TA
BCR#6	A**T**G CTC **T**GG TAG GGC **T**GG TA
BCR#7	**T**CC GCA A**T**C CTC AAA AC**T** CC
BCR#8	**T**TC TCA T**T**G GAG C**T**G CAG TC
BCR#9	**T**TG TCC CGG AAC A**T**G CGG **T**A
BCR#10	A**T**C TGC G**T**C TCC A**T**G GAA GG
BCR#11	AC**T** CGC TN**T** AGT GGA C**T**C CA
BCR#12	A**T**C TGC **T**GA CTC G**T**C AGC AC
BCR#13	**T**TG TGG A**T**C TCG **T**AG AGC TC
BCR#14	**T**CA GG**T** TCT CGG AGA **T**TT CT
BCR#15	**T**GG CAT C**T**T TGT TGC **T**TC TG
ABL1#1	GCC A**T**T TTT GG**T** TTG GGC **T**T
ABL1#2	**T**TG ACT GGC G**T**G ATG **T**AG TT
ABL1#3	G**T**A ATG G**T**A CAC CC**T** CCC TT
ABL1#4	A**T**G ATG A**T**G AAC CAA C**T**C GG
ABL1#5	A**T**A ATG GAG CG**T** GGT GA**T** GA
ABL1#6	A**T**A GAC AG**T** GGG CT**T** GTT GC
ABL1#7	A**T**C TCC CAC **T**T GTCG **T**AG TT
ABL1#8	**T**TG TGC **T**TC ATG G**T**G ATG TC
ABL1#9	**T**CC TCC **T**TC AAG GTC **T**TC AC
ABL1#10	C**T**T TCA TGA C**T**G CAG C**T**T CT
ABL1#11	AGG **T**TC CCG **T**AG GTC A**T**G AA
ABL1#12	A**T**C TGA G**T**G GCC ATG **T**AC AG
ABL1#13	**T**AC CT**T** CAC CAA G**T**G GTT CT
ABL1#14	A**T**T TGA **T**GG GGA AC**T** TGG CT
ABL1#15	A**T**G GAG AAC **T**TG TTG **T**AG GC

Cells flow cytometer data were obtained using the BD FACS Canto II equipped with air cooled green (488 nm solid state; 20 mW laser output) and red (633 nm HeNe; 17 W output) lasers, and the data were analyzed with the FCS Express 6 (De Novo Software). The fluorescent detectors/filters relevant for this study were FITC (530 +/- 30 nm) and APC (660 +/- 30 nm), respectively. Acquisition was done with a flow rate of 10 µL/min. FSC and SSC were used for detection of forward light scatter parameter and side light scatter parameter. The machine was cleaned before and after each measurement. The photomultiplier tube (PMT) voltage for each parameter was adjusted in order to have the cells displayed in the centre of the investigating plot.

After hybridization and wash, HEK293T or HEK293T-eGFP were immediately analysed by flow cytometer, the samples were maintained in buffer during acquisition.
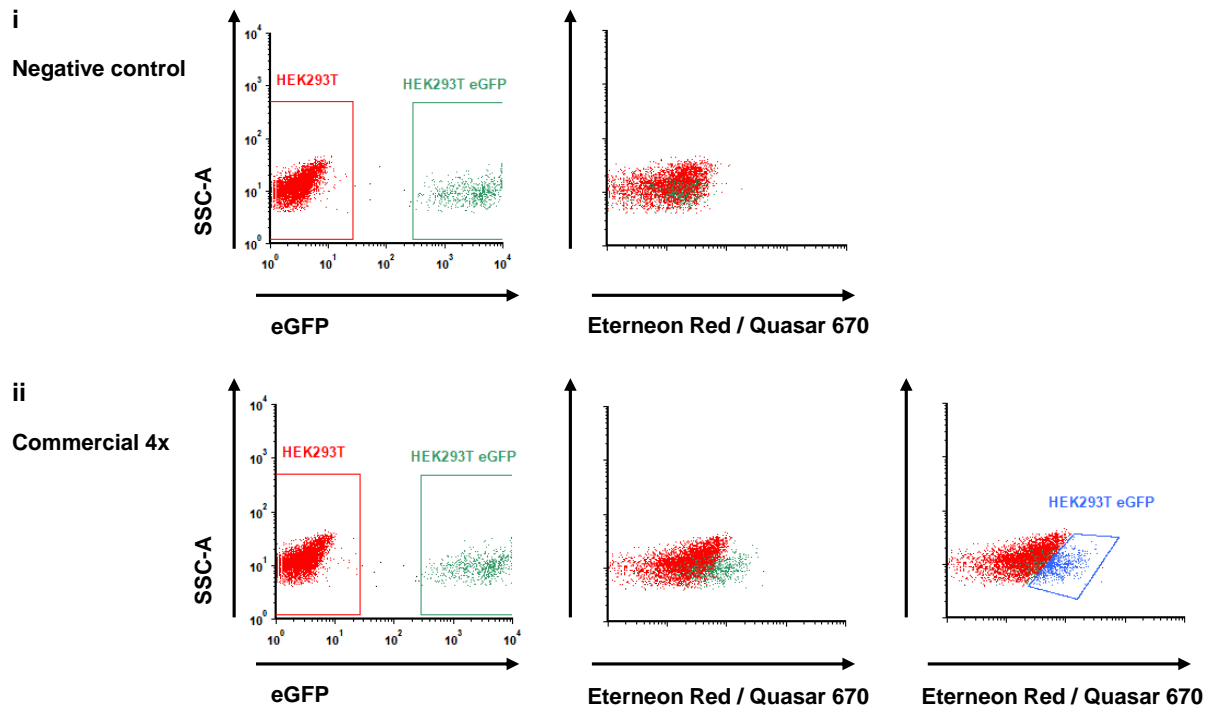
Figure S8. **FACS Dot plot comparison of 30x1 and 10x3 probes.** (i) The negative control of the experiment, where FISH was performed replacing the set of probes with the same volume of H_2O. A population of HEK293T and HEK293TeGFP was mixed in a ratio 20:1, and analyzed by FACS. The eGFP dot-plot reports the correct separation of the HEK293T and HEK293-eGFP. The Eterneon Red dot plot shows a mix of the two populations, as expected. (ii) Same result as the 10x3 probes was obtained with the 30x1 probes set by increasing the concentration four fold to 0.2 ng DNA/µl.

Cells and viruses

The rhesus monkey epithelial cell line MA104 stably expressing NSP5-EGFP [3] was cultured in DMEM (Dulbecco's modified Eagle medium, GlutaMax-I, 4.5 g/L glucose, ThermoFisher), supplemented with 10 % fetal bovine serum (FBS), 1x MEM non-essential amino acids solution (Sigma), 1 mM sodium pyruvate (Sigma) and 500 µg/ml G418 (Roche).

Bovine rotavirus A strain RF (G6P6[1]) was a generous gift from Dr. Ulrich Desselberger (University of Cambridge, UK). It was grown in MA104 Clone 1 cell line (ATCC CRL-2378.1), as described previously [4].

For RNA imaging experiments, MA104 cells expressing NSP5-EGFP were seeded into Ibidi 8-well µ-slides and allowed to reach 90 % confluency prior to the infection. Confluent MA104 cell monolayers were rinsed twice with DMEM medium without FBS for 10 minutes to remove any residual FBS, and were subsequently infected with trypsin-activated bovine rotavirus at multiplicity of infection, MOI = 20.

Rotavirus RNA imaging using RNA FISH

Rotavirus-infected and mock-infected MA104 controls were fixed with 4 % (v/v) paraformaldehyde in nuclease-free phosphate saline buffer (PBS) for 10 min at room temperature. Samples were then washed twice with PBS, and the fixed cells were permeabilized with 70 % (v/v) ethanol (200 proof) in RNAse-free water and stored in ethanol at +4 °C for at least 12 hours prior to hybridization. Permeabilized cells were then re-hydrated for 5 min in a pre-hybridization buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0 in nuclease-free water, 10 % v/v formamide, supplemented with 2 mM vanadyl ribonucleoside complex). Re-hydrated samples were hybridized with an equimolar mixture of 10 RNA FISH DNA probes specific to bovine rotavirus strain RF gene segment 7 (62.5 nM final concentration, see SI Table 1) in a total volume of 200 µl of the hybridization buffer (Stellaris RNA FISH hybridization buffer, SMF-HB1-10, Biosearch Technologies, supplemented with 10 % v/v deionized formamide). After 4 hours of incubation at 37 °C in a humidified chamber, samples were briefly rinsed with the wash buffer (300 mM NaCl, 30 mM trisodium citrate, pH 7.0, 10 % v/v formamide in nuclease-free water, after which a fresh aliquot of 300 µl of the wash buffer was applied to each well and incubated twice at 37 °C for 30 min. After washes, nuclei were briefly stained with 300 nM 4',6-diamidino-2-phenylindole (DAPI) solution in 300 mM NaCl, 30 mM trisodium citrate, pH 7.0) and the samples were finally rinsed with and stored in the same buffer without DAPI prior to imaging.

Fluorescent microscopy imaging

HEK cells were imaged on a Zeiss Cell Observer SD microscope equipped with a Yokogawa spinning disk unit CSU-X1 using a 1.40 NA 100x Plan-apochromat oil immersion objective (Zeiss). RNA FISH probes were imaged using 532/561 nm and a BP 690 filter, and eGFP signal was imaged with 488 nm laser excitation and a BP 525/50 filter, respectively. DAPI staining was visualized with 405 nm excitation using a BP 525/50 filter. Data analysis yielding the labeling efficiency was performed using ImageJ.

Rotavirus RNA FISH probes imaging was carried out using a Nikon Ti-E widefield microscope equipped with a 100x 1.40 NA Plan Apochromat oil immersion objective, a CoolLED pE 4000 light source, a Zyla camera and a Quad DAPI/FITC/CY3/CY5 HC filter set.

Z-series spanning the full cell volume with a step size of 0.5 µm at 30 % power with 55 ms exposure times at 385 nm for DAPI, 58 % power with 500-ms exposure times at 470 nm for eGFP and at 35 % power with 500-ms exposures at 550-nm for TAMRA were acquired and the final images represent maximum intensity projections calculated using ImageJ.

Table S4. Into each of the DNA strands designed, three C8-alkyne-dU building blocks were inserted at former dT positions using phosphoramidites developed before[1]. Sequences of the 10 oligonucleotide RNA FISH probes listed below were generated for the rotavirus A gene segment 7 (gene accession no. KF729693.1) using the Stellaris RNA FISH probe designer (https://www.biosearchtech.com/stellaris-designer). The position of the modifications is highlighted in blue.

Number	Sequence 5' -> 3'
1	A**T**T GTG GTA **T**AT TCA A**T**A CCA
2	AG**T** CCA TTA **T**TC TCG TTA **T**TG
3	GT**T** TTG CGC A**T**T TAT TA**T** GGT
4	AAG G**T**A TCT **T**TC CAT TCA G**T**A
5	GT**T** CTT GTG **T**AG AGT CA**T** ATT
6	A**T**A GCG TTA **T**GT CCA T**T**G GAT
7	ACC AAG **T**GT TTG **T**GT ATT **T**AA
8	TG**T** AAT CAC **T**AA CTT C**T**C CGT
9	CTG **T**AT GAC **T**GC TAC GTT C**T**C
10	G**T**G GTG TAG T**T**G TTG GA**T** CAG

References:

1. Gierlich J, Burley GA, Gramlich PME, Hammond DM, Carell T: **Click Chemistry as a Reliable Method for the High-Density Postsynthetic Functionalization of Alkyne-Modified DNA**. *Org Lett* 2006, **8**:3639–3642.

2. Arrigucci R, Bushkin Y, Radford F, Lakehal K, Vir P, Pine R, Martin D, Sugarman J, Zhao Y, Yap GS, et al.: **FISH-Flow, a protocol for the concurrent detection of mRNA and protein in single cells using fluorescence in situ hybridization and flow cytometry**. *Nat Protoc* 2017, **12**:1245–1260.

3. Eichwald C, Rodriguez JF, Burrone OR: **Characterization of rotavirus NSP2/NSP5 interactions and the dynamics of viroplasm formation**. *J Gen Virol* 2004, **85**:625–634.

4. Cheung W, Gill M, Esposito A, Kaminski CF, Courousse N, Trugnan G, Keshavan N, Desselberger U, Chwetzoff S, Lever A: **Rotaviruses associate with cellular lipid droplet components to replicate in viroplasms, and compounds disrupting or blocking lipid droplets inhibit viroplasm formation and viral replication**. *J Virol* 2010, **84**:6782–6798.

8. Glossary

6HB: 6-helix bundle

α-Gal A: α-Galactosidase A

A: Adenine

AGEs: advanced glycation end products

AFM: Atomic force microscopy

AGE: Agarose gel electrophoresis

BTTAA: 2-(4-((bis((1-(tert-butyl)-1H-1,2,3-triazol-4-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)acetic acid

BTTES: tris(triazolylmethyl)amine

C: Cytosine

CM: chainmails

CPG: controlled pore glass

CuAAC: Cupper(I) Catalyzed Azide-Alkine Cycle addition

DMT: 4,4'-dimethoxytrityl

DNA: The deoxyribonucleic acid

DNase I: deoxyribonuclease I

DNS: DNA structures

dsDNA: double stranded DNA filament

EdU: 5-Ethynil-deoxy-Uridine

EGFP: Enhanced green fluorescent protein

FACS: fluorescent activated cell sorting

FBS: fetal bovine serum

FISH: Fluorescent In Situ Hybridization

G: Guanine

GOs: Gene oligonucleotides

ITN: Innovative Training Network

IUTP: 5-iodouridine triphosphate

LAMP-1: Lysosomal-associated membrane protein

LMU: Ludwig Maximilians Universität

LNP: Lipid Nanoparticle

mRNA: messenger RNA

ncDNS: non-covalent structures

GalNac: N-acetylgalactosamine

OD: octahedral

ODN: Oligonucleotides

PAGE: Polyacrylamide gel electrophoresis

PCA: DNA polymerase cycling assembly

PCR: Polymerase chain reaction

PABP: poly(A) binding protein

RO: rectangular DNA origami

SBO: Square box DNA origami

siRNA: Short interfering RNA

SR-A: scavenger receptors of class A

SR-B: scavenger receptors of class B

SR-E: scavenger receptors of class E

SPAAC: Strain promoted azide alkyne cycloaddition

ssDNA: single strand DNA filament

T: Thymine

TdT: terminal deoxynucleotidyl transferase

TEM: transmission electron microscopy

TD: tetrahedral

tRNA: transfer RNA

THPTA: Tris((1-hydroxy-propyl-1H-1,2,3-triazol-4-yl)methyl)amine

TBTA: Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine

U: Uracil

UTR: Untranslated region

9. References

(1) Cassinelli, V.; Oberleitner, B.; Sobotta, J.; Nickels, P.; Grossi, G.; Kempter, S.; Frischmuth, T.; Liedl, T.; Manetto, A. One-Step Formation of "Chain-Armor"-Stabilized DNA Nanostructures. *Angew. Chemie - Int. Ed.* **2015**, *54*, 7795–7798.

(2) Li, H.; Labean, T. H.; Leong, K. W. Nucleic Acid-Based Nanoengineering: Novel Structures for Biomedical Applications. *Interface Focus* **2011**, *1*, 702–724.

(3) Raniolo, S.; Vindigni, G.; Ottaviani, A.; Unida, V.; Iacovelli, F.; Manetto, A.; Figini, M.; Stella, L.; Desideri, A.; Biocca, S. Selective Targeting and Degradation of Doxorubicin-Loaded Folate-Functionalized DNA Nanocages. *Nanomedicine Nanotechnology, Biol. Med.* **2018**, *14*, 1181–1190.

(4) Fu, Y.; Zeng, D.; Chao, J.; Jin, Y.; Zhang, Z.; Liu, H.; Li, D.; Ma, H.; Huang, Q.; Gothelf, K. V; *et al.* Single-Step Rapid Assembly of DNA Origami Nanostructures for Addressable Nanoscale Bioreactors. *J. Am. Chem. Soc.* **2013**, *135*, 696–702.

(5) Jiang, Q.; Song, C.; Nangreave, J.; Liu, X.; Lin, L.; Qiu, D.; Wang, Z. G.; Zou, G.; Liang, X.; Yan, H.; *et al.* DNA Origami as a Carrier for Circumvention of Drug Resistance. *J. Am. Chem. Soc.* **2012**, *134*, 13396–13403.

(6) Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440*, 297–302.

(7) Linko, V.; Ora, A.; Kostiainen, M. A. Review DNA Nanostructures as Smart Drug-Delivery Vehicles and Molecular Devices. *Trends Biotechnol.* **2015**, *33*, 586–594.

(8) Zahid, M.; Kim, B.; Hussain, R.; Amin, R.; Park, S. H. DNA Nanotechnology: A Future Perspective. *Nanoscale Res. Lett.* **2013**, *8*, 119.

(9) Simmel, F. C. Three-Dimensional Nanoconstruction with DNA. *Angew. Chemie - Int. Ed.* **2008**, *47*, 5884–5887.

(10) Chandrasekaran, A. R.; Pushpanathan, M.; Halvorsen, K. Evolution of DNA Origami Scaffolds. *Mater. Lett.* **2016**, *170*, 221–224.

(11) El-Sagheer, A. H.; Brown, T. Synthesis and Polymerase Chain Reaction Amplification of DNA Strands Containing an Unnatural Triazole Linkage. *J. Am. Chem. Soc.* **2009**, *131*, 3958–3964.

(12) Ducani, C.; Kaul, C.; Moche, M.; Shih, W. M.; Högberg, B. Enzymatic Production of "monoclonal Stoichiometric" Single-Stranded DNA Oligonucleotides. *Nat. Methods* **2013**, *10*, 647–652.

(13) Chalmers, F.; Curnow, K. Scaling Up the Ligase Chain Reaction-Based Approach to Gene Synthesis. *Biotechniques* **2001**, *30*, 249–252.

(14) Stemmer, W. P. C.; Crameri, A.; Kim, D. H.; Brennan, T. M.; Heyneker, H. L. Single-Step Assembly of a Gene and Entire Plasmid from Large Numbers of Oligodeoxyribonucleotides. *Gene* **1995**, *164*, 49–53.

(15) Cassinelli, V.; Oberleitner, B.; Sobotta, J.; Nickels, P.; Grossi, G.; Kempter, S.; Frischmuth, T.; Liedl, T.; Manetto, A. One-Step Formation of "Chain-Armor"-Stabilized DNA Nanostructures. *Angew. Chemie Int. Ed.* **2015**, n/a-n/a.

(16) Cho, Y.; Lee, J. B.; Hong, J. Controlled Release of an Anti-Cancer Drug from DNA Structured Nano-Films. *Sci. Rep.* **2014**, *4*, 4078.

(17) Smith, D.; Schüller, V.; Engst, C.; Rädler, J.; Liedl, T. Nucleic Acid Nanostructures for Biomedical Applications. *Nanomedicine (Lond).* **2013**, *8*, 105–121.

(18) Vindigni, G.; Raniolo, S.; Ottaviani, A.; Falconi, M.; Franch, O.; Knudsen, B. R.; Desideri, A.; Biocca, S. Receptor-Mediated Entry of Pristine Octahedral DNA Nanocages in Mammalian Cells. *ACS Nano* **2016**, *10*, 5971–5979.

(19)   Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L. P.; *et al.* Self-Assembly of a Nanoscale DNA Box with a Controllable Lid. *Nature* **2009**, *459*, 73–76.

(20)   Sahin, U.; Kariko, K.; Tureci, O. MRNA-Based Therapeutics - Developing a New Class of Drugs. *Nat. Rev. Drug Discov.* **2014**, *13*, 759–780.

(21)   Matsui, A.; Uchida, S.; Ishii, T.; Itaka, K.; Kataoka, K. Messenger RNA-Based Therapeutics for the Treatment of Apoptosis-Associated Diseases. *Sci. Rep.* **2015**, *5*, 1–10.

(22)   Meng, Z.; O'Keeffe-Ahern, J.; Lyu, J.; Pierucci, L.; Zhou, D.; Wang, W. A New Developing Class of Gene Delivery: Messenger RNA-Based Therapeutics. *Biomater. Sci.* **2017**, *5*, 2381–2392.

(23)   Profile, S. E. E. MRNA : Fulfilling the Promise of Gene Therapy. **2015**.

(24)   Derosa, F.; Smith, L.; Shen, Y.; Huang, Y.; Pan, J.; Xie, H.; Yahalom, B.; Heartlein, M. W. Improved Efficacy in a Fabry Disease Model Using a Systemic MRNA Liver Depot System as Compared to Enzyme Replacement Therapy. *Mol. Ther.* **2019**, *27*, 878–889.

(25)   Tavernier, G.; Andries, O.; Demeester, J.; Sanders, N. N.; De Smedt, S. C.; Rejman, J. MRNA as Gene Therapeutic: How to Control Protein Expression. *J. Control. Release* **2011**, *150*, 238–247.

(26)   Zhong, Z.; Mc Cafferty, S.; Combes, F.; Huysmans, H.; De Temmerman, J.; Gitsels, A.; Vanrompay, D.; Portela Catani, J.; Sanders, N. N. MRNA Therapeutics Deliver a Hopeful Message. *Nano Today* **2018**, *23*, 16–39.

(27)   Leonhardt, C.; Schwake, G.; Stögbauer, T. R.; Rappl, S.; Kuhr, J. T.; Ligon, T. S.; Rädler, J. O. Single-Cell MRNA Transfection Studies: Delivery, Kinetics and Statistics by Numbers. *Nanomedicine Nanotechnology, Biol. Med.* **2014**, *10*, 679–688.

(28)   Kaczmarek, J. C.; Kowalski, P. S.; Anderson, D. G. Advances in the Delivery of RNA Therapeutics: From Concept to Clinical Reality. *Genome Med.* **2017**, *9*, 60.

(29)   Dowdy, S. F. Overcoming Cellular Barriers for RNA Therapeutics. *Nat. Biotechnol.* **2017**, *35*, 222–229.

(30)   Sabnis, S.; Kumarasinghe, E. S.; Salerno, T.; Mihai, C.; Ketova, T.; Senn, J. J.; Lynn, A.; Bulychev, A.; Mcfadyen, I.; Chan, J.; *et al.* A Novel Amino Lipid Series for MRNA Delivery : Improved Endosomal Escape and Sustained Pharmacology and Safety in Non-Human Primates. *Mol. Ther.* **2018**, *26*, 1509–1519.

(31)   Lorenz, C.; Fotin-mleczek, M.; Roth, G.; Becker, C.; Dam, T. C.; Verdurmen, W. P. R.; Brock, R.; Probst, J.; Lorenz, C.; Fotin-mleczek, M.; *et al.* A n d e s i o s c i e n c e o n o t d i s t r i b u t E. **2011**, *6286*.

(32)   Bhosle, S. M.; Loomis, K. H.; Kirschman, J. L.; Blanchard, E. L.; Vanover, D. A.; Zurla, C.; Habrant, D.; Edwards, D.; Baumhof, P.; Pitard, B.; *et al.* Unifying in Vitro and in Vivo IVT MRNA Expression Discrepancies in Skeletal Muscle via Mechanotransduction. *Biomaterials* **2018**, *159*, 189–203.

(33)   Willoughby, J. L. S.; Chan, A.; Butler, J. S.; Nair, J. K.; Racie, T.; Shulga-morskaya, s.; Nguyen, T.; Qian, K.; Yucius, K.; Charisse, K.; *et al.* Evaluation of GalNAc-SiRNA Conjugate Activity in Pre-Clinical Animal Models with Reduced Asialoglycoprotein Receptor Expression. **2018**, *26*, 105–114.

(34)   Kwon, S. Single-Molecule Fluorescence in Situ Hybridization : Quantitative Imaging of Single RNA Molecules. **2013**, *6696*.

(35)   Hesse, S.; Manetto, A.; Cassinelli, V.; Fuchs, J.; Ma, L.; Raddaoui, N.; Houben, A. Fluorescent Labelling of in Situ Hybridisation Probes through the Copper-Catalysed Azide-Alkyne Cycloaddition Reaction. *Chromosom. Res.* **2016**, *24*, 299–307.

(36)   Henegariu, O.; Bray-Ward, P.; Ward, D. C. Custom Fluorescent-Nucleotide Synthesis as an Alternative Method for Nucleic Acid Labeling. *Nat. Biotechnol.* **2000**, *18*, 345–348.

(37)   Baerlocher, G. M.; Mak, J.; Tien, T.; Lansdorp, P. M. Telomere Length Measurement by Fluorescence

in Situ Hybridization and Flow Cytometry: Tips and Pitfalls. *Cytometry* **2002**, *47*, 89–99.

(38)   Femino, A. M.; Fay, F. S.; Fogarty, K.; Singer, R. H. Visualization of Single RNA Transcripts in Situ. *Science (80-.).* **1998**, *280*, 585 LP – 590.

(39)   Gaspar, I.; Wippich, F.; Ephrussi, A. Enzymatic Production of Single-Molecule FISH and RNA Capture Probes. **2017**, 1582–1591.

(40)   Raj, A.; van den Bogaard, P.; Rifkin, S. a; van Oudenaarden, A.; Tyagi, S. Imaging Individual MRNA Molecules Using Multiple Singly Labeled Probes. *Nat. Methods* **2008**, *5*, 877–879.

(41)   Wilson, J. L. Biochemistry; Third Edition (Stryer, Lubert). *J. Chem. Educ.* **1988**, *65*, A337.

(42)   WATSON, J. D.; CRICK, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.

(43)   Michelson, A. M.; Alexander, S.; Todd, R. X x x *. **1954**, 808–815.

(44)   Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Crystal Structure Analysis of a Complete Turn of B-DNA. *Nature* **1980**, *287*, 755–758.

(45)   RK, S.; S, S.; F, F.; KB, M.; GT, H.; HA, E.; N, A. Enzymatic Amplification of Beta-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. *Science (80-.).* **1985**, *230*, 1350–1354.

(46)   Caruthers, M. H.; Barone, A. D.; Beaucage, S. L.; Dodds, E. F.; Fisher, E. F.; McBride, L. J.; Matteucci, M.; Stabinsky, Z.; Tang, J. Y. Chemical Synthesis of Deoxyoligonucleotides by the Phosphoramidite Method. *Methods Enzymol.* **1987**, *154*, 287–313.

(47)   Holtkamp, S.; Kreiter, S.; Selmi, A.; Simon, P.; Koslowski, M.; Huber, C. Modification of Antigen-Encoding RNA Increases Stability , Translational Efficacy , and T-Cell Stimulatory Capacity of Dendritic Cells. **2019**, *108*, 4009–4018.

(48)   Jemielity, J.; Fowler, T.; Zuberek, J.; Stepinski, J.; Lewdorowicz, M.; Niedzwiecka, A.; Stolarski, R.; Darzynkiewicz, E.; Rhoads, R. E. Novel " Anti-Reverse " Cap Analogs with Superior Translational Properties. **2003**, 1108–1122.

(49)   Pardi, N.; Hogan, M. J.; Pelc, R. S.; Muramatsu, H.; Andersen, H.; DeMaso, C. R.; Dowd, K. A.; Sutherland, L. L.; Scearce, R. M.; Parks, R.; *et al.* Zika Virus Protection by a Single Low-Dose Nucleoside-Modified MRNA Vaccination. *Nature* **2017**, *543*, 248–251.

(50)   Richner, J. M.; Himansu, S.; Dowd, K. A.; Butler, S. L.; Salazar, V.; Fox, J. M.; Julander, J. G.; Tang, W. W.; Shresta, S.; Pierson, T. C.; *et al.* Modified MRNA Vaccines Protect against Zika Virus Infection. *Cell* **2017**, 1–12.

(51)   Magadum, A.; Kaur, K.; Zangi, L. MRNA-Based Protein Replacement Therapy for the Heart. *Mol. Ther.* **2019**, *27*, 785–793.

(52)   Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie - Int. Ed.* **2001**, *40*, 2004–2021.

(53)   Meldal, M.; Tomøe, C. W. Cu-Catalyzed Azide - Alkyne Cycloaddition. *Chem. Rev.* **2008**, *108*, 2952–3015.

(54)   Rostovtsev, V. V; Green, L. G.; Fokin, V. V; Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chemie - Int. Ed.* **2002**, *41*, 2596–2599.

(55)   Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on Solid Phase:-Triazoles by Regiospecific Copper (I)-Catalyzed 1, 3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **2002**, *67*, 3057–3064.

(56)   Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. Copper(I)-

Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates. *J. Am. Chem. Soc.* **2005**, *127*, 210–216.

(57)   Soriano Del Amo, D.; Wang, W.; Jiang, H.; Besanceney, C.; Yan, A. C.; Levy, M.; Liu, Y.; Marlow, F. L.; Wu, P. Biocompatible Copper(I) Catalysts for in Vivo Imaging of Glycans. *J. Am. Chem. Soc.* **2010**, *132*, 16893–16899.

(58)   Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano Del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. Increasing the Efficacy of Bioorthogonal Click Reactions for Bioconjugation: A Comparative Study. *Angew. Chemie - Int. Ed.* **2011**, *50*, 8051–8056.

(59)   Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. Copper-Free Click Chemistry for Dynamic in Vivo Imaging. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16793–16797.

(60)   van den Bosch, S. M.; Rossin, R.; Renart Verkerk, P.; ten Hoeve, W.; Janssen, H. M.; Lub, J.; Robillard, M. S. Evaluation of Strained Alkynes for Cu-Free Click Reaction in Live Mice. *Nucl. Med. Biol.* **2013**, *40*, 415–423.

(61)   Agard, N. J.; Prescher, J. A. J. a.; Bertozzi, C. R.; Nicholas, J. SI. A Strain-Promoted [3+2] Azide-Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

(62)   Holub, J. M.; Garabedian, M. J.; Kirshenbaum, K. Peptoids on Steroids: Precise Multivalent Estradiol-Peptidomimetic Conjugates Generated via Azide-Alkyne [3 + 2] Cycloaddition Reactions. *QSAR Comb. Sci.* **2007**, *26*, 1175–1180.

(63)   Hammond, D. M.; Manetto, A.; Gierlich, J.; Azov, V. A.; Gramlich, P. M. E.; Burley, G. A.; Maul, M.; Carell, T. DNA Photography: An Ultrasensitive DNA-Detection Method Based on Photographic Techniques. *Angew. Chemie - Int. Ed.* **2007**, *46*, 4184–4187.

(64)   Chan, T. R.; Fokin, V. V. Polymer-Supported Copper(I) Catalysts for the Experimentally Simplified Azide-Alkyne Cycloaddition. *QSAR Comb. Sci.* **2007**, *26*, 1274–1279.

(65)   Chittaboina, S.; Xie, F.; Wang, Q. One-Pot Synthesis of Triazole-Linked Glycoconjugates. *Tetrahedron Lett.* **2005**, *46*, 2331–2336.

(66)   Gramlich, P. M. E.; Wirges, C. T.; Manetto, A.; Carell, T. Postsynthetic DNA Modification through the Copper-Catalyzed Azide-Alkyne Cycloaddition Reaction. *Angew. Chemie - Int. Ed.* **2008**, *47*, 8350–8358.

(67)   Gierlich, J.; Burley, G. a.; Gramlich, P. M. E.; Hammond, D. M.; Carell, T. Click Chemistry as a Reliable Method for the High-Density Functionalisation of Alkyne-Modified Oligodeoxyribonucleotides. *Org. Lett.* **2006**, *8*, 3639–3642.

(68)   Ducani, C.; Kaul, C.; Moche, M.; Shih, W. M.; Högberg, B. Enzymatic Production of "monoclonal Stoichiometric" Single-Stranded DNA Oligonucleotides. *Nat. Methods* **2013**, *10*, 647–652.

(69)   Xiong, A.-S. A Simple, Rapid, High-Fidelity and Cost-Effective PCR-Based Two-Step DNA Synthesis Method for Long Gene Sequences. *Nucleic Acids Res.* **2004**, *32*, e98–e98.

(70)   Hughes, R. A.; Miklos, A. E.; Ellington, A. D. *Gene Synthesis: Methods and Applications*; 2011; Vol. 498.

(71)   Perrault, S. D.; Shih, W. M. Virus-Inspired Membrane Encapsulation of DNA Nanostructures to Achieve in Vivo Stability. *ACS Nano* **2014**, *8*, 5132–5140.

(72)   Ranallo, S.; Amodio, A.; Idili, A.; Porchetta, A.; Ricci, F. Electronic Control of DNA-Based Nanoswitches and Nanodevices. *Chem. Sci.* **2015**, *00*, 1–6.

(73)   Evans, C. G.; Winfree, E. Chem Soc Rev Physical Principles for DNA Tile Self-Assembly. **2017**, 3808–3829.

(74) Kallen, K.-J. J.; Theß, A.; Thess, A. A Development That May Evolve into a Revolution in Medicine: MRNA as the Basis for Novel, Nucleotide-Based Vaccines and Drugs. *Ther. Adv. vaccines* **2014**, *2*, 10–31.

(75) Kimoto, M.; Kawai, R.; Mitsui, T.; Harada, Y.; Sato, A.; Yokoyama, S.; Hirao, I. Site-Specific Incorporation of Fluorescent Probes into RNA by Specific Transcription Using Unnatural Base Pairs. *Nucleic Acids Symp. Ser. (Oxf).* **2005**, 287–288.

(76) Seo, Y. J.; Malyshev, D. A.; Lavergne, T.; Ordoukhanian, P.; Romesberg, F. E. Site-Specific Labeling of DNA and RNA Using an Efficiently Replicated and Transcribed Class of Unnatural Base Pairs. *J. Am. Chem. Soc.* **2011**, *133*, 19878–19888.

(77) Svitkin, Y. V.; Cheng, Y. M.; Chakraborty, T.; Presnyak, V.; John, M.; Sonenberg, N. N1-Methyl-Pseudouridine in MRNA Enhances Translation through EIF2α-Dependent and Independent Mechanisms by Increasing Ribosome Density. *Nucleic Acids Res.* **2017**, *45*, 6023–6036.

(78) Li, B.; Luo, X.; Dong, Y. Effects of Chemically Modified Messenger RNA on Protein Expression. *Bioconjug. Chem.* **2016**, *27*, 849–853.

(79) Yamada, N. A.; Rector, L. S.; Tsang, P.; Carr, E.; Scheffer, A.; Sederberg, M. C.; Aston, M. E.; Ach, R. A.; Tsalenko, A.; Sampas, N.; *et al.* Visualization of Fine-Scale Genomic Structure by Oligonucleotide-Based High-Resolution FISH. *Cytogenet. Genome Res.* **2011**, *132*, 248–254.

(80) Gall, J. G. The Origin of in Situ Hybridization - A Personal History. *Methods* **2016**, *98*, 4–9.

(81) Rouhanifard, S. H.; Mellis, I. A.; Dunagin, M.; Bayatpour, S.; Jiang, C. L.; Dardani, I.; Symmons, O.; Emert, B.; Torre, E.; Cote, A.; *et al.* ClampFISH Detects Individual Nucleic Acid Molecules Using Click Chemistry–Based Amplification. *Nat. Biotechnol.* **2019**, *37*, 84–94.

(82) Grant, G. P. G.; Qin, P. Z. A Facile Method for Attaching Nitroxide Spin Labels at the 5' Terminus of Nucleic Acids. *Nucleic Acids Res.* **2007**, *35*, e77–e77.

(83) Schmitz, G. G.; Walter, T.; Seibl, R.; Kessler, C. Nonradioactive Labeling of Oligonucleotides in Vitro with the Hapten Digoxigenin by Tailing with Terminal Transferase. *Anal. Biochem.* **1991**, *192*, 222–231.