



Using DNA Origami As A Protein Scaffold To Enhance Metabolic Production *In Vivo*

McMaster BioMod Team

Abstract

Self-assembling DNA nanostructures hold great promise in controlling the spatial organization and orientation of enzymatic complexes. However, their usage *in vivo* has been limited by the harsh conditions needed for assembly. We propose a modular toolkit to allow both researchers and future International Genetically Engineered Machine (iGEM) participants to easily build DNA nanostructures that can be decorated with proteins in a wide range of cells. Our proposed solution uses a novel reverse-transcriptase based method to synthesize necessary double-stranded DNA (dsDNA) in the cellular environment and then uses the sequence-specific recognition properties of transcription activator-like (TAL) domains of effector proteins to assemble them into complex 2D or 3D shapes. Proteins of interest could then be fused with the TAL domains to bring them into proximity of each other and control their spatial orientation. The proposed technology can be used to increase the efficiency of complex enzyme cascades or to create chimeric signalling pathways by linking signals to downstream messengers. The modular nature allows the toolkit to be easily integrated into any synthetic biology project and could drastically increase the efficiency of existing enzymatic pathways.

Introduction

DNA Nanostructures

In 2006, Paul Rothemund introduced a novel method to create “DNA origami” through directed folding of a long single-stranded DNA (ssDNA) “scaffold” into arbitrary shapes using short ssDNA segments as “staples”¹. This method revolutionized DNA nanotechnology by offering fine control over the creation of complex shapes. To date, various groups have successfully applied modifications of this technique to create both 2D and 3D shapes, such as various icons, monoliths, and nanoflasks^{1–4}. These nanostructures have seemingly limitless utility in a variety of applications. As drug delivery vehicles, aptamer-gated DNA nano-robots loaded with antibody fragments can change their configuration to deliver cargo to specific target cells⁵. As biosensors, DNA nanostructures have successfully provided readouts of enzymatic activity and pH^{6,7}. The functionalization of DNA nanostructures can be expanded even further through creating DNA-protein arrangements.

Despite significant methodological advances, very few groups contend with the challenge of producing DNA nanostructures *in vivo*, preferring to create and apply such structures extracellularly. Developing DNA origami, decorated with functional proteins/small molecules that can be produced within living organisms would expand the potential for DNA nanostructures to be applied in a wide range of biological applications as discussed below.



In Vivo DNA Nanostructures: Scaffolds as a Proof-of-Concept

As a proof-of-concept application, we will use our toolkit to create an *in vivo* DNA nanostructure scaffold for enzymatic reactions. By spatially fixing enzymes relative to one another, cellular scaffolds increase the efficiency of signaling cascades and biosynthetic pathways by shuttling substrates, stabilizing reaction intermediates and reducing diffusion time^{8,9}. While *ex vivo* methods for engineering synthetic scaffolds have been well developed, there is still a growing need to express them *in vivo* to produce “cell-factories”¹⁰. However, existing nanostructure strategies currently being researched contains innate flaws that lower their potential when expressed in living bacteria, as outlined in Table 1. In brief, the main challenges currently faced by the *in vivo* synthesis of engineered scaffolds are due to (1) instability of backbone structures and (2) unpredictability of scaffold structure upon expression *in vitro*. Ideally, DNA nanostructures expressed *in vivo* offer unique advantages that tackle both of these issues, as they are not only highly stable structures that maintain rigidity, but also modular in the possible shapes, which allows for optimized intermediate diffusion¹.

Strategy	Challenges
Fusion Proteins With Linkers	Inconsistent expression and folding as number of enzymes increase ^{11,12}
Protein Scaffolds	Inconsistent expression and folding as number of enzymes increase ^{13–16}
ssDNA	Unpredictable shape of scaffold due to lack of rigidity ^{10,17}
Plasmid DNA	Supercoiling of plasmids reduce enzymatic function ^{18,19}
RNA	Chemical and thermodynamic instability of RNA ^{20,21}

Table 1: Overview of the limitations of existing *in vivo* scaffolding technologies.

Current DNA nanostructure technology is based on Rothemund’s *ex vivo* assembly method, which requires harsh salt solutions and high temperatures for assembly⁴. Moreover, many groups chemically synthesize their DNA strands, which can come at a high cost for large-scale applications. Therefore, implementing DNA nanostructures that can be synthesized and self-assembled within living cells is a unique challenge in synthetic biology. At its core, our proposal is thus a modular *in vivo* assembly method for DNA origami. We will adapt a reverse-transcription based technique to achieve rapid synthesis of DNA strands in *Escherichia coli* (*E. coli*) using a plasmid encoding the DNA origami subunit sequences⁵. We then propose to use TAL-specific to the DNA sequences in these strands for two purposes (1) to hold these dsDNA strands together into a DNA nanostructure and (2) to create fusion proteins to enzymes or functional groups to allow the origami structure to serve as a scaffold. After assembly, the assembled DNA nanostructure is not limited to serving scaffolding purposes, and could be programmed to match existing DNA nanostructure biological applications such as creating molecular machines^{6,7}, DNA computation⁸, and even switchable structures that resemble riboswitches⁹. Critically, our system will be completely modular. This genetic circuit will be cloned in an easy expression system that allows future iGEM teams and researchers to adapt the technology for their specific uses.

The extensive utility of *in vivo* DNA nanostructure assembly can be seen by the implications of creating DNA scaffolds. As proteins scaffolds act naturally as hubs to aggregate signals and then activating specific downstream pathways, a scaffold could be used to alter responses to signalling inducers in a similar manner to how promoters could react to different inputs²². Compared to the transcription factor-promoter



system, a modular scaffolding-based signalling pathway has several advantages such as rapid response to inducer, ready amplification to signal, responsive shut-off, and ease of linking input to phenotype. For example, Lim and colleagues used a protein scaffold and successfully linked mating to the high-osmolarity stress MAPK pathway in *S.cerevisiae*²³. Our modular system would readily enable the establishment and testing of these chimeric signalling pathways. Furthermore, a modular DNA scaffold could also assemble metabolic and synthetic pathways into complexes using the idea of enforced proximity. Previously, scaffolding has been used to prevent loss of low-abundance intermediates, to increase the effective concentration of reactants and to control metabolic flux by controlling the scaffold assembly, as shown in a carbamoyl phosphate synthase, polyketide synthase, and tryptophan synthase pathway model^{24–26}. It was shown that scaffolding could strikingly increase the production of mevalonate pathway by nearly 100-fold¹⁶. Our modular DNA scaffolding system could be used to assemble theoretically any enzymes complexes in complex pathways and improve yield. Thus, it would allow both academic and industrial researchers to optimize and enable the biosynthesis of important compounds. Theoretically, our system could be readily adapted to improve the efficiency of any enzymatic pathways by offering a scaffolding platform to increase yields in processes such as biosynthesis, signal transduction, and metabolism¹⁰.

Rationale

While many groups have explored the uses of DNA nanostructures assembled *ex vivo*, little work has been done on DNA origami assembled *in vivo*. Present *ex vivo* assembly methods require high amounts of chemically synthesized DNA, high salt concentrations and high temperatures to create nanostructures. These conditions are not suitable for living cells, so a new method must be developed for DNA nanostructure assembly *in vivo*. Only by having such a tool readily available can fundamental applications like: DNA scaffolds, molecular machines and riboswitch-like structures be produced inside living cells to show their biological utility.

Aims

In order to expand the toolkit of synthetic biologists with respect to DNA origami in this project, we propose to (1) develop a method for the *in vivo* assembly of dsDNA nanostructures in *E. coli* and (2) create a DNA origami scaffold for enzymes to improve pathway efficiency as a proof-of-concept

Approaches & Methodology

Developing methods for *in vivo* assembly of dsDNA nanostructures in *E. coli*

In this project, we propose harnessing the ability of Human Immunodeficiency Virus Reverse-Transcriptase (HIVRT) to produce dsDNA from non-coding RNA (ncRNA) encoded on a plasmid. Counterintuitively, making enough dsDNA for *in vivo* DNA nanostructure assembly has proven to be challenging. We will repurpose the method that Elbaz et al. have utilized to make ssDNA *in vivo*²⁷. A t-RNA^{Lys} genetic part, which both recruits HIVRT and acts as a transcriptional terminator, will be introduced in front of



the sequence encoding the DNA nanostructure strand. Following the transcription of the RNA encoding the DNA nanostructure subunits, t-RNA^{Lys} will recruit the two HIVRT subunits and the Monomeric Murine Leukemia Reverse Transcriptase (MLRT) to reverse transcribe the RNA back into dsDNA. These tools allow the rapid synthesis of large amounts of dsDNA *in vivo*.

The dsDNA will be folded using TAL domains. The DNA nanostructures encode sites specific to the domain of TAL effector proteins originally discovered in plants. Praetorius and Dietz have used the TAL proteins to form the staples of the DNA nanostructure and thus was able to mold the dsDNA into complex 3D structures²⁸. Thus, by coexpressing the TAL proteins in conjunction with HIVRT, MLRT, and the plasmid encoding the dsDNA subunits, we will be able to assemble complex nanostructures *in vivo*.

We will first validate the key design features of our project *in vitro* as described by Han and colleagues²⁹. *In vitro* replicability will be established by PCR amplification of both strands of desired DNA nanostructures. Upon synthesis, the PCR products will be annealed to form their dsDNA sequences, which can then be imaged under Atomic Force Microscopy (AFM) after exposure to a concoction of HIVRT, MLRT and several TAL proteins. We expect that PCR amplification will produce dsDNA structures of identical morphology.

To test our system *in vivo*, we will transform our plasmid-based expression system into *E. coli* Cells will then be lysed and dsDNA will be purified. Native gel electrophoresis will be employed on extracted dsDNA alongside those generated *in vitro*. We expected similar pre-folding and post-folding migration patterns for the dsDNA generated *in vitro* and in living cells. This will be validated by AFM imaging to confirm the expected morphology.

Modelling for modularity in the dsDNA origami shape

Distinct from usual DNA nanostructure projects, our approach using pre-annealed dsDNA and TAL domains eliminate the need to model DNA folding and base-pair interactions. Instead, the *in silico* guided design efforts will need to focus on the structural considerations of staple proteins binding to the dsDNA template.²⁸ utilized electron density projections modelling to determine desired conformations of the various staple proteins on the dsDNA and predict which of these conformations are optimal for protein-DNA binding. Further, Praetorius and Dietz also provides a good resource for understanding how to model staple protein binding to the dsDNA template²⁸. Thus, we will combine these existing tools to guide the design of our DNA nanostructure.

Creating a DNA origami scaffold for enzymes to improve pathway efficiency as a proof-of-concept

We propose creating a dsDNA scaffold upon which enzymes can be assembled into a particular three-dimensional arrangement to enhance enzymatic pathway efficiency. Our DNA nanostructure is assembled from pieces of dsDNA held together by TAL effectors that act as staple proteins. Taking advantage of these proteins being present on the DNA scaffold, we will create fusion proteins between the TAL effectors and our enzymes of interest. We plan on creating this fusion by drawing on the methods of Praetorius and



Dietz, who tested the suitability of eleven different peptide linkers to fuse two TAL proteins and create a double-TAL staple²⁸. These linkers were based upon the prototypical flexible glycine-serine, or “GS” linker, from which they found a SAGA (Ser-Ala-Gly-Ser) linker to be ideal. They further used this short flexible linker to create an N-terminal fusion between a TAL protein and GFP. We will therefore link each enzyme of interest to the N-terminal of the TAL at the desired location on the DNA nanostructure using a SAGA linker. Since the optimal linker for GFP and a functional enzyme may differ, we will compare the enzymatic activities between the purified and fusion proteins to ensure minimal change in activity. In the event that enzymatic function is affected due to the short SAGA linker, we will test longer flexible linkers made of glycine, serine and alanine to ensure optimal length from the DNA nanostructure³⁰

Modelling for Enzymatic Scaffolding

We will produce a model that can accurately predict the enzyme-substrate kinetics for the overall DNA scaffold in order to optimise its design. The random nature of molecules moving around and interacting inside of a cell is best captured using a stochastic method, for which several parameters must be considered^{31–33}. These include proximity (how close subsequent cascade enzymes are to each other), spatial orientation (arranging geometry of active sites), substrate channelling (what path the substrate will likely take from enzyme to enzyme, and what protein bridges it may use for the purpose)^{31,32}. Other considerations unique to DNA scaffolds are pH changes in the active site and the negative charge of the DNA³³. We also aim to specifically quantify to what extent the DNA scaffold increases the efficiency or yield of the pathway under consideration. By bringing all these parameters together, we hope to create a model that can: (1) determine the optimal shape among a subset of DNA origami nanostructures for a given enzymatic pathway and (2) predict the extent to which a particular scaffold will increase pathway efficiency.

Resources & Personnel

Recent work on DNA nanostructures has repurposed DNA to increase the efficiency of enzymatic function *ex vivo*. We have therefore assembled a multidisciplinary team of genetic engineers, biochemists, physicists, chemical engineers, and bioinformaticians to address the critical gap to bring this concept into an *in vivo* setting. The McMaster BioMod team is ideally positioned to carry out this project. In 2017, we participated in the iGEM competition in Boston, MA, and captured a bronze medal. The team members all have extensive lab experience in many of the techniques needed for the current project, such as genetic cloning, protein and plasmid purification, mathematical modelling, etc. We are currently supervised by Dr. Zeinab Hosseini-Doust in the Department of Chemical Engineering at McMaster University. Collaborations with other faculty members of various departments at McMaster University has given us many insights in their special field of expertise. Together, this team is poised to make a transformative contribution in this exciting area.



Potential Limitations

We recognize that the extensive cloning involved in this project could pose problems that could delay the progress of the project. If the proposed restriction enzyme-based cloning methods do not successfully form the desired lentiviral construct, we will use an alternative approach, such as golden gate cloning, gibson assembly, or direct plasmid synthesis. As for potential difficulties with the sensitive microscopy techniques like AFM to validate this project, we will be working together with individuals at McMaster University who have an expertise in structural biology and imaging.

We also recognize that there are likely pathways that cannot be optimized with our system due to inherent incompatibilities. For example, enzymes with charge-sensitive active or allosteric sites that are inactivated by proximity to the DNA scaffold are contraindicated with our approach³³. Due to constraints on both the size of plasmids that can be expressed as the backbone and the number of plasmids that can express staple proteins, there are limits to the complexity of the DNA nanostructure and the number of enzymes attached to it. Tuning the expression of these individual parts to ensure an optimal concentration of each for scaffold formation is something else that needs to be carefully considered and tested. Nonetheless, we believe that there is clear value to the addition of being able to produce dsDNA scaffolds *in vivo*.

Conclusion

We propose developing a novel method for the *in vivo* assembly of dsDNA nanostructures to allow the translation of developing DNA origami applications such as scaffolding and DNA nanorobots to the setting of live cells. To overcome current production bottlenecks of the harsh assembly conditions and the high cost of large-scale DNA synthesis, we utilize (1) a novel reverse transcriptase-based method to generate large amounts of dsDNA *in vivo* and (2) fusion proteins of TAL effectors and enzymes to create the desired shape while forming an enzyme scaffold. Therefore, our proposal allows the production of dsDNA origami in a scalable and cost-effective manner. Importantly, we expect this method to give the *in vivo* assembly of dsDNA nanostructures sufficient modularity to be easily used by future researchers and compatible with existing design frameworks. Thus, the proposed technology should bridge the gap that limits the application of DNA nanostructures to living biological systems.

References

- [1] Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
- [2] Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009).
- [3] Han, D. *et al.* DNA origami with complex curvatures in three-dimensional space. *Science* **332**, 342–346 (2011).
- [4] Chandrasekaran, A. R., Anderson, N., Kizer, M., Halvorsen, K. & Wang, X. Beyond the Fold: Emerging Biological Applications of DNA Origami. *Chembiochem* **17**, 1081–1089 (2016).



- [5] Douglas, S. M., Bachelet, I. & Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335**, 831–834 (2012).
- [6] Tintore, M., Gallego, I., Manning, B., Eritja, R. & Fabrega, C. DNA origami as a DNA repair nanosensor at the single-molecule level. *Angew. Chem. Int. Ed. Engl.* **52**, 7747–7750 (2013).
- [7] Kuzuya, A. *et al.* Nanomechanical DNA origami pH sensors. *Sensors (Basel)* **14**, 19329–19335 (2014).
- [8] Good, M. C., Zalatan, J. G. & Lim, W. A. Scaffold proteins: hubs for controlling the flow of cellular information. *Science* **332**, 680–686 (2011).
- [9] Lee, M. J. *et al.* Engineered synthetic scaffolds for organizing proteins within the bacterial cytoplasm. *Nat. Chem. Biol.* **14**, 142–147 (2018).
- [10] Chen, R. *et al.* Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Curr. Opin. Biotechnol.* **28**, 59–68 (2014).
- [11] Zhang, Y. *et al.* Using unnatural protein fusions to engineer resveratrol biosynthesis in yeast and Mammalian cells. *J. Am. Chem. Soc.* **128**, 13030–13031 (2006).
- [12] Albertsen, L. *et al.* Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. *Appl. Environ. Microbiol.* **77**, 1033–1040 (2011).
- [13] Tsai, S. L., Oh, J., Singh, S., Chen, R. & Chen, W. Functional assembly of minicellulosomes on the *Saccharomyces cerevisiae* cell surface for cellulose hydrolysis and ethanol production. *Appl. Environ. Microbiol.* **75**, 6087–6093 (2009).
- [14] Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C. & Shively, J. M. Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nat. Rev. Microbiol.* **6**, 681–691 (2008).
- [15] Choudhary, S., Quin, M. B., Sanders, M. A., Johnson, E. T. & Schmidt-Dannert, C. Engineered protein nano-compartments for targeted enzyme localization. *PLoS ONE* **7**, e33342 (2012).
- [16] Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753–759 (2009).
- [17] Wilner, O. I., Shimron, S., Weizmann, Y., Wang, Z. G. & Willner, I. Self-assembly of enzymes on DNA scaffolds: en route to biocatalytic cascades and the synthesis of metallic nanowires. *Nano Lett.* **9**, 2040–2043 (2009).
- [18] Conrado, R. J. *et al.* DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic Acids Res.* **40**, 1879–1889 (2012).
- [19] Higgins, N. P. & Vologodskii, A. V. Topological Behavior of Plasmid DNA. *Microbiol Spectr* **3** (2015).
- [20] Guo, P., Haque, F., Hallahan, B., Reif, R. & Li, H. Uniqueness, advantages, challenges, solutions, and perspectives in therapeutics applying RNA nanotechnology. *Nucleic Acid Ther* **22**, 226–245 (2012).



- [21] Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333**, 470–474 (2011).
- [22] Bhattacharyya, R. P., Remenyi, A., Yeh, B. J. & Lim, W. A. Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* **75**, 655–680 (2006).
- [23] Park, S. H., Zarrinpar, A. & Lim, W. A. Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299**, 1061–1064 (2003).
- [24] Miles, E. W., Rhee, S. & Davies, D. R. The molecular basis of substrate channeling. *J. Biol. Chem.* **274**, 12193–12196 (1999).
- [25] Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M. & Rayment, I. Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product. *Biochemistry* **36**, 6305–6316 (1997).
- [26] An, S., Kumar, R., Sheets, E. D. & Benkovic, S. J. Reversible compartmentalization of de novo purine biosynthetic complexes in living cells. *Science* **320**, 103–106 (2008).
- [27] Elbaz, J., Yin, P. & Voigt, C. A. Genetic encoding of DNA nanostructures and their self-assembly in living bacteria. *Nat Commun* **7**, 11179 (2016).
- [28] Praetorius, F. & Dietz, H. Self-assembly of genetically encoded DNA-protein hybrid nanoscale shapes. *Science* **355** (2017).
- [29] Han, D. *et al.* Single-stranded DNA and RNA origami. *Science* **358** (2017).
- [30] Chen, X., Zaro, J. L. & Shen, W. C. Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* **65**, 1357–1369 (2013).
- [31] Fu, J. *et al.* Assembly of multienzyme complexes on DNA nanostructures. *Nat Protoc* **11**, 2243–2273 (2016).
- [32] Fu, J., Liu, M., Liu, Y., Woodbury, N. W. & Yan, H. Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures. *J. Am. Chem. Soc.* **134**, 5516–5519 (2012).
- [33] Zhang, Y., Tsitkov, S. & Hess, H. Proximity does not contribute to activity enhancement in the glucose oxidase–horseradish peroxidase cascade. *Nature Communications* **7** (2016).