Designing 3D Wireframe DNA Nanoparticles for Programmable Innate Immune Activation

by

Rebecca R. Du

B.S. Biological Engineering
California Institute of Technology, 2016

Submitted to the Department of Biological Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biological Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2022

Signature of Author: _____
Department of Biological Engineering
May 13, 2022

Certified by: _____
Mark Bathe
Professor of Biological Engineering
Thesis Supervisor

Accepted by: _____
Katherina Ribbeck
Associate Professor of Biological Engineering
Graduate Program Chair

Designing 3D Wireframe DNA Nanoparticles for Programmable Innate Immune Activation

by

Rebecca R. Du

Submitted to the Department of Biological Engineering
on May 13, 2022, in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Biological Engineering

Abstract

DNA nanotechnology harnesses the predictability and specificity of canonical base pairing interactions to enable the synthesis of complex DNA nanodevices capable of interacting with their environment in a controllable manner. Wireframe DNA origami are a class of highly customizable DNA nanoparticles that can be folded into near-arbitrary 2D or 3D geometries upon which ligands can be attached and organized with nanoscale precision. Due to their versatility, DNA origami have been designed for a wide range of applications in a variety of fields ranging from chemistry to computer science, and there is substantial interest in harnessing DNA nanostructures for therapeutic applications. However, an understanding of how the immune system responds to wireframe DNA nanostructures is currently lacking. Furthermore, it is unclear how controllable design parameters such as scaffold sequence, nanoparticle geometry, or ligand organization influence immune activation.

In my thesis, I begin by describing a method for scalable bioproduction of circular single-stranded DNA that provides a high degree of sequence control and opens up opportunities for scaffold engineering. I demonstrate the design and production of two different scaffolds in both shaker flask and bioreactor setups, then validate the application of this method towards DNA nanotechnology by characterizing the successful folding of wireframe DNA origami from each scaffold. Using nanoparticles folded from one of these engineered scaffolds, I next investigate the molecular mechanisms by which wireframe DNA nanoparticles interact with innate immune receptors. I start by characterizing immunological recognition of unmodified wireframe nanoparticles and find that they induce Type I IFN expression primarily through cGAS-STING, while TLR9 is very minimally activated. I then enhance the ability of these nanoparticles to activate TLR9 by attaching discrete copy numbers of immunostimulatory CpG motifs at precise spatial positions and show that activation levels can be controllably tuned by changing the nanoscale organization of CpGs on the nanoparticle. Finally, I delve further into how design parameters such as nanoparticle geometry, inter-CpG distance, CpG copy number, and spatial presentation can be used to modulate innate immune activation. Overall, this work identifies specific nanoparticle properties that impact immunological recognition and sheds light on design principles that enable the production of immunomodulatory DNA origami.

Thesis Supervisor: Mark Bathe
Title: Professor of Biological Engineering

Acknowledgments

To my thesis advisor, Mark Bathe: thank you for your guidance throughout these past six years, both in your capacity as a scientific advisor and as a mentor. You have shown me the value of the 30,000 foot view, taught me to tackle challenges—both scientific and otherwise— meticulously and conscientiously, and impressed upon me the importance of thinking critically and evaluating ideas from different perspectives. You have provided me with a space to learn through failures and successes alike, and my time in your lab has been a truly irreplaceable experience.

To all those at MIT and beyond who have supported my scientific endeavors: thank you for sharing your time and expertise; without you, this work would have been quite different. To my committee members Darrell Irvine and Jim Collins: thank you for asking the important questions at committee meetings and for sharing not just scientific but also life advice. You have helped shape the trajectory of my graduate career and I'm very grateful to have had the opportunity to learn from you. To my collaborators at the NIH NCL, Marina Dobrovolskaia and Ed Cedrone: thank you for hosting me at the NCL. Even though I was only there for a brief period, it was a wonderful learning experience. Thank you as well for always being willing to make time for scientific discussions; they have helped much more than I have been able to express. To Jaime Cheah at the Koch HTS: thank you for getting me started on tissue culture and for going through so many rounds of troubleshooting with me. Although those experiments didn't work out, I nevertheless learned an enormous amount from you.

To everyone in the Bathe lab who I crossed paths with throughout my graduate years, whether we overlapped for several months or for several years: thank you for sharing your experiences and expertise with me, for the many insightful conversations, both scientific and otherwise, and for all your support throughout the past few years. My graduate experience would not have been the same without every single one of you. To my cohort, who accompanied me through all the challenges of the first two years of graduate school, from rotations to settling down in a lab, from taking written quals to passing our thesis proposals: thank you for your companionship and support! Those first two years were undeniably difficult, but we struggled through together, and not without making some lasting memories along the way. And to those in my cohort that have become some

of my closest friends here at MIT, particularly Molly and Joseph: I'm so glad to have ended up in the same lab as both of you, and to have gotten to tackle the challenges of grad school together. You've both been amazing sources of moral support, fantastic labmates and wonderful friends, and I couldn't have done it without you.

To my friends outside of MIT: even though we've all gone our separate ways, we've managed to keep in touch across time zones and continents and for that I am so unbelievably grateful. I appreciate every text, every video chat, every chaotic online gaming session, and every in-person meetup we've been able to coordinate. Each of you has made my graduate experience brighter with your presence in my life, and I'm looking forward to rejoining you all in the "real world" very soon.

To my family: thank you for your endless love, support, and encouragement. Thank you for supporting me through all my experiences in life that have led me to where I am today. Thank you for everything! I love you.

To Ben: thank you for being my best friend and partner in crime for almost ten years now. Thank you for sticking with me through all the ups and downs of grad school, for supporting me and reminding me to take better care of myself and for always managing to put a smile on my face, no matter how stressed or tired I might be. I'm excited to see where our next adventure will take us!

Table of Contents

Chapter 3: Programming innate immune stimulation with 3D wireframe DNA origami

Chapter 4: Conclusion

Appendix

List of Figures

Appendix

List of Tables

13

Chapter 1

Introduction

Thesis Overview

The development of DNA nanotechnology in recent decades has enabled the rational design and synthesis of highly programmable DNA-based structures capable of sensing and interacting with their environments in a myriad of ways. Although a vast spectrum of architectures has been developed, ranging from static DNA tiles to dynamic therapeutic delivery vehicles that 'unlock' in the correct environment, the fundamental principle underlying the design of all DNA nanostructures is the same as the principle which dictates the iconic structure of the double helix: the predictable and specific interactions between complementary nucleotides. The preferential binding of adenine (A) to thymine (T) and cytosine (C) to guanine (G) has not only been utilized in nature for hundreds of millions of years to write and regulate the genetic information contained in all living creatures but has also been co-opted in recent decades to enable the use of DNA as a controllable nanomaterial.

DNA nanoparticles are assembled from several to several hundred oligonucleotides which hybridize to each other according to these well-defined base pairing interactions. DNA origami, a sub-branch of DNA nanotechnology, relies on the folding of a long single-stranded scaffold into specified shapes by the binding of numerous short staple strands to multiple complementary sequences along the scaffold, bringing regions that are far apart in sequence space into close spatial proximity. By programming the sequence complementarity of their composite oligos, either manually or using automated design tools, it is possible to create 2D or 3D nanoparticles of near-arbitrary geometry. Furthermore, the addressability of DNA assemblies enables attachment of ligands with nanometer-scale control over their spatial positioning and orientation. These unique advantages, along with the relative ease of DNA oligo synthesis, have spurred the increasing usage

of DNA as a nanomaterial and established DNA origami as a versatile platform for the development of complex nanodevices suitable for applications in a wide variety of fields requiring precision engineering and a high degree of spatial control.

In this chapter, I will provide a brief introduction to several areas of ongoing research which are relevant to my thesis work. I begin with a brief history of DNA nanotechnology, outline some limitations of DNA origami design and synthesis, and summarize several technological advancements that seek to tackle these challenges. Additionally, as nucleic acid nanodevices are increasingly being designed for therapeutic purposes, it is therefore important to understand the manner in which they interact with the immune system. As a result, in the second half of this chapter, I transition to a discussion of pattern recognition receptors within the innate immune system, focusing specifically on several DNA-sensing immune receptors which are most likely to mediate nanoparticle-induced immune responses. Finally, I describe the motivation for my thesis work and lay out the structure of the subsequent chapters of this thesis.

The development of DNA nanotechnology

Since the pioneering of the field of DNA nanotechnology by Ned Seeman in 1982, there have been numerous scientific and technological advancements in this area which have enabled the design and synthesis of custom DNA nanodevices of near-arbitrary geometry and designed with nanoscale precision that are capable of interacting with the environment in programmable ways[1].

The biological principle underlying the development of the field of DNA nanotechnology is the fact that branch junctions, which DNA assembles into during genetic recombination, can be

redesigned in order to achieve specific outcomes. Holliday junctions, a type of natural branched DNA structure composed of 4 strands of DNA arranged in a '+' shape, are intrinsically unstable due to sequence complementarity that enables the sliding of the junction along the composite strands. Seeman showed that DNA oligos could be made to preferentially associate into immobile Holliday junctions by minimizing internal sequence symmetry, and further postulated that macromolecular networks of nucleic acids could be rationally constructed by assembling multiple Holliday junctions together using sticky end ligation[1]. Seeman and colleagues followed up on this work by designing and synthesizing a 4-arm junction, then expanded upon this to assemble highly branched structures ranging from immobile 5-arm junctions up to 12-arm junctions[2,3].

Shortly after Seeman described this immobile Holliday junction, Fu and Seeman demonstrated that two Holliday junctions could be joined together to form a double crossover (DX) molecule, in which two double helices are held together either in parallel or in the more energetically favorable antiparallel conformation by two DNA crossovers[4,5]. The design of this DX 'tile' then inspired the creation of constructs such as a rigid 4x4 tile, a six-point star tile and a six-helix bundle, all of which can either exist as individual units or self-assemble into macroscopic, periodic structures[6–8]. Apart from periodic structures, Yan *et al.* also designed self-assembled, aperiodic DNA lattices, which were formed by the alignment of several DX tiles around a long scaffold DNA strand and could be extended further into ribbon lattices by functionalizing the individual units with sticky ends[9].

DNA origami

In 2006, Paul Rothemund published a paper introducing the concept of DNA origami, wherein he described the assembly of 2D structures ranging from rectangular shapes to smiley faces from a single long 'scaffold' strand and hundreds of shorter 'staple' strands[10]. To design these DNA origami, the desired shape must first be visualized as parallel cylinders, which are simplified representations of the DNA helices. These helices are constrained to be integer number of turns in length, where each turn of the helix is ~10.67 bases in length, and are held together using double crossovers, which are designated locations where staple strands that are partially hybridized to one helix cross over to a programmed region on adjacent helix containing a complementary sequence. These crossovers are separated by ~1.5 turns, or ~15-16 nucleotides, to ensure that all of the staples are aligned within the same plane throughout the entire origami. This also enables the routing of a single scaffold strand through the shape such that it forms the second strand in every helix in the design. Once the staple crossovers and scaffold routing are fixed, given a scaffold sequence, the staple sequences can be determined as well. DNA origami of near-arbitrary shapes can be synthesized through this method, and additionally, by simply reprogramming the staple routing, many different geometries can be folded from the same scaffold.

DNA origami synthesis proceeds via the folding of the long single-stranded scaffold into prescribed geometries through the programmed hybridization interactions with numerous short single-stranded staple strands. Each staple binds to a unique set of sequences on the scaffold, bringing regions that may be far apart in sequence space into close physical proximity. The cooperative binding of the staple strands to their complementary scaffold sequences folds the

scaffold into the target geometry over the course of a thermal annealing process in which the temperature of the folding mixture is raised to 95C and gradually cooled to 25C over several hours.

This general DNA origami design has been used to create many different classes of single-layer, 3D DNA structures. One example of this is the folding of rigid 2D planar origami into 3D geometries, such as a box with a controllable lid, by the careful placement of single-stranded scaffold 'hinges.' Another extension of this method has been to design crossovers such that they would not be separated by an integer number of half-turns and therefore would not be arranged on a single plane, introducing a controlled curvature. A demonstration of this method was the creation of an origami nanotube by Douglas *et al.* in 2007[11].

Additionally, besides the folding of DNA nanoparticles from tightly packed sheets of parallel double helices, wireframe DNA origami can be created by designating multi-arm junctions with controlled angles as vertices and pairs of antiparallel DNA helices as edges[12]. While the development of wireframe origami enables the construction of increasingly complex architectures, designing these structures manually becomes correspondingly more difficult, prompting the creation of many different computer-aided design algorithms.

The first design algorithms developed required some level of technical understanding of DNA origami design principles on the part of the user. Software such as caDNAno and Tiamat enabled the automated assignment of staple sequences based on a scaffold routing pattern, but manual input of the routing and crossover design was necessary, therefore restricting the routine use of these design tools to those well-versed in the field[13,14]. In comparison, the next wave of design tools

were relatively more accessible and less technically demanding, as these programs, such as TALOS, PERDIX, METIS, and DAEDALUS, fully automated the scaffold routing and staple design for any user-defined shape[15–18]. Finally, the most recently developed algorithms, Adenita and ATHENA, integrated design features from previous generations of software to provide the most streamlined and user-friendly experience for DNA origami design[19,20].

DNA origami scaffold and staple synthesis

Apart from the development of more user-friendly design softwares, recent advancements in scaffold and staple production have also enabled the design of increasingly customizable wireframe DNA origami. Synthesis of short staple oligos, which are typically between 30-100 nucleotides in length, can be accomplished rapidly and at low cost through a variety of chemical and enzymatic approaches, including asymmetric PCR from a double-stranded DNA template[21], in which one primer is added to the reaction mixture in molar excess over the second primer to selectively amplify a single strand of the template DNA, as well as *de novo* solid-phase oligonucleotide synthesis, in which oligos of up to ~100 nucleotides can be synthesized at costs of several cents per nucleotide and on the scale of tens to hundreds of nanomoles per oligo[22]. However, production of kilobase-length scaffolds at the milligram quantities required for *in vitro* cell-based and *in vivo* assays is still somewhat prohibitive. Solid-phase synthesis is currently limited to per-base incorporation efficiencies of 99 – 99.5%, which severely restricts the length of oligos that can reliably be produced using this method. While the expected yield for oligonucleotides of up to 100 nucleotides is between 37 – 60%, the expected yield for oligos of 200 nucleotides drops to 13 – 37% and even further to <0.01–1% for a 1 kilobase oligo. Additionally, although it has been demonstrated that enzymatic approaches such as asymmetric PCR, exonuclease digestion, rolling

circle amplification or biotin-streptavidin separation can be used to generate ssDNA sequences of 10kb+, the primer design is often complex, the protocols require optimization on a case-by-case basis, and the reaction efficiency tends to be limited by physical constraints such as polymerase processivity and streptavidin dissociation[23]. Furthermore, as each of these methods involves the presence of dsDNA, the downstream purification process required to ensure the isolation of pure ssDNA from residual dsDNA contamination introduces yet another complexity to the ssDNA production process.

In order to circumvent the aforementioned limitations surrounding kilobase-length production of ssDNA, the field of DNA nanotechnology has traditionally relied heavily on a commercially available 7492 nucleotide variant of the circular single-stranded M13 bacteriophage genome. The high phage titers that can be achieved within bacteriophage-infected *E. coli* cultures allows for 1-10 mg of purified ssDNA per liter of shaker flask culture, while high-density bioreactor fermentation of bacteriophage-infected *E. coli* has enabled the rapid and efficient production of up to several hundred milligrams of M13 ssDNA in one liter of bacterial culture. As the purified ssDNA from these biofermentations could theoretically be used directly as scaffolds for DNA origami, this eliminates the need for a dsDNA intermediate and enables the production of *in vivo* quantities of DNA nanostructures. However, while this approach represents a significant improvement over previous methods, the reliance on the M13 bacteriophage genome severely limits the sequence space available for nanoparticle design and the presence of coding sequences within the genome may be undesirable for a variety of purposes, most notably therapeutic applications. To address this issue, helper plasmid systems have been developed which harness

the M13 bacteriophage replication machinery to allow for the production of circular ssDNA with near-complete sequence control.

M13 bacteriophage

M13 is a filamentous, lysogenic bacteriophage which is natively composed of a 6407 nucleotide single-stranded circular DNA genome encoding 11 viral proteins enveloped in a protein capsid comprised of 2700 copies of the p8 major coat protein and 5 copies each of the minor coat proteins p3, p6, p7 and p9. Upon successful infection of *E. coli* cells carrying the F pilus, M13 phage replication and occurs continuously within the host cell. While the viral reproduction process introduces an additional metabolic burden on the host cell, slowing its growth and lengthening the doubling time, infection of *E. coli* with M13 nevertheless does not induce lysis.

M13 replication is initiated upon delivery of the single-stranded phage genomic DNA (+ strand) into the cytoplasm of the host cell through binding of p3 to the *E. coli* F pilus. Bacterial enzymes convert the single-stranded viral genome into its double-stranded replicative form (RF DNA), which consists of the (+) strand and its complementary (–) strand. DNA polymerases can then bind to and replicate the newly formed RF DNA, from which mRNA transcription is immediately initiated, while rolling circle amplification, mediated by p2 and p10, occurs in parallel to produce additional copies of single-stranded (+) strand DNA. This cycle of replication continues until a threshold concentration of phage protein p5 is reached, upon which numerous copies of p5 begin binding to (+) strand ssDNA, sequestering it and preventing its conversion into dsDNA[24].

The formation of this p5-ssDNA complex also serves as a signal for the phage assembly complex, consisting of p1, p4 and p11, to assemble the major coat protein p8 and tail virion minor coat proteins p7 and p9 around the ssDNA genome while simultaneously removing p5. Phage capsid formation is complete when the head virion minor coat proteins p3 and p6 are attached to the capsid and interact to form the p3-p6 complex, which is necessary for correct termination of phage assembly[25]. Besides mediating the packaging of the ssDNA genome into viral capsids, the assembly complex is also critical for export of fully synthesized phage from the host cell. The inner membrane spanning protein p1 interacts with p4, an outer membrane spanning protein, and p11, an inner membrane-anchored periplasmic protein, to form a channel through the host cell membrane through which newly synthesized phage are extruded[26].

Engineering M13

M13 has been extensively studied and characterized in the several decades since it was discovered in the 1960s, while its simplistic structure, limited number of encoded proteins, and ease of amplification have proven highly beneficial for genetic engineering. Indeed, the method of phage display, for which the 2018 Nobel Prize in Chemistry was awarded, takes advantage of the fact that coat proteins of phages such as M13 can be readily engineered to display proteins or peptides of interest on the surface of phage particles while the corresponding gene remains encapsulated within the phage. Although any of the coat proteins could hypothetically be used for phage display, the two most frequently used coat proteins are the major coat protein p8 and the minor coat protein p3. While p8 display still typically uses bacteriophage vectors, engineered p3 is more commonly expressed from phagemid vectors[27].

Phagemids are plasmids which contain a bacteriophage-derived f1 origin of replication (f1 ori) and packaging signal in addition to the typical plasmid origin of replication. Due to the presence of the two origins of replication, phagemids are not only capable of double-stranded replication like standard plasmids but are also imparted with the capability to be replicated as single-stranded DNA and subsequently packaged into viral capsids for export from the host bacterial cells. While the majority of the M13 genome consists of coding and regulatory sequences, phagemids need only to contain the two origins of replication and, optionally, an antibiotic resistance gene. This allows for significantly greater flexibility in terms of sequence design, as custom sequences of up to several kilobases in length can easily be inserted into the phagemid using traditional cloning methods such as Gibson or Golden Gate. However, as the phagemids themselves lack all phage protein coding genes, they are unable to create functional phage and therefore behave like regular plasmids in uninfected bacterial cells, only regaining their full functionality in cells which have been infected with bacteriophage or helper phage, or alternatively, co-transformed with helper plasmids.

Helper phages are bacteriophages which have been engineered to have disrupted packaging signals and often contain a plasmid origin of replication as well as an additional antibiotic resistance gene. The most frequently used, commercially available helper phage is M13KO7, which contains a low copy number plasmid origin of replication (p15a) and a kanamycin resistance gene inserted into the M13 origin of replication. While the packaging signal functionality is not completely abolished, and therefore production and export of M13KO7 particles is still possible in the absence of phagemid DNA, phagemids containing a wild-type f1 origin and packaging signal are preferentially packaged and exported from the cells. However, this methodology produces heterogenous populations containing both the engineered phagemid and M13KO7, which not only

reduces the yield of phagemid particles but also requires additional downstream purification steps to isolate the target phagemid ssDNA from the similarly single-stranded M13KO7 genome.

One method that has been employed to circumvent these limitations is the use of 'helper plasmids,' plasmids which contain all of the requisite M13 protein-coding regions but have had the M13 origin of replication excised[28]. The removal of the packaging signal contained within the origin of replication ensures that the helper plasmid cannot be packaged and exported from the host cells, thereby enabling the production of pure phagemid ssDNA. The M13cp helper plasmid was designed in this manner and additionally contains a chloramphenicol resistance gene as well as a copy of the p15a origin of replication to maintain low copy numbers of the helper plasmid within the host cell. Since this method has been reported, many unique approaches have been developed that leverage this helper plasmid system combined with innovative phagemid design in order to produce pure ssDNA of custom sequence.
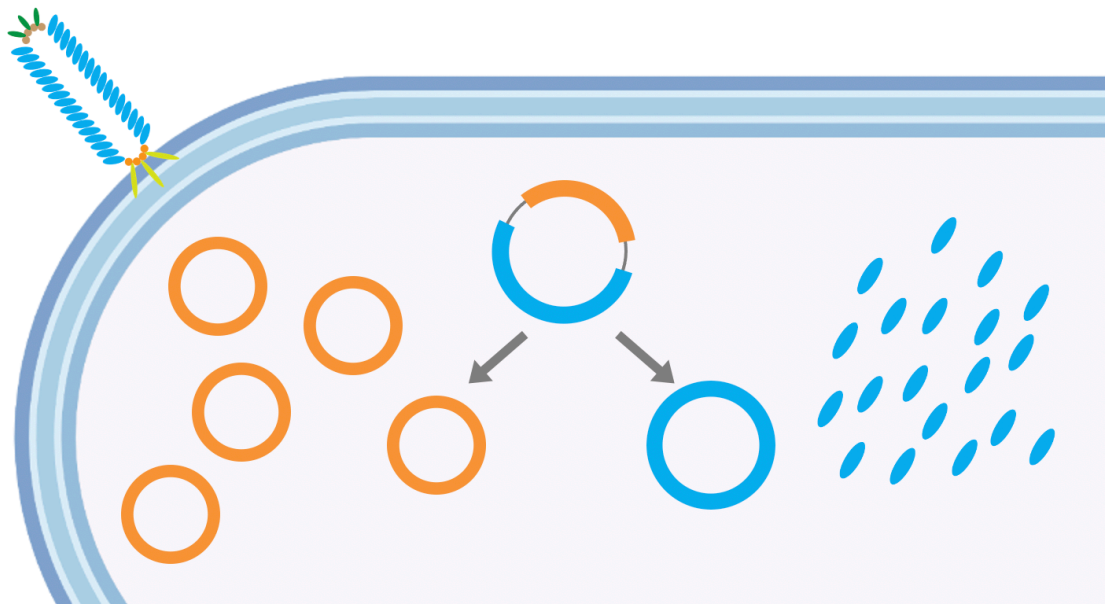
Figure 1.1. Schematic of engineered phagemid and helper plasmid within an *E. coli* cell. The f1 origin (orange) and protein coding genes (blue) of the M13 bacteriophage genome can be split into

two separate functional plasmids which can operate independently. This enables the design and production of custom sequence phagemids without the presence of M13 protein coding genes. Phagemids can be replicated, packaged and exported from cells that have been co-transformed with the helper plasmid (blue).

M13-based ssDNA production methods

Phagemids have much greater sequence customizability compared to the M13 bacteriophage genome. However, the plasmid and bacterial origins of replication, along with the antibiotic resistance gene, comprise a 2-3kb sequence fixed region. The minimum possible length of the phagemid is constrained by the size of this fixed region, thus somewhat limiting the utility of this method for the design of scaffolds that require complete sequence control. To increase the degree of control over phagemid sequence customizability, several attempts have been made to minimize the size of this fixed region by modifying the f1 origin of replication.

Dotto *et al.* demonstrated that the f1 origin is used for both initiation and termination of (+) strand synthesis, and that these while these initiation and termination domains are overlapping, they are sufficiently distinct such that they can be individually inactivated without significantly affecting the function of the other[29]. Specthrie *et al.* followed up on this research by sequentially inserting the minimal sequences for the (+) strand origin of replication, the packaging signal, and a truncated (+) strand origin of replication into an existing plasmid vector[30]. The truncated origin of replication was recognized as the site for termination of ssDNA synthesis as well as for recircularization, and therefore the phagemid design enables the production of circular ssDNA containing only the sequence flanked by the initiator and the terminator. Co-transformation of this phagemid and helper phage into *E. coli* demonstrated the successful packaging of circular ssDNA of as little as 221 nucleotides in length into 'microphage' particles.

More recently, Nafisi *et al.* constructed a phagemid, termed 'pScaf,' which integrated the findings from Dotto *et al.* and Specthrie *et al.* and further optimized the design in order to enable the construction of custom length ssDNA scaffolds and improve the purity of the resulting ssDNA product[31]. Their pScaf vector, similarly to Specthrie *et al.*'s design, also had a fixed sequence containing an M13 origin of replication, a packaging signal, and a terminator derived from a second, truncated M13 origin of replication. However, a cloning site was added between the initiation and termination domains to facilitate insertion of additional custom sequences, and the sequence of the terminator was optimized through a mutational screen in order to minimize its ability to serve as an initiator of ssDNA replication, therefore reducing the amount of off-target sequences produced via this method.

An alternative method developed by the Dietz lab utilizes a custom-sequence phagemid with helper plasmids or helper phages, but instead of minimizing the length of the fixed sequence, they integrated self-excising DNAzymes into the phagemid backbone[32]. DNAzyme self-excision is induced by adding $ZnCl_2$ to the purified phagemid ssDNA, and sequences that were inserted in between two DNAzyme cassettes are separated from the rest of the phagemid backbone. However, unlike the previous methods, DNAzyme excision results in linear products, which may be more susceptible to nuclease degradation. Nevertheless, this approach enables the simultaneous production of DNA origami scaffold and staple oligos, streamlining the DNA origami production process.

DNA nanotechnology and the immune system

The emergence of technologies such as those described above, which facilitate the production of ssDNA scaffolds of custom length and sequence, has contributed to the development of DNA nanotechnology for a variety of applications in fields ranging from materials science and biophysics to molecular computing. One particular avenue of research which has benefited significantly from these technological advancements has been the design of DNA nanodevices for therapeutic and prophylactic applications. The deployment of these nanodevices in *in vivo* environments, whether in model organisms or in humans, means that they must be designed not only to perform a particular therapeutic function but also with their potential interactions with the immune system in mind. The ease at which DNA nanodevices can be functionalized with hybridization or click chemistry methods and the nanoscale precision at which ligands can be attached to the scaffold have inspired the design of nanostructures carrying aptamers or antibodies for active targeting or immune cell stimulation. However, while many efforts have been undertaken to engineer DNA nanostructures for activation or enhancement of the adaptive immune response, the primarily nucleic acid composition of these nanostructures suggests that they themselves may intrinsically act as ligands for specific receptors within the innate immune system.

Nucleic acid sensing pathways in the innate immune system

The innate immune system has evolved a variety of pattern recognition receptors (PRRs) to detect and respond to evolutionarily conserved molecular patterns characteristic of pathogenic microorganisms (pathogen associated molecular patterns; PAMPs) or cell and tissue damage (damage associated molecular patterns; DAMPs). Of these PRRs, a subset has become specialized in detecting foreign or exogenous nucleic acids that are typically associated with viral infection,

and these PRRs can be further differentiated based on their subcellular localization. One group of innate immune sensors is located within the endosome and is expressed primarily by immune cells such as dendritic cells (DCs), B cells, and macrophages, while the other group is located within the cytoplasm and can be found in almost all cell types. Upon binding to their cognate ligand, all PRRs within both of these groups trigger downstream signaling pathways and induce the activation of transcription factors including TNF-α and interferon regulatory factor 3 (IRF3), which upregulate the expression of Type I interferons (IFNs) and downstream IFN-stimulated genes (ISGs)[33,34]. This antiviral defense not only activates the innate immune system but is also critical for priming the antigen-specific adaptive immune response.
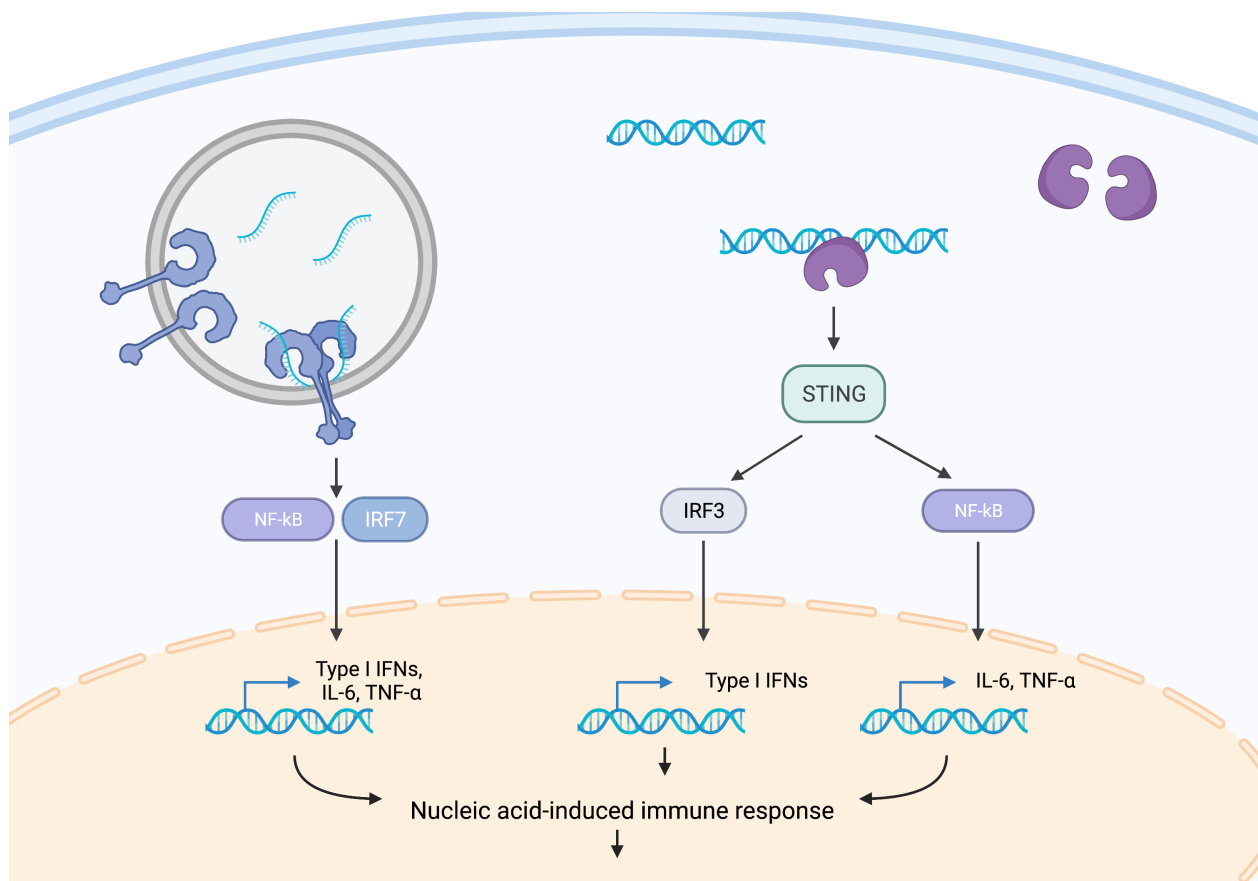
Figure 1.2. Potential nanoparticle-responsive innate immune DNA-sensing pathways. The TLR9 pathway is shown on the left and the cGAS-STING pathway is shown on the right. Figure created with BioRender.com.

Innate immune sensing: TLRs

One major class of PRRs which sense and respond to nucleic acids is the Toll-like receptor (TLR) family[35]. TLRs were the first class of PRRs to be identified, and while many other types of PRRs such as RIG-I-like receptors (RLRs), Nod-like receptors (NLRs) and cytosolic DNA sensors like cGAS have been discovered in the years since, TLRs remain the most extensively characterized. TLRs are single transmembrane proteins that are composed of three structural domains: an N-terminal ectodomain with leucine-rich repeats (LRRs) that is responsible for PAMP recognition and binding, a transmembrane domain, and a C-terminal cytoplasmic Toll/IL-1 receptor (TIR) domain involved in the triggering of downstream signaling. TLRs are either homo- or heterodimers, with each of the monomers in the N-terminal ectodomain exhibiting a horseshoe-like shape which overlap slightly upon ligand binding and dimer activation. Dimer formation brings the TIR domains of the two monomers into close proximity, and TIR-domain-containing adaptor proteins such as the myeloid differentiation primary response protein 88 (MyD88) and the TIR-domain-containing adaptor-inducing IFNβ (TRIF) are recruited to trigger formation of the Myddosome signaling complex. Successful myddosome formation initiates the activation of downstream signaling pathways, which ultimately leads to the expression of Type I IFNs and other inflammatory cytokines involved in the innate immune response.

Humans have 10 TLRs, which can be divided into two groups according to localization. The first subfamily of TLRs is localized on the cell surface membrane and consists of TLR1, TLR2, TLR4, TLR5, TLR6, and TLR10, while the second subfamily of TLRs is localized within the endosomal membrane and includes TLR3, TLR7, TLR8, and TLR9. While the cell surface membrane TLRs primarily respond to proteins and lipids, which are found in high abundance in microbial

membranes, endosomal TLRs are involved in nucleic acid recognition. Of the four endosomal TLRs, three of them detect RNA: TLR3 responds to double-stranded RNA (dsRNA), while TLR7 and TLR8 recognize single-stranded RNA (ssRNA). The last endosomal TLR, TLR9, recognizes single-stranded DNA containing high proportions of unmethylated cytosine-phosphate-guanine (CpG) dinucleotides[36].

DNA sensing: TLR9

TLR9 is synthesized and initially localized in the endoplasmic reticulum (ER) in quiescent cells and is actively recruited to endosomes upon cellular internalization of CpG DNA[37]. TLR9 is escorted from the ER by UNC93B1, which facilitates loading of TLR9 into COPII vesicles and its subsequent trafficking to the Golgi. UNC93B1 and TLR9 then travel together to the cellular surface, upon which AP-2 is recruited and TLR9 is internalized into endosomes via clathrin-mediated endocytosis[38]. Upon localization of TLR9 into endolysosomes, the acidic environment facilitates the cleavage of the TLR9 ectodomain by cathepsins and endopeptidases to generate the functional, mature form of TLR9[39,40].

Mature, inactive TLR9 exists a horseshoe-like monomer form. However, a recently determined crystal structure of activated TLR9 bound to agonistic CpG oligos has demonstrated that binding of CpG DNA to TLR9 induces the formation of a 2:2 complex consisting of two CpG oligos and two TLR9 monomers, in which the bases in each of the CpG DNA oligos interact with the N-terminal end of one TLR9 monomer while the phosphate backbones interface with the C-terminal side of the other TLR9 monomer[41]. Additionally, it has been shown that there is another binding site within each of the TLR9 monomers that binds to oligos which have a cytosine at the second

position relative to the 5' end (5'-xCx) and functions cooperatively with the CpG binding site to promote TLR9 dimerization and subsequent activation[42]. While the binding of two 5'-xCx oligos is not required for TLR9 activation, it has been shown that their presence enhances TLR9 activation levels.

The distance between the two CpG binding sites within the TLR9 dimer can be estimated from the freely accessible protein database (PDB) crystal structure of the activated TLR9 dimer in complex with CpG DNA oligos and was found to be approximately 7nm[42]. As it has been shown that multiple residues within both binding pockets interact with the CpG bases as well as the phosphate backbone, we hypothesize that the structure of the activated TLR9-CpG complex is likely to be fairly constrained, and therefore, for any set of two CpG oligos bound to a TLR9 dimer, the inter-ligand distance of the two oligos can be approximated to be 7nm. This suggests that pairs of CpG oligos fixed to a surface at distances of 7nm apart may be optimally positioned for TLR9 binding and may therefore demonstrate enhanced TLR9 activation, and in fact, there has been a recent study which has investigated the effects of 'optimized' versus highly disfavorable spatial positioning of CpG oligos on the strength of TLR9 signaling[43].

One interesting question which arises due to the close proximity of the two CpG binding pockets is whether the TLR9 dimer may be activated by binding of a single CpG-containing ssDNA to both monomers simultaneously. Depending on the length of the oligo, it is certainly not impossible for this to occur, particularly as the distance between bases in ssDNA has been found to be approximately 0.65nm while the persistence length of ssDNA has been measured to be in the range of 1.98 ± 0.72nm, indicating high flexibility even at the length scales relevant to TLR9 dimer

binding[44–46]. However, the orientation of the residues within the binding pockets that interface with the cytosine and guanine nucleotides suggests that a substantial amount of contortion may be required in order for a single CpG oligo to bind to both sites at once. Therefore, if there is a sufficient concentration of CpG oligos within the endosomal compartment, it is more likely that a TLR9 dimer would be bound by two separate CpG-containing oligos rather than one.

The same question can be posed with respect to the 5'-xCx binding sites within the TLR9 dimer as well, as the distance between each set of the CpG binding sites and the 5'-xCx binding sites on the same side of the TLR9 dimerization interface was also measured to be approximately 7nm. Given this distance, it has been hypothesized that a single DNA oligo could theoretically bind to both sites simultaneously provided that at least 10 nucleotides separate the two CpG motifs. However, as the bottom of the 5'-xCx binding site is separated from the CpG binding site by the dimerization interaction of the two TLR9 monomers, this would require the CpG motif closer to the 5' end of the oligo to bind to the 5'-xCx motif before looping down to facilitate the binding of the CpG motif closer to the 3' end of the oligo to the CpG binding pocket. Because of these structural constraints, Ohto *et al.* postulated that it is more likely that each of these sites independently binds a separate CpG DNA, resulting in a 2:2:2 complex. This is further supported by the fact that the 10-nucleotide spacing requirement is not observed in the sequences of CpG DNA that have been experimentally optimized for TLR9 activation[47].

As CpG dinucleotides are not entirely absent from mammalian genomes, there are two major factors which prevent TLR9 from mounting unwanted immune responses towards self-DNA: 1) ~70–80% of cytosines in mammalian CpG dinucleotides are methylated, and 2) CpG dinucleotides

occur at a significantly lower rate in mammalian genomes compared to bacterial or viral genomes, as methylated CpG sites have a high mutation rate due to spontaneous deamination of the cytosine into a thymine[48,49]. Additionally, while the presence of a CpG dinucleotide is minimally sufficient for TLR9 activation, it has been shown that the surrounding sequence context has a significant impact on the potency of TLR9 activation[50].

As of now, three major classes of CpG oligodeoxyribonucleotides (ODNs) have been identified, each of which has distinct sequences and elicits differential immune responses. Class A ODNs, alternatively named Class D ODNs, contain a central palindromic CpG with a poly-G motif at the ends and are characterized by strong activation of NK cells and pDCs but weak activation of B cells. These ODNs induce high Type I IFN secretion as well as TNF-α and IL-12 expression, but are weak activators of the TLR9-dependent NF-κB pathway. In contrast, Class B ODNs, alternatively named Class K ODNs, containing CpG motifs and a phosphorothioate backbone strongly activate B cells and induce strong TLR9-dependent NF- κB signaling, culminating in high expression of pro-inflammatory cytokines such as IL-6, IL-10 and IL-12[51]. Finally, Class C ODNs combine the characteristics of Class A and B ODNs, containing CpG motifs at the 5' end, a CpG-containing palindromic sequence at the 3' end and a fully phosphorothioate backbone. Their immunostimulatory properties are also a combination of those of Class A and B ODNs, with the ability to induce strong Type I IFN production as well as NK cell activation and B cell proliferation[52,53].

In addition to the identification of the structural features differentiating the major CpG ODN classes, more systematic studies have also been undertaken to determine the sequence motifs

which impact TLR9 activation and to identify the minimum sequence requirements for strong stimulation of the TLR9 pathway. By analyzing and performing mutational studies on the sequences of known strong TLR9 agonists ODN2006 and ODN2395, Pohar *et al.* identified the properties of ODNs which influenced their ability to activate TLR9 as well as the minimal sequence motifs of Class B ODNs which induced strong activation of the TLR9 pathway in B cells as well as PBMCs[54]. Oligos satisfying those requirements must contain at least two CpG dinucleotides separated by 6-10 nucleotides and be a minimum of 21 nucleotides in length. Additionally, it was shown that the nucleotides adjacent to the CpG dinucleotides also strongly define the immunostimulatory capability of the CpG motif. The presence of a thymine immediately upstream of the first CpG motif at the 5' end of the oligo contributes substantially to CpG oligo potency, as substitution of the T to a different base resulted in a significantly reduced ability to activate TLR9. The presence of a poly-T at the 3' end of CpG ODNs, downstream of the second CpG dinucleotide, also significantly improves the ability of the oligo to induce TLR9 activation. Finally, the distance between the two CpG dinucleotides is yet another modulator of ODN potency, as the lengthening or shortening of the spacer outside of the aforementioned range decreases the immunostimulatory capability of the oligo. In conclusion, the authors determined that the sequence requirements for oligos which have the ability to strongly activate human TLR9 are as follows: the ODN length must be 21 nucleotides of longer, with at least two CpG dinucleotides separated by a spacer of between 6-10 nucleotides where the 5' CpG must be immediately preceded by a T and the 3' CpG must be followed by a poly-T tail. These findings not only facilitated the rational design of potent TLR9 agonists, including the ones that will be discussed in later chapters of this thesis, but also shed light on the sequence contexts within which CpG dinucleotides could be considered strongly immunostimulatory.

Thus, based on these sequence requirements, we can begin to formulate hypotheses about the TLR9-activating capability of nucleic acid nanostructures. A nanoparticle folded from a circular kilobase-length ssDNA scaffold and numerous short linear ssDNA staples, if not designed with the express purpose of avoiding TLR9 activation in mind, will almost certainly contain some number of CpG dinucleotides and potentially several of these optimized CpG sequences as well. While the interactions of TLR9 dimers with the ssDNA scaffold may be somewhat limited due to its circular structure and the relative rigidity of the CpG binding site locations within the TLR9 dimer, the linear ssDNA staples are subject to no such constraints and the high concentration of these staples within a nanoparticle folding mixture may facilitatee the interaction of staples with TLR9 dimers within the endosomal compartment. Even if the CpG motifs within these staples do not satisfy the sequence requirements for optimal potency, the binding of multiple staples to the same set of TLR9 dimers may be able to cooperatively achieve similarly enhanced levels of TLR9 activation. Additionally, while the intact ssDNA scaffold would hypothetically have limited interactions with TLR9 dimers, degradation of the scaffold over time by nucleases such as DNase I and II will most likely generate shorter linear ssDNA fragments which may interact with TLR9 in a similar manner to the staple oligos[55]. Therefore, when designing nucleic acid nanoparticles, particularly for applications in which the nanoparticles will be interfacing with the immune system, the presence of immunostimulatory motifs within the scaffold and staple sequences should be taken into consideration.

Innate immune sensing: Cytosolic nucleic acid sensors

Unlike the endosomal nucleic acid sensors, all of which belong to the TLR family, there are a variety of different nucleic acid sensors localized within the cell cytosol[34,56]. The RNA-sensing

cytosolic receptors belong to the RIG-I-like receptor (RLR) family, which consist of retinoic acid inducible gene I (RIG-I), melanoma differentiation associated gene 5 (MDA5), and laboratory of genetics and physiology 2 (LGP2)[57,58]. Upon binding to double-stranded RNA in the cytosol, DExD/H box helicases signal through mitochondrial antiviral signalling protein (MAVS) and IRF3/IRF7 to induce the production of Type I IFNs. MAVS activation also promotes apoptosis through the IRF3/IRF7 pathway, and additionally contributes to the production of IL-1β and its associated pro-inflammatory response via activation of the NOD-, LRR- and pyrin domain-containing 3 (NLRP3) inflammasome.

There are also a number of DNA-sensing receptors localized within the cytoplasm, specifically cyclic GMP-AMP synthetase (cGAS), absent in melanoma 2 (AIM2), and IFNɣ-inducible protein 16 (IFI16)[34]. These detect pathogenic dsDNA as well as self-derived dsDNA that leaked into the cytoplasm as a result of cellular damage or death. Activation of AIM2 by dsDNA binding induces the recruitment of apoptosis-associated speck-like protein containing a CARD (ASC) and subsequent formation of the AIM2 inflammasome, which promotes pyroptosis, an inflammatory form of cell death, and IL-1β maturation via activation of caspase-1[59]. AIM2-related protein IFI16, a primarily nucleus-associated protein which can be found in small quantities in the cytoplasm, is also involved in inflammasome activation following detection of cytoplasmic dsDNA[60]. Finally, cGAS signals through stimulator of IFN genes (STING) to induce Type I IFN production as well as apoptosis via the IRF3 pathway.

DNA sensing: the cGAS-STING pathway

Levels of intracellular dsDNA are kept in check by the presence of DNases such as DNase II, localized in the lysosome, and three prime repair exonuclease 1 (TREX1), localized within the cytoplasm[33]. However, if there is an excess of cytoplasmic dsDNA, such as what would occur during a bacterial or viral infection, cytosolic DNA sensors such as cGAS will be activated. The binding of cGAS to dsDNA stimulates the production of an isomer of cyclic GMP-AMP (cGAMP), specifically 2'3'-cGAMP, which functions as a secondary messenger that then binds to and activates STING[61]. STING, an ER-localized adaptor protein, translocates to the Golgi upon activation, where it then recruits TBK1 to phosphorylate IRF3, which then relocalizes into the nucleus and activates transcription of Type I IFNs, particularly IFNβ, as well as the NF- κB - dependent expression of pro-inflammatory cytokines such as IL-6 and TNF-α.

In its inactive form, cGAS exists within the cytosol as a loosely associated homodimer. However, upon binding of each cGAS monomer to a dsDNA oligo, the activation loops within each monomer undergo a conformational change to enable the formation of an activated 2:2 complex[62]. Interestingly, it has been found that the dimer complex is not very strongly associated, which is consistent with the fact that the protein-protein interface is mediated by very few residues. The presence of the two dsDNA oligos is therefore critically important as it serves to crosslink the cGAS dimer, with each monomer interacting with both of the oligos at two binding sites and stabilizing the dimer in the catalytically active form[63].

While cGAS binding to dsDNA is sequence independent, as cGAS interacts with dsDNA via the phosphate backbone rather than individual nucleotides, it has been shown that dsDNA oligo length

is a critical factor in the activation of cGAS-STING signaling. Although short dsDNAs of ~20 nucleotides can bind to cGAS and interact with both binding sites, it is not sufficient for activation of the cGAS-STING pathway in human cells, potentially because of the rapid dissociation of individual 2:2 cGAS:dsDNA complexes[64]. In contrast, as dsDNA oligos of lengths over the minimum threshold of ~45 bases are long enough to bind to multiple cGAS dimers simultaneously, they can mediate the formation of higher order cGAS:dsDNA oligomers in which cGAS dimers assemble cooperatively and with increased affinity in a ladder-like formation along dsDNA oligos arranged in a quasi-continuous form. Thus, the oligo length plays a crucial role in the cooperative formation of higher order cGAS complexes and the efficacy of cGAS-mediated downstream signaling.

Besides dsDNA length, another factor which impacts the magnitude of cGAS activation is DNA structuring. It has been shown that DNA-bending proteins such as mitochondrial transcription factor A (TFAM) and high-mobility group box protein 1 (HMGB1) are able to enhance cGAS detection of DNA, potentially by structuring dsDNA into U-shapes[65]. The binding of these proteins to dsDNA results in the duplex folding over on itself, thereby bringing the two halves of the oligo close together. The arrangement of these two duplexes in close spatial proximity, positioned correctly for efficient cGAS binding, enables not just individual dimer activation but also the nucleation of larger oligomers. This structuring effect may also explain some of the length dependence of cGAS activation, as the binding of two long dsDNA dimers to an initial cGAS dimer prearranges them with the correct spatial distancing and orientation to facilitate efficient attachment of additional dimers. Interestingly, it has been demonstrated that short 12-20-mer dsDNA oligos can strongly activate cGAS if they are flanked by unpaired guanine-rich ssDNA

overhangs, although the mechanism underlying this method of cGAS activation has not yet been elucidated[64].

In summary, cGAS activation and subsequent downstream signaling through the cGAS-STING pathway are sequence independent but heavily dependent on both dsDNA length and duplex prestructuring. When these results are viewed through the lens of DNA origami design, they raise some interesting questions about how wireframe DNA origami may interact with cGAS *in vivo* and how cGAS recognition may contribute to overall nanoparticle immunogenicity. Wireframe DNA origami are composed of edges which contain two parallel dsDNA duplexes arranged in close proximity and are therefore structurally similar to dsDNA oligos bound to a single cGAS dimer. As a result, these types of DNA nanostructures may be intrinsically cGAS activating. Additionally, the edge lengths of wireframe DNA nanostructures can be easily programmed or modulated by redesigning the particle geometry or switching scaffold sequences. This enables the design of DNA nanostructures with edges lengths of ~45bp or longer to enhance the potency of cGAS activation, or alternatively the design of nanoparticles with edge lengths of ~20bp or less to minimize or eliminate nanoparticle-mediated cGAS-STING activation. Another variable which may influence the cGAS activating ability of DNA nanostructures is their geometry. While the wireframe edges contain parallel dsDNA duplexes, and may therefore mediate the assembly of several cGAS dimers along the length of each edge, nanoparticles of suitable geometry may also have the potential to further enhance cGAS activation by facilitating the formation of higher-order cGAS oligomers spanning the edges of multiple nanoparticles organized in a superstructure. As cGAS is one of the primary PRRs which respond to cytosolic DNA, it is critical to design DNA

nanoparticles with a clear understanding of their potential interactions with cGAS and the impact those interactions may have on the innate immune response.

DNA sensing: AIM2

Besides cGAS, there is one more cytosolic DNA sensor, AIM2, that plays a critical role in the innate immune response. Unlike cGAS, which activates the STING pathway and ultimately induces the production of Type I IFNs as well as NF- κB -dependent pro-inflammatory cytokines, AIM2 interacts with ASC to activate caspase-1, a proteolytic enzyme which then processes the inactive pro-IL-1β and pro-IL-18 into their mature, active forms[66]. Additionally, this AIM2-ASC interaction induces formation of the pyroptosome, a large assembly of oligomerized ASC dimers which, upon recruitment and activation of caspase-1, results in an inflammatory form of cell death known as pyroptosis[67].

Unlike TLR9 and cGAS, AIM2 does not undergo a conformational change upon binding to dsDNA to become activated. Instead, in the absence of its ligand, AIM2 exists in an autoinhibited state due to interactions between its HIN and Pyrin (PYD) domains[68]. These interactions are significantly weakened upon dsDNA engagement, thus releasing the autoinhibition of AIM2 and enabling oligomer formation. AIM2 recognition of dsDNA, much like cGAS, is sequence independent, as the binding pocket engages the dsDNA backbone rather than specific bases. Additionally, as the electrostatic interactions between the dsDNA ligand and the binding pocket are nonspecific, AIM2 can bind anywhere along an intact dsDNA oligo, facilitating the formation of massive signaling complexes composed of multiple AIM2 receptors. It has been shown that a 16-bp dsDNA oligo can simultaneously accommodate the HIN domains of four AIM2 receptors. Furthermore, it is

known that oligos of lengths between 70-80bp are required for optimal AIM2-mediated interferon induction[60]. This suggests that a dsDNA oligo of ~80bp can accommodate up to 20 AIM2 receptors simultaneously and serve as a platform for the assembly of signaling complexes that are sufficiently large to enable enhanced activation of the downstream immune response. Interestingly, it has also been demonstrated that AIM2 is not activated by 45mer dsDNA oligos, suggesting that there is a threshold length below which AIM2 oligomerization may not occur[34].

Given the dual-duplex structure of each edge of wireframe DNA nanoparticles, we hypothesize that AIM2 may respond to these nanoparticles in much the same way that they respond to dsDNA oligos. Specifically, the length dependence of AIM2 activation suggests that nanoparticles of edge lengths below 45 nucleotides may not be sufficient to induce AIM2 activation, while at the opposite end of the spectrum, nanoparticles with edge lengths of approximately 70 or 80 nucleotides may have the correct structural composition to induce optimal AIM2 activation. Alternatively, as crystal structures of HIN:DNA complexes show that dsDNA oligos are arranged pseudo-continuously via head-to-tail stacking, where HIN domains are arrayed around the double helix in a staircase-like fashion, it is also possible that AIM2:nanoparticle interactions may mediate the formation of pseudo-superstructures, therefore reducing the impact of nanoparticle edge length on AIM2 activation. One additional interesting aspect that can be further explored is the effect of the dual duplexes on AIM2 activation. AIM2 oligomerization is facilitated by the assembly of multiple HIN domains along a single dsDNA oligo, but it is not known how the presence of an additional duplex may perturb this system. As AIM2:DNA binding is nonspecific and highly flexible, it is possible that the presence of the additional duplex may be able to coordinate the formation of even larger AIM2 complexes. However, it is also possible that the convergence of

multiple dual-duplex edges at each nanoparticle vertex may sterically inhibit HIN domain binding and prevent AIM2 oligomerization. Thus, although DNA nanostructures are composed of duplex edges which, in isolation, may act as ligands for AIM2, the structuring of these edges into a 3D geometry may significantly change the manner in which they interact with immune receptors and influence their overall immunostimulatory effect.

Thesis goals and approach

The goal of this thesis is to elucidate the immunostimulatory effects of DNA origami and to begin to understand the ways in which these nanostructures can be designed for controllable modulation of the immune response. While several groups have previously reported on the immunogenicity of DNA nanostructures, because the term 'DNA origami' encompasses such a vast array of geometries and architectures, all of which interact with the immune system in different ways, the immunostimulatory profile of wireframe DNA origami is still very poorly understood. Furthermore, there is still a lack of understanding of how specific nanoparticle design parameters influence the potency of immune activation. Additionally, despite the fact that immune adjuvants have been displayed on several DNA nanodevices for immunostimulatory purposes, there has not yet been an investigation into how the nanoscale organization of such adjuvants can be harnessed to induce controllable magnitudes of immune activation.

In Chapter 2, I describe the development of a method for the scalable bioproduction of pure single-stranded DNA containing custom sequences, enabling the design of scaffolds enriched for or absent of immunostimulatory motifs in order to modulate the interaction of the resulting DNA origami with the innate immune system. Utilizing a helper plasmid system in combination with a

novel phagemid design, I demonstrate the production of circular single-stranded DNA without the presence of helper plasmid or dsDNA contamination. Additionally, I perform optimization assays for several bioproduction parameters in order to improve the yields that are achievable within shaker flask setups, and furthermore, demonstrate the scalability of this method with the production of ~1.5mg/L of pure ssDNA extracted from a 10L bioreactor production. This method substantially improves the scale at which DNA origami scaffolds containing user-defined custom sequences can be produced and streamlines the downstream purification process, reducing the number of steps required between the conceptualization of a DNA origami design and the final origami synthesis.

In Chapter 3, I first characterize the innate immunostimulatory capabilities of wireframe DNA origami. Initially focusing on two different nanoparticles, a pentagonal bipyramid (PB84) and an icosahedron (ICO42), as well as their respective unstructured controls, I explore the effects of geometry and nanostructuring within the context of reporter cells for two particularly relevant immune pathways, TLR9 and cGAS-STING and show that structured nanoparticles are minimally TLR9 activating, but strongly activate the cGAS-STING pathway. Additionally, we see that the unstructured controls behave differently compared to their folded counterparts, showing that nanostructuring can have an impact on immunostimulation. As neither nanostructure activated the TLR9 pathway, I chose to use TLR9 reporter cells as a model system for testing nanoparticle-mediated immune modulation. I then designed and synthesized several PB84 variants displaying a known TLR9 agonist, ODN2006, at various copy numbers and with a variety of spatial organizations on the surface of the nanoparticle and evaluated the potency of TLR9 activation induced by each construct. I demonstrate that the magnitude of TLR9 activation can be

controllably modulated by varying the valency of CpG-containing overhangs on PB84, and I show the generalizability of this method by recapitulating these results with a different nanoparticle geometry, ICO42. Furthermore, I show that TLR9 activation can also be tuned by varying the spatial organization of a constant number of CpG overhangs, and additionally demonstrate that displaying CpG overhangs on wireframe DNA nanostructures enables TLR9 activation at significantly lower concentrations than would be necessary for equivalent concentrations of the free CpG oligos, suggesting that the presentation of CpG oligos with pre-arranged spatial configurations enhances TLR9 binding and activation. Finally, I test these engineered TLR9-activating constructs in PBMCs to evaluate their effects on downstream Type I IFN production and show that cytokine expression can be controllably tuned by these constructs as well. These results begin to elucidate the impact of different nanoparticle design parameters on immune activation and demonstrate that wireframe DNA origami can be designed for programmable modulation of the innate immune response.

In Chapter 4, I review the findings and implications of my thesis work and introduce a few questions that have cropped up over the course of these studies. I then discuss some potential explanations and suggest several directions in which this research could be extended in the future, including the design of siRNA-DNA nanoparticles for immunostimulation and efficient gene silencing, which I elaborate on in the Appendix. Finally, I close with a brief outlook on how this thesis work may be leveraged to enhance the therapeutic capabilities and clinical translatability of wireframe DNA origami.

References

(1)   Seeman, N. C. Nucleic Acid Junctions and Lattices. *Journal of Theoretical Biology* **1982**, *99* (2), 237–247. https://doi.org/10.1016/0022-5193(82)90002-9.

(2)   Wang, Y.; Mueller, J. E.; Kemper, B.; Seeman, N. C. Assembly and Characterization of Five-Arm and Six-Arm DNA Branched Junctions. *Biochemistry* **1991**, *30* (23), 5667–5674. https://doi.org/10.1021/bi00237a005.

(3)   Wang, X.; Seeman, N. C. Assembly and Characterization of 8-Arm and 12-Arm DNA Branched Junctions. *J Am Chem Soc* **2007**, *129* (26), 8169–8176. https://doi.org/10.1021/ja0693441.

(4)   Fu, T. J.; Seeman, N. C. DNA Double-Crossover Molecules. *Biochemistry* **1993**, *32* (13), 3211–3220. https://doi.org/10.1021/bi00064a003.

(5)   Lu, M.; Guo, Q.; Seeman, N. C.; Kallenbach, N. R. Parallel and Antiparallel Holliday Junctions Differ in Structure and Stability. *Journal of Molecular Biology* **1991**, *221* (4), 1419–1432. https://doi.org/10.1016/0022-2836(91)90942-Y.

(6)   Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H. DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires. *Science (1979)* **2003**, *301* (5641), 1882–1884. https://doi.org/10.1126/science.1089389.

(7)   He, Y.; Tian, Y.; Ribbe, A. E.; Mao, C. Highly Connected Two-Dimensional Crystals of DNA Six-Point-Stars. *J Am Chem Soc* **2006**, *128* (50), 15978–15979. https://doi.org/10.1021/ja0665141.

(8)   Mathieu, F.; Liao, S.; Kopatsch, J.; Wang, T.; Mao, C.; Seeman, N. C. Six-Helix Bundles Designed from DNA. *Nano Letters* **2005**, *5* (4), 661–665. https://doi.org/10.1021/nl050084f.

(9)   Yan, H.; LaBean, T. H.; Feng, L.; Reif, J. H. Directed Nucleation Assembly of DNA Tile Complexes for Barcode-Patterned Lattices. *Proceedings of the National Academy of Sciences* **2003**, *100* (14), 8103–8108. https://doi.org/10.1073/pnas.1032954100.

(10)   Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440* (7082), 297–302. https://doi.org/10.1038/nature04586.

(11)   Douglas, S. M.; Chou, J. J.; Shih, W. M. DNA-Nanotube-Induced Alignment of Membrane Proteins for NMR Structure Determination. *Proceedings of the National Academy of Sciences* **2007**, *104* (16), 6644–6648. https://doi.org/10.1073/pnas.0700930104.

(12)   Zhang, F.; Jiang, S.; Wu, S.; Li, Y.; Mao, C.; Liu, Y.; Yan, H. Complex Wireframe DNA Origami Nanostructures with Multi-Arm Junction Vertices. *Nature Nanotechnology* **2015**, *10* (9), 779–784. https://doi.org/10.1038/nnano.2015.162.

(13)   Douglas, S. M.; Marblestone, A. H.; Teerapittayanon, S.; Vazquez, A.; Church, G. M.; Shih, W. M. Rapid Prototyping of 3D DNA-Origami Shapes with CaDNAno. *Nucleic Acids Research* **2009**, *37* (15), 5001–5006. https://doi.org/10.1093/nar/gkp436.

(14)   Williams, S.; Lund, K.; Lin, C.; Wonka, P.; Lindsay, S.; Yan, H. Tiamat: A Three-Dimensional Editing Tool for Complex DNA Structures; 2009; pp 90–101. https://doi.org/10.1007/978-3-642-03076-5_8.

(15)   Jun, H.; Shepherd, T. R.; Zhang, K.; Bricker, W. P.; Li, S.; Chiu, W.; Bathe, M. Automated Sequence Design of 3D Polyhedral Wireframe DNA Origami with Honeycomb Edges. *ACS Nano* **2019**, acsnano.8b08671. https://doi.org/10.1021/acsnano.8b08671.

(16)   Jun, H.; Wang, X.; Bricker, W. P.; Bathe, M. Automated Sequence Design of 2D Wireframe DNA Origami with Honeycomb Edges. *Nature Communications* **2019**, *10* (1), 5419. https://doi.org/10.1038/s41467-019-13457-y.

(17)   Jun, H.; Zhang, F.; Shepherd, T.; Ratanalert, S.; Qi, X.; Yan, H.; Bathe, M. Autonomously Designed Free-Form 2D DNA Origami. *Science Advances* **2019**, *5* (1). https://doi.org/10.1126/sciadv.aav0655.

(18)   Veneziano, R.; Ratanalert, S.; Zhang, K.; Zhang, F.; Yan, H.; Chiu, W.; Bathe, M. Designer Nanoscale DNA Assemblies Programmed from the Top Down. *Science (1979)* **2016**, *352* (6293), 1534–1534. https://doi.org/10.1126/science.aaf4388.

(19)   de Llano, E.; Miao, H.; Ahmadi, Y.; Wilson, A. J.; Beeby, M.; Viola, I.; Barisic, I. Adenita: Interactive 3D Modelling and Visualization of DNA Nanostructures. *Nucleic Acids Research* **2020**, *48* (15), 8269–8275. https://doi.org/10.1093/nar/gkaa593.

(20)   Jun, H.; Wang, X.; Parsons, M. F.; Bricker, W. P.; John, T.; Li, S.; Jackson, S.; Chiu, W.; Bathe, M. Rapid Prototyping of Arbitrary 2D and 3D Wireframe DNA Origami. *Nucleic Acids Research* **2021**, *49* (18), 10265–10274. https://doi.org/10.1093/nar/gkab762.

(21)   Veneziano, R.; Shepherd, T. R.; Ratanalert, S.; Bellou, L.; Tao, C.; Bathe, M. In Vitro Synthesis of Gene-Length Single-Stranded DNA. *Scientific Reports* **2018**, *8* (1), 6548. https://doi.org/10.1038/s41598-018-24677-5.

(22)   Kosuri, S.; Church, G. M. Large-Scale de Novo DNA Synthesis: Technologies and Applications. *Nature Methods* **2014**, *11* (5), 499–507. https://doi.org/10.1038/nmeth.2918.

(23)   Nehdi, A.; Samman, N.; Aguilar-Sánchez, V.; Farah, A.; Yurdusev, E.; Boudjelal, M.; Perreault, J. Novel Strategies to Optimize the Amplification of Single-Stranded DNA. *Frontiers in Bioengineering and Biotechnology* **2020**, *8*. https://doi.org/10.3389/fbioe.2020.00401.

(24)   Lee, B.-Y.; Lee, J.; Ahn, D. J.; Lee, S.; Oh, M.-K. Optimizing Protein V Untranslated Region Sequence in M13 Phage for Increased Production of Single-Stranded DNA for Origami. *Nucleic Acids Research* **2021**, *49* (11), 6596–6603. https://doi.org/10.1093/nar/gkab455.

(25)   Gailus, V.; Ramsperger, U.; Johner, C.; Kramer, H.; Rasched, I. The Role of the Absorption Complex in the Termination of Filamentous Phage Assembly. *Research in Microbiology* **1994**, *145* (9), 699–709. https://doi.org/10.1016/0923-2508(94)90042-6.

(26)   Smeal, S. W.; Schmitt, M. A.; Pereira, R. R.; Prasad, A.; Fisk, J. D. Simulation of the M13 Life Cycle I: Assembly of a Genetically-Structured Deterministic Chemical Kinetic Simulation. *Virology* **2017**, *500*, 259–274. https://doi.org/10.1016/j.virol.2016.08.017.

(27)   Filamentous Bacteriophage: Biology, Phage Display and Nanotechnology Applications. *Current Issues in Molecular Biology* **2011**. https://doi.org/10.21775/cimb.013.051.

(28)   Chasteen, L.; Ayriss, J.; Pavlik, P.; Bradbury, A. R. M. Eliminating Helper Phage from Phage Display. *Nucleic Acids Research* **2006**, *34* (21), e145–e145. https://doi.org/10.1093/nar/gkl772.

(29)   Dotto, G. P.; Horiuchi, K.; Zinder, N. D. Initiation and Termination of Phage F1 Plus-Strand Synthesis. *Proceedings of the National Academy of Sciences* **1982**, *79* (23), 7122–7126. https://doi.org/10.1073/pnas.79.23.7122.

(30)   Specthrie, L.; Bullitt, E.; Horiuchi, K.; Model, P.; Russel, M.; Makowski, L. Construction of a Microphage Variant of Filamentous Bacteriophage. *Journal of Molecular Biology* **1992**, *228* (3), 720–724. https://doi.org/10.1016/0022-2836(92)90858-H.

(31)   Nafisi, P. M.; Aksel, T.; Douglas, S. M. Construction of a Novel Phagemid to Produce Custom DNA Origami Scaffolds. *Synthetic Biology* **2018**, *3* (1). https://doi.org/10.1093/synbio/ysy015.

(32)   Praetorius, F.; Kick, B.; Behler, K. L.; Honemann, M. N.; Weuster-Botz, D.; Dietz, H. Biotechnological Mass Production of DNA Origami. *Nature* **2017**, *552* (7683), 84–87. https://doi.org/10.1038/nature24650.

(33)   Paludan, S. R.; Bowie, A. G. Immune Sensing of DNA. *Immunity* **2013**, *38* (5), 870–880. https://doi.org/10.1016/j.immuni.2013.05.004.

(34) Schlee, M.; Hartmann, G. Discriminating Self from Non-Self in Nucleic Acid Sensing. *Nature Reviews Immunology* **2016**, *16* (9), 566–580. https://doi.org/10.1038/nri.2016.78.

(35) Kawasaki, T.; Kawai, T. Toll-Like Receptor Signaling Pathways. *Frontiers in Immunology* **2014**, *5*. https://doi.org/10.3389/fimmu.2014.00461.

(36) Bauer, S.; Kirschning, C. J.; Häcker, H.; Redecke, V.; Hausmann, S.; Akira, S.; Wagner, H.; Lipford, G. B. Human TLR9 Confers Responsiveness to Bacterial DNA via Species-Specific CpG Motif Recognition. *Proceedings of the National Academy of Sciences* **2001**, *98* (16), 9237–9242. https://doi.org/10.1073/pnas.161293498.

(37) Latz, E.; Schoenemeyer, A.; Visintin, A.; Fitzgerald, K. A.; Monks, B. G.; Knetter, C. F.; Lien, E.; Nilsen, N. J.; Espevik, T.; Golenbock, D. T. TLR9 Signals after Translocating from the ER to CpG DNA in the Lysosome. *Nature Immunology* **2004**, *5* (2), 190–198. https://doi.org/10.1038/ni1028.

(38) Lee, B. L.; Moon, J. E.; Shu, J. H.; Yuan, L.; Newman, Z. R.; Schekman, R.; Barton, G. M. UNC93B1 Mediates Differential Trafficking of Endosomal TLRs. *Elife* **2013**, *2*. https://doi.org/10.7554/eLife.00291.

(39) Marongiu, L.; Gornati, L.; Artuso, I.; Zanoni, I.; Granucci, F. Below the Surface: The Inner Lives of TLR4 and TLR9. *Journal of Leukocyte Biology* **2019**, *106* (1), 147–160. https://doi.org/10.1002/JLB.3MIR1218-483RR.

(40) Tan, X.; Sun, L.; Chen, J.; Chen, Z. J. Detection of Microbial Infections Through Innate Immune Sensing of Nucleic Acids. *Annual Review of Microbiology* **2018**, *72* (1), 447–478. https://doi.org/10.1146/annurev-micro-102215-095605.

(41) Ohto, U.; Shibata, T.; Tanji, H.; Ishida, H.; Krayukhina, E.; Uchiyama, S.; Miyake, K.; Shimizu, T. Structural Basis of CpG and Inhibitory DNA Recognition by Toll-like Receptor 9. *Nature* **2015**, *520* (7549), 702–705. https://doi.org/10.1038/nature14138.

(42) Ohto, U.; Ishida, H.; Shibata, T.; Sato, R.; Miyake, K.; Shimizu, T. Toll-like Receptor 9 Contains Two DNA Binding Sites That Function Cooperatively to Promote Receptor Dimerization and Activation. *Immunity* **2018**, *48* (4), 649-658.e4. https://doi.org/10.1016/j.immuni.2018.03.013.

(43) Comberlato, A.; Koga, M. M.; Nüssing, S.; Parish, I. A.; Bastings, M. M. C. Spatially Controlled Activation of Toll-like Receptor 9 with DNA-Based Nanomaterials. *Nano Letters* **2022**, acs.nanolett.2c00275. https://doi.org/10.1021/acs.nanolett.2c00275.

(44) Chi, Q.; Wang, G.; Jiang, J. The Persistence Length and Length per Base of Single-Stranded DNA Obtained from Fluorescence Correlation Spectroscopy Measurements Using Mean Field Theory. *Physica A: Statistical Mechanics and its Applications* **2013**, *392* (5), 1072–1079. https://doi.org/10.1016/j.physa.2012.09.022.

(45) Ambia-Garrido, J.; Vainrub, A.; Pettitt, B. M. A Model for Structure and Thermodynamics of SsDNA and DsDNA near a Surface: A Coarse Grained Approach. *Computer Physics Communications* **2010**, *181* (12), 2001–2007. https://doi.org/10.1016/j.cpc.2010.08.029.

(46) Roth, E.; Glick Azaria, A.; Girshevitz, O.; Bitler, A.; Garini, Y. Measuring the Conformation and Persistence Length of Single-Stranded DNA Using a DNA Origami Structure. *Nano Letters* **2018**, *18* (11), 6703–6709. https://doi.org/10.1021/acs.nanolett.8b02093.

(47) Pohar, J.; Yamamoto, C.; Fukui, R.; Cajnko, M.-M.; Miyake, K.; Jerala, R.; Benčina, M. Selectivity of Human TLR9 for Double CpG Motifs and Implications for the Recognition of Genomic DNA. *The Journal of Immunology* **2017**, *198* (5), 2093–2104. https://doi.org/10.4049/jimmunol.1600757.

(48) Li, E.; Zhang, Y. DNA Methylation in Mammals. *Cold Spring Harbor Perspectives in Biology* **2014**, *6* (5), a019133–a019133. https://doi.org/10.1101/cshperspect.a019133.

(49)   Walsh, C. P.; Xu, G. L. Cytosine Methylation and DNA Repair. In *DNA Methylation: Basic Mechanisms*; Springer-Verlag: Berlin/Heidelberg; pp 283–315. https://doi.org/10.1007/3-540-31390-7_11.

(50)   Pohar, J.; Kužnik Krajnik, A.; Jerala, R.; Benčina, M. Minimal Sequence Requirements for Oligodeoxyribonucleotides Activating Human TLR9. *The Journal of Immunology* **2015**, *194* (8), 3901–3908. https://doi.org/10.4049/jimmunol.1402755.

(51)   Krieg, A. M.; Yi, A.-K.; Matson, S.; Waldschmidt, T. J.; Bishop, G. A.; Teasdale, R.; Koretzky, G. A.; Klinman, D. M. CpG Motifs in Bacterial DNA Trigger Direct B-Cell Activation. *Nature* **1995**, *374* (6522), 546–549. https://doi.org/10.1038/374546a0.

(52)   Hartmann, G.; Battiany, J.; Poeck, H.; Wagner, M.; Kerkmann, M.; Lubenow, N.; Rothenfusser, S.; Endres, S. Rational Design of New CpG Oligonucleotides That Combine B Cell Activation with High IFN-α Induction in Plasmacytoid Dendritic Cells. *European Journal of Immunology* **2003**, *33* (6), 1633–1641. https://doi.org/10.1002/eji.200323813.

(53)   Vollmer, J.; Weeratna, R.; Payette, P.; Jurk, M.; Schetter, C.; Laucht, M.; Wader, T.; Tluk, S.; Liu, M.; Davis, H. L.; Krieg, A. M. Characterization of Three CpG Oligodeoxynucleotide Classes with Distinct Immunostimulatory Activities. *European Journal of Immunology* **2004**, *34* (1), 251–262. https://doi.org/10.1002/eji.200324032.

(54)   Pohar, J.; Kužnik Krajnik, A.; Jerala, R.; Benčina, M. Minimal Sequence Requirements for Oligodeoxyribonucleotides Activating Human TLR9. *The Journal of Immunology* **2015**, *194* (8), 3901–3908. https://doi.org/10.4049/jimmunol.1402755.

(55)   Baranovskii, A. G.; Buneva, V. N.; Nevinsky, G. A. Human Deoxyribonucleases. *Biochemistry (Moscow)* **2004**, *69* (6), 587–601. https://doi.org/10.1023/B:BIRY.0000033731.50496.01.

(56)   Wu, J.; Chen, Z. J. Innate Immune Sensing and Signaling of Cytosolic Nucleic Acids. *Annual Review of Immunology* **2014**, *32* (1), 461–488. https://doi.org/10.1146/annurev-immunol-032713-120156.

(57)   Yoneyama, M.; Kikuchi, M.; Natsukawa, T.; Shinobu, N.; Imaizumi, T.; Miyagishi, M.; Taira, K.; Akira, S.; Fujita, T. The RNA Helicase RIG-I Has an Essential Function in Double-Stranded RNA-Induced Innate Antiviral Responses. *Nature Immunology* **2004**, *5* (7), 730–737. https://doi.org/10.1038/ni1087.

(58)   Yoneyama, M.; Kikuchi, M.; Matsumoto, K.; Imaizumi, T.; Miyagishi, M.; Taira, K.; Foy, E.; Loo, Y.-M.; Gale, M.; Akira, S.; Yonehara, S.; Kato, A.; Fujita, T. Shared and Unique Functions of the DExD/H-Box Helicases RIG-I, MDA5, and LGP2 in Antiviral Innate Immunity. *The Journal of Immunology* **2005**, *175* (5), 2851–2858. https://doi.org/10.4049/jimmunol.175.5.2851.

(59)   Wan, D.; Jiang, W.; Hao, J. Research Advances in How the CGAS-STING Pathway Controls the Cellular Inflammatory Response. *Frontiers in Immunology* **2020**, *11*. https://doi.org/10.3389/fimmu.2020.00615.

(60)   Unterholzner, L.; Keating, S. E.; Baran, M.; Horan, K. A.; Jensen, S. B.; Sharma, S.; Sirois, C. M.; Jin, T.; Latz, E.; Xiao, T. S.; Fitzgerald, K. A.; Paludan, S. R.; Bowie, A. G. IFI16 Is an Innate Immune Sensor for Intracellular DNA. *Nature Immunology* **2010**, *11* (11), 997–1004. https://doi.org/10.1038/ni.1932.

(61)   Motwani, M.; Pesiridis, S.; Fitzgerald, K. A. DNA Sensing by the CGAS–STING Pathway in Health and Disease. *Nature Reviews Genetics* **2019**, *20* (11), 657–674. https://doi.org/10.1038/s41576-019-0151-1.

(62)   Zhang, X.; Wu, J.; Du, F.; Xu, H.; Sun, L.; Chen, Z.; Brautigam, C. A.; Zhang, X.; Chen, Z. J. The Cytosolic DNA Sensor CGAS Forms an Oligomeric Complex with DNA and Undergoes

Switch-like Conformational Changes in the Activation Loop. *Cell Reports* **2014**, *6* (3), 421–430. https://doi.org/10.1016/j.celrep.2014.01.003.

(63) Li, X.; Shu, C.; Yi, G.; Chaton, C. T.; Shelton, C. L.; Diao, J.; Zuo, X.; Kao, C. C.; Herr, A. B.; Li, P. Cyclic GMP-AMP Synthase Is Activated by Double-Stranded DNA-Induced Oligomerization. *Immunity* **2013**, *39* (6), 1019–1031. https://doi.org/10.1016/j.immuni.2013.10.019.

(64) Herzner, A.-M.; Hagmann, C. A.; Goldeck, M.; Wolter, S.; Kübler, K.; Wittmann, S.; Gramberg, T.; Andreeva, L.; Hopfner, K.-P.; Mertens, C.; Zillinger, T.; Jin, T.; Xiao, T. S.; Bartok, E.; Coch, C.; Ackermann, D.; Hornung, V.; Ludwig, J.; Barchet, W.; Hartmann, G.; Schlee, M. Sequence-Specific Activation of the DNA Sensor CGAS by Y-Form DNA Structures as Found in Primary HIV-1 CDNA. *Nature Immunology* **2015**, *16* (10), 1025–1033. https://doi.org/10.1038/ni.3267.

(65) Andreeva, L.; Hiller, B.; Kostrewa, D.; Lässig, C.; de Oliveira Mann, C. C.; Jan Drexler, D.; Maiser, A.; Gaidt, M.; Leonhardt, H.; Hornung, V.; Hopfner, K.-P. CGAS Senses Long and HMGB/TFAM-Bound U-Turn DNA by Forming Protein–DNA Ladders. *Nature* **2017**, *549* (7672), 394–398. https://doi.org/10.1038/nature23890.

(66) Fernandes-Alnemri, T.; Yu, J.-W.; Datta, P.; Wu, J.; Alnemri, E. S. AIM2 Activates the Inflammasome and Cell Death in Response to Cytoplasmic DNA. *Nature* **2009**, *458* (7237), 509–513. https://doi.org/10.1038/nature07710.

(67) Fernandes-Alnemri, T.; Wu, J.; Yu, J.-W.; Datta, P.; Miller, B.; Jankowski, W.; Rosenberg, S.; Zhang, J.; Alnemri, E. S. The Pyroptosome: A Supramolecular Assembly of ASC Dimers Mediating Inflammatory Cell Death via Caspase-1 Activation. *Cell Death & Differentiation* **2007**, *14* (9), 1590–1604. https://doi.org/10.1038/sj.cdd.4402194.

(68) Jin, T.; Perry, A.; Jiang, J.; Smith, P.; Curry, J. A.; Unterholzner, L.; Jiang, Z.; Horvath, G.; Rathinam, V. A.; Johnstone, R. W.; Hornung, V.; Latz, E.; Bowie, A. G.; Fitzgerald, K. A.; Xiao, T. S. Structures of the HIN Domain:DNA Complexes Reveal Ligand Binding and Activation Mechanisms of the AIM2 Inflammasome and IFI16 Receptor. *Immunity* **2012**, *36* (4), 561–571. https://doi.org/10.1016/j.immuni.2012.02.014.

Chapter 2

Bioproduction of pure, kilobase-scale

single-stranded DNA

ABSTRACT

Scalable production of kilobase single-stranded DNA (ssDNA) with sequence control has applications in therapeutics, gene synthesis and sequencing, scaffolded DNA origami, and archival DNA memory storage. Biological production of circular ssDNA (cssDNA) using M13 addresses these needs at low cost. However, one unmet goal is to minimize the essential protein coding regions of the exported DNA while maintaining its infectivity and production purity to produce sequences less than 3,000 nt in length, relevant to therapeutic and materials science applications. Toward this end, synthetic miniphage with inserts of custom sequence and size offers scalable, low-cost synthesis of cssDNA at milligram and higher scales. Here, we optimize growth conditions using an E. coli helper strain combined with a miniphage genome carrying only an f1 origin and a β-lactamase-encoding (bla) antibiotic resistance gene, enabling isolation of pure cssDNA with a minimum sequence genomic length of 1,676 nt, without requiring additional purification from contaminating DNA. Low-cost scalability of isogenic, custom-length cssDNA is demonstrated for a sequence of 2,520 nt using a bioreactor, purified with low endotoxin levels (<5 E.U./ml). We apply these exonuclease-resistant cssDNAs to the self-assembly of wireframe DNA origami objects and to encode digital information on the miniphage genome for biological amplification.

INTRODUCTION

Kilobase-length, single-stranded DNA (ssDNA) is essential to numerous biotechnological applications including sequencing (1), cloning (2), homology directed repair templating for gene editing (3), DNA-based digital information storage (4,5), and scaffolded DNA origami (6-9). Specifically, scaffolded DNA origami enables the fabrication of custom structured nanoscale objects with application to nanoscale lithography (10,11), light harvesting and nanoscale energy transport (12-15), metal nanoparticle casting (16), and therapeutic delivery (17,18). In this approach, a long ssDNA scaffold is folded via self-assembly into arbitrary, user-specified shapes by slow annealing in the presence of complementary short oligonucleotide "staples". These staples are designed using Watson-Crick base-pair complementarity to the scaffold, forcing sequences that are far apart in sequence space to be close in physical space. Fully automated, top-down computational design of scaffolded DNA origami nanostructures has now been enabled by sequence design algorithms in both 2D and 3D (19-23), enabling the democratization of otherwise complex scaffolded DNA origami design that previously excluded non-experts. However, therapeutic and materials science applications that require large-scale, low-cost scaffold with custom length and sequence requirements are still hindered by limitations in production of circular, isogenic scaffolds on the 1–3 kb scale (24-26).

The most common low-cost source of native cssDNA for scaffolded DNA origami is the 7,249 base single-stranded M13mp18 phage genome, which has allowed production of up to 410 mg of cssDNA per liter of *E. coli* growth through fed-batch fermentation (27). Innovative approaches to achieving custom bacterial scaffolds (25) and staple production were recently implemented (26). However, numerous applications require increased purity of circular custom scaffolds, such as in

53

therapeutic applications of DNA origami, that would additionally require synthetic staples for chemical functionalization and stabilization (28). Further, elimination of the antibiotic selection marker from the produced linear (26) or circular (25) ssDNA does not allow for subsequent re-infection without this selective control (25), potentially limiting downstream biological applications where reinfection would be useful, such as for biological sequence amplification.

These preceding advances employed helper plasmid systems where the M13 coding sequences are sub-cloned onto a double-stranded, low-copy number vector that is co-transformed with a phagemid containing an ssDNA origin of replication (e.g., f1 origin) that allows for the synthesis and packaging of ssDNA. The most commonly used helper plasmid system is M13KO7 (29), which maintains a packaging sequence, albeit with a mutated packaging signal to reduce the packaging frequency. This system has shown utility in phage display (30-32), and has also been applied to produce a phagemid that encodes a 2,404-nt sequence containing an cssDNA f1 origin, a dsDNA pUC origin, and an ampicillin selection marker (24). However, the preceding phagemid cssDNA was contaminated by other DNAs, both from the dsDNA phagemid and M13KO7 (24). Importantly, isolation of the target ssDNA from these DNA impurities would require subsequent purification steps for further bioproduction scale-up, and would potentially introduce high background in sequencing application, with possible errors in phage propagation, and lower yields in DNA origami applications. To overcome these potential problems with the helper system, the ssDNA origin of replication and packaging signal was entirely removed from a helper plasmid (*E. coli* str. M13cp) (33), thereby enabling the biological production of isogenic cssDNA without DNA impurities. Indeed, recent advances in engineered phage systems have utilized this strategy (25,26), however, purification would ultimately be required for DNAzyme-based approaches to

bacterial scaffold as well as staple production (26) and optimization for pure cssDNA production (34) without genomic or plasmid contamination is still required (25,26). Thus, there remains a critical need for scalable production of isogenic cssDNA at the 1–3 kb length scale that maintains replication capacity.

To address this need, in the present work we show that isogenic miniphage production of cssDNA is scalable by fermentation using the *E. coli* str. SS320 with the M13cp helper plasmid (generous gift of Dr. Andrew Bradbury, Los Alamos National Lab). Three miniphages were synthesized using both classic restriction and restriction-free (RF) cloning (35,36). The miniphage presented here maintain the selection marker and origin of replication, which allows for the reinfection of the phage in culture while reducing the occurrence of contaminating dsDNA because they do not contain a double-strand origin of replication, similar to the natural M13 phage. Monitoring phage yields and growth rates of the bacteria, we identify an 8-hour timepoint after inoculation that yields maximal cssDNA production with no detectible DNA contamination. Silica-column-based DNA extraction techniques from clarified media yielded 2 mg of pure cssDNA per liter of culture with < 5 endotoxin units per milliliter of sample. We demonstrate that the cssDNA material is practical for the generation of custom length circular scaffolds for wireframe DNA origami with partial sequence control, with additional applications to write-once, read-many archival DNA data storage.

RESULTS

Proof-of-concept circular ssDNA synthesis and production

A variant of extension-overlap, restriction-free cloning (35,36) using long ssDNA (**Figure 2.1a**) was applied for the *de novo* assembly of a miniphage genome containing only an f1 origin of replication and an ampicillin resistance selection marker (phPB52). Two kilobase-scale megaprimer ssDNAs were generated using asymmetric PCR (aPCR; (37)) using 5′-phosphorylated primers: a top-strand megaprimer encoding the f1 origin sequence (427 nt; (38)) and a bottom-strand megaprimer encoding an ampicillin resistance cistron (*bla*; 1,249 nt) (**Figure 2.1b, Figure A.3**). The kilobase primers were synthesized such that the two sequences contained a complementary sequence of 20 nt on each of the 5′ and 3′ ends (**Figure 2.1a**). The two megaprimers were mixed at equimolar concentration and completed to dsDNA using PCR, followed by enzymatic ligation. The ligated plasmid (**Figure A.4**) was transformed into chemically competent *E. coli* str. M13cp (33) and dual selected on ampicillin and chloramphenicol with no detectable toxicity due to the miniphage, resulting in normal colony shape and size. Two out of eight colonies screened were found to be of the exact sequence desired. Liquid culture was inoculated and grown in a shaker flask, after which the culture was centrifuged to separate the phage-containing media from the bacterial pellet. Phage in the clarified media were visualized by TEM, showing the anticipated size of and homogeneity (**Figure 2.1c and Figure A.3**). The cssDNA from this phage was isolated using silica column purification and showed 88% cssDNA purity according to agarose gel imaging (**Figure 2.1d**), with an approximate yield of 0.5 mg per liter of bacterial growth, while the bacterial pellet showed helper plasmid, dsDNA intermediate phage DNA, and cssDNA.
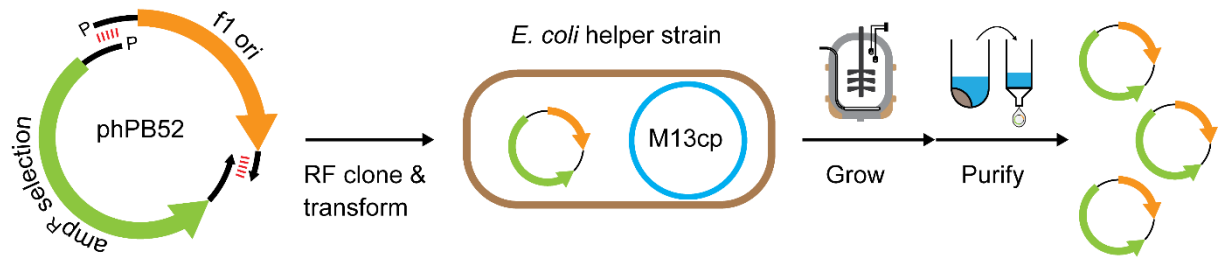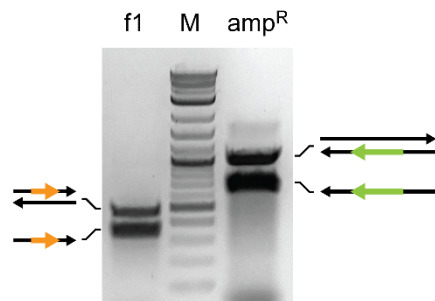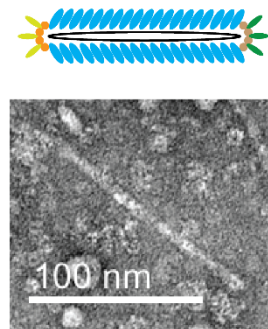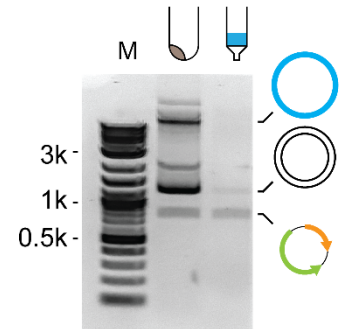
Figure 2.1. Scalable bacteriophage production of isogenic cssDNA. (**a**) Miniphage phPB52 was assembled using restriction-free (RF) cloning was used for miniphage phPB52 assembly and transformed into *E. coli* containing the M13cp helper plasmid for production of isogenic cssDNA. (**b**) aPCR was used to generate the two ssDNA megaprimers for RF cloning encoding the f1 origin of replication (f1 ori) and the *bla* ampicillin selection marker (Selection). (**c**) Phage particles from clarified media were visualized by TEM. (**d**) DNA purification from the bacterial pellet and the clarified media show mostly pure cssDNA in the media and cssDNA and dsDNA phagemid, and helper plasmid in the bacterial pellet. See **Figure A.3** for uncropped images.

Having generated a phage containing only the f1 origin and a resistance gene, demonstrated to be stably produced and exported to the media from the helper strain, we next sought to generate a second phage with a synthetic fragment of DNA that is orthogonal in sequence to bacterial and phage genomes. A fragment of length 844 nt was ligated between the f1 origin and the *bla* cistron using standard restriction cloning to generated a plasmid of size 2,520 nt (phPB84; **Figure 2.2a** and **Figure A.4**). This plasmid was transformed into the helper strain and the produced phage was

purified and its sequence verified by primer walking with Sanger sequencing (**Appendix Table 5 and 6**).

Development of a shaker-flask production method

In order to obtain milligram-scale production of cssDNA with high genetic purity of the final material, we used a shaker flask setup (**Figure 2.1a**) to vary the growth time, the *E. coli* strain, the growth media, and the media pH to determine optimal conditions. We found the highest and purest yield of cssDNA production occurred at the 8-hour timepoint after inoculation, near the end of log phase, with production falling off thereafter and the appearance of dsDNA contaminations in the media visualized at the 12-hour timepoint (**Figure 2.2b** and **Figure A.5**). Two strains were tested for production: DH5a F′Iq (Invitrogen) and the SS320 strain (Invitrogen). Both express the F pili and are commonly used for phage production, and each was transformed with the M13cp helper plasmid purified from *E. coli* str. M13cp. Strain SS320 showed approximately double the cssDNA yield (**Figure 2.2c** and **Figure A.6**) and was therefore chosen as the strain for further optimization of growth conditions. Terrific broth (TB) and 2× yeast extract tryptone (2×YT) media for bacterial growth and cssDNA production were both evaluated for phage growth while also monitoring dsDNA contamination using agarose gel analysis (**Figure 2.2d** and **Figure A.6**). Notably, TB had significant dsDNA contamination by the 8-hour timepoint (**Figure A.6**), and 2×YT was therefore chosen as the optimal media for batch production. Next, we investigated the pH sensitivity of the production of phage material (34), which exhibited a three-fold increase in yield at pH 6.8 and 7.2 compared to pH 8 (**Figure 2.2e** and **Figure A.6**).

58
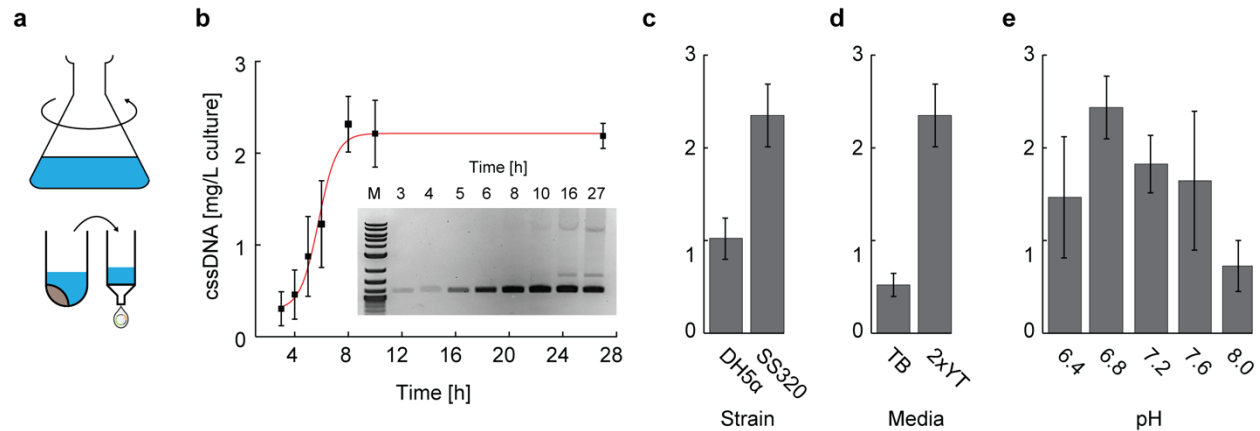
Figure 2.2. Shaker flask production of pure cssDNA. (**a**) Shaker flask growth of phPB84 was used to optimize conditions for phage amounts and purity. (**b**) Time-course assay of cssDNA production of phPB84, with cssDNA yield calculated by absorbance at 280 nm and purity adjusted by agarose gel band intensity, showing maximum yield and purity at the 8-hour timepoint. The 16 h time-point is from a separate culture, and therefore is not included in the plot. (**c**) Comparison between DH5a F′Iq and SS320 showing two-fold yield increases in the SS320 strain. (**d**) Comparison between growth media showing five-fold improved cssDNA yield in 2×YT after 8 hours of production. (**e**) Comparison of five pH values for cssDNA production, controlled by use of 100 mM HEPES-NaOH. Error bars indicate standard deviation of triplicate experiments. See **Figure A.5** and **A.6** for uncropped triplicate measurements of all experiments.

Establishment of a batch fermentation protocol

Having identified the optimal growth conditions in the shaker flask setup, we next identified conditions for scale-up in a batch fermenter process (**Figure 2.3a**) using a Stedium Sartorius 5L fermenter (Sartorius, Germany). Shaker flask conditions were transferred to the bioreactor setup including using 2×YT media, while pH 7.0 was controlled using external phosphoric acid and ammonium hydroxide. The growth curve was monitored using O.D.600 absorbance measurements and the pH and dissolved oxygen were monitored by calibrated probes. Each timepoint was additionally monitored for cssDNA and dsDNA production using agarose gel analysis (**Figure**

2.3b and **Figure A.7**), showing maximal cssDNA yield at the 8-hour timepoint, as with the shaker flask, with minimal contaminating dsDNA up to the 12-hour timepoint (**Figure A.7**). Extraction of 900 mL of media for phage purification was carried out at the 8-hour timepoint and processed using a silica-column based approach specifically designed to reduce endotoxin levels (EndoFree Megaprep Kit, Qiagen, MD). Gel band intensity analysis after kit purification showed no detectable dsDNA contamination (**Figure 2.3c**). Sanger sequencing by primer walking verified the sequence of the phage DNA (**Appendix Table 5 and 6**). The kit-based purification yielded 2 mg of cssDNA/L of culture, matching the yield from phenol-chloroform extraction. Endotoxins were tested using a colorimetric assay (ToxinSensor Chromogenic LAL Endotoxin Assay Kit, GenScript, NJ), showing the final product yielded endotoxin levels at 1.1 ± 0.1 E.U./ml per cssDNA concentration of 10 nM, similar to endotoxin reduction by Triton-X114 ((39); **Figure A.8**). Circularization of the produced cssDNA was verified by incubation with exonuclease I, showing no detectable degradation after 30 min (**Figure 2.3d**).
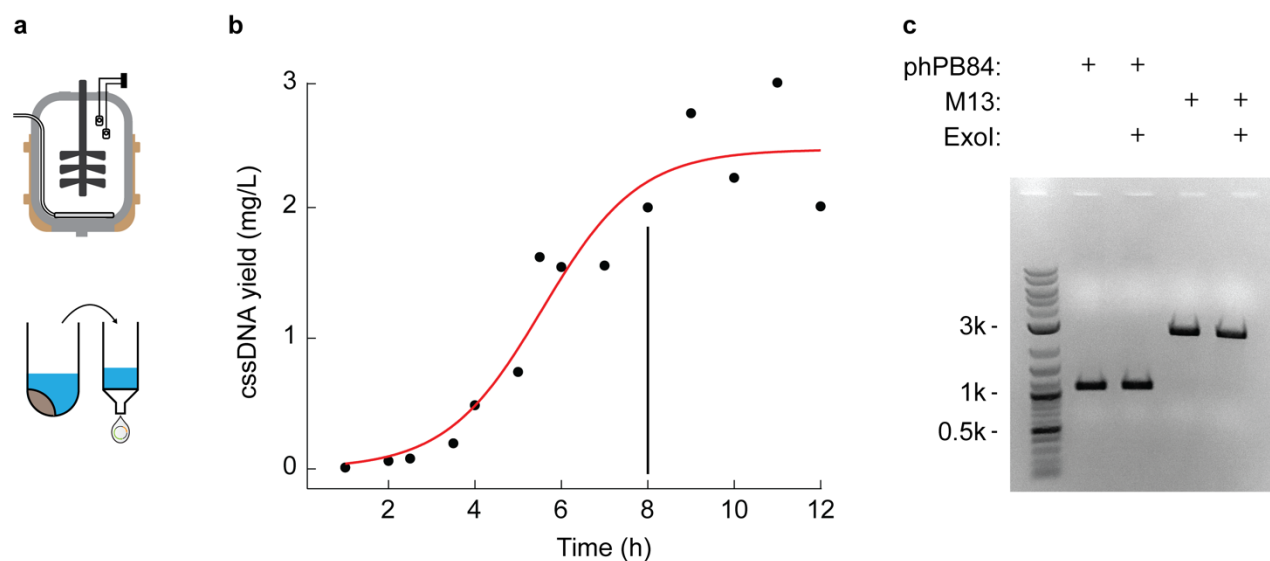
Figure 2.3. Batch fermenter production of pure cssDNA. (**a**) Scalable production in a stirred-tank bioreactor. (**b**) Time-course assay of cssDNA yield based on agarose gel band intensity analysis (**Figure A.7**), with the 8-hour timepoint used for 900 mL cssDNA purification. (**c**) Silica-column DNA purification from the PEG-precipitated phPB84 phage showed no detectible dsDNA contamination, similar in purity to commercially available M13mp18, yielding 2 mg of DNA per liter of culture at the 8-hour timepoint. Stability from exonuclease I (ExoI) degradation after 30 min coincubation indicates the ssDNA is circular.

Application of synthesized scaffolds to DNA origami production

Having implemented a method for milligram-scale production of isogenic miniphage cssDNA, we next applied the method to produce custom length single-stranded DNA scaffold with partial sequence control for application to wireframe scaffolded DNA origami (**Figure 2.4**). We used the DAEDALUS design algorithm (19) to design a DNA-scaffolded pentagonal bipyramid with a 52-bp edge length (1,580 nt scaffold length) using the smallest phPB52 phage genome sequence (1,676 nt) and a second DNA-scaffolded pentagonal bipyramid with an 84-bp edge length (2,520 nt scaffold length) using the phPB84 phage genome sequence (2,520 nt). Notably, any DNA origami with scaffold lengths larger than 1,676 nt can have perfectly matched phage genome

lengths, as exemplified in the pentagonal bipyramid with an 84-bp edge length. DNA origami object folding was characterized using agarose gel mobility shift assays and transmission electron microscopy (TEM), which confirmed monodispersed object sizes with near quantitative yield of self-assembly (**Figure 2.4** and **Figure A.9** and **A.10**).
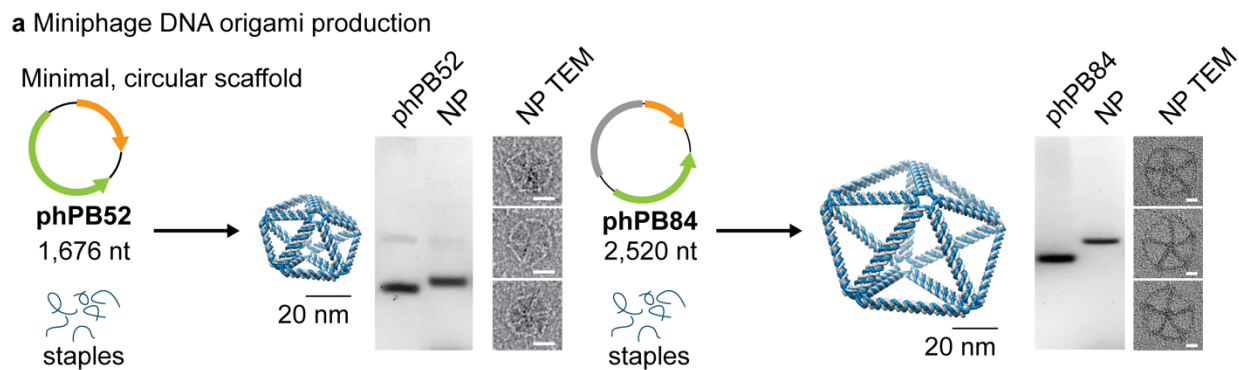
Figure 2.4. Applications of scalable, isogenic miniphage production. Two pentagonal bipyramids of 52-bp and 84-bp edge-lengths were folded using the phPB52 and phPB84 as scaffolds, respectively. Agarose gel shift mobility assays and TEM were used to validate the folding of the scaffold to the expected design. See **Figure A.9** and **A.10** for uncropped gel images and example full field TEM micrographs, respectively. Figure adapted from original Figure 4 of this work.

DISCUSSION

We applied the *E. coli* str. M13cp strain for scalable bioproduction of pure cssDNA, which has the capabilities of generating isogenic material for biotechnological applications including scaffolded DNA origami and digital information archival and amplification, amongst other uses. The method employed here to direct purification of phage cssDNA without additional dsDNA contamination allows for new technology development in synthetic cssDNA sequence production that can be made bio-orthogonal and scalable, enabling future application to novel therapeutics and materials. Additional advances in the scaffolded DNA origami field are applicable to this strain, including the incorporation of DNAzymes (26) that would allow for greater control over the sequence and size of the produced linear ssDNA. However, in the approach used here, maintenance of the f1 origin and the selection marker in the produced phage allows for reinfection across the culture, which is important for subsequent biological amplification such as needed in phage display and archival information storage. Moreover, circularization blocks exonuclease activity, which may prove important for therapeutic applications (41). In the future, improved understanding of phage biology should enable new approaches to excising specific coding sequences from M13 to generate engineered systems specifically designed for production of cssDNA.

The yields from the bioreactor approach used here were somewhat lower than wild type phage production that has been extensively optimized (26,27). This is due in part to the loss of the native feedback control over gene expression in the phage genome (42), the use of batch fermentation as opposed to a fed-batch approach that would allow for higher cell density (27), and plasmid loss due to ampicillin selection. Interestingly, we were not able to obtain clones of kanamycin or chloramphenicol selection cistrons on the vector purely under the control of the f1 origin. This

may be due to the use of a single-stranded promoter, which might be overcome by alternative single-strand-specific promoters (43). This resistance insertion would then allow for fed-batch scale-up, leading to significantly improved yields.

Increased cssDNA production yields, together with advances in custom sequence design (25,26) and bio-orthogonality (44,45) with and without protein coding sequences, suggest that our approach is amenable to therapeutic applications in which ssDNA are used in circular (46) or linear (47) forms. In particular, scalable production of pure ssDNA at lower costs could enable yields required for therapeutic dosages of kilobase-length HDR template strands (48), a strategy that is further enabled by applying a DNAzyme approach for linearization (26). Scalable biological production of scaffolded DNA origami now matches production amounts from solid-state DNA synthesis commonly used for staple production, so that scaffolded DNA origami nanoparticles may now be produced at reasonable cost for mouse and higher animal therapeutic studies. Staple sequences synthesized with modifications to improve staple stability may further enhance nanoparticle lifetimes (28).

MATERIALS AND METHODS

Plasmid assembly by single-stranded DNA. All sequences of phage genomes and primers are contained in **Appendix Tables 5** and **6**. The sequence of the f1 origin of replication was ordered from Integrated DNA Technologies (IDT, Inc., Coralville, IA) as a gBlock™ with 20 nt primers flanking the 5′ and 3′ sides designed to have a calculated melting temperature of 57ºC (50). Double stranded DNA was generated by amplification of the synthetic gBlock f1 sequence with Phusion™ polymerase (New England Biolabs, Inc., Ipswitch, MA). The beta-lactamase (*bla*) ampicillin resistance gene with its promoter and terminator sequences were amplified from pUC19 using Phusion™ polymerase and 5′ and 3′ primers extended on their 5′ by the complementary pair of the f1 gBlock fragment. In each case, the PCR-amplified material was purified by ZymoClean agarose gel purification (Zymo Research, Inc., Irvine, CA) and column cleanup (Qiagen miniprep spin purification kit, Qiagen, Inc., Germany). Single-stranded DNA was generated using asymmetric production with 200 ng of purified dsDNA and 1 μM 5′-phosphorylated primer and Accustart HiFi polymerase (QuantaBio, Inc., Beverly, MA) in 1× Accustart HiFi buffer with 2 mM $MgCl_2$, and cycled 25 times, as previously described (51). The ssDNA was gel- and column-purified. The two ssDNA products were then mixed in a 1:1 molar ratio and the ssDNA was converted to dsDNA using Phusion polymerase, column purified, and ligated using T4 DNA ligase (NEB) in 1× T4 DNA ligation buffer with 30 ng of amplified DNA incubated at room temperature overnight.

E. coli strains M13cp (33), DH5α F′Iq (Thermo Fisher, Inc., Waltham, MA), and SS320 (Lucigen, UK) were each made competent by washing log-phase grown cells in ice cold 100 mM $CaCl_2$. 20 μL of competent cells were transformed with 2 μL of phagemid DNA ligation mix. Cells were incubated on ice for 30 minutes, heat shocked at 42ºC for 45 seconds, and then put on ice. Pre-

warmed SOB media was added and the cell culture was shaken at 37ºC for 1 hour. 100 µL of cells were plated evenly across a Luria-Agar (LA) media plate made with 100 µg/mL ampicillin and 15 µg/mL chloramphenicol.

Individual colonies were selected and grown in 5 mL of Terrific Broth (TB) supplemented with 1% glycerol for overnight at 37ºC. Bacteria was removed by centrifuging at 4,000 rpm for 10 minutes. Supernatant was removed and placed in a new 1.5 mL spin column and spun at 4,000 rpm for an additional 10 minutes. 1 µL of the clarified supernatant was added to 20 µL of nuclease-free water and heated to 95ºC for 5 minutes, after which 1 µL of the heated solution was added to a Phusion PCR mix containing enzyme, buffer, nucleotides, and forward and reverse primers used to generate the plasmid. Positive colonies were determined by the presence of the PCR amplicon as visualized by agarose gel, and the purified phage were sent for Sanger sequencing. Of the eight colonies chosen, two were shown to have the correct sequence. The bacterial pellet was processed to purify all containing DNA by alkaline lysis and column purification (Qiagen miniprep spin kit, Qiagen, Inc., Germany).

Synthetic phage production. Phage producing colonies, as judged by positive PCR, gel visualization, and sequencing results, were grown overnight in 4 mL 2×YT supplemented with 100 µg/mL ampicillin, 15 µg/mL of chloramphenicol and 5 µg/mL of tetracycline (Sigma-Aldrich, Inc.) in 15 mL culture tubes shaken at 200 RPMs at 37ºC. The following day, the cultures were diluted to an O.D.600 of 0.05 in 2×YT supplemented with 100 µg/mL ampicillin, 15 µg/mL of chloramphenicol and 5 µg/mL of tetracycline and grown between 3 h to 27 h for time course experiments, and 8 h for media, pH, and strain optimization experiments. For pH optimization, the

pH was controlled by addition of 100 mM HEPES-NaOH to the 2×YT media. Strain-specific antibiotics were used as recommended by the manufacture. After the chosen time point, the cultures were spun down at 4,000 RPMs for 15 minutes, after which the supernatant was removed to a fresh tube and spun at 4,000 RPMs for an additional 15 minutes, and filtered using a 0.45 μm cellulose acetate filter. For gel and nanodrop quantification, 400 μL of the clarified media was lysed by addition of Qiagen Buffer P1 supplemented with Proteinase K (20 μg/mL final; Sigma) and RNase A/T1 and incubated at 37ºC for 1 h, followed by addition of Qiagen Buffer P2 and heating to 70ºC for 15 minutes, and letting return to room temperature. Qiagen Buffer N3 was then added and precipitant was centrifuged. One volume of 100% ethanol was added to the supernatant and applied to a Qiagen spin column, and purified. The purified eluate DNA concentration was determined by A280 absorbance from a NanoDrop 2000 (Thermo Fisher) for each time point and condition tested, and ran on a 1% agarose gel in 1×Tris-Acetate-EDTA (TAE) stained with SybrSafe (Thermo Fisher) for visualization of the product. The ssDNA purity was judged by ImageJ (52) intensity analysis and the amount of ssDNA from the time point or condition was adjusted by this purity multiplied by the total amount of DNA found from A280 absorbance.

For milligram-scale production of synthetic miniphage, a Stedium Sartorius fermenter was used for growing 5 L of culture. An overnight culture was grown in 2×YT supplemented with 100 μg/mL Ampicillin and 15 μg/mL of chloramphenicol and 5 μg/mL of tetracycline and diluted to O.D. 600 of 0.05 for inoculating 5 L of media. The growth media for the batch fermentation was also 2×YT supplemented with 100 μg/mL Ampicillin and 15 μg/mL of chloramphenicol and 5 μg/mL of tetracycline. Oxygen and pH were monitored throughout the growth, and the pH was automatically adjusted with phosphoric acid and ammonium hydroxide, with a constant agitation

of 800 RPM. Time points were taken approximately every hour and samples were processed as above for the shaker flask. At 8 h, 900 mL of liquid culture was removed for processing. For milligram-scale purification of ssDNA, 900 mL of liquid culture bacteria was pelleted by centrifuging twice at 4,000 × g for 20 min, followed by 0.45 μm cellulose acetate filtration. Phage from clarified media were precipitated by adding 6% w/v of polyethylene glycol-8000 (PEG-8000) and 3% w/v of NaCl and stirring continuously at 4ºC for 1 h. Precipitated phage were collected by centrifuging at 12,000 × g for 1 h, and the PEG-8000 supernatant was removed completely, and pellet was resuspended in 30 mL of 10 mM Tris-HCl pH 8, 1 mM Ethylenediaminetetraacetic acid (EDTA) buffer (TE buffer). The phage was then processed using an EndoFree Maxiprep (Qiagen, Germany) column-based purification, following the manufacturer's protocol with two adjustments. First, proteinase K (20 μg/mL final) was added to EndoFree Buffer P1 and incubated at 37ºC for 1 h before addition of EndoFree Buffer P2 and incubation at 70ºC for 10 min. The lysed phage was returned to room temperature before proceeding. Second, after removal of endotoxins, 0.2 v/v of 100% ethanol was added to the clarified sample, before applying to the EndoFree Maxiprep column to increase ssDNA binding. All other steps remained the same, and the cssDNA was eluted in 1 mL of endotoxin-free TE buffer. The amount of collected DNA was judged by absorbance at A280, and the purity was judged by running on a 1% agarose gel in 1× TAE stained with ethidium bromide.

Endotoxin amounts were tested using the ToxinSensor chromogenic LAL endotoxin assay kit (GenScript, Piscataway, NJ) following the manufacturer's protocol. The cssDNA phPB84 was diluted to 10 nM in endotoxin-free water, with absorbance read on an Evolution 220 UV/Vis spectrophotometer (Thermo Fisher). Stability from exonuclease I degradation was tested by

incubating cssDNA phPB84 with exonuclease I in 1× exonuclease buffer (NEB) at 37ºC for 30 min. The reaction was quenched by incubating the reaction at 80ºC for 15 min, and was subsequently ran on a 1% agarose gel in 1× TAE stained with ethidium bromide.

DNA origami assembly. DNA purified from phages phPB52 and phPB84 were used to fold a pentagonal bipyramid with edge length 52 base pairs and 84 base pairs, respectively. Staples for each object were generated from the automated scaffold routing and staple design software DAEDALUS (19). Staples were synthesized by IDT, and listed in **Appendix Table 7** and **8**. To fold the nanoparticles, 20 nM of bacterially-produced and purified scaffold was incubated with 20-molar excess of staples in 1×TAE buffer with 12 mM $MgCl_2$. The objects were annealed over 13 hours from 95ºC to 24ºC as previously described (19), and the folded particles were run on 1% agarose gel in 1×TAE buffer with 12 mM $MgCl_2$ with the respective cssDNA scaffolds for reference. The folded nanoparticles were purified using a 100 kDA MWCO spin concentrator (Amicon) for a total of five 5-fold buffer exchanges as purification for TEM.

Transmission electron microscopy. The structured DNA pentagonal bipyramid with 52-base-pair and 84-base-pair edge length assembled using the phage-produced scaffold was visualized by transmission electron microscopy (TEM). 200 μL of folded reaction was purified from excess staples and buffer exchanged into 20 mM Tris-HCl pH 8.0 and 8 mM $MgCl_2$ using a 100kDa MWCO spin concentrator (Amicon, Merck Millipore, Billerica, MA). The concentration was subsequently adjusted to 5 nM. Carbon film with copper grids (CF200H-CU; Electron Microscopy Sciences Inc., Hatfield, PA) were glow discharged and the sample was applied for 60 seconds. The sample was then blotted from the grid using Whatman 42 ashless paper, and the grid was placed

on drop of freshly prepared 1% uranyl-formate with 5mM NaOH for 10 s (53). Remaining stain was wicked away using Whatman 42 paper and dried before imaging. The grid was imaged on a Technai FEI with a Gatan camera.

REFERENCES

1.   Messing, J., Crea, R. and Seeburg, P.H. (1981) A system for shotgun DNA sequencing. *Nucleic Acids Res*, **9**, 309-321.

2.   Zoller, M.J. and Smith, M. (1982) Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. *Nucleic Acids Res*, **10**, 6487-6500.

3.   Chen, F., Pruett-Miller, S.M., Huang, Y., Gjoka, M., Duda, K., Taunton, J., Collingwood, T.N., Frodin, M. and Davis, G.D. (2011) High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat Methods*, **8**, 753-755.

4.   Church, G.M., Gao, Y. and Kosuri, S. (2012) Next-generation digital information storage in DNA. *Science*, **337**, 1628.

5.   Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E.M., Sipos, B. and Birney, E. (2013) Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, **494**, 77-80.

6.   Dietz, H., Douglas, S.M. and Shih, W.M. (2009) Folding DNA into twisted and curved nanoscale shapes. *Science*, **325**, 725-730.

7.   Douglas, S.M., Dietz, H., Liedl, T., Hogberg, B., Graf, F. and Shih, W.M. (2009) Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, **459**, 414-418.

8.   Rothemund, P.W. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature*, **440**, 297-302.

9.   Sharma, J., Chhabra, R., Cheng, A., Brownell, J., Liu, Y. and Yan, H. (2009) Control of self-assembly of DNA tubules through integration of gold nanoparticles. *Science*, **323**, 112-116.

10.   Diagne, C.T., Brun, C., Gasparutto, D., Baillin, X. and Tiron, R. (2016) DNA Origami Mask for Sub-Ten-Nanometer Lithography. *ACS Nano*, **10**, 6458-6463.

11.   Surwade, S.P., Zhao, S. and Liu, H. (2011) Molecular lithography through DNA-mediated etching and masking of SiO2. *J Am Chem Soc*, **133**, 11868-11871.

12.   Dutta, P.K., Levenberg, S., Loskutov, A., Jun, D., Saer, R., Beatty, J.T., Lin, S., Liu, Y., Woodbury, N.W. and Yan, H. (2014) A DNA-Directed Light-Harvesting/Reaction Center System. *Journal of the American Chemical Society*, **136**, 16618-16625.

13.   Dutta, P.K., Varghese, R., Nangreave, J., Lin, S., Yan, H. and Liu, Y. (2011) DNA-Directed Artificial Light-Harvesting Antenna. *Journal of the American Chemical Society*, **133**, 11985-11993.

14.   Hemmig, E.A., Creatore, C., Wünsch, B., Hecker, L., Mair, P., Parker, M.A., Emmott, S., Tinnefeld, P., Keyser, U.F. and Chin, A.W. (2016) Programming Light-Harvesting Efficiency Using DNA Origami. *Nano Letters*, **16**, 2369-2374.

15.   Pan, K., Boulais, E., Yang, L. and Bathe, M. (2014) Structure-based model for light-harvesting properties of nucleic acid nanostructures. *Nucleic Acids Res*, **42**, 2159-2170.

16.   Sun, W., Boulais, E., Hakobyan, Y., Wang, W.L., Guan, A., Bathe, M. and Yin, P. (2014) Casting inorganic structures with DNA molds. *Science*, **346**, 1258361.

17.   Douglas, S.M., Bachelet, I. and Church, G.M. (2012) A logic-gated nanorobot for targeted transport of molecular payloads. *Science*, **335**, 831-834.

18.   Zhao, Y.X., Shaw, A., Zeng, X., Benson, E., Nystrom, A.M. and Hogberg, B. (2012) DNA origami delivery system for cancer therapy with tunable release properties. *ACS Nano*, **6**, 8684-8691.

19. Veneziano, R., Ratanalert, S., Zhang, K., Zhang, F., Yan, H., Chiu, W. and Bathe, M. (2016) Designer nanoscale DNA assemblies programmed from the top down. *Science*, **352**, 1534.

20. Benson, E., Mohammed, A., Gardell, J., Masich, S., Czeizler, E., Orponen, P. and Hogberg, B. (2015) DNA rendering of polyhedral meshes at the nanoscale. *Nature*, **523**, 441-444.

21. Douglas, S.M., Marblestone, A.H., Teerapittayanon, S., Vazquez, A., Church, G.M. and Shih, W.M. (2009) Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res*, **37**, 5001-5006.

22. Jun, H., Shepherd, T.R., Zhang, K., Bricker, W.P., Li, S., Chiu, W. and Bathe, M. (2018) Automated Sequence Design of 3D Polyhedral Wireframe DNA Origami with Honeycomb Edges. *Submitted*.

23. Jun, H., Zhang, F., Shepherd, T., Ratanalert, S., Qi, X., Yan, H. and Bathe, M. (2019) Autonomously designed free-form 2D DNA origami. *Sci Adv*, **5**, eaav0655.

24. Brown, S., Majikes, J., Martinez, A., Giron, T.M., Fennell, H., Samano, E.C. and LaBean, T.H. (2015) An easy-to-prepare mini-scaffold for DNA origami. *Nanoscale*, **7**, 16621-16624.

25. Nafisi, P.M., Aksel, T. and Douglas, S.M. Construction of a novel phagemid to produce custom DNA origami scaffolds. *Synthetic Biology*, **3**.

26. Praetorius, F., Kick, B., Behler, K.L., Honemann, M.N., Weuster-Botz, D. and Dietz, H. (2017) Biotechnological mass production of DNA origami. *Nature*, **552**, 84-87.

27. Kick, B., Praetorius, F., Dietz, H. and Weuster-Botz, D. (2015) Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami. *Nano Lett*, **15**, 4672-4676.

28. Conway, J.W., McLaughlin, C.K., Castor, K.J. and Sleiman, H. (2013) DNA nanostructure serum stability: greater than the sum of its parts. *Chem Commun (Camb)*, **49**, 1172-1174.

29. Vieira, J. and Messing, J. (1987) Production of single-stranded plasmid DNA. *Methods Enzymol*, **153**, 3-11.

30. Pasqualini, R. and Ruoslahti, E. (1996) Organ targeting in vivo using phage display peptide libraries. *Nature*, **380**, 364-366.

31. Winter, G., Griffiths, A.D., Hawkins, R.E. and Hoogenboom, H.R. (1994) Making antibodies by phage display technology. *Annu Rev Immunol*, **12**, 433-455.

32. Ferrara, F., Kim, C.Y., Naranjo, L.A. and Bradbury, A.R. (2015) Large scale production of phage antibody libraries using a bioreactor. *MAbs*, **7**, 26-31.

33. Chasteen, L., Ayriss, J., Pavlik, P. and Bradbury, A.R. (2006) Eliminating helper phage from phage display. *Nucleic Acids Res*, **34**, e145.

34. Reddy, P. and McKenney, K. (1996) Improved method for the production of M13 phage and single-stranded DNA for DNA sequencing. *Biotechniques*, **20**, 854-856, 858, 860.

35. Bryksin, A.V. and Matsumura, I. (2010) Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques*, **48**, 463-465.

36. van den Ent, F. and Lowe, J. (2006) RF cloning: a restriction-free method for inserting target genes into plasmids. *J Biochem Biophys Methods*, **67**, 67-74.

37. Veneziano, R., Shepherd, T.R., Ratanalert, S., Bellou, L., Tao, C. and Bathe, M. (2018) In vitro synthesis of gene-length single-stranded DNA. *Sci Rep*, **8**, 6548.

38. Dotto, G.P., Horiuchi, K. and Zinder, N.D. (1984) The functional origin of bacteriophage f1 DNA replication. Its signals and domains. *J Mol Biol*, **172**, 507-521.

39. Hahn, J., Wickham, S.F., Shih, W.M. and Perrault, S.D. (2014) Addressing the instability of DNA nanostructures in tissue culture. *ACS Nano*, **8**, 8765-8775.

40. Miller, A. (1953). The Crucible. Act II, Scene 2.
41. Lovett, S.T. (2011) The DNA Exonucleases of Escherichia coli. *EcoSal Plus*, **4**.
42. Smeal, S.W., Schmitt, M.A., Pereira, R.R., Prasad, A. and Fisk, J.D. (2017) Simulation of the M13 life cycle II: Investigation of the control mechanisms of M13 infection and establishment of the carrier state. *Virology*, **500**, 275-284.
43. Masai, H. and Arai, K. (1997) Frpo: a novel single-stranded DNA promoter for transcription and for primer RNA synthesis of DNA replication. *Cell*, **89**, 897-907.
44. Kozyra, J., Ceccarelli, A., Torelli, E., Lopiccolo, A., Gu, J.Y., Fellermann, H., Stimming, U. and Krasnogor, N. (2017) Designing Uniquely Addressable Bio-orthogonal Synthetic Scaffolds for DNA and RNA Origami. *ACS Synth Biol*, **6**, 1140-1149.
45. Rovner, A.J., Haimovich, A.D., Katz, S.R., Li, Z., Grome, M.W., Gassaway, B.M., Amiram, M., Patel, J.R., Gallagher, R.R., Rinehart, J. *et al.* (2015) Recoded organisms engineered to depend on synthetic amino acids. *Nature*, **518**, 89-93.
46. Seidl, C.I. and Ryan, K. (2011) Circular single-stranded synthetic DNA delivery vectors for microRNA. *PLoS One*, **6**, e16925.
47. Davis, L. and Maizels, N. (2016) Two Distinct Pathways Support Gene Correction by Single-Stranded Donors at DNA Nicks. *Cell Rep*, **17**, 1872-1881.
48. Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L. and Corn, J.E. (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol*, **34**, 339-344.
49. Clark, J.R. and March, J.B. (2006) Bacteriophages and biotechnology: vaccines, gene therapy and antibacterials. *Trends Biotechnol*, **24**, 212-218.
50. SantaLucia, J., Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A*, **95**, 1460-1465.
51. Veneziano, R., Shepherd, T.R., Bellou, L., Ratanalert, S., Tao, C. and Bathe, M. (2017) Enzymatic synthesis of gene-length single-stranded DNA. *bioRxiv*.
52. Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, **9**, 671-675.
53. Castro, C.E., Kilchherr, F., Kim, D.-N., Shiao, E.L., Wauer, T., Wortmann, P., Bathe, M. and Dietz, H. (2011) A primer to scaffolded DNA origami. *Nature Methods*, **8**, 221-229.

Chapter 3

Programming innate immune stimulation with 3D wireframe DNA origami

ABSTRACT

Wireframe DNA origami have programmable structural and sequence features that enable them to be developed for a range of prophylactic and therapeutic applications. However, the immunological properties of these nanomaterials, including the roles of their geometric shape and CpG content, remain largely unknown. Here, we investigate the immunostimulatory properties of 3D wireframe DNA origami on the TLR9 and cGAS-STING pathways using reporter cells and primary immune cells. Our results suggest that wireframe DNA origami induce a cGAS-STING dependent immune response, but are minimally TLR9 activating, despite the presence of numerous CpG dinucleotides within the scaffold and staples. However, displaying CpG-containing oligos from wireframe DNA origami results in robust TLR9 pathway activation and enhancement of the downstream immune response, which is substantiated by an increase in Type I and Type III IFN production in peripheral blood mononuclear cells. We then investigate the impacts of CpG copy number, inter-CpG spacing, and spatial patterning on the strength of TLR9 activation, and find that signaling intensity is correlated with CpG valency and density of clustering. Our results suggest that key properties of immunostimulatory wireframe DNA origami can be programmed to modulate immune pathway activation controllably, offering an improved understanding of how nanoscale organization of immunostimulatory oligos influences their interactions with the innate immune system, and highlighting design parameters that may prove useful for downstream biological applications.

INTRODUCTION

The field of nucleic acid nanotechnology has over the past several decades developed technologies enabling the fabrication of programmable DNA-based assemblies of prescribed size, geometry, rigidity, and chemical composition[1–9]. These nanomaterials now comprise a toolbox for the design and fabrication of nanodevices capable of interacting with diverse cellular environments[10–12]. One approach to designing nucleic acid nanostructures on the 10–100nm scale is the concept of DNA origami, wherein programmed regions of a long, single-stranded DNA scaffold that are far apart in sequence space are brought into spatial proximity through the hybridization of small, single-stranded DNA staples. Since the introduction of this strategy by Paul Rothemund[13], who used this approach to pioneer the fabrication of 2D rectilinear, bricklike assemblies of parallel DNA duplexes, it has been generalized to enable the design of 2D and 3D wireframe duplexes, which can adopt arbitrary spatial orientations to fabricate polyhedral geometries[1,3–5,9]. Furthermore, recent development of fully automated sequence design algorithms facilitates the design and fabrication of wireframe structures composed of dual-duplex or six-helix-bundle edges with variable mechanical properties, allowing for the production of virus-like geometries and other more complex nucleic acid nanoparticle (NANP) shapes for diverse biophysical and cellular applications[1–4,8,9]. Scalable scaffold production strategies using M13 bacteriophage engineering and bioproduction have also now enabled control over scaffold sequence composition and length, expanding the accessible design space for such wireframe NANPs[14–16].

The ability to chemically functionalize NANPs also permits the attachment of therapeutic nucleic acid (TNA) cargo such as siRNA and miRNA, as well as small molecules, aptamers, peptides, and proteins with nanometer-level spatial control for active targeting or immune cell stimulation[10,12,17]

and enhances the potential of NANPs to interface with biological systems *in vitro* and *in vivo* [18,19].

However, the immunostimulatory properties of DNA-based NANPs, which are fundamental to

their *in vivo* application, have only been examined to a limited extent[20,21]. Specifically, Hong et al.

identified the primary properties of structured DNA assemblies involved in immunological

recognition[22], and Schüller et al. demonstrated that nonfunctionalized DNA origami tubes did not

trigger cytokine production or induce dendritic cell (DC) activation[23]. However, a systematic

investigation of the immunostimulatory properties of wireframe DNA origami has yet to be

performed, and while hypotheses can be formulated about the immune response towards such

nanostructures based on their physical properties, the effect of nanostructuring and 3D nanoscale

organization of intrinsically immunogenic oligonucleotides on immune recognition remain open

questions.

The innate immune system contains several pattern recognition receptors (PRRs) that are

responsible for recognizing evolutionarily conserved pathogen-associated or damage-associated

molecular patterns (PAMPs or DAMPs). Once activated, these PRRs invoke innate immune

recognition while simultaneously activating the adaptive immune response[24,25]. Binding of PRRs

to their corresponding ligand triggers activation of downstream signaling pathways, ultimately

resulting in the production of Type I interferons (IFNs) and other proinflammatory cytokines that

are essential for initiation of a host of immune functions[26,27]. Cyclic GMP-AMP synthase (cGAS)

is one such PRR that responds to cytosolic double-stranded DNA (dsDNA) from both endogenous

sources such as certain forms of cell death as well as exogenous DNA from pathogens[28–30] in a

sequence-independent but length-dependent manner[30–33]. Because wireframe DNA NANP edges

are composed of duplex DNA, it is likely that if these NANPs are internalized into the cytosol of

cells that express the cGAS-STING pathway, they may trigger cGAS binding and downstream immune activation.

Toll-like Receptors (TLRs) are a family of transmembrane PRRs that sense and respond to a wide variety of ligands, including nucleic acids. TLRs are composed of an N-terminal PAMP-binding ectodomain, a transmembrane domain and a C-terminal Toll IL-1 receptor domain (TIR), and are subdivided into two groups according to whether they localize at the cell surface membrane or within the endosomal membrane[34,35]. The endosomal TLR subfamily, which consists of TLR3, TLR7/8, and TLR9, is involved in nucleic acid recognition, and while TLR3 and TLR7/8 recognize dsRNA and ssRNA, respectively, TLR9 responds to single-stranded DNA (ssDNA) containing unmethylated cytosine-phosphate-guanine dinucleotides (CpGs)[36–38]. Each of the monomers within the TLR9 homodimer undergoes a conformational change upon binding to a CpG-containing ssDNA oligo, enabling the formation of a 2:2 TLR9 monomer:CpG oligo complex in which each monomer interacts with both CpGs[39,40]. Activated TLR9 homodimer formation brings the two C-terminal TIR domains into close proximity, allowing for assembly of TIR-containing adaptor proteins (TIRAP) and MyD88 adaptor proteins into the myddosome, successful formation of which is critical for initiation of the downstream signaling cascade and production of Type I IFNs and pro-inflammatory cytokines[39,41–43]. Thus, the presence of CpG dinucleotides within the scaffold and staples of DNA-based NANPs may elicit TLR9 signaling and enable innate immune activation. Indeed, the therapeutic potential of the immune cascade triggered by TLR9 activation has prompted investigations into the optimization of CpG motifs, which are CpG dinucleotides placed within a particular sequence context that enhances their TLR9 activation efficacy, as well as into the application of TLR9 agonists as anti-cancer drugs, vaccine adjuvants, and combination

therapies, a few of which, such as MGN-1703 and SD-101, are in ongoing clinical development[44–48].

DNA nanotechnology has also been used to stimulate innate immune signaling through TLR activation. For example, the Ding group[50] developed a cancer vaccine nanodevice in which CpG loops were hybridized within an antigen-displaying DNA origami tube to enable TLR9 activation following pH-triggered conformational change of DNA locks. This study followed earlier work by the Liedl group[23] that placed CpG oligos onto immune inert DNA origami tubes to induce CpG-dependent cytokine production and immune cell activation, as well as work by Li et al. wherein they demonstrated that small self-assembled DNA tetrahedra functionalized with multivalent CpG motifs could induce enhanced secretion of inflammatory cytokines via activation of the TLR9 pathway[49]. Recently, Comberlato et al. fixed CpG dimer pairs at distances of 7nm or 38nm on a 2D nano-disk and showed that the 7nm dimer pair, which matched the distance between binding sites in the TLR9 dimer, induced increased immune activation, demonstrating the important of interligand distance[50]. However, while these studies evaluated changes in overall cytokine or surface marker expression levels induced by NANP delivery, there has not yet been an investigation of the relative contributions of individual pathways in NANP-induced innate immune responses, nor has there been a systematic evaluation of the impacts of specific NANP properties on innate immune activation. Previous studies using alternative nanotechnologies to DNA-based materials have also demonstrated that incorporating CpGs into nanostructures can influence the resulting immune response[51,52]. However, those studies were limited by the inability to precisely control ligand presentation and stoichiometry, so the relative impacts of nanoscale organization of

TLR ligands such as spatial positioning and valency on the intensity of pathway activation remain poorly understood.

Here, to address the aforementioned gaps in our knowledge, we characterized the immunological properties of unmodified 3D wireframe origami NANPs by first examining the effects of NANP geometry and structuring on isolated immune pathways using reporter cell lines expressing TLR9 and cGAS-STING, and subsequently assessing cytokine induction in human PBMCs. We then probed the relative impacts of immunostimulatory motif valency and spatial distancing on immune pathway activation by presenting discrete copy numbers of CpG oligos at precise locations on DNA NANPs and evaluating the effects of these design parameters on the magnitude of TLR9 activation and interferon production. Finally, we interrogated the impact of the wireframe NANP itself on the ability of CpG oligos to activate TLR9 as well as how nanoscale organization of CpGs may be used to tune the intensity of immune pathway activation. Insights from our studies reveal design principles that may help enable the rational design of nucleic acid nanostructures with programmable immunostimulatory profiles for a variety of prophylactic and therapeutic applications.

RESULTS AND DISCUSSION

Design and characterization of wireframe DNA NANPs

Wireframe DNA origami NANPs were designed using DAEDALUS and folded from a bacterially produced synthetic circular scaffold and an excess of short single-stranded staples via thermal annealing[16]. In order to systematically evaluate the immunostimulatory properties of these NANPs, which are composed of dual-duplex DNA edges and additionally contain 262 CpG dinucleotides scattered throughout the scaffold and staples (Table S1), we focused on multiple variations of one primary NANP, a pentagonal bipyramid with 84 base pairs per edge (PB84). As a comparative construct to investigate the effects of geometry on TLR9 and cGAS-STING activation, we additionally fabricated an icosahedron with 42 base pairs per edge (ICO42), which was designed using the same scaffold as PB84 to eliminate potential effects of sequence variation on innate immune signaling (Figure 3.1A). In addition to the unmodified NANPs that we used to probe the intrinsic immunogenicity of these wireframe structures, we also fabricated PB84 and ICO42 constructs in which select staples were extended on the 3' end to expose a 20-nt single-stranded overhang comprising the CpG sequence ODN 2006/7909, a class B CpG ODN containing three repeats of an optimized hexamer known to strongly activate human TLR9. We used this approach to display CpG motifs at defined spatial locations on our NANP constructs in order to interrogate the effects of CpG copy number, spacing, and organization on TLR9 activation (Figure 3.1B). Validation of proper NANP self-assembly was first performed using agarose gel electrophoresis, in which successful folding was indicated by an upward shift of the folded NANP band relative to the scaffold band, with this shift enhanced by the addition of ssDNA overhangs. NANP monodispersity was evaluated using dynamic light scattering (DLS), and the hydrodynamic diameter of each sample was measured to validate batch to batch consistency in NANP folding

(Figure 3.1C). Fluorimetry using Cy5-labeled oligos hybridized to CpG overhangs confirmed the capability of this approach to present the expected copy numbers of CpGs on each NANP variant. A gel clot endotoxin test was used to verify that the endotoxin level of each sample was below a threshold of 1.5 EU/mL prior to its application in cell assays.
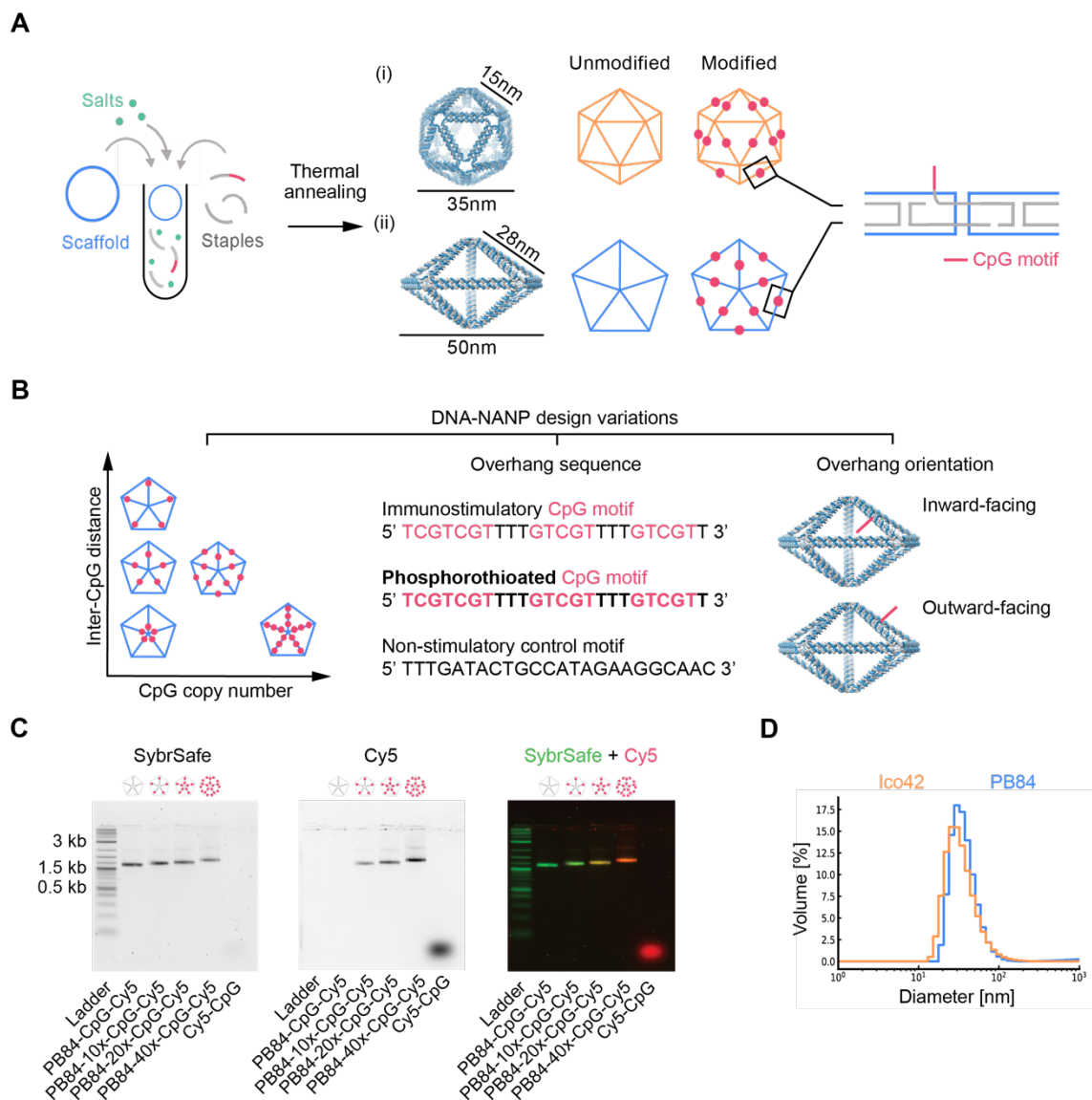
Figure 3.1. NANP design and characterization. (A) DNA NANPs of varying sizes and geometries were fabricated with staples containing ssDNA CpG motif overhangs to control the copy number of CpGs displayed from NANP edges: (i) Icosahedron displaying up to 30 CpG

overhangs; (ii) pentagonal bipyramid displaying up to 40 CpG overhangs. Red circles denote locations where CpG overhangs are displayed on the NANP. (B) Design variations were used to explore the effects of CpG overhang copy number, inter-CpG distance, orientation, and sequence. (C) A fluorescent agarose gel shift assay was used to analyze the quality of NANP folding after spin column purification. Pentagonal bipyramids displaying 0, 10, 20, or 40 ssDNA overhangs hybridized to complementary Cy5-modified oligos show gel shifts corresponding to their increasing molecular weight on the SybrSafe channel, while the gel image taken in the Cy5 channel exhibits an increase in band intensity due to the successive increase in Cy5-CpG copy number on each NANP. (D) Dynamic light scattering (DLS) was used to evaluate NANP hydrodynamic diameter and polydispersity. Representative DLS measurements of an unmodified pentagonal bipyramid and an icosahedron are shown.

Innate immune stimulation by unmodified DNA NANPs *in vitro*.

We first sought to characterize the baseline immunostimulatory properties of NANPs by investigating the effect of unmodified PB84 and ICO42 on TLR9 and cGAS-STING pathway activation *in vitro*. HEK-Blue reporter cells expressing stably transfected human TLR9 were incubated with 10nM PB84 and ICO42, and to assess the impact of NANP structuring, we also tested equivalent concentrations of the scaffold (phPB84) as well as the PB84 and ICO42 staple sets as non-structured controls. Consistent with results from previous literature[22], NANPs that were not complexed with lipofectamine prior to incubation with reporter cells failed to elicit an innate immune response. Thus, for all experiments described herein, all constructs were co-complexed with lipofectamine to enable transfection, and a lipofectamine-only control was also included in each assay (Figure 3.4).
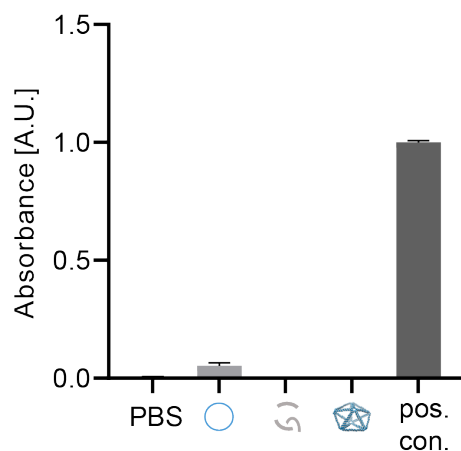
Figure 3.2. NANP formulations do not significantly activate the TLR9 pathway in the absence of lipofectamine complexation. 10nM equivalent of PB84 scaffold, staples, and folded PB84 were not complexed with lipofectamine prior to incubation with HEK-Blue TLR9 cells for 24 hours. No activation of TLR9 was observed for the staple and PB84 samples, and very minimal TLR9 activation was induced in response to the scaffold sample. Data show the average absorbance of samples in triplicate with standard error, where n = 3 biologically independent assays.

After incubation of TLR9 cells with all NANP samples for 24 hours, we evaluated TLR9 activation and found that structured constructs resulted in minimal activation of the TLR9 pathway, whereas the unstructured scaffold induced slightly stronger TLR9 activation (Figure 3.3A). Importantly, the phPB84 scaffold and the PB84/ICO42 staple sets from which the NANPs were folded contained several hundred CpG dinucleotides, although only three sets of these CpG dinucleotides could be considered strongly stimulatory CpG motifs according to previously identified sequence requirements[47,48]. Notwithstanding, this result still suggests that such CpGs, which are typically inaccessible in fully folded NANPs and might not be sufficiently exposed to induce TLR9 engagement and downstream pathway activation even after NANP degradation by DNases, may induce low levels of TLR9 activation when presented in unstructured formulations. The effects of

NANP sequence composition on immune activation and the impact of structuring on the shielding of immunostimulatory motifs may therefore be an important question for future investigations.

A

TLR9

B
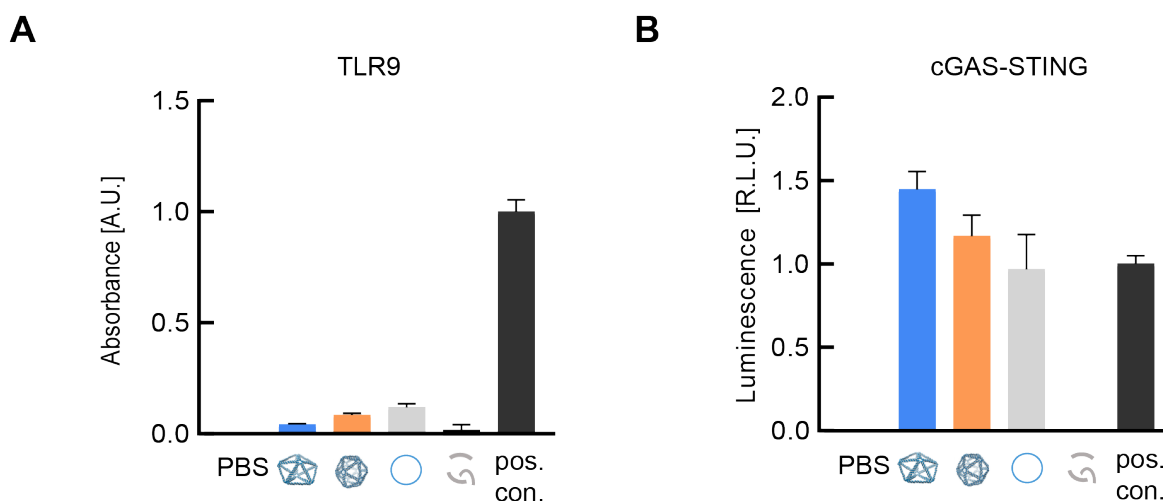
cGAS-STING

Figure 3.3. Cellular immunostimulation by unfunctionalized NANPs *in vitro*. (A) Activation of the TLR9 pathway in HEK-Blue hTLR9 cells or (B) the cGAS-STING pathway in THP1-Dual cells by PB84, Ico42, and scaffold versus staples negative controls. Readouts for the TLR9 and cGAS-STING assays were normalized to the positive controls (pos. con.) ODN2006 and TNF-α, respectively. Data were analyzed separately for each donor and all IFN expression levels were normalized to the corresponding positive control, ODN2216. Data in all assays were calculated from samples in triplicate, where n = 3 biologically independent assays.

Unlike TLR9, which requires specific sequence contexts for binding and activation, cGAS responds to cytosolic dsDNA in a primarily sequence-independent manner. As a result, while neither of the unmodified PB84 nor ICO42 constructs induced strong TLR9 signaling, we hypothesized that they might instead elicit a substantial cGAS-STING response. To test this, we transfected PB84, ICO42, and their respective unstructured controls into THP1-Dual reporter cells. After a 24-hour incubation, we observed that cells treated with structured NANPs strongly activated the cGAS-STING pathway, while staples alone did not induce cGAS activation (Figure

3.3B). Surprisingly, we observed moderate levels of immune activation in response to transfected phPB84 scaffold. As some sections of the scaffold have potential for self-dimerization, this might have occurred during incubation, resulting in stretches of dsDNA becoming available for cGAS binding. To verify that the observed responses to these NANPs were indeed mediated by the TLR9 and cGAS-STING pathways, we performed the same set of experiments in the Null 1 and THP1-KO-cGAS cell lines, which do not express TLR9 and have a stable cGAS knockout, respectively (Figure 3.4). All formulations failed to elicit a response, confirming the role of these innate immune pathways in the NANP-induced responses.

A

Null 1 (no TLR9)

B

KO-cGAS
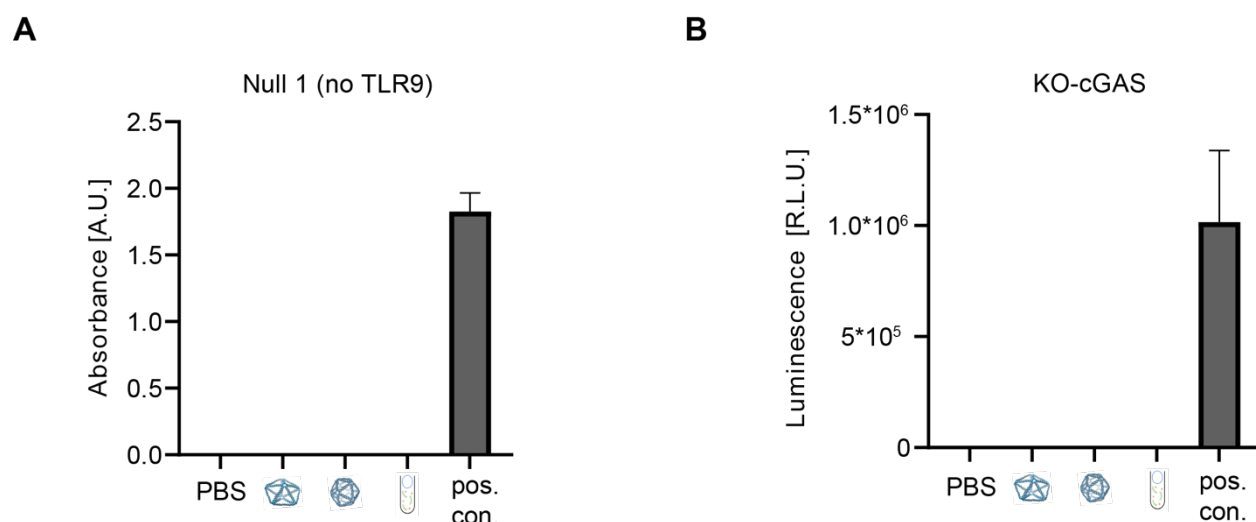
Figure 3.4. Responses of control reporter cells to transfected NANPs. Pathway activation is not observed in response to transfection of PB84, ICO42 or the unfolded control into (A) HEK293T cells without transfected TLR9 and (B) cGAS-knockout THP1-Dual cells. Data show the average absorbance of samples in triplicate with standard error, where n = 3 biologically independent assays.

Interferon response towards unmodified DNA NANPs in primary cells

Having evaluated the responses of isolated innate immune pathways to DNA-based NANPs, we next sought to investigate the effects of these nanostructures on PBMCs, which contain many critical immune cell subtypes such as lymphocytes, monocytes, and dendritic cells and therefore represent a more complete model system of the innate immune response. We transfected PB84, ICO42, scaffold, and staples into PBMCs and recovered cell supernatant after a 20-hour incubation[53]. The immunostimulatory activity of each construct was quantified using a multiplex enzyme-linked immunosorbent assay (ELISA) to measure secretion of IFNα, IFNβ, IFNω, and IFNλ, interferons which are known to be expressed in response to intracellular nucleic acids[24,25]. We observed increases in production of all interferons in response to both PB84 and ICO42, as well as weaker but consistently elevated levels of IFN expression in response to the pHPB84 scaffold (Figure 3.5). Interestingly, we found that the staples alone induced much higher levels of IFN production compared to the other NANP formulations, even though all of the constructs have the same sequence composition. This may be due to the fact that the staples are much smaller than the folded constructs and the scaffold, potentially allowing for higher transfection efficiency and therefore higher concentrations of CpG-containing oligos available for TLR9 binding. This result highlights the ability of nanostructuring to modulate the properties of the individual components: once these highly immunostimulatory staples are folded into a wireframe NANP, whether because of a decrease in transfection efficiency or reduced accessibility of CpG dinucleotides within the scaffold and staples to TLR9, the folded NANP is much less immunostimulatory than the sum of its parts.
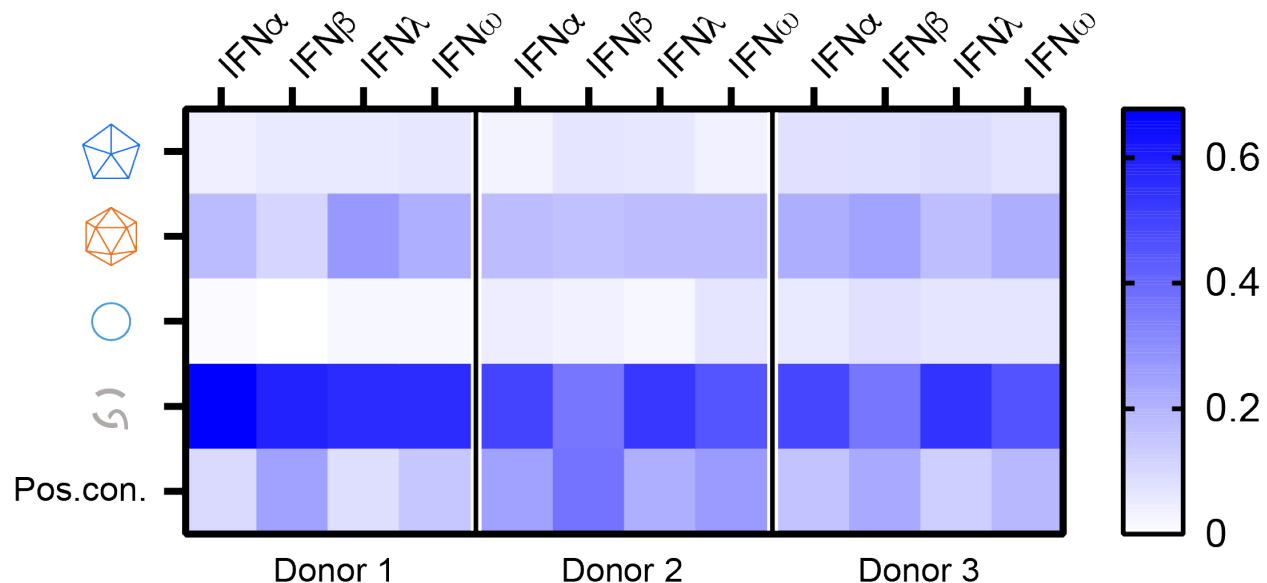
Figure 3.5. Interferon expression in response to unmodified DNA NANPs in PBMCs. NANP-induced expression levels of IFNα, IFNβ, IFNλ, and IFNω in PBMCs. Data were analyzed separately for each donor and all IFN expression levels were normalized to the corresponding positive control, ODN2216. Data in all assays were calculated from samples in triplicate, where n = 3 biologically independent assays.

Taken together with the results from the reporter cell lines, these data suggest that while the immune response towards the unstructured scaffold and staples can be attributed towards both TLR9 and cGAS-STING, the immune response towards the structured NANPs is largely independent of the TLR9 pathway and may primarily be mediated by cGAS-STING. This suggests that the immunostimulatory profile of the individual NANP components is affected by sequence composition and self-dimerization potential, whereas properties which affect cGAS activation such as NANP duplex edge length might have the greatest influence on the immune response towards intact, unmodified DNA NANPs.
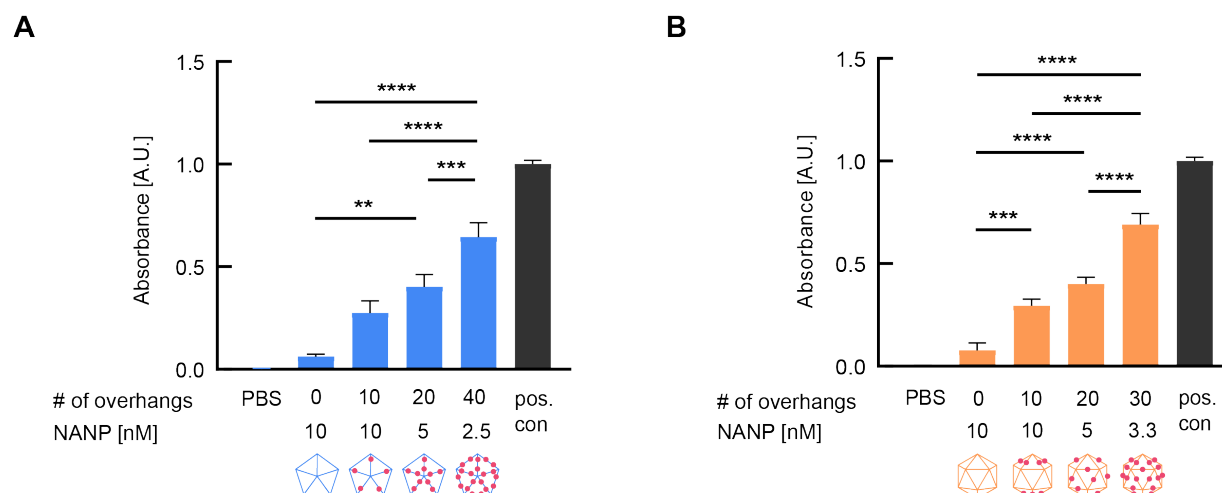
Figure 3.6. NANPs displaying CpG overhangs (CpG-OHs) activate TLR9 at controllable magnitudes. The strength of TLR9 activation corresponds to the copy numbers of CpG-OHs and/or densities of CpG-OH organization on (A) PB84 and (B) Ico42 NANPs. The overall concentration of CpG-OHs is constant across samples. For all assays, absorbances were normalized with the ODN2006 positive control (pos. con.). Data show the average absorbance of samples in triplicate with standard error, where n = 3 biologically independent assays. P values are from a one-sided analysis of variance (ANOVA) with correction for multiple comparisons (*: $P \leq 0.05$; **: $P \leq 0.01$, ***: $P \leq 0.001$, ****: $P \leq 0.0001$). All unlabeled pair-wise comparisons are not significant.

Designing DNA NANPs for controllable immunostimulation

We next investigated the ability of 3D wireframe DNA NANPs to trigger programmable activation of the innate immune response through rationally designed display of immunostimulatory motifs. Since we did not observe significant stimulation of the TLR9 pathway by unmodified structured PB84 in previous reporter assays, we reasoned that we could treat this structure as immunologically inert in the context of TLR9 reporter cells. This NANP could then be controllably functionalized with immunostimulatory motifs to systematically investigate the relative impacts of various parameters of nanoscale display without any confounding levels of baseline TLR9

activation. To this end, we fabricated PB84 variants displaying 0, 10, 20, or 40 copies of CpG-OH distributed evenly across the exterior of the NANP (Figure 3.1B). Nanostructures displaying corresponding copy numbers of CpG-free overhangs (CpG-f-OH) of identical length and GC content served as comparative controls.

We transfected each of the above formulations into HEK-Blue TLR9 cells, and following a 24-hour incubation period, we verified that PB84 displaying 0 copies of CpG-OH did not activate TLR9. In contrast, each of the constructs displaying 10, 20, or 40 copies of CpG-OH induced TLR9 activation, and furthermore, the strength of TLR9 activation was directly correlated with the CpG-OH copy number on the nanostructure (Figure 3.6A). Importantly, the total concentration of CpG motifs was held constant across these samples, indicating that the magnitude of TLR9 activation is dependent not only on the CpG concentration, but also on the valency of CpG motifs per NANP. We also found that all constructs displaying CpG-f-OH, regardless of copy number per NANP, did not induce significantly higher levels of TLR9 activation compared with the unmodified nanostructure, which shows that changes in non-CpG-containing total DNA content have no effect on TLR9 signaling (Figure 3.7).
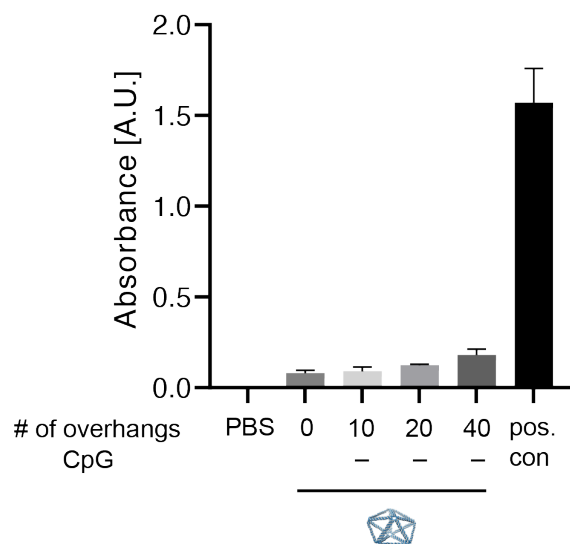
Figure 3.7. The presence of CpG motifs within the ssDNA overhangs is critical for TLR9 activation. Regardless of the copy number of CpG-free overhangs displayed on PB84, there is no significant difference between the level of TLR9 activation induced by those constructs compared to nonmodified PB84. Data show the average absorbance of samples in triplicate with standard error, where n = 3 biologically independent assays.

To test whether this method of controllable immunostimulation was applicable to more than a single type of NANP, we also fabricated ICO42 displaying 0, 10, 20, or 30 copies of CpG-OH. As with PB84, we observed that the magnitude of TLR9 activation was dependent on the copy number of CpG-OH presented by the NANP (Figure 3.6B). However, the magnitude at which the TLR9 pathway was activated by ICO42 and PB84 constructs displaying identical copy numbers of CpG-OH was not the same, suggesting that the spatial organization of CpG overhangs on the nanostructures may influence TLR9 activation levels as well.
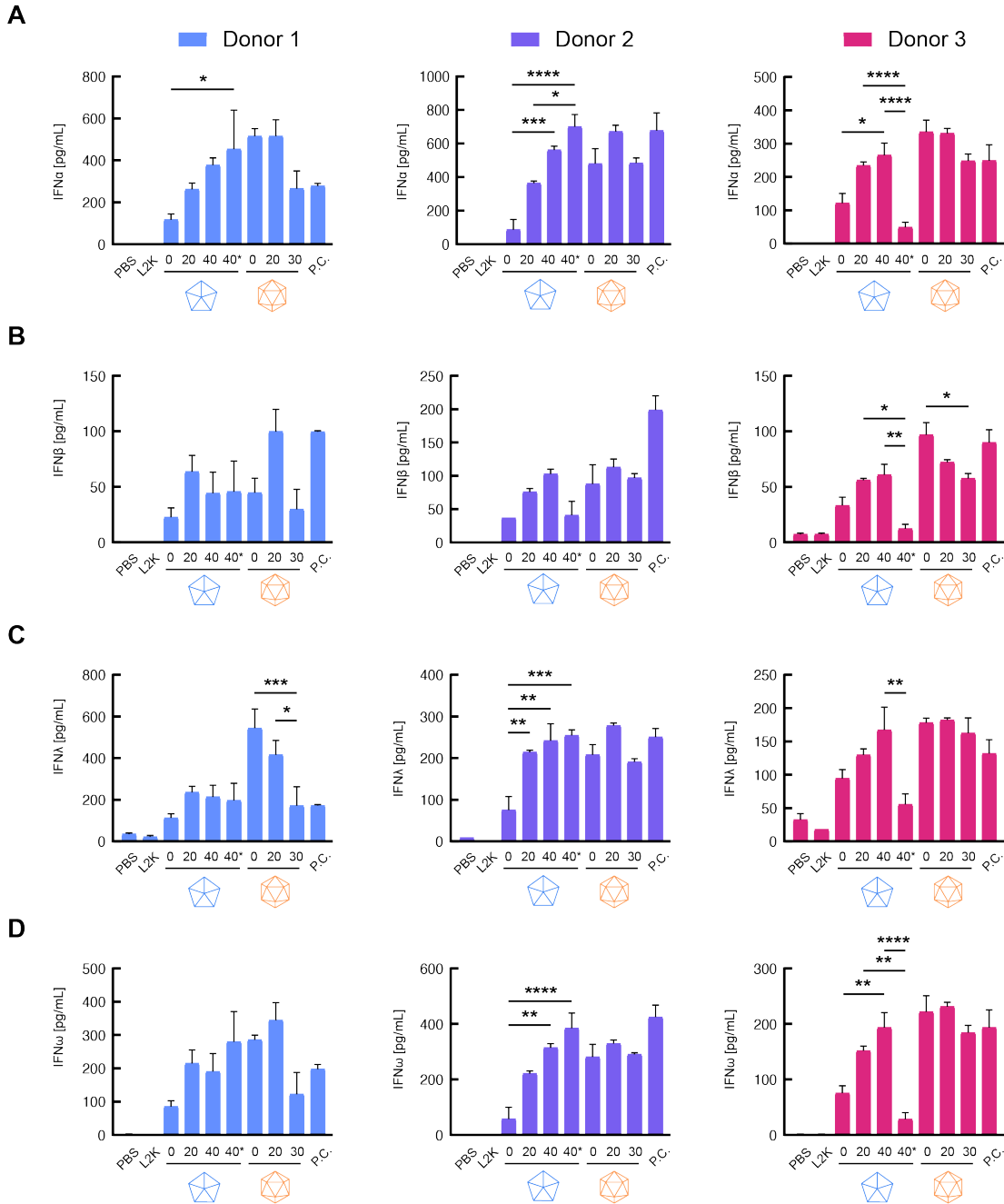
Figure 3.8. Innate immune response of PBMCs in response to engineered immunostimulatory NANPs. PBMCs were transfected with 10nM of NANPs and incubated for 24 hours prior to multiplex ELISA analysis of interferon expression. Higher expression of (A) IFNα, (B) IFNλ, (C) IFNβ, and (D) IFNω was observed in response to PB84 displaying 20 or 40 CpG-OHs compared to the unmodified PB84, while the addition of 20 or 30 CpG-OHs to ICO42 had minimal effect in some cases and reduced interferon expression in others. ODN2216 was used

as the assay positive control (pos. con.). Each bar represents averaged triplicate data from a single donor, where n = 3 donors.

To investigate whether the ability of these NANPs to controllably modulate immune pathway activation was reproducible outside of reporter cell lines, we evaluated the capability of these NANPs to induce interferon production in primary immune cells. PB84 with 0, 20, and 40 copies of CpG-OH and ICO42 with 0, 20, and 30 copies of CpG-OH were transfected into PBMCs, where the total concentration of CpG-OHs in each sample was held constant at 10nM. After incubation for 20 hours, subsequent multiplex ELISA analysis of cell supernatant showed that for all analytes tested, CpG-OH-displaying PB84 constructs induced significantly higher interferon expression compared to their unmodified counterparts (Figure 3.8). Additionally, interferon production was generally higher in response to PB84 displaying 40 copies of CpG-OH compared to PB84 with 20 copies of CpG-OH. Interestingly, unlike what was seen in the reporter cell assays, the increase in IFN response towards ICO42 displaying CpG-OHs appeared to saturate at 20 CpG-OHs. In most cases, ICO42 displaying 30 copies of CpG-OH induced similar levels of IFNs to the unmodified ICO42, and in a few cases, IFN expression in response to ICO42 with 30 CpG-OHs was actually lower than the unmodified NANP. Because the baseline immunostimulation induced by ICO42, likely mediated by cGAS-STING, was already quite high, it's possible that the addition of 20 CpG overhangs elicited an increase in TLR9 activation while minimally affecting cGAS binding, whereas the addition of 30 CpG overhangs significantly hindered the ability of cGAS to bind to ICO42, so much so that the increase in TLR9 activation was not sufficient to compensate for the loss of cGAS-STING activation. These results imply that CpG valency and spatial organization can indeed be tuned to controllably modulate the strength of the NANP-mediated innate immune response, but the specific levels of each parameter may need to be optimized for different NANP

93

geometries. However, as with the reporter cell assays, NANPs that were not complexed with lipofectamine did not elicit cytokine production, suggesting that these nanostructures may not be natively taken up by any immune cell populations within PBMCs at sufficient levels to activate immune signaling, and therefore may be largely immunologically inert in the absence of functionalizations enabling cellular uptake and internalization, minimizing the potential for off-target immunostimulation.

Investigating the parameters controlling DNA NANP-mediated immune activation

Now that we have demonstrated that these NANPs can be modified with CpGs to tune TLR9 activation and modulate the resulting innate immune response, we sought to elucidate the roles of the NANP design parameters contributing to this controllable immunostimulation. Because we saw that NANP structuring significantly affected the intensity of TLR9 signaling, we started by further investigating the role of NANPs in TLR9 activation. As precise control over directionality and spatial positioning is one unique advantage of wireframe DNA origami, we began by testing whether the orientation of CpG-OH presentation on NANPs affected TLR9 activation. Interestingly, we found that there were no significant differences in TLR9 signaling between PB84 constructs displaying CpG-OHs on the exterior versus interior of the NANP (Figure 3.9B). This may be the result of the flexibility of the 20-nt ssDNA overhang and the wireframe structure of NANPs; because only the duplex edges are solid, the 3' terminus of the inward-facing overhangs may in fact end up on the exterior of the NANP, and vice versa.
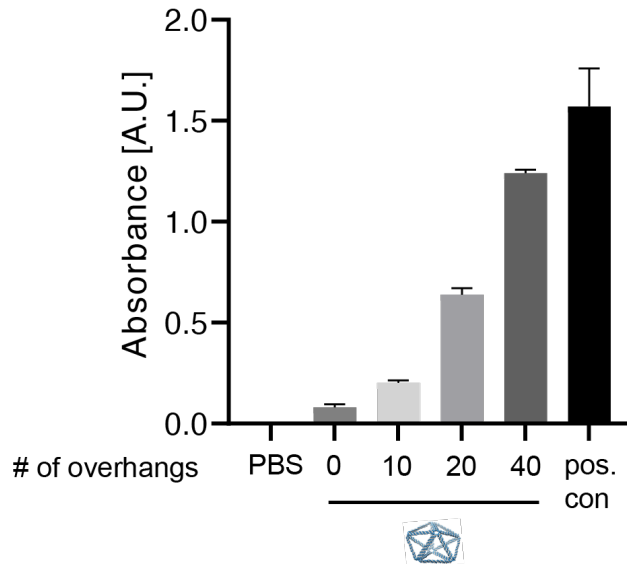
Figure 3.9. The ability of CpG overhangs to activate TLR9 is not affected by orientation.
TLR9 activation levels increase monotonically with CpG-OH copy number regardless of whether
the overhangs are inward-facing or outward-facing (Figure 3.6).

To further understand the contribution of the NANP itself to TLR9 activation, we treated HEK-
Blue TLR9 cells with free CpG-OH and found that there was no TLR9 signaling in response to
free CpG-OH at any of the tested concentrations, even though the same concentrations of CpG-
OH displayed on NANPs induced strong TLR9 activation (Figure 3.10). To test whether this lack
of activation may be due to degradation of free CpG-OHs by intracellular DNases, we incubated
HEK-Blue TLR9 cells with phosphorothioate-stabilized free CpG-OHs (pCpG-OHs). Unlike the
non-stabilized CpG-OHs, the magnitude of TLR9 activation was correlated with the concentration
of pCpG-OHs delivered (Figure 3.10).

Absorbance [A.U.]									

CpG [nM]	PBS	100	200	400	100	200	400	20	40	80	pos. con.
phosphorothioated		—	—	—	+	+	+	+	+	+	
(NANP)		+	+	+	—	—	—	+	+	+	
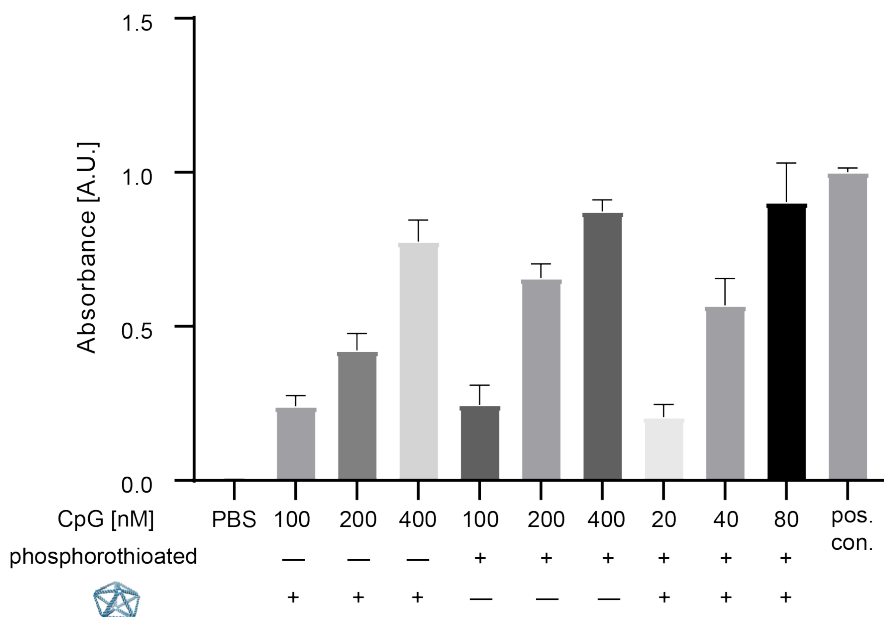
Figure 3.10. Effect of attachment to NANPs and phosphorothioate stabilization on the ability of CpG overhangs to activate TLR9. Phosphodiester CpG overhangs displayed on PB84 induce similar levels of TLR9 activation to corresponding concentrations of free phosphorothioated CpG overhangs. Additionally, overhangs that are both phosphorothioated and attached to PB84 induce similar levels of TLR9 activation at 10-fold lower concentrations compared to their phosphodiester or non-PB84-attached counterparts. Data show the average absorbance of samples in triplicate with standard error, where n = 3 biologically independent assays.

As phosphorothioate stabilization of CpG-OHs was able to rescue their ability to activate TLR9, this suggests that CpG oligo degradation significantly impacts TLR9 activation, corroborating our earlier hypothesis. Interestingly, we found that when CpG-OHs were displayed on NANPs, they were able to induce similar levels of TLR9 activation as their free pCpG-OH counterparts. These results demonstrate that in terms of the ability to activate the TLR9 pathway, attachment of CpG-OHs to DNA NANPs produces a similar effect to phosphorothioate stabilization, suggesting either that NANP-bound CpG-OHs may be less susceptible to DNase degradation, or alternatively that the effects of NANP-mediated CpG-OH clustering may be able to compensate for the limited

stability of phosphodiester CpG overhangs. Finally, we tested the effect of the proximity between

the CpG motif and the NANP by varying the location of the CpG along the ssDNA overhang.
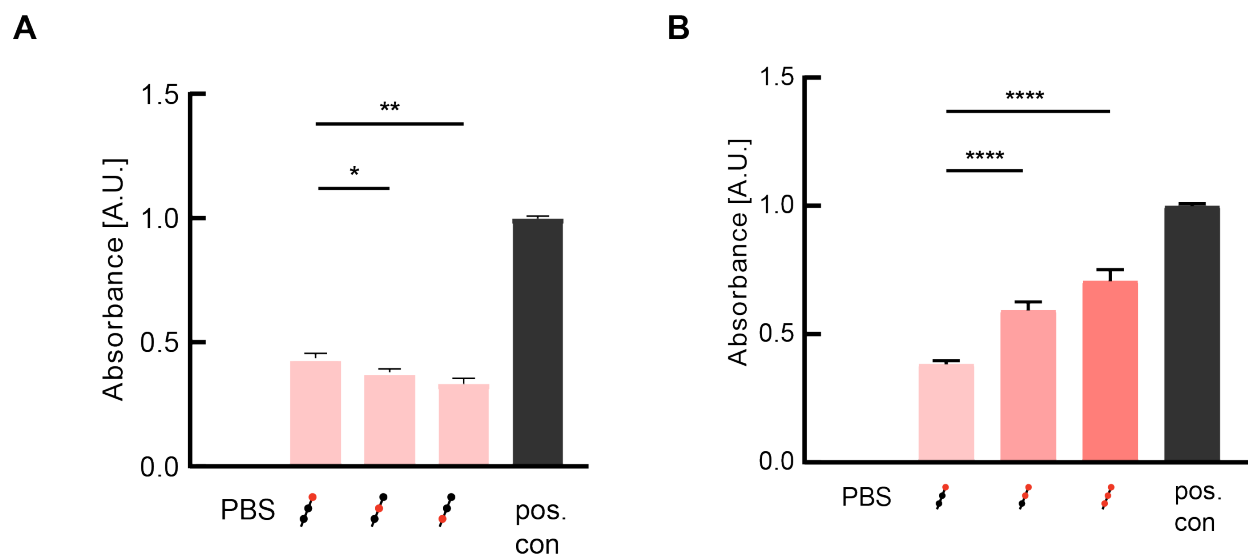
Figure 3.11. The location and number of CpG motifs within ssDNA overhangs affects the strength of TLR9 activation. (A) The magnitude of TLR9 signaling decreases as the CpG motif (denoted by the orange circle) is shifted closer to the nanoparticle. (B) Increasing the number of CpG motifs on each overhang results in a corresponding increase in TLR9 activation. The total concentration of overhangs is kept constant in each sample. Data show the average absorbance of samples in triplicate with standard error, where n = 3 biologically independent assays. P values are from a one-sided analysis of variance (ANOVA) with correction for multiple comparisons (*: P ≤ 0.05; **: P ≤ 0.01, ***: P ≤ 0.001, ****: P ≤ 0.0001). All unlabeled pair-wise comparisons are not significant.

We found that NANPs where the CpG motif was furthest away from the NANP induced significantly higher levels of TLR9 activation compared to NANPs where the CpG motif was spatially closer along the ssDNA overhang (Figure 3.11). This suggests that the NANP may provide a steric hindrance towards CpG-TLR9 binding when there is not sufficient distance

between the dual-duplex edge of the NANP and the CpG motif, which highlights spatial distancing as another parameter that can be tuned to influence the magnitude of TLR9 activation.

In the previous assays, the CpG-OH valency, inter-CpG distance, and pattern of presentation were all unique for each of the NANPs tested. As such, any or all of these properties could theoretically be a potential determinant of the strength of TLR9 activation. However, as these variables are interdependent, it is difficult to evaluate any of these properties in isolation using the current system. To partially circumvent this technical limitation, we reduced the complexity of the system by holding the CpG-OH copy number constant and analyzing the effects of the remaining parameters on the TLR9 pathway. As the crystal structure of an activated TLR9 dimer in complex with two CpG oligos has been published, we measured the inter-CpG distance using PyMOL and determined it to be approximately 7-8nm. Consequently, we fabricated three PB84 variants, each displaying 10 copies of CpG-OH, in which the distances between adjacent CpG overhangs were varied across constructs (Figure 3.12A). The distance between adjacent CpG overhangs on one construct was 7nm, matching the measured distance between TLR9-bound CpG oligos, while the inter-CpG distances on the other two constructs were 14nm and 20nm, distances which are significantly larger than the theoretical 'optimal' spacing. Upon transfecting these samples into HEK-Blue TLR9 cells, we observed that the NANP in which the distance between adjacent CpG-OHs was 7nm induced a significantly higher level of TLR9 activation compared to the constructs in which the spacing between adjacent CpG-OHs was significantly larger than the measured distance between CpG binding pockets.
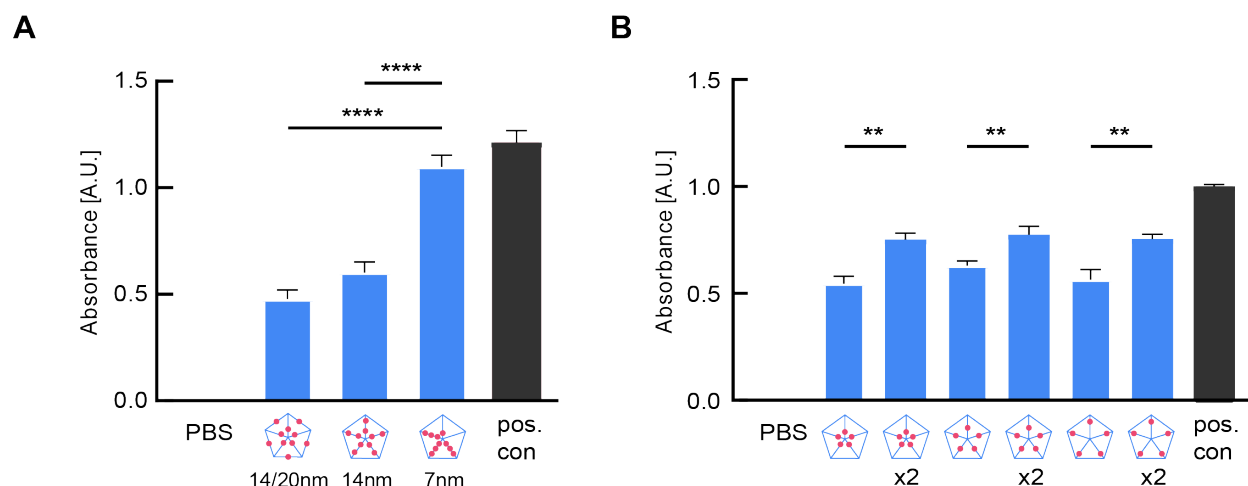
Figure 3.12. TLR9 activation is dependent on CpG-OH valency, inter-CpG distance, and spatial distribution of CpG-OHs. HEK-Blue TLR9 cells were incubated with NANPs displaying 100nM CpG for 24 hours. (A) TLR9 activation measured in response to stimulation with PB84 displaying 10 CpG-OHs with different spacings between adjacent CpGs shows that the strongest TLR9 activation is induced by the 7nm spacing construct. (B) Comparison of PB84 constructs in which either one or both sides of the NANP displays 5 CpG-OHs. TLR9 activation is significantly stronger for constructs in which both sides of PB84 are modified to display CpG-OHs. Absorbances were normalized with the ODN2006 positive control (pos. con.). Data show the average absorbance of samples in triplicate with standard error, where n = 3 biologically independent assays. P values are from a one-sided analysis of variance (ANOVA) with correction for multiple comparisons (*: $P \leq 0.05$; **: $P \leq 0.01$, ***: $P \leq 0.001$, ****: $P \leq 0.0001$). All unlabeled pair-wise comparisons are not significant.

Additionally, in order to investigate the effect of spatial distribution of CpG-OHs across different faces of the NANP, we synthesized PB84 constructs with five CpG-OHs displayed at low, medium, and high clustering densities on either one or both sides of the NANP. Interestingly, we found that there was a significant difference in TLR9 signaling between each of the five CpG-OH constructs and their corresponding 10 CpG-OH counterparts, even though the total CpG concentration and the density of CpG-OH motifs was held constant for each pair of NANPs (Figure 3.12B). As

clustering of multiple activated TLR dimers in close proximity has been shown in previous studies to mediate enhanced downstream signaling via stable myddosome formation, the ability of the 10 CpG-OH construct to coordinate binding of more dimers than its five CpG-OH counterpart may explain the observed increase in TLR9 activation levels.

CONCLUDING DISCUSSION

We characterized baseline immunostimulatory effects of wireframe DNA NANPs in reporter cell lines and show that the immune response towards these nanostructures is primarily driven by the cGAS-STING pathway, while intact nanostructures do not induce a TLR9 response. We then functionalized these nanostructures with immunostimulatory CpG overhangs and demonstrated that both TLR9 activation levels and the downstream production of Type I and III IFNs can be controllably modulated by adjusting critical design parameters. These parameters included CpG valency, clustering density, and pattern of presentation, as well as NANP geometry and CpG motif stabilization. Lastly, we investigated several physical parameters regulating intra- versus inter-TLR9 dimer binding, noting considerable future work to be performed to fully understand the mechanistic basis of immune signaling by NANPs.

Specifically, one open question is how the lipofectamine-complexed NANPs are able to escape from endosomes and activate cytosolic cGAS-STING. As suggested in previous studies, lipofectamine may help NANPs avoid lysosomal degradation by evading active transport and enable intact NANPs to escape through transient pores in the endosomal membrane[54]; alternatively, NANPs may experience partial degradation within the acidic environment prior to endosomal escape, but not so much that cGAS-STING cannot be activated. Another question is how NANP

immunogenicity may change with nuclease- or pH-mediated NANP degradation over time. While cGAS-STING activation levels will likely decrease as the dual-duplex edges of the NANP are progressively degraded, the heightened activation of the TLR9 pathway by the unfolded scaffold and staple formulations compared to the folded constructs suggests that TLR9 activation may increase over time, highlighting the importance of a future investigation into the effects of scaffold and staple sequence composition on NANP immunogenicity. Lastly, another important question that would be interesting to explore more thoroughly in future studies is whether wireframe DNA NANPs primarily modulate TLR9 activation levels by enhancing activation of individual TLR9 dimers or coordinating activation of clusters of dimers, which might be resolved by super-resolution fluorescence imaging[55–57].

Taken together, the nanoscale precision at which ligands can be displayed on DNA nanostructures provides a platform to investigate this dual binding regime hypothesis. Because of the flexibility of ssDNA, rather than an optimal distance between adjacent CpG-OHs, there likely exists a range of inter-CpG distances that allows for effective TLR9 dimer activation as well as a minimal distance between sets of CpG-OHs that allows for multi-TLR9 dimer binding. This also implies that there must be a threshold distance between CpG overhangs above which intra-TLR9 dimer binding is no longer possible and binding, if any occurs, must take place across different TLR9 dimers. Further exploration of this hypothesis using our DNA NANP platform may deepen our understanding of TLR pathway activation and allow for the rational design of nanostructures that can efficiently coordinate the formation of activated TLR9 clusters for enhanced signaling activation.

While nucleic acid nanodevices have been applied to various therapeutic areas, their immunostimulatory properties have been explored to a lesser degree. Such fundamental understanding of NANP immunological properties is important to inform viable therapeutic, prophylactic, and diagnostic capabilities of these novel nanomaterials. By establishing a deeper understanding of the baseline immunostimulatory properties of unmodified wireframe DNA NANPs and the relative impacts of various NANP design parameters on DNA-NANP-mediated tunable TLR9 activation, we can work towards enabling the design of DNA nanostructures with specific immunostimulatory profiles for therapeutic outcomes.

MATERIALS AND METHODS

Scaffold Synthesis.

The circular ssDNA scaffolds used to fold all variants of PB84 and ICO42 described in the paper were produced using a previous published method of bacterial production[16]. *E. coli* SS320 cells (Lucigen) containing the circular phagemid comprising the target scaffold sequence and a M13cp helper plasmid, provided by Dr. Andrew Bradbury (Los Alamos National Laboratory), were grown overnight in 25mL 2 x YT broth (manufacturer, cat. code) supplemented with 100 ug/mL ampicillin (Sigma-Aldrich, cat. code A5354-10ML), 15 ug/mL chloramphenicol (Sigma-Aldrich, cat. code C0378-25G) and 5 ug/mL tetracycline (Sigma-Aldrich, cat. code T7660) in a 200mL flask shaken at 200 RPM at 37C. After 16 hours, the overnight was diluted to an OD600 of 0.05 in 2 x YT broth containing the same supplements and grown for 8 hours at 200 RPM at 37C.

Scaffold Purification.

Bacterial cultures were centrifuged at 4000 x g for 30 minutes at 4C, transferred into a clean bottle, and subjected to an identical centrifugation step. The clarified supernatant was then filtered with a 0.45 um cellulose acetate filter (Sigma-Aldrich, cat. code CLS430516-12EA), poured into a sterile 750mL bottle along with 6% w/v PEG-8000 (Sigma-Aldrich, cat. code P2139-500G) and 3% w/v NaCl, and stirred continuously for 16 hours at 4C. Following PEG precipitation, phage was harvested by centrifuging the solution at 20,000 x g for 30 hour at 4C and discarding the supernatant. The phage-containing pellet was then processed using an Endofree Plasmid Giga Kit (Qiagen, cat. code 12391), with the following adjustments: a final concentration of 20 ug/mL Proteinase K was added to Buffer P1, and the solution was incubated at 37C for 1 hour prior to addition of Buffer P2. Following Buffer P2 addition, the solution was heated to 70C for 10 minutes

and allowed to cool back to room temperature before proceeding with the rest of the standard protocol. Finally, after addition of Buffer ER, 200mL of 100% ethanol was added to improve ssDNA binding. All remaining steps are left unchanged. The concentration of purified ssDNA was determined by measuring absorbance at 280nm using a Nanodrop, and the purity of the sample was evaluated by running it on a 1% agarose gel in 1X TAE stained with SybrSafe. A quantitative measure of sample endotoxin level was determined using Endosafe LAL cartridges with the Endosafe nexgen-PTS system, and Rapid single-test LAL vials (Charles River Laboratories) were used to establish qualitatively whether sample endotoxin levels were below a pre-specified threshold. Both endotoxin assays were performed according to manufacturer protocols.

Endotoxin purification.

In the cases when additional endotoxin purification was needed, chilled 10% v/v Triton X-114 was added to the sample to a final concentration of 2% v/v Triton X-114. The sample was placed on a rocker at 4C for 30 minutes, then transferred to a rocker at 37C for 5 minutes. Lastly, the sample was centrifuged at 30790xg for 30 minutes at 37C. The sample was then carefully removed from the centrifuge and the top layer of the phase-separated sample was gently pipetted into an endotoxin-free microcentrifuge tube. This process was repeated until sample endotoxin levels were below the desired threshold for subsequent assays.

Synthesis and Purification of DNA NANPs.

NANPs were designed using DAEDALUS, with modifications to add ssDNA overhangs being carried out in Tiamat or UCSF Chimera. Briefly, to design inward- or outward-facing overhangs, the locations of nick positions were shifted by the required number of bases to ensure that the 3'

ends of staples would be pointing perpendicularly inward or outward with respect to the center of the nanostructure. No additional nicks were created and all new nick positions were located such that at least 8 bases separated the nick position from the location of the nearest crossover. The modified NANP designs were exported into Excel, and the staples chosen to be functionalized with ssDNA overhangs were extended by concatenating the sequence of the overhang to the 3' end of the existing staple sequence (Supplementary Table 2). To fold NANPs, a circular exact size bacterially produced ssDNA scaffold was mixed with a 5X excess of ssDNA staples, 1X TAE, 12mM MgCl$_2$, and nuclease-free water, and folded through a 13 hour thermal annealing process in which the temperature was gradually ramped down from 95C to 25C as described in previous literature[9]. To remove the excess staples and buffer exchange into sterile PBS, nanostructures were pipetted into Amicon Ultra 100kDa MWCO centrifugal filters (Sigma-Aldrich, cat. code UFC810024) and spun at 1000xg for 30 minutes at room temperature for up to 5 rounds. In between each round, the flow-through was discarded and additional PBS was added. Staples were purchased through Integrated DNA Technologies (IDT) or synthesized in-house using a Dr. Oligo synthesizer following the recommended protocol and purified using a size exclusion column with a Waters HPLC. Phosphoramidites for in-house synthesis of the phosphorothioated CpG overhang-containing staples (Supplementary Table 2) were ordered from Glen Research.

Gel Shift Assay, Dynamic Light Scattering, Fluorimetry.

Agarose gel electrophoresis was used to analyze degree of NANP folding. 1.5% low melt agarose was dissolved in buffer containing 1X TAE and 12 mM MgCl$_2$ and slowly heated to boiling, then cooled back to room temperature. SybrSafe was mixed into the gel according to manufacturer protocols and the gel was allowed to set for at least 45 minutes at room temperature or 4C prior to

gel electrophoresis. 100ng of DNA NANP sample was mixed with gel loading buffer (final concentration 1X) with water to make up the necessary volume. The gel was placed in a chamber containing pre-chilled gel running buffer consisting of 1X TAE and 12mM $MgCl_2$, and once the samples were loaded, the gel was run at 85V for 120 minutes at 4C. Images were captured using a Typhoon. DLS was used for validation of NANP monodispersity: samples were diluted to 50nM in PBS, and 50uL of each sample was loaded into a plastic cuvette to be evaluated on a ZetaSizer. Samples were analyzed in triplicate using single angle scattering. Quantification of CpG copy number on NANPs was measured using a Tecan Spark. A Cy5-conjugated ssDNA oligo purchased from IDT was hybridized to folded NANPs via a 4 hour thermal annealing ramp during which the sample temperature was decreased incrementally from 37C to 25C. Fluorescently labeled NANPs were purified using the spin purification method described above and concentration was determined using UV-Vis measurements taken on a Nanodrop (specify) with absorbances set to 260nm and 647nm. A fluorescent calibration curve was produced using serial dilutions of the free CpG overhang hybridized to the Cy5-conjugated ssDNA oligo and measured with the Tecan Spark to produce a regression line with which the coverage of CpGs on each NANP variant was determined.

Reporter Cell Assays.

HEK-Blue TLR9 reporter cells and THP1-Dual reporter cells were purchased from Invivogen and cultured according to manufacturer protocols. One day prior to cell assays, TLR9 cells were detached from the flask, centrifuged at 250x*g* for 5 minutes at room temperature and resuspended in fresh growth medium, consisting of Dulbecco's Modified Eagle Medium with 4.5 g/L glucose (Sigma-Aldrich, cat. code D6546-6X500ML), 10% heat-inactivated FBS (Fisher Scientific, cat.

code SH3007002HI), 100 U/mL penicillin, 100 ug/mL streptomycin (Life Technologies, cat. code 15140148), 100 ug/mL Normocin (manufacturer, cat. code) and 2mM L-glutamine (Fisher Scientific, cat. code SH3003401). Cells were counted using a Nexcelom Cellometer Auto 2000 and 80,000 cells in 160uL media were seeded into a flat-bottom 96-well plate. Cells were allowed to adhere for 24 hours prior to removal of growth medium from each well and subsequent addition of 160uL of prewarmed HEK-Blue Detection Medium. 40uL of NANP sample was added to each well, with each sample being tested in triplicate, and cells were incubated at 37C and 5% CO_2 for 20-24 hours. TLR9 activation levels were quantified by reading the absorbance of the 96-well plate at 620-655nm using a spectrophotometer and normalized to a PBS negative control. For cGAS-STING assays, THP1-Dual reporter cells were centrifuged at 250xg for 5 minutes at room temperature and resuspended in fresh test medium, consisting of RPMI 1640 (Life Technologies, cat. code 11875093), 2mM L-glutamine, 25mM HEPES, 10% heat-inactivated FBS, 100 U/mL penicillin, and 100 ug/mL streptomycin. 100,000 cells in 160uL media were seeded into a flat-bottom 96-well plate and 40uL of NANP sample was added to each well, with each sample being tested in triplicate. Cells were incubated at 37C and 5% CO_2 for 20-24 hours, after which 20uL of cell supernatant was carefully transferred into an opaque white 96-well plate. 50uL of QUANTI-Luc assay solution (Invivogen, cat. code rep-qlc1) was added to each row of the 96-well plate, the plate was gently tapped several times on each side to mix, and luminescence readings were quantified using a luminometer set to a 0.1 second reading time. Positive controls for TLR9 and cGAS-STING assays were not complexed with lipofectamine prior to incubation with PBMCs.

PBMC Isolation, Stimulation, and Multiplex ELISA.

Research donor blood was obtained from anonymous healthy donor volunteers under the IRB approved NCI-at-Frederick Protocol OH9-C-N046 and the IRB approved exemption E-3359. PBMCs were isolated from three separate human donor buffy coats following the NCL protocol ITA-10 and tested following protocol ITA-27[53]. Briefly, blood was mixed at a 1:1 ratio with fresh PBS and gently layered onto Ficoll-Paque (Fisher Scientific, cat. code 45001751) at a ratio of 3 mL Ficoll-Paque to 4mL diluted blood in a 50mL Falcon tube. Tubes were centrifuged at 900x*g* for 30 minutes with minimum acceleration and with the brakes off, after which the PBMC layer was gently pipetted into a new 50mL tube. PBMCs were washed three times with HBSS (Life Technologies, cat. code 24020117) and resuspended in complete RPMI. Cells were counted using a Nexcelom Cellometer Auto 2000 and 1.25×10^6 cells in 160uL media were seeded into a flat-bottom 96-well plate. NANP samples were complexed with Lipofectamine 2000 (Life Technologies, cat. code 11668027) at a ratio of 4:1and incubated at room temperature for 30 minutes, after which OptiMEM (Life Technologies, cat. code 31985062) was added to dilute the NANPs to the correct working concentration. 40uL of the lipofectamine-complexed NANP sample was added to each well, with each sample being tested in triplicate, and cells were incubated at 37C and 5% CO_2 for 20-24 hours. 150uL of cell supernatant was carefully transferred into a new 96-well plate, and the plate was flash frozen and stored at -80C. Supernatant was then shipped on dry ice to Quansys Biosciences, which analyzed production of IFNα, IFNβ, IFNω, and IFNλ using multiplex ELISA.

References

(1) Jun, H.; Wang, X.; Parsons, M. F.; Bricker, W. P.; John, T.; Li, S.; Jackson, S.; Chiu, W.; Bathe, M. Rapid Prototyping of Arbitrary 2D and 3D Wireframe DNA Origami. *Nucleic Acids Research* **2021**, *49* (18), 10265–10274. https://doi.org/10.1093/nar/gkab762.

(2) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera-- A Visualization System for Exploratory Research and Analysis. *Journal of Computational Chemistry* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.

(3) Benson, E.; Mohammed, A.; Bosco, A.; Teixeira, A. I.; Orponen, P.; Högberg, B. Computer-Aided Production of Scaffolded DNA Nanostructures from Flat Sheet Meshes. *Angewandte Chemie International Edition* **2016**, *55* (31), 8869–8872. https://doi.org/10.1002/anie.201602446.

(4) Jun, H.; Shepherd, T. R.; Zhang, K.; Bricker, W. P.; Li, S.; Chiu, W.; Bathe, M. Automated Sequence Design of 3D Polyhedral Wireframe DNA Origami with Honeycomb Edges. *ACS Nano* **2019**, acsnano.8b08671. https://doi.org/10.1021/acsnano.8b08671.

(5) Benson, E.; Mohammed, A.; Gardell, J.; Masich, S.; Czeizler, E.; Orponen, P.; Högberg, B. DNA Rendering of Polyhedral Meshes at the Nanoscale. *Nature* **2015**, *523* (7561), 441–444. https://doi.org/10.1038/nature14586.

(6) Dietz, H.; Douglas, S. M.; Shih, W. M. Folding DNA into Twisted and Curved Nanoscale Shapes. *Science* **2009**, *325* (5941), 725–730. https://doi.org/10.1126/science.1174251.

(7) Castro, C. E.; Kilchherr, F.; Kim, D.-N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H. A Primer to Scaffolded DNA Origami. *Nature Methods* **2011**, *8* (3), 221–229. https://doi.org/10.1038/nmeth.1570.

(8) Douglas, S. M.; Marblestone, A. H.; Teerapittayanon, S.; Vazquez, A.; Church, G. M.; Shih, W. M. Rapid Prototyping of 3D DNA-Origami Shapes with CaDNAno. *Nucleic Acids Research* **2009**, *37* (15), 5001–5006. https://doi.org/10.1093/nar/gkp436.

(9) Veneziano, R.; Ratanalert, S.; Zhang, K.; Zhang, F.; Yan, H.; Chiu, W.; Bathe, M. Designer Nanoscale DNA Assemblies Programmed from the Top Down. *Science* **2016**, *352* (6293), 1534–1534. https://doi.org/10.1126/science.aaf4388.

(10) Veneziano, R.; Moyer, T. J.; Stone, M. B.; Wamhoff, E.-C.; Read, B. J.; Mukherjee, S.; Shepherd, T. R.; Das, J.; Schief, W. R.; Irvine, D. J.; Bathe, M. Role of Nanoscale Antigen Organization on B-Cell Activation Probed Using DNA Origami. *Nature Nanotechnology* **2020**, *15* (8), 716–723. https://doi.org/10.1038/s41565-020-0719-0.

(11) Lee, H.; Lytton-Jean, A. K. R.; Chen, Y.; Love, K. T.; Park, A. I.; Karagiannis, E. D.; Sehgal, A.; Querbes, W.; Zurenko, C. S.; Jayaraman, M.; Peng, C. G.; Charisse, K.; Borodovsky, A.; Manoharan, M.; Donahoe, J. S.; Truelove, J.; Nahrendorf, M.; Langer, R.; Anderson, D. G. Molecularly Self-Assembled Nucleic Acid Nanoparticles for Targeted in Vivo SiRNA Delivery. *Nature Nanotechnology* **2012**, *7* (6), 389–393. https://doi.org/10.1038/nnano.2012.73.

(12) Douglas, S. M.; Bachelet, I.; Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **2012**, *335* (6070), 831–834. https://doi.org/10.1126/science.1214081.

(13) Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440* (7082), 297–302. https://doi.org/10.1038/nature04586.

(14) Praetorius, F.; Kick, B.; Behler, K. L.; Honemann, M. N.; Weuster-Botz, D.; Dietz, H. Biotechnological Mass Production of DNA Origami. *Nature* **2017**, *552* (7683), 84–87. https://doi.org/10.1038/nature24650.

(15) Engelhardt, F. A. S.; Praetorius, F.; Wachauf, C. H.; Brüggenthies, G.; Kohler, F.; Kick, B.; Kadletz, K. L.; Pham, P. N.; Behler, K. L.; Gerling, T.; Dietz, H. Custom-Size, Functional, and

Durable DNA Origami with Design-Specific Scaffolds. *ACS Nano* **2019**, *13* (5), 5015–5027. https://doi.org/10.1021/acsnano.9b01025.

(16) Shepherd, T. R.; Du, R. R.; Huang, H.; Wamhoff, E.-C.; Bathe, M. Bioproduction of Pure, Kilobase-Scale Single-Stranded DNA. *Scientific Reports* **2019**, *9* (1), 6121. https://doi.org/10.1038/s41598-019-42665-1.

(17) Knappe, G. A.; Wamhoff, E.-C.; Read, B. J.; Irvine, D. J.; Bathe, M. *In Situ* Covalent Functionalization of DNA Origami Virus-like Particles. *ACS Nano* **2021**, *15* (9), 14316–14322. https://doi.org/10.1021/acsnano.1c03158.

(18) Liu, X.; Xu, Y.; Yu, T.; Clifford, C.; Liu, Y.; Yan, H.; Chang, Y. A DNA Nanostructure Platform for Directed Assembly of Synthetic Vaccines. *Nano Letters* **2012**, *12* (8), 4254–4259. https://doi.org/10.1021/nl301877k.

(19) Irvine, D. J.; Hanson, M. C.; Rakhra, K.; Tokatlian, T. Synthetic Nanoparticles for Vaccines and Immunotherapy. *Chemical Reviews* **2015**, *115* (19), 11109–11146. https://doi.org/10.1021/acs.chemrev.5b00109.

(20) Hong, E.; Halman, J.; Shah, A.; Cedrone, E.; Truong, N.; Afonin, K.; Dobrovolskaia, M. Toll-Like Receptor-Mediated Recognition of Nucleic Acid Nanoparticles (NANPs) in Human Primary Blood Cells. *Molecules* **2019**, *24* (6), 1094. https://doi.org/10.3390/molecules24061094.

(21) Surana, S.; Shenoy, A. R.; Krishnan, Y. Designing DNA Nanodevices for Compatibility with the Immune System of Higher Organisms. *Nature Nanotechnology* **2015**, *10* (9), 741–747. https://doi.org/10.1038/nnano.2015.180.

(22) Hong, E.; Halman, J. R.; Shah, A. B.; Khisamutdinov, E. F.; Dobrovolskaia, M. A.; Afonin, K. A. Structure and Composition Define Immunorecognition of Nucleic Acid Nanoparticles. *Nano Letters* **2018**, *18* (7), 4309–4321. https://doi.org/10.1021/acs.nanolett.8b01283.

(23) Schüller, V. J.; Heidegger, S.; Sandholzer, N.; Nickels, P. C.; Suhartha, N. A.; Endres, S.; Bourquin, C.; Liedl, T. Cellular Immunostimulation by CpG-Sequence-Coated DNA Origami Structures. *ACS Nano* **2011**, *5* (12), 9696–9702. https://doi.org/10.1021/nn203161y.

(24) Schlee, M.; Hartmann, G. Discriminating Self from Non-Self in Nucleic Acid Sensing. *Nature Reviews Immunology* **2016**, *16* (9), 566–580. https://doi.org/10.1038/nri.2016.78.

(25) Paludan, S. R. Activation and Regulation of DNA-Driven Immune Responses. *Microbiology and Molecular Biology Reviews* **2015**, *79* (2), 225–241. https://doi.org/10.1128/MMBR.00061-14.

(26) Paludan, S. R.; Bowie, A. G. Immune Sensing of DNA. *Immunity* **2013**, *38* (5), 870–880. https://doi.org/10.1016/j.immuni.2013.05.004.

(27) Wu, J.; Chen, Z. J. Innate Immune Sensing and Signaling of Cytosolic Nucleic Acids. *Annual Review of Immunology* **2014**, *32* (1), 461–488. https://doi.org/10.1146/annurev-immunol-032713-120156.

(28) Li, X.; Shu, C.; Yi, G.; Chaton, C. T.; Shelton, C. L.; Diao, J.; Zuo, X.; Kao, C. C.; Herr, A. B.; Li, P. Cyclic GMP-AMP Synthase Is Activated by Double-Stranded DNA-Induced Oligomerization. *Immunity* **2013**, *39* (6), 1019–1031. https://doi.org/10.1016/j.immuni.2013.10.019.

(29) Chen, Q.; Sun, L.; Chen, Z. J. Regulation and Function of the CGAS–STING Pathway of Cytosolic DNA Sensing. *Nature Immunology* **2016**, *17* (10), 1142–1149. https://doi.org/10.1038/ni.3558.

(30) Sun, L.; Wu, J.; Du, F.; Chen, X.; Chen, Z. J. Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway. *Science* **2013**, *339* (6121), 786–791. https://doi.org/10.1126/science.1232458.

(31) Zhou, W.; Whiteley, A. T.; de Oliveira Mann, C. C.; Morehouse, B. R.; Nowak, R. P.; Fischer, E. S.; Gray, N. S.; Mekalanos, J. J.; Kranzusch, P. J. Structure of the Human CGAS–DNA Complex

Reveals Enhanced Control of Immune Surveillance. *Cell* **2018**, *174* (2), 300-311.e11. https://doi.org/10.1016/j.cell.2018.06.026.

(32)   Civril, F.; Deimling, T.; de Oliveira Mann, C. C.; Ablasser, A.; Moldt, M.; Witte, G.; Hornung, V.; Hopfner, K.-P. Structural Mechanism of Cytosolic DNA Sensing by CGAS. *Nature* **2013**, *498* (7454), 332–337. https://doi.org/10.1038/nature12305.

(33)   Andreeva, L.; Hiller, B.; Kostrewa, D.; Lässig, C.; de Oliveira Mann, C. C.; Jan Drexler, D.; Maiser, A.; Gaidt, M.; Leonhardt, H.; Hornung, V.; Hopfner, K.-P. CGAS Senses Long and HMGB/TFAM-Bound U-Turn DNA by Forming Protein–DNA Ladders. *Nature* **2017**, *549* (7672), 394–398. https://doi.org/10.1038/nature23890.

(34)   Kawai, T.; Akira, S. The Role of Pattern-Recognition Receptors in Innate Immunity: Update on Toll-like Receptors. *Nature Immunology* **2010**, *11* (5), 373–384. https://doi.org/10.1038/ni.1863.

(35)   Kawasaki, T.; Kawai, T. Toll-Like Receptor Signaling Pathways. *Frontiers in Immunology* **2014**, *5*. https://doi.org/10.3389/fimmu.2014.00461.

(36)   Krieg, A. M.; Yi, A.-K.; Matson, S.; Waldschmidt, T. J.; Bishop, G. A.; Teasdale, R.; Koretzky, G. A.; Klinman, D. M. CpG Motifs in Bacterial DNA Trigger Direct B-Cell Activation. *Nature* **1995**, *374* (6522), 546–549. https://doi.org/10.1038/374546a0.

(37)   Cervantes, J. L.; Weinerman, B.; Basole, C.; Salazar, J. C. TLR8: The Forgotten Relative Revindicated. *Cellular & Molecular Immunology* **2012**, *9* (6), 434–438. https://doi.org/10.1038/cmi.2012.38.

(38)   Majer, O.; Liu, B.; Barton, G. M. Nucleic Acid-Sensing TLRs: Trafficking and Regulation. *Current Opinion in Immunology* **2017**, *44*, 26–33. https://doi.org/10.1016/j.coi.2016.10.003.

(39)   Latz, E.; Verma, A.; Visintin, A.; Gong, M.; Sirois, C. M.; Klein, D. C. G.; Monks, B. G.; McKnight, C. J.; Lamphier, M. S.; Duprex, W. P.; Espevik, T.; Golenbock, D. T. Ligand-Induced Conformational Changes Allosterically Activate Toll-like Receptor 9. *Nature Immunology* **2007**, *8* (7), 772–779. https://doi.org/10.1038/ni1479.

(40)   Ohto, U.; Ishida, H.; Shibata, T.; Sato, R.; Miyake, K.; Shimizu, T. Toll-like Receptor 9 Contains Two DNA Binding Sites That Function Cooperatively to Promote Receptor Dimerization and Activation. *Immunity* **2018**, *48* (4), 649-658.e4. https://doi.org/10.1016/j.immuni.2018.03.013.

(41)   Deguine, J.; Barton, G. M. MyD88: A Central Player in Innate Immune Signaling. *F1000Prime Reports* **2014**, *6*. https://doi.org/10.12703/P6-97.

(42)   Marongiu, L.; Gornati, L.; Artuso, I.; Zanoni, I.; Granucci, F. Below the Surface: The Inner Lives of TLR4 and TLR9. *Journal of Leukocyte Biology* **2019**, *106* (1), 147–160. https://doi.org/10.1002/JLB.3MIR1218-483RR.

(43)   Bonham, K. S.; Orzalli, M. H.; Hayashi, K.; Wolf, A. I.; Glanemann, C.; Weninger, W.; Iwasaki, A.; Knipe, D. M.; Kagan, J. C. A Promiscuous Lipid-Binding Protein Diversifies the Subcellular Sites of Toll-like Receptor Signal Transduction. *Cell* **2014**, *156* (4), 705–716. https://doi.org/10.1016/j.cell.2014.01.019.

(44)   Krieg, A. M. Toll-like Receptor 9 (TLR9) Agonists in the Treatment of Cancer. *Oncogene* **2008**, *27* (2), 161–167. https://doi.org/10.1038/sj.onc.1210911.

(45)   Chen, H. C.; Sun, B.; Tran, K. K.; Shen, H. Effects of Particle Size on Toll-like Receptor 9-Mediated Cytokine Profiles. *Biomaterials* **2011**, *32* (6), 1731–1737. https://doi.org/10.1016/j.biomaterials.2010.10.059.

(46)   Luchner, M.; Reinke, S.; Milicic, A. TLR Agonists as Vaccine Adjuvants Targeting Cancer and Infectious Diseases. *Pharmaceutics* **2021**, *13* (2), 142. https://doi.org/10.3390/pharmaceutics13020142.

(47)   Pohar, J.; Kužnik Krajnik, A.; Jerala, R.; Benčina, M. Minimal Sequence Requirements for Oligodeoxyribonucleotides Activating Human TLR9. *The Journal of Immunology* **2015**, *194* (8), 3901–3908. https://doi.org/10.4049/jimmunol.1402755.

(48)   Vollmer, J.; Weeratna, R.; Payette, P.; Jurk, M.; Schetter, C.; Laucht, M.; Wader, T.; Tluk, S.; Liu, M.; Davis, H. L.; Krieg, A. M. Characterization of Three CpG Oligodeoxynucleotide Classes with Distinct Immunostimulatory Activities. *European Journal of Immunology* **2004**, *34* (1), 251–262. https://doi.org/10.1002/eji.200324032.

(49)   Li, J.; Pei, H.; Zhu, B.; Liang, L.; Wei, M.; He, Y.; Chen, N.; Li, D.; Huang, Q.; Fan, C. Self-Assembled Multivalent DNA Nanostructures for Noninvasive Intracellular Delivery of Immunostimulatory CpG Oligonucleotides. *ACS Nano* **2011**, *5* (11), 8783–8789. https://doi.org/10.1021/nn202774x.

(50)   Comberlato, A.; Koga, M. M.; Nüssing, S.; Parish, I. A.; Bastings, M. M. C. Spatially Controlled Activation of Toll-like Receptor 9 with DNA-Based Nanomaterials. *Nano Letters* **2022**, acs.nanolett.2c00275. https://doi.org/10.1021/acs.nanolett.2c00275.

(51)   Leleux, J. A.; Pradhan, P.; Roy, K. Biophysical Attributes of CpG Presentation Control TLR9 Signaling to Differentially Polarize Systemic Immune Responses. *Cell Reports* **2017**, *18* (3), 700–710. https://doi.org/10.1016/j.celrep.2016.12.073.

(52)   Jin, J.-O.; Park, H.; Zhang, W.; de Vries, J. W.; Gruszka, A.; Lee, M. W.; Ahn, D.-R.; Herrmann, A.; Kwak, M. Modular Delivery of CpG-Incorporated Lipid-DNA Nanoparticles for Spleen DC Activation. *Biomaterials* **2017**, *115*, 81–89. https://doi.org/10.1016/j.biomaterials.2016.11.020.

(53)   Dobrovolskaia, M. A.; Afonin, K. A. Use of Human Peripheral Blood Mononuclear Cells to Define Immunological Properties of Nucleic Acid Nanoparticles. *Nature Protocols* **2020**, *15* (11), 3678–3698. https://doi.org/10.1038/s41596-020-0393-6.

(54)   Cardarelli, F.; Digiacomo, L.; Marchini, C.; Amici, A.; Salomone, F.; Fiume, G.; Rossetta, A.; Gratton, E.; Pozzi, D.; Caracciolo, G. The Intracellular Trafficking Mechanism of Lipofectamine-Based Transfection Reagents and Its Implication for Gene Delivery. *Scientific Reports* **2016**, *6* (1), 25879. https://doi.org/10.1038/srep25879.

(55)   Latz, E.; Schoenemeyer, A.; Visintin, A.; Fitzgerald, K. A.; Monks, B. G.; Knetter, C. F.; Lien, E.; Nilsen, N. J.; Espevik, T.; Golenbock, D. T. TLR9 Signals after Translocating from the ER to CpG DNA in the Lysosome. *Nature Immunology* **2004**, *5* (2), 190–198. https://doi.org/10.1038/ni1028.

(56)   Pohar, J.; Yamamoto, C.; Fukui, R.; Cajnko, M.-M.; Miyake, K.; Jerala, R.; Benčina, M. Selectivity of Human TLR9 for Double CpG Motifs and Implications for the Recognition of Genomic DNA. *The Journal of Immunology* **2017**, *198* (5), 2093–2104. https://doi.org/10.4049/jimmunol.1600757.

(57)   Guo, S.-M.; Veneziano, R.; Gordonov, S.; Li, L.; Danielson, E.; Perez de Arce, K.; Park, D.; Kulesa, A. B.; Wamhoff, E.-C.; Blainey, P. C.; Boyden, E. S.; Cottrell, J. R.; Bathe, M. Multiplexed and High-Throughput Neuronal Fluorescence Imaging with Diffusible Probes. *Nature Communications* **2019**, *10* (1), 4377. https://doi.org/10.1038/s41467-019-12372-6.

Chapter 4

Conclusions and Future Directions

CONCLUSIONS

In the decades since the first description of the construction of an immobile Holliday junction by Seeman in 1982, DNA nanotechnology has come a long way. Numerous advancements in the field such as the development of the DX tile, the introduction of DNA origami, and the creation of automated design algorithms, have facilitated the design and synthesis of increasingly complex DNA-based nanodevices capable of sensing, responding to, and interacting with their environments in a controllable manner. The high degree of nanoparticle customizability and the nanoscale precision of ligand functionalization has led to the use of DNA nanoparticles for a vast range of applications in fields such as chemistry, biology, materials science, and computer science. However, the application of DNA nanostructures towards therapeutic directions, specifically, has exposed several limitations in existing scaffold production technologies and has highlighted the need for a deeper understanding of how DNA origami design may influence its immunological recognition.

The M13 bacteriophage genome has long been the most prevalent source of ssDNA for DNA origami synthesis. Not only is it fully sequenced, but its length (~6.4kb) is also sufficient to fold most nanostructures and it is extremely easy to propagate in cultures of *E. coli*. However, there is very little sequence control and the presence of protein coding genes is less than optimal for therapeutic purposes. Thus, there is a need for the field to move away from the use of phage-derived scaffolds and investigate methods of custom scaffold production. In Chapter 2 of my thesis, I describe the development of a novel method of ssDNA bioproduction which is both scalable and enables a high degree of sequence customization. We constructed a phagemid containing a user-defined custom sequence and show that when the phagemid is co-transformed into cells with the

M13cp helper plasmid, this system successfully produces pure phagemid ssDNA in a shaker flask system. Following several rounds of growth condition optimization, we test the scalability of this method with a 10L bioreactor batch fermentation and demonstrate the production of milligram-scale quantities of pure, circular kilobase-length ssDNA with customized sequences. We next show the ease of downstream purification without the need for removal of contaminating dsDNA or helper plasmids, and finally, we validate the use of these scaffolds for DNA nanotechnology by folding the circular ssDNA products into programmed geometries, characterized by gel shift assays and TEM. This method enables the production of DNA origami with custom sequences at the scales necessary for clinical applications while simultaneously streamlining the downstream purification to maximize yield and reduce the required processing time. Additionally, the high degree of sequence control facilitates the synthesis of rationally designed scaffolds containing functional sequences, such as mRNA for the production of therapeutically relevant proteins, recognition sequences for RNA capture, or immunostimulatory motifs for targeted immune activation.

The development of the scaffold production technology described in Chapter 2 enabled the synthesis of DNA origami at the milligram scales required for *in vitro* cell assays and *in vivo* studies, thereby greatly expanding the set of questions that could be explored using this platform. In Chapter 3 of my thesis, I take advantage of this expanded scaffold production capability to delve into several questions about how wireframe DNA nanoparticles interact with the immune system and how they might be engineered for controllable immune activation. Despite previous cytokine-level characterization of the immunostimulatory profiles of other classes of DNA nanostructures, there has not yet been a systematic investigation into the immunological recognition towards

wireframe DNA origami. Additionally, although DNA nanodevices have previously been used to display immune adjuvants, there is still a lack of understanding of how DNA origami design parameters and ligand spatial organization impact the resulting immune response, both in terms of the activation of individual innate immune pathways and in terms of the downstream interferon response. To address these questions, I began by characterizing the response of the TLR9 and cGAS-STING pathways towards a ssDNA scaffold as well as two different geometries folded from this scaffold, a pentagonal bipyramid (PB84) and an icosahedron (ICO42). I found that the structured nanoparticles induced minimal TLR9 activation while the unstructured scaffold control resulted in stronger TLR9 signaling, suggesting that the CpG motifs within the scaffold that are required for TLR9 activation may be inaccessible upon folding. Additionally, cGAS was strongly activated by both folded PB84 and ICO42. This level of activation was higher than what was induced by the unstructured scaffold, reinforcing the idea that nanostructuring plays a role in immunological recognition and suggesting that the immune response towards unmodified wireframe DNA nanostructures is primarily cGAS-mediated. I then tested these constructs in PBMCs, a model system that is commonly used to study the innate immune response, and found that the unmodified, structured nanoparticles induced strong expression of all Type I IFNs as well as IFNλ, a Type III IFN, while the unstructured controls induced moderate levels of interferon expression.

Upon elucidating the baseline immunostimulatory profile of unmodified wireframe DNA nanoparticles, I then sought to engineer nanostructures capable of controllably tuning the immune response. I started by using the TLR9 reporter cell system, which exhibited minimal to no activation in response to unmodified DNA nanoparticles. By attaching increasing copy numbers

of TLR9-agonist CpG overhangs to PB84, I was able to controllably activate TLR9, where the magnitude of TLR9 activation was directly correlated with the number of CpG overhangs displayed on the nanoparticle. Interestingly, this result was consistently achieved even when the total CpG concentration was kept constant across samples, indicating that parameters such as the valency of CpGs per nanoparticle or the spatial organization of the CpG overhangs may be responsible for differential TLR9 activation. To assess whether this result was reproducible outside of the context of TLR9 reporter cells, I incubated PBMCs with PB84 variants displaying 0, 20, or 40 copies of CpG overhangs, and showed that the magnitude of interferon expression was also directly correlated with the copy number of CpG overhangs displayed on the nanoparticle. However, when ICO42 variants displaying 0, 20 or 30 copies of CpG overhangs were tested in the same system, the interferon response towards these constructs was either constant, or in some cases, inversely correlated. As the unmodified ICO42 construct already induced interferon expression levels on par with PB84 displaying 40 CpG overhangs, it is possible that this result is due to the well-characterized 'hook effect' observed frequently in immunoassays; further assay optimization may be required to eliminate this possibility.

To probe the molecular mechanisms that underlie TLR9-nanoparticle interactions, I then designed constructs in which the valency was kept constant but the organization of CpG overhangs varied. When these nanoparticles were evaluated in TLR9 reporter cells, I found that there were significant differences between the magnitude of the TLR9 response towards different constructs, suggesting that CpG nanoscale organization plays an important role in TLR9 binding and activation, although additional studies will be required to elucidate the specific parameters involved and their relative impacts on TLR9 signaling. I also tested PB84 variants in which CpG overhangs were displayed

in identical spatial patterns on either one or both sides of the nanoparticle, and demonstrated that the PB84 constructs displaying CpG overhangs on both sides induced significantly stronger TLR9 activation than their single-side counterparts, even when the CpG concentration was kept constant. As the distance between CpG overhangs displayed on opposite sides of PB84 are too far apart in 3D space to be bound to the same TLR9 dimer according to measurements taken from crystal structures of TLR9 dimer:CpG oligo constructs, this result is consistent with a model in which PB84 can be engaged by more than one TLR9 dimer simultaneously, thereby enhancing downstream signaling activation by the clustering of multiple activated dimers. Overall, these studies begin to uncover the relative impacts of various nanoparticle design parameters on immunological recognition and demonstrate the feasibility of designing wireframe DNA nanoparticles for controllable immunomodulation.

FUTURE DIRECTIONS

In this thesis, I have demonstrated the development of a novel method of scalable ssDNA bioproduction and took steps towards characterizing and designing immunomodulatory wireframe DNA nanoparticles. However, much work can still be done to advance the research described here, and in the following section, I will elaborate on several interesting and important open questions that remain.

With the development of ssDNA production technologies that enable the design and synthesis of custom sequence scaffolds, it is now possible to rationally design scaffolds for immunostimulation or immune evasion. One particularly relevant application would be the enrichment of CpG motifs or the elimination of CpG dinucleotides within the scaffold to control the degree of potential TLR9 recognition. However, in order to synthesize CpG-rich or CpG-free scaffolds for wireframe DNA origami, there are several design considerations to keep in mind. First, the GC content must be kept within the standard 40-60% to ensure proper nanoparticle folding, which means that instead of directly adding or removing CpGs, the arrangement of Gs and Cs within the scaffold must be rearranged to increase or reduce the incidence of CpG dinucleotides. Next, after a sufficient level of enrichment or depletion is reached, the scaffold should be checked for palindromic regions or repeated sequences which would not only increase the difficulty of phagemid construction but may also cause nanoparticle misfolding if staple sequences are not orthogonal. Additionally, prior to the synthesis of CpG-enriched scaffolds to enhance nanoparticle-mediated TLR9 activation, there is another step in the design process, which is to consider the locations of the CpG motifs in geometric space within the folded nanoparticle. As fully intact nanoparticles do not activate TLR9, it is important to place these CpG motifs at locations that are relatively more likely to expose

single-stranded regions to engage in TLR9 binding. One such set of regions are potentially the 3'
or 5' ends of staples within nanoparticle edges and their respective complementary scaffold regions.
It has been demonstrated that the terminal base pairs in DNA duplexes are susceptible to transient
fraying events, which temporarily exposes short ssDNA fragments. Additionally, as staples are
linear, there is no phosphodiester bond between the 5' and 3' ends. Therefore, the locations of
staple termini along the nanoparticle are equivalent to single-stranded nicks, which are known to
be locations of reduced stability due to the potential for DNA unwinding as well as increased
susceptibility to degradation. However, as the process of nanoparticle degradation is not well
understood, these locations may not behave as expected. Thus, it is important to include controls
such as CpG-enriched scaffolds in which the motifs are scattered in random locations across the
folded nanoparticle in order to systematically investigate the effects of immunostimulatory or
immune evasive scaffold design.

Additional future work in the scaffold design direction could focus on minimizing the fixed region
of the phagemid to enhance the degree of sequence control over the produced scaffold. Currently,
the fixed region is 1676 nucleotides in length and includes the f1 origin of replication containing
the associated packaging signal and the beta-lactamase antibiotic resistance gene. While larger
scaffolds can easily be produced from this minimal phagemid by the introduction of additional
custom sequences, this restricts the minimal sequence length that can be synthesized via this
method. One potential workaround would be to remove the beta-lactamase gene, which comprises
~1.2kb of the fixed region; however, the loss of the antibiotic resistance may lead to substantial
reduction in ssDNA yields, as there would no longer be selective pressure on the bacteria to retain
the phagemid. Alternatively, as demonstrated by Specthrie *et al.* and Dotto *et al.*, an additional,

truncated f1 origin can be introduced into the phagemid to act as a terminator. The advantage of this method is that the phagemid itself can be substantially larger than the target sequence, allowing for the retention of the antibiotic resistance gene as well as the addition of other functional elements as needed. However, this method requires substantial optimization, as the termination and initiation domains within the f1 origin overlap, resulting in trace production of incorrect ssDNA products. If engineered properly, however, this would reduce the size of the fixed region to ~300-400 nucleotides, further increasing the degree of sequence control afforded by this method and facilitating the synthesis of an expanded range of scaffold sizes.

While much work has been done in this thesis to characterize the innate immune response towards unmodified nanoparticles, there are still some questions that remain unanswered. For example, one cytosolic DNA sensor that was not investigated in this thesis is AIM2. Therefore, one question that naturally follows is: do unmodified nanoparticles mediate AIM2 activation, and if so, is inflammasome activation and/or pyroptosis observed in response? It has been demonstrated that dsDNA of lengths shorter than ~50bp were not sufficient to induce AIM2 activation, while optimal activation was achieved with dsDNA oligos of lengths between 70– 80bp. As each edge of dual-duplex wireframe DNA origami is a ligand for AIM2, the effects of nanoparticle edge length of AIM2 activation would be an interesting question to pursue in AIM2-expressing reporter cells which eliminate potential cross-talk from other innate immune pathways. Similarly, while I have shown that unmodified nanostructures activate cGAS-STING more strongly than their unstructured counterparts, the effect of nanoparticle geometry, edge length, and number of edges remains unknown. As cGAS activation, much like AIM2 activation, is known to have a strong length dependence, nanoparticle edge length may strongly influence the magnitude of cGAS

activation. In addition, the number of edges directly impacts the number of dsDNA oligos available for cGAS binding and may contribute to the potency of nanoparticle-mediate cGAS signaling as well. Further investigation of how nanoparticle properties such as geometry and edge length influence cGAS and AIM2 activation will contribute to our understanding of immunological recognition of DNA nanostructures and help to guide the designs of future generations of immunomodulatory nanoparticles.

There also remains a lack of understanding of if or how wireframe DNA nanoparticles can be internalized into immune cells in the absence of a transfection reagent or internalization ligand. As shown in Figure 3.2, no immune activation was observed when cells were treated with nanoparticles that had not been previously complexed with lipofectamine. While these results suggest that there is no uptake, since endocytosed nanoparticles would be localized within the highly CpG-responsive endosomal compartment, it does not definitively rule out that possibility. As PBMCs encompass a variety of immune cell populations including B cells, monocytes, and dendritic cells, one method that can be used to investigate the uptake of nanoparticles by immune cells is to perform flow cytometry analysis on PBMCs incubated with fluorescent nanoparticles. Using lipofectamine-complexed nanoparticles as a positive control, we can identify subpopulations of cells that are able to take up nanoparticles in the absence of a transfection reagent. If any cells are identified, we can then follow up with multiplex ELISA analysis of nanoparticle-induced interferon expression to determine if or how the expression pattern varies compared to lipofectamine-complexed nanoparticles. Additionally, fluorescent microscopy could be performed using the same set of fluorescently labeled nanostructures to not only investigate nanoparticle internalization but also visualize the cellular compartments where nanoparticles localize.

Finally, while the work described in this thesis has begun to explore the effects of various nanoparticle design and ligand presentation parameters on immunological recognition, there is much work left to do. One question which can be investigated on a relatively short timescale is: how does spatial distance between CpG overhangs affect the potency of TLR9 activation? As mentioned in the introduction, an analysis of the crystal structure of activated TLR9 in complex with CpG oligos showed that the distance between the two DNA binding pockets is approximately 7– 8nm. This suggests that displaying a pair of CpG overhangs on a wireframe nanoparticle at an inter-CpG distance of ~7nm may facilitate 'optimal' TLR9 dimer binding and activation. However, given that ssDNA is highly flexible on these length scales, it is likely that there is not an 'optimal' distance but rather a range of distances within which TLR9 activation is enhanced. Additionally, if this is true, it suggests that there may be a thresholding effect wherein if the inter-CpG distance is above a certain threshold, they can no longer be engaged by the same TLR9 dimer, resulting in a substantial drop in the magnitude of TLR9 activation. To investigate this question, a set of PB84 variants can be synthesized, each of which is displaying a pair of CpG overhangs at different inter-ligand spacings. Starting from a minimal distance of <7nm, the distance between the overhangs should be incrementally increased up to 50nm, the maximum spacing possible on the PB84. As each overhang is 24 nucleotides long, with the 3'-most CpG motif on the overhang located 5 nucleotides from the 3' end, the closest possible distance between the two CpG motifs is 24nm, which is substantially larger than the estimated distance between binding pockets. Thus, it should be possible to design a set of PB84 variants to probe the maximal range of distances for individual TLR9 dimer binding.

The same set of PB84 constructs could theoretically be used to investigate a different question, which is whether our nanoparticles are able to interact with multiple TLR9 dimers simultaneously. As mentioned previously, the length of PB84 is 50nm, while the furthest CpG motif on each nanoparticle is 13nm away from the nanoparticle. This means that the largest possible inter-CpG distance for two CpG motifs displayed on the same pentagonal bipyramid is 76nm, a distance which, according to measurements of the TLR9 dimer crystal structure, should allow at least two TLR9 dimers to bind to the nanoparticle simultaneously. We can begin to test this hypothesis using the CpG dimer pair constructs described in the previous paragraph. Once the inter-CpG distance is too large for a single TLR9 dimer to bind, we should theoretically see a drop in TLR9 activation as a result. However, if there is either no drop in activation for any of the constructs tested, or if the TLR9 activation ramps back up after a decrease, that would imply that there must be some nanoparticles upon which two different TLR9 dimers are binding, and furthermore that each TLR9 dimer must be interacting with CpG overhangs on two different nanoparticles. One potential way this hypothesis can be tested is by functionalizing the two CpG overhangs on each PB84 with a set of FRET pairs. As the distance between binding pockets for a single TLR9 dimer is known to be ~7nm, which is within the acceptable distance for FRET, if the two overhangs are bound to the same dimer, FRET should occur, and acceptor fluorescence emission can be measured. While FRET may also occur in the absence of binding, the fluorophores would most likely travel in and out of the acceptable range for FRET due to the flexibility of the ssDNA overhang and the random movement of oligos in solution, and it should therefore be possible to differentiate between bound and unbound states. Once the distance between CpG overhangs is too large for a single dimer to bind, TLR9 dimers will either no longer be activated, in which case no FRET should occur, or some proportion of dimers will bind to CpG overhangs on two different nanoparticles and become

activated. In the second scenario, because there will be a mixture of dimers binding to two donor fluorophores, two acceptor fluorophores, and one donor and one acceptor fluorophore, the relative level of acceptor fluorescence emission will be reduced compared to samples in which all TLR9 dimers are bound to correct FRET pairs. This method may be able to allow us to differentiate between TLR9 dimer activation mediated by CpG overhangs on the same nanoparticle or on different nanoparticles and begin to elucidate the effects of inter-CpG spacing on TLR9 activation.

While there are still many questions remaining about how nanoscale organization of ligands and nanoparticle design parameters influence immunological recognition, future work, such as the experiments highlighted above, will contribute to our understanding of how wireframe DNA nanoparticles interface with the immune system and help guide the design of immunomodulatory nanostructures tailored to address specific prophylactic and therapeutic challenges.

Appendix

Design of wireframe DNA NANPs for RNAi

Overview

The versatility of wireframe DNA origami as a platform for nanoscale engineering can be harnessed to enable DNA nanoparticles to have additional functionality besides immunomodulation. While the focus of this thesis has been on DNA sensing pathways within the innate immune system, one natural extension of this work would be to design RNA/DNA nanoparticles, whether by integration of RNA into the scaffold or staples or by the attachment of RNA oligos to the folded nanostructure, to enable the nanoparticle to activate RNA sensing pathways as well. One class of RNAs that has been researched extensively and has demonstrated significant therapeutic potential is a short, noncoding, functional RNA known as small interfering RNA (siRNA). These siRNAs operate within the RNA interference (RNAi) system and regulate gene expression by selectively degrading target mRNA containing a complementary sequence. Additionally, it has been shown that siRNA is capable of activating the endosomal RNA-sensing PRRs TLR3, TLR7, and TLR8, thus operating as both a therapeutic nucleic acid as well as an immune adjuvant. In this section, I describe preliminary research into the development of siRNA/DNA nanostructures for targeted gene silencing as well as enhanced innate immunostimulation. I will introduce the design of DNA nanoparticles capable of modular siRNA attachment, show proof-of-concept characterization of siRNA-DNA nanoparticle synthesis, and close with a brief discussion of future directions for this work.

Small Interfering RNA

Small interfering RNAs are short double-stranded noncoding RNAs that operate within the RNAi pathway, an endogenous biological pathway that induces sequence-specific inhibition of gene

expression by degrading or inhibiting translation of mRNA transcripts[9]. These siRNAs can be delivered exogenously or produced endogenously by Dicer, a ribonuclease-like enzyme which binds to and processes long double-stranded RNAs (dsRNAs) or short hairpin RNAs (shRNAs) into mature siRNAs, which are subsequently loaded into the RNA induced silencing complex (RISC)[9-11]. One strand of the siRNA, termed the "guide" or more commonly "antisense" strand, is selectively incorporated into RISC based on the relative thermodynamic stability of the two ends of the siRNA duplex[12-15]. The complex containing the properly loaded antisense strand then binds to mRNA transcripts based on sequence complementarity, after which Argonaute, a protein within RISC, cleaves the bound mRNA transcript with a RNaseH-like mechanism of action[16]. Importantly, the antisense strand does not dissociate from RISC upon mRNA cleavage. Instead, the entire complex detaches from the cleaved mRNA and is immediately available to bind to other mRNA transcripts, thereby facilitating gene silencing in a catalytic manner[9].

Advantages of siRNA

From the mechanism of action described above, it is clear that siRNAs have several characteristics which give them an advantage over other types of therapeutic molecules. The first advantage is that siRNA-induced gene silencing is highly intolerant of sequence mismatches—with the exception of unique cases which will be elaborated on below, siRNA functionality is significantly reduced if even single base-pair mismatches between the siRNA and mRNA transcript are present, which decreases the risk of off-target gene silencing[12,17,18]. Additionally, because RISC-incorporated siRNAs function catalytically, even low doses of siRNA may be sufficient to produce a significant effect since each molecule of siRNA molecule can cleave numerous copies of the target mRNA transcript. Finally, because the binding of siRNA to its target transcript is entirely

sequence-dependent and the design principles behind siRNA functionality have been well characterized, it is theoretically possible to design a set of siRNAs to efficiently silence any actively transcribed gene.

Limitations of siRNA

While the stringent sequence specificity requirements of siRNA-induced gene silencing suggests that they should be inherently nonfunctional in the absence of their exact mRNA target, siRNA-induced off-target effects have been observed. Several papers have demonstrated that, rather than siRNAs being permissive to sequence mismatches, the "seed region" of the antisense strand, which corresponds to bases 2-8, can exhibit behavior similar to that of microRNAs (miRNAs) if there is perfect complementarity between the seed region and the 3' UTR of mRNA transcripts[17,18,22]. As miRNAs frequently regulate the expression of multiple genes simultaneously, the ability of siRNAs to exhibit miRNA-like behavior poses a challenge to optimizing sequence specificity and minimizing off-target gene silencing effects. Additionally, while sequence mismatches typically eliminate gene silencing efficacy, single- or double-nucleotide mismatches in bases 13-17 have been occasionally found to induce low levels of gene knockdown, particularly if the mismatch is between an adenine and a cytosine[18,22].

The delivery of high concentrations of exogenous siRNA may also result in the saturation of native RNAi machinery, namely Dicer and RISC. As Dicer is required to process siRNA and miRNA precursors into their mature forms, and because both of these RNAi molecules must be complexed with RISC to induce gene silencing, therapeutic siRNAs may directly compete with native

miRNAs for function, potentially leading to dysregulation of critical cellular functions such as differentiation and proliferation[23-27].

DNA Origami for Multiplexed siRNA Delivery

While the drawbacks discussed above are certainly significant, it is critical to note that they are heavily sequence- and concentration-dependent. This may be circumvented by multiplexed delivery of multiple siRNA sequences targeting the same mRNA transcript, which allows the total concentration of siRNAs targeting the transcript to be maintained while lowering the concentration of each individual siRNA sequence, decreasing the magnitude of concentration-dependent off-target effects correspondingly. Additionally, this strategy can be utilized to test the combinatorial effects of multiple siRNAs on the efficacy of target gene silencing and may be helpful for identifying combinations of siRNAs that induce high levels of gene silencing at low concentrations, thereby reducing the risk of Dicer and RISC saturation as well.

Wireframe DNA origami can easily be modified with single-stranded overhangs of any sequence to allow attachment of molecules such as targeting ligands to direct cell- or tissue-type localization, shielding molecules to enhance stability *in vivo*, and small molecule drugs or therapeutically relevant oligos such as siRNAs or miRNAs[41-43]. Additionally, the sequences of each overhang can be distinct and orthogonal, enabling the straightforward multiplexing of many different siRNA sequences with controlled stoichiometry. In this section, I show preliminary work into the design and synthesis of siRNA/DNA nanoparticles for synergistic gene knockdown and immunostimulation.

Design of siRNA/DNA nanoparticles

Wireframe DNA nanoparticles were designed using DAEDALUS, which generates an atomic model of the structure and provides the list of staple sequences needed to fold the user-specific scaffold sequence into the target nanostructure[37]. The nanoparticle chosen for the initial siRNA/DNA nanoparticle design was a pentagonal bipyramid with an edge length of 84 bases, which is sufficiently large to enable the attachment of up to 45 siRNA molecules onto each nanoparticle, where each of the 15 edges can attach a maximum of 3 evenly distributed siRNAs. Staples upon which siRNAs were to be hybridized were extended by 20-23 nucleotides on the 3' ends, exposing a single-stranded 'linker' sequence[10,12]. The pentagonal bipyramid was then folded from a circular scaffold sequence and a specified number of these modified staples through a thermal annealing process, resulting in a folded nanostructure displaying precise copy numbers of single-stranded linker sequences at defined locations. In the same way, to facilitate the attachment of siRNAs onto the DNA nanoparticle, the complement of the linker sequence was appended to the 3' end of the siRNA sense strand. The resulting double-stranded siRNA, with the linker attached, was then hybridized to the pentagonal bipyramid through a second thermal annealing step.

As an initial proof-of-concept to demonstrate that siRNA could be successfully attached to DNA nanoparticles with this method, a pentagonal bipyramid was functionalized with multiple copies of a 23-nucleotide linker sequence, where each sequence has an identical sequence wherein the 3 5'-most bases served as a spacer between the siRNA and the nanoparticle and the remaining 20 bases comprised the region for siRNA linker binding. The siRNA tested was a well-characterized sequence which targeted the gene encoding murine programmed death-ligand 1 (PD-L1), a protein

with immunosuppressive function which is a promising target for cancer immunotherapy. The exact sequence was shared with our lab by collaborators in the Tannous lab at MGH, and the knockdown efficacy was additionally evaluated using algorithms such as the Whitehead Institute's siRNA selection program. The 3' end of the PD-L1 siRNA sense strand was modified to include the 20-base complementary linker sequence, and the 5' end of the sense strand was functionalized with an Atto488 fluorophore. Additionally, a scrambled version of the anti-PD-L1 siRNA was synthesized with an Atto647 fluorophore on the 5' end of the sense strand and an identical 20-base linker on the 3' end as a comparative control.

Pentagonal bipyramid nanoparticles were folded with 0, 10, 20, and 40 overhangs in order to test whether the overhangs would interfere with folding and to additionally determine whether a visible difference between each set of nanoparticles could be identified with a quick agarose gel assay. Each of the four samples was run on a 2% agarose gel (Appendix Figure 1), along with unfolded scaffold as a control. We saw that each nanoparticle formulation exhibited a distinct shift upwards relative to the unfolded scaffold, and furthermore resulted in a clean, tight single band on the gel, suggesting a monodisperse population of folded nanoparticle. This indicates that the addition of any number of overhangs does not result in dimer or aggregate formation, which is critical for maintaining high yield and eliminates the need for additional downstream purification.
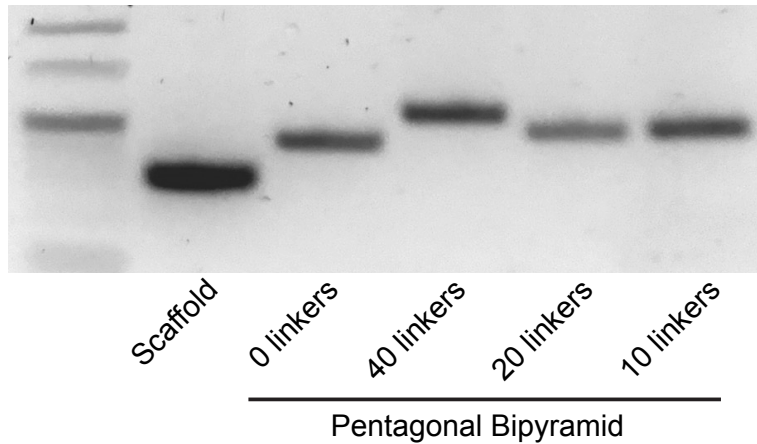
Appendix Figure 1. Agarose gel demonstrating attachment of linkers onto DNA nanoparticles. A gel shift assay was performed on a 2% agarose gel to assess folding. Representations of the samples in each lane are shown immediately below the gel. Lane 1: Unfolded circular scaffold. Lanes 2-5: pentagonal bipyramid with 0, 40, 20, and 10 linkers. The upwards shift of the nanoparticle lanes relative to the unfolded scaffold indicates successful folding.

Next, the scrambled anti-PD-L1 siRNA was annealed to five different pentagonal bipyramid constructs, with 0, 10, 20, 23, and 40 linkers, respectively, and binding efficacy was assessed using a fluorescent colocalization assay. Samples were run on an agarose gel, and a Typhoon imager was used to excite the Atto647 fluorophore conjugated to the siRNA and capture images of the resulting bands (Appendix Figure 2). We found that only the lanes containing siRNA-annealed nanoparticles were visible in the Atto647 channel, demonstrating that each of the 10, 20, 23, and 40 linker pentagonal bipyramid constructs were successfully hybridized with siRNA. In contrast, the siRNA could not bind to the 0 linker pentagonal bipyramid and the unfolded scaffold, showing that the presence of the linker on the nanoparticle is necessary for siRNA binding, as expected. The signal intensities in each of the siRNA-annealed nanoparticle lanes were additionally used to determine siRNA coverage, and we found that the fluorescent intensity of each band was proportional to the number of linkers on the corresponding nanoparticle.
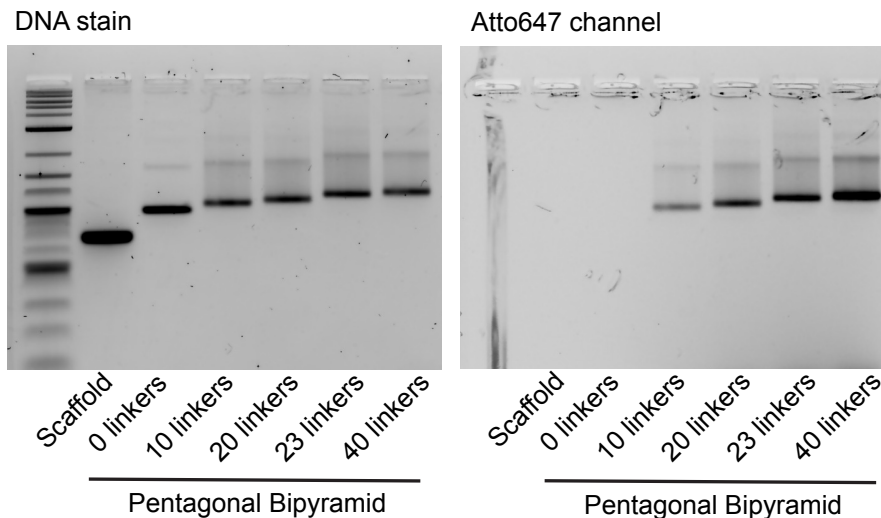
Appendix Figure 2. Attachment of siRNA onto DNA nanoparticles is linker-dependent and proportional to linker copy number. Unfolded scaffold and pentagonal bipyramid constructs with 0, 10, 20, 23, and 40 linkers were run on a 2% agarose gel and imaged in a DNA-stain specific channel (left) or in the Atto647 channel (right). Intensity of the band in the gel on the right is proportional to the concentration of annealed siRNA.

As we have demonstrated that siRNA can be hybridized to DNA nanoparticles in a linker-specific manner and with precise stoichiometry, the attachment of different linker sequences should facilitate the controllable binding of multiple unique nucleic acids, such as multiplexed delivery of siRNA or the simultaneous display of therapeutic siRNAs as well as immunostimulatory CpGs. Additionally, the addressability of DNA nanoparticles can be harnessed to attach additional functional molecules such as targeting or internalization ligands, antigens, or small molecules to these wireframe DNA nanoparticles as well. Future work in this direction can potentially combine the immunomodulatory design principles investigated in earlier chapters of this thesis with the siRNA functionalization method demonstrated above in order to design and synthesize DNA nanoparticles engineered for targeted gene silencing as well as innate immune activation. However, the addition of dsRNA onto these DNA nanoparticles raises the question of whether RNA-sensing

innate immune pathways such as RIG-I or TLR3 may now respond to these nanostructures and how the activation of these pathways may affect the overall nanoparticle-induced immune response, a question which should be investigated carefully in order to fully understand their new immunostimulatory characteristics. These multifunctional wireframe DNA nanostructures have the potential to be a unique class of therapeutic delivery vehicle with intrinsic yet controllable immune adjuvant properties, and the additional knowledge gained from future research in these directions will hopefully provide valuable insights into the capabilities of engineered DNA nanostructures as prophylactics or therapeutics.

Redesign of DNA Nanoparticles for Linker Attachment.

All nanoparticle constructs were designed using DAEDALUS and modified in Tiamat. The high-level staple routing remained unchanged from the initial DAEDALUS output, and in the cases when nick repositioning was required for linker attachment, the locations of nick positions would be shifted the minimal number of bases. Modified nick positions were kept on the same side of the staple as the original nick, and all nicks were located at least 8 bases from the location of the nearest crossover. All staple sequences were exported into Excel, and staples chosen for linker attachment were modified by appending the selected linker sequence to the 3' end of the current staple sequence (Appendix Table 4). Linker sequences were designed following published protocols to ensure efficient Dicer recognition and cleavage as well as optimized loading into RISC[10,12]. Modified staples with the appended linker sequence were then BLASTed against the scaffold and all other staples to minimize the possibility of unexpected crosstalk. Staples were purchased through Integrated DNA Technologies (IDT).

siRNA Synthesis and Annealing.

siRNA oligos were designed using specialized algorithms such as ThermoFisher's RNAi Designer or the Whitehead Institute's siRNA Selection Program and purchased through IDT with full 2'OMe modifications or synthesized in-house using a Dr. Oligo synthesizer following the recommended protocol and purified using a size exclusion column with a Waters HPLC (Appendix Table 4). 2'OMe and 2'F phosphoramidites for in-house synthesis of siRNA were ordered from Glen Research. Purity of synthesized oligos was characterized with a 12-18% denaturing polyacrylamide gel. The sense and antisense strands of the siRNA oligos were incubated for 30 minutes at 37C and cooled to 25C over the course of 4 hours. The annealing temperature for each distinct set of siRNA sequences was determined using IDT's Oligo Analyzer or an equivalent software.

siRNA/DNA Nanoparticle Production.

Nanoparticles were folded from circular bacterially-produced scaffold ssDNA mixed with a 5X excess of ssDNA staples in a folding buffer consisting of 1X TAE, 12mM MgCl$_2$, as well as nuclease-free water. The folding mixture was transferred into 96-well plates in which each well was filled to a maximum of 95uL, and the nanoparticle was folded through a thermal annealing process during which the temperature was initially raised to 95C and gradually ramped down to 25C over a period of 13 hours as described in previous literature[37]. Folded nanoparticles were transferred into Amicon Ultra 100kDa MWCO centrifugal filters and centrifuged at 1000x*g* for 30 minutes at room temperature for up to 5 rounds in order to remove excess staples. Flow-through was discarded and fresh PBS was added to the filter in between each round of centrifugation. Purified nanoparticles were mixed with a 5X excess of double-stranded siRNA with the

complementary linker sequence and hybridized in a final thermal annealing step, in which the sample is initially heated to 65C and cooled to 25C over a period of 8 hours. The annealed siRNA/DNA nanoparticle is purified away from excess siRNA using the spin purification protocol detailed above and concentrated to the desired level. The final siRNA/DNA nanoparticles were characterized using a 1.5%-2% agarose gel made with 1X TAE and 12 mM $MgCl_2$ and visualized with SybrSafe or SybrGold. 100ng of each nanoparticle sample tested was mixed with gel loading buffer and water to obtain a final concentration of 1X gel loading buffer. The gel was placed in a gel running chamber containing 1X TAE and 12mM $MgCl_2$, and after sample loading, the gel was run at 65V for 90-120 minutes at 4C. Images were captured using a Typhoon.

REFERENCES

1   Strebhardt, K. & Ullrich, A. Paul Ehrlich's magic bullet concept: 100 years of progress. *Nat Rev Cancer* **8**, 473-480, doi:10.1038/nrc2394 (2008).

2   Korsmeyer, R. Critical questions in development of targeted nanoparticle therapeutics. *Regenerative Biomaterials* **3**, 143-147, doi:10.1093/rb/rbw011 (2016).

3   Shigdar, S. *et al.* Aptamer Therapeutics: The 21st Century's Magic Bullet of Nanomedicine. *The Open Conference Proceedings Journal* **1**, 118-124, doi:10.2174/22102892010010100118 (2010).

4   Hendricks, W. P. D., Yang, J., Sur, S. & Zhou, S. B. Formulating the magic bullet: barriers to clinical translation of nanoparticle cancer gene therapy. *Nanomedicine* **9**, 1121-1124, doi:10.2217/nnm.14.63 (2014).

5   Liu, D. X., Mori, A. & Huang, L. ROLE OF LIPOSOME SIZE AND RES BLOCKADE IN CONTROLLING BIODISTRIBUTION AND TUMOR UPTAKE OF GM1-CONTAINING LIPOSOMES. *Biochimica Et Biophysica Acta* **1104**, 95-101, doi:10.1016/0005-2736(92)90136-a (1992).

6   Levchenko, T. S., Rammohan, R., Lukyanov, A. N., Whiteman, K. R. & Torchilin, V. P. Liposome clearance in mice: the effect of a separate and combined presence of surface charge and polymer coating. *International Journal of Pharmaceutics* **240**, 95-102, doi:10.1016/s0378-5173(02)00129-1 (2002).

7   Li, S. D. & Huang, L. Pharmacokinetics and biodistribution of nanoparticles. *Molecular Pharmaceutics* **5**, 496-504, doi:10.1021/mp800049w (2008).

8   Wilhelm, S. *et al.* Analysis of nanoparticle delivery to tumours. *Nature Reviews Materials* **1**, 12, doi:10.1038/natrevmats.2016.14 (2016).

9   Hammond, S. M. Dicing and slicing - The core machinery of the RNA interference pathway. *Febs Letters* **579**, 5822-5829, doi:10.1016/j.febslet.2005.08.079 (2005).

10   Vermeulen, A. *et al.* The contributions of dsRNA structure to Dicer specificity and efficiency. *Rna-a Publication of the Rna Society* **11**, 674-682, doi:10.1261/rna.7272305 (2005).

11   Bartel, D. P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297, doi:10.1016/s0092-8674(04)00045-5 (2004).

12   Birmingham, A. *et al.* A protocol for designing siRNAs with high functionality and specificity. *Nature Protocols* **2**, 2068-2078, doi:10.1038/nprot.2007.278 (2007).

13   Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21-and 22-nucleotide RNAs. *Genes & Development* **15**, 188-200, doi:10.1101/gad.862301 (2001).

14   Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias (vol 115, pg 209, 2003). *Cell* **115**, 505-505, doi:10.1016/s0092-8674(03)00893-6 (2003).

15   Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208, doi:10.1016/s0092-8674(03)00759-1 (2003).

16   Parker, J. S., Roe, S. M. & Barford, D. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *Embo Journal* **23**, 4727-4737, doi:10.1038/sj.emboj.7600488 (2004).

17   Birmingham, A. *et al.* 3 ' UTR seed matches, but not overall identity, are associated with RNAi off-targets (vol 3, pg 199, 2006). *Nature Methods* **4**, 533-533, doi:10.1038/nmeth0607-533 (2007).

18   Dahlgren, C. *et al.* Analysis of siRNA specificity on targets with double-nucleotide mismatches. *Nucleic Acids Research* **36**, 7, doi:10.1093/nar/gkn190 (2008).

19   Chakraborty, C., Sharma, A. R., Sharma, G., Doss, C. G. P. & Lee, S. S. Therapeutic miRNA and siRNA: Moving from Bench to Clinic as Next Generation Medicine. *Molecular Therapy-Nucleic Acids* **8**, 132-143, doi:10.1016/j.omtn.2017.06.005 (2017).

20   Xu, C. F. & Wang, J. Delivery systems for siRNA drug development in cancer therapy. *Asian Journal of Pharmaceutical Sciences* **10**, 1-12, doi:10.1016/j.ajps.2014.08.011 (2015).

21   *Clinicaltrials.gov.*

22   Du, Q., Thonberg, H., Wang, J., Wahlestedt, C. & Liang, Z. C. A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Research* **33**, 1671-1677, doi:10.1093/nar/gki312 (2005).

23   Uprichard, S. L. The therapeutic potential of RNA interference. *Febs Letters* **579**, 5996-6007, doi:10.1016/j.febslet.2005.08.004 (2005).

24   Kanasty, R. L., Whitehead, K. A., Vegas, A. J. & Anderson, D. G. Action and Reaction: The Biological Response to siRNA and Its Delivery Vehicles. *Molecular Therapy* **20**, 513-524, doi:10.1038/mt.2011.294 (2012).

25   Khan, A. A. *et al.* Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nature Biotechnology* **27**, 549-U592, doi:10.1038/nbt.1543 (2009).

26   Castanotto, D. *et al.* Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Research* **35**, 5154-5164, doi:10.1093/nar/gkm543 (2007).

27   Grimm, D. *et al.* Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**, 537-541, doi:10.1038/nature04791 (2006).

28   Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chemistry* **3**, 103-113, doi:10.1038/nchem.957 (2011).

29   Elbaz, J. *et al.* DNA computing circuits using libraries of DNAzyme subunits. *Nature Nanotechnology* **5**, 417-422, doi:10.1038/nnano.2010.88 (2010).

30   Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585-1588, doi:10.1126/science.1132493 (2006).

31   Qian, L. & Winfree, E. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **332**, 1196-1201, doi:10.1126/science.1200520 (2011).

32   Zhang, D. Y. & Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *Journal of the American Chemical Society* **131**, 17303-17314, doi:10.1021/ja906987s (2009).

33   Fern, J. & Schulman, R. Design and Characterization of DNA Strand-Displacement Circuits in Serum-Supplemented Cell Medium. *Acs Synthetic Biology* **6**, 1774-1783, doi:10.1021/acssynbio.7b00105 (2017).

34   Panyutin, I. G. & Hsieh, P. THE KINETICS OF SPONTANEOUS DNA BRANCH MIGRATION. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 2021-2025, doi:10.1073/pnas.91.6.2021 (1994).

35   Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302, doi:10.1038/nature04586 (2006).

36   Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes (vol 459, pg 414, 2009). *Nature* **459**, 1154-1154, doi:10.1038/nature08165 (2009).

37   Veneziano, R. *et al.* Designer nanoscale DNA assemblies programmed from the top down. *Science* **352**, 8, doi:10.1126/science.aaf4388 (2016).

38   Monia, B. P., Johnston, J. F., Sasmor, H. & Cummins, L. L. Nuclease resistance and antisense activity of modified oligonucleotides targeted to Ha-ras. *Journal of Biological Chemistry* **271**, 14533-14540, doi:10.1074/jbc.271.24.14533 (1996).

39   Geary, R. S., Norris, D., Yu, R. & Bennett, C. F. Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Advanced Drug Delivery Reviews* **87**, 46-51, doi:10.1016/j.addr.2015.01.008 (2015).

40   Conway, J. W., McLaughlin, C. K., Castor, K. J. & Sleiman, H. DNA nanostructure serum stability: greater than the sum of its parts. *Chemical Communications* **49**, 1172-1174, doi:10.1039/c2cc37556g (2013).

41   Douglas, S. M., Bachelet, I. & Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **335**, 831-834, doi:10.1126/science.1214081 (2012).

42   Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459**, 73-U75, doi:10.1038/nature07971 (2009).

43   Li, S. P. *et al.* A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. *Nature Biotechnology* **36**, 258-+, doi:10.1038/nbt.4071 (2018).

44   Lee, H. *et al.* Molecularly self-assembled nucleic acid nanoparticles for targeted in vivo siRNA delivery. *Nature Nanotechnology* **7**, 389-393, doi:10.1038/nnano.2012.73 (2012).

45   Bujold, K. E., Hsu, J. C. C. & Sleiman, H. F. Optimized DNA "Nanosuitcases" for Encapsulation and Conditional Release of siRNA. *Journal of the American Chemical Society* **138**, 14030-14038, doi:10.1021/jacs.6b08369 (2016).

46   Afonin, K. A. *et al.* Triggering of RNA Interference with RNA-RNA, RNA-DNA, and DNA-RNA Nanoparticles. *Acs Nano* **9**, 251-259, doi:10.1021/nn504508s (2015).

47   Høiberg, H. C. *et al.* An RNA Origami Octahedron with Intrinsic siRNAs for Potent Gene Knockdown. *Biotechnology Journal*, doi: 10.1002/biot.201700634 (2018).

48   Kortylewski, M. *et al.* In vivo delivery of siRNA to immune cells by conjugation to a TLR9 agonist enhances antitumor immune responses. *Nature Biotechnology* **27**, 925-U988, doi:10.1038/nbt.1564 (2009).

49   Ovcharenko, D. High throughput in vitro RNAi screening: From cell lines to primary cells. *Faseb Journal* **19**, A262-A262 (2005).

50   Chasteen, L., Ayriss, J., Pavlik, P. & Bradbury, A. R. M. Eliminating helper phage from phage display. *Nucleic Acids Research* **34**, 11, doi:10.1093/nar/gkl772 (2006).

Appendix Table 1. Scaffold sequence used for all NANPs. CpG dinucleotides are displayed in red, CpG motifs which satisfy the minimal hTLR9 activation sequence requirements are bolded.

Name	Sequence
PB84/ ICO42	GAGCGCAACGCAATTAATGTGCGCCCTGTAGCGGCGCATTAAGCGCGGC GGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAG CGCCCGCTCCTTT**CGCTTTCTTCCCTTCCTTTCT**CGCCA**CGT**TCGCCGG CTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAG TGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCAC GTAGTGGGCCATCGCCCTGATAGACGGTTTTT**CGCCCTTTGACGTTGGA GTCCACGT**TCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCA ACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGG CCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAC AACCGGGGTACATATGATTGGGGTCTGACGCTCAGTGGAACGAAAACTC ACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA TCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTC AGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGA TACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAG CCAGCCGGAAGGGCCGAGCGCAtAAGTGGTCCTGCAACTTTATCCGCCTC CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGT**TCGCCA GTTAATAGTTTGCGCAACGT**TGTTGCCATTGCTACAGGCATCGTGGTGT CACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTC GGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCAT GGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGAT GCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGT ATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGC GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA CCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTT TCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTAT TGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA GTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA AAATAGGCGTATCACGAGGCCCTTTCGTCGAATTCGTCGTCGTCCCCTCA AACTCTTGGGTGGAGAGGCTATTCGTTTAAGGTCACATCGCATGTAATTT ACTTATTCTCTGTTGTTGAGCCACCCGGGCGCCAGATTTTGTTTAAAGCTT TGTCTCTTAGTTTGTATAGACAGATTCAGAGTGCAAGGTTTCGTTCGCTCG TACCTGGTTTTCCCTGGTTCTTCACAGATAGGATTTGACTTTCTACAACAC TTATGCGGCTTCCTACCCGTTTGAAGGCCGATACAGGTGCTGCGCAAAAT GCGGGCGAACATAGAGTATCAAAACAACGCCTTCTAATCTAGGAATATA GGGAAGATACGTATTTGCTACCATGCTTTCTTGGGTCATTAACGACCAAC CTCTTTTCTTTTAAAGTAGGATTGCACAATGAATGAATACACGTGGTCCG

ATAACTGACCAAGTAACATGGTTATCACTaGATGTCCGCCAGACGTGTGC
AAACCAACCCGGGAGTTACGTCACTAATCCTTCGCTACGTCGTGAAGATA
TTTACTTGTGAATATCGAGGGTAATAAGATAATAGACTGTGACTAGTATT
GCCAGACTGTCGCTACCTGCAACACATAACTATCCTGAGGTTACTGCATA
GTACTGATTACACCCGAGTCAAAATTTCTAACTTCTAACATGTACCTAGT
AACCAGCTCAATAATTATGTCAGAATATAGCTCTGGGAACCCTCGGACAA
TTATGATACACGGTATTAATATCTTGCTTGCGTTAGCCACTTCTCATCTTT
GGATACCGATTCTATTTTGCATAGCAGTTCCTTTTACACATATAAGAATTT
CGCCATAGGTATGCTGCAG

Appendix Table 2. Staple sequences for all PB84 variants. Sequences are given in 5' to 3' orientation, and modified staple sets for each variant are provided in their own sub-table. Not all staples are listed for each variant. For unlisted staples (check staple number), the original staple from the **PB84, nonmodified** table was used.

PB84, nonmodified. CpG dinucleotides are displayed in red, CpG motifs which satisfy the minimal hTLR9 activation sequence requirements are bolded.

Name	Sequence
PB84_staple2	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGG
PB84_staple3	TGCGCTCGCCGCTACAGGGCGCACATTAAGCAGGACCACTTA
PB84_staple4	GCTGGTTTACACCCGCCGCGCTTAATGCGGCCCTTCCGGCTG
PB84_staple5	GTCACGCTGCGTTTTTCGTAACCACCATTGCTGATAATTTTTATCTGGAGCCCACGACGGGGATTTTTGTCAGGCAAC
PB84_staple6	GTGTAGCGAGCCGGCGAACGTGGCGAGAAAGGGCGCTGGCAA
PB84_staple7	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATAT
PB84_staple8	TCTGACATGTCCGAGGGTTCCCAGGAAAGGAG
PB84_staple9	CATTGGTAACTTTTTTGTCAGACCAAGATTTAGAGCTTTTTTTGACGGGGAA
PB84_staple10	TACTTTAGGGAACCCTAAAGGGAGCCCCCGTTTACTCATATA
PB84_staple11	TTCATTTTGGTGCCGTAAAGCACTAAATCATTGATTTAAAAC
PB84_staple12	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGA
PB84_staple13	GTTGAGTGTTGTTTTTTTCCAGTTTGCACTACGTGAATTTTTCCATCACCCT
PB84_staple14	CGATGGCCGAACAAGAGTCCACTATTAAACCGTCTATCAGGG
PB84_staple15	GCGAAAAAGAACGTGGACTCCAACAATCAGTA
PB84_staple16	CTATGCAGATTTTGACTCGGGTGTGTCAAAGG
PB84_staple17	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGACC
PB84_staple18	TATAAATCTAGGCCGAAATCGGCATGTTGTAG
PB84_staple19	AAAGTCAATAGGAAGCCGCATAAGAAATCCCT
PB84_staple20	AAGATCCTTTTTTTTTTGATAATCTCTTCGCGTTAAATTTTTTTTTGTTAA

PB84_staple21	GGTTGTAAATGACCAAAATCCCTTAACGTTCATATGTACCCC
PB84_staple22	GACCCCAAGAGTTTTCGTTCCACTCCCGTATT
PB84_staple23	GACGCCGGTGTGGCGCGGTATTATGAGCGTCA
PB84_staple24	ACTGATTAAGTATGGATGAA
PB84_staple25	GGTGCCTCCGAAATAGACAGATCGAAAAGCAT
PB84_staple26	CTTACGGATACTCACCAGTCACAGCTGAGATA
PB84_staple27	GTTATCTAGGTGAGCGTGGGTCTCGCGGTCTCCCGTATCGTA
PB84_staple28	GGTAAGCCATCATTGCAGCACTGGAACCGGAG
PB84_staple29	CTGAATGACGCCTTGATCGTTGGGGGCCAGAT
PB84_staple30	TCCCGGCAACATTTTTATTAATAGACTACCTATGGCGTTTTTAAATTCTTAT
PB84_staple31	AATGTGCGGGCGAACTACTTACTCTAGCTCACTTTTCGGGGA
PB84_staple32	TGTTTATTACGTTGCGCAAACTATTAACTCGGAACCCCTATT
PB84_staple33	TCAAATATATGCCTGTAGCAATGGCAACATTTCTAAATACAT
PB84_staple34	AACGACGAGCGTTTTTTGACACCACGGTATCCGCTCATTTTTTGAGACAATA
PB84_staple35	AGCCATACCATCATGTAACT
PB84_staple36	CATTTCCGTGTTTTTTCGCCCTTATTTTTTTTGCACATTTTTACATGGGGGA
PB84_staple37	CATTTTGCAGGACCGAAGGAGCTAACCGCCCCTTTTTTGCGG
PB84_staple38	CTCACCCACTTACTTCTGACAACGATCGGCTTCCTGTTTTTG
PB84_staple39	AAGTAAAAGAGTGATAACACTGCGGCCAAGAAACGCTGGTGA
PB84_staple40	CTCGGTCGCCGTTTTTCATACACTATATGCAGTGCTGTTTTTCCATAACCAT
PB84_staple41	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTA
PB84_staple42	GCAAGAGCAAAGTTCTGCTA
PB84_staple43	GATGCTGAAGATTTTTTCAGTTGGGTTCCAATGATGATTTTTGCACTTTTAA

PB84_staple44	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAA
PB84_staple45	CTTGAGAGCTGGATCTCAACAGCGTGAATCTG
PB84_staple46	TCTATACAACGAAACCTTGCACTCGTAAGATC
PB84_staple47	GAGTATTCAAACCCTGATAA
PB84_staple48	AAGAGTATATGCTTCAATAATATTACATGCGA
PB84_staple49	TGTGACCTAGAGAATAAGTAAATTGAAAAAGG
PB84_staple50	TTATCGGACCATTTTTCGTGTATTCAGGTTTCTTAGATTTTTCGTCAGGTGG
PB84_staple51	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAAT
PB84_staple52	TTGGTCGTCGTGATACGCCTATTTTTATAAAAAGAAAAGAGG
PB84_staple53	AAGCATGGACGAATTCGACGAAAGGGCCTTAATGACCCAAGA
PB84_staple54	CAAAGCTTTAATTTTTACAAAATCTGCAAGAGTTTGATTTTTGGGGACGACGTAGCAAATACGTTTTTTATCTTCCCT
PB84_staple55	TCTCCACCGCGCCCGGGTGGCTCAACAACTAAACGAATAGCC
PB84_staple56	AACTAAGAGAGTACGAGCGA
PB84_staple57	TGCGCAGCACCTTTTTTGTATCGGCCTGAAGAACCAGTTTTTGGAAAACCAG
PB84_staple58	ATCCTATCTGTTCAAACGGG
PB84_staple59	GCCCGCATTTATATTCCTAG
PB84_staple60	CTATGTTCATTAGAAGGCGTTGTTATTACCCT
PB84_staple61	CGATATTCCACAGTCTATTATCTTTTGATACT
PB84_staple62	ACTGCTATATAACCATGTTACTTGGTCAGATGTGTAAAAGGA
PB84_staple63	GGTATCCAACGTCTGGCGGACATCTAGTGGCAAAATAGAATC
PB84_staple64	GCTAACGCAACTCCCGGGTTGGTTTGCACAAGATGAGAAGTG
PB84_staple65	GTAGCGAAGGATTTTTTTAGTGACGTAAGCAAGATATTTTTTAATACCGTG
PB84_staple66	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATATC

PB84_staple67	CTGGTTACTAGTTTTTGTACATGTTAATAGTTATGTGTTTTTTTGCAGGTAG
PB84_staple68	TAACCTCAGGGAAGTTAGAA
PB84_staple69	AATTATTGAGTATCATAATT

PB84, CpG Motif Test.

Name	Sequence
PB84_CpG_motiftest_staple2v1	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGGTGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple7v1	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple12v1	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGATGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple17v1	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGACCTGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple28v1	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple41v1	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTATGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple44v1	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAATGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple49v1	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple51v1	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATTGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple66v1	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATATCTGGCATTGTACCATTCTAAGGCTA
PB84_CpG_motiftest_staple2v2	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGGTGGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftest_staple7v2	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTGGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftest_staple12v2	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGATGGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftest_staple17v2	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGACCTGGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftest_staple28v2	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTGGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftest_staple41v2	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTATGGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftest_staple44v2	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAATGGCATTGTACCATTCTTTGTCGTT

PB84_CpG_motiftes t_staple49v2	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTGGCATTGT ACCATTCTTTGTCGTT
PB84_CpG_motiftes t_staple51v2	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATT GGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftes t_staple66v2	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATAT CTGGCATTGTACCATTCTTTGTCGTT
PB84_CpG_motiftes t_staple2v3	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCG GTGGCATTGTTGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple7v3	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTGGCATTGT TGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple12v3	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGA TGGCATTGTTGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple17v3	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGAC CTGGCATTGTTGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple28v3	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTGGCATTGT TGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple41v3	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGT ATGGCATTGTTGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple44v3	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAA TGGCATTGTTGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple49v3	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTGGCATTGT TGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple51v3	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATT GGCATTGTTGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple66v3	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATAT CTGGCATTGTTGTCGTTTTAGGCTA
PB84_CpG_motiftes t_staple2v4	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCG GTCGTCGTTTTCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple7v4	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTCGTCGTTT TCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple12v4	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGA TCGTCGTTTTCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple17v4	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGAC CTCGTCGTTTTCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple28v4	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTCGTCGTTTT CCATTCTAAGGCTA
PB84_CpG_motiftes t_staple41v4	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGT ATCGTCGTTTTCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple44v4	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAA TCGTCGTTTTCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple49v4	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTCGTCGTTT TCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple51v4	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATT CGTCGTTTTCCATTCTAAGGCTA
PB84_CpG_motiftes t_staple66v4	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATAT CTCGTCGTTTTCCATTCTAAGGCTA

PB84_CpG_motiftest_staple2v5	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGGTGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple7v5	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple12v5	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGATGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple17v5	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGACCTGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple28v5	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple41v5	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTATGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple44v5	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAATGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple49v5	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple51v5	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATTGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple66v5	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATATCTGGCATTGTTGTCGTTTTGTCGTT
PB84_CpG_motiftest_staple2v6	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGGTCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple7v6	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple12v6	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGATCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple17v6	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGACCTCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple28v6	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple41v6	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTATCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple44v6	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAATCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple49v6	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple51v6	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATTCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple66v6	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATATCTCGTCGTTTTGTCGTTTTAGGCTA
PB84_CpG_motiftest_staple2v7	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGGTCGTCGTTTTCCATTCTTGTCGTT
PB84_CpG_motiftest_staple7v7	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTCGTCGTTTTCCATTCTTGTCGTT
PB84_CpG_motiftest_staple12v7	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGATCGTCGTTTTCCATTCTTGTCGTT

PB84_CpG_motiftes t_staple17v7	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGAC CTCGTCGTTTTCCATTCTTGTCGTT
PB84_CpG_motiftes t_staple28v7	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTCGTCGTTTT CCATTCTTGTCGTT
PB84_CpG_motiftes t_staple41v7	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGT ATCGTCGTTTTCCATTCTTGTCGTT
PB84_CpG_motiftes t_staple44v7	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAA TCGTCGTTTTCCATTCTTGTCGTT
PB84_CpG_motiftes t_staple49v7	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTCGTCGTTT TCCATTCTTGTCGTT
PB84_CpG_motiftes t_staple51v7	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATT CGTCGTTTTCCATTCTTGTCGTT
PB84_CpG_motiftes t_staple66v7	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATAT CTCGTCGTTTTCCATTCTTGTCGTT
PB84_CpG_motiftes t_staple2v8	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCG GTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple7v8	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTCGTCGTTT TGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple12v8	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGA TCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple17v8	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGAC CTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple28v8	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTCGTCGTTTT GTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple41v8	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGT ATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple44v8	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAA TCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple49v8	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTCGTCGTTT TGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple51v8	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATT CGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_motiftes t_staple66v8	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATAT CTCGTCGTTTTGTCGTTTTGTCGTT

PB84, CpG-containing, Outward-facing.

Name	Sequence
PB84_CpG_out _staple2	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGGTC GTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out _staple3	TGCGCTCGCCGCTACAGGGCGCACATTAAGCAGGACCACTTATC GTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out _staple4	GCTGGTTTACACCCGCCGCGCTTAATGCGGCCCTTCCGGCTGTC GTCGTTTTGTCGTTTTGTCGTT

PB84_CpG_out_staple6	GTGTAGCGAGCCGGCGAACGTGGCGAGAAAGGGCGCTGGCAATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple7	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple8	TCTGACATGTCCGAGGGTTCCCAGGAAAGGAGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple10	TACTTTAGGGAACCCTAAAGGGAGCCCCCGTTTACTCATATATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple11	TTCATTTTGGTGCCGTAAAGCACTAAATCATTGATTTAAAACTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple12	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple14	CGATGGCCGAACAAGAGTCCACTATTAAACCGTCTATCAGGGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple15	GCGAAAAAGAACGTGGACTCCAACAATCAGTATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple17	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGACCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple18	TATAAATCTAGGCCGAAATCGGCATGTTGTAGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple19	AAAGTCAATAGGAAGCCGCATAAGAAATCCCTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple21	GGTTGTAAATGACCAAAATCCCTTAACGTTCATATGTACCCCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple22	GACCCCAAGAGTTTTCGTTCCACTCCCGTATTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple25	GGTGCCTCCGAAATAGACAGATCGAAAGCATTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple26	CTTACGGATACTCACCAGTCACAGCTGAGATATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple27	GTTATCTAGGTGAGCGTGGGTCTCGCGGTCTCCCGTATCGTATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple28	GGTAAGCCATCATTGCAGCACTGGAACCGGAGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple29	CTGAATGACGCCTTGATCGTTGGGGGCCAGATTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple31	AATGTGCGGGCGAACTACTTACTCTAGCTCACTTTTCGGGGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple33	TCAAATATATGCCTGTAGCAATGGCAACATTTCTAAATACATTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple37	CATTTTGCAGGACCGAAGGAGCTAACCGCCCCTTTTTTGCGGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple39	AAGTAAAAGAGTGATAACACTGCGGCCAAGAAACGCTGGTGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple41	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTATCGTCGTTTTGTCGTTTTGTCGTT

Name	Sequence
PB84_CpG_out_staple44	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple45	CTTGAGAGCTGGATCTCAACAGCGTGAATCTGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple46	TCTATACAACGAAACCTTGCACTCGTAAGATCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple48	AAGAGTATATGCTTCAATAATATTACATGCGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple49	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple51	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple52	TTGGTCGTCGTGATACGCCTATTTTTATAAAAAGAAAAGAGGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple53	AAGCATGGACGAATTCGACGAAAGGGCCTTAATGACCCAAGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple55	TCTCCACCGCGCCCGGGTGGCTCAACAACTAAACGAATAGCCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple60	CTATGTTCATTAGAAGGCGTTGTTATTACCCTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple61	CGATATTCCACAGTCTATTATCTTTTGATACTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple62	ACTGCTATATAACCATGTTACTTGGTCAGATGTGTAAAAGGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple64	GCTAACGCAACTCCCGGGTTGGTTTGCACAAGATGAGAAGTGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_out_staple66	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATATCTCGTCGTTTTGTCGTTTTGTCGTT

PB84, CpG-free, Outward-facing.

Name	Sequence
PB84_CpGfree_out_staple2	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCGGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple3	TGCGCTCGCCGCTACAGGGCGCACATTAAGCAGGACCACTTACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple4	GCTGGTTTACACCCGCCGCGCTTAATGCGGCCCTTCCGGCTGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple6	GTGTAGCGAGCCGGCGAACGTGGCGAGAAAGGGCGCTGGCAACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple7	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATATCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple8	TCTGACATGTCCGAGGGTTCCCAGGAAAGGAGCCCGATACTGCCATAGACGGCAA

PB84_CpGfree_out_staple10	TACTTTAGGGAACCCTAAAGGGAGCCCCCGTTTACTCATATACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple11	TTCATTTTGGTGCCGTAAAGCACTAAATCATTGATTTAAAACCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple12	TCTAGGTGAATCAAGTTTTTGGGGTCGATAATTTAAAAGGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple14	CGATGGCCGAACAAGAGTCCACTATTAAACCGTCTATCAGGGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple15	GCGAAAAAGAACGTGGACTCCAACAATCAGTACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple17	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGACCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple18	TATAAATCTAGGCCGAAATCGGCATGTTGTAGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple19	AAAGTCAATAGGAAGCCGCATAAGAAATCCCTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple21	GGTTGTAAATGACCAAAATCCCTTAACGTTCATATGTACCCCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple22	GACCCCAAGAGTTTTCGTTCCACTCCCGTATTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple25	GGTGCCTCCGAAATAGACAGATCGAAAAGCATCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple26	CTTACGGATACTCACCAGTCACAGCTGAGATACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple27	GTTATCTAGGTGAGCGTGGGTCTCGCGGTCTCCCGTATCGTACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple28	GGTAAGCCATCATTGCAGCACTGGAACCGGAGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple29	CTGAATGACGCCTTGATCGTTGGGGGCCAGATCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple31	AATGTGCGGGCGAACTACTTACTCTAGCTCACTTTTCGGGGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple33	TCAAATATATGCCTGTAGCAATGGCAACATTTCTAAATACATCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple37	CATTTTGCAGGACCGAAGGAGCTAACCGCCCCTTTTTTGCGGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple39	AAGTAAAAGAGTGATAACACTGCGGCCAAGAAACGCTGGTGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple41	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple44	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple45	CTTGAGAGCTGGATCTCAACAGCGTGAATCTGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple46	TCTATACAACGAAACCTTGCACTCGTAAGATCCCCGATACTGCCATAGACGGCAA

PB84_CpGfree_out_staple48	AAGAGTATATGCTTCAATAATATTACATGCGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple49	TGTGACCTAGAGAATAAGTAAATTGAAAAAGGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple51	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAATCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple52	TTGGTCGTCGTGATACGCCTATTTTTATAAAAAGAAAAGAGGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple53	AAGCATGGACGAATTCGACGAAAGGGCCTTAATGACCCAAGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple55	TCTCCACCGCGCCCGGGTGGCTCAACAACTAAACGAATAGCCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple60	CTATGTTCATTAGAAGGCGTTGTTATTACCCTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple61	CGATATTCCACAGTCTATTATCTTTTGATACTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple62	ACTGCTATATAACCATGTTACTTGGTCAGATGTGTAAAAGGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple64	GCTAACGCAACTCCCGGGTTGGTTTGCACAAGATGAGAAGTGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_out_staple66	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATATCCCCGATACTGCCATAGACGGCAA

PB84, CpG-containing, Inward-facing.

Name	Sequence
PB84_CpG_in_staple2	GGCGGATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple3	ACTTATGCGCTCGCCGCTACAGGGCGCACATTAAGCAGGACCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple4	GGCTGGCTGGTTTACACCCGCCGCGCTTAATGCGGCCCTTCCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple6	GGCAAGTGTAGCGAGCCGGCGAACGTGGCGAGAAAGGGCGCTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple7	GTTCCCAGGAAAGGAGCGGGCGCTAGGAAGGGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple8	AAGAAAGCAGCTATATTCTGACATGTCCGAGGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple10	ATATATACTTTAGGGAACCCTAAAGGGAGCCCCCGTTTACTCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple11	AAAACTTCATTTTGGTGCCGTAAAGCACTAAATCATTGATTTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple12	AAGGATCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAATCGTCGTTTTGTCGTTTTGTCGTT

PB84_CpG_in_staple14	CAGGGCGATGGCCGAACAAGAGTCCACTATTAAACCGTCTATTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple16	ACTCCAACAATCAGTACTATGCAGATTTTGACTCGTCGTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple17	AGACCGAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple18	CGCATAAGAAATCCCTTATAAATCTAGGCCGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple19	AATCGGCATGTTGTAGAAAGTCAATAGGAAGCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple21	ACCCCGGTTGTAAATGACCAAAATCCCTTAACGTTCATATGTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple23	GTTCCACTCCCGTATTGACGCCGGTGTGGCGCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple25	AGTCACAGCTGAGATAGGTGCCTCCGAAATAGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple26	ACAGATCGAAAAGCATCTTACGGATACTCACCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple27	TCGTAGTTATCTAGGTGAGCGTGGGTCTCGCGGTCTCCCGTATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple28	TCGTTGGGGGCCAGATGGTAAGCCATCATTGCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple29	AGCACTGGAACCGGAGCTGAATGACGCCTTGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple31	GGGGAAATGTGCGGGCGAACTACTTACTCTAGCTCACTTTTCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple33	TACATTCAAATATATGCCTGTAGCAATGGCAACATTTCTAAATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple37	TGCGGCATTTTGCAGGACCGAAGGAGCTAACCGCCCCTTTTTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple39	GGTGAAAGTAAAAGAGTGATAACACTGCGGCCAAGAAACGCTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple41	CAGTAAGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple44	CCGAAGAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple45	TTGCACTCGTAAGATCCTTGAGAGCTGGATCTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple46	CAACAGCGTGAATCTGTCTATACAACGAAACCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple48	AGTAAATTGAAAAGGAAGAGTATATGCTTCATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple49	ATAATATTACATGCGATGTGACCTAGAGAATATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple51	GCAATCCTACTTTGGTTAATGTCATGATAATAATTTCATTGTTCGTCGTTTTGTCGTTTTGTCGTT

154

PB84_CpG_in_staple52	AGAGGTTGGTCGTCGTGATACGCCTATTTTTATAAAAAGAAATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple53	CAAGAAAGCATGGACGAATTCGACGAAAGGGCCTTAATGACCTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple55	TAGCCTCTCCACCGCGCCCGGGTGGCTCAACAACTAAACGAATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple60	ATTATCTTTTGATACTCTATGTTCATTAGAAGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple61	GCGTTGTTATTACCCTCGATATTCCACAGTCTTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple62	AAGGAACTGCTATATAACCATGTTACTTGGTCAGATGTGTAATCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple64	AAGTGGCTAACGCAACTCCCGGGTTGGTTTGCACAAGATGAGTCGTCGTTTTGTCGTTTTGTCGTT
PB84_CpG_in_staple66	ATATCTTCACGACCGACAGTCTGGCAATACTAGTACAAGTAATCGTCGTTTTGTCGTTTTGTCGTT

PB84, CpG-free, Inward-facing.

Name	Sequence
PB84_CpGfree_in_staple2	GGCGGATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple3	ACTTATGCGCTCGCCGCTACAGGGCGCACATTAAGCAGGACCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple4	GGCTGGCTGGTTTACACCCGCCGCGCTTAATGCGGCCCTTCCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple6	GGCAAGTGTAGCGAGCCGGCGAACGTGGCGAGAAAGGGCGCTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple7	GTTCCCAGGAAAGGAGCGGGCGCTAGGAAGGGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple8	AAGAAAGCAGCTATATTCTGACATGTCCGAGGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple10	ATATATACTTTAGGGAACCCTAAAGGGAGCCCCCGTTTACTCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple11	AAAACTTCATTTTGGTGCCGTAAAGCACTAAATCATTGATTTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple12	AAGGATCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple14	CAGGGCGATGGCCGAACAAGAGTCCACTATTAAACCGTCTATCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple16	ACTCCAACAATCAGTACTATGCAGATTTTGACCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple17	AGACCGAGATAGGATCAGCTCATTTTTTAACCAAAAAGAATCCCGATACTGCCATAGACGGCAA

PB84_CpGfree_in_staple18	CGCATAAGAAATCCCTTATAAATCTAGGCCGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple19	AATCGGCATGTTGTAGAAAGTCAATAGGAAGCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple21	ACCCCGGTTGTAAATGACCAAAATCCCTTAACGTTCATATGTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple23	GTTCCACTCCCGTATTGACGCCGGTGTGGCGCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple25	AGTCACAGCTGAGATAGGTGCCTCCGAAATAGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple26	ACAGATCGAAAAGCATCTTACGGATACTCACCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple27	TCGTAGTTATCTAGGTGAGCGTGGGTCTCGCGGTCTCCCGTACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple28	TCGTTGGGGGCCAGATGGTAAGCCATCATTGCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple29	AGCACTGGAACCGGAGCTGAATGACGCCTTGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple31	GGGGAAATGTGCGGGCGAACTACTTACTCTAGCTCACTTTTCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple33	TACATTCAAATATATGCCTGTAGCAATGGCAACATTTCTAAACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple37	TGCGGCATTTTGCAGGACCGAAGGAGCTAACCGCCCCTTTTTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple39	GGTGAAAGTAAAAGAGTGATAACACTGCGGCCAAGAAACGCTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple41	CAGTAAGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple44	CCGAAGAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple45	TTGCACTCGTAAGATCCTTGAGAGCTGGATCTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple46	CAACAGCGTGAATCTGTCTATACAACGAAACCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple48	AGTAAATTGAAAAAGGAAGAGTATATGCTTCACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple49	ATAATATTACATGCGATGTGACCTAGAGAATACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple51	GCAATCCTACTTTGGTTAATGTCATGATAATAATTTCATTGTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple52	AGAGGTTGGTCGTCGTGATACGCCTATTTTTATAAAAAGAAACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple53	CAAGAAAGCATGGACGAATTCGACGAAAGGGCCTTAATGACCCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple55	TAGCCTCTCCACCGCGCCCGGGTGGCTCAACAACTAAACGAACCCGATACTGCCATAGACGGCAA

PB84_CpGfree_in_staple60	ATTATCTTTTGATACTCTATGTTCATTAGAAGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple61	GCGTTGTTATTACCCTCGATATTCCACAGTCTCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple62	AAGGAACTGCTATATAACCATGTTACTTGGTCAGATGTGTAACCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple64	AAGTGGCTAACGCAACTCCCGGGTTGGTTTGCACAAGATGAGCCCGATACTGCCATAGACGGCAA
PB84_CpGfree_in_staple66	ATATCTTCACGACCGACAGTCTGGCAATACTAGTACAAGTAACCCGATACTGCCATAGACGGCAA

Appendix Table 3. Staple sequences for all ICO42 variants. Sequences are given in 5' to 3' orientation, and staple sets for each variant are provided in their own sub-table.

ICO42, nonmodified. CpG dinucleotides are displayed in red, CpG motifs which satisfy the minimal hTLR9 activation sequence requirements are bolded.

Name	Sequence
ICO42_staple2	CCCTAAAGTTGCGTTGCGCTCCTGCAGCACACTAAATCGGAA
ICO42_staple3	CCGCTACAGGGTTTTTCGCACATTAAGGAGCCCCCGATTTTTTTTAGAGCTT
ICO42_staple4	CAAGATATAAGATGAGAAGTGGCTATGCGGCTTAAACGCAAG
ICO42_staple5	GAAGGGAAGAATTTTTAGCGAAAGGAGGGCGCTGGCATTTTTAGTGTAGCGGCGTAACCACCATTTTTCACCCGCCGC
ICO42_staple6	TTAGAAATTGGTTACTAGGTACATCTGCGTCACGGTTAGAAG
ICO42_staple7	TCTATTATACAGTCTGGCAATACTCGCTAGCGGGAGTCACAG
ICO42_staple8	CGAGAAAGGACGGGGAAAGCCGGCTCAGGGTCTAGAACGTGG
ICO42_staple9	TGGGGTCGAGGTTTTTTGCCGTAAAGTACCTATGGCGTTTTTAAATTCTTATGCTTTAAACAATTTTTAATCTGGCGC
ICO42_staple10	TCATTTTTCATCACCCTAATCAAGTTTTTTTGTTAAATCAGC
ICO42_staple11	GCGATGGCCCATTTTTCTACGTGAACTAACCAATAGGTTTTTCCGAAATCGG
ICO42_staple12	AGATAGGGTTGTTTTTAGTGTTGTTCCAAGAGTCCACTTTTTTATTAAAGAAAACGTCAAAGGTTTTTGCGAAAAACC
ICO42_staple13	GAAGGATTAAGTAAATATCTTCACACTCCCGTGGGACGTAGC
ICO42_staple14	TCAACAGCGAGTGGGTTACATCGATGGAACAGTTACTGGATC
ICO42_staple15	ATAGACCGCAAAATCCCTTATAAATGAGTTAACGTCAAAAGA
ICO42_staple16	GGCAACTATGGTTTTTATGAACGAAAGTTGTAATTCGTTTTTCGTTAAATTT
ICO42_staple17	GAGATAGGACCCCAATCATATGTACCCCGTAGACAGATCGCT
ICO42_staple18	TTTCGTTCCACTTTTTTGAGCGTCAGTGCCTCACTGATTTTTTTAAGCATTG
ICO42_staple19	TTTAGATTGATTTTTTTTAAAACTTCTAAAAGGATCTTTTTTAGGTGAAGATAATCTCATGACTTTTTCAAAATCCCT
ICO42_staple20	ATGAGCACCGCCCCGAAGAACGTTTTGATCCTTTTTCCAATG
ICO42_staple21	GGCTGGCTACCACTTATGCGCTCGTAATTATTTTGCCCTTCC

158

ICO42_staple22	ATATATACGTAACTGTCAGACCAAGATGGGGCCAGTTTACTC
ICO42_staple23	CAACAGAGTATCTACACGACGGGGAGTCACCGGGTGGCTCAA
ICO42_staple24	TAAGCCCTCCCTTTTTGTATCGTAGTAATAAGTAAATTTTTTTACATGCGAT
ICO42_staple25	ACGAATAGCCTTTTTTCTCCACCCAAACCGGAGCTGATTTTTATGAAGCCATGCGGTATCATTTTTTTGCAGCACTGG
ICO42_staple26	CGTGACACGAGCCGGTGAGCGTGGGTCTCACCAAACGACGAG
ICO42_staple27	GGAGGCGGATATTTTTAAGTTGCAGGGGTTTATTGCTTTTTTGATAAATCTGCACGATGCCTGTTTTTTAGCAATGGC
ICO42_staple28	CGGGCAAGCGCGGTATTATCCCGTTGGATTAGACATTGACGC
ICO42_staple29	CGCAAACTATTTTTTTAACTGGCGAATAGCTTCCCGGTTTTTCAACAATTAA
ICO42_staple30	CAGTAAGAAAAAGCATCTTACGGATACTCCTACTTGGCATGA
ICO42_staple31	GCTTTTTTATCGGAGGACCGAAGGCGTTGAACAAAGCTAACC
ICO42_staple32	CGACGACGAACTCGCCTTGATCGTTGGGAGAGTTTGAGGGGA
ICO42_staple33	GCACAACATGGTTTTTGGGATCATGTAAT**TCGACGAATTTTTAGGGCCTCG**TTAATAATGGTTTTTTTTCTTAGACGT
ICO42_staple34	TTTTCGGGGAATTTTTATGTGCGCGGGCCAACTTACTTTTTTTCTGACAACG
ICO42_staple35	TTATTTTTACCATGAGTGATAACACTGCGAACCCCTATTTGT
ICO42_staple36	GAATTATGCAGTTTTTTGCTGCCATACTAAATACATTTTTTTTCAAATATGTA
ICO42_staple37	TTCAATAATATTTTTTTGAAAAAGGACTTATTCCCTTTTTTTTTTTGCGGCATTGAGTACTCATTTTTCCAGTCACAG
ICO42_staple38	TGACTTGGTTTTGCCTTCCTGTTTTTGCTACTATTCTCAGAA
ICO42_staple39	TTTTAAAGTTCTTTTTTGCTATGTGGAGCAACTCGGTTTTTTCGCCGCATACCACCCAGAAACTTTTTGCTGGTGAAA
ICO42_staple40	CTGAAGATCAGTTTTTTTGGGTGCACGGTAAGATCCTTTTTTTGAGAGTTTT
ICO42_staple41	TAGTGATAGTTTGCACACGTCTGGAGATGGTAAACGGACATC
ICO42_staple42	GTGTCGCCAGAGTATGAGTATTCACCACGTCGGAACATTTCC
ICO42_staple43	ATAAATGCTCCGCTCATGAGACAATGGTAAAGCATAACCCTG
ICO42_staple44	TTGATACTATTCCTAGATTAGAAGGGCACCAGGTGCGTTGTT

ICO42_staple45	TGTCATGAGATACGCCTATTTTTATGTAGAGTGTTAGGTTAA
ICO42_staple46	GCGAACGAAGAACCAGGGAAAACCCTTAAGTGACAGGTACGA
ICO42_staple47	ACTGCTATCTATACAAACTAAGAGACAAAATGTGTAAAAGGA
ICO42_staple48	AAAGTCAAATCTTTTTCTATCTGTGAAACCTTGCACTTTTTTCTGAATCTGT
ICO42_staple49	GCAAAATAGAATTTTTTCGGTATCCATAATACCGTGTTTTTTATCATAATTGAAACGGGTAGGTTTTTAAGCCGCATA
ICO42_staple50	AGAGCTATGCAGCACCTGTATCGGCCTTCTCCGAGGGTTCCC
ICO42_staple51	GCAAATACGTATTTTTTCTTCCCTATCTATGTTCGCCTTTTTCGCATTTTGC
ICO42_staple52	ATTCTGACATATTTTTATTATTGAGCTTTGAC**TCGGGTTTTTTGTAATCAGTTTGGTCGT**TAATTTTTTGACCCAAGA
ICO42_staple53	AAAAGAGGACTATGCAGTAACCTCAGGATCCTACTTTAAAAG
ICO42_staple54	AGTGACGTAACTTTTTTCCCGGGTTGACCATGTTACTTTTTTTGGTCAGTTATGTATTCATTCTTTTTATTGTGCAAT
ICO42_staple55	AGTTATGTGTTTTTTTGCAGGTAGCGCTTATTACCCTTTTTTCGATATTCAC

ICO42, CpG-containing, Outward-facing.

Name	Sequence
ICO42_CpG_out_staple2	CCCTAAAGTTGCGTTGCGCTCCTGCAGCACACTAAATCGGAATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple4	CAAGATATAAGATGAGAAGTGGCTATGCGGCTTAAACGCAAGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple6	TTAGAAATTGGTTACTAGGTACATCTGCGTCACGGTTAGAAGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple7	TCTATTATACAGTCTGGCAATACTCGCTAGCGGGAGTCACAGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple8	CGAGAAAGGACGGGGAAAGCCGGCTCAGGGTCTAGAACGTGGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple10	TCATTTTTCATCACCCTAATCAAGTTTTTTTGTTAAATCAGCTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple13	GAAGGATTAAGTAAATATCTTCACACTCCCGTGGGACGTAGCTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple14	TCAACAGCGAGTGGGTTACATCGATGGAACAGTTACTGGATCTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple15	ATAGACCGCAAAATCCCTTATAAATGAGTTAACGTCAAAAGATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out_staple17	GAGATAGGACCCCAATCATATGTACCCCGTAGACAGATCGCTTCGTCGTTTTGTCGTTTTGTCGTT

ICO42_CpG_out _staple20	ATGAGCACCGCCCCGAAGAACGTTTTGATCCTTTTTCCAATGTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple21	GGCTGGCTACCACTTATGCGCTCGTAATTATTTTGCCCTTCCTCG TCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple22	ATATATACGTAACTGTCAGACCAAGATGGGGCCAGTTTACTCTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple23	CAACAGAGTATCTACACGACGGGGAGTCACCGGGTGGCTCAAT CGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple26	CGTGACACGAGCCGGTGAGCGTGGGTCTCACCAAACGACGAGT CGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple28	CGGGCAAGCGCGGTATTATCCCGTTGGATTAGACATTGACGCTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple30	CAGTAAGAAAAAGCATCTTACGGATACTCCTACTTGGCATGATC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple31	GCTTTTTTATCGGAGGACCGAAGGCGTTGAACAAAGCTAACCTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple32	CGACGACGAACTCGCCTTGATCGTTGGGAGAGTTTGAGGGGAT CGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple35	TTATTTTTACCATGAGTGATAACACTGCGAACCCCTATTTGTTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple38	TGACTTGGTTTTGCCTTCCTGTTTTTGCTACTATTCTCAGAATCG TCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple41	TAGTGATAGTTTGCACACGTCTGGAGATGGTAAACGGACATCTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple42	GTGTCGCCAGAGTATGAGTATTCACCACGTCGGAACATTTCCTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple43	ATAAATGCTCCGCTCATGAGACAATGGTAAAGCATAACCCTGT CGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple44	TTGATACTATTCCTAGATTAGAAGGGCACCAGGTGCGTTGTTTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple45	TGTCATGAGATACGCCTATTTTTATGTAGAGTGTTAGGTTAATC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple46	GCGAACGAAGAACCAGGGAAAACCCTTAAGTGACAGGTACGAT CGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple47	ACTGCTATCTATACAAACTAAGAGACAAAATGTGTAAAAGGAT CGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple50	AGAGCTATGCAGCACCTGTATCGGCCTTCTCCGAGGGTTCCCTC GTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_out _staple53	AAAAGAGGACTATGCAGTAACCTCAGGATCCTACTTTAAAAGT CGTCGTTTTGTCGTTTTGTCGTT

ICO42, CpG-free, Outward-facing.

Name	Sequence
ICO42_CpGfree_out_staple2	CCCTAAAGTTGCGTTGCGCTCCTGCAGCACACTAAATCGGAACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple4	CAAGATATAAGATGAGAAGTGGCTATGCGGCTTAAACGCAAGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple6	TTAGAAATTGGTTACTAGGTACATCTGCGTCACGGTTAGAAGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple7	TCTATTATACAGTCTGGCAATACTCGCTAGCGGGAGTCACAGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple8	CGAGAAAGGACGGGGAAAGCCGGCTCAGGGTCTAGAACGTGGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple10	TCATTTTTCATCACCCTAATCAAGTTTTTTTGTTAAATCAGCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple13	GAAGGATTAAGTAAATATCTTCACACTCCCGTGGGACGTAGCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple14	TCAACAGCGAGTGGGTTACATCGATGGAACAGTTACTGGATCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple15	ATAGACCGCAAAATCCCTTATAAATGAGTTAACGTCAAAAGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple17	GAGATAGGACCCCAATCATATGTACCCCGTAGACAGATCGCTCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple20	ATGAGCACCGCCCCGAAGAACGTTTTGATCCTTTTTCCAATGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple21	GGCTGGCTACCACTTATGCGCTCGTAATTATTTTGCCCTTCCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple22	ATATATACGTAACTGTCAGACCAAGATGGGGCCAGTTTACTCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple23	CAACAGAGTATCTACACGACGGGGAGTCACCGGGTGGCTCAACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple26	CGTGACACGAGCCGGTGAGCGTGGGTCTCACCAAACGACGAGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple28	CGGGCAAGCGCGGTATTATCCCGTTGGATTAGACATTGACGCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple30	CAGTAAGAAAAAGCATCTTACGGATACTCCTACTTGGCATGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple31	GCTTTTTTATCGGAGGACCGAAGGCGTTGAACAAAGCTAACCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple32	CGACGACGAACTCGCCTTGATCGTTGGGAGAGTTTGAGGGGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple35	TTATTTTTACCATGAGTGATAACACTGCGAACCCCTATTTGTCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple38	TGACTTGGTTTTGCCTTCCTGTTTTTGCTACTATTCTCAGAACCGATACTGCCATAGACGGCAA

ICO42_CpGfree_out_staple41	TAGTGATAGTTTGCACACGTCTGGAGATGGTAAACGGACATCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple42	GTGTCGCCAGAGTATGAGTATTCACCACGTCGGAACATTTCCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple43	ATAAATGCTCCGCTCATGAGACAATGGTAAAGCATAACCCTGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple44	TTGATACTATTCCTAGATTAGAAGGGCACCAGGTGCGTTGTTCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple45	TGTCATGAGATACGCCTATTTTTATGTAGAGTGTTAGGTTAACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple46	GCGAACGAAGAACCAGGGAAAACCCTTAAGTGACAGGTACGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple47	ACTGCTATCTATACAAACTAAGAGACAAAATGTGTAAAAGGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple50	AGAGCTATGCAGCACCTGTATCGGCCTTCTCCGAGGGTTCCCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_out_staple53	AAAAGAGGACTATGCAGTAACCTCAGGATCCTACTTTAAAAGCCCGATACTGCCATAGACGGCAA

ICO42, CpG-containing, Inward-facing.

Name	Sequence
ICO42_CpG_in_staple2	CGGAACCCTAAAGTTGCGTTGCGCTCCTGCAGCACACTAAATTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple4	AAGTGGCTATGCGGCTTAAACGCAAGCAAGATATAAGATGAGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple6	AGGTACATCTGCGTCACGGTTAGAAGTTAGAAATTGGTTACTTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple7	GCAATACTCGCTAGCGGGAGTCACAGTCTATTATACAGTCTGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple8	AAGCCGGCTCAGGGTCTAGAACGTGGCGAGAAAGGACGGGGATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple10	TCAGCTCATTTTTCATCACCCTAATCAAGTTTTTTTGTTAAATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple13	GTAGCGAAGGATTAAGTAAATATCTTCACACTCCCGTGGGACTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple14	TACATCGATGGAACAGTTACTGGATCTCAACAGCGAGTGGGTTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple15	CTTATAAATGAGTTAACGTCAAAAGAATAGACCGCAAAATCCTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple17	TCGCTGAGATAGGACCCCAATCATATGTACCCCGTAGACAGATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple20	AGAACGTTTTGATCCTTTTTCCAATGATGAGCACCGCCCCGATCGTCGTTTTGTCGTTTTGTCGTT

ICO42_CpG_in_staple 21	CTTCCGGCTGGCTACCACTTATGCGCTCGTAATTATTTTGC CTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 22	CAGACCAAGATGGGGCCAGTTTACTCATATATACGTAACT GTTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 23	CTCAACAACAGAGTATCTACACGACGGGGAGTCACCGGG TGGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 26	ACGAGCGTGACACGAGCCGGTGAGCGTGGGTCTCACCAA ACGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 28	TATCCCGTTGGATTAGACATTGACGCCGGGCAAGCGCGGT ATTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 30	CTTACGGATACTCCTACTTGGCATGACAGTAAGAAAAAGC ATTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 31	ACCGAAGGCGTTGAACAAAGCTAACCGCTTTTTTATCGGA GGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 32	GGGGACGACGACGAACTCGCCTTGATCGTTGGGAGAGTTT GATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 35	TTTGTTTATTTTTACCATGAGTGATAACACTGCGAACCCCT ATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 38	CAGAATGACTTGGTTTTGCCTTCCTGTTTTTGCTACTATTC TTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 41	ACGTCTGGAGATGGTAAACGGACATCTAGTGATAGTTTGC ACTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 42	AGTATTCACCACGTCGGAACATTTCCGTGTCGCCAGAGTA TGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 43	TGAGACAATGGTAAAGCATAACCCTGATAAATGCTCCGCT CATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 44	ATTAGAAGGGCACCAGGTGCGTTGTTTTGATACTATTCCT AGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 45	TATTTTTATGTAGAGTGTTAGGTTAATGTCATGAGATACGC CTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 46	GGAAAACCCTTAAGTGACAGGTACGAGCGAACGAAGAAC CAGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 47	AAGGAACTGCTATCTATACAAACTAAGAGACAAAATGTGT AATCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 50	TTCCCAGAGCTATGCAGCACCTGTATCGGCCTTCTCCGAG GGTCGTCGTTTTGTCGTTTTGTCGTT
ICO42_CpG_in_staple 53	GTAACCTCAGGATCCTACTTTAAAAGAAAAGAGGACTATG CATCGTCGTTTTGTCGTTTTGTCGTT

ICO42, CpG-free, Inward-facing.

Name	Sequence
ICO42_CpGfree_in_staple2	CGGAACCCTAAAGTTGCGTTGCGCTCCTGCAGCACACTAAATCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple4	AAGTGGCTATGCGGCTTAAACGCAAGCAAGATATAAGATGAGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple6	AGGTACATCTGCGTCACGGTTAGAAGTTAGAAATTGGTTACTCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple7	GCAATACTCGCTAGCGGGAGTCACAGTCTATTATACAGTCTGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple8	AAGCCGGCTCAGGGTCTAGAACGTGGCGAGAAAGGACGGGGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple10	TCAGCTCATTTTTCATCACCCTAATCAAGTTTTTTTGTTAAACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple13	GTAGCGAAGGATTAAGTAAATATCTTCACACTCCCGTGGGACCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple14	TACATCGATGGAACAGTTACTGGATCTCAACAGCGAGTGGGTCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple15	CTTATAAATGAGTTAACGTCAAAAGAATAGACCGCAAAATCCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple17	TCGCTGAGATAGGACCCCAATCATATGTACCCCGTAGACAGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple20	AGAACGTTTTGATCCTTTTTCCAATGATGAGCACCGCCCCGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple21	CTTCCGGCTGGCTACCACTTATGCGCTCGTAATTATTTTGCCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple22	CAGACCAAGATGGGGCCAGTTTACTCATATATACGTAACTGTCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple23	CTCAACAACAGAGTATCTACACGACGGGGAGTCACCGGGTGGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple26	ACGAGCGTGACACGAGCCGGTGAGCGTGGGTCTCACCAAACGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple28	TATCCCGTTGGATTAGACATTGACGCCGGGCAAGCGCGGTATCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple30	CTTACGGATACTCCTACTTGGCATGACAGTAAGAAAAAGCATCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple31	ACCGAAGGCGTTGAACAAAGCTAACCGCTTTTTTATCGGAGGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple32	GGGGACGACGACGAACTCGCCTTGATCGTTGGGAGAGTTTGACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple35	TTTGTTTATTTTTACCATGAGTGATAACACTGCGAACCCCTACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple38	CAGAATGACTTGGTTTTGCCTTCCTGTTTTTGCTACTATTCTCCCGATACTGCCATAGACGGCAA

ICO42_CpGfree_in_staple41	ACGTCTGGAGATGGTAAACGGACATCTAGTGATAGTTTGCACCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple42	AGTATTCACCACGTCGGAACATTTCCGTGTCGCCAGAGTATGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple43	TGAGACAATGGTAAAGCATAACCCTGATAAATGCTCCGCTCACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple44	ATTAGAAGGGCACCAGGTGCGTTGTTTTGATACTATTCCTAGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple45	TATTTTTATGTAGAGTGTTAGGTTAATGTCATGAGATACGCCCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple46	GGAAAACCCTTAAGTGACAGGTACGAGCGAACGAAGAACCAGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple47	AAGGAACTGCTATCTATACAAACTAAGAGACAAAATGTGTAACCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple50	TTCCCAGAGCTATGCAGCACCTGTATCGGCCTTCTCCGAGGGCCCGATACTGCCATAGACGGCAA
ICO42_CpGfree_in_staple53	GTAACCTCAGGATCCTACTTTAAAAGAAAAGAGGACTATGCACCCGATACTGCCATAGACGGCAA

Appendix Table 4. Sequences for siRNA-DNA nanoparticles. Sequences are given in 5' to 3' orientation.

PB84 with linker. Linker sequence is displayed in bold.

| Name	Sequence
PB84_linker_staple2	ATAAAGTTTTGCGTTGCGCTCCTGCAGCATGGATGGAGGCG**GCCCGATACTGCCATAGACGGCAA**
PB84_linker_staple3	TGCGCTCGCCGCTACAGGGCGCACATTAAGCAGGACCACTTA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple4	GCTGGTTTACACCCGCCGCGCTTAATGCGGCCCTTCCGGCT**GCCCGATACTGCCATAGACGGCAA**
PB84_linker_staple6	GTGTAGCGAGCCGGCGAACGTGGCGAGAAAGGGCGCTGGCAA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple7	CGGGCGCTAGGAAGGGAAGAAAGCAGCTATAT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple8	TCTGACATGTCCGAGGGTTCCCAGGAAAGGAG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple10	TACTTTAGGGAACCCTAAAGGGAGCCCCCGTTTACTCATATA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple11	TTCATTTTGGTGCCGTAAAGCACTAAATCATTGATTTAAAAC**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple12	TCTAGGTGAATCAAGTTTTTTGGGGTCGATAATTTAAAAGGA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple14	CGATGGCCGAACAAGAGTCCACTATTAAACCGTCTATCAGGG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple15	GCGAAAAAGAACGTGGACTCCAACAATCAGTA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple17	GAGATAGGATCAGCTCATTTTTTAACCAAAAAAGAATAGAC**CCCCGATACTGCCATAGACGGCAA**
PB84_linker_staple18	TATAAATCTAGGCCGAAATCGGCATGTTGTAG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple19	AAAGTCAATAGGAAGCCGCATAAGAAATCCCT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple21	GGTTGTAAATGACCAAAATCCCTTAACGTTCATATGTACCCC**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple22	GACCCCAAGAGTTTTCGTTCCACTCCCGTATT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple25	GGTGCCTCCGAAATAGACAGATCGAAAGCAT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple26	CTTACGGATACTCACCAGTCACAGCTGAGATA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple27	GTTATCTAGGTGAGCGTGGGTCTCGCGGTCTCCCGTATCGTA**CCCGATACTGCCATAGACGGCAA**

|	
PB84_linker_staple28	GGTAAGCCATCATTGCAGCACTGGAACCGGAG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple29	CTGAATGACGCCTTGATCGTTGGGGGCCAGAT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple31	AATGTGCGGGCGAACTACTTACTCTAGCTCACTTTTCGGGGA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple33	TCAAATATATGCCTGTAGCAATGGCAACATTTCTAAATACAT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple37	CATTTTGCAGGACCGAAGGAGCTAACCGCCCCTTTTTTGCGG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple39	AAGTAAAAGAGTGATAACACTGCGGCCAAGAAACGCTGGTGA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple41	AGAGAATTTCTCAGAATGACTTGGTTGAGTGGCATGACAGTA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple44	GAACGTTTGCACGAGTGGGTTACATCGAATTTTCGCCCCGAA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple45	CTTGAGAGCTGGATCTCAACAGCGTGAATCTG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple46	TCTATACAACGAAACCTTGCACTCGTAAGATC**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple48	AAGAGTATATGCTTCAATAATATTACATGCGA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple49	TGTGACCTAGAGAATAAGTAAATTGAAAAAGG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple51	CCTACTTTGGTTAATGTCATGATAATAATTTCATTGTGCAAT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple52	TTGGTCGTCGTGATACGCCTATTTTTATAAAAAGAAAAGAGG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple53	AAGCATGGACGAATTCGACGAAAGGGCCTTAATGACCCAAGA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple55	TCTCCACCGCGCCCGGGTGGCTCAACAACTAAACGAATAGCC**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple60	CTATGTTCATTAGAAGGCGTTGTTATTACCCT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple61	CGATATTCCACAGTCTATTATCTTTTGATACT**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple62	ACTGCTATATAACCATGTTACTTGGTCAGATGTGTAAAAGGA**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple64	GCTAACGCAACTCCCGGGTTGGTTTGCACAAGATGAGAAGTG**CCCGATACTGCCATAGACGGCAA**
PB84_linker_staple66	TTCACGACCGACAGTCTGGCAATACTAGTACAAGTAAATATC**CCCGATACTGCCATAGACGGCAA**

siRNA sequences. 2'o-methylated, 5' Atto488. m = 2'O-Me, r = RNA. 2-nt 3' overhang and linker sequence are in bold.

Mouse PD-L1 siRNA, sense strand with linker, 5' Atto488.
/5ATTO488N/mGmCmGmAmAmUmCmAmCmGmCmUmGmAmAmAmGmUmC**mUmUrGrCrCrGrUrCrUrArUrGrGrCrArGrUrArUrC**
Mouse PD-L1 siRNA, antisense strand.
mGmAmCmUmUmUmCmAmGmCmGmUmGmAmUmUmCmGmCmUmU
Mouse PD-L1 scrambled siRNA, sense strand with linker, 5' Alexa647.
/5ATTO647NN/mGmGmAmAmCmGmGmCmUmAmCmUmCmGmUmAmAmCmA**mUmUrGrCrCrGrUrCrUrArUrGrGrCrArGrUrArUrC**
Mouse PD-L1 scrambled siRNA, antisense strand.
mUmGmUmUmAmCmGmAmGmUmAmGmCmCmGmUmUmCmCmUmU
GAPDH1.
Antisense: 5' AAAUCCGUUGACUCCGACC (**DUDU**) 3'
Sense: 5' GGUCGGAGUCAACGGAUUU (**UUGCCGUCUAUGGCAGUAUC**) 3'
Scrambled GAPDH1.
Antisense: 5' CCGCUGCUACUCAAAGACU (**DUDU**) 3'
Sense: 5' AGUCUUUGAGUAGCAGCGG (**UUGCCGUCUAUGGCAGUAUC**) 3'
GAPDH2.
Antisense: 5' CAGAGAUGAUGACCCUUUU (**DUDU**) 3'
Sense: 5' AAAAGGGUCAUCAUCUCUG (**UUGCCGUCUAUGGCAGUAUC**) 3'
Scrambled GAPDH2.
Antisense: 5' CGUAAGAACCAGUCUUGUU (**DUDU**) 3'
Sense: 5' AACAAGACUGGUUCUUACG (**UUGCCGUCUAUGGCAGUAUC**) 3'
GAPDH3.
Antisense: 5' GGCCAUCCACAGUCUUCUG (**DUDU**) 3'
Sense: 5' CAGAAGACUGUGGAUGGCC (**UUGCCGUCUAUGGCAGUAUC**) 3'
Scrambled GAPDH3.
Antisense: 5' UCACGCGAGCGUACUCUUC (**DUDU**) 3'
Sense: 5' GAAGAGUACGCUCGCGUGA (**UUGCCGUCUAUGGCAGUAUC**) 3'

Supplementary Information: Chapter 2
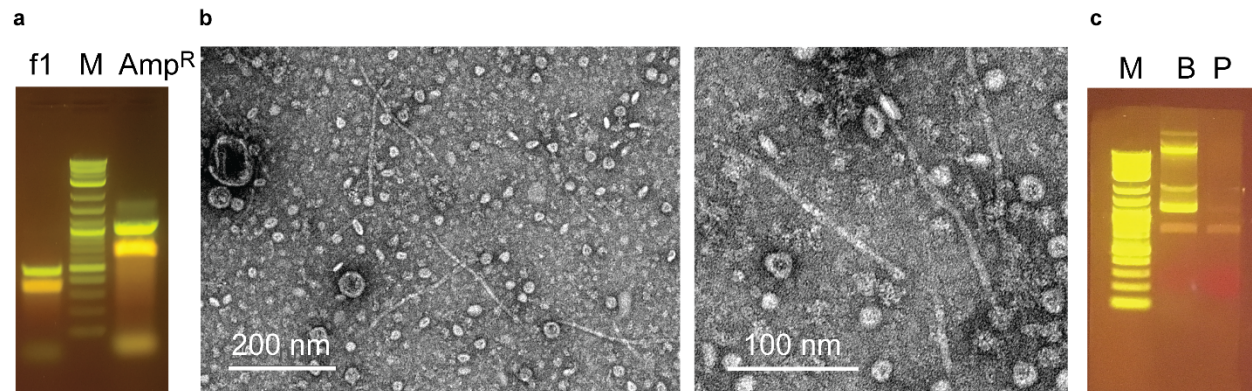
Figure A.3. Single-strand DNA synthesis and phage assembly. (**a**) Full agarose gel from **Figure 1b**, showing asymmetric PCR products of f1 origin-containing synthetic gBlock and bla-selection marker-containing ssDNAs. (**b**) Wide-field TEM micrograph showing assembled phage in the clarified media, rod-shaped and approximately 200 nm in size. (**c**) Full agarose gel from **Figure 1d** showing DNA purification results from the bacterial pellet and the purified phage from the media. M: Marker. Scale bar is shown for reference.
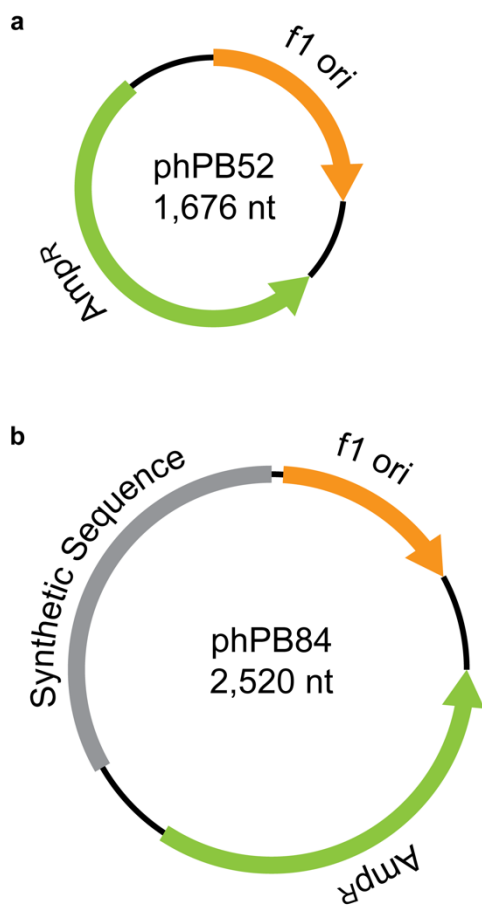
Figure A.4. Plasmid maps of constructs generated. (**a**) phPB52 and (**b**) phPB84.
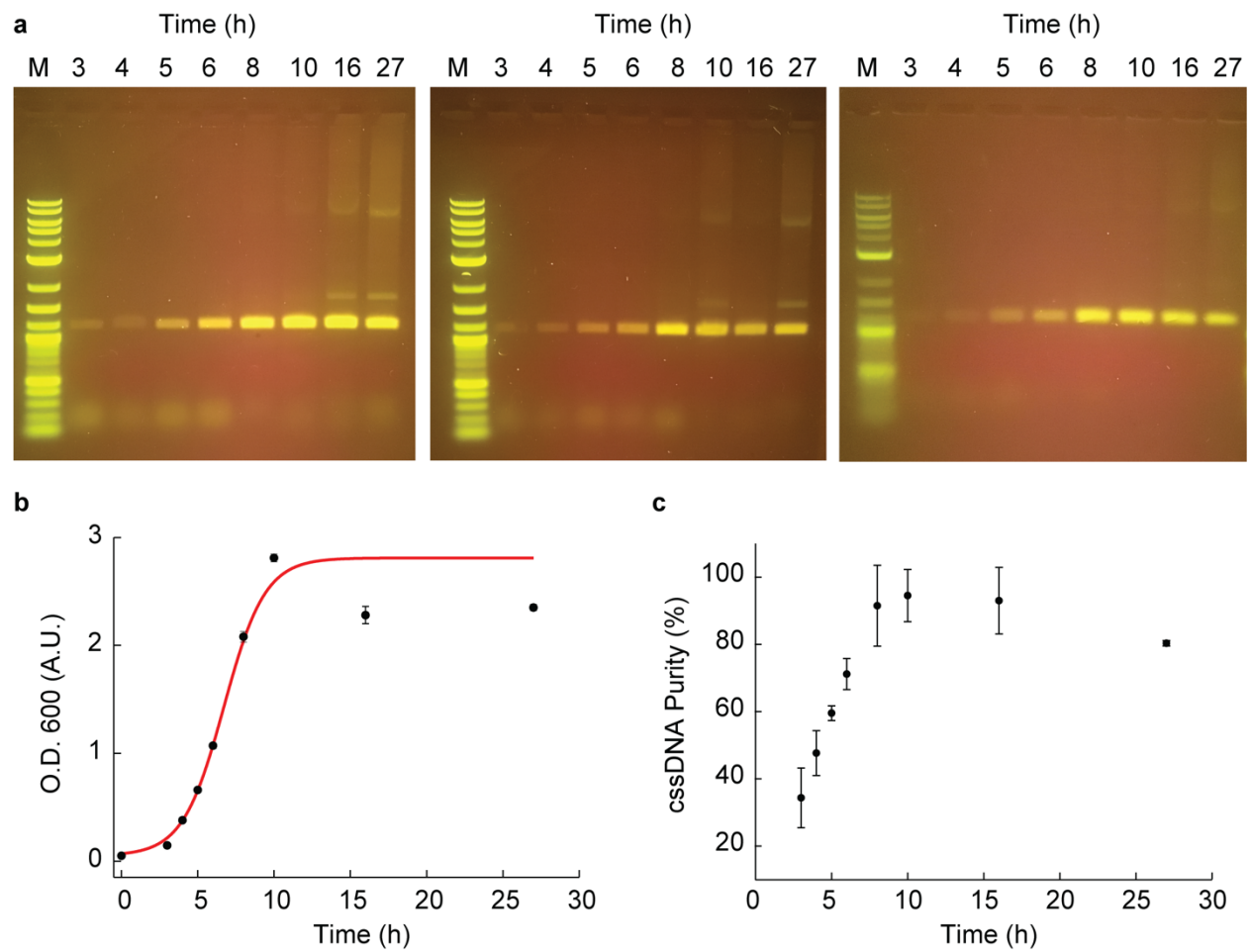
Figure A.5. Triplicate time course assays for optimizing phPB84 miniphage production. (a) DNA was prepared from column purification of processed, cleared supernatant of bacterial growths at the indicated times. Growth was in 2×YT media using the SS320 strain transformed with phPB84 and the M13cp helper plasmid. (**b**) O.D.600 time course monitoring the growth of the bacterial culture, inoculated at the 0 time point to have an O.D.600 of 0.05. (**c**) Percent purity was measured based on gel percent intensity of the cssDNA compared to total lane intensity, measured by ImageJ (52)

a

M   DH5α F'Iq

b

Terrific Broth

Time [h]

M   3   4   6   8   12   16   M   8   8   8

Time [h]

c

M   pH 6.4   M   pH 6.8   M   pH 7.2   M   pH 7.6   M   pH 8.0
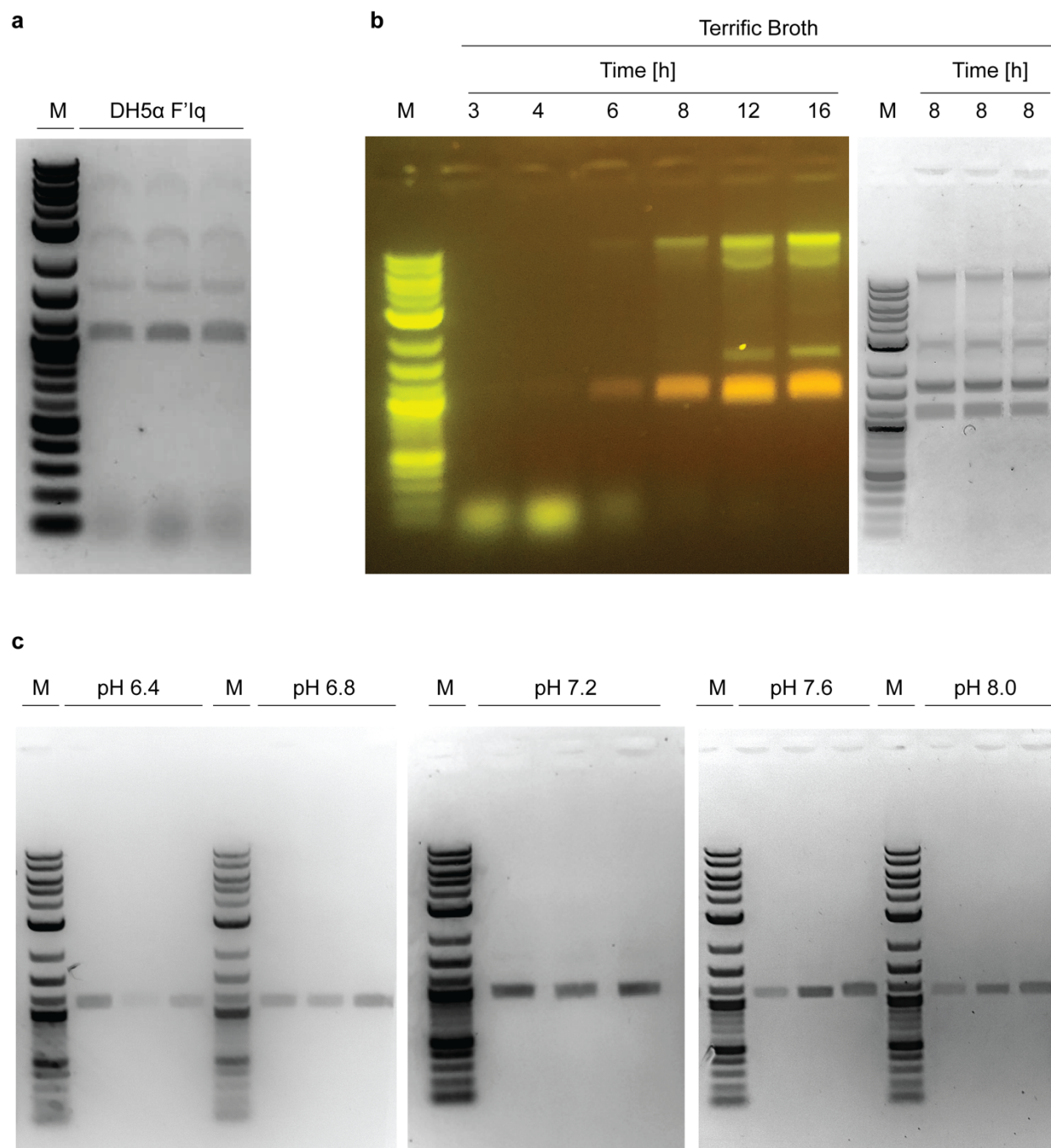
Figure A.6. Agarose gels showing triplicate measurements of cssDNA prepared from the cleared media for optimizing conditions. (**a**) Independent colonies of the DH5a F'Iq (Invitrogen) strain with helper plasmid transformed with phPB84 and grown to the 8 h time point in 2×YT media. (**b**) Agarose gel showing processed cssDNA from a time course assay and triplicate growths at the 8 h time point with Terrific Broth as the growth media using the SS320 strain. (**c**) Agarose gel showing triplicate cssDNA prepared from the cleared media using 2×YT media at pHs indicated, buffered with 100 mM HEPES-NaOH at the 8 h time point using the SS320 strain.

173
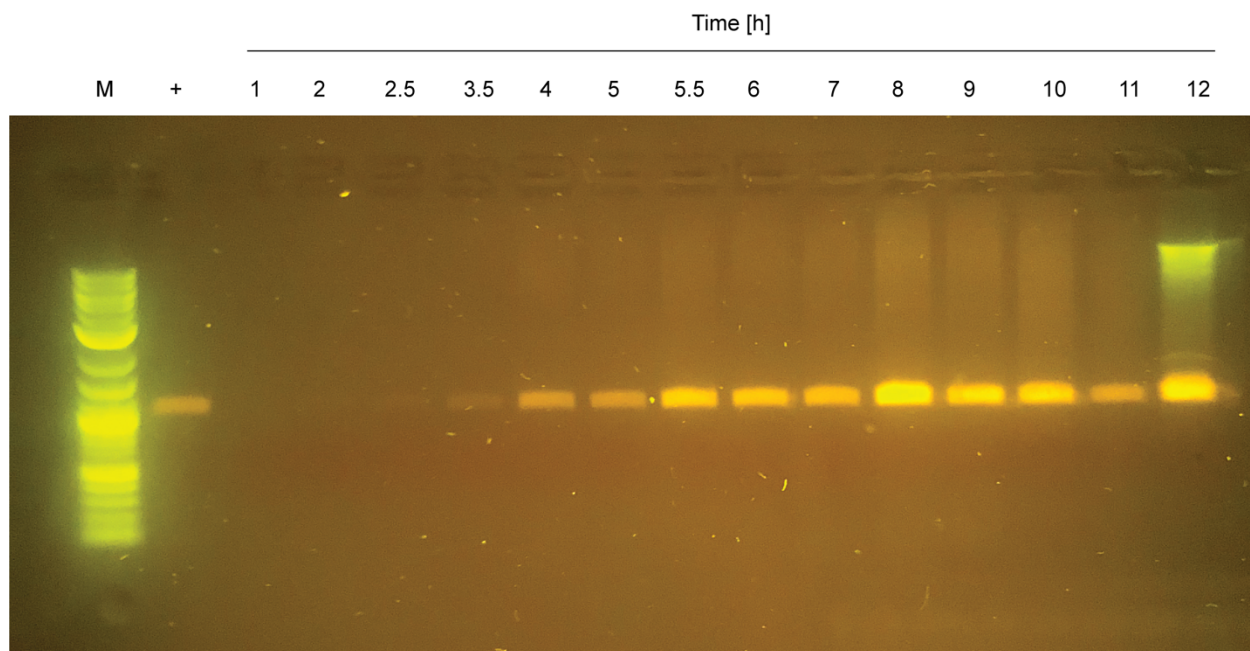
Figure A.7. Agarose gel of time course production of cssDNA phPB84 in a bioreactor. Growth was in 2×YT media using the SS320 strain transformed with phPB84 and the M13cp helper plasmid with pH controlled to 7.0. M: marker; +: phPB84 from shaker flask growth.
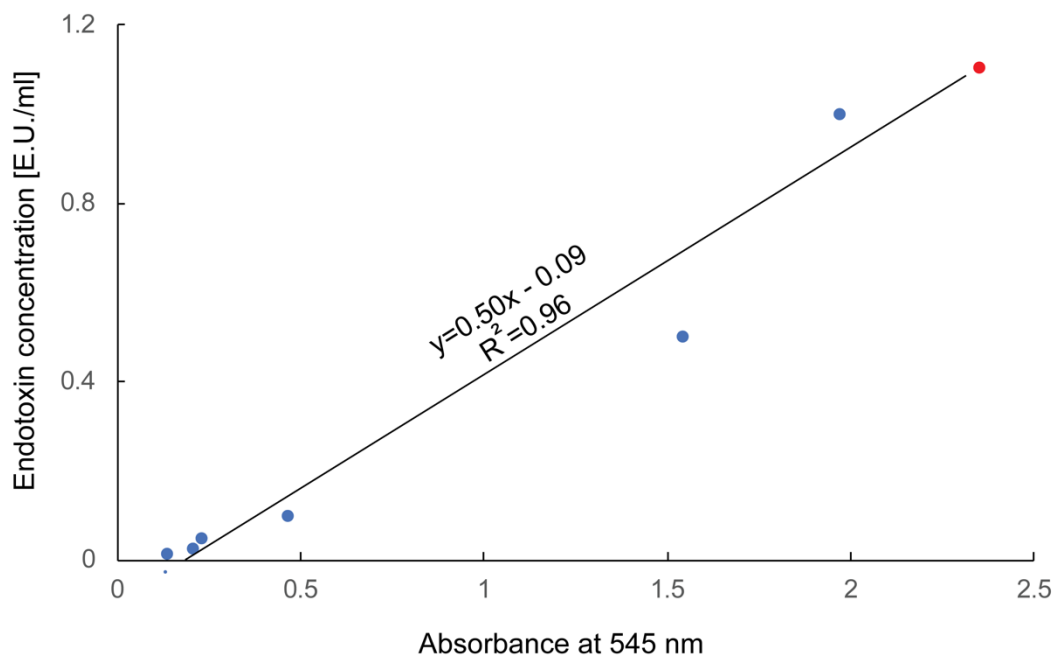
Figure A.8. Endotoxin screening using standard curve with the 10 nM scaffold data point shown. The ToxinSensor Chromogenic LAL Endotoxin Assay Kit (GenScript) was used to assay endotoxin levels in the prepared DNA. The eluted DNA was diluted to 10 nM from the Endofree Maxiprep purification kit (Qiagen) and tested against a standard curve by following the manufacturer's protocol monitoring absorbance at 545 nm, shown as blue dots. The endotoxin level was calculated to be 1.1 ± 0.1, shown as a red dot.
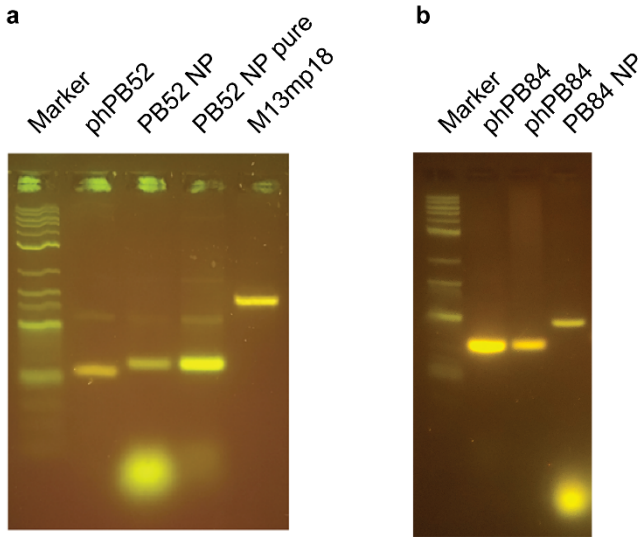
Figure A.9. Agarose gel analysis of folding of nanoparticles. (**a**) Uncropped gel of a folded pentagonal bipyramid of 52-bp edge length. (**b**) Uncropped gel of a folded pentagonal bipyramid of 84-bp edge length.
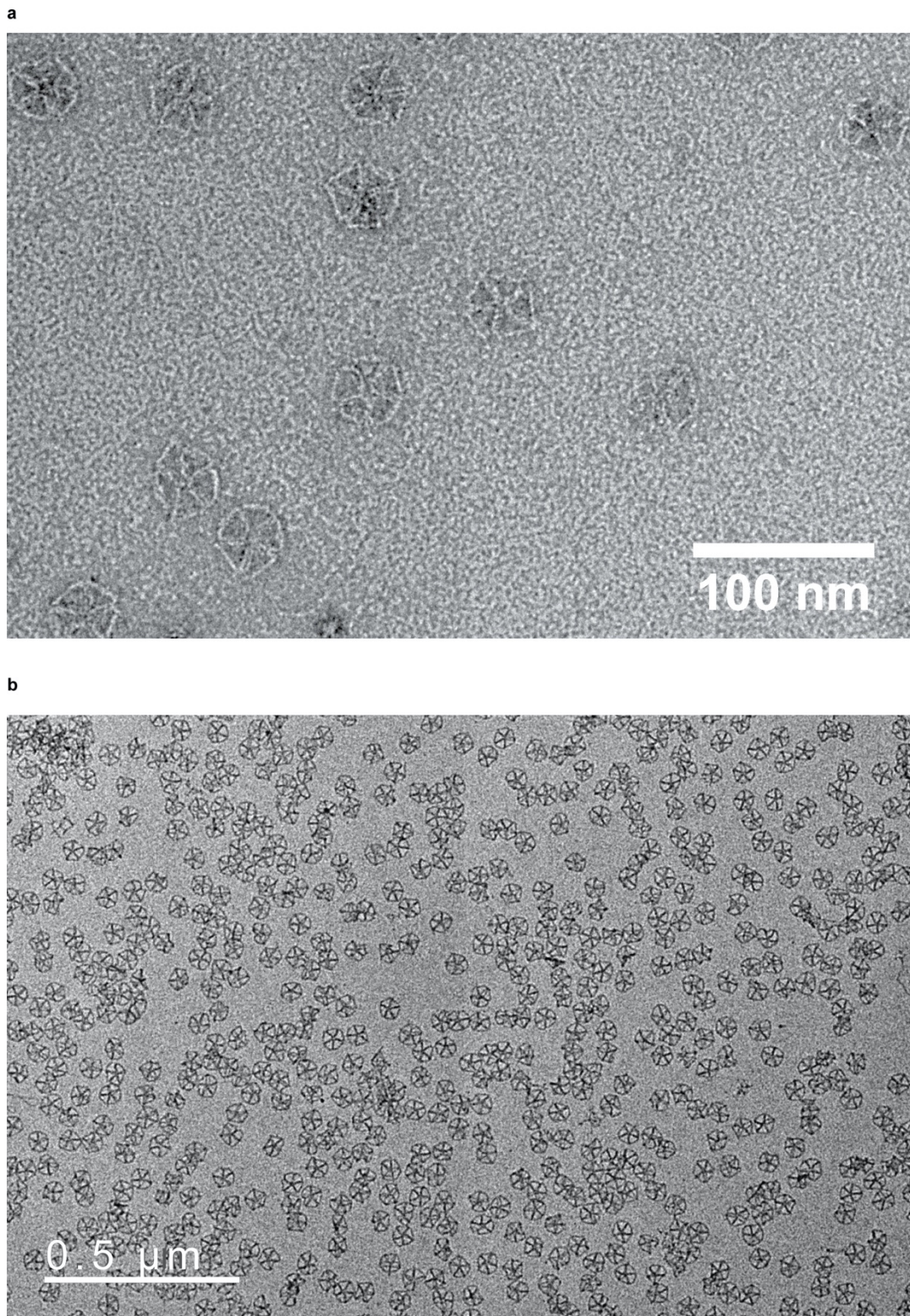
Figure A.10. Wide-field TEM micrographs of scaffolded DNA nanoparticles. (**a**) Pentagonal bipyramid with 52-bp edge lengths and (**b**) Pentagonal bipyramid with 84-bp edge lengths are shown. Scale bars are shown for reference.

Appendix Table 5. Sequences of each of the phages and amplicons generated in this study.

phPB52	GAGCGCAACGCAATTAATGTGCGCCCTGTAGCGGCGCATTAAGCGCGG CGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCC TAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCC GGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT TTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATG GTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGAC GTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACA ACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCC GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAAC GCGAATTACAACCGGGGTACATATGATTGGGGTCTGACGCTCAGTGGA ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTA AAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGT GAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG ATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT GGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGG AAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC CATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCA TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGT TGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAG TAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT ACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT CTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTT TAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAA GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCA AAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACG GAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTT ATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACC TGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAG GCGTATCACGAGGCCCTTTCGTC
phPB84	GAGCGCAACGCAATTAATGTGCGCCCTGTAGCGGCGCATTAAGCGCGG CGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCC TAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCC GGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT TTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATG GTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGAC GTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACA ACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCC GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAAC

	GCGAATTACAACCGGGGTACATATGATTGGGGTCTGACGCTCAGTGGA ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTA AAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGT GAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG ATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAtAAGTG GTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCC ATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCAT TCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT GTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGT AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCT TGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGG ATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCA ACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTAT CAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAA ATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTG ACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGC GTATCACGAGGCCCTTTCGTCGAATTCGTCGTCGTCCCCTCAAACTCTT GGGTGGAGAGGCTATTCGTTTAAGGTCACATCGCATGTAATTTACTTAT TCTCTGTTGTTGAGCCACCCGGGCGCCAGATTTTGTTTAAAGCTTTGTC TCTTAGTTTGTATAGACAGATTCAGAGTGCAAGGTTTCGTTCGCTCGTA CCTGGTTTTCCCTGGTTCTTCACAGATAGGATTTGACTTTCTACAACACT TATGCGGCTTCCTACCCGTTTGAAGGCCGATACAGGTGCTGCGCAAAA TGCGGGCGAACATAGAGTATCAAAACAACGCCTTCTAATCTAGGAATA TAGGGAAGATACGTATTTGCTACCATGCTTTCTTGGGTCATTAACGACC AACCTCTTTTCTTTTAAAGTAGGATTGCACAATGAATGAATACACGTGG TCCGATAACTGACCAAGTAACATGGTTATCACTaGATGTCCGCCAGACG TGTGCAAACCAACCCGGGAGTTACGTCACTAATCCTTCGCTACGTCGTG AAGATATTTACTTGTGAATATCGAGGGTAATAAGATAATAGACTGTGA CTAGTATTGCCAGACTGTCGCTACCTGCAACACATAACTATCCTGAGGT TACTGCATAGTACTGATTACACCCGAGTCAAAATTTCTAACTTCTAACA TGTACCTAGTAACCAGCTCAATAATTATGTCAGAATATAGCTCTGGGA ACCCTCGGACAATTATGATACACGGTATTAATATCTTGCTTGCGTTAGC CACTTCTCATCTTTGGATACCGATTCTATTTTGCATAGCAGTTCCTTTTA CACATATAAGAATTTCGCCATAGGTATGCTGCAG
F1ori amplicon	/5phos/GAGCGCAACGCAATTAATGTGCGCCCTGTAGCGGCGCATTAAGC GCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC

Top Strand	GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTT CGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTC CGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTG ATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTT GACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGA ACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTT GCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTT AACGCGAATTACAACCGGGGTACATATGATT
AmpR amplicon Bottom Strand	/5phos/ACATTAATTGCGTTGCGCTCGACGAAAGGGCCTCGTGATACGCC TATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGG TGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTC TAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAA ATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTC CGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCT CACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGT GCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTT GAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAG TTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCA ACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCA CCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA TGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTC TGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACA TGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATG AAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGG CAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTC CCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACC ACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCT GGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCA GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCA CTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTT AGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGAT CCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCC ACTGAGCGTCAGACCCCAATCATATGTACCCCGGTTGTA

Appendix Table 6. Primers used for cloning and sequence validation.

SeqPrimer1	CAATCATATGTACCCCGGTTGT
SeqPrimer2	GACGAAAGGGCCTCGTGATA
SeqPrimer3	ACCGAAGGAGCTAACCGCTT
SeqPrimer4	CCCGATTTAGAGCTTGACGG
SeqPrimer5	TGCTATGCAAAATAGAATCG
SeqPrimer6	ACCAGGGAAAACCAGGTAC
SeqPrimer7	GCGCTCCTGCAGGCTGAAAAGGTGGCATCAATAG
SeqPrimer8	CTGCAGCATACCTATGGCGAAATTCTTATATGTG
SeqPrimer9	ACCAGGGAAAACCAGGTAC
SeqPrimer10	CTGAATGAAGCCATACCAAACGA
SeqPrimer11	GACGAAAGGGCCTCGTGATA
SeqPrimer12	ACATTAATTGCGTTGCGCTC
F1gBlock_forward	/5phos/GAGCGCAACGCAATTAATGT
F1gBlock_reverse	ATCATATGTACCCCGGTTGT
AmpR_forward	ACAACCGGGGTACATATGATTGGGGTCTGACGCTCAGTGGAA
AmpR_reverse	/5phos/ACATTAATTGCGTTGCGCTCGACGAAAGGGCCTCGTGATA
PB84ins_PstI_reverse	CTGCAGCATACCTATGGCGAAATTCTTATATGTG
F1A_PstI_for	GGTATGCTGCAGGAGCGCAACGCAATTAATGT
F1A_EcoRI_rev	GACGACGAATTCGACGAAAGGGCCTCGTGATA

Appendix Table 7. Staples for 52-bp edge-length pentagonal bipyramid nanoparticle from phPB52 sequence

pbip52_1-1293-V	TCTCAACAGCGTTTTTGTAAGATCCTAACATGGGGGATTTTTTCATGTAACT
pbip52_1-1520-E	TTTTTCTAAAAGTGGGTTACATCGAACTGGACTATTTGTTTA
pbip52_1-1510-E	TACATTCAAATGGGTGCACG
pbip52_2-1261-E	GCCCCGAAGAGAGCTAACCGCTTTTTTGCACTGAGAGTTTTC
pbip52_2-1251-E	ACGTTTTCCAAGGACCGAAG
pbip52_3-1088-V	GTAAGAGAATTTTTTTATGCAGTGCTCTTACTTCTGATTTTTCAACGATCGGATGATGAGCACTTTTTTTTTAAAGTT
pbip52_3-783-E	TCTGCGCTGGCGGATAAAGTTGCAACGGATG
pbip52_3-1096-E	GCATGACAACAGAAAGCATCTTGGACCACT
pbip52_4-880-V	CAACGTTGCGCTTTTTAAACTATTAAGACAGATCGCTTTTTTGAGATAGGTG
pbip52_4-1043-E	GCGGCCAAGCCATAACCATGAGTGCTGTAGC
pbip52_4-888-E	AATGGCAAGTGACACCACGATGCATAACACT
pbip52_5-484-E	ATCTCATGACCGCTACAGGGCGCCACTTTTCCCTTTTTGATA
pbip52_5-474-E	CAAAATCCCTCTTAAAGCGC
pbip52_6-516-V	TAAAAGGATCTTTTTTAGGTGAAGATGGGGAAATGTGTTTTTCGCGGAACCC
pbip52_6-951-E	TTGGGAACCGTTTAAAACTTCATTTTTAATTCGCCTTGATCG
pbip52_6-941-E	GAGCTGAATGTTTAGATTGA
pbip52_7-568-V	ACCAAGTTTACTTTTTTCATATATACAAGCCATACCATTTTTAACGACGAGC
pbip52_7-419-E	TTAAATTTTGAGCGTCAGACCCCAATTGGTA
pbip52_7-576-E	ACTGTCAGCCTCACTGATTAAGCAATTCGCG
pbip52_8-828-V	ACAATTAATAGTTTTTACTGGATGGACGGCCCTTCCGTTTTTGCTGGCTGGT
pbip52_8-848-E	CTTACTCTAGCAACTATGGATGAACGAAATACTGGCGAACTA
pbip52_8-838-E	CTTCCCGGCAGGGAGTCAGG
pbip52_9-1190-V	GACGCCGGGCATTTTTAGAGCAACTCTTGGTTGAGTATTTTTCTCACCAGTC
pbip52_9-1345-V	TAAAAGATGCTTTTTTGAAGATCAGTTATGTATCCGCTTTTTTCATGAGACA
pbip52_9-1353-E	GGTGAAAGTGTTTTTGCTCACCCAGTATTAT
pbip52_9-1198-E	CCCGTATTCTGCTATGTGGCGCGGAAACGCT
pbip52_10-157-E	GCCCCCGACACTAAATCGGAACCCCGCCCTT
pbip52_10-1406-E	ATTCCCTTTCAACATTTCCGTGTTAAAGGGA

pbip52_11-1449-V	AAAAAGGAAGATTTTTGTATGAGTATTTTTGCGGCATTTTTTTTTGCCTTCC
pbip52_11-44-V	GTAACCACCACTTTTTACCCGCCGCGTAACGTGAGTTTTTTTTTCGTTCCACTTGTTAAATCATTTTTGCTCATTTTT
pbip52_11-52-E	CGCTGCGCGGGCGCTGGCAAGTGTGCTTCAA
pbip52_11-1457-E	TAATATTGATAACCCTGATAAATAGCGGTCA
pbip52_12-308-V	AACAAGAGTCCTTTTTACTATTAAAGCGTCTATCAGGTTTTTGCGATGGCCC
pbip52_12-328-E	TGAGTGTTGTTGGCGAGAAAGGAAGGGAAGACGAGATAGGGT
pbip52_12-318-E	TCCAGTTTGGCCGGCGAACG
pbip52_13-202-V	TGGGGTCGAGGTTTTTGCCGTAAAGTTTAGAGCTTGTTTTTACGGGGAAAG
pbip52_13-210-E	AAGTTTTTACTACGTGAACCATCATATTCTC
pbip52_13-1147-E	AGAATGACGGTCGCCGCATACACCCCTAATC
pbip52_14-360-V	TTATAAATCAATTTTTAAGAATAGACAAGCGAAAGGATTTTTGCGGGCGCTA
pbip52_14-679-E	CCTCCCGTCAGCACTGGGGCCAGAAATCGGC
pbip52_14-368-E	AAAATCCCTAACCAATAGGCCGATGGTAAGC
pbip52_15-724-V	GTGGGTCTCGCTTTTTGGTATCATTGATCGTAGTTATTTTTCTACACGACG
pbip52_15-263-E	CGAAAAACAACGTGGACTCCAACGCTGGAGC
pbip52_15-732-E	CGGTGAGCTTATTGCTGATAAATTCAAAGGG

Appendix Table 8. Staples for 84-bp edge-length pentagonal bipyramid nanoparticle from phPB84 sequence.

PB84_st1	GCGAAATTAGACAAAGCTTTAAACAAAATCTG
PB84_st2	AAACTAAGCTTATATGTGTAAAAGGAACTGC
PB84_st3	TATGCAAATTGCACTCTGAATCTGTCTATAC
PB84_st4	CAGGTACGAGCTTTTTGAACGAAACCATAGAATCGGTTTTTTATCCAAAGATCCGTGTATCATTTTTTAATTGTCCGA
PB84_st5	GGGAAAACTCAAATCCTATCTGTGTTTCTTA
PB84_st6	GACGTCAGTCATGATAATAATGGAAGAACCA
PB84_st7	GGTTAATGGTGGCACTTTTCGGGGAAATGTGC
PB84_st8	TGAGCTGGTTATTTTTCTAGGTACATCCGCATAAGTGTTTTTTTGTAGAAAG
PB84_st9	GTAGGAAGGTTAGAAGTTAGAAATTTTGACT
PB84_st10	CGGGTGTACACCTGTATCGGCCTTCAAACGG
PB84_st11	TTGCGCAGATCAGTACTATGCAGTAACCTCAG
PB84_st12	GATAGTTATGATACTCTATGTTCGCCCGCATT
PB84_st13	ATTCATTCATTTTTTTGTGCAATCCTCTAGATTAGAATTTTTGGCGTTGTTT
PB84_st14	CTATATTCACTTTAAAAGAAAAGAGGTTGATACGTATCTTCC
PB84_st15	CCGCTCATGAAAGCATGGTAGCAAGTCGTTAA
PB84_st16	TGACCCAAGAGACAATAACCCTGAATATGTAT
PB84_st17	CCACGTGTAAGGATTAGTGACGTAACTCCCGG
PB84_st18	GTTGGTTTCATGTTACTTGGTCAGTTATCGGA
PB84_st19	GTGATAACGCACACGTCTGGCGGAAGAGAAT
PB84_st20	TGTGTTGCAGGTTTTTTAGCGACAGTAAATATCTTCATTTTTCGACGTAGCG
PB84_st21	ATTCACAAGTCTGGCAATAC
PB84_st22	CCCCAATCTCTTATTACCCTCGATTAGTCACA
PB84_st23	GTCTATTAATATGTACCCCGGTTGGCGTCAGA
PB84_st24	ATAATTATGGGTTCCCAGAGCTATCGAGATA
PB84_st25	GGGTTGAGATCAAAAGAATAGACATTCTGAC
PB84_st26	TATTAATAGAGAAGTGGCTAACGCTGACGGG
PB84_st27	GAAAGCCGCCCCGATTTAGAGCTAAGCAAGA
PB84_st28	AAGGGAGCGCGAACGTGGCGAGAAAGGAAGGG
PB84_st29	GCGCCCGGTTGCGCTCCTGCAGCATACCTATG
PB84_st30	ACAGGGCGCACTTTTTATTAATTGCGGTGGCTCAACATTTTTACAGAGAATA

PB84_st31	GCGCCGCTCTGGCGAACTACTTACTCTAGCGCCGCGCTTAAT
PB84_st32	CACGCTGCGCGTAACCACCACACCCTTCCCGG
PB84_st33	CAACAATTAATAGACTGGATGGAGGTAGCGGT
PB84_st34	GCGGATAAAGGCTGGCAAGT
PB84_st35	GAAAGGAGCGGTTTTTGCGCTAGGGCTTGCAGGACCATTTTTCTTATGCGCT
PB84_st36	AAGAAAGCGTAAAGCACTAAATCGGAACCCTA
PB84_st37	GCCCTCCCGTATTTTTTCGTAGTTATGTTTTTTGGGGTTTTTTCGAGGTGCC
PB84_st38	CTAATCAACTACACGACGGGGAGTCAGGCGTGAACCATCACC
PB84_st39	TCTATCAGGGCGATGGCCCACTACAACTATGG
PB84_st40	ATGAACGAAATAGACAGATCGCTGAAAAACCG
PB84_st41	AGATAGGTGCTCAAAGGGCG
PB84_st42	TAAATCAGCTCTTTTTATTTTTTAACTAAAGAACGTGTTTTTGACTCCAACG
PB84_st43	TCCACTATCAATAGGCCGAAATCGGCAAAATC
PB84_st44	CCTTATAATGTTGTTCCAGTTTGGAACAAGAG
PB84_st45	ATTTTTGTCGTGAGTTTTCGTTCCACTGATAATTCGCGTTAA
PB84_st46	CTCACTGATTATTTTTAGCATTGGTACTCATGACCAATTTTTAATCCCTTAA
PB84_st47	TTGATAATACTGTCAGACCAAGTTTACTCATA
PB84_st48	TATACTTAAAAGGATCTAGGTGAAGATCCTTT
PB84_st49	TTAATTTAAGATTGATTTAAAACTACTTCTG
PB84_st50	GATGGTAACGGCCCTTCCGGCTGGCTGGTTTA
PB84_st51	TTGCTGATGGTATCATTGCAGCACTGGGGCCA
PB84_st52	CCCCGAAGAACTTTTTGTTTTCCAATCAACGTTGCGCTTTTTAAACTATTAA
PB84_st53	AATGGCAAGATGAGCACTTTTAAAGTTCTGCT
PB84_st54	ATGTGGCGGCGTGACACCACGATGCCTGTAGC
PB84_st55	AACGACGACGGTATTATCCCGTATTGACGCC
PB84_st56	GGGCAAGACGGAGCTGAATGAAGCCATACCA
PB84_st57	CGGAGGACCGATTTTTAGGAGCTAACCTCGCCTTGATTTTTTCGTTGGGAACGCAACTCGGTCTTTTTGCCGCATACA
PB84_st58	TCATGTAACGCTTTTTTGCACAACGAGCGTG
PB84_st59	GGTCTCGCAAATCTGGAGCCGGTATGGGGGA
PB84_st60	ACAACGATAACACTGCGGCCAACTTTCATTT
PB84_st61	CAGTCACAGAATTTTTAAGCATCTTAGCTGCCATAACTTTTTCATGAGTGAT

PB84_st62	TATGCAGTCGGATGGCATGACAGTACATCTA
PB84_st63	GTACTCACCTATTCTCAGAATGACGCATTTT
PB84_st64	GCCTTCCTATTCCCTTTTTTGCGTTGGTTGA
PB84_st65	AGAGTTTTCGAGTAAATTAC
PB84_st66	ATCTCAACAGCGGTAAGATCCTTGATGCGATG
PB84_st67	TGACCTTAAACGAATAGCCTCTCCCGAACTGG
PB84_st68	AGGGGACGGGGTGCACGAGTGGGTTACATACCCAAGAGTTTG
PB84_st69	AAAAGATGCTGTTTTTAAGATCAGTTACGACGAATTCTTTTTGACGA AAGGG
PB84_st70	GTGAAAGTGTATGAGTATTCAACATTTCCGTG
PB84_st71	TCGCCCTTGTTTTTGCTCACCCAGAAACGCTG
PB84_st72	CCTATTTGTTTTTTTTATTTTTCTAAAATAATATTGATTTTTAAAAGG AAGA
PB84_st73	TAAATGCTTCATACATTCAA
PB84_st74	GCGGAACCCCTCGTGATACGCCTATTTTTATA