NANOSCALE PHOTONIC DEVICES FABRICATED USING DNA

NANOSTRUCTURES

by

Hieu Trung Bui

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Electrical Engineering

Boise State University

August 2011

© 2011

Hieu Trung Bui

ALL RIGHTS REVERSED

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Hieu Trung Bui

Thesis Title: Nanoscale Photonic Devices Fabricated Using DNA Nanostructures

Date of Final Oral Examination: 14 June 2011

The following individuals read and discussed the thesis submitted by student Hieu Trung Bui, and they evaluated his presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

William B. Knowlton, Ph.D.	Chair, Supervisory Committee
Bernard Yurke, Ph.D.	Member, Supervisory Committee
Wan Kuang, Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by William B. Knowlton, Ph.D., Chair of the Supervisory Committee. The thesis was approved for the Graduate College by John R. Pelton, Ph.D., Dean of the Graduate College.

ACKNOWLEDGMENTS

This thesis would not be possible without the help and support of many individuals. I would like to thank Dr. William B. Knowlton, my graduate and research adviser, who gave me my first research opportunity as an undergraduate. Dr. Knowlton allowed me to continue my graduate studies working in his lab, which helped me grow and learn. His guidance and encouragement has been so valuable, and allowed me to get this thesis done. My special thanks also goes to Dr. Bernard Yurke, my research adviser, who taught and inspired me about science. He is the catalyst behind the ideas of this thesis and provided support to develop and expand on the original concept. I am grateful to Dr. Wan Kuang, Dr. Elton Graugnard, Dr. Jeunghoon Lee, and Dr. William Hughes for all the support amidst the many demands on their time.

It has been my privilege to work with and learn from each and every one of my research assistant fellows, past and present. I especially appreciate Yerpeng Tan, Ross Bulter, Craig Onodera, Carson Kidwell, Glen Purnell, Stephanie Barnes, and Donald Kellis for actively contributing to this work. I thank all the members of the Nanoscale Materials and Device Group, past and present, for making the place as friendly and engaging as it is.

Most importantly, I would like to acknowledge my family for their love and support. My mom and dad have been my foundation. Their sacrifices are immeasurable. I

iv

thank Trang Phuong Le for her love and support through this entire process. She is the driving force in my life; I could not have asked for a better companion.

The materials in Chapter 4, Appendix D, and Appendix E have been reported in the following article: Hieu Bui, Craig Onodera, Carson Kidwell, YerPeng Tan, Elton Graugnard, Wan Kuang, Jeunghoon Lee, William B. Knowlton, Bernard Yurke, William L. Hughes, "Programmable Periodicity of Quantum Dot Arrays with DNA Origami Nanotubes," Nano Letters 10, no. 9 (2010): 3367-3372. Link to the article can be found in the following: <u>http://pubs.acs.org/doi/abs/10.1021/nl101079u</u> or DOI: 10.1021/nl101079u.

ABSTRACT

In the field of DNA nanotechnology, self-assembly is being advanced as the key technology for the creation of nanoscale structures. Popular and effective DNA nanotechnology methods of producing nanoscale structures are branched DNA junctions and DNA origami. DNA nanostructures have recently been employed as scaffolds for the bottom-up arrangement of proteins, as well as semiconductor and metallic nanoparticles, with nanometer precision. Such structures are expected to exhibit unique optical properties and may enable new photonic devices. Conversely, the majority of photonic devices for optical waveguide are fabricated using top-down processes. However, the cost and controllability of complex nanostructures using top-down processes imposes significant challenges. As an alternative to top-down processes, work will be presented demonstrating the use of DNA self-assembly processes to fabricate nanoscale photonic devices for optical waveguide. To fabricate photonic devices using DNA self-assembly, DNA nanostructures were used as scaffolds to configure light emitting molecules so as to create a near-field energy transfer waveguide. Spectrophotometry was used to characterize the device operation. The spectral results indicate that DNA nanostructures functionalized with light emitting molecules can transfer energy through a three molecule system with 28% efficiency, demonstrating the potential of using DNA nanostructures for future photonic devices for optical waveguide.

vi

TABLE OF CONTENTS

ACKNOWLEDGMENTS iv
ABSTRACTvi
LIST OF FIGURES
LIST OF ABBREVIATIONS/SYMBOLS
CHAPTER 1: INTRODUCTION 1
CHAPTER 2: DNA TILE DIFFUSIVE WAVEGUIDE
2.1 Experimental
2.1.1. Background on DNA
2.1.2 Design
2.1.3 Materials and Methods15
2.2 Results and Discussion
CHAPTER 3: DNA ORIGAMI NANOTUBE DIFFUSIVE WAVEGUIDE
3.1 Experimental
3.1.1 Design
3.1.2 Materials and Methods
3.2 Results and Discussion
CHAPTER 4: DNA ORIGAMI NANOTUBE QUANTUM DOT ARRAYS 44
4.1 Experimental
4.1.1 Design

4.1.2 Materials and Methods 45
4.2 Results and Discussion 46
CHAPTER 5: SUMMARY & FUTURE WORK 52
REFERENCES
APPENDIX A
FRET Background and Fluorophores
APPENDIX B
PAGE Gel Filtration and Agarose Gel Filtration
APPENDIX C
A Least-Squares Curve Fitting73
APPENDIX D
6-Helix Bundle Nanotube Design, Sequence Generator, Sequence List, Position of Biotin-Labeled Staple Strands, and Statistical Analysis of Quantum Dot Arrays76
APPENDIX E
Statistical Analysis of Quantum Dot Arrays
APPENDIX F
Experimental Equipment93

LIST OF TABLES

Table 1.1 Electromagnetic spectrum ¹	. 1
Table 2.1: Measured excitation and emission maxima of the dyes used in this work	14
Table 2.2: DNA sequences used to construct dye-labeled DNA tiles	15
Table A.1: Absorbance (Abs) and emission (Em) maxima of three fluorophores from different buffer conditions	66
Table D.1: Name and sequence for the 170 staple strands used for the 6-helix DNA nanotube	83

LIST OF FIGURES

 Fig.1.1: Schematics of photonic devices for optical wave guiding using DNA duplex approaches by (a) Haustein <i>et al.</i>³ (b) Ohya <i>et al.</i>⁴ (c) Vyawahare <i>et al.</i>⁵ (d) Heilemann <i>et al.</i>⁶. The dots with different colors represent the different fluorophores which are incorporated into the duplex DNA
Fig.1.2: (a) Schematic of an immobile four-arm DNA junction; (b) Formation of two- dimensional lattice from a four-arm junction with sticky ends, X and Y are sticky ends, X' and Y' are their complementary sticky ends, respectively.
Fig.1.3: Schematic of a DNA origami arbitrary shape adopted from Ref.17. The scaffold is depicted in a long black strand, whereas the staples are displayed in short colored strands
Fig.1.4: Arrangement of fluorophores (different colored dots) on the DNA origami rectangle adopted from Ref.19
Fig.2.1: A schematic showing the three major constituents of a nucleotide: a nitrogenous base (adenine), a pentose sugar (ribose), and one or more phosphates. The schematic is adopted from Ref.20
Fig.2.2: Schematics of the nitrogenous bases that distinguish DNA nucleotides. (a) Purines - adenine and guanine; (b) Pyriminides - thymine and cytosine. These schematics were adopted from Ref.20
Fig.2.3: (a) A diagrammatic view of DNA double helix. The sugar-phosphate backbones of the double helix are represented by colored ribbons. The bases attached to the sugar deoxyribose are on the inside of the helix. (b) An enlarged view of two base pairs. Note that the two DNA strands run in opposite directions defined by the 5' and 3' groups of deoxyribose. The bases on opposite strands form pairs because of hydrogen bonds. Cytosine pairs with guanine; thymine pairs with adenine ²⁰
Fig.2.4: Dye-labeled (dots) strands within the tile. FAM is attached to the left strand (labeled F), TAM is attached to the middle strand (labeled T), and Cy5 is attached to the right strand (labeled C). X, Y, and Z strands are illustrated as the straight lines that act as the scaffold strands. Arrowheads indicate 3' end of strands

Fig.2.5: The s	pectral overlap between fluorophores (shaded areas) are shown with excitation and emission spectra of FAM (solid), TAM (dashed), and Cy5 (dotted) fluorophores
Fig.2.6: Intera	ction of two FRET pairs (i.e., left and middle fluorophore pair and middle and right fluorophore pair) demonstrating double FRET. Upon excitation, the first donor, FAM, relaxes to the ground state, thus exciting the center fluorophore, TAM, which first acts as an acceptor. This fluorophore then becomes the donor of the second FRET pair that transfers the energy to the third fluorophore, Cy5. Ideally, only the last fluorophore relaxes by emitting a longer wavelength photon ³
Fig.2.7: Schen	natic of various expected and observed fluorescence emission spectra in the three fluorophore system; λ_1 , λ_2 , and λ_3 are the emission peaks of the input, intermediate, and output fluorophores, respectively
Fig.2.8: FRET	remission spectrum from the tiles with all three fluorophores. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)
Fig.2.9: FRET	r emission spectrum from the tiles missing a TAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)
Fig.2.10: FRE	T emission spectrum from the tiles missing a FAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)
Fig.2.11: FRE	T emission spectrum from the tiles missing a Cy5 fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)
Fig.2.12: TAN	A fluorescence from DNA tiles containing FAM-TAM (FT) and FAM- TAM-Cy5 (FTC). Each bar in the graph was created using five trial measurements from the same solution. The TAM fluorescence from emission spectra was normalized for molecular concentration. There is a considerable drop in TAM fluorescence when the Cy5 is present, indicating FRET behavior
Fig.3.1: Dye-l	abeled (dots) staple strands within the nanotube. FAM is attached to the left staple strand (labeled F), TAM is attached to the middle strand (labeled T), and Cy5 is attached to the right strand (labeled C)
Fig.3.2: AFM	height image acquired under ambient conditions for DNA origami nanotubes deposited on an atomically flat mica surface (a); a nanotube's length (b) and diameter (c) of the dotted circle

Fig.3.3: FRET	Temission spectrum from the nanotubes with all three fluorophores. The inset illustrates a schematic of the structure. (Note: excitation wavelength $= 480 \text{ nm}$)
Fig.3.4: FRET	remission spectrum from the nanotubes missing TAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)
Fig.3.5: FRET	r emission spectrum from the nanotubes missing FAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)
Fig.3.6: FRET	r emission spectrum from the nanotubes missing Cy5 fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)
Fig.3.7: TAM	fluorescence from DNA origami nanotubes containing FAM-TAM (FT) and FAM-TAM-Cy5 (FTC). Each bar in the graph was made using five trial measurements from the same solution. The TAM fluorescence from emission spectra was normalized for molecular concentration. There is a considerable drop in TAM fluorescence when the Cy5 is present, indicating FRET behavior
Fig.4.1: Schen	natics, AFM images at low magnification (upper) and high magnification (lower), and cross-sectional (upper) and axial (lower) height profiles of functionalized DNA origami nanotubes with 9 biotin binding sites with: (a-e) no attached nanoparticles; (f-j) attached streptavidin; (k-o) attached streptavidin-conjugated quantum dots. The dashed lines in the high magnification AFM images indicate the location of the cross-sectional profiles. Axial profiles represent the average of multiple profiles across the width of the nanotube. (Reprinted with permission from Ref.22. Copyright 2010 American Chemical Society)
Fig.4.2: High	magnification AFM images of streptavidin-conjugated quantum dots attached to functionalized DNA origami nanotubes with: (a) 5 binding sites, 71 nm period; (b) 9 binding sites, 43 nm period; (c) 15 binding sites, 29 nm period; and (d) 29 binding sites, 14 nm period. All scale bars are 100 nm. Note (c) and (d) have fewer attached quantum dots than available binding sites. In addition, the diameter of quantum dots varies between images because of variation in tip radii between scans. (Reprinted with permission from Ref.22. Copyright 2010 American Chemical Society).
Fig.A.1: Illust	ration of the dipole-dipole coupling between two transition dipole moments of two given fluorophores and typical values of the factor (κ^2) for specific dipole orientations ³⁰

Fig.A.2: Transitie th	on energy diagram of the FRET process. The energy is corresponding to e vertical direction
Fig.A.3: Illustrati	ion of variation in energy transfer efficiency and the corresponding uorescence emission spectrum of the two fluorophores system
Fig.A.4: Spectral	characteristics of a fluorophore
Fig A.5: Molecul	lar structures of FAM, TAM, and Cy5; the three fluorophores used in the onstruction of DNA-based diffusive waveguides ²¹
Fig.B.1: 10% PA	GE gel in 1x TAE Mg ²⁺ dye-labeled DNA tiles with (1) FAM, (2)
TA	AM, (3) Cy5, (4) FAM-TAM, (5) FAM-Cy5, (6) TAM-Cy5, and (7)
FA	AM-TAM-Cy5
Fig.B.2: 2% agar	rose gel in 1x TAE Mg ²⁺ dye-labeled nanotubes with (1) FAM, (2) TAM
(3	b) Cy5, (4) FAM-TAM, (5) FAM-Cy5, (6) TAM-Cy5, and (7) FAM-
T ₂	AM-Cy5. The majority of DNA nanotubes migrate as a single band in
ag	garose-gel electrophoresis. This population presumably represents well
fo	formed nanotubes, whereas slower migrating species apparent on the gel
pr	resumably represent misfolded structures
Fig.D.1: Two-dir	mensional layout of the scaffold and staple strands of the DNA nanotube
an	and 3D schematic. (a) Layout of the scaffold showing nucleotide numbers
at	the crossovers. (b) Staple layout for the left end of the tube. The staple
m	otif is shown in columns 11-13. In helix 3, staples in columns 4, 7, and
10	are extended with sticky-ends labeled A, B, and C. (c) Staple layout in
th	e middle of the tube. The M13mp18 scaffold begins and ends in helix 1,
co	plumn 43. Scaffold crossovers are located at the ends and in columns 39
an	and 41. A, B, and C sticky-ends are added to staples 41, 45, and 48 of
he	elix 3. (d) Staple layout for the right end of the tube. Sticky-ends are
ad	lded to staples in helix 3 in columns 77, 80, and 83. Four nucleotides
re	main at the end of each helix. (e) Schematic of the formed tube
ill	lustrating the A, B, and C sticky-ends along helix 3 of the formed
na	anotube. (Reprinted with permission from Ref.22. Copyright 2010
At	merican Chemical Society)
Fig.D.2: Biotin-la	abeled DNA origami nanotube arrays (a) 29 particles, (b) 15 particles,
(c) 9 particles, (d) 5 particles
Fig.E.1: Histogra	ams (bars) and calculated binomial distributions (lines) for the number of
at	tached quantum dots for DNA nanotubes with (a) 5, (b) 9, (c) 15, and (d)
29	biotin binding sites. Data for each histogram were compiled from AFM
in	hage analysis for over 225 separate nanotubes, with the exact number, N,
sh	hown for each histogram. The average attachment probabilities, p, used
to	generate the calculated binomial distributions are indicated for each

case. (Reprinted	with permission	from Ref.22.	Copyright 2010 A	merican
Chemical Society	y)			89

ig.E.2: Histograms (bars) and calculated geometric distributions (lines) for nearest-
neighbor (N-N) separation of bound quantum dot pairs for DNA
nanotubes with (a) 5, (b) 9, (c) 15, and (d) 29 biotin binding sites. The
numbers of separations. N. measured for each case are provided in the
figures along with the average attachment probabilities n N-N
separation of zero indicates two nearest neighbors with a separation less
then one helf of a period. (Deprinted with permission from Def 22
than one-nall of a period. (Reprinted with permission from Ref.22.
Copyright 2010 American Chemical Society)
ig.F.1: Eppendorf Centrifuge 5418, used for filtering DNA solutions
ig.F.2: Eppendorf Mastercycler Personal, used for annealing DNA solutions
ig.F.3: Hoefer gel electrophoresis apparatus, used for purifying DNA solutions
ig.F.4: Agilent Varian Spectrophotometry, used for measuring fluorescence
ig E.5: Veeco atomic force microscope multimode used for characterizing topography
of DNA nanostructures
$\mathbf{U} \mathbf{D} \mathbf{U} \mathbf{U} \mathbf{U} \mathbf{U} \mathbf{U} \mathbf{U} \mathbf{U} U$

LIST OF ABBREVIATIONS/SYMBOLS

DNA	deoxyribonucleic acid
FRET	Förster resonance energy transfer
AFM	atomic force microscopy
FAM	6-carboxyfluorescein
TAM	carboxyl-tetramethyl-rhodamine
Cy5	cyanine
nm	nanometer
λ	wavelength
<i>h</i>	
υ	spatial frequency
<i>eV</i>	electron-volt
<i>N</i> _A	Avogadro's number
μ	dipole moment
<i>R</i> ₀	Förster radius
<i>E</i>	energy transfer efficiency
τ	fluorescence lifetime
<i>n</i>	refractive index
Φ	quantum yield

CHAPTER 1: INTRODUCTION

The goal of photonic devices is to use and control the interaction of light with matter to perform functions; it is analogous to the use and control the interaction of electrons with matter-to-perform functions in electronic devices. A wave packet of light, or a photon, is electromagnetic radiation of frequencies in the range from 1 THz to 10 PHz, corresponding to wavelengths between ~300 μ m and ~30 nm in free space. This electromagnetic range is generally divided into infrared, visible, and ultraviolet regions, as indicated in Table 1.1. The primary interest in the applications of photonic devices is in a narrow range of visible and near-infrared wavelengths. This spectral range for applications is largely determined by the properties of materials used for photonic devices¹.

Wave region	Frequency	Wavelength	Devices
Radio	kHz-MHz-GHz	km-m-cm	Electronic devices
Microwave	1 GHz-1 THz	300 mm-300 µm	Microwave devices
Optical			
Infrared	1 THz-430 THz	300 µm-700 nm	
Visible	430 THz-750 THz	700 nm-400 nm	Photonic devices
Ultraviolet	750 THz-10 PHz	400 nm-30 nm	
X-ray	10 PHz-10 EHz	30 nm-300 pm	
Gamma ray	10 EHz and above	300 pm and shorter	

Table 1.1	Electromagnetic spectrum ¹
-----------	---------------------------------------

Current photonic devices for optical wave guiding can nearly achieve nanometer size by using advanced top-down fabrication methods. However, top-down processes are

costly and increasingly complex. Advanced techniques that can attain nanometer size with reasonable costs will play an important role in the production of nanoelectronic, nanomechanical, and nanooptical devices. Unlike top-down techniques, which create sophisticated devices by etching individual molecules away from bulk materials, bottom-up techniques exploit molecular self-assembly by specific placement of individual molecules and requires less time and effort to achieve².

One bottom-up approach with the potential of lower cost and synthesis time for fabricating photonic devices for optical wave guiding at nanometer scale is DNA selfassembly. Using DNA self-assembly, nanostructures are formed with sub-10 nanometer resolution. Recently, several groups have reported methods of fabricating photonic devices for optical wave guiding using duplex DNA as scaffolds and multiple fluorophores as light transmitting elements, which are illustrated in Fig.1.1. The photon energy is selectively injected into one end of the duplex DNA and it is detected at the opposite end. The transport of photon energy is possible due to the fluorescence (or Förster) resonance energy transfer (FRET) between fluorophores with spectral overlap and close proximity (typically 2 - 10 nm); the FRET process is discussed in detail in Appendix A.1. For these wave guiding devices (hereafter referred to as FRET-based waveguide), the efficiency is in the range from 17% to $40\%^{3,4,5,6}$. It is worth noting here that a key distinction between FRET-based waveguides and traditional waveguides is the way in which photon energy propagates in the devices. The photon energy is diffusively transferred in FRET-based waveguides, whereas it is coherently transferred in traditional waveguides⁵.



Fig.1.1: Schematics of photonic devices for optical wave guiding using DNA duplex approaches by (a) Haustein *et al.*³ (b) Ohya *et al.*⁴ (c) Vyawahare *et al.*⁵ (d) Heilemann *et al.*⁶. The dots with different colors represent the different fluorophores which are incorporated into the duplex DNA.

Unlike traditional waveguides, the FRET-based waveguides discussed in this thesis are built upon DNA oligomers. Current synthetic DNA oligomers are limited to roughly 200 nucleotides in length. Although such oligomers could in principle be used to create a FRET-based waveguide, however, the persistence length would be limited to 50 nm⁷. Such a duplex DNA waveguide would be difficult to couple with larger electrodes and would exhibit a high degree of curvature, reducing the effective length of the waveguide. In order to extend beyond the size limitations imposed by current synthetic DNA oligomers and create longer FRET-based waveguides, a much longer DNA scaffold is needed.

One approach to fabricating a longer DNA scaffold can be realized by incorporating more than a single duplex DNA into the design structure. One example is an immobile DNA branch junction⁸ which has been demonstrated by Ned Seeman, as shown in Fig.1.2 (a). This junction is a stable analogue of the Holliday junction⁹. By adding "sticky ends" to the four-arm junction, a two-dimensional lattice¹⁰ can be formed, as illustrated in Fig.1.2 (b); a sticky end is a non-binding nucleotide in the DNA structure but it can potentially bind to any complementary nucleotide outside of its own structure. The use of branched intermediates allow one to form connected structures from DNA^{11,12}, as well as periodic^{13,14} and aperiodic^{15,16} arrays. Using branched DNA junctions, multiple fluorophores can be easily incorporated into the design to overcome the limitation of a single duplex DNA.



Fig.1.2: (a) Schematic of an immobile four-arm DNA junction; (b) Formation of two-dimensional lattice from a four-arm junction with sticky ends, X and Y are sticky ends, X' and Y' are their complementary sticky ends, respectively.

Another approach to overcome the limitations of a single duplex DNA scaffold is a method described by Paul Rothemund in 2006. The method he called DNA origami¹⁷ produces nanoscale DNA structures using numerous short "staple" strands of DNA to direct folding of a long "scaffold" strand into a flat array of antiparallel helices, as shown in Fig.1.3. A "scaffold" strand is a long, single-stranded genomic DNA, obtained from a bacteriophage M13mp18, which is harvested from infected bacteria. "Staple" strands are synthetic oligonucleotides, typically less than 100 nucleotides in length, that hybridize with the scaffold strand in strategic locations. One half or end of the staple strand hybridizes in one location while the other half hybridizes in another location. As many as 170 staple strands can interact with the scaffold strand in this manner, which will eventually fold the DNA into a desired shape (i.e., DNA origami). Since its introduction, the use of DNA origami has grown dramatically. Currently, DNA origami can provide not only arbitrary 2D nanostructures but also nano-sized breadboards, a term coined by Bernard Yurke (as cited in Ref.17), for the arraying of nanomaterials and 3D nanostructures, such as hollow polyhedrons or even more complicated nano-objects¹⁸.



Fig.1.3: Schematic of a DNA origami arbitrary shape adopted from Ref.17. The scaffold is depicted in a long black strand, whereas the staples are displayed in short colored strands.

Realizing the advantages of DNA origami, one of which is the nano-sized breadboards for arraying nanomaterials, Stein *et al.* reported the first construction of FRET-based waveguides using DNA origami rectangles they decorated with several different fluorophores¹⁹, as illustrated in Fig.1.4. Using a single-molecule technique, they successfully demonstrated that the photon energy was diffusively transferred with up to 36% efficiency.



Fig.1.4: Arrangement of fluorophores (different colored dots) on the DNA origami rectangle adopted from Ref.19.

To overcome the limitation of a single duplex DNA and to take advantage of both branched DNA junctions as well as DNA origami, this thesis explains the efforts towards the design of FRET-based waveguides and the investigation of the device's operation by optical means. Specifically, Chapter 2 reports the design of DNA tiles as scaffolds for the arrangement of fluorophores and the investigation of FRET on dye-labeled DNA tiles. Chapter 3 describes the design of DNA origami nanotubes as scaffolds for the arrangement of fluorophores and the investigation of FRET on dye-labeled DNA origami nanotubes. Chapter 4 explains the design of DNA origami nanotubes as scaffolds for the arrangement of semiconductor nanoparticle arrays and the surface topography characterization of nanoparticle arrays via atomic force microscopy (AFM).

CHAPTER 2: DNA TILE DIFFUSIVE WAVEGUIDE

2.1 Experimental

The experimental section is organized in the following manner: (1) a brief background on DNA is discussed, (2) an explanation of the design procedure is given, and (3) materials and methods used to characterize the design are described.

2.1.1. Background on DNA

It is useful to provide a brief background on the molecular structure of DNA prior to examining DNA design. The following provides this information. Nucleotides are molecules that when joined together form the structural units of RNA and DNA. Each nucleotide is comprised of approximately 20 atoms, such as carbon, nitrogen, and oxygen, and contains three components: a five-carbon sugar (either ribose or deoxyribose), a nitrogenous base, and one or more phosphate groups as shown in Fig.2.1. The nitrogenous bases are heterocyclic aromatic rings with a variety of substituents²⁰. There are two classes of base as shown in Fig.2.2: the bicyclic purines (a) and the monocylic pyrimidines (b). A chain of a definite nucleotide sequences is called a singlestranded DNA or a DNA oligomer with spacing between nucleotides of ~ 0.43 nm. An oligomer can be synthesized or harvested from an infected bacteriophage (i.e., M13mp18).



Fig.2.1: A schematic showing the three major constituents of a nucleotide: a nitrogenous base (adenine), a pentose sugar (ribose), and one or more phosphates. The schematic is adopted from Ref.20.



Fig.2.2: Schematics of the nitrogenous bases that distinguish DNA nucleotides. (a) Purines - adenine and guanine; (b) Pyriminides - thymine and cytosine. These schematics were adopted from Ref.20.

Formation of a DNA double helix occurs when a single-stranded oligomer binds to its complementary sequence through hydrogen bonding between base pairs. The most common binding between bases occurs when adenine binds to thymine (A-T) or cytosine binds to guanine (C-G), shown in Fig.2.3. In addition to possessing a complementary sequence, formation of a double helix requires the two sequences to be formed with opposite directionality, as determined by the phosphate bonds that connect the deoxyribose sections of each nucleotide. In the form of DNA used in this work, the DNA double helix rotates through 360° every 10.5 nucleotides with a spacing between nucleotides of ~0.34 nm and double helix diameter of ~2 nm. The angular separation of adjacent nucleotides in a DNA double helix is ~ 34° .



Fig.2.3: (a) A diagrammatic view of DNA double helix. The sugar-phosphate backbones of the double helix are represented by colored ribbons. The bases attached to the sugar deoxyribose are on the inside of the helix. (b) An enlarged view of two base pairs. Note that the two DNA strands run in opposite directions defined by the 5' and 3' groups of deoxyribose. The bases on opposite strands form pairs

because of hydrogen bonds. Cytosine pairs with guanine; thymine pairs with adenine²⁰.

2.1.2 Design

In the design assessment process, designs were evaluated based on the following design constraints: (1) the design simplicity, (2) the ability to control the distance between fluorophores, (3) the ease to optimize the device operation, (4) the ability to perform troubleshooting as a design feedback mechanism, and (5) the design flexibility for future enhancement. The subsequent section examines the designs that were ultimately chosen using this design assessment approach.

Using the sequence-dependent binding properties of DNA (i.e., G binds to C, A binds to T), a DNA tile was designed consisting of three parallel double helices bound together with six crossovers, as illustrated in Fig.2.4. The tile is composed of six strands; a strand is a short segment of DNA composed of known nucleotide sequences. Three straight strands (i.e., X, Y, and Z) are composed of 42 nucleotides and are used as scaffold strands; recall that a scaffold strand is the backbone for forming DNA nanostructures. Three strands are in the shape of an S (i.e., F, T, and C) are composed of 42 nucleotides and used as staple strands, which is a small segment of DNA for holding the DNA nanostructures in place. The staple strands are programmed to be complementary to three separate 14-nucleotide regions of the scaffold strand. It is predicted that the staple strands self-assemble with the scaffold strand into the tile shape of three parallel double helices. Arrowheads indicate the three prime (3') end of strands where the "3' end" is the third carbon positioning in the sugar structure of DNA. In its

ideal form, the tile dimension is 14.28 nm x 6 nm, assuming a spacing of 0.34 nm per base pair.



Fig.2.4: Dye-labeled (dots) strands within the tile. FAM is attached to the left strand (labeled F), TAM is attached to the middle strand (labeled T), and Cy5 is attached to the right strand (labeled C). X, Y, and Z strands are illustrated as the straight lines that act as the scaffold strands. Arrowheads indicate 3' end of strands.

For studying FRET-based waveguides, DNA tiles were used to construct a stable sequential arrangement of fluorescent dyes separated by a regulated distance. The F, T, and C strands were labeled with three different fluorophores: 6-carboxyfluorescein (FAM), carboxy-tetramethyl-rhodamine (TAM), and cyanine (Cy5), respectively. The three fluorophores were chosen such that they formed a linear chain in the tile, ordered in terms of absorption energy from the primary donor (FAM) to the final acceptor (Cy5) via the intermediate acceptor-donor (TAM). The expected distance between fluorophores is 2.38 nm, equivalent to the distance of a chain of 7 nucleotides.

Fig.2.5 illustrates the spectral overlap (not to scale) between fluorophores and the spectral characteristics of each fluorophore used in this study. For the FAM-TAM fluorophore pair, the excitation spectrum of TAM (dark, dashed) partially overlaps with

the emission spectrum of FAM (gray, solid). The excitation spectrum of Cy5 (dark, dotted) coincides in part with the emission spectrum of TAM (gray, dashed). The excitation (dark) and emission (gray) spectra for each fluorophore are also plotted in Fig.2.5. The maximum wavelength values of excitation and emission spectra for FAM, TAM, and Cy5 are listed in Table 2.1.



Fig.2.5: The spectral overlap between fluorophores (shaded areas) are shown with excitation and emission spectra of FAM (solid), TAM (dashed), and Cy5 (dotted) fluorophores.

Table 2.1: Measured excitation and emission maxima of the dyes used in this work

	Dye-labeled oligomer			
	Excitation peak (nm)	Emission peak (nm)		
FAM	496	517		
TAM	561	583		
Cy5	656	667		

2.1.3 Materials and Methods

The dye-labeled strands and their complementary strands were purchased from Integrated DNA Technologies²¹ and the sequences are listed in Table 2.2. Fluorophores were attached as follows: FAM to the 3' end of strand F; TAM to the 7th base of strand T; Cy5 to the 28th base of strand C. All strands were also available unlabeled, which allowed for the construction tiles with different fluorophore combinations. The other dyelabeled DNA tiles were synthesized for control experiments; the schematics are not shown.

Name	Sequence (5' to 3')	Length (bp)
F	TGA CAA CAA CCA TCG GCT TGA GAT GGT TAA GCG AAC CAG ACC	42
Т	CGA GTA GTA AAT TGG CCC ACG CAT AAC CAG AGG CTG AGA CTC	42
С	ATC AGT AGC GAC AGA CAT GAA AGT ATT AGA TAT ATT CGG TCG	42
Х	CGA CCG AAT ATA TCC AAT TTA CTA CTC GGG TCT GGT TCG CTT	42
Y	AAC CAT CTC AAG CCG GTT ATG CGT GGG CTA ATA CTT TCA TGT	42
Ζ	CTG TCG CTA CTG ATG AGT CTC AGC CTC TGA TGG TTG TTG TCA	42

Table 2.2: DNA sequences used to construct dye-labeled DNA tiles

To synthesize DNA tiles containing FAM, TAM, and Cy5, equimolar amounts of the DNA strands were mixed in a solution of $1 \times TAE$, Mg²⁺ (40 mM tris, 20 mM acetic acid, 2 mM ethylenediaminetetracetic acid [EDTA], and 12.5 mM magnesium acetate; pH 8.0). TAE, magnesium acetate tetrahydrate, and laboratory grade water [Milli-Q Water, Millipore] were purchased from Sigma Aldrich). The solution was annealed by heating the samples to 90 °C for 20 min, followed by a slow cooling to room temperature (~ 2.0 h) using a thermal cycler (Eppendorf Mastercycler Personnal). To remove malformed dye-labeled DNA tiles, gel electrophoresis was employed; the detail of experimental approach is discussed in Appendix B.1. To measure the emission of fluorophores, fluorescence measurements were performed using a Cary Eclipse fluorescence spectrophotometer (Agilent Inc.). A 100 μ L of a 100 nM DNA tile solution was placed in a 350 μ L special optical glass (SOG) cell with 3 mm path length and kept at ambient temperature during the measurement. All emission spectra were collected using an excitation wavelength of 480 nm. The spectral bandwidths for excitation and emission monochromators were both chosen to be 10 nm. The emission spectra of the dye-labeled DNA tiles were recorded over the range of 500 to 800 nm.

2.2 Results and Discussion

To understand the energy transport in dye-labeled DNA tiles, the transition energy diagram of two FRET pairs (i.e., left and middle fluorophore pair and middle and right fluorophore pair) is illustrated in Fig.2.6, assuming that the distance between fluorophores is less than 10 nm. Note that single FRET occurs between a pair of fluorophores, whereas multiple FRET (e.g., double and triple FRET) occurs between multiple pairs of fluorophores (e.g., two FRET pairs and three FRET pairs). Fig.2.6 illustrates double FRET. Upon excitation from an incident photon with energy greater than the energy band gap of the first donor, a single FRET event occurs when the first donor (FAM) quickly relaxes to the ground state, releasing energy to and subsequently exciting the center fluorophore (TAM), which first acts as acceptor. This fluorophore in turn becomes the donor of the second FRET pair that transfers the energy to the third fluorophore (Cy5). Under optimal conditions, only the last fluorophore relaxes by emitting a longer wavelength photon (due to the Stokes shift effect²). For dye-labeled DNA tiles, the FAM fluorophore acts as an optical input, absorbing the incident radiation

and transferring it as an excitation energy into the waveguide. The TAM fluorophore functions as a diffusive optical transmission channel between FAM (the input fluorophore) and Cy5 (the output fluorophore). Here, the term "diffusive" is used to describe the optical energy transfer due to its gradient (i.e., the energy is not conserved). The efficiency of energy transmission through the waveguide can be assessed by exciting the FAM fluorophore and measuring the emission from the Cy5 molecule.



Fig.2.6: Interaction of two FRET pairs (i.e., left and middle fluorophore pair and middle and right fluorophore pair) demonstrating double FRET. Upon excitation, the first donor, FAM, relaxes to the ground state, thus exciting the center fluorophore, TAM, which first acts as an acceptor. This fluorophore then becomes the donor of the second FRET pair that transfers the energy to the third fluorophore, Cy5. Ideally, only the last fluorophore relaxes by emitting a longer wavelength photon³.

In the three-fluorophore system designed for the studies described in this thesis, the fluorescence emission spectra are measured and used to calculate the performance of the FRET system. In theory, FRET is a radiationless energy transfer process, so ideally, a single emission spectrum of the output fluorophore should be observed, as illustrated in Fig.2.7 ("Ideal FRET"), this spectrum corresponds to hundred percent energy transfer efficiency. However, it is most typical to observe an emission spectrum with the combination of the input, intermediate and output fluorophores, as illustrated in Fig.2.7 ("Typical FRET"), this spectrum corresponds to imperfect energy transfer efficiency as some of the energy is lost via emission of radiation. In addition, Fig.2.7 illustrates two other cases in which no energy transfer should occur as labeled "No FRET." In one case, the three-fluorophore system has a missing input fluorophore, as indicated with a dashed black line; this spectrum results in no emission. In the other case, the three-fluorophore system has a missing intermediate fluorophore, as indicated by the blue curve; this spectrum results a single emission spectrum of the input fluorophore because the energy transfer path is disrupted due to the missing intermediate fluorophore.



Fig.2.7: Schematic of various expected and observed fluorescence emission spectra in the three fluorophore system; λ_1 , λ_2 , and λ_3 are the emission peaks of the input, intermediate, and output fluorophores, respectively.

FRET-based waveguides were successfully synthesized using DNA tiles

containing a sequential chain of FAM, TAM, and Cy5 fluorophores. All control devices

were also synthesized with various fluorophore combinations. The solution of DNA tiles containing FAM, TAM, and Cy5 fluorophores was analyzed by using spectrophotometry. The solution was illuminated at 480 nm (2.58 eV), an excitation wavelength of FAM fluorophore. The bulk fluorescence shows three peaks in the emission spectrum, as illustrated in Fig.2.8. These peaks correspond to the emission of FAM fluorophore (left peak at 517 nm or 2.40 eV), TAM fluorophore (middle peak at 583 nm or 2.13 eV), and Cy5 fluorophore (right peak at 667 nm or 1.86 eV).



Fig.2.8: FRET emission spectrum from the tiles with all three fluorophores. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

Assuming perfect FRET along the fluorophore chain, only the emission spectrum of the last fluorophore (i.e., Cy5) should be observed. Instead, an emission spectrum with three peaks is observed. Since only the FAM is efficiently excited by 480 nm (direct absorption from the other fluorophores at 480 nm is negligible), the resulting three emission peaks must exclusively result from two successive FRET events between the fluorophores composing the structure, as observed previously with duplex DNA⁵. For

each fluorophore, a portion of the excitation energy is emitted into free space (hence, the three emission peaks), and a portion is directly transferred to the adjacent fluorophore. The partial emission of the excitation energy is an indication of an inefficiency in the FRET process.

In order to determine the overall efficiency of direct energy transfer from FAM to Cy5, a control experiment was conducted using DNA tiles missing the intermediate TAM fluorophore. It is predicted that the TAM fluorophore should in theory absorb energy from the FAM fluorophore and transfer its absorption energy to the Cy5 fluorophore. By removing the TAM fluorophore, the peak at 583 nm (Fig.2.8) should not appear. The schematic of the fluorophore-labeled DNA tile is illustrated in the inset of Fig.2.9. The solution of DNA tiles containing only FAM and Cy5 (i.e., without TAM) was illuminated at 480 nm by the incident excitation light source. As expected, the bulk fluorescence does not show an emission peak of TAM, as indicated in Fig.2.9. Instead, the bulk fluorescence shows a dominate peak by the FAM fluorophore at 517 nm and a small peak by the Cy5 fluorophore at 667 nm, as illustrated in Fig.2.9. With the absence of the TAM fluorophore, most of the FAM excitation energy is emitted into free space while an insignificant percentage of the FAM fluorophore energy transfers to the Cy5 fluorophore. If FAM was substantially exciting Cy5 (i.e., FRET was occurring), then the Cy5 peak would be similar to the peak in Fig.2.8, demonstrating that direct excitation of Cy5 is taking place. Because the Cy5 emission peak is significantly reduced in this experiment, there is little direct excitation from FAM to Cy5. Hence, the exclusion of TAM demonstrates the disruption of FAM to Cy5 FRET.



Fig.2.9: FRET emission spectrum from the tiles missing a TAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

Another control experiment was performed using DNA tiles excluding the input FAM fluorophore for several reasons. The first reason is to determine the extent to which the FAM acts as the energy input channel. The second reason is to eliminate the possibility of direct excitation of the TAM or Cy5 by the 480 nm incident excitation light source. It is predicted that the FAM fluorophore, after absorbing an incident photon, should transfer energy to the TAM fluorophore. By removing the FAM fluorophore and using the same FAM excitation light source at 480 nm, the peak at 517 nm (Fig.2.8), should not appear. Additionally, both the excitation of the TAM and FRET from TAM to Cy5 should be significantly interrupted resulting in minimal fluorescence peaks at 583 nm and 667 nm. Experimentally, this was confirmed. The solution was illuminated at 480 nm, the excitation wavelength of FAM. As expected, the bulk fluorescence does not show
an emission peak of FAM, as indicated in Fig.2.10. Instead, the bulk fluorescence shows two insignificant peaks of TAM fluorophore at 583 nm and of Cy5 fluorophore at 667 nm, as illustrated in Fig.2.10. These results indicate that (1) without the input FAM fluorophore, the incident 480 nm photon from the excitation light source cannot be coupled into the waveguide thereby negating the occurrence of FRET, and (2) the direct excitation of TAM or Cy5 fluorophore is negligible.



Fig.2.10: FRET emission spectrum from the tiles missing a FAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

To determine the extent to which Cy5 acts as the energy output channel, a control experiment was conducted using DNA tiles missing Cy5 fluorophore. The Cy5 fluorophore should in theory absorb energy from the TAM fluorophore (resulting from the first FRET event as indicated in Fig.2.6) and emit energy, resulting in an expected spectrum, as indicated in Fig.2.8. Hence, it is predicted that by removing the Cy5 fluorophore, the peak at 667 nm (Fig.2.8) should not appear. In the control experiment,

the solution was illuminated at 480 nm by the incident excitation light source. As expected, the bulk fluorescence does not show an emission peak of Cy5, as indicated in Fig.2.11. Instead, the bulk fluorescence shows two peaks in the emission spectrum that correspond to the emission of the FAM fluorophore (left peak at 517 nm) and TAM fluorophore (right peak at 583 nm). These results indicate that (1) Cy5 emission peak observed in Fig.2.8 is the consequence of energy transfer from TAM, and (2) the emission spectrum observed in this control experiment is the result of the first FRET event between FAM and TAM fluorophores.



Fig.2.11: FRET emission spectrum from the tiles missing a Cy5 fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

It is useful to briefly summarize the experimental results as presented previously.

The emission spectra of the dye-labeled DNA tile solutions were measured under equal excitation conditions with: (1) all three fluorophores (Fig.2.8), (2) missing the intermediate fluorophore (Fig.2.9), (3) missing the input fluorophore (Fig.2.10), and (4) missing the output fluorophore (Fig.2.11). With all three fluorophores, a single emission

peak of the output fluorophore (667 nm) is expected under perfect FRET conditions. Rather, the observed emission spectrum shows a pronounced peak of the output fluorophore (667 nm) and emission peaks of the input (517 nm) and intermediate (583 nm) fluorophores (Fig.2.8). Although the observed emission spectrum is the result of the FRET process, it is extremely difficult to obtain the perfect FRET conditions due to the use of different fluorophores with different quantum efficiency, spectral overlap, and dipole orientation². Clearly understanding the physics of the FRET process is an ongoing effort in the literature. In theory, without the intermediate fluorophore (i.e., TAM), no emission peak of the output fluorophore (667 nm) and an emission peak of the input fluorophore (517 nm) are expected resulting in a disruption of the FRET process. In reality, a minimal emission peak of the output fluorophore (667 nm) and a pronounced emission peak of the input fluorophore (517 nm) are observed (Fig.2.9), indicating that only a small percentage of the input fluorophore energy transfers to the output fluorophore and thus FRET is significantly but not fully disrupted. Without the input fluorophore (i.e., FAM), no emission peak of the output fluorophore (667 nm) as well as the intermediate fluorophore (583 nm) are theoretically expected. Indeed, the observed fluorescence spectrum (Fig.2.10) shows no indication of emission from both the output and intermediate fluorophores. Without the output fluorophore (i.e., Cy5), no emission peak of the output fluorophore (667 nm) is theoretically expected. Instead, the observed emission spectrum (Fig.2.11) shows emission peaks of the input (517 nm) and intermediate (583 nm) fluorophores without the emission peak of the output fluorophore (667 nm); this also confirms the first FRET event according to the design.

From the experimental fluorescence results as displayed in Fig.2.8, the dominant peak at 667 nm of the Cy5 fluorophore can only be explained through double FRET events (Fig.2.6) for several reasons. First, the direct excitation of Cy5 fluorophore is not possible at the 480 nm incident excitation wavelength. Second, FRET via direct excitation of the TAM fluorophore cannot solely account for this substantial increase, as indicated in Fig.2.10. And finally, a direct transfer of energy from the FAM fluorophore to the Cy5 fluorophore is insignificant, as indicated in Fig.2.9.

To determine the overall efficiency of the three fluorophore system, least-squares curve fitting (details in Appendix C.1) was performed for fluorophore-labeled DNA tiles in three configurations: (1) three fluorophores (FAM, TAM, Cy5), (2) missing output fluorophore (Cy5), and (3) excluding intermediate and output fluorophores (TAM, Cy5). As a result of the fitting procedure (Appendix A.1 and C.1), the contribution of each fluorophore to the overall spectrum is calculated. The energy transfer efficiency from the input fluorophore to the intermediate fluorophore is determined by³

$$E_1 = 1 - \frac{k_{FT}}{k_F} \tag{2.1}$$

where k_{FT} is the fluorescent contribution of input fluorophore in the presence of the intermediate fluorophore and k_F is the fluorescent contribution of input fluorophore in the absence of the intermediate fluorophore. The energy transfer efficiency from the intermediate fluorophore to the output fluorophore is determined by³

$$E_2 = 1 - \frac{k_{FTC}}{k_{FT}} \tag{2.2}$$

where k_{FTC} is the fluorescent contribution of intermediate fluorophore in the presence of the output fluorophore and k_{FT} is the fluorescent contribution of intermediate fluorophore in the absence of the output fluorophore. The overall efficiency is determined by³

$$E = E_1 \times E_2 \tag{2.3}$$

Hence, Equation (2.3) indicates a total amount of the energy transfer to the output fluorophore from the input fluorophore via the intermediate fluorophore.

By analyzing the fluorescence spectra for DNA tiles with three fluorophores (Fig.2.8) and missing Cy5 fluorophore (Fig.2.11), the TAM to Cy5 efficiency can be calculated, providing more proof that the energy transfer occurs from the FAM to the Cy5 via the TAM intermediary. A least-squares fitting procedure (details in Appendix C.1) was used to extract the contribution of the TAM to the fluorescence spectra in Fig.2.8 and Fig.2.11. The results are summarized in Fig.2.12. The decrease in fluorescence of the TAM is clearly observed when Cy5 is proximal, demonstrating that energy transfer to the Cy5 from the TAM is occurring. The FAM fluorescence for the two structures was roughly equal, indicating that after absorbing an incident photon, the FAM fluorophore emitted the same photon energy into free space independent of the Cy5 fluorophore's presence. Using Equation (2.2), the energy transfer efficiency from TAM to Cy5 was determined to be 57.8%.



Fig.2.12: TAM fluorescence from DNA tiles containing FAM-TAM (FT) and FAM-TAM-Cy5 (FTC). Each bar in the graph was created using five trial measurements from the same solution. The TAM fluorescence from emission spectra was normalized for molecular concentration. There is a considerable drop in TAM fluorescence when the Cy5 is present, indicating FRET behavior.

By analyzing the fluorescence spectra for DNA tiles with the FAM fluorophore only and with FAM and the TAM fluorophores, the FAM to TAM FRET efficiency can be calculated. Applying Equation (2.1), the energy transfer efficiency from FAM to TAM was determined to be 44.3%. Through analysis of the spectral data for the double FRET process on dye-labeled DNA tiles and Equation (2.3), an efficiency of ~25.3% was determined. This efficiency is most likely reasonable as it is comparable to the calculated value reported by Haustein *et al.*³ with duplex DNA.

From the fluorescence results and subsequent discussion described in this section, it is evident that FRET-based waveguides were successfully designed and fabricated using fluorophore-labeled DNA tiles. The optical results indicate that the photonic energy was diffusively transferred along the sequential chain of fluorophores via double FRET events. It has been established that the photonic energy transfer performance of FRETbased waveguides using fluorophore-labeled DNA tiles is comparable to FRET-based waveguides using fluorophore-labeled duplex DNA. As a result, more complex FRETbased waveguides can be built beyond the use of duplex DNA, perhaps not only using DNA tiles, but various branched DNA junctions can be implemented as well.

CHAPTER 3: DNA ORIGAMI NANOTUBE DIFFUSIVE WAVEGUIDE

3.1 Experimental

3.1.1 Design

DNA origami nanotubes were previously designed and fabricated²² with the idea of being used as nanoparticle scaffolds. In this work, DNA nanotubes that incorporate fluorophores provide another design concept for building larger area waveguides. The following section explores the designs for building FRET-based waveguides using fluorophore-labeled DNA origami nanotubes.

DNA origami nanotubes were designed using the principles reported by Mathieu *et al.*²³ and Douglas *et al.*²⁴, where the circular single-stranded M13mp18 DNA molecule was folded into a six-helix nanotube bundle using the DNA origami method developed by Rothemund¹⁷. The design reported here uses 170 unique staple strands to fold the single-stranded M13mp18 scaffold, resulting in DNA nanotubes with blunt ends. The nanotube design is illustrated and described in detail in Appendix D.1.

For studying FRET-based waveguides, DNA origami nanotubes were used to construct a stable sequential arrangement of fluorescent dyes separated by a distance dictated to maximize the probability of FRET. Three different fluorophores identical to those of the dye-labeled DNA tiles were incorporated at specific locations on three of the unique staple strands used to fold the DNA origami nanotube scaffold as illustrated in Fig.3.1. The fluorophore locations were chosen to form a descending fluorophore chain relative to absorption energy from the primary donor (FAM) to the final acceptor (Cy5) via the intermediate acceptor-donor (TAM). The designed distance between fluorophores is ~3.1 nm; equivalent to a chain of 7 nucleotides in horizontal direction and 2 nm in vertical direction.



Fig.3.1: Dye-labeled (dots) staple strands within the nanotube. FAM is attached to the left staple strand (labeled F), TAM is attached to the middle strand (labeled T), and Cy5 is attached to the right strand (labeled C).

3.1.2 Materials and Methods

The dye-labeled staple strands used were identical to those of the dye-labeled DNA tiles. Staple strands used to fold the long M13mp18 strand are listed in Appendix D.3. All strands were purchased from Integrated DNA Technologies²¹. All strands were available unlabeled, which permitted a variety of fluorophore combinations to be incorporated in the nanotubes. DNA origami nanotubes with a variety of dye combinations were synthesized for control experiments (schematics are not shown). To

synthesize the nanotubes, M13mp18 viral DNA was combined with 170 staple strands (i.e., regular staple strands and dye-labeled staple strands) in a molar ratio of 1:5 in a solution of $1\times$ TAE, Mg²⁺. The concentration of the M13mp18 genomic DNA was approximately 50 nM. The volume of prepared solution was 30 µL. All DNA strands were used without further purification. The DNA solution was thermally annealed at 90°C for 20 minutes and subsequently cooled to 20°C at approximately 1°C per minute using a thermal cycler. Following nanotube synthesis, the solution was filtered using gel electrophoresis to remove excess staple strands and fluorophores. Appendix B.2 shows the details of the procedure to perform the gel electrophoresis for fluorophore-labeled DNA origami nanotubes.

To assess the formation of nanotubes, AFM was performed. During AFM sample preparation, 5 μ L of nanotube solution was dispersed onto freshly cleaved mica with 20 μ L of 1×TAE, Mg²⁺ buffer and allowed to adsorb onto the surface for 5 minutes. Then, the surface was washed with Milli-Q water and dried with forced nitrogen gas. Images of nanotubes were acquired using AFM (Veeco Multimode PicoForce with a Nanoscope IV controller) under ambient conditions, in AC mode / tapping mode, using silicon cantilever-based tips (Nanosensors PPP-NCH). Cantilevers had a nominal spring constant of 42 N/m with a range of 10 – 130 N/m.

To measure the emission of dye-labeled DNA origami nanotubes, 100 μ L of a 2 nM DNA nanotube solution was placed in a 350 μ L special optical glass (SOG) cell with a 3 mm path length and kept at ambient temperature during the measurement. Fluorescence measurements were performed similar to the method that was described for the dye-labeled DNA tiles in Chapter 2.

3.2 Results and Discussion

AFM was a primary method to confirm that the nanotube synthesis was successful. DNA origami six-helix nanotubes were designed to be 412 nm in length and 6 nm in diameter. Using AFM Fig.3.2, the length was measured to be 418 ± 5 nm and the height was measured to be 2.55 ± 0.5 nm (Fig.3.2). This measured length is in good agreement with the design length. The diameter differences between the design and the experiment suggests that the nanotubes were compressed either by the drying processes or by the tapping force inducing from AFM's probe tip as previously observed by Weisonhorn *et al.*²⁵.



Fig.3.2: AFM height image acquired under ambient conditions for DNA origami nanotubes deposited on an atomically flat mica surface (a); a nanotube's length (b) and diameter (c) of the dotted circle.

Understanding the energy transport in dye-labeled DNA origami nanotubes can be explained in the schematic shown in Fig.2.6. In brief, the incident photons are absorbed

by the input fluorophore and transferred to the output fluorophore through the intermediate fluorophore through double FRET events. FRET-based waveguides were successfully synthesized using DNA origami nanotubes containing a sequential chain of FAM, TAM, and Cy5 fluorophores. DNA origami nanotubes labeled with various fluorophore combinations were synthesized for control purposes. Spectrophotometry was used to analyze the solution of DNA origami nanotubes containing fluorophores. The solution of fluorophore-labeled DNA origami nanotubes with three fluorophores (FAM, TAM, and Cy5) was excited at 480 nm by the incident light source. The ensemble fluorescence spectrum displays three emission peaks, as indicated in Fig.3.3. The left peak at 517 nm (2.40 eV) is the emission of FAM fluorophore; the middle peak at 583 nm (2.13 eV) is the emission of TAM fluorophore.



Fig.3.3: FRET emission spectrum from the nanotubes with all three fluorophores. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

It is predicted that under the perfect FRET condition, the ensemble fluorescence spectrum illustrated in Fig.3.3 should in theory display a single emission peak at 667 nm (Cy5). Rather, the fluorescence spectrum with three peaks is observed. The direct excitation by the incident light source at 480 nm of the TAM and Cy5 should be negligible and only the FAM is efficiently excited at 480 nm. Hence, the resulting fluorescence spectrum must arise from two successive FRET events between the fluorophores contained in the structures, as was also observed for the spectrum (Fig.2.8) of the dye-labeled DNA tiles.

A control experiment was carried out for nanotubes missing the intermediate TAM fluorophore to determine the percentage of direct energy transfer from FAM to Cy5. It is anticipated that the TAM fluorophore should absorb energy from the FAM fluorophore and transfer its absorption energy to the Cy5 fluorophore. By excluding the TAM fluorophore, the emission peak at 583 nm in Fig.3.3 should disappear and the path to directly transfer energy from FAM to Cy5 should be interrupted. The schematic of the design is illustrated in the inset of Fig.3.4. The dye-labeled DNA origami nanotubes solution was excited at the incident light source wavelength of 480 nm. As expected, no emission peak of TAM appears in the bulk fluorescence spectrum as shown in Fig.3.4. Instead, a dominant peak at 517 nm (FAM) and an insignificant peak at 667 nm (Cy5) are observed. The control experimental results indicate that the majority of the FAM excitation energy is emitted into free space and an insignificant percentage of the FAM fluorophore energy transfers to the Cy5 fluorophore in the absence of the intermediate TAM fluorophore. Hence, the interruption of FRET was achieved.



Fig.3.4: FRET emission spectrum from the nanotubes missing TAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

Another control experiment was performed for fluorophore-labeled DNA origami nanotubes without the input fluorophore (FAM) to eliminate the possibility that TAM or Cy5 are excited directly at 480 nm by the incident light source. In theory, the FAM fluorophore after absorbing an incident photon should only be transferring energy to excite the subsequent fluorophore. By excluding the FAM fluorophore, the emission peak at 517 nm (Fig.3.3) should disappear and no emission should be observed due to direct excitation of TAM and/or Cy5. The schematic of the control experiment is illustrated in the inset of Fig.3.5. The solution of fluorophore-labeled DNA origami nanotubes was excited at 480 nm. As expected, the ensemble fluorescence spectrum displays no emission peak at 517 nm as indicated in Fig.3.5. Rather, two insignificant peaks of TAM fluorophore at 583 nm and of Cy5 fluorophore at 667 nm are observed. The experimental results indicate that the direct excitation of TAM and/or Cy5 is insignificant in the absence of the FAM fluorophore.



Fig.3.5: FRET emission spectrum from the nanotubes missing FAM fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

A final control experiment was conducted for fluorophore-labeled DNA origami nanotubes without the Cy5 fluorophore to determine the extent to which Cy5 acts as the energy output channel. It is anticipated that the Cy5 should be excited from the TAM fluorophore (resulting from the FRET between FAM and TAM, as indicated in Fig.2.6) and emit energy resulting in an expected 667 nm emission peak, as depicted in Fig.3.3. By removing the Cy5 fluorophore, the emission peak at 667 nm (Fig.3.3) should disappear and the emission peak of TAM fluorophore should release more energy into free space than the 583 nm emission peak, as shown in Fig.3.3. The schematic of the control experiment is illustrated in the inset of Fig.3.6. The solution of fluorophorelabeled DNA origami nanotubes was illuminated at 480 nm. As expected, no emission peak of Cy5 is observed in the ensemble fluorescence spectrum, as indicated in Fig.3.6. Instead, two emission peaks of FAM fluorophore at 517 nm and of TAM fluorophore at 583 nm are observed in the ensemble fluorescence spectrum. The experimental results indicate that (1) Cy5 emission peak observed in Fig.3.3 is the consequence of the energy transfer from the TAM fluorophore, and (2) the fluorescence spectrum exhibited in this control experiment demonstrates the first FRET event between FAM and TAM fluorophores on DNA nanotube origami.



Fig.3.6: FRET emission spectrum from the nanotubes missing Cy5 fluorophore. The inset illustrates a schematic of the structure. (Note: excitation wavelength = 480 nm)

A brief summary of all experimental results presented in the previous section is discussed in the following section. The fluorophore-labeled DNA origami nanotube

solutions were characterized under the same excitation conditions while measuring the fluorescence emission spectrum of the following fluorophore configuration on nanotubes: all three fluorophores (Fig.3.3), without the intermediate fluorophore (Fig.3.4), without the input fluorophore (Fig.3.5), and without the output fluorophore (Fig.3.6). It is predicted that a single emission peak of the output fluorophore (Cy5) is expected in fluorophore-labeled DNA origami nanotubes with three fluorophores. Instead, the observed emission spectrum displays a dominant peak of the output fluorophore (667 nm) and emission peaks of the input (517 nm) and intermediate (583 nm) fluorophores (Fig.3.3). Although the observed spectrum is the result of the FRET process, it is extremely difficult to obtain the perfect FRET conditions, as mentioned previously. In theory, no emission peak of the output fluorophore (Cy5) and a pronounced peak of the input fluorophore (FAM) are expected in the absence of the intermediate fluorophore (TAM) in fluorophore-labeled DNA origami nanotubes. In fact, a minimal emission peak of the output fluorophore (667 nm) and a dominant emission peak of the input fluorophore (517 nm) are observed (Fig.3.4), indicating that only a small amount of the input fluorophore energy transfers to the output fluorophore. By excluding the input fluorophore (i.e., FAM) from the fluorophore-labeled DNA origami nanotubes, it is predicted that no emission peak of the intermediate (583 nm) and output (667 nm) fluorophores is anticipated. Indeed, the observed fluorescence spectrum (Fig.3.5) shows no indication of emission from both the intermediate and output fluorophores, indicating that the direct excitation of the intermediate or output fluorophore is negligible in the absence of the input fluorophore. Without the output fluorophore (i.e, Cy5), no emission peak of the output fluorophore (667 nm) is theoretically expected from fluorophorelabeled DNA origami nanotubes. Hence, the observed fluorescence spectrum (Fig.3.6) displays no emission peak of the output fluorophore. Instead, the spectrum shows emission peaks of the input (517 nm) and intermediate (583 nm) fluorophores. In summary, the experiments demonstrate the first observation of FRET waveguide on DNA origami nanotubes.

The configuration difference between the fluorophore-labeled DNA origami nanotubes in Fig.3.3 and Fig.3.6 results in the presence of an output fluorescence peak that appears in Fig.3.3 but not in Fig.3.6. It is then predicted that the input fluorescence intensity in Fig.3.3 should be same or less than the input fluorescence intensity in Fig.3.6. Experimentally, the input fluorescence intensity in Fig.3.3 is higher than the input fluorescence intensity in Fig.3.6. This difference could be attributed to the systematic errors, such as (1) the variation in the concentrations, (2) the variation in the solutions, and (3) the variation in the instrument.

It is worth noticing that there is a slight variation in the ratios of the peak height fluorescence emission spectra in Fig.2.11 and Fig.3.6. From spectral analysis, the input emission peak is higher than the intermediate emission peak in Fig.2.11, whereas the input emission peak is lower than the intermediate emission peak in Fig.3.6. The difference in the emission peak in fluorescence spectra can be explained based on the energy transfer efficiency, as illustrated in Fig.A.3. Three cases can be examined to illustrate this point. In one case, no emission peak of the intermediate fluorophore corresponds to zero percent efficiency and thus no FRET occurs. For another case, no emission peak of the input fluorophore and a single emission peak of the intermediate fluorophore correspond to one hundred percent efficiency, which indicates perfect FRET. In the last case, a combination of both emission peaks of the input and intermediate fluorophores can occur and is most typical. This last case corresponds to the variation in the efficiency such that (a) if the emission peak of the input fluorophore is higher than the emission peak of the intermediate fluorophore, then the energy transfer efficiency is typically lower than 50% (e.g., Fig.2.11), and (b) if the emission peak of the input fluorophore, then the energy transfer efficiency is lower than the emission peak of the intermediate fluorophore is lower than the emission peak of the intermediate fluorophore, then the energy transfer efficiency is typically greater than 50% (e.g., Fig.3.6). Hence, the last case provides a general indication of the efficiencies expected in Fig.2.11 and Fig.3.6. And from the spectral analysis, this is found to be the case. Spectral analysis of the data in both figures Fig.2.11 and Fig.3.6, respectively, determined that the energy transfer efficiencies are 44.3% and 66.2%.

As FRET is a diffusive process and thus energy is lost during the process, quantification of the FRET efficiency is necessary to establish a base-line value. To quantitatively determine the overall FRET efficiency occurring between the input fluorophore to the output fluorophore via the intermediate fluorophore, a least-squares fitting procedure (details in Appendix C.1) was used. Fluorophore-labeled DNA origami nanotubes with all three fluorophores and without the output fluorophore were analyzed together to provide the quantitative efficiency. The results are summarized in Fig.3.7. The reduction in fluorescence of the intermediate fluorophore (TAM) is clearly observed when the output fluorophore (Cy5) is abutting the intermediate fluorophore, demonstrating that energy transfer to the output fluorophore from the input fluorophore (FAM) occurs via the intermediate fluorophore. Using the efficiency calculation presented previously in Chapter 2, an efficiency of ~28.3% was determined through analysis of the spectral data. This efficiency is comparable to the value reported for fluorophore-labeled DNA tiles as established in Chapter 2.



Fig.3.7: TAM fluorescence from DNA origami nanotubes containing FAM-TAM (FT) and FAM-TAM-Cy5 (FTC). Each bar in the graph was made using five trial measurements from the same solution. The TAM fluorescence from emission spectra was normalized for molecular concentration. There is a considerable drop in TAM fluorescence when the Cy5 is present, indicating FRET behavior.

From the design specification, the distance between fluorophores on the DNA tiles is smaller than the distance of those on the DNA origami nanotubes. In addition, the rotational angle between the fluorophores is zero for the DNA tile structure and non-zero for the DNA origami nanotube structure. Since the energy transfer efficiency is inversely proportional to the effective distance between the fluorophores and directly proportional to the cosine of the rotational angle between fluorophores (i.e., FRET orientation factor - see Appendix A.1), then one would expect that the energy transfer efficiency should be

greater for the fluorophore-labeled DNA tiles. However, the experimental results indicate that the energy transfer efficiency in fluorophore-labeled DNA origami nanotubes is greater than that in fluorophore-labeled DNA tiles. This difference may be explained in terms of the dynamics of the DNA structures. The energy transfer efficiency calculations assume that the DNA tiles and nanotubes are rigid, non-dynamic structures. Since the fluorescence measurements are performed in solution, the DNA structures are not rigid at all but are rotationally active due to thermal and concentration fluctuations. Hence, the more structurally rigid a DNA structure, the less rotational and torsional activity that will be experienced by a DNA structure. The less rotational torsional activity a structure experiences, the less deviation the structure will encounter in the effective distance and FRET orientation factor, which in theory will provide a greater energy transfer efficiency. Cursory finite element analyses were performed using CanDo²⁶ on the DNA tile and DNA origami nanotube to examine their stiffness or rigidity. It was found that the DNA origami nanotube was more rigid than the DNA tile. Hence, the fact that the DNA tile is less rigid than the DNA origami nanotube may explain why the fluorophore-decorated DNA tile has a lower energy transfer efficiency than fluorophore-decorated DNA origami nanotube.

From the fluorescence results and subsequent discussion described in this section, it is evident that FRET-based waveguides were successfully designed and fabricated using fluorophore-labeled DNA origami nanotubes. The optical results indicate that the photonic energy was diffusively transferred along the sequential chains of fluorophores via double FRET events. It is also apparent that FRET-based waveguides can be extended beyond DNA tiles or branched DNA junctions. As a result, more complex FRET-based waveguides can be built using not only DNA origami nanotubes but various DNA origami nanostructures.

CHAPTER 4: DNA ORIGAMI NANOTUBE QUANTUM DOT ARRAYS

"Reproduced in part with permission from Hieu Bui, Craig Onodera, Carson Kidwell, YerPeng Tan, Elton Graugnard, Wan Kuang, Jeunghoon Lee, William B. Knowlton, Bernard Yurke, William L. Hughes, "Programmable Periodicity of Quantum Dot Arrays with DNA Origami Nanotubes," Nano Letters 10, no. 9 (2010): 3367-3372. Copyright 2010 American Chemical Society." Link to the article can be found in the following: <u>http://pubs.acs.org/doi/abs/10.1021/n1101079u</u> or DOI: 10.1021/n1101079u.

4.1 Experimental

A brief background on the properties of semiconductor quantum dots prior to examining the design is provided in this section. Quantum dots are now popularly employed for biomolecular and cellular imaging as a result of their photophysical stability and their intense fluorescent emission. Moreover, their large Stokes' shifts facilitate the detection of emitted light without collecting scatter excitation light²⁷. Their emission wavelength is also tunable by controlling their particle size²⁷. These properties have made quantum dots one of the most promising fluorescence labeling agents²⁸. In contrast, fluorescent fluorophores are also commonly used for labeling biomolecules because of their usability and wide variety. However, emission from fluorescent dyes is usually weaker than that from quantum dots due to their low extinction coefficients²⁹. Hence, an effort to design and fabricate quantum dot-labeled DNA origami nanotubes was carried out. In the following sections, the design and fabrication DNA origami as scaffolding for the periodic arrangements of quantum dots are related. Much of the work described in this chapter has been subsequently published by the author²².

4.1.1 Design

The following designs were chosen because of (1) the ability to precisely arrange nanoparticles via the inherent programmability of DNA and the advanced DNA labeling techniques, and (2) semiconducting quantum dots are predicted to be the ideal fluorescent donors for later FRET studies.

To incorporate quantum dot binding sites, prior to nanotube synthesis, selected staple strands were extended with a 2.2 nm tether consisting of 5 thymine nucleotides and modified with biotin at the 3' end. The resulting DNA nanotubes possessed precisely spaced biotin binding sites for controlled positioning of streptavidin-conjugated quantum dots (hereafter referred as QD) along the length of the nanotube. To test controlled nanoparticle patterning, four distinct DNA nanotubes were synthesized with evenly spaced binding sites designed to attach 5, 9, 15, or 29 QDs in order to form arrays with periodicities of 71, 43, 29, or 14 nm, respectively. The biotin-labeled DNA nanotubes were designed by functionalizing the appropriate staple strands as described detail in Appendix D.4.

4.1.2 Materials and Methods

The nanotubes were synthesized by combining M13mp18 viral DNA with unmodified and biotin-labeled staple strands in a molar ratio of 1:10:10 in a solution of 1xTAE, Mg^{2+} . The concentration of the M13mp18 viral DNA was approximately 50 nM. The volume of prepared solution was 80 µL. Staple strands used to fold the long M13mp18 strand are listed in Appendix D.3 and D.4. All DNA strands were used without further purification. To form nanotubes, the DNA solution was thermally annealed at 90°C for 20 minutes, then cooled to 20°C at approximately 1°C per minute using a thermal cycler. After the nanotubes were synthesized, the solution was centrifuged using a centrifugal filter (Amicon-Ultra-0.5-100k, Millipore Inc.) at 14,000 g for 10 minutes to remove excess staple strands and small, unbound DNA fragments. To assess the successful formation of functional DNA origami nanotubes, AFM was used as described in the Materials and Methods section in Chapter 3.

4.2 Results and Discussion

Fig.4.1 shows DNA origami nanotubes with 9 biotin binding sites as synthesized (a-e), after functionalization with streptavidin (f-j), and after functionalization with streptavidin-conjugated quantum dots (k-o). Fig.4.1a illustrates the biotin-labeled nanotube structure, while (b,c) show low and high magnification AFM height images, respectively. The dashed line in (c) indicates the location of the cross-sectional height profile in (d). From this profile, a nanotube height of ~ 2.6 nm is measured. The mean nanotube height ranged from 1.7 ± 0.4 to 3.5 ± 0.1 nm. The axial profile shown in (e) emphasizes relative height variations along the nanotube length. The mean nanotube length was measured to be 436 ± 14 nm from 100 samples and was independent of the imaging conditions.



Fig.4.1: Schematics, AFM images at low magnification (upper) and high magnification (lower), and cross-sectional (upper) and axial (lower) height profiles of functionalized DNA origami nanotubes with 9 biotin binding sites with: (a-e) no attached nanoparticles; (f-j) attached streptavidin; (k-o) attached streptavidin-conjugated quantum dots. The dashed lines in the high magnification AFM images indicate the location of the cross-sectional profiles. Axial profiles represent the average of multiple profiles across the width of the nanotube. (Reprinted with permission from Ref.22. Copyright 2010 American Chemical Society).

Once biotin-labeled DNA nanotubes were verified via AFM, the accessibility and

reactivity of the biotin attachment sites were examined by combining a 1 nM solution of

biotin-labeled nanotubes with pure, lyophilized streptavidin (Sigma Aldrich). The components were allowed to react for 2 hours at room temperature prior to AFM characterization. Fig.4.1f illustrates the biotin-labeled DNA nanotubes with attached streptavidin. The successful attachment of 9 streptavidin molecules is clearly observed by comparing the small scan AFM images without streptavidin in (c) and with streptavidin in (h). The cross-sectional profile in (i), obtained at an apparent streptavidin site, reveals a height increase of ~0.5 nm relative to the nanotube shown in (d). The axial profile in (j) clearly displays 9 peaks with a periodicity of 45 nm, very close to the expected value of 43 nm. While the measured height increase at a streptavidin site was ~0.5 nm, the mean height of free streptavidin, dispersed onto freshly-cleaved mica, was measured to range from 0.7 ± 0.2 to 2.3 ± 0.5 nm under various imaging conditions. The streptavidin heights measured here are consistent with previous studies²⁵.

CdSe/ZnS core/shell streptavidin conjugated quantum dots (Qdot 585, Invitrogen) with an average diameter of 15-20 nm were chosen to test nanoparticle attachment. To ensure a high attachment yield, a 1 nM solution of functionalized DNA nanotubes was combined at room temperature with a 200 nM solution of quantum dots for 2 hours. The reacted DNA nanotubes, with attached quantum dots, were dispersed onto mica surface and dried as described above. Fig.4.1k illustrates the attachment of the quantum dots to the biotin-labeled DNA nanotubes. Fig.4.1(l,m) respectively show low and high magnification AFM height images of the DNA nanotubes with attached quantum dots. When compared to (c) and (h), quantum dots attach to biotin-labeled DNA nanotubes with the same periodic spacing. The cross-sectional profile across an apparent quantum dot in (n) yields a height of 5.5 nm, nearly twice the measured height of the nanotube

without attached particles. The mean height of free quantum dots, dispersed onto freshlycleaved mica, was measured to range from 4.7 ± 0.7 to 5.5 ± 0.6 nm under various imaging conditions. Although the diameter of the streptavidin-conjugated quantum dots is ~20 nm in solution according to manufacture specifications, the AFM height measurements of the dehydrated quantum dots corresponding to the approximate diameter of the CdSe/ZnS core/shell quantum dot, as measured by TEM (data not shown).

To illustrate the flexibility of the design and confirm control over nanoparticle attachment, functionalized DNA nanotubes were synthesized with 5, 9, 15, and 29 biotin attachment sites to enable the formation of quantum dots arrays with periodicities of 71, 43, 29, and 14 nm, respectively. These nanotubes were reacted with quantum dots and dispersed onto mica in the same manner described above. Fig.4.2 shows height AFM images of quantum dots attached to DNA nanotubes with (a) 5, (b) 9, (c) 15, and (d) 29 biotin binding sites. Successful attachment to each biotin binding site was observed for nanotubes with 5 or 9 available sites; however, attached quantum dots were not observed at each site for nanotubes with 15 and 29 available sites, respectively. The average distance between two adjacent quantum dots were measured to be approximately 71 ± 3 , 49 ± 4 , 46 ± 5 , and 31 ± 4 nm for nanotubes with 5, 9, 15, and 29 available biotin binding sites, respectively. The measured spacing for 5 and 9 binding sites agree well with the predicted periods of 71 and 43 nm. However, the arrays seen in Fig.4.2(c,d) formed with a reduced number of quantum dots, and consequently, a larger spacing than expected, i.e. 29 and 14 nm, respectively.



Fig.4.2: High magnification AFM images of streptavidin-conjugated quantum dots attached to functionalized DNA origami nanotubes with: (a) 5 binding sites, 71 nm period; (b) 9 binding sites, 43 nm period; (c) 15 binding sites, 29 nm period; and (d) 29 binding sites, 14 nm period. All scale bars are 100 nm. Note (c) and (d) have fewer attached quantum dots than available binding sites. In addition, the diameter of quantum dots varies between images because of variation in tip radii between scans. (Reprinted with permission from Ref.22. Copyright 2010 American Chemical Society).

Functionalized DNA origami nanotubes were designed with biotin-labeled staple strands spaced evenly along the axis of the nanotubes. The nanotubes were synthesized and combined with streptavidin-conjugated quantum dots to form nanoparticle arrays with controlled periodicities. AFM results of the synthesized arrays revealed successful attachment of quantum dots at locations along the nanotube axes that corresponded to available biotin binding sites. Statistical analysis is described in detail in Appendix E.10. However, the analysis indicates that steric hindrance strongly affects the arrays with smaller distance separation between nanoparticles. Although steric hindrance seems to affect nanoparticle attachment, it is predicted that with molecular particle size (e.g., fluorescent fluorophores), hybrid FRET-based waveguides can be built to study the photonic energy transport between quantum dot (inorganic light emitting semiconductor) and fluorophore (organic fluorescent dyes).

CHAPTER 5: SUMMARY & FUTURE WORK

In summary, two FRET-based waveguides and a systematic approach to fabricate four different nanoparticle arrays have been demonstrated. FRET-based waveguides have been successfully designed, fabricated, and characterized using DNA tiles, DNA origami nanotubes, and fluorophore-labeled DNA. It has been established that the photonic energy was diffusively transferred from one end of the devices to the other through the FRET processes. The limitation of the persistent length of a single duplex DNA has been overcome by using larger DNA scaffolds (e.g., branched DNA junctions and DNA origami). The strengths of these approaches are (1) DNA materials are capable of selfassembly, molecular recognition, and programmability, (2) the flexibility of choosing fluorophores to form the energy cascade as the driving force in order to transfer photon energy through FRET, (3) the synthetic simplicity, and (4) the design flexibility of the structure for future enhancement for more complex circuitries. However, the current fluorescence measurements only provided the average representation of all FRET-based waveguides contained in the tested solution. To gain more insight on the complex photophysical behavior of the FRET-based waveguides, time-resolved fluorescence spectroscopy, single-molecule fluorescence spectroscopy, or total internal reflection fluorescence microscopy is needed.

In addition to FRET-based waveguides, it has been demonstrated that four different nanoparticle arrays have been successfully designed, fabricated, and characterized using DNA origami nanotubes, biotin-labeled DNA, and streptavidinconjugated quantum dots. The ability to control nanoparticle patterning has been explored. In the quantum dot-labeled DNA origami nanotube alone, there are 170 unique staple strands that can be functionalized by various means at either end, or even within the strand itself. Hence, the possibilities for variations in nanoparticle arrangements are significantly enormous.

The significance of this thesis work includes (1) the use of different DNA nanostructures (DNA tile and DNA origami nanotube) to build FRET-based waveguides with comparable efficiencies and (2) the ability to systematically design and incorporate various arrays of semiconductor nanoparticles onto the same DNA origami nanotubes. To minimize the design time and costs, the same fluorophore-labeled DNA strands can be used in two different DNA template designs. Ultimately, the results from this thesis create more opportunities to investigate novel near-field optical interactions between nanoparticles (i.e., organic and/or inorganic) where such interactions are largely unexplored.

The future work from the presented studies can be pursued on developing (1) the extension of the energy transfer to longer distance on the entire DNA origami nanostructures, (2) the insertion of a photo switches to enable the controlled switching of the photonic waveguide, and (3) the construction of novel optical transistors. To improve the current performance of FRET-based waveguides, the distance between fluorophores and the alignment of dipole orientation can be investigated. FRET-based waveguides using DNA nanostructures as scaffolding (e.g., tiles and nanotubes) have been demonstrated with fluorophores and is underway using quantum dots. The goal of this

device concept is to build broadband input FRET-based waveguides by utilizing the large absorption cross section of quantum dots and their ability to be excited in the ultraviolet spectrum.

REFERENCES

1. Liu, J.-M., *Photonic devices*. Cambridge University Press: 2005.

2. Pistol, C.; Dwyer, C.; Lebeck, A. R., Nanoscale Optical Computing Using Resonance Energy Transfer Logic. *Micro, IEEE* **2008**, *28*, 7-18.

3. Haustein, E.; Jahnz, M.; Schwille, P., Triple FRET: A tool for Studying Long-Range Molecular Interactions. *ChemPhysChem* **2003**, *4*, 745-748.

4. Ohya, Y.; Yabuki, K.; Tokuyama, M.; Ouchi, T., Construction and Energy Transfer Behavior of Sequential Chromophore Arrays on an Oligo-DNA - Assembly. *Supramolecular Chemistry* **2003**, *15*, 45-45.

5. Vyawahare, S.; Eyal, S.; Mathews, K. D.; Quake, S. R., Nanometer-scale Fluorescence Resonance Optical Waveguides. *Nano Letters* **2004**, *4*, 1035-1039.

6. Heilemann, M.; Tinnefeld, P.; Sanchez Mosteiro, G.; Garcia Parajo, M.; Van Hulst, N. F.; Sauer, M., Multistep Energy Transfer in Single Molecular Photonic Wires. *Journal of the American Chemical Society* **2004**, *126*, 6514-6515.

7. Seeman, N. C., DNA in a material world. *Nature* **2003**, *421*, 427-431.

8. Seeman, N. C., Nucleic acid junctions and lattices. *Journal of Theoretical Biology* **1982**, *99*, 237-247.

9. Holliday, R., A Mechanism for Gene Conversion in Fungi. *Genetics Research* **1964**, *5*, 282-304.

10. Seeman, N. C., DNA NANOTECHNOLOGY: Novel DNA Constructions. *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, 225-248.

11. Chen, J.; Seeman, N. C., Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **1991**, *350*, 631-633.

12. Seeman, N. C.; Kallenbach, N. R., DNA Branched Junctions. *Annual Review of Biophysics and Biomolecular Structure* **1994**, *23*, 53-86.

13. Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998**, *394*, 539-544.

14. Mao, C.; Sun, W.; Seeman, N. C., Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy. *Journal of the American Chemical Society* **1999**, *121*, 5437-5443.

15. LaBean, T. H.; Yan, H.; Kopatsch, J.; Liu, F.; Winfree, E.; Reif, J. H.; Seeman, N. C., Construction, Analysis, Ligation, and Self-Assembly of DNA Triple Crossover Complexes. *Journal of the American Chemical Society* **2000**, *122*, 1848-1860. 16. Mao, C.; LaBean, T. H.; Reif, J. H.; Seeman, N. C., Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature* **2000**, *407*, 493-496.

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

18. Kuzuya, A.; Komiyama, M., DNA origami: Fold, stick, and beyond. *Nanoscale* **2010**, *2* (3), 310-310.

19. Stein, I. H.; Steinhauer, C.; Tinnefeld, P., Single-Molecule Four-Color FRET Visualizes Energy-Transfer Paths on DNA Origami. *Journal of the American Chemical Society* **2011**, *0*.

20. Trudy, M.; McKee, J., *Biochemistry: The Molecular Basis of Life*. McGrau-Hill Science/Engineering/Math: 2002.

21. IDT_Website <u>http://www.idtdna.com</u>.

22. Bui, H.; Onodera, C.; Kidwell, C.; Tan, Y.; Graugnard, E.; Kuang, W.; Lee, J.; Knowlton, W. B.; Yurke, B.; Hughes, W. L., Programmable Periodicity of Quantum Dot Arrays with DNA Origami Nanotubes. *Nano Letters* **2010**, *10*, 3367-3372.

23. Mathieu, F.; Liao, S.; Kopatsch, J.; Wang, T.; Mao, C.; Seeman, N. C., Six-Helix Bundles Designed from DNA. *Nano Letters* **2005**, *5*, 661-665.

24. 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*, 6644-6648.

25. Weisenhorn, A. L.; Schmitt, F. J.; Knoll, W.; Hansma, P. K., Streptavidin binding observed with an atomic force microscope. *Ultramicroscopy* **1992**, *42-44*, 1125-1132.

26. CanDo <u>http://cando.dna-origami.org/</u>.

27. Kashida, H.; Sekiguchi, K.; Liang, X.; Asanuma, H., Accumulation of Fluorophores into DNA Duplexes To Mimic the Properties of Quantum Dots. *Journal of the American Chemical Society* **2010**, *132*, 6223-6230.

28. Michalet, X.; Pinaud, F. F.; Bentolila, L. A.; Tsay, J. M.; Doose, S.; Li, J. J.; Sundaresan, G.; Wu, A. M.; Gambhir, S. S.; Weiss, S., Quantum Dots for Live Cells, in Vivo Imaging, and Diagnostics. *Science* **2005**, *307*, 538-544.

29. Johal, M. S., Understanding Nanomaterials. CRC Press: 2011.

30. Tsurumi, T.; Hirayama, H.; Vacha, M.; Taniyama, T., *Nanoscale physics for materials science*. CRC Press: 2009.

31. Valeur, B., *Molecular fluorescence: principles and applications*. Wiley-VCH: 2002.

32. Dietrich, A.; Buschmann, V.; Müller, C.; Sauer, M., Fluorescence resonance energy transfer (FRET) and competing processes in donor-acceptor substituted DNA strands: a comparative study of ensemble and single-molecule data. *Reviews in Molecular Biotechnology* **2002**, *82*, 211-231.

33. IDT_FluorescentDyes

http://cdn.idtdna.com/Support/Technical/TechnicalBulletinPDF/Spectra_of_DNA_conjug_ated_fluorescent_dyes.pdf.

34. IDT_ModificationHome

http://www.idtdna.com/Catalog/Modifications/Modificationhome.aspx.

35. Mendenhall, W.; Scheaffer, R., *Mathematical Statistics with Applications*. Duxbury Press: North Scituate, MA, 1973.
APPENDIX A

FRET Background and Fluorophores

A.1 FRET Background

Fluorescence resonance energy transfer (FRET) is the radiationless transfer of energy between two weakly interacting dipoles – in the context of this thesis, the dipoles are referred to as the two fluorophores. FRET is also referred to as a Förster energy transfer after Theodor Förster, who first derived the equation for the energy transfer rate³⁰. The following explains the relationships between (1) dipole-dipole interaction, (2) the FRET relative orientation, (3) the energy transfer rate, (4) Förster radius, and (5) the FRET efficiency.

The interaction energy in the transition dipole moments between two fluorophores through electrostatic means is determined by²⁹

$$V(R,\theta_1,\theta_2,\theta_T) = \frac{\mu_1 \mu_2}{4\pi\varepsilon_0 R^3} \left(\cos(\theta_T) - 3\cos(\theta_1)\cos(\theta_2)\right)$$
(A.1)

where μ_1 and μ_2 are the dipole moments of the two fluorophores, respectively. Two fluorophores are placed on the same DNA structure at a distance *R* and at an angle θ_T from one to the other as illustrated in Fig.A.1. Each fluorophore has an angle relative to the DNA structure itself (i.e. θ_1 and θ_2). From Equation (A.1), the FRET relative orientation constant of the two fluorophores to the interaction energy can be expressed by³⁰

$$\kappa = \cos(\theta_T) - 3\cos(\theta_1)\cos(\theta_2) \tag{A.2}$$

The quantity κ^2 ranges from 0 (perpendicular transition moments) to 4 (collinear transition moments) as illustrated in Fig.A.1. For instance, when the transition moments

are parallel, but opposed, $\kappa^2 = 1$. In the literature³¹, the average value of κ^2 is 2/3, which corresponds to the assumption that the fluorophores are free to randomly rotate.



Fig.A.1: Illustration of the dipole-dipole coupling between two transition dipole moments of two given fluorophores and typical values of the factor (κ^2) for specific dipole orientations³⁰

By invoking the general form of Fermi's golden rule for the transition rate between two fluorophores³¹ and using Equation (A.1), Förster derived an expression for the rate constant k_{ET} for dipole-dipole induced energy transfer (as cited in Ref.32):

$$k_{ET} = \frac{9000 \cdot \ln 10 \cdot \kappa^2 \cdot \Phi_D}{128 \cdot \pi^5 \cdot n^4 \cdot N_A \cdot \tau_D \cdot R^6} \int_0^\infty F_D(\lambda) \cdot \varepsilon_A(\lambda) \cdot \lambda^4 \cdot d\lambda$$
(A.3)

Equation (A.3) expresses the rate constant for energy transfer in measureable spectroscopic quantities such as: the refractive index of the medium, *n*; the fluorescence quantum yield of the donor, Φ_D ; its fluorescence lifetime, τ_D ; Avogadro's number, N_A ; the normalized fluorescence spectrum of the donor, $F_D(\lambda)$; the absorption spectrum of the acceptor, expressed by its extinction coefficient, $\varepsilon_A(\lambda)$; and the wavelength λ in nm.

Equation (A.3) can be written in terms of the Förster critical transfer radius R_0 , the distance at which the transfer efficiency (as cited in Ref.32) equals 50%:

$$k_{ET} = \frac{1}{\tau_D} \left(\frac{R_0}{R}\right)^6 \tag{A.4}$$

where R_0 is given by³⁰:

$$R_0^6 = \frac{9000 \cdot \ln 10 \cdot \kappa^2 \cdot \Phi_D}{128 \cdot \pi^5 \cdot n^4 \cdot N_A} \int_0^\infty F_D(\lambda) \cdot \varepsilon_A(\lambda) \cdot \lambda^4 \cdot d\lambda$$
(A.5)

The FRET efficiency is related to the distance, R, between the fluorophores and is given by³²:

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \tag{A.6}$$

The FRET efficiency between fluorophores defines the transfer rate, or latency, of an exciton passing through the system. This efficiency also defines important systemlevel properties such as power consumption, heat dissipation requirements, gain, and the signal-to-noise ratio². Four important parameters relate specific fluorophore properties (and relative positions on a nanostructure) to FRET efficiency: fluorophore separation, spectral overlap, Förster radius, and FRET relative orientation or rotational angle between fluorophores.

To define the Förster distance from the transfer rate Equation (A.3), we let

$$\gamma = \frac{9000 \cdot \ln 10 \cdot \kappa^2 \cdot \Phi_D}{128 \cdot \pi^5 \cdot n^4 \cdot N_A} \int_0^\infty F_D(\lambda) \cdot \varepsilon_A(\lambda) \cdot \lambda^4 \cdot d\lambda$$
(A.7)

Equation (A.3) can be rewritten as

$$k_{ET} = \frac{\gamma}{\tau_D R^6} \tag{A.8}$$

When $R = R_0$, Equation (A.8) simplifies to

$$k_{ET} = \frac{1}{\tau_D} \tag{A.9}$$

which states that the transfer rate between a donor and an acceptor is equal to the fluorescence rate of a donor in the absence of an acceptor.

Equate Equation (A.8) and (A.9) when $R = R_0$, we obtain

$$\frac{\gamma}{\tau_D R_0^6} = \frac{1}{\tau_D} \tag{A.10}$$

$$\gamma = R_0^6 \tag{A.11}$$

From Equation (A.11) and (A.7), it has been shown that the Förster distance is equivalent to the expression in Equation (A.5).

Fig.A.2 illustrates the transition energy diagram of the FRET process between fluorophores. The donor fluorophore is first excited by the absorption of a photon with energy hv. The excited-state donor energy, also called an exciton, is transferred to the acceptor fluorophore, which becomes excited via FRET and, through spontaneous decay

of the excited state, emits a photon with lower energy² hv'. The energy of the acceptor excited state is always lower than the energy of the donor excited state. Without additional energy, this constrains FRET to a single direction: from donor to acceptor². Förster derived the FRET efficiency transfer rate on the basis of classical charge dipoledipole coupling and quantum mechanics; it is in the 10⁻¹¹ second to 10⁻⁹ second time scale².



Fig.A.2: Transition energy diagram of the FRET process. The energy is corresponding to the vertical direction.

Fig.A.3 illustrates the variation in energy transfer efficiency shifts and its corresponding fluorescence emission spectrum of the input and intermediate fluorophores. It is predicted that (1) no intermediate emission peak (λ_2) corresponds to zero percent energy transfer (i.e., no FRET), (2) no input emission peak (λ_1) and a single emission peak of the intermediate fluorophore (λ_2) correspond to one hundred percent energy transfer (i.e., ideal FRET), and (3) a combination between input and intermediate emission peak is higher than the intermediate emission peak, then the energy transfer such that (a) if the input emission peak is higher than the intermediate emission peak is lower than the intermediate emission peak is lo



Fig.A.3: Illustration of variation in energy transfer efficiency and the corresponding fluorescence emission spectrum of the two fluorophores system.

A.2 Fluorophores

Fluorophores (or dyes) absorb light at a characteristic wavelength and re-emit light at a lower energy, or longer wavelength, as illustrated in Fig.A.4. The wavelength where photon energy is most efficiently captured is defined as the absorbance maximum whereas the wavelength where light is most efficiently released is defined as the emission maximum³³. Each fluorophore has four (possibly unique) dipoles: two permanent and two transient. The permanent dipoles coincide to the ground state (S0) and the excited state (S1), and the transient dipoles describe the transitions between those two states. The transient dipoles include an absorption dipole, which appears during the transition from the ground state to the excited state, and an emission dipole, which appears during the transition from the excited state back to the ground state². The transition from a ground state dipole to an excited state dipole typically occurs in less than 10⁻¹⁵ seconds because of the purely electronic nature of this process².



Fig.A.4: Spectral characteristics of a fluorophore.

The molecules 6-carboxyfluorescein (FAM), carboxy-tetramethyl-rhodamine (TAM) and Cy5 are fluorescent dyes and are exclusively used in this work as the energy

transfer elements. FAM is a derivative of fluorescein dye family, TAM is a derivative of rhodamine dye family, and Cy5 is a derivative of cyanine dye family. Dye molecular structures are illustrated in Fig A.5.



Fig A.5: Molecular structures of FAM, TAM, and Cy5; the three fluorophores used in the construction of DNA-based diffusive waveguides²¹.

Table A.1 lists the absorbance and emission maxima of FAM, TAM, and Cy5

measured in different buffer conditions as well as in different structural templates (e.g.,

individual dye or dye-labeled DNA oligonucleotides)

	Unconjugated dyes	Conjugated dyes	Conjugated dyes
	Abs(nm)/Em(nm)*	Abs(nm)/Em(nm)**	Abs(nm)/Em(nm)***
FAM	492/515	492/514	496/517
TAM	565/580	557/578	561/583
Cy5	643/667	648/657	656/667

Table A.1: Absorbance (Abs) and emission (Em) maxima of three fluorophores from different buffer conditions

* 10 mM Tris, 50 mM KCl, 5 mM MgCl2, pH 8.3³⁴ ** 50 mM Tris, 50 mM KCL, 5 mM MgCl2, pH 8.0³³ *** 40 mM Tris, 20 mM Acetic acid, 2 mM EDTA, 12.5 mM Magnesium acetate, pH 8.0 Bogh *et al.* reported that the absorbance or emission properties of fluorescent dyes are affected by their environment, including solvent, pH, and conjugation to other macromolecules³³.

APPENDIX B

PAGE Gel Filtration and Agarose Gel Filtration

B.1 PAGE Gel Filtration

The following description outlines the procedure for the PAGE gel filtration process used in this thesis work. A typical gel image of the assemblies formed by tiles with 1) FAM, 2) TAM, 3) Cv5, 4) FAM-TAM, 5) FAM-Cv5, 6) TAM-Cv5, and 7) FAM-TAM-Cy5 is illustrated in Fig.B.1. 10% polyacrylamide (PAGE) in 1x TAE Mg²⁺ buffer filled into the gel cask. Each gel lane was filled with 16 µL of about 1 µM dye-labeled DNA tiles and 4 μ L of bromophenol blue loading buffer. The PAGE gel was run in 1x TAE Mg²⁺ buffer at 15 V/cm for 2.0 h. The gel was imaged with an AlphaImager© (Alpha Inotech, San Leandro, CA). The fluorescence AlphaImager was used to show the emission response of FAM, TAM, Cy5, and the combination of dyes. Discrete bands are apparent and can be assigned to DNA tiles labeling with FAM, TAM, Cy5, and so on. Judging from the band position, all bands closed to the wells were well-formed dyelabeled DNA tiles, all bands away from the wells were malformed dye-labeled DNA tiles, and all bands at the right of the gel were the mixture of dye-labeled oligomers and their complementary strands. The structures from these bands were excised from the gel and analyzed by spectrophotometer.



Fig.B.1: 10% PAGE gel in 1x TAE Mg²⁺ dye-labeled DNA tiles with (1) FAM, (2) TAM, (3) Cy5, (4) FAM-TAM, (5) FAM-Cy5, (6) TAM-Cy5, and (7) FAM-TAM-Cy5

B.2 Agarose Gel Filtration

The following description outlines the procedure on how to perform agarose gel filtration. 2% agarose in 1x TAE Mg²⁺ buffer was microwave heated to boiling for 90 seconds and cooled for 5 minutes and filled into the gel cask. Each gel lane was filled with 16 μ L of about 15 nM DNA nanotubes and 4 μ L of bromophenol ficol loading buffer. The agarose gel was run in 1x TAE Mg²⁺ buffer at 8.33 V/cm for 90 minutes. After running, the gel was imaged with the AlphaImager. A typical gel image is shown in Fig.B.2. Judging from the band position, all middle bands were well-formed dye-labeled nanotubes; all bands on the right were the mixture of dye-labeled staple strands and unlabeled staple strands. The structures from these bands were excised from the gel and analyzed by spectrophotometer. To excise the band from the gel, a scalpel was used to cut all gel bands labeled "nanotubes" (Fig.B.2). Each gel band was crushed into small pieces using a micropestle (Eppendorf Inc.) and 1x TAE Mg²⁺ buffer (200 μ L) was subsequently added. The dye-labeled DNA origami nanotube solutions were kept at ambient conditions and the structures were left to diffuse from the crushed gel pieces into the buffer over the course of 12 hours. The nanotube solutions were separated from the crushed gel pieces using a pipette (Eppendorf Inc.).



Fig.B.2: 2% agarose gel in 1x TAE Mg²⁺ dye-labeled nanotubes with (1) FAM, (2) TAM, (3) Cy5, (4) FAM-TAM, (5) FAM-Cy5, (6) TAM-Cy5, and (7) FAM-TAM-Cy5. The majority of DNA nanotubes migrate as a single band in agarose-gel electrophoresis. This population presumably represents well formed nanotubes, whereas slower migrating species apparent on the gel presumably represent misfolded structures.

APPENDIX C

A Least-Squares Curve Fitting

C.1 A Least-Squares Curve Fitting

A Matlab script was used to determine the contribution of individual fluorophores to the bulk fluorescence measurement. The code was written based upon the least squares curve-fitting equation adopted from Ref.5:

$$S(\lambda) = f \cdot F(\lambda) + t \cdot T(\lambda) + c \cdot C(\lambda)$$
(C.1)

where $S(\lambda)$ is the data spectrum and $F(\lambda)$, $T(\lambda)$, and $C(\lambda)$ are the emission spectra of FAM, TAM, and Cy5 fluorophores, respectively. The three constant *f*, *t*, and *c* were determined using a least-squares fit in the regions which the spectra were measured.

Matlab Script:

```
%%
                                       %%
%%
         A least square curve fitting (BSU)
                                      %%
%%
                                       %%
clear all, close all
load filename.csv %format of csv: wavelength, intensity, intensity, intensity, intensity
I=filename;
                  %extract data for FAM from I
FAM=[I(:,1),I(:,2)];
TAM=[I(:,1),I(:,3)];
                   %extract data for TAM from I
                  %extract data for CY5 from I
CY5=[I(:,1),I(:,4)];
DATA=[I(:,1),I(:,5)];
                 %extract spectrum data from I
%%% begin fitting
tic
k=1e12;
sum=0;
for f=0:0.1:1
 for t=0:0.1:1
   for c=0:0.1:1
     for i=1:length(FAM)
       S=(f*FAM(i,2)+t*TAM(i,2)+c*CY5(i,2))^2;
       D=DATA(i,2)^2;
       diff=abs(D-S);
       sum=sum+diff;
     end
     if sum < k
       s1=f;
       s2=t;
       s3=c;
       k=sum:
     end
     sum=0;
   end
```

end end toc

% display fitting parameters s1, s2, s3 B=[FAM(:,1),FAM(:,2)*s1+TAM(:,2)*s2+CY5(:,2)*s3]; % display the result fitting curve and the data plot(DATA(:,1),DATA(:,2),'r',B(:,1),B(:,2),'k')

APPENDIX D

6-Helix Bundle Nanotube Design, Sequence Generator, Sequence List, Position of Biotin-Labeled Staple Strands, and Statistical Analysis of Quantum Dot Arrays

D.1. 6-Helix Bundle Nanotube Design

The design of the 6-helix bundle nanotube is provided in the following. As shown in Fig.D.1(a), the scaffold strand is arranged into six numbered helices with the ends of the M13mp18 located in the middle of helix 1. From the 5' end, the scaffold strand proceeds to the left with crossovers at the ends of the nanotube and two staggered crossovers near the middle. The nucleotide numbers for the crossovers are indicated in Fig.D.1(a). Staple strands are grouped into 86 columns and numbered from the left end, as shown in Fig.D.1(b). Staples strands are labeled according to the helix and column location of their 5' end. The first column of staple strands begins 14 nucleotides from the left end scaffold crossovers. Columns 11-13 show the 3 column repeating motif in which each staple strand consists of three 14 nucleotide domains complementary to a section of the M13mp18 scaffold strand and spans 3 helices. Although not used in the current study, three random 20 nucleotide sticky-ends, label A, B, and C, are added to staples in columns 4, 7, and 10 of helix 3. For each sticky-end, the helix 3 domain complementary to M13mp18 is lengthened by 7 nucleotides and the adjacent staple domain is correspondingly shortened, as illustrated in the figure.

The staple strand layout in the middle of the nanotube is shown in Fig.D.1(c). The 5' and 3' ends of the M13mp18 scaffold are located in helix 1 and staple column 43. Mid-nanotube scaffold crossovers are located in staple columns 39 and 41. The same A, B, and C sticky-ends are added to staples in columns 41, 45, and 48 of helix 3. Fig.D.1(d) illustrates the staple layout for the right end of the nanotube. A, B, and C sticky-ends are added to staple of columns 77, 80, and 83 of helix 3. Four nucleotides remain

77

unhybridized at the end of each helix. A schematic of the formed DNA nanotube with A, B, and C sticky-ends is shown in Fig.D.1(e).



Fig.D.1: Two-dimensional layout of the scaffold and staple strands of the DNA nanotube and 3D schematic. (a) Layout of the scaffold showing nucleotide numbers at the crossovers. (b) Staple layout for the left end of the tube. The staple motif is shown in columns 11-13. In helix 3, staples in columns 4, 7, and 10 are extended with sticky-ends labeled A, B, and C. (c) Staple layout in the middle of the tube. The M13mp18 scaffold begins and ends in helix 1, column 43. Scaffold crossovers are located at the ends and in columns 39 and 41. A, B, and C sticky-ends are added to

staples 41, 45, and 48 of helix 3. (d) Staple layout for the right end of the tube. Sticky-ends are added to staples in helix 3 in columns 77, 80, and 83. Four nucleotides remain at the end of each helix. (e) Schematic of the formed tube illustrating the A, B, and C sticky-ends along helix 3 of the formed nanotube. (Reprinted with permission from Ref.22. Copyright 2010 American Chemical Society).

D.2 Sequence Generator

To facilitate generation of the staple strands required to form a 6-helix DNA nanotube, a Perl script was written to layout the scaffold sequence into the geometric raster pattern according to a given set of turning points²². The script accepts the M13mp18 sequence as an input, divides the sequence into regions demarcated by turning points, and then outputs the scaffold and its complement according to the designed pattern. The complementary sequence is then easily divided into staple strands according to the desired motif.

Perl Script:

```
*****
      #
                     DNA Origami Design Helper Program
                                                                 #
      #
                                                                 #
                     Boise State University - May 2008
      #
                                                                 #
      *****
      use strict;
      use warnings;
      use English;
      use Carp;
      exit(main()):
                                          # a nice technique for avoiding global
variables in Perl
      sub main {
          my $DNA SEQUENCE FILENAME = 'M13mp18BP.txt';
          # Read the dna sequence data from the file named in $DNA SEQUENCE FILENAME.
          my $DNA SEQUENCE FILE;
          if (!open($DNA SEQUENCE FILE, "< $DNA SEQUENCE FILENAME")) {
             print "Unable to open data file $DNA SEQUENCE FILENAME\n";
             return 1;
                                                     # failure
          my @seqLines = <$DNA SEQUENCE FILE>;
                                                     # read the dna sequence data
lines into
                                                      # the seqLines array
          close $DNA SEQUENCE FILE;
          { # Temporarily change the record separator so that we can
              # get rid of the line ending characters that are found in
              # text files created on Windows.
             local $INPUT RECORD SEPARATOR = "\r\n";
             chomp (@seqLines);
          }
          my $n = scalar(@seqLines);
          print "Read $n lines from file $DNA SEQUENCE FILENAME.\n";
          # Convert the file data into an array of bases.
          shift @seqLines;
                                                    # discard the header line
```

map { \$ =~ s/[^GATC]+//g } @seqLines; # discard everything that isn't a G, A, T, or C my \$dnaSequence = join '', @seqLines; # concatenate the lines into one big string print 'dnaSequence has ' . length(\$dnaSequence) . " bases.\n"; # The scaffold structure is described as a set of lines and turnaround points. # Data structure: arefSubsequenceLines -> arefSubsequenceRanges -> [startPos, endPosl my \$arefSubsequenceLines = [[[595, 1], [7249, 6637]], [[596, 1162], [5996, 6636]], [[1729, 1163], [5995, 5355]], [[1730, 2268], [4686, 5354]], [[2807, 2269], [4685, 4017]], [[2808, 4016], 1,]; # Use the position data to divide the DNA sequence into the # base pair sequences that make up each line of the structure. my @designLines; foreach my \$arefSubsequenceRanges (@\$arefSubsequenceLines) { my \$designLine = ""; foreach my \$arefSubsequenceRange (@\$arefSubsequenceRanges) { my (\$p1, \$p2) = @\$arefSubsequenceRange;i8i8sx print "p1=\$p1, p2=\$p2; "; if (\$p1 <= \$p2) { # don't need to reverse the subsequence? # (i.e., should appear left to right?) \$designLine .= "*" if length(\$designLine) && \$p1 != 1; # separate subsequences # with an asterisk \$designLine .= substr \$dnaSequence, \$p1-1, \$p2-\$p1+1; else { # must reverse the subsequence because it must # appear right to left \$designLine .= "*" if length(\$designLine) && \$p2 != 1; # separate subsequences # with an asterisk \$designLine .= reverse substr \$dnaSequence, \$p2-1, \$p1-\$p2+1; } print "length(designLine) =" . length(\$designLine) . "\n"; } push @designLines, \$designLine; print "\n"; } my \$DESIGN OUTPUT FILE; if (!open(\$DESIGN OUTPUT FILE, "> tubedesign.txt")) { print "Unable to open output file tubedesign.txt\n"; return 1; # failure } my \$COMP DESIGN OUTPUT FILE; if (!open(\$COMP DESIGN OUTPUT FILE, "> comptubedesign.txt")) { print "Unable to open output file comptubedesign.txt\n"; return 1; # failure } # Format the output lines as needed for the next design step. foreach my \$designLine (@designLines) { my \$outputLine = \$designLine; \$outputLine =~ s/([GATC]{7})*([GATC]{7})/ \$1*\$2 /; # create a 7+7 subsequence at # the turnaround

points

```
$outputLine =~ s/([GATC]{14})/$1 /g;
                                                                              # break everything
else into
                                                                              # 14-base
subsequences
                $outputLine =~ s/ {3,}/ /g;
                                                                              # make sure there
are no more
                                                                              # than 2
consecutive spaces
                $outputLine =~ s/ $//;
                                                                              # get rid of the
dangling spaces
                                                                              # at the end of
the line
                 # Create the complements.
                my $complementLine = $outputLine;
                $complementLine =~ s/G/c/g;
                $complementLine =~ s/C/g/g;
                $complementLine =~ s/A/t/g;
                $complementLine =~ s/T/a/g;
                $complementLine = uc $complementLine;
                print $DESIGN_OUTPUT_FILE "$outputLine\n$complementLine\n\n";
print $COMP_DESIGN_OUTPUT_FILE "$complementLine\n\n";
            }
            close $DESIGN OUTPUT FILE;
            close $COMP_DESIGN_OUTPUT_FILE;
            return 0;
                                                        # success
        }
```

D.3 Sequence List

The following shows the staple strand sequence list for forming the 6-helix bundle

nanotube. Staple strands include 9 strands with 69 nucleotides, 9 strands with 35

nucleotides, and 152 strands with 42 nucleotides, the total of 170.

Table D.1: Name and sequence for the 170 staple strands used for the 6-helix DNA nanotube

Helix	Column	Sequence
1	2	GCCAGAGGGGGTAAAGACTCCTTATTACAACGCAAAGACACC
1	5	CAATACTGCGGAATAACGCAATAATAACATAGAAAATTCATA
1	8	AAATGCTTTAAACATAAGCAGATAGCCGCGACATTCAACCGA
1	11	AAAAATCAGGTCTTAAATAGCAATAGCTAAATTATTCATTAA
1	14	GCGGATTGCATCAACAAGAATTGAGTTAGCCATTTGGGAATT
1	17	CAAATATCGCGTTTAGTCAGAGGGTAATTTACCATTAGCAAG
1	20	GGAAGCAAACTCCAGAAGCGCATTAGACATAGCAGCACCGTA
1	23	TTGCTCCTTTTGATTGAAAATAGCAGCCTTAGCGTCAGACTG
1	26	GCTTAATTGCTGAACCCAATCCAAATAAATAGCCCCCCTTATT
1	29	ATATGCAACTAAAGGCCTAATTTGCCAGTCACCGGAACCAGA
1	32	AACAGTTGATTCCCTTTATCCTGAATCTCCGCCACCCTCAGA
1	35	ACCATTAGATACATCCTTAAATCAAGATGAGCCGCCACCAGA
1	38	TATATTTTCATTTGAGGCGTTTTAGCGAACAGGAGTTAGACT
1	41	TCTACTAATAGTAGCAAATCAGATATAGATCCTTTGCCCGAA
1	44	GCAAGGCAAAGAATTTTATTTTCATCGTATTATCATTTTGCG
1	47	GCATAAAGCTAAATATTAAACCAAGTACATTATCATCATATT
1	50	TAATACTTTTGCGGATCAATAATCGGCTAATATAATCCTGAT
1	53	AAAATTTTTAGAACAAAAATAATATCCCAGGGTTAGAACCTA
1	56	GTAATGTGTAGGTAAGAACGCGCCTGTTAGAAATAAAGAAAT
1	59	GACAGTCAAATCACTCTGTCCAGACGACTGAATATACAGTAA
1	62	TGATAAATTAATGCAGTAATAAGAGAATAACGGATTCGCCTG
1	65	TACAAAGGCTATCAAACAACGCCAACATGCGCAGAGGCGAAT
1	68	AAGAGAATCGATGACCAACGCTCAACAGAGATGATGAAACAA
1	71	CATATGTACCCCGGTTTAGTATCATATGTAACAATTTCATTT
1	74	GAAGATTGTATAAGATAAGAATAAACACATAAATCAATATAT
1	77	TTTGTTAAAATTCGTAATGGTTTGAAATCGTCGCTATTAATT
1	80	TTTTAACCAATAGGTTTCAAATATATTTAGCGATAGCTTAGA
1	83	CCTTCCTGTAGCCATGATGCAAATCCAAATTTATCAAAATCA
2	2	TATCATAACCCTCGCGTCTTTCCAGACGGTACAAACTACAAC
2	5	CATAACGCCAAAAGTTGCTAAACAACTTCCAATAGGAACCCA
2	8	TCAGTTGAGATTTAAAGGAACAACTAAACCACCCTCAGAGCC
2	11	AACGAACTAACGGATGAAAATCTCCAAAGGTTTAGTACCGCC
2	14	TATACCAGTCAGGAGTATCGGTTTATCAATATAAGTATAGCC
2	17	ATCATTGTGAATTAAGCTTGATACCGATTTTTGCTCAGTACC
2	20	CGAGTAGTAAATTGGCCCACGCATAACCAGAGGCTGAGACTC
2	23	TCATTCAGTGAATAGAGTTAAAGGCCGCTGCCTATTTCGGAA
2	26	AGAACCGGATATTCAAAGACAGCATCGGGTGCCTTGAGTAAC
2	29	GGCGCATAGGCTGGTTGAGGACTAAAGAGATGATACAGGAGT
2	32	TGACCAACTTTGAAGGGTAAAATACGTATCTCTGAATTTACC
2	35	GCCGGAACGAGGCGCAAAAGAGGCAAAACAAACAAATAAAT
2	38	GATAAATTGTGTCGCCCAGCGATTATACAGAAGTAGTTGAGG
2	41	TTTGCGTATTGGGCTCTTTTCACCAGTGTAATAGATTAGAGC
2	44	CCAGCTGCATTAATCGCCTGGCCCTGAGTTGAGGAAGGTTAT
2	47	GTTGCGCTCACTGCTTGCCCCAGCAGGCAATCAATATCTGGT
2	50	AGCCTGGGGTGCCTATCGGCAAAATCCCATCTAAAGCATCAC
2	53	CACAATTCCACACAGGGTTGAGTGTTGTGCCTGCAACAGTGC
2	56	ATCATGGTCATAGCAAGAACGTGGACTCAGCAGAAGATAAAA

2	59	СТСАСТСТАСАСССАСССАТСССССТАССССТААААСАТ
2	55	
2	62	CGTTGTAAAACGACTTTTTGGGGTCGAGCAATATTTTTGAAT
2	65	AGGCGATTAAGTTGAAAGGGAGCCCCCGAGAACCCTTCTGAC
2	68	TCTTCGCTATTACGAACGTGGCGAGAAACACACGACCAGTAA
2	71	TTCAGGCTGCGCAAGCTAGGGCGCTGGCAATCGTCTGAAATG
2	74	
2	77	
2	/ /	
2	80	GCATCGTAACCGTGAGAATCAGAGCGGGAATAACATCACTTG
2	83	GGATTGACCGTAATTTTAGACAGGAACGATCACGCAAATTAA
3	1	ATCTAAAGTTTTGTTTTACCAGACGACGGCAAAAGAAGTTTT
3	4	GAACGCACTTGGTCTACTGAATGAATTTTCTGTATGGGATTGAATTACGAGGCATGACTGGATAGCGTC
3	7	GTGACATACCTTCGGAGCATTTTCAGCGGAGTGAGAATAGAGGAATACCACATTCATT
3	10	
2	12	
2	10	
3	10	TGAATTTCTTAAACCCTTATGCGATTTTAGCCCGAAAGACTT
3	19	TGACAACAACCATCGGCTTGAGATGGTTAAGCGAACCAGACC
3	22	CTGAGGCTTGCAGGAGGCTTGCCCTGACGAGAGTACCTTTAA
3	25	CACCCTCAGCAGCGATTACCCAAATCAAGCGGATGGCTTAGA
3	28	CGGCTACAGAGGCTCTGACCTTCATCAATCAACATGTTTTAA
3	31	AGTTTCCATTAAACAGAGACAGATGAAGTTTCCATTCCA
3	31	
2	27	
3	37	ACTCATCHTIGACCAAATCCGCGACCTGATAACCTGTTIAGC
3	40	GAACGCACTTGGTCTACTGAGAAACAAAGTACAATGGTTTTGCCAGGGCGGAGATAAGGTGGCATCAAT
3	43	GTGACATACCTTCGGAGCATCAACAGCTGATTGCCCTTCACGAATCGGCCAACGCAATAAATCATACAG
3	46	CGCTTCACGAGGTTACAATGCAGCAAGCGGTCCACGCTGGTCCGCTTTCCAGTCGAATAAAGCCTCAGA
3	49	ATGGTGGTTCCGAAAATGAGTGAGCTAAACATTATGACCCTG
3	52	AATAGCCCGAGATAACATACGAGCCGGATCAACGCAAGGATA
3	55	ΔGΔGTCCΔCTΔTTΔTGTTTCCTGTGTGTGΔΔTGCCΔΔTGCCTGΔ
2	50	
2	50	
3	61	CACCCAAATCAAGTGGCCAAGTGCCAAGCTCAACCGTTCTAGC
3	64	TAAATCGGAACCCTGGTAACGCCAGGGTATTTTTGAGAGATC
3	67	GGGGAAAGCCGGCGCCAGCTGGCGAAAGAGTCTGGAGCAAAC
3	70	
	/0	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT
3	73	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCCAAAAACAG
3	73 76	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGGTACAAGATCGCACTCCAGTAAACGTTAATAT
3 3 3	73 76 79	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGGCAAAGCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGCTACAAGATCGCACTCCAGTAAACGTTAATAT
3 3 3	73 76 79	CGAAAGGAGCGGGCCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGCGGCAGCAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGGTCATCTGGCCAGTTTGTAAATCAGCTCATT CCCTTCACCACCACCACCACCACTACAACGTCCCCACTAATCCACCTCACCACTAATCAGCTCATT
3 3 3 3	73 76 79 82	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGCTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG
3 3 3 3 3 3	73 76 79 82 85	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG
3 3 3 3 <u>3</u> 4	73 76 79 82 85 1	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCCAACATATAAAAGAGCAGTATGTTAGCA
3 3 3 3 <u>3</u> 4 4	73 76 79 82 85 1 4	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCCGGAAACCAGGCAAAGCCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG
3 3 3 3 <u>3</u> 4 4 4	73 76 79 82 85 1 4 7	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCAATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGCGCGCACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATAAAAGGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG
3 3 3 3 3 3 4 4 4 4 4	73 76 79 82 85 1 4 7 10	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCAATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTAAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCCTCAGAACCGCGTAAATATGACGGATCTTACCGAAGCC
3 3 3 3 3 3 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13	CGAAAGGAGCGGGCCCGGTTGGGAAGGGCACTAGCAATGCCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGGTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCCAATAATAAG
3 3 3 3 3 3 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCAATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGGCGCGCTACAGGGCGGCGCACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGATAACGTGCTTTCCTCGTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGCCATTAATAAG ACCCCCCATTACCTCGCCCCCAATAATAAG
3 3 3 3 3 3 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16	CGAAAGGAGCGGGCCTGTTGGGAAGCGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGCTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTTCGTCACCGACTGGAGCCCAATAATAAG AGGCGGATAAGTGCCCCCACATAGAGCACATAATAC
3 3 3 <u>3</u> 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22	CGAAAGGAGCGGGCCTGTTGGGAAGCGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGGTCAGCTTGCAGATTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAACAAAGTAACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGGAGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCCCCGACTTGAAGCCCATAAATAAC AGGCGGATAAGTGCCCCCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAGGATCAATGAAACCATCGGGGGAAATTAACTG
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22	CGAAAGGAGCGGGGCCTGTTGGGAAGGGCACTAGCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCGCTACAGGGCGGCGCACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACCAGAAGAGAAT
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25	CGAAAGGAGCGGGGCCTGTTGGGAAGGGCACTAGCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGCGCACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTCAAAGACAAAAGGGAACTAACCCAAAAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACCAGAGAGAAT AGTGCCCGTATAAAACGGCATTTTCGGTCGAAACGATTTTTG
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28	CGAAAGGAGCGGGGCCCTGTTGGGAAGGGCACTAGCAATGCCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATAAAAGGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCAGAACGCGTAAATATGACAGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATGACGGACCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGGCCCAATAATAC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTCCGGTCGAAACGATTTTTTG GTACTGGTAATAAGTTTCATAATCAAAATTACAAAATAAACA
3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31	CGAAAGGAGCGGGGCCTGTTGGGAAGGGCACTAGCAATGCCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATCTCCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCAGAACGCGTAAATATGACAGGAACAAAGTTACCAG ACCCTCCAGAACCGCGTAAATATGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCCCCAATAATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTCCGGTCGAAACGATTTTTTG GTACTGGTAATAAGTTTCATAATCAAAATTACAAAATAAACA GTTCCAGTAAGCGTCGCCTCCCTCAGAGTACCAACGCTAACG
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34	CGAAAGGAGCGGGCCTGTTGGGAAGCGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGCGCACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTCGTCGGATCTCCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCCCAATAATAC CTCAAGAGAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTCCGGTCGAAACGATTTTTG GTACTGGTAATAAGGTTTCATAATCAAAATAAACA GTTCCAGTAAGCGTCGCCTCCCTCAGAGTACCAACGCTAACG CTCATTAAAGCCAGAGCCACCACCCTCATAGTTGCTATTTG
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37	CGAAAGGAGCGGGGCCTGTTGGGAAAGCGCAAAGCCCCAAAACAAG GAACGCACTTGGTCTACTGAGCGCGCGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCGCGCAAGGCCGATAAAGGGCGGATAAGGTCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTAATAAGGTCATCAAAATTACAAAATAACA GTTCCAGTAAGCGTCGCCTCCCCAGGAGCAACGACTATTTG GTACTGGTAATAAGTTTCATAATCAAAATTACAAAATAACA GTTCCAGTAAGCGCGCCCCCCCCCC
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40	CGAAAGGAGCGGGGCCTGTTGGGAAGCGCAAAGCCCCAAAAACAAG CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCGCTACAGGGCGGCGCACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCACCCTCATTTTCAAAGACAAAAGGGAACCAAAGTTACCAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTGTCCCTAATACAAATAAACA GTTCCAGTAAGGTTCGCCCCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCTGCGCCCCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCGTGGCCCCCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCGAGGCCCACCACCACCACGCTAACG CTCATTAAAGCAGAGCCCACCACCACCACCACCACGCTAACG CTCATTAAAGCAGAGCCCACCCCCCCTATAGTTGCCTTTTG CAGGTCAGCAGATTCGCCGCCACCACCACCACCCCCCTATCGC CCCCAATACATCAACGCCCCCCCCTATAACCACCCCCCCC
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40	CGAAAGGAGCGGGGCCTGTTGGGAAGCGCACTAGCAATGCCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTCAAAGACAAAAGGCAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTTCGGTCGAAACGATTTTTG GTACTGGTAATAAGGTTCCTCCCACAGGAGCAACGACTTTTTG GTACTGGTAATAAGGTTCCTCCCTCAGAGTACCAACGGCTAACG CTCATTAAAGCGAGGCCCCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCGAGGCCCCCCCCCC
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43	CGAAAGGAGCGGGCCTGTTTGGGAAGGGCACTAGCATGTCAAT CGCTGCCGGTAACCCGGAAACCCAGGCAAAGCCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACCAAGAGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTCGTCGGATTCCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTCAAAGACAAAAGGGAACAAGTTACCAG ACCCTCAGAACCGCGTAAATATGACGGATCTTACCGAAGC CGGAATAGGTGTATCGTCACCGACTTGAAGGCACCATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTTCGGTCGAAACGATTTTTG GTACTGGTAATAAGTTTCATAATCAAAATTACAAATAAACA GTTCCAGTAAGCGCCGCCTCCTCAGAGTACCAACGCTAATTG CCTCATTAAAGCCAGAGCCACCCACGCTAAGTGCCTATTTG CAGGTCAGACGATCGCCCCCCCCATAGTTGCCTATTTG CAGGTCAGAAGGCACCACCCCCCCATAGTGCCTATTTG CTCATTAAAGCCAGAGCCACCACCCTCATAGTGCCTATTTG CAGGTCAGACGATTCGCCGCCACCCCCCATAGTGCCTATTTG CAGGTCAGACGATCACCACCGCCACCCTCCCGACTGC CGTCAATAAGATCAACTGGCACTTGACCTCCCGACTTGC CGTCAATAGATAATACAACTGGTATTAAAAGCTTATCCGGT CTAAAAATATCAAACTGGTAATAAAAGGCTTATCCGGT CTAAAAATATCTTTAAAAGTTGGATAACAGGAATCATTACCG
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46	CGAAAGGAGCGGGCCTGTTGGGAAAGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCGCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCGCGCTACAAGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACCAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAAGGCACATTGGGTCGAAACGATTATTTG GTACTGGTAATAAGGCCACCTCAGAGTACCAACGGCTAACG CTCATTAATCCAAAGGTTTCCACTAAGATATACAAATAACA GTTCCAGTAAGCGCCCCCCCCAGAGTACCACCGCTAACG CTCATTAAAGCCAGGCCCCCCCCAGAGTACCACCGCTACG CCTCATTAAAGCCAGGCCCCCCCCAGAGTACCACCGCTACG CTCATTAAAGCCAGGCCCCCCCCAGAGTACCACCGCTACG CTCATTAAAGCCAGAGCCCCCCCCAGAGTACCACCGCTACG CTCATTAAAGCCAGAGCCCCCCCCCAAGAGTACCACGCTACG CTCATTAAAGCCAGAGCCCCCCCCAGAATGACCACCGCACTGC CGTCAATAAGATAATACAACTTGAGTACCACCGCACTGC CGTCAATAGATAATACAACTTGGATATTAAAAGGCTTATCCGGT CTAAAATATCTTTAAAAGTTTGAGTAACAGGAATCATTACCG CAGTTGGCAAATCACCAGAAGGAGCGGACCACCTCATCGAGA
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49	CGAAAGGAGCGGGCCTGFTGGGAAGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCGCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACCATACCTTCGGAGCATGATAAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCACCCTCAGTATCGCGGTAATATGAGGGACCAAAGTTACCAG ACCCTCCAGAACCGCGTAAATATGACGGAACTAACCGAAGC CGGAATAGGTGTATCGTCACCGACTTGAAGCGCCAATAATAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGACAGAGCACCAGTAGCACCATGAGCGCCTAATACG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTTCGGTCGAAACGATTTTTG GTACTGGTAATAAGTTTCATAATCAAAATTACAAAATAAACA GTTCCAGTAAGCGCGCCCCCCCCAGAGTACCAACGCTAACG CCTCATTAAAGCGCAGGCGCCACCACCCTCATAGTGCTATTTG CAGGTCAGACGAGGCCACCACCCTCATAGTGCTATTTG CAGGTCAGACGATGCCCCCCCACAGGAACCAACGCTAACG CTCAATAAGATAATACAACTCGTATTAAAAGGCTTATCCGGC CGTCAATAGATAATACAACTCGTATTAAAAGGCTTATCCGGT CTAAAATATCTTTAAAAGTTGACATACGGAATCACTGCGC CGTCAATAGATAATACAACTGGAACCGACCCTCATCGAGA CTTGCTGAAACCCCAGAAGGAGCGGACCGCCCCCCTCATACGGC CAGTTGGCAAATACACACCGGAACGAATCACTCCGGT CTAAAATATCTTTAAAAGTTGACAGCACCATCATCCGG CAGTTGGCAAATACACCCAGAGGAGCGGACCGCACTCATCGAGA CTTGCTGAACCCCAGAAGGAGCGGACCGACTCATCCTTATCG
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52	CGAAAGGGGGCCTGTTGGGAAGGGCAATGCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCACCCTCATTTTCAAAGACCAAAAGGGAACAAAGTTACCAG ACCCTCAGAACGCGGTAAATATTGACGGATCTTACCGAAGC CGGAATAGGTGTATCGTCACCAGTAGCACAAAGGCACAAAGTAACG AGCGGGATAAGTGTCCCACCAGTAGCACCATGAGCGCAATAATA AGGCCGGATAAGTCCACCAGTAGCACCATGAGCGCAATAATC CTCAAGAGAGGATCAATGAAACCATCGGGGAGAATTAACTG CTCAAGAGAGGATCAATGAAACCATCGGGCGAAACGATTTTTG GTACTGGTAATACAAGTTTCATAATCAAAATTACAAAATAAACA GTTCCAGTAAGCGCCGCCTCCCTCAGAGTACCACGGCTAATTC CTCAATAAGCCAGAGCCACCACCCCCATAGTGCCATTG GTACTGGTAATAAGTTTCATAATCAAAATTACAAAATAAACA GTTCCAGTAAGCGTCGCCCCCCCAAGAGTACCACGCCTACG CGCAATAGGATAAGACCACCGCCACCCCCCATAGTGCCATTG CAGGTCAGACGATCGCCCCCCCCATAGTGCCATTTG CAGGTCAGACGATCCACCGCCACCCCCCATAGTGCCATTGC CGTCAATAGATAACAACTCGTATTAAAAGGCTTATCCGG CGTCAATAGATAACAACTCGTATTAAAAGGCTTATCCGG CAATAGATAATACAACTCGTATTAAAAGGCTATTACCG CAGTTGGCAAATCACAGAAGGAGCGGACCCACCTCATCGAGA CTTGCTGAAACCCCAGAAGGAGCGGACCCACCCTCATCGAGA CTTGCTGAAACCCCAGAAGGAGCGGACCCACCCTCATCGAGA CTTGCGAAATCACCAGAAGGAGCGGACCCACCCTCATCGAGA
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55	CGAAAGGAGCGGGCCTGTTGGGAAACAGGCCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACCATACCTTCGGAGCATGATAACGTGCTTTCCTCGTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTACCAG ACCCTCAGAACCGCGTAAATATGACGGAACCAAGGTACCAG ACCCTCAGAACCGCGTAAATATGACGGAACCATGACCAG ACCCTCAGAACCGCGTAAATATGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGCACTGAGCCCCAATAATAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTTCGGTCGAAACGATTTTTG GTACTGGTAATAAGGTCCACCACGAGGACCAACGCTAACG CTCATTAAAGCGCGCCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCGCGCCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCGCGCCCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCGCGCCCCCCCCAGAGTACCAACGCTAACG CTCAATAAGATAATACAACTCGTATTAAAAGGCTTATCCGGT CTAAAAATACTTTAAAAGTTGGATACAAGGAATCATTACCG CAGTCGGAAATCACACCGGAAGGAACCACTCCGGAC CGTCAATAGATAATACAACTCGTATTAAAAGGCTTATCCGGT CTAAAATACTTTAAAAGTTGGATACAGGAATCATTACCG CAGTTGGCAAATCACCAGAAGGAGCGGACCCACTCACTCGAGA CTTGCTGAAACCCAGAAGGAAGCGGACCGACTCATCCTAATGAA CACGCTGAGGCCATTCGCAGTATCAACGAATCATTACCA CACGCTGAGGCCATTGCCAGAATCATTCCTTATCA CACGCTGAGGCCATTGCACGTAATCAACGAATCATTACCA
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58	CGAAAGGAGCGGCCCGGTAGCGAAGCGCGCTACAGGCCCAAAAACAG GAACGCACTTGGTCACCGGAAACCAGGCCGCTACAGGCCCCAAAAACAG GAACGCACTTCGGTCTACTGAGCGCCCGCTACAGGCCGCTACAAGACCGCCCCGTTAAATCAGCTCATT GTGACATACCTTGGGAGCATGTATAAAGGGCGTTTCCTCGTTCATCTGCCAGTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGGCCACATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATGACGGAACCATCAGCAACAAAG ACCCTCAGAACCGCGTAAATATGACGGAACTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTTGACGCCCAATAATAAG AGCCGGATAAGTCCACCACGGGCGCAATTAACG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAACGGCATTTTCGGTCGAAACGATTATTG GTACTGGTAAAAAGGGCATCATCAGGAGCAATTAACAG GTTCCAGTAAGCGCCCCCCCCAGAACGATTATTTG CTCAAGAGAAGCATCAGCGCCACTACACGCTAATTACG CTCATTAAAGCGCCGCCACCACCCTCATAGTGCTATTTTG GTACTGGTAATAAGGCACCTCCAGAGCACATGACG CTCATTAAAGCGCGCCCCCCCCAATAATAACAG GTTCCAGAAGCAACCACCGCCACCACCCTCATAGTGCTATTTG CAGGTCAGACGATTCGCCGCCACCCCCCAATGATGCCACACG CTCAATAGATAATACAACTCGTATTAAAAGGCTTATCCGGT CTAAAATACTTTTAAAAGTTTGAGTAACAGGAATCATTACCG CGTCAATAGATAATACAACTCGTATTAAAAGGCATTACCGG CCTCATTAGCAGAGCCACCACCCCCCATACTTACCGGT CTAAAATACCATCACCAGAAGGAGCGGACCCACTCATCGAGA CTTGCTGAAACCACCGGAATCAATCACGGACCACTCATCGGA CCTTGCGGAAATCACCAGAAGGAGCGGACCCACTCATCGAGA CTTGCTGAAACCATCGCAATCATCATCTTTCCTTACCA CACGCTGAAGGCCATTCCCAGAAGAACCAACAACA
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 9 52 55 58 61	CGAAAGGAGCGGCCCGGTAGCAAGGGCACTAGCAAGGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTCCTTTCCTCGTTCATCTGCCAGTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAGGGAACAAAGTTACCAG ACCACCCTCAGAACGCGTAATATTGACGAGATACCCAG ACCCCCCAGAACGCGTAAATATGGCGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGACACCATCGGGGGACATTAACTG CCTATTATTCTGAAAATCAAGTTGCCTTTTACAGAGAGAAT AGTGCCCGTATAAAGGCCACCATGAGCGCTAATATC GTCCAGTAAGGGCCCCCCCCCC
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 55 8 1 40 43 46 49 52 55 8 6 4	CGAAAGGACCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAAGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTAAGGTCACGTTAATTCGCGTCTGG TGAGAAGTTTTTCGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCAGAACCGCGTAAATATTGACGGATCTTACCGAAGC CGGAATAGGTGTATCGTCACCAGTAGCACCATGAGCGCCAATAATAC AGCCGGATAAGTGCACCAGTAGCACCATGAGCGCCAATAATAC CTCAAGAGGATCAATGAAACCATCGGGGAGATTAACTG CCTATTATTCTGAAAATGAACCATCGGGGAGATTAACTG CCTATTATTCGAAAATGAACCATCGGGGAGACTAATCG GTACTGGTAATAAGGTTTCCTCAAGAGGACCAACGGCACTAACG CTCATTAAAGCGCGCCCCCCCAGAGCACCAAGGAGCACAAGG CTCATTAAAGCGCGCCCCCCCCAGAGTACCCCGACTACG CTCATTAAAGCGCGCGCCCCCCCCAAAAATTACAAAATAAACA GTTCCAGTAAGGGCGCCCCCCCCCATGACGCTAATTG CTCAAGAGGATCAATGAACCACCGAGAACCAATGGCTAATG CTCAATAAAGCTGCGCCCCCCCCAATGACCTATTTG CAAGTCAGACGATTCACCAACGATCATTACCGACCAACG CTCATTAAAGCTGCGCCACCACCCCCAATGACCTATTTG CAGGTCAGACGATCACCGACACCACCGCACTCATCGGGT CTAAAATATCTTTAAAAGTTGGACTACCCGACATCATTACCG CAGTTGGCAAATCACAGAATCATTAACGGCATCTTACCGACA CTTGCTGAAACCACGAATCATCATCGTCTTTCCTTATCA CACGCTGAGAGGCGATTGCCGCACCACTCATCTACCGACA CTGCTGAAACCACGAATACAATAACAACAACA CCCCTTTAAAATAGGTTTAACGCCAACACTCAACAACAACA CGCCTATAAAATAGGTTTAACGTCAGAGAACAATAAACAACA
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 67	CGAAAGGACCGGGCCTGTTGGGAAGGCACTAGCATGTCAAT CGCTGCGCGTAACCCGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAAGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTTGTAAATCAGCTCATT CGCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGGTACCAGTTTTCTCTCGATCTCCGTAAATGTGAGCGGAG GCCTGTAGCATTCCCAACATATAAAAGAGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCTCCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGGTGTATCGTCACCGACTGGAGGCCCAATAATAAG ACCACCCCCATATCGTCACCGACTGGAGGCCCAATAATAAG AGGCGGATAAGTGCACCCAGTAGCGCTAATATC CTCAAGAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATGAAACCATCGGGGCGAAATTAACTG CTCATTAAAGGCATTTCCGTCGAACGACTATTTTTG GTACTGGTAATAAGGTCATCAAGTTTCCCTTTTACAAGAGAAAT AGTGCCCGTATAAACGGCACTTTCGGTCGAAACGATTTTTG GTACTGGTAATAAGGTCACCACCCTCATAGTGCTATTTG CAGGTCAGACGATTCGCCCCCCCAGAGTACCACCGCTAACG CTCCATTAAAGCCAGGGCCACCACCCTCATAGTGCTATTTG CAGGTCAGACGATTGCGCCCCCCCTCATAGTGCTATTTG CAGGTCAGACGATTGCGCCCCCCCCAAAGGCTCACCGGT CTAAAATACTTTTAAAAGTTTGAGATACAGGAATCATTACCG CACGTGGAAATCACCAGCGCCACCCCCTATAGTGCTTTCCTGGTG CACGTGGAAATCACCAGAGGCGACGCACTCATCCGGT CTAAAATACTTTTAAAAGTTTGAGTAACAGGAATCATTACCG CACGTGGAAATCACCAGAGGAGCGACGCACTCATCCAGAA CTTGCTGAACCCTCGAGAAGGACCGACCTCATCTCCGACT CCCATTAAAAGTTCGACAATATAGAATCAACAATTACAA GCTCATTAAGGCTTTACCGCAAAACTATCAACAATTACAACAACA GCTATTAGGTATCACGGCAGAACAATATAAAGAACAACA GCTATTAGGCAAATTAAAAGGATTTCACGACAATAACAACA GCTATTAGGCTAAAACAATACAACAACAACAACAACAACAACAACAACA
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 776 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 55 58 61 64 67 7	CCAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGGTAACCGGGAAACCAGGCAAGCCCAGCAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAAGGCCCCAAAAACAG GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTAAATCAGCTCATT CGCTCACCAGGGTTACAATGCAGGAGCCCGATTAAAGGGCAGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGGTGTTTTTCGCGAATCTCCGTAAATGTGAGCAG GCCTGTAACACTTTTTGTCACAATAAAAGGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATAAAAGGGAACAAAGTTACCAG ACCACCCTCATTTCCAACATATAAAAGGGAACAAAGTTACCAG ACCACCTCCAGAACCGGCTAAATATGAGCGATCTTACCGAAGCC CGGAATAGTGACCACCGCTACAAAGGGAACAAAGTTACCAG ACCCTCCAGAACCGCGTAAATATTGACGGATCTTACCGAAGCC CGGAATAGTGCACCACCAGTAGCACCATGAGCCCAATAATAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCCCAATAATAG AGGCCGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATGGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATGGCGCAAATAACAG GTTCCCGGTAATAAGTTCCACGACACCATGACGCCTAATATG GTACTGGTAAAAACAAGTTTCCGGCGAGAACTATTTG GTACTGGTAAAGCGCCCCCCCCCAAGATCACACGCTAACG CTCATTAAAACGGCCACCACCACCCTCAAGATCACACGCTAACG CTCAATAAAGCCAGAGCCACCACCCTCAAGATCACACGCTAACGG CTCAATAAGCCAGAGCCACCACCCCCCAAGATCACCGCTAACG CTCAATAAGTTTGAAAACTGGAATCAACAGGAACCATTACCG CGTCAATAAGATATACAACTGGAAACAGGAATCATTACCG CAGTGGCAAATCACCAGAAGGACGCACCCACCTCATCGGAA CTTGCTGAAACATAACAAGTTGAACAGGAATCATTACCG CAGCTGAAACACCAGAAGGACGGACGCACTCATCGGAA CTTGCTGAAACCACCAGAAGGACGCAACCATTACCGA CAGGCTGAGAGCCATTCTGAAAATGGAATCCTTACTACACAATTACA CACGCTGAAGCCATTCTGAAAATGGAATCCTTACTACACAATACAA CGCCATTAAAATACGCCAGAAGCACGAATCATTACAACAATTACAA CGCCATTAAAAATAGGTTTAACGACAAGAACCATTACCGACA CTGAAAGCGTAAGACAAGTTACCAGAACAATTAAAAGAACCAACA CGCATTAAAAATAGGATTACAAAATCGTAATTAAGAACCAACA CTGAAAGCGTAAGACAAGTTACCAAAAATCGTAATTAGGACCAACCA
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 776 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70	CCAAAGGAGCGGGCCTGTTGGGAAGGGCACTAGCATGTCAAT CGCTGCGCGTAACCGGGAAACCAGGCAAGCCCACTAGCACAGAGCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTAGTAACGGTCATT CCCTTCACGAGGTTACAATGCAGGAGCCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAQAAGGACCGGTTACAATGCAGGAGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG GCCTGTAGCATTCCCAACATATAAAGGCCACTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCCTCCAGAACCGCTAAATATGGAGCGCACTATGTTACCAG ACCCTCCAGAACCGCTAAATATGGCAGCCCAATAATAAG ACCCTCCAGAACGCGTAAATATGGCACTTACCGAAGC CGGAATAGGTGTATCGTCACCGACTTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGAATCAATGAAACATCGGGGGAGAATTAACTG CTCATTAATCTGAAAATCAAGTTTGCGTTTTCGGGGAGAATTAACTG CTCATTAAAGGCCACTACTGAGCCCACACGCTAACG GTTCCAGTAAAAGGCCACTCCTCAGAGTACCAACGCTAACG CTCCATTAAAGCGCCGCCCCCCCCAGAGTACCAACGCTAACG CTCCATTAAAGCGCACTCCGCCAGCATCACCGCCAACGCTAACG CTCCATAAAGCGCACTCCGCCAGCATCACCGCCACCTCATGCC CGCCAATAGATAATCACACTCGTATTAAAAGGCTTATCCGGT CTAAAATATCTTTAAAAGTTTGGATAACAGGAATCATTACCG CAGTGGCAAATCACCAGAGGGCGCGCCGCCCCCCCCCACGGT CTAAAATATCTTTAAAAGTTTGGATAACAGGAATCATTACCG CAGTGGCAAATCACCAGAGGGACGGACCCACTCATCGAGA CTTGCTGAACCCCGAGATGCACCACGCTAATCGACACCGAC CCCCATTAAAACCCAAGGGAATCATTCAACAATAACA GCCTATAAGTTTCGAAAACTATCAACAATAACA GCCTATTAGGCAAGTTCACAGAAACAATAAACAACA GGCTATTAGGCAAGTTACACAAAACAATAAACAACA CTGAAAGGGACATTCACCGGAAACCAATAAACAACA CTGAAAGGGACATCCACCGAAAACCAATAAAAGAACAACAACA CTGAAAGGCAAGTTCACCAGAAGAACAATAAAAGAACAACAACAACAACAACAACAACAACAACAACAA
3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73	CCAAAGGAGCGGGCCTGTTGGGAAAGGCACTACCATGTCAAT CGCTGCGCGTAACCGGGAAACCAGGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGGCGCCGCTACAGGGCGCGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTCGTCATCTGCCAGTTGTAAATCAGCTCATT CCCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAAGGATTTTCCCAACATATAAAAGAGCAGTATGTGAGCA GCCTGTAACACTTTTTCGTCACAATCAGGAATACCCAAAAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCACCCTCATTTTCAAAGACAAAAGGGAACAAAGTTACCAG ACCCCCCAGAACGCGTAAATATGACGGATCTTACCGAAGC CGGAATAGGTGTACCGCCACTGAAGCCCAATAATAAG AGGCGGATAAGTGCCACCAGTAGCACCATGAAGCCCAATAATAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGGAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTATTATTCTGAAAATCAAGTTTGCCTTTTACCAGAGGAAT AGTGCCGTATAAAGGGCACTACCGGGGAGAATTAACTG CTCAATGAAGGGCACCACCACCCTCATGAGCCCTAACG GTTCCAGTAAAAGGGCACTTTCGGTCGAAACGATTTTTG GTACTGGTAATAAGGTTCAATAAACAATTACAAGTTTCCCTAATG CTCAATAAGCGAGGCCACCCCCCCACAGGCCAACGCTAACG CTCATTAAAGCGAGGCCACCACCCTCATAGTTGCTATTTG CAGGTCAGACGATCGCCGCCAGCATTCACCTCAGAGTACTTACCG CGTCAATAGATAATACAACTGTATTAAAAGGCTTATCCGGT CTAAATACCTTTAAAAGTTCGACACCGCACTCATCGGGA CTTGCTGAACCCAGAAGGAGCGGACGCACCACCTCATCGAGA CTTGCTGAACCCAGAAGGAGCGGACGCACCACCTCATCGAGA CTTGCTGAACCCAGAAGGAGCGGACGCACCACCTCATCGAGA CTGCTGAACCCCAGATCCACGTATATAAAGGATTCCCTAATCA CACCGCTAAAAATGGCTTTACAGAAATCATCACACAATAGAT CCCCATTAAAATGGTTTAACGTCAAATTAAAAGTACCAAATAGAT CCCCATTAAAATAGGTTTAACGTCAAATATAAAGTACCGACA CTGAAAGCGTAAGACAATTAAAAGGAACAATATAAAGTACCGACA CTGAAAGCGTAAGACAACAATAAAACAACA GGCTATTAGCTTTCGGAAAACAATATAAAGTACCGACA TAAAAGGGACATTCACCTGAGCAAACAATATAAAGTACCGACA CTGAAACCGTAAGACAAGTACCAAAATATAAAGTACCGACA CTCAAAGCGTAAGACAAATAACAACAACAACAACACA GGCTATTAGCATTGGAAAACAATACAAATACAAATTACAATCG TCAAGGAAATCCCTAATGGAAACAATAACAACAACAACACA GATTATTACATTGGAAAACAATACAAATACAAATTACAATCC TCAAAGGAAATCCCTAAGGAACAATATAAAAGGCTTAATTAC
3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 776 79 82 85 1 4 7 10 13 16 19 22 52 8 31 34 37 40 43 46 49 55 58 61 64 67 70 73 76	CCAAAGGAGCGGGCCTGTTGGGAAGGGCACTACCATGTCAAT CGCTGCGCGTAACCGGGAAACCAGGCAAAGCCCCAAAACAG GAACGCACTTGGTCTACTGTGGGCCCGCTACAGGCCGCTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTGGGAGCATGTATAACGTGCTTTCCTCGTTCATCTGCCAGTTGTAAATCAGCTCATT CCCTTCACGAGGTTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCTGG TGAGAAGTGTTTTTCCTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACAATATAAAGAGCACTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGGAATACCCAAAAG ACCCTCCGAACCGCGTAAAATGTGAGGGCACAAAGTTACCAG ACCCTCCGAACCGCGTAAAATATGACGGATCTTACCGAGCC CGGAATAGGTGTATCGTCACCGACTGAAGCGCTAATATC CTCAAGAGAGGATCAATGAAGACCATCGGGGACAATATACG CCTATTATCTGAAAATCAGATTGCCTTTACCAGAGGAAT AGGCGGATAAGTGCCACCAGTAGCACCATGGGGCCTAATATC CTCAAGAGAGGATCAATGAAACCATCGGGGACAATATACG CCTATTATCTGAAAATCAAGTTTGCGTTTACCAAGAGAGAT ASTGCCGGTAATAAGGCCACTTCGGGACAACGATTTTTG GTACTGGTAATAAGTTCATAATCAAAATTACAAAATAAACA GTTCCAGTAAGGCCACCCCCCCAGAGTACCAACGCTAACG CTCATTAAAGCCAGAGCCACCACCTCAAGGTACCAACGCTAACG CTCATTAAAGCCAGAGCCACCACCTCAAGGTACCAACGCTAACG CTCATAAAGCCAGAGCCACCACCTCAAGGTAACCAATTACCG CGTCAATAGATAATCAAACTCGTATTAAAGGCATATTCG CCAGGTCAGACGATTCTTCAGGAAGCACCACCACCGACTTGC CGCTCAATAGATAATCAAACTCGTATTAAAGGCATCTCCGGA CTGCAAAAGAACAATCTCAAGAGGAGCGCACCACCACCAATGCACCACCGAC CAGGTCGGAAATCACCAGAAGGGAGCGCACCACCATCATTACCG CAGGTGGAAACCACTCGAATGCACACCACCACGAATCCAATTACCG CAGGTGGGAAATCCCCAGAAGGAAGCAATCATTACCG CAGGTGGGAAATCACCAGAAGGAGGACCACTAATTAACGA CAGAGGTAGGCGGTTTCCACGTAAAACTATCAACAATGGA CGCAATTAAAAATAGGTTAACAGAACAATATAAAGAACAA CGCAATTAAAAAAAGGTTACAAAATCGTAATTAAGGAA CAGAGGTAAGCCAATTCACCGACACCACAACCAA CTGAAACGGTAAGACAAGTTACAAGATAGGACTAATTAGG ATATTTACAATGACTATAATAATGGAATCCTAATTAGG ATATATTACATTGCTTCCGGAAAAGATAGGCTTAATTGG GATATTTACATTGCATTAATAATACTGGTATATCAAAATC CCAGAACAATTATATTACCTCGTTAATCGACAATATAATCC CCAGAACAATTATATTGCTTCTGTAAAACCGACCGGGAT
3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 76 79 82 85 1 4 7 10 13 16 19 22 55 8 1 40 43 46 49 52 55 8 61 64 67 70 73 76 79	CGAAAGGAGCGGGACCTGTTGGGAAGGGCACTACCATGTCATT GCGTGCGCGTAACCGGAAACCGCGCAAAGCCCCAAAAACAG GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGTACAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCATT CGCTTCACGAGGTTACTACAATGCAGGAGGCCGATTAAAGGGATGGGATAGGTCACGTTAATTCGCGTCAGT TGAGAAGCACTTTTGGTCGGATTCTCCGTAAATGTGAGCGAG GCCTGTAGCATTCCCAACATATAAAAGGCAGTATGTTAGCA TGTACCGTAACACTTTTTGTCACAATCAGAATTACCCAAAAG ACCACCCTCATTTTGTCACAATCAGGATCTTACCGAAGC CGGAATAGGTGTATCGTCACAGACAGAGGACCATAATAAG ACCACCCCTCATTTTGTCACGACACAGGGACCATAATAAG ACCACCCGGAAACGGCCACCAGTAGCACCATGAAGGCCCAATAATAAG ACCCCCAGAACGGCCACCAGTAGCACCATGAAGCCCCAATAATAAG ACCGCCGGATAAGTGCCACCAGTAGCACCATGAGCGCCAATAATAG AGGCCGGATAAGTGCCACCAGTAGCACCATGAGCGCTAATATC CTCAAGAGAAGGATCAATGAAACCATCGGGGAGAATTAACTG CCTCATTATTCTGAAAATCAACATCGGGGAGAATTATTTG GTACTGGTAATAGGTTTCATATCGAAACGATGTTTTTG GTACTGGTAATAGGTTTCATATCCAAAATTACAAAATTAACA GTTCCAGTAAGCGTGCCCCCCTCATAGTTGCCTTTTCCGACGC CCTCATTAAACGGCGCCCCCCTCATAGTTGCCTTATCG CCTCAATAGATAATACAACGGCAATGACTATCCGG CTCAATAGATAATACAACTCGTATTAAAGGAATCATTACCG CCTCAATAGATAATACAACTCGTATTAAAAGGCACTATCCGG CCTCAATAGATAATACAACGGAATCATTACCGG CCTCAATAGATAATACAACGGAATCATTACCGG CCTCAATAGATAATACAACGGAATCATTACCGG CAGTGGCAAATCACCGAAGGACCGACCCACTCATCGAGA CTTGCTGAAGCCTCAGAGGCACGACCACTCATCCGAGA CTGCTGAAGCCACCCCCAGAATCATCAACAGGAACCATTACCG CAGGTGAGAGCCATTCGTGAAAAATGGAATCATTACCGAC CCGCATTAAAATAGGTTTAACGTAACAGGAACAATATAACAACA GGCCATTAAATGTTTCAACAGAAACATATAAACAACA CGCCATTAAAATAGGTTAAACGTAACAGAACAATAACAACA CCGAAAAATATTACCAACGAAACAATAACGAAATATTAACGACAA CTGAAAGCGAAATCACTAAGAACAATAACGAAATATTAACCAACGA CTGAAAGGGACATTCACCTGAAAAATGCAAATAACAAATC CCAGAACAATTATTACCTTAATGCAACAGTACCTAATTAC CCAGAACAATTATTACCTAATGCAACAGTACCGACA
3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	73 776 79 82 85 1 4 7 10 13 16 19 22 52 8 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79 82	CGAAAGGAGCGGGCCTGTTGGGAAGGGCACTACCATGTCATT GCGCTGCCGCTAACCGGAAACCGCCGAAAACAC GAACGCACTTGGTCTACTGAGCGCCGCTACAGGGCGGATCAAGATCGCACTCCAGTAAACGTTAATAT GTGACATACCTTCGGAGCATGTATAACGTGCTTTCCTGGTCATCTGCCAGTTGTAAATCAGCTCATT CGCTGACGAGGTTACAATGCGAGGCCGATTAAAGGGATGGGCACGTTAATTCGCGCTCTGG TGAGAAGTGTTTTCGTCGGATTCTCCGTAAATGTGAGCGGAG GCCTGTAGCATTCCCAACATATAAAAGGCCGTAAATGGAGCGAG ACCCCCCACATTTTTGTCACAATCAGGAATAACCGAAG ACCACCCCCCTCATTTTGTCACGAAGGACCAAGGTACCCAG ACCCCCGGAATAGGCACCAAGGACCAAGGTACCCAG ACCCCCGGAATAGGCACCAGTGAAGCTCACCGAGGC CGGATAGGGCGTAATATGGACGCACTGAAGGCACCAATAATAG AGGCGGATAAGTGCCACCAGTAGCACCATGAGGCCCCATAATATC CTCAAGAGAGGATCAATGAAACCACGGGGAGAATTAACTG CCTATTATTCTGAAAAGATTGCCCTTTTACCAGAGGAAT AGTGCCCGTATAAACGGCATTTTCGGTCGAAACGATTATTG GTACTGGTAATAGGTTCCCTCTTACACGAGGAATTAACTG CTCCAGTAAGCGTCGCCCCCCTCAAGTGCCCACCACGCTACG CTCATTAAGCGTCGCCCCCCCTCAAGTTGCCATTTTG CAGGTCAGACGATCGCCGCCACCCTCATGATGTCCTTTTCCGGT CTAAAATATCTTTAAAAGTTTGCGTTTAAAAGGTTCCCGCCACCTGC CGTCATAAGCGAACCACCGGAACGAATTAACCG CTCATTAAGCTCAGAGCCACCCCTCATAGTTGCCTTTTCG CAGGTCAGACGATCCCCGCAGCATTGACCTCCCGGACTGC CTCAATAGATAATACAACTGGTATAAAGGAATCATTACCG CAGTTGGCAAATCACCGGAAATTAACAGGAATCATTACCG CAGTGGCAAATCACCAGAGGCACCCACTCATCCGGA CTTGCTGAAGCCTCAGATGGCAATCATAGGATCCTAACTGG CCCCATTAGAATAAGGTTTAACGGACGAATCATACCG CAGGGGGAGGCGCTTGCCACCGAATAAAGGACCGACACA CGCCATTAGAATAAGGTTTAACGTAAACGGACTAATTACGA CAGAGGGGAGGCGCTTCCGCGAAATAAAGGACCGACA CTGGAAGCCAATCACCAGAAGAATAAAGGATCCGACA CTGGAAGCCAATCACCGAGAACAATAAAGGACCGACCAACAACAACAA GGCTATTAGCTTTGCGAGAAACAATATAAAGTACCGACA CTAGAAGCGTAAGCCAAGTACAAAATCGTAATTGA CAAAGGGAAATCCTAAGGAACAGTACCGAATCAAATTGA CCAGAACAATATTATGGAAACAATATAAGGATCGGAACCAAATTGA CCAGAACAATATTATGGAAACAATATAAAGGCTAAATTGA CCAGAACAATATTATGGAAACAGTACCGGCAATCAAAATTA CCAGAACAATATTATGGAAACAGTACCGGAATCATAATTA CCAGAACAATATTATGCTTCGTAAAACGTCACAATATAACGCCGCGTGGAT CCTGAGAAGAAACCTAAAGGAACAGTACCGGAATCATAATTA

5	3	ACGGAATAAGTTTAGAGTTTCGTCACCATTAGTAA
5	6	TGGTTTACCAGCGCCAGGGATAGCAAGCTCAACAG
5	9	TTGAGGGAGGGAAGCACCCTCAGAACCGGGAATTG
5	12	AGGTGAATTATCACCACCGTACTCAGGAAAAAAGGCTCCAAA
5	15	AGAGCCAGCAAAATCGTCGAGAGGGTTGGCTTGCTTTCGAGG
5	18	GCCGGAAACGTCACTAGGATTAGCGGGGAGTTGCGCCGACAA
5	21	ATCAGTAGCGACAGACATGAAAGTATTAGATATATTCGGTCG
5	24	TAGCGCGTTTTCATCAGTTAATGCCCCCCTTTTGCGGGATCGT
5	27	AGCGTTTGCCATCTTTTTAACGGGGTCAAACGAGGGTAGCAA
5	30	GCCACCACCGGAACCATACATGGCTTTTCTTTTCATGAGGA
5	33	ACCGCCACCTCAGAATGGAAAGCGCAGATGCCACTACGAAG
5	36	ACCACCACAGAGCGGCCTTGATATTCAGAATACACTAAAAC
5	39	TTACAAACAATTCGACATTTGAGGATTTCAAGCGC
5	42	CGTTATTAATTTTAGGAGCACTAACAACAGACGGG
5	45	GAACAAAGAAACCAACAGTTGAAAGGAAAGAGTTG
5	48	CCTGATTATCAGATAATATCAAACCCTCGAAAATCCTGTTTG
5	51	TGTTTGGATTATACGCAGCAAATGAAAATTATAAATCAAAAG
5	54	CCATATCAAAATTATCAGTATTAACACCTCCAGTTTGGAACA
5	57	TGCGTAGATTTTCACCGAACGAACCACCCAACGTCAAAGGGC
5	60	CAGTACCTTTTACAAATGCGCGAACTGAACTACGTGAACCAT
5	63	ATTGCTTTGAATACATACGTGGCACAGAGTGCCGTAAAGCAC
5	66	TATTCATTTCAATTTGGCCAACAGAGATATTTAGAGCTTGAC
5	69	ACATCAAGAAAACAGCAGATTCACCAGTGGAAGGGAAGAAAG
5	72	GAATTACCTTTTTTACATTTTGACGCTCAAGTGTAGCGGTCA
5	75	GTGAGTGAATAACCCCGCCAGCCATTGCGCTTAAT
5	78	AATTTTCCCTTAGACTCAAACTATCGGCCGAGCAC
5	81	TTAAGACGCTGAGACTTCTTTGATTAGTAGCTAAA
5	0.4	
J	04	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC
6	3	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC
5 6 6	3 6	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA
5 6 6 6	3 6 9	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA
5 6 6 6 6	84 3 6 9 12	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA
5 6 6 6 6 6	84 3 6 9 12 15	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT
5 6 6 6 6 6 6 6	3 6 9 12 15 18	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA
6 6 6 6 6 6 6 6	3 6 9 12 15 18 21	TAGGTCTGAGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACGGCTCAT AACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATATTCGAGCTTCATAATTTCAACTTTA AACACCAGGACAGGTCAGGATTAGAGAACACCCAGAA
6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24	TAGGTCTGAGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAGAGATTAAGAGGAAAGAACGGCTCAT AACCCTGAACAATGTACGAGGCTTCATAATTTCAACTTTA AACACCTGAACAATGTCGGCTTCATAATTTCAACTTTA AACACCTGAACAATGTCGGGATTAGAGAACACCAGAA TTTAACGTCAAAAAAGGTCAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAAGAGGTCATTTTCCGTAACAAAGCTGC
6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27	TAGGTCTGAGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAATGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAAATTTCAACTTTA AACACCCTGAACAATATTCGAGCTTCATAATTTCAACTTTA AACACACTGAACAATATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGGCAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTCGTAACAAAGCTGC GCCATATTATTATATAATGCTGTAGCGAGTAACCTTGACA
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30	TAGGTCTGAGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGGCCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGTCATTTTCGTAACAAAGCTGC GCCATATTATTATTATAATGCTGTAGCGAGTAACTTGACA AGCGTCTTTCCCAGATACGGTGTCTCGGAACGAGCCA
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33	TAGGTCTGAGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGGCAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTCGTAACAAAGCTGC GCCATATTATTATTATAATGCTGTAGCGAGTAACGTGTACAGACCA AGCGTCTTTCCAGATACGGTGTCTGGAACGATGACGAGCCA CACCCAGCTCACAATAATTCTGCGAACGATAAGGGAACCAACA
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 33 36	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGGCCATAAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGACCAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGTCATTTTCGTAACAAAGCTGC GCCATATTATTATTATAATGCTGTAGCGAGTAACGGTGACAGACCA AGCGTCTTTCCAGATACGGTGTCTGGAACGATAAGGGAACCAGAC CACCCAGCTACAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAGTTTTGAAGTTCGCAACGATAAGGCAACCGAAC GGGAGTTTTGAAGTTCGCAACGATAAGGTCACTTTA
5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGGCTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTATA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGACAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTATTATAATGCTGTAGCGAGTAACGGTACAGACCA AGCGTCTTTCCAGAATAGTCTGGCAACGGTGTACAGACCA CACCCAGCTACAATAATTCGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTTCGCAACGATAAGGTCACCCAC GGGAGGTTTTGAAGTCGCGAACGATAAGGCAACCGAAC GGGAGGTTTTGAAGTCGCAAATGGTCACTCCATGTTACTTA ATTCTAAGAACGCGGGGCCGAACTGGATCATCATCGCCT
966666666666666666	3 6 9 12 15 18 21 24 27 30 33 36 39 42	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTCGAGGATAAGAGACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACACAGGACAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTATATATGATGCTGTAGGATAAGGAACCAGACA AGCGTCTTTCCAGAATAGTGTGAGCGAGTAACGGTGACCAGACCA CACCCAGCTACAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTTCGCAAATGGTCACTCCATGTTACTTA ATTCTAAGAACGCGGGGCGCGAGCTGAATTGTATCATCGCCT CGCCCAATAGCAAGATAACACCCGGGGAGCGGAGAGGCGG
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATATTCGAGCTTCATAATTTCAACTATA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGACAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTCGTAACAAAGCTGC GCCATATTATTATTATAATGCTGTAGCGAGTAACACCAGAC AGGGTCTTTCCAGAATAGTGTGTCGGAACGATAAGGTACCAGACCA CACCCAGCTACAATAATTCGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTTCGCAAATGGTCACCCAGTACAGACCA GGGAGGTTTTGAAGTTCGCAAATGGTCACCCAGACCA GGGAGGTTTTGAAGTTCGCAAATGGTCACCCAGTACAGCCGAC GGGAGGCGCGGGGCGCGAGCTGAATGTATCATCGCCT CCCCCAATAGCAATAACATCCGCGGGGAGAGCGG ACACCCAGTTAAGCATTAACATCCGCGGGAACCGGCG ACACCCAAGCAGTTAACATCCGCGGGAAACCTGTCGTG
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGAAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGACAGGTCATGTATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTATTATAATGCTGTAGCGAGTGTACAGACCA AGCGTCTTTCCAGAATAGTGTCGGAACGGTGTACAGACCA CACCCAGCTACAATAATTCGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTTCGCAAATGGTCACCCAGTGTACTTA ATTCTAAGAACGCGGGGCGCGAGCTGAATTGTATCATCACCGCG GCCCAATAGCAAGTAGCATTAACATCGCGGGGAGAGGCG ACACCCAAGCCGTTTAGCAAAATTAAGCGGAAACCTGTCGTG TTCCAAGAACCGGTTGACACCAAAATTAAGCGAAACCTGTCGTG
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTCATAATTTCAACTTTA AACACTAAAAACAGGACAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTTATTATAATGCTGTAGCGAGTAACTCTGGAC AGCGTCTTTCCAGATACGGTGTCTGGAACGGTGTACAGACCA AGCCGACCAACAATAATCTGCGCAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTTCGCAAATGGTCACTCCATGTTACTTA ATTCTAAGAACGCGGGGCGCGAGCTGAATTGTATCATCGCCT CGCCCCAATAGCAAGTAGCATTAACATCCGCGGGGAGAGGCGG ACAAGCAAGCCGTTTAGCAAAATTAAGCGGAAACCTGTCGTG TTCCAAGAACGGGTCGGATGACCAAAACTCACACATTAATTGC CCCATGTAGAAACCAGAGACCCTTTATTAGCAAACTGATAA
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACCATGTCACGAAACGACGACGACACGACTCA AAGAGAAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTCATAATTTCAACTTTA AACACTAAAAAAGAGGTCATTATAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTTATTATAATGCTGTAGCGAGTAACTCTGGAC AGCGTCTTTCCAGATACGGTGTCTGGAACGGTGACAGACCA AGCCGAGCTACAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAAGTTTGAAGTTCGCAAATGGTCACTCCATGTTACTTA ATTCTAAGAACGCGGGGGCGCGAGCTGAATTGTATCATCGCCT CGCCCAATAGCAAGTAGCATTAACATCCGCGGGGAGAGGCGG ACCCCAGCTTAGCAAAATTAACGCGGAAACCTGTCGTG TTCCAAGAAACCGAGTGGATGTACCAAAACTCACACATTAATGC GCATGTAGAAACCAAGAAGAAGCCTTTATTAGCATAAAGTGTAA AACCCCAAACCAAGAAAGCAAGAAGCCTTTATTAACTTCCGCT
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACCATGTACCTGACATATAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTTATTATAATGCTGTAGCGAGTAACTTGACA AGCGTCTTTCCAGATACGGTGTCTGGAACGGTGACAGACCA AGCGGTCTTTGCAAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTACGCAAATGGTCACTCCATGTTACTTA ATTCTAAGAACGCGGGGCGCGAGCTGAATTGTATCATCGCCT CGCCCAATAGCAAGTAGCATTAACATCCGCGGGGAGAGGCGG ACACCAAGCAGGTCGGTTGTACCAAAACCTACCATTAATGC GCATGTAGAAACCAGAGAAGCCTTTATTAACATCCGCTG TTCCAAGAAACCAGAGAAGACCTATATTTAAAATTGTATCCGCT TTCCAAGAAACCAGAGAAGCCTCTATATTTAAATTGTATCCGCT TCCAAGAAACCAGAGAAGCCTCATATATTTAAAATGGTAAAAGTGTAA AAGTCCTGAACAAGGAATCCAAAACGAGACGCTCATATATTTAAATGTATCCGCT
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACCATGTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGACAGGTCAGGTCAGAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTTATTATAATGCTGTAGCGAGTAACTTGACA AGCGTCTTTCCAGATACGGTGTCTGGAACGGTGACAGACCA AGCGGTCTTTCCAGATACGGTGTCTGGAACGGTGACAGACCA AGCGGTCTTTCCAGATAAGTGGCACCCACCA CACCCAGCTACAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGATGCACAACACTCACATAATGTACATCACCA AGCCCCAATAGCAAGAGCATTAACATCCGCGGGGAGAGGCGG ACAAGCAAGCCGTTTAGCAAAATTAACCGCGAAACCTGTCGTG TTCCAAGAAACCAGAGAAGCCTTTATTAACATCCGCT CGCATGTAACAAGAGGTACAAAATTAAGCGAAACCTGAAAGTGTAA AAGCCCTGAAACCAGAGAAGCCTTTATTAAAATGTATACCGCT TGTTCAGCTAAAGCATAACAAAAGGAGCTCGAATTCGTA AAGTCCTGAAACCAGAGATCAAAAGGAACCAGAAGCTCGCAATCGTA AAGTCCTGAAACCAGAGATCAAAAGGAGCTCGAATCGCAA
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACCCTGACATATAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTA AACACCCTGAACAAGGTCAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTTATTATAATGCTGTAGCGAGTAACTTTGACA AGCGTCTTTCCAGATACGGTGTCTGGAACGGTGACAGACCA CACCCAGCTACAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTGCGCAAATGGTCACTCCATGTTACTTA ATTCTAAGAACGCGGGGGCGCGAGCTGAATTGTATCATCGCCT CGCCCAATAGCAAGTAGCATTAACATCCGCGGGGAGAGGCGG ACAAGCAAGCCGTTTAGCAAAATTAACGCGGAAACCTGTCGTG TTCCAAGAACCAGAGAAGCCTCTTATTAAACTCACATTAATTGC GCAATGTAACAAGCCTCATATATTTTAAAATGGTATCCGCT TGTTCAGCAAAACCAAGAAGCCTCATATATTTAACATCGCGTA AAAGTCCTGAAAAGCAAGAAGCTCAAAAGGAGCTCGAATTCGTA AAAGTCCTGAAAAGAGAATCAAAAGGAGCTCGAATTGGATCGCT TGTTCAGCAAAACCAAGAATCAAAAGGAGCTCGAATTCGCAGCCTGCAA AAGGTAAAACCAAGAATCAAATGATATTGCATACGCCTGCAG GCCATTTTCCGAGCCGGAGAGGCTGTTTCCCCAGCTGCAG
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 45 51 57 60 63 66	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCAGAGGCGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTCAGAAAACGAGGGTAGAAAGATCA AGCAAGAAACAATGTCCGGACATATTAATCTACGTTAATAA AGAGAGCCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGGGCTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTTATATATATGCTGTAGCGAGTAATCTTGACA AGCGTCTTTCCAGATACGTGTCTGGAACGGTGTACAGACCA CACCCAGCTACAATAATTCGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTTCGCAAATGGTCACTCCATGTACAGACCA GGGAGGTTTTGAAGTCGCAAATGGTCACTCCCATGTACTTA ATTCTAAGAACGCGGGCGCGAGCTGAATTGTATCATCGCT CGCCCAATAGCCAGTAGCATTAACACCCGGGGAAACCTGTCGTG TTCCAAGAACCAGAGAAGCCTTTATTAGCATAAAGTGTAA AAGCCCTGAACAGCCTCATATATTTTAAATTGTTATCCGCT TGTCCAGCAAAACCAGAGAGCCTTTATTAGCATAAAGTGATA AAGGCCTGAACAGCCTCATATATTTTAATTTTAACTGTATCCGCT TGTCCAGCAAAACCAAGACTCCAAAAGGAGCCCGAATCGTA AAGGCCTGAACAGCCCCATATATTTTAATTTTAATTGCAGCCTGCAG GGCATTTCGGACAGGAGCCCCATATATTTTAATTTTAACTGCAGCCAGC
9 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 45 51 54 57 60 63 66 69	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGACACCCAAAAGATTAGAGGAAAGACTGGCTCAT AACACCCTGAACAATAATTCGAGCTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTCATTATTAGAGAAACACCAGAA TTTAACGTCAAAAAAAAGAGGTCATTTTTCGTAACAAAGCTGC GCCATATTATTTATATATGCTGTAGCGAGTAACTTGACA AGCGGTCTTTCCAGAATAGGTGCTCGGAACGGTAACCTGACA AGCGACTTTACAATAATTCTGCGAACGGTGTACCAGCCA CACCCAGCTACAATAATTCTGCGAACGGTGACCCAGCCA GGGAGGTTTTGAAGTTGCCAAATGGTCACTCCATGTACTTA ATTCTAAGAACGGGGGCGCGAGCTGAATTGTATCACGCT CGCCCAATAGCAAGATACACACGCGGGAAACCTGTCGTG CTCCAAGAACCAGAGAAGCCTTTATTAGCATAAAGTGTAA AAGGCCAAACCAGAGAAGCCTTTATTAGCATCAAAGTGTAA AAGGCCTGAACAGGAGCCCGAGCTGAATTGGTA AAAGCCCGAACAGCCCCATATATTTTAGATATTGCAGCCTGCAG GCATGTAGAAACAAGAAGCCCCATATATTTTAGATATTTGCAGCCTGCAG GGCATTTCCAGACAAGCCCCAAAACTCCAGAGAGGCTCGAATTCGTA AAAGCCAGAACCAGAGAGCCCCATATATTTGCAGCCCGCA
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGGCGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTCCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGACAGG
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 45 51 54 57 60 63 66 69 72 75	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAGTCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAGCACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACATAAAAACAGGACAGG
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78	TAGGTCTIGAGAGACAAAAGAGTCTGTCCCTTACGCCAGAATCC AACTGGCATGATTATAGTAAATGTTTAAGTAAAGAGCAACAC AAGGAAACCGAGGACCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGACTGGCTCAT AACACCCGAACAATAATTCGAGCTTCATAATTTCAACTATA AACATAAAAACGGACAGGTCAGGATTAGAGAAACACCAGAA TTTAACGTCAAAAAAGAGGTCATTTTCGTAACAAAGCTGC GCCATATTATTATTATATAGCTGTAGCGAGTAAGGAACACCAGAA AGCGTCTTTCCAGATACGGTGTCTGGAACGGTGTACAGACCA CACCCAGCTACAATAATTCTGCGAACGATTAGGGAACCGACCA CACCCAGCTACAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAGGTTTTGAAGTTCGCAAATGGTCACTCCATGTTACTTA ATTCTAAGAACGCGGGGCGCGAGCTGAATTGTACCATCGCT CGCCCAATAGCAAGAGCACTTAACATCGCGGGGAGAGGCGG ACAAGCAAGCCGTTTAGCAAAATTAAGCGGAAACTCACGCTG CGCCAATAGCAAGGAGCGTTATTATACATCGCGGGGAGAGGCGG ACAAGCAAGCGGTCGGTTGTACCAAAACTCACATTAATTGC GCATGTAGAAACCAGAGAAGCCTTTATTAGCATAAAGTGTAA AAGGTCCGAACAACCCAGAGAGCCCTATTTTTGCATGCCTGCGT TGTTCAGCTAAAGCAGAAGCCCTATATTTTAAATTGTTATCGCT AAAGGTAAAGCAAGCCGCGAAGCCTGAATTGTACATCGCA GGCATTTCGAGCAAAATCACAATATAAGAGCAGCCCGCCACGAG GCAATTTCGAGCAGAATCCAATATGATATTTGCATGCCCGCGA AAAGGTAAAGTAATCACCAATAGCATAATTTTGCATGCCTGCC
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 45 45 54 57 60 63 66 69 72 75 78 81	TAGGTCTIGAGAGACAAAAGAGTCTGTCCCTTACGCCAGAATCC AACTGGCATGATTATAGTAAATGTTTAAGTAAGGCAACAC AAGGAAACCGAGGCGTCATAAATGTTTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGGAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAGAACTGGCTCAT AACACCGTGAACAATAATTCGAGGTCAGGATTAGAGAAACACCGGAA TTTAACGTCAAAAAAGAGGTCAGTTAGAGAAACACCAGAA TTTAACGTCAAAAAAAGAGGTCATTTTCGTAACAAAGCTGC GCCATATTATTATTATATAGCTGTAGCGAGTAACTCTGACA AGCGTCTTCCAGATACGGTGTCGGAACGGTGACAGACCGAAC CACCCAGCTACAATAATTCTGCGAACGATAAGGTAACCTGGCA CACCCAGCTACAATAATTCTGCGAACGATAAGGGAACCGAAC GGGAGGTTTGAAGTTCGCAAAGGTCACTCCATGTTAC ATTCTAAGAACGCGGGGCGCGGAGCTGAATTGTATCATCGCCT CGCCCAATAGCAAGGAGCAGTAAACTCCGCGGGGAGAGCCGG ACAAGCAAGCCGTTTAGCAAAATTAACCGCGGAGACCCGTCGCG TTCCAAGAACGGGTCGGTGTACCAAAACTCACATTAATTGC GCCTGAACAAGCCAGGAGCCTGTATTATAGCAGAAACCTGCTG TGTCAGCTAATGCAAGAGCCTTTATTAGCATAAAGTGTAA AAGTCTGAACAAGCCCCCATATATTTTAAATTGTATCCGCT TGTCCAGCTAATGCAAGATCAAAAGTAACTCACATTAATTGC GGCATGTAGAAACCAGAGAGCCTTTATTAGCATAAAGTGTAA AAAGGTAAAGTAATCAACAATAGAAGTAGCTCGCGCGA GGCATTTCGAGCAAGGATCCAAAACTTACGCTGCAG GGCATTTTCGAGCAAGAGTCCAAAACTTACCGCCAGCA AAAGGTAAAGTAATCAAAATGAAGTAGTCTTTCCCAGCT GGCATATTGGAGCATGAAATTCAAAAGGGAGCCGAATCCGTA AAAGGTAAAGTAATCAAAATGAAATTTACCAGGCGCAATCGTA AAAGGTAAAGCAGTAATCAAAATAGAACGGCCATTCGCA GAATCGCCATATTGGTAATCGTAAAGATCGGTGCGGGCC CTAGAAAAAGCCTGTTGATAATTTAAATTTCCCAGCTTCGGCA AAATAAGGCGTTAAACAATATTAAATTTTGCAGGCCCACTTCCGGC TTCCGACCTAAATTAAAT
9	3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78 81 84	TAGGTCTGAGAGACAAAAGAGTCTGTCCGTACGCCAGAATCC AACTGGCATGATTATAGTAAAATGTTTAAGTAAGAGCAACAC AAGGAAACCGAGGACGTCATAAATATTCAACTAATGCAGATA CTTTTTAAGAAAAGGTTCAGAAAACGAGGGTAGAAAGATTCA AGCAAGAAACAATGTACCCTGACTATTAATCTACGTTAATAA AGAGAGATAACCCAAAAGATTAAGAGGAAAGAACTGGCTCAT AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACCCTGAACAATAATTCGAGCTTCATAATTTCAACTTTA AACACTAAAAACAGGACAGG

D.4 Position of Biotin-Labeled Staple Strands

To form functionalized DNA nanotubes with 29 attachment sites for streptavidin conjugated nanoparticles, all 29 staple strands of helix 4 (H4-C1 to H4-C85) were modified by adding a 5 thymine tether to the 3' end followed by a biotin molecule. During synthesis of the nanotubes, these strands were substituted for the corresponding unmodified staple strands. Note that the strands are labeled in helix 4 by the location of their 5' end, but the biotin modified 3' ends are located in helix 6. To synthesize nanotubes with 15, 9 and 5 available binding sites, the subsets of the helix 4 staple strands were substituted. The column numbers of the substituted staple strands are listed below.

For 15 binding sites, every other staple of helix 4 was substituted:

1, 7, 13, 19, 25, 31, 37, 43, 49, 55, 61, 67, 73, 79, 85.

For 9 binding sites, every third staple of helix 4 was substituted, starting with column 7:

7, 16, 25, 34, 43, 52, 61, 70, 79.

For 5 binding sites, every fifth staple of helix 4 was substituted, starting with column 5:

13, 28, 43, 58, 73.



Fig.D.2: Biotin-labeled DNA origami nanotube arrays (a) 29 particles, (b) 15 particles, (c) 9 particles, (d) 5 particles

APPENDIX E

Statistical Analysis of Quantum Dot Arrays

E.1 Statistical Analysis of Quantum Dot Arrays

Histograms of the number of streptavidin-conjugated quantum dots attached to biotin-labeled DNA origami nanotubes with (a) 5, (b) 9, (c) 15, and (d) 29 binding sites are shown in Fig.E.1. The histogram data were gathered from AFM image analyses for over 225 individual nanotubes for each case. In theory, it is predicted that the number of particle attachment in 5, 9, 15, and 29 available binding sites should peak at 5, 9, 15, and 29, respectively. Experimentally, the number of particle attachment in 5, 9, 15, and 29 available binding sites peak at 4, 7, 10, and 17, respectively. The histogram data suggest that attachment to each available binding site is much more likely for nanotube functionalized with 5 or 9 binding sites than for 15 and 29 binding sites.



Fig.E.1: Histograms (bars) and calculated binomial distributions (lines) for the number of attached quantum dots for DNA nanotubes with (a) 5, (b) 9, (c) 15, and (d) 29 biotin binding sites. Data for each histogram were compiled from AFM image analysis for over 225 separate nanotubes, with the exact number, N, shown for each histogram. The average attachment probabilities, p, used to generate the calculated

binomial distributions are indicated for each case. (Reprinted with permission from Ref.22. Copyright 2010 American Chemical Society).

Assuming that quantum dot binding events occur with an equal average attachment probability, p, for each site, the attachment histograms would be expected to follow a binomial distribution, P(m), given by

$$P(m) = \frac{n!}{m!(n-m)!} \cdot p^{m} \cdot (1-p)^{n-m}$$
(E.1)

where *n* is the given number of available biotin binding sites per nanotube and *m* is the number of attached quantum dots per nanotube³⁵. The average attachment probability, *p*, is given by

$$p = \frac{\sum attached \ QD}{\sum available \ sites}$$
(E.2)

where the numerator is the total number of attached quantum dots, and the denominator is the total number of available attachment sites³⁵.

Applying Equation (E.2), the average attachment probabilities were calculated from the histogram data to be 0.77, 0.76, 0.65, and 0.64 for 5, 9, 15, and 29 sites, respectively. The solid lines in Fig.E.1 plot the calculated binomial distribution of Equation (E.1) for each case. Overall, the calculated distributions agree with the histogram data, indicating equal average attachment probabilities. However, the histogram data in (c) and (d) display a slight shift toward lower attachment relative to the calculated distribution data, providing some evidence for steric hindrance or site bridging.



Fig.E.2: Histograms (bars) and calculated geometric distributions (lines) for nearest-neighbor (N-N) separation of bound quantum dot pairs for DNA nanotubes with (a) 5, (b) 9, (c) 15, and (d) 29 biotin binding sites. The numbers of separations, N, measured for each case are provided in the figures, along with the average attachment probabilities, p. N-N separation of zero indicates two nearest neighbors with a separation less than one-half of a period. (Reprinted with permission from Ref.22. Copyright 2010 American Chemical Society).

For evidence of steric hindrance or site bridging, the nearest-neighbor separation distances, projected along the nanotube axis, were measured for pairs of bound quantum dots. In theory, the nearest-neighbor separation histogram would be expected to follow a geometric distribution peaked at the designed nanotube binding site periodicity in the absence of steric hindrance or site bridging. The geometric distribution, P(l), of nearest-neighbor separations is given by³⁵

$$P(l) = p \cdot (1-p)^{l-1}$$
(E.3)

where l is the integer number of periods between nearest-neighbors and p is the average attachment probabilities obtained from Equation (E.2). Fig.E.2 shows histograms of the nearest-neighbor separations and the geometric distributions for each of the four cases. For each case, the measured nearest-neighbor separation distances were normalized to represent the number of designed binding sites periods between particles. The data were sorted into bins of width a centered on the *nth* period, where a is the designed nanotube periodicity and n is an integer. Hence, nearest-neighbor separations of less than a/2 were indicated as a zero separation. Experimentally, the nearest-neighbor separation histograms are peaked at 1, 1, 2, and 3 for 5, 9, 15, and 29 binding sites; "1" means that the nearest-neighbor separation agrees with the designed binding site periodicity, "2" and "3" mean that the nearest-neighbor separations are twice and triple the designed binding site periodicity, respectively. The analysis demonstrates that the calculated geometric distributions match the data for the nanotubes with 5 and 9 available sites, but deviate significantly for the nanotubes with 15 and 29 available sites. Hence, the data indicate that steric hindrance or site bridging is reducing the number of quantum dots attached to the nanotubes.

APPENDIX F

Experimental Equipment
F.1 Experimental Equipment



Fig.F.1: Eppendorf Centrifuge 5418, used for filtering DNA solutions

	AC	
	eppendorf Ma	stercycler personal
-	Detailined 0	ALESCO
1		Exit 7 8 9
1	WHITE IS A REAL FOR THE PARTY OF THE PARTY O	,

Fig.F.2: Eppendorf Mastercycler Personal, used for annealing DNA solutions



Fig.F.3: Hoefer gel electrophoresis apparatus, used for purifying DNA solutions



Fig.F.4: Agilent Varian Spectrophotometry, used for measuring fluorescence



Fig.F.5: Veeco atomic force microscope multimode, used for characterizing topography of DNA nanostructures.