PNA-Polypeptide Assembly

in a 3D DNA Nanocage for Building

Artificial Catalytic Centers

by

Justin David Flory

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2014 by the
Graduate Supervisory Committee:

Petra Fromme, Chair
Hao Yan
Daniel Buttry
Giovanna Ghirlanda

ARIZONA STATE UNIVERSITY

May 2014

ABSTRACT

Proteins and peptides fold into dynamic structures that access a broad functional landscape, however, designing artificial polypeptide systems continues to be a great challenge. Conversely, deoxyribonucleic acid (DNA) engineering is now routinely used to build a wide variety of two dimensional and three dimensional (3D) nanostructures from simple hybridization based rules, and their functional diversity can be significantly expanded through site specific incorporation of the appropriate guest molecules. This dissertation describes a gentle methodology for using short (8 nucleotide) peptide nucleic acid (PNA) linkers to assemble polypeptides within a 3D DNA nanocage, as a proof of concept for constructing artificial catalytic centers. PNA-polypeptide conjugates were synthesized directly using microwave assisted solid phase synthesis or alternatively PNA linkers were conjugated to biologically expressed proteins using chemical crosslinking. The PNA-polypeptides hybridized to the preassembled DNA nanocage at room temperature or 11 °C and could be assembled in a stepwise fashion. Time resolved fluorescence anisotropy and gel electrophoresis were used to determine that a negatively charged azurin protein was repelled outside of the negatively charged DNA nanocage, while a positively charged cytochrome c protein was retained inside. Spectroelectrochemistry and an in-gel luminol oxidation assay demonstrated the cytochrome c protein remained active within the DNA nanocage and its redox potential decreased modestly by 10 mV due to the presence of the DNA nanocage. These results demonstrate the benign PNA assembly conditions are ideal for preserving polypeptide structure and function, and will facilitate the polypeptide-based assembly of artificial catalytic centers inside a stable DNA nanocage. A prospective application of assembling multiple cyclic γ-PNA-peptides to

mimic the oxygen-evolving complex (OEC) catalytic active site from photosystem II (PSII) is described. In this way, the robust catalytic capacity of PSII could be utilized, without suffering the light-induced damage that occurs by the photoreactions within PSII via triplet state formation, which limits the efficiency of natural photosynthesis. Therefore, this strategy has the potential to revolutionize the process of designing and building robust catalysts by leveraging nature's recipes, and also providing a flexible and controlled artificial environment that might even improve them further towards commercial viability.

DEDICATION

I dedicate this dissertation to my loving wife Jennifer Flory and son Preston Flory, whose love and support carried me through the demanding efforts required for research. I also dedicate this dissertation to my incredible parents, Beth Flory, Pierre Dansereau, John Flory and Debra Flory, for pushing me to do my best and strive for excellence, while at the same time supporting me and encouraging balance in life. In addition, this dissertation is dedicated to my amazing in-laws, Cathie Anderson and Bill Anderson, who assisted my family in countless ways and always pushed me to publish. I also thank Keren Salgado Mendez for her dedication to support us by watching Preston whenever we needed it. I also dedicate this to all of my amazing science teachers and professors I have had throughout the years; notably, Mr. Richardson from Davis Senior High School, and Dr. Joseph Tenn and Dr. Lynn Cominsky from Sonoma State University.

ACKNOWLEDGEMENTS

I am very grateful for the support from my advisor Dr. Petra Fromme and for her insightful feedback and guidance during the up and down research process and careful review of my work. I also appreciated the leeway she gave me to tryout my own ideas, and to make mistakes, which gave me a much deeper understanding of the subject matter of our project as well as the research processes itself.

I also very much appreciate the help and guidance from my faculty collaborators, Dr. Giovanna Ghirlanda, Dr. Hao Yan and Dr. Yan Liu. Their regular feedback and suggestions to overcome technical challenges presented in my research was critical. I also would like to acknowledge their intellectual input to conceive of and develop the proposed application for using DNA nanostructures to assemble artificial catalytic centers.

I am also very grateful for the feedback from other members of the Center for Bio-Inspired Solar Fuel Production. In particular, Dr. Raimund Fromme, Chad Simmons, Dong Wang and Dr. Kimberly Rendek actively participated in our regular meetings, which helped shape the way I think about scientific problems. Dr. Devens Gust and Alexander Melkozernov were also very supportive of my work, providing opportunities to attend conferences to present my work and obtain financial support to sponsor an undergraduate student working with me. Trey Johnson began working with me on my research project as an undergraduate student and continued well after graduating. He quickly became thoroughly invested in the project and blissfully complied with my sometimes unreasonable requests in order to keep the project moving forward. I am truly grateful for his contributions, as well as for giving me the rewarding experience of mentoring and guiding his research activities.

A large part of my research involved biophysical characterization using fluorescence based techniques. I am forever grateful to Dr. Su Lin for her phenomenal ability to teach fluorescence based methods, discuss and interpret complex results, as well as for her patience with the time it took me to understand and apply the main concepts.

I am also grateful to all members of the Fromme lab for their support, including preparing for my oral examination, helping me out with other various things and keeping the mood enjoyable. I want to specifically thank Dr. Ingo Grotjohann for his help conceiving new experiments, discussing and interpreting results, and for his humor. The countless hours of experiments in the lab would have been significantly more difficult and less enjoyable without Jay-How Yang. I very much appreciate his dedication to keeping the lab in good order, making it easier for me to focus on the difficult challenges presented by the science, and for his humor and conversation to keep the mood enjoyable.

I also would like to thank many people for their help teaching me new techniques. Dr. Jeannette Nangreave and Chad Simmons were instrumental in teaching me about DNA nanotechnology. Dr. Sandip Shinde helped teach me about peptide synthesis, functionalization and purification. Dr. Jose M. Martin-Garcia, Jay-How Yang, Zhen Gong, Allan Lee, Dr. James Zook, Domingo Meza-Aguilar, Dr. Katerina Dörner helped to teach me about protein techniques.

I also want to thank many experts in their fields whom let me consult with them. Specifically I want to thank Zach Laughrey and John Lopez in the ASU Protein Chemistry Lab for sharing their expertise on MALDI, peptide synthesis and purification, circular dichroism and their general consultation and interpretation of results. I want to thank

Anindya Roy for his consultation on organic synthesis and Dayn Sommers for consultation on spectroscopic redox titrations and other scientific discussions.

I am grateful to Dr. James Allan for taking time to listen to my concerns, and provide sound advice for positioning myself to succeed after graduation.

I am also very grateful for the financial support from the Center for Bio-Inspired Solar Fuel Production, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001016. This support has been critical to develop of the research methods contained in this dissertation and to my continued growth and development as a scientist.

TABLE OF CONTENTS

LIST OF TABLES xiii

LIST OF FIGURES xiv

LIST OF ABBREVIATIONS xix

PREFACE xxi

CHAPTER

1 INTRODUCTION 1

Bionanotechnology: Understanding and reengineering nature's molecular machines 1

DNA nanostructures for engineering polypeptides 4

PNA for linking polypeptides to DNA 14

Specific Aims 17

Outline and organization 18

2 MATERIALS AND METHODS 21

PNA-peptide synthesis, purification and functionalization 22

PNA coupling 22

Amino acid coupling 23

Fluorescent dye coupling 24

PNA-peptide acetylation 25

Cleavage 26

PNA-peptide purification and mass identification by MALDI 26

PNA maleimide functionalization 28

γ-PNA functionalization and purification 30

CHAPTER Page

γ-PNA Cy5 labeling 30

γ-PNA Cy5 purification 31

Cytochrome c preparation 32

Fluorescent labeling of cytochrome c with TMR 32

Thiolation of cytochrome c using SPDP 33

Azurin expression and purification 34

PNA-protein conjugation, purification and characterization 36

PNA-cytochrome c conjugation 36

PNA-azurin conjugation, purification and characterization 40

TMR labeling of PNA-azurin 42

ssDNA toehold mediated PNA and PNA-protein purification 42

Toehold DNA design 42

PNA-FAM purification 43

PNA-CytC purification 45

DNA nanocage preparation 47

ssDNA purification 47

DNA nanocage assembly 48

PNA-polypeptide assembly with the DNA nanocage 52

PNA-peptide incubation with the DNA nanocage 52

PNA-protein incubation with the DNA nanocage 54

PNA-peptide hybridization efficiency 56

PNA-polypeptide hybridization kinetics 58

CHAPTER Page

DNA-PNA-polypeptide thermal stability and fluorescence characterization 61

Sample preparation 61

Steady state fluorescence spectra 62

Dissociation temperature (T_D) curves 63

Donor quantum yield 66

Calculating the absorption spectra and extinction coefficients for the overlap integral of TMR labeled PNA-proteins 66

Fluorescence lifetime 69

Fluorescence anisotropy 74

Förster radius, energy transfer efficiency and distance calculations 75

Förster distance uncertainty calculations 76

Molecular structures, models and chemical structures 77

Molecular structures of DNA nanocage, PNA, cytochrome c and azurin ... 77

PNA-protein chemical linker structures 78

Cytochrome c activity and redox potential 78

Luminol oxidation activity 78

Redox potential titration with ferri/ferrocyanide 79

Circular dichroism of cytochrome c 83

3 PNA-PEPTIDE ASSEMBLY IN A 3D DNA NANOCAGE AT ROOM TEMPERATURE 85

Abstract 85

Introduction 86

CHAPTER Page

Results and Discussion 89

DNA-PNA-peptide complex design 89

PNA-peptide synthesis 93

PNA-peptide assembly with the DNA nanocage 95

DNA-PNA-peptide thermal stability 100

Fluorescence characterization: lifetime, anisotropy and energy transfer 103

Conclusion 111

4 LOW TEMPERATURE ASSEMBLY OF FUNCTIONAL 3D DNA-PNA-PROTEIN COMPLEXES 113

Abstract 113

Introduction 114

Results and Discussion 117

DNA-PNA-protein complex design 117

PNA-protein conjugation, purification and fluorescent labeling 120

PNA-protein purification by toehold-mediated DNA strand displacement 124

PNA-protein assembly with the DNA nanocages 129

Fluorescence characterization: anisotropy, lifetime and energy transfer 132

Cytochrome c activity 138

Cytochrome c redox potential 139

Cytochrome c secondary structure 142

Conclusion 143

CHAPTER Page

5 PURIFICATION AND ASSEMBLY OF CY5 LABELED γ-PNAS INTO A 3D DNA NANOCAGE 145

Abstract 145

Introduction 145

Results and Discussion 147

DNA-PNA complex design 147

Fluorescently labeled and unlabeled DNA nanocage assembly 150

Fluorescent labeling and purification of γ-PNA 152

Assembly of the DNA nanocages with fluorescently labeled γ-PNAs 157

Conclusion 158

6 PROSPECTIVE: BUILDING A PEPTIDE-BASED ARTIFICIAL OXYGEN-EVOLVING COMPLEX WITHIN A 3D DNA NANOCAGE USING γ-PNA LINKERS 160

Abstract 160

Discussion 161

Natural Photosynthesis 161

Photosystem II: Nature's solar-driven water oxidation catalyst 164

Artificial Photosynthesis 166

Potential design of an artificial oxygen-evolving complex 167

Potential strategy to form cyclic γ-PNA-peptide conjugates 172

Potential strategy to assemble and operate the artificial OEC 175

Conclusion 177

CHAPTER Page

7 Summary and Outlook 179

Summary 179

Outlook 180

REFERENCES 183

APPENDIX 196

A PERMISSION TO REPRODUCE SELECTED FIGURES 196

B CO-AUTHOR APPROVAL 208

BIOGRAPHICAL SKETCH 210

LIST OF TABLES

Table Page

2-1. Fitting parameters and time to reach 99% hybridization for PNA-protein binding kinetics with DNA nanocage. 61

2-2. Lifetime components and calculated average lifetimes of the donor and acceptor fluorophores. 71

3-1. Fitting parameters and time to reach 99% hybridization for PNA1-GPG-TMR binding kinetics with DNA nanocage. 100

3-2. Lifetime components and calculated average lifetimes of the donor and acceptor. 104

5-1. DNA, PNA and peptide strand names and sequences. 149

5-2. Complex nomenclature with list of constituent DNA and γ-PNA(-peptide) strand names. 150

5-3. Progress of γ-PNA-Cy5 elution from gel blocks. 156

5-4. Fitting parameters for of γ-PNA1-Cy5 and γ-PNA3-Cy5 elution. 157

LIST OF FIGURES

Figure Page

1-1. Hierarchy of organization in a plant. 3

1-2. DNA Holliday junction and examples of 2D and 3D DNA nanostructures made of synthetic oligonucleotides. 5

1-3. 2D and 3D DNA nanostructures made using DNA origami. 7

1-4. Arranging polypeptides with DNA nanostructures 9

1-5. DNA tetrahedron assembly, protein encapsulation and cryo Electron Microscopy density. 12

1-6. Chemical structures of PNA, DNA and peptide backbones, and a microwave assisted solid phase peptide (and PNA) synthesizer. 15

1-7. Applications with PNA. 16

2-1. UV absorbance of Fmoc deprotection during PNA synthesis. 23

2-2. Chemical structures of FAM and TMR. 25

2-3. MALDI mass spectrum of the synthesized PNA-peptides. 27

2-4. RP-HPLC chromatograms of the synthesized PNA-peptides. 28

2-5. MALDI-MS spectrum and RP-HPLC of PNA-Gly-SMCC. 29

2-6. Quantification of SPDP thiolation of cytochrome c by MALDI-MS and UV-Vis. 34

2-7. MALDI-MS spectrum of SPDP thiolated cytochrome c conjugated to PNA-SMCC. 37

2-8. Native PAGE (7%) at 4 °C of the DNA nanocage incubated with PNA-CytC. . 38

2-9. MALDI-MS spectrum and SDS-PAGE of azurin conjugated to PNA-SMCC. .. 41

Figure Page

2-10. Native PAGE (7%) at 4 ^{0}C of toehold mediated ssDNA purification of PNA. .. 44

2-11. IEX-FPLC of toehold mediated ssDNA purification of PNA-TMR. 45

2-12. IEX-FPLC of toehold mediated ssDNA purification of PNA-CytC. 46

2-13. SE-FPLC chromatograms of the DNA nanocage with and without PNA-peptides. 49

2-14. DNA nanocage assembly and PNA incubation. 51

2-15. Native PAGE of the DNA nanocage incubated with PNA1-GPG at 4 °C. 54

2-16. Time course PNA-protein hybridization with the DNA nanocage. 60

2-17. Temperature dependence of FAM Fluorescence. 64

2-18. Fluorescence spectra of the TMR labeled PNA-proteins bound to the FAM labeled DNA nanocage as a function of temperature. 65

2-19. Calculated dye labeled PNA-protein absorption spectra and spectral changes upon hybridization to the DNA nanocage. 68

2-20. Donor (FAM) fluorescence decay kinetics attached to DNA nanocage with and without the acceptor (TMR) labeled PNA-proteins hybridized to the DNA nanocage. 70

2-21. Decay associated spectra (DAS) of the FAM labeled DNA nanocage hybridized to either the TMR labeled PNA-azurin or PNA-cytochrome c. 73

2-22. Visible spectra of cytochrome c, PNA-CytC and DNA-PNA-CytC. 82

2-23. Circular dichroism of cytochrome c, PNA-CytC and DNA-PNA-CytC. 84

3-1. Schematic of the designed dye labeled DNA-PNA-peptide complex. 88

Figure Page

3-2. DNA and PNA sequences and schematic showing their arrangement in the complex. 91

3-3. Complex nomenclature with list of constituent DNA and PNA-peptide strand names. 92

3-4. Native PAGE of the DNA nanocage hybridized to increasing molar excess of the PNA1-peptide labeled with the fluorescent dyes FAM and TMR. 96

3-5. Absorbance of the labeled PNA1-peptide hybridized to the DNA nanocage. 98

3-6. Time course PNA1-peptide hybridization with the DNA nanocage. 99

3-7. Fluorescence spectra of the TMR and FAM labeled PNA1-peptide bound to the FAM and TMR labeled DNA nanocage as a function of temperature. 101

3-8. Fluorescence decay of the DNA nanocage with two fluorescently labeled PNA-peptides. 104

3-9. Fluorescence Decay and Anisotropy of FAM and TMR dyes attached to a complex of a PNA-peptide inside the DNA nanocage. 106

3-10. Förster distance measurements between FAM and TMR dyes attached to a complex of a DNA nanocage with 0, 1 or 2 PNA-peptides inside. 108

4-1. Proposed design to separately assemble two proteins into a DNA nanocage using a PNA linker. 117

4-2. DNA and PNA sequences and schematic showing their arrangement in the complex. 119

4-3. Complex nomenclature with list of required DNA and PNA-protein strands for assembly. 120

Figure Page

4-4. Scheme used to produce the following dye labeled and unlabeled PNA-protein conjugates: PNA-CytC, PNA-CytC-TMR, PNA-Az, PNA-Az-TMR. 122

4-5. Scheme for purifying PNA-protein using toehold-mediated DNA strand displacement. 123

4-6. DNA sequences for toehold mediated PNA purification. 125

4-7. Results for PNA-protein conjugate purification by toehold-mediated strand displacement. 127

4-8. Gel showing the binding of PNA-cytochrome c, PNA-azurin or PNA-FAM to the DNA nanocage. 130

4-9. Fluorescence decay and anisotropy of two TMR dye labeled proteins hybridized to a DNA nanocage with PNA. 132

4-10. Steady State Fluorescence Spectra, Molecular Model of the DNA-PNA-protein complexes and linker chemical structure. 135

4-11. In-gel activity of PNA-cytochrome c bound to the DNA nanocage..... 139

4-12. Spectroelectrochemical results of PNA-cytochrome when bound to the DNA nanocage. 140

5-1. Schematics showing the DNA and PNA sequence arrangement in the complexes. 148

5-2. Native PAGE (7%) of labeled and unlabeled DNA nanocages for binding two γ-PNAs. 151

5-3. Scheme for labeling γ-PNAs with Cy5-DBCO using copper-free click chemistry. 152

Figure Page

5-4. Purification of Cy5 labeled γ-PNAs using denaturing PAGE and time course elution from gel blocks. 154

5-5. Characterization of purified Cy5 labeled γ-PNA1 by MALDI-MS and native PAGE..... 155

5-6. Characterization of purified Cy5 labeled γ-PNA3 by MALDI-MS and Native PAGE..... 156

5-7. Gel showing the binding of the unlabeled and labeled γ-PNAs to the DNA nanocage. 158

6-1. Photosynthetic electron transport chain. 162

6-2. Photosystem II structure and electron transport chain. 164

6-3. Artificial system to create fuel from sunlight and water. 167

6-4. Peptides identified from PSII for assembling an artificial OEC. 169

6-5. Model and schematic of assembling the PSII active site within a DNA nanocage using γ-PNA linkers. 171

6-6. Proposed conjugation scheme to form cyclic γ-PNA-peptides and DNA-peptides for assembling the artificial OEC. 173

LIST OF ABBREVIATIONS

Chemicals

DBCO, dibenzocyclooctyl; EtBr, ethidium bromide; TCEP, tris(2-carboxyethyl)phosphine; TFA, trifluoroacetic acid; TIS, triisopropylsilane; (Sulfo) SMCC, (sulfo)succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate; SPDP, N-Succinimidyl 3-(2-pyridyldithio)-propionate; EDTA, ethylenediaminetetraacetic acid; TMR, tetramethylrhodamine; FAM, fluorescein; Cy3, cyanine-3; Cy5, cyanine-5; NHS, N-hydroxysuccinimide; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; DTT, dithiothreitol; DEAE, diethyl-aminoethanol; AEM, aminoethylmaleimide; TAE, tris acetic acid EDTA; TBE, tris boric acid EDTA; ACN, acetonitrile; HATU, 2-(1H-7-Azabenzotriazol-1-yl)--1,1,3,3-tetramethyl uronium hexafluorophosphate methanaminium; DIEA, diisopropyletheyla-mine; NMP, N-methylpyrrolidone; PyBOP, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate; HBTU, O-Benzotriazole-N,N,N’,N’-tetramethyl-uronium-hexafluoro-phosphate; HOBt, Hydroxybenzotriazole; APS, ammonium persulfate; TEMED, tetramethylethylenediamine, ATBS, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid).

Molecules

DNA, deoxyribonucleic acid; cDNA, complementary DNA; ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; RNA, ribonucleic acid; PNA, peptide nucleic acid; γ-PNA, gamma modified PNA; PSII, photosystem II; OEC, oxygen-

evolving complex, CytC, cytochrome c; Az, azurin; Da, Dalton; nt, nucleotide; bp, base pair; Ac, acetylated; PDB, protein data bank.

Methods and Instruments

SPPS, solid phase peptide synthesis or solid phase PNA systhesis; PAGE, poly-acrylamide gel electrophoresis; SDS-PAGE, sodium dodecyl sulfate PAGE; RP-HPLC, reverse-phase high pressure liquid chromatography, MALDI-MS, matrix assisted laser desorption ionization mass spectrometry; SE-FPLC, size-exclusion fast protein liquid chromatography; IEX-FPLC, ion-exchange fast protein liquid chromatography; FRET, Förster resonance energy transfer; VV, vertical excitation and vertical emission; VH, vertical excitation and horizontal emission; FWHM, full width at half-maximum; IRF, instrument response function; DAS, decay associated spectra; CCD, charge coupled device; AU, absorbance unit; CD, circular dichroism; UV-Vis, ultraviolet-visible; Ex/Em, excitation and emission; T_D, dissociation temperature ; TCSPC, time-correlated single photon counting; D, donor; A, acceptor; DA, donor-acceptor; E_T, energy transfer efficiency; Q_Y or Φ, quantum yield; R_0, Förster radius; r, Förster distance; PSA, post source artifact.

Other

MWCO, molecular weight cutoff; IDT, Integrated DNA Technologies, 2D, two dimensional; 3D, three dimensional.

PREFACE

Significant portions of the text used in this dissertation were used directly from manuscripts published or submitted for publication with permission from the publishers shown in appendix A.

Although the majority of the work described in this dissertation in Chapter 3, 4 and 5 was done by Justin Flory, the PhD candidate, some of the work was carried out by others as a part of Center for Bio-Inspired Solar Fuel Production, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001016. A breakdown of the collaborative work done by other members of the center, in close collaboration with the PhD candidate, is described as follows, and their contributions are greatly appreciated and so by acknowledged here. The figures containing the major results provided by the collaborator is listed for convenience. The PhD candidate used the results provided by the collaborators to produce the figures in this dissertation.

Trey Johnson (undergraduate student, working under the supervision of Justin Flory): γ-PNA + Cy5 conjugation and PAGE purification and elution from gel blocks (Figure 5-3, Figure 5-4); assembly and purification of labeled and unlabeled DNA nanocages for γ-PNA (Figure 5-7) and some DNA nanocages for PNA-protein assembly (Figure 2-14); purity assessment by native PAGE and MALDI of γ-PNA-Cy5 (Figure 5-5, Figure 5-6); hybridization of unlabeled and Cy5 labeled γ-PNA with DNA nanocages (Figure 5-7); some purification of ssDNA for DNA nanocages. Figure 5-3 was prepared by Trey Johnson.

Chad Simmons: PNA + Azurin conjugation and purification (Figure 2-9); DNA nanocage molecular model shown in Figure 6-5; DNA-PNA nanocage molecular model shown in Figure 2-18, Figure 4-10 and Figure 4-12; molecular models of azurin and cytochrome c shown in Figure 4-1, Figure 4-4, Figure 4-5, Figure 4-10 and Figure 4-12.

Ingo Grotjohann: Prepared Figure 6-4 and the molecular model of photosystem II used in Figure 1-1 and Figure 6-2.

Alessio Adreoni: Expression and purification of azurin.

CHAPTER 1 – Introduction

Due to the interdisciplinary nature of the work presented in this dissertation, I would like to first explore the broader context that links together the various disciplines and fields, and how this relates to the my background, interests and ambitions. The first section of the introduction provides insight into how nanotechnology is used in biological systems (i.e., bionanotechnology) and the potential impact of making improvements at the nanoscale. The next section introduces DNA nanotechnology, emphasizing some of the work that has been done for organizing peptides and proteins, but underscoring the need to develop methodologies for trying to assemble highly interconnected DNA-polypeptide complexes within appropriately sized (<4 nm radius) nanostructures in order to mimic protein catalytic active sites. The next section introduces Peptide Nucleic Acid (PNA) as a compelling hybrid molecule that is ideal for developing a completely new type of methodology for connecting a number of peptides within small DNA nanoscaffolds. The introduction proceeds to describe the specific aims of the work described in this dissertation, and concludes by describing the outline and organization of the work contained herein.

Bionanotechnology: Understanding and reengineering nature's molecular machines

Engineering is a powerful methodology, unique to humankind, which uses knowledge of basic scientific principles to design and build new and transformative technologies that have not existed previously except in the mind of the designer. As an applied physics undergraduate student taking a cellular and molecular biology course at Sonoma State University (CA), I gazed in amazement at some of nature's remarkable

'molecular machines' and wondered how we might be able to use them and improve them to benefit society. Biological systems carry out a huge diversity of sophisticated tasks on the molecular level, such as self-assembling and repairing materials, self-reproducing, transporting cargo, transmitting and receiving signals, capturing and converting different forms of energy among many others, all which are carefully controlled through complex feedback mechanisms. At the core of each of these extraordinary capabilities are proteins and other biomolecules that operate on the nanoscale (i.e., bionanotechnology), and through a remarkable set of organizational structures, seamlessly drive macroscopic functions. Figure 1-1 demonstrates the hierarchical level of organization in a plant, showing the careful ordering from the nanoscale world of molecules and proteins (0.1–100 nm), to the mesoscale world of chloroplasts and cells (0.1–1000 μm), and finally to the macroscale world of leafs and plants (0.001–100 m).

The fact that the molecules and proteins within a plant have a profound impact on the functionality of a leaf underscores how research and engineering at the level of bionanotechnology can impact our world in a wide variety of ways to meet the current and future needs of humankind. Furthermore, these incredible molecular machines function and survive using widely available resources, such as sunlight, water and oxygen, and are constructed using the earth's abundant elements (e.g., C, H, N, O, P, S). Thus biology provides an ideal recipe for sustainable processes to manufacture goods, such as fuels, food and plastics, and to perform services, such as waste water remediation and CO_2 sequestration. These processes could help to mitigate the current conflicts between human development and the environment, such as accumulation of non-biodegradable waste, air and water contamination, greenhouse gas emission and climate change.

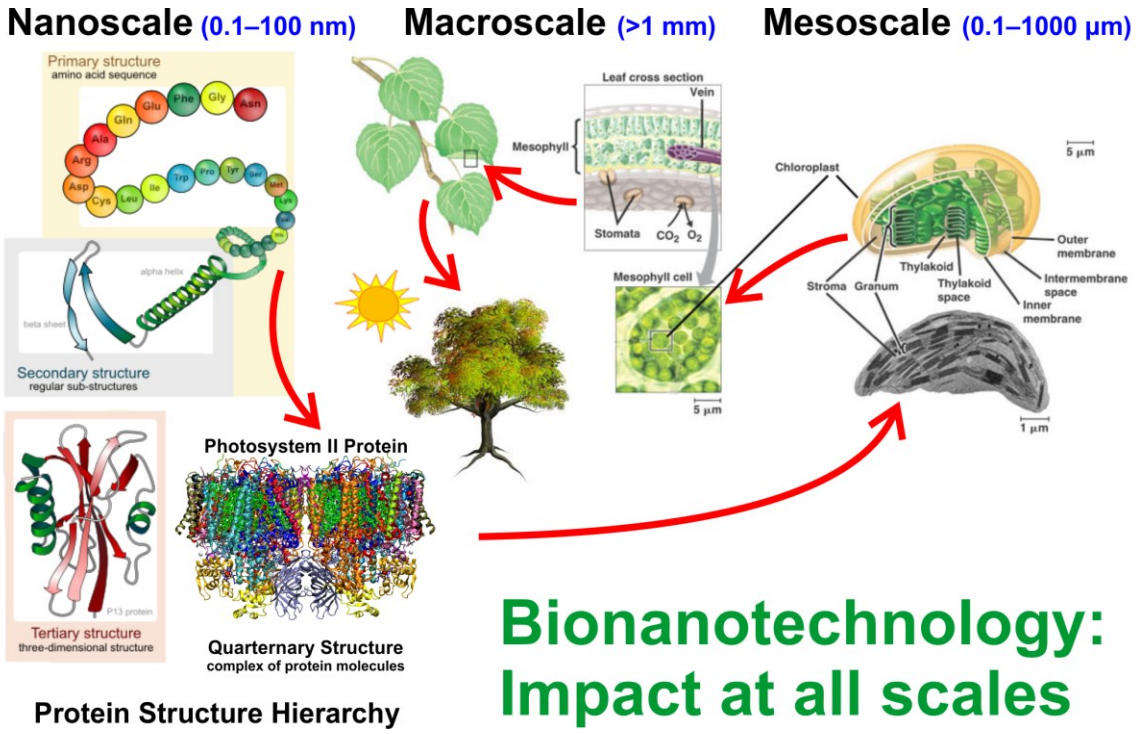


Figure 1-1. Hierarchy of organization in a plant. [Left, Nanoscale] Hierarchy of protein structures, which serve as nature's molecular machines on the nanoscale: 1) amino acids are joined to make polypeptides; 2) polypeptides self-assemble into stable structural motifs, such as an alpha helix or beta sheet; 3) these motifs foldup to form functional proteins; 4) some proteins interact to form complexes with other proteins and biomolecules with more sophisticated capabilities.[1] [Right, Mesoscale] In plants, photosynthetic proteins operate within the membranes of stacked thylakoids in the chloroplast, which reside in the mesophyll cells in plant leaves.[2] [Center, Macroscale] Trees arrange the leaves in a canopy to capture a large area of sunlight and convert light energy into the chemical energy for growth and proliferation. Photosystem II image: crystal structure solved by Umena et al (PDB accession code 3ARC).[3] The image was created from 3ARC by Ingo Grotjohann.

With the intention of trying to find ways to mimic nature (i.e., biomimicry) or even improve on nature (i.e., biological engineering) to address critical needs for sustainable manufacturing, I joined the interdisciplinary Biological Design Graduate Program at Arizona State University (ASU). Bionanotechnolgy spans a range of scientific disciplines, including physics, chemistry, biology and engineering, so a training program that emphasizes interdisciplinary research would be critical to navigating the complex inter-

faces between these disciplines. ASU is also one of the world's leading institutions for photosynthesis research and DNA nanotechnology, making it a perfect fit for cutting edge research in bionanotechnology. I was immediately drawn to a challenging project that looked to leverage significant advancements in engineering DNA nanostructures to reengineer some of nature's molecular machines (i.e., peptides and proteins) that were important for solar energy conversion in photosynthesis. With funding from the Department of Energy for an Engineering Frontier Research Center, the Center for Bio-Inspired Solar Fuel production was facilitating the research necessary to build bio-inspired catalysts for sustainable fuel production. However, in order to use DNA nanoscaffolds for assembling bio-inspired catalysts, a number of new tools needed to be developed and characterized to pave the way for this promising line of research, and as such, is the primary focus of the work in this dissertation.

DNA nanostructures for engineering polypeptides

The field of structural DNA nanotechnology has transformed our thinking about the purpose and possibilities of canonical base pairing biopolymers, which can have a much broader structural landscape than what has evolved in natural organisms. Deoxyribonucleic acid (DNA) is a ubiquitous biological polymer evolved as a stable means for storing the information required to synthesize ribonucleic acid (RNA) and protein molecules, which are used by the organism to do useful work. Two DNA strands are normally intertwined in a helical conformation and held together through bonding between complementary bases, as originally discovered by Watson and Crick:[4] adenosine (A) forms two hydrogen bonds with thymine (T), (or uracil (U) in the case of RNA) and cytosine

(C) forms three hydrogen bonds with guanine (G). When chromosomal DNA is copied, strand breaks can occur at the same point in the original and copied strands, which allows the strands to crossover, creating what is called a Holliday junction,[5] as show in Figure 1-2A. Because the sequences in the original and copied strands are identical, the Holliday junction can migrate up and down both strands.

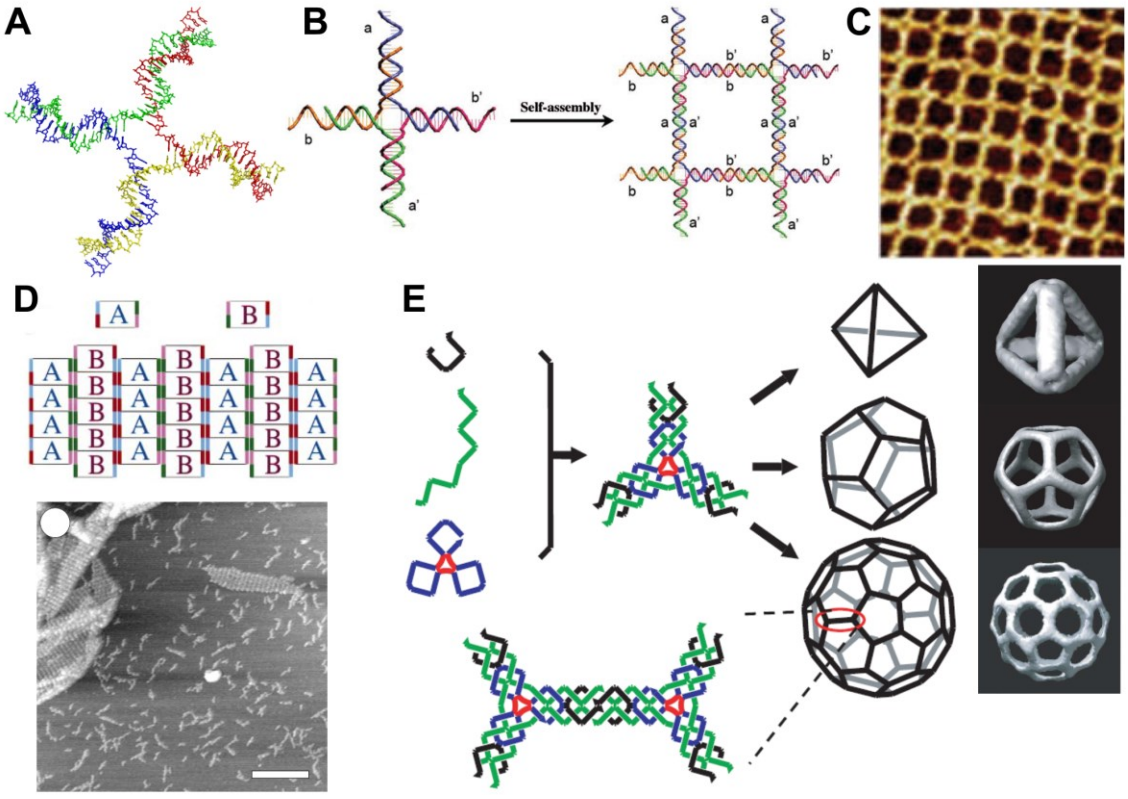


Figure 1-2. DNA Holliday junction and examples of 2D and 3D DNA nanostructures made of synthetic oligonucleotides. [A] Natural crossover junction formed between two DNA strands during recombination. Image from PDB accession code 3CRX.[6] [B] Model showing a crossover tile with sticky ends a and b and complementary sticky ends a' and b'[7] that assemble to form a periodic lattice as shown in [C].[8] [D] Non-self-complementary tiles A and B are preassembled and then mixed to form periodic lattices.[9] Scale bar 300 nm. [E] (left) Schematic showing how short oligonucleotides are preassembled into crossover motifs and subsequently mixed to form large 3D nanostructures, such as a tetrahedron, dodecahedron and Buckyball. (right) Cryo transmission electron micrograph of each 3D shape.[10] Figures reproduced with permission from: [B] ref. [7], © Wiley 2006, [C] ref. [8], © AAAS 2003, [D] ref. [9], © AAAS 1998, [E] ref. [9], © NPG 2008.

In a pioneering effort, Nadrian Seeman theorized[11] and later showed[12] that similar junctions could be assembled using shorter strands of DNA (oligonucleotides), but by restricting the complementary regions to only part of the sequence, stable 2D and 3D nanostructures could be engineered. This work launched a new field now called 'structural DNA nanotechnology'; the reader is referred to several excellent reviews of the progress made in this dynamic field and the challenges that remain.[13–15] Over the ensuing three decades, the rules for constructing DNA nanostructures have been refined,[16] but is still are largely governed by simple Watson-Crick base pairing between complementary DNA sequences, making them significantly more straightforward to engineer than proteins or peptides.[17,18]

Seeman's stable crossover structure set the stage for building a wide assortment of nanostructures assembled using mostly short [<100 nucleotides (nt)] oligonucleotides that could be readily synthesized. One common motif developed early on was a crossover tile, which structurally mimicked a small Holliday junction (Figure 1-2B),[7] but included short single stranded domains called 'sticky ends' (a and b in Figure 1-2B) that were complementary to sticky ends somewhere else on the same tile (a' and b' in Figure 1-2B).[7] In this way the tiles would spontaneously assemble together into large lattices, which could be further decorated with proteins, nanoparticles or other guest molecules with highly controlled nanometer spacing, as shown in Figure 1-2C.[8] More complex tiles, in which strands crossover two[9] or three times[19] (i.e., double or triple crossover) were soon developed, along with non-self-complementary tiles (A and B in Figure 1-2D) that allowed each tile to be preassembled in high yield and subsequently mixed to build more complex periodic patterns, as shown in Figure 1-2D.[9] This technique was later extended to build

3D structures, such as the tetrahedron, dodecahedron and Buckyball shown in Figure 1-2E,[10] by using other variations on the preassembled tile. One of the capstone accomplishments of this era of the field, was the rational design and assembly of a 3D DNA crystal by the group of Seeman.[20]

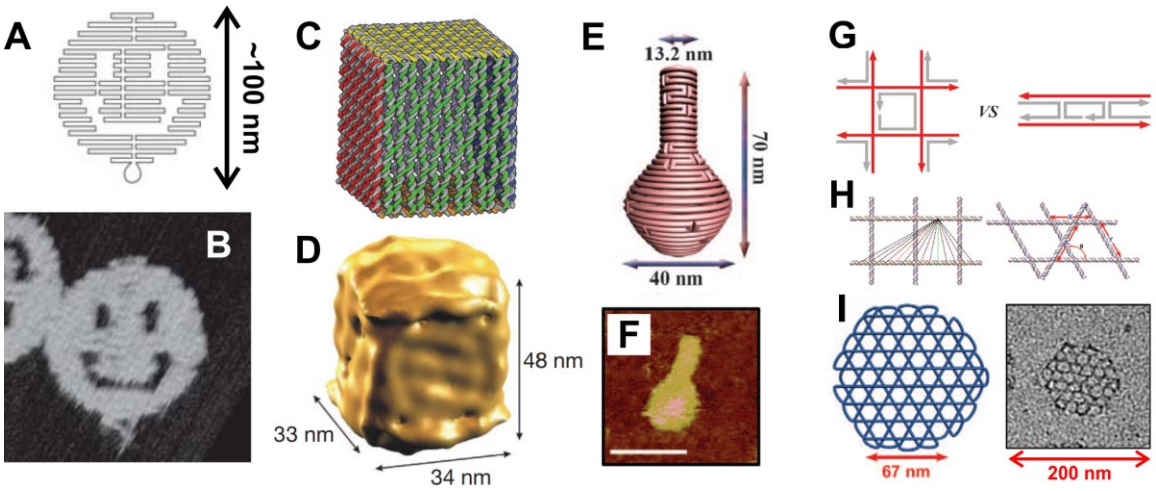


Figure 1-3. 2D and 3D DNA nanostructures made using DNA origami. [A] Schematic [A] and AFM image [B] of a 2D DNA smiley face made from a >7000nt M13 bacteriophage DNA strand folded together like 'origami' and held together by over 200 short oligonucleotide 'staple' strands.[21] Schematic [C] and reconstructed cyro transmission electron micrograph (TEM) [D] of a hollow 3D DNA box made using DNA origami.[22] DNA cylinder model [E] and TEM [F] of a DNA nanoflask made using DNA origami.[23] Scale bar for [F] is 75 nm. [G] Basic crossover motif used in DNA origami (right) and a four arm junction (left) for increasing the angle of helical junctions [H] for designing more complex gridiron structures. [A] A hexagonal gridiron schematic (left) and TEM image (right).[24] Figures reproduced with permission from: [A] and [B] ref. [21], © NPG 2006, [C] and [D] ref. [22], © NPG 2009, [E] and [F] ref. [23], © AAAS 2011, [G], [H] and [I] ref. [24], © AAAS 2013.

One of the most transformative advancements in the field, that significantly expanded the size and complexity of controlled DNA nanostructure assembly was Paul Rothemund's seminal paper describing what he called 'single-stranded DNA origami'.[21] DNA origami uses a long (>7000nt) strand of DNA from the M13 bacteriophage, which can be carefully folded together into a massive number of different shapes and structures,

which are held together by over 200 short oligonucleotides called 'staple strands'.[21] Figure 1-3[21] shows the outline of the M13 strand as it is woven together into a smiley face [A] and the corresponding atomic force microscopy (AFM) image [B] of the assembled structure. Due to the large number of unique staple strands, each one could be uniquely functionalized to systematically arrange a large number of guest molecules almost anywhere on this molecular pegboard. This technique was later extended to build remarkable 3D shapes, including a hollow box shown in Figure 1-3C, D.[22] The group of Hao Yan made significant advancements to DNA origami based structural design by establishing design principles for producing complex curved 3D shapes, such as the DNA nanoflask shown in Figure 1-3E, F.[25] Furthermore, the Yan group showed that the basic DNA origami crossover pattern, shown in Figure 1-3G (right), could be replaced by a four arm junction Figure 1-3G (left), resembling the Holliday junction, which allowed for a much broader junction angle between the joined helices (Figure 1-3H) and thus more complex gridiron-like structures (Figure 1-3I).[24]

Because the functional diversity of unmodified DNA is limited, much effort has been focused on introducing functional groups. DNA scaffolds have been used to introduce polymers to form foldamers[26] or block copolymers,[27] and metals to arrange nanoparticles,[28–30] form nanowires,[8,31] and nanomagnets.[32] DNA nanostructures have also been developed for hybridization based biosensing, such as for unlabeled detection of RNA.[33] Novel developments have been shown to introduce conformational dynamics into normally static DNA nanostructures by introducing 'fuel' DNA strands that competitively replace strands in an existing structure such that it can switch to a new conformation,[34,35] and even 'walk' across a DNA surface.[36] DNA strands can also be cop-

ied in huge quantities and with high fidelity using enzymes *in vitro* by means of polymerase chain reaction, helping to reduce the potential cost of using DNA nanostructures in commercial applications. The group of Hao Yan made significant progress in this direction, by designing DNA nanostructures that form spontaneously after *in vitro* enzymatic amplification,[37] and later showed that genes encoding DNA nanostructures could be transformed and assembled within living cells.[38]

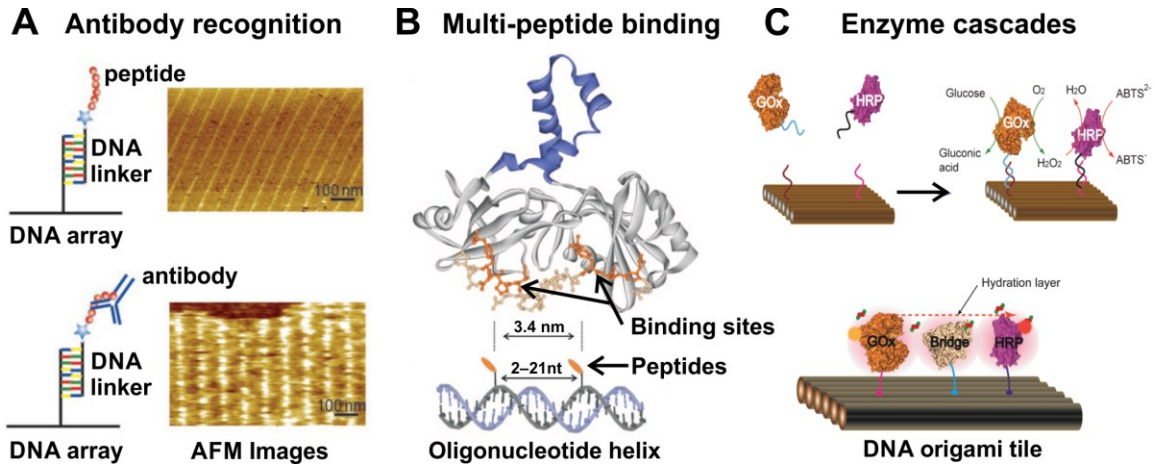


Figure 1-4. Arranging polypeptides with DNA nanostructures . [A] (Left) Schematics showing a peptide epitope conjugated to a DNA linker attached to a 2D DNA array (top) and then with the complementary antibody bound (bottom). (Right) AFM images of the same constructs showing rows of peptides (top) and the bound antibodies (bottom) on the DNA surface.[39] [B] Schematic showing the crystal structure of a peptide binding protein Syk kinase, the peptide binding sites, and an oligonucleotide helix presenting two peptides with various nucleotide (nt) separations.[40] [C] (top) Assembly of an enzyme cascade on a DNA origami tile where glucose oxidase (GOx) produces H_2O_2 using glucose, which is used by the adjacent enzyme horseradish peroxidase (HRP) to produce a detectable substrate $ATBS^{-}$. The rate of catalysis in the multi-enzyme system was enhanced at shorter separation distances and using the water shell from a non-catalytic protein.[41] Figures reproduced with permission from: [A] ref. [39], © Wiley 2007, [B] ref. [40], © Wiley 2011, [C] ref. [41], © ACS 2012.

DNA scaffolds offer a flexible, biocompatible workbench to explore the wide variety of functions performed by proteins in nature, such as catalysis, scaffolding, molecular transport and molecular recognition among others. Methods for manipulating proteins

and peptides on 1D and 2D DNA nanostructures have been utilized to organize proteins on 2D arrays,[8] to study antibody recognition of peptide epitopes,[39] for selectively arranging multiple enzymes,[42] investigating DNA binding proteins that recognize single stranded aptamer sequences[43,44] or specific DNA sequences in the major groove of the DNA helix,[45] assembling complementary protein subunits[46] or protein ligands,[40,47] and for investigating the spatial dependence of enzyme cascades.[48,49] Some of these applications using DNA to arrange polypeptides are shown in Figure 1-4.[39–41] 3D DNA nanostructures have also been used to encapsulate[50] and arrange proteins[51–53] as well as to organize membrane proteins for NMR structure determination.[54] The relative ease with which DNA nanostructures are now constructed, compared to the challenges associated with protein and peptide engineering,[17,18] raises the possibility of using DNA to investigate and modulate polypeptide function.

Small 3D DNA nanostructures (radius <4 nm, <100 kDa) are particularly interesting because they allow each spatial dimension to be controlled on the scale of typical peptide and protein intra and intermolecular interactions. While a number of techniques are available for identifying protein-protein interactions ,[55,56] many of these interactions are dependent on the method used to discover them or are non-functional,[57] so new methods are needed to validate and further characterize the specific properties of each interaction.[58] Detailed characterization of proteins and protein complexes may require relatively high molar concentration, such as for spectroelectrochemistry[59] or circular dichroism,[60] which are not practical with large DNA origami nanostructures (> 4 MDa).[21] Although, DNA provides a spatially precise scaffold for arranging polypeptides, its biological origin makes it more prone to naturally interact with the target molecules. For example, some

proteins contain specific DNA binding domains,[61] which may strongly interact with the DNA scaffold, while other proteins may be influenced by non-specific interactions with the negatively charged backbone or hydrophobic nucleobase core. For these reasons, a better understanding of the interactions between several proteins and the DNA scaffold is needed to establish guidelines for using DNA scaffolds to study proteins.

Small DNA nanostructures are also better suited to investigate intramolecular protein interactions (i.e., interactions occurring within a protein), which could further elucidate relationships between a proteins structure and function. Some proteins carry out multiple functions, such as photosystem II, which performs solar-driven water oxidation in photosynthetic organisms, and involves light capture, photochemical conversion of light into electrical energy, and catalysis. Small DNA nanostructures are an intriguing platform to decouple and investigate these processes independently, and even optimize and reengineer them. However, this type of extensive polypeptide manipulation may require assembling a number of polypeptides within a small DNA cavity, which is not practical using current conjugation methods. Furthermore, high resolution structural information, which elucidates the orientation of the DNA helices, is essential to properly orient the polypeptides within the DNA nanostructure. However, due to the challenges associated with obtaining high resolution DNA structures, few are available.[62–64]

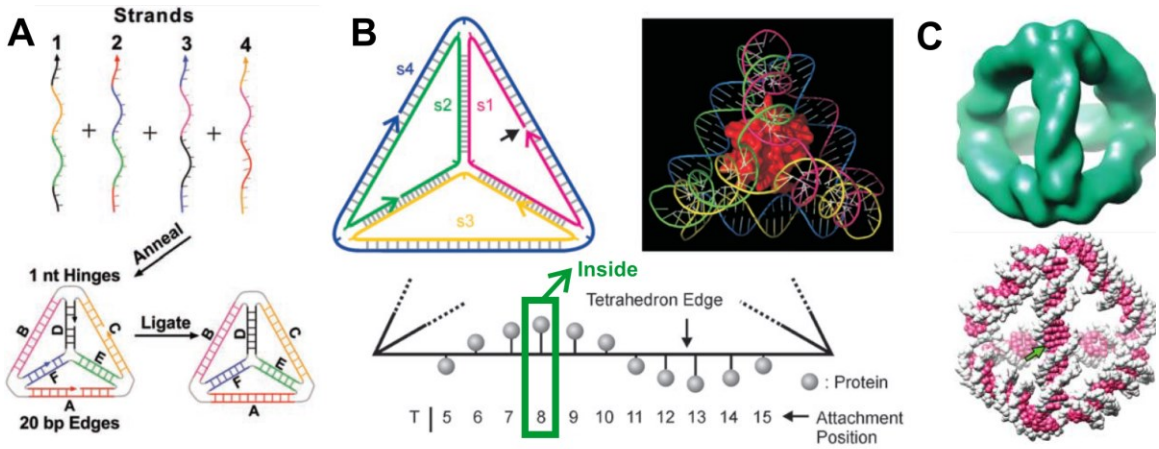


Figure 1-5. DNA tetrahedron assembly, protein encapsulation and cryo Electron Microscopy density. [A] Schematic showing the assembly of a 3D DNA tetrahedron made from four 63nt DNA strands (top) each containing a 20nt domain complementary to a domain of matching color on each of the other strands. After annealing, the 5' and 3' ends of each strand can be enzymatically ligated to improve stability.[65] [B] A strand schematic (left) and molecular model (right) of the DNA tetrahedron encapsulating the protein cytochrome c. The protein was attached at each of the indicated base positions (5–15) and position 8 was found to orient the protein inside the tetrahedron.[50] [C] Electron density (top) and molecular model (bottom) of the DNA tetrahedron. The green arrow points to the major groove in the center of the helix facing outwards.[64] Figures reproduced with permission from: [A] ref. [65], © AAAS 2005, [B] ref. [50], © Wiley 2006, [C] ref. [64], © ACS 2009.

One particularly well characterized small DNA nanostructure is an 78 kDa DNA tetrahedron with 7 nm edges, as originally designed by Goodman, Turberfield and coworkers.[65] The tetrahedron is assembled by mixing each of the four 63nt oligonucleotides at low concentration (< 1 μM) followed by high temperature annealing (i.e., from >90 ºC to < 25 ºC). The assembly process is shown in Figure 1-5,[65] along with an optional ligation step to connect the ends of each strand post-assembly to provide additional stability. The authors were able to identify the location of the major and minor grooves and approximate base orientations through a series of DNA linking[65] and protein encapsulation experiments,[50] along with electron density from cryo-electron microscopy.[64]

Figure 1-5B[50] shows a sequence schematic of the DNA tetrahedron along with the various base positions where the protein cytochrome c was attached. Base position 8 was found to orient the protein inward such that it was fully encapsulated, as shown in the molecular model in Figure 1-5B.[50] Cryo electron micrograph electron density (Figure 1-5C top)[64] was subsequently used to identify the location of the major and minor grooves and further refine the base orientations from their previous studies. A molecular model of the DNA tetrahedron is shown in Figure 1-5C (bottom).[64] Together, this structural information provides an excellent guide for determining the base positions to introduce other guest molecules with well-defined spatial orientations.

Biological molecules can be challenging to work with because their structure and function can be disturbed through solution additives, such as organic solvents or metal salts, or when exposed to elevated temperatures. The temperature sensitivity is particularly challenging because of the typical high temperature annealing step required during DNA nanostructure assembly. Smaller DNA nanostructures can be assembled quite rapidly (e.g., 0.5–5 min), and sometimes from lower initial starting temperatures (e.g., 54 °C),[50] which may limit thermal damage to the guest molecule. However, large DNA nanostructures can take up to a week to assemble properly.[25] In order to decouple the complex and lengthy assembly process of large DNA nanostructures, the staple strands can be lengthened at defined positions in order to provide an unhybridized domain for capturing a secondary 'helper strand' post-assembly under mild conditions (e.g., <40 °C), which is conjugated to a guest molecule.[29] In this approach, the sequence of the helper strand provides a unique molecular recognition domain for placing the guest molecule on the DNA nanoscaffold, and allows a single robust bio-orthogonal chemical coupling

method to be used for conjugating a number of guest molecules each to a unique helper strand.

The two step process, of first preassembling the DNA nanostructure followed by assembling guest molecule(s) using helper strands, works well with large DNA nanostructures when the molecules are more widely spaced (>5–10 nm). To allow for sufficient thermal stability at room temperature, typical DNA helper strands sequences are at least 15–20nt long, which form 4.9–6.5 nm long double-stranded DNA (dsDNA) helices upon hybridization. However, such long linker sequences put significant constraints on how the linker can be used in small DNA nanostructures (radius <4 nm). The linker length could be shortened by increasing the GC content of the sequence, however this would limit the sequence variability needed to attach more than a few guest molecules. While conjugation strategies using nucleic acid linkers provide a large sequence space for directing their assembly, there is a need to increase the thermal stability of these linkers for assembling guest molecules within small DNA nanostructures.

PNA for linking polypeptides to DNA

Peptide nucleic acid (PNA)[66] is a DNA analog, which contains a peptide-like backbone and nucleobase side chains, allowing it to hybridize with a complementary DNA, RNA or PNA stand. The chemical structure of a peptide, PNA and DNA backbone is shown in Figure 1-6A. Because PNA lacks the negative charges contained on the DNA backbone, it can bind with higher stability to a complementary DNA strand because the backbone repulsion is absent; this is especially true at low salt concentrations where the negative DNA phosphates are poorly shielded. Purine rich PNA sequences also tend to

have significantly higher thermal stability because the PNA backbone is more flexible than the DNA backbone and allows for greater pi-pi stacking between adjacent nucleobases (P.E. Nielsen, personal communication, January 9th, 2013). This remarkable stability allows it to bind to double stranded DNA and displace the DNA strand with the same sequence, by so called strand invasion.[66] Like peptides, PNAs are linked by amide bonds and can be synthesized using similar solid phase protocols,[67] thus allowing for seamless synthesis of PNA-peptide conjugates.[68] A microwave assisted solid phase peptide synthesizer, which can also be used to synthesize PNA strands and PNA-peptide conjugates, is shown in Figure 1-6B.[69]

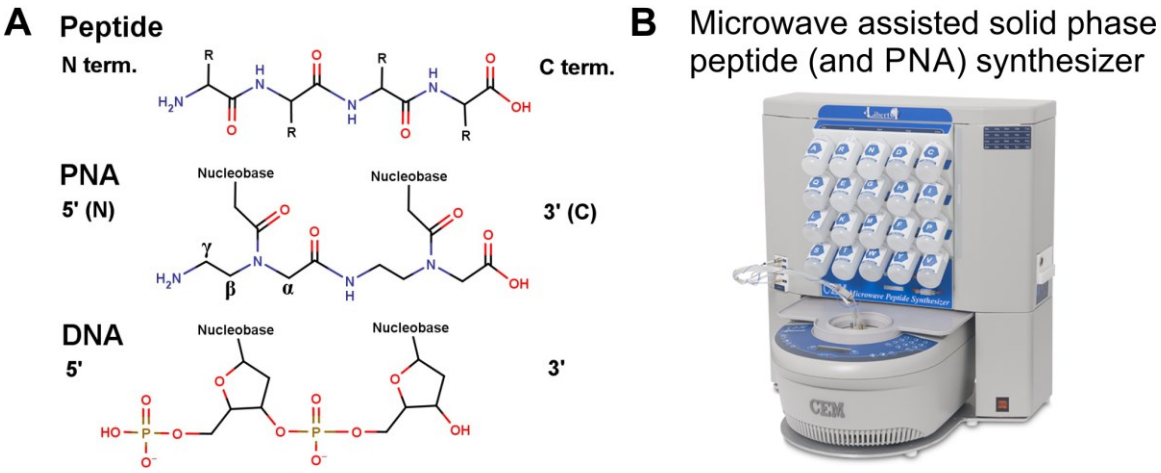


Figure 1-6. Chemical structures of PNA, DNA and peptide backbones, and a microwave assisted solid phase peptide (and PNA) synthesizer. [A] Chemical structure of a peptide, PNA and DNA strand. The α, β and γ substitution positions are shown on the PNA backbone. [B] Microwave assisted solid phase synthesizer for peptides and PNA.[69]

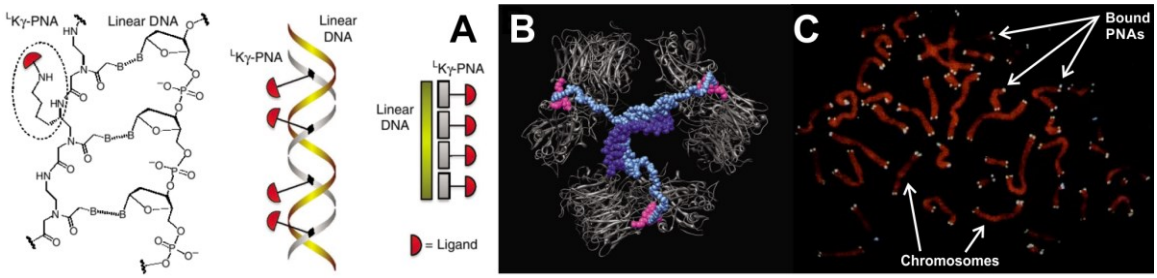


Figure 1-7. Applications with PNA. [A] (left) Chemical structure of a PNA-DNA duplex with a tripeptide ligand attached to a lysine residue at the gamma position of the PNA backbone. Ribbon (center) and cartoon (right) diagrams showing the same PNA-DNA duplex with multiple ligands attached to the PNA strand.[47] [B] Molecular model showing three ligands from the PNA-DNA duplex from [A] binding to the receptors on the integrin protein.[47] [C] Fluorescein (white) labeled PNA strands hybridized *in situ* to chromosome telomeres counterstained with propidium iodide (orange).[70] Figures adapted with permission from: [A] and [B] ref. [47], © NPG 2012, [C] ref. [70], © Oxford 1996.

PNA has been used for a number of other applications, which have recently been reviewed,[71] including a review focused on chemical modifications to PNA.[72] Its synthetic backbone makes it highly resistant to nucleases and proteases[73] and useful for targeting DNA[70] and RNA,[74] as well as using it as a molecular probe[75] and biosensor.[76,77] Figure 1-7C[70] shows fluorescently labeled PNA strands hybridized *in situ* to terminal repeating nucleotide sequences (telomeres) on each chromosome. PNA has also been assembled within DNA scaffolds to investigate applications in bionanotechnology.[47,78,79] Substitutions to the gamma position on the PNA backbone (see Figure 1-6) have been shown to increase the solubility and thermal stability of the PNA strand,[80] as well as provide a handle to introduce functional groups without affecting its binding to a complementary strand.[81] Figure 1-7A[47] shows the chemical structure, ribbon and cartoon representation of a PNA strand functionalized at the gamma position with a lysine side chain, which serves as the attachment point for a tripeptide ligand. A complementary DNA strand, with repeating complementary sequences for the ligand containing PNA strand, increases

the valency of the tripeptide ligand to enhance binding to the receptors on the integrin protein. Figure 1-7B[47] shows a molecular model of the ligand containing DNA-PNA duplex bound to the integrin protein.[47]

PNA is an intriguing synthetic polymer for bridging the diverse functional landscape of polypeptides with the straight-forward construction of nucleic acid nanostructures. PNA-peptide conjugates could be synthesized directly, which would facilitate hybridization based strategies for organizing peptides within DNA nanostructures. Because of PNAs increased thermal stability compared to DNA, only short sequences of PNA (8nt) are necessary, making them ideally suited for introducing a number of different peptides within small DNA nanostructures. When the PNA strands are functionalized at the gamma position, it would allow the peptide to be attached at any position along the backbone or even at multiple locations,[47,82] providing flexibility for arranging the peptides within the DNA nanostructure. Furthermore, conjugating the peptides at both ends to the PNA strand, to form cyclic PNA-peptides, would transfer additional structural support from the underlying DNA framework.

Specific Aims

Although a significant amount of work has been done manipulating proteins and peptides with DNA nanoscaffolds, relatively little is known about how the DNA nanostructures affect the function of the proteins and how the polypeptides interact with the DNA nanostructures. Furthermore, the work to date has not tried to spatially confine the polypeptides to the scale where intramolecular interactions within a protein dominate (i.e., <5 nm), or to precisely control the 3D orientation of the attached polypeptides.

The specific aims for the work in this dissertation are described as follows:

1. **Design** a DNA nanocage that can bind and assemble multiple PNA-polypeptides.
2. **Synthesize** PNA-peptide conjugates using microwave-assisted solid phase peptide (or PNA) synthesis (SPPS).
3. **Label** PNA-polypeptides with fluorescent dyes to study their assembly, thermal stability and arrangement in the DNA nanocage.
4. **Conjugate** PNA to biologically expressed proteins.
5. **Purify** PNA-protein conjugates from unreacted PNA and protein.
6. **Assemble** PNA-polypeptides within a DNA nanocage using a gentle and flexible method that preserves function and provides temporal control over the assembly.
7. **Validate** the proper placement of the PNA-polypeptides within the DNA nanocage.
8. **Characterize the stability** of the PNA-polypeptides within the DNA nanocage.
9. **Characterize the interactions** of the polypeptides with the DNA nanocage.
10. **Characterize the redox activity** of the polypeptides in the DNA nanocage.
11. **Characterize the structure** of the polypeptides in the DNA nanocage.
12. **Design a protein active site mimic** using a DNA-PNA-peptide complex.

Outline and organization

This dissertation will describe a series of methods developed for constructing multiple PNA-polypeptide conjugates, which utilize short sequences (8nt) of PNA, in order to assemble polypeptides at specific locations within a 3D DNA nananocage based on the sequence of each PNA strand. This is the first time PNA strands have been used to assemble polypeptides within 3D DNA nanostructures, which has significant advantages

over traditional methods that directly conjugate polypeptides to DNA. Chapter 3 describes a DNA nanocage design that facilitates binding multiple PNA strands, and develops a synthesis protocol to directly synthesize PNA-peptide conjugates using SPPS. The peptides were designed to be short, so that the arrangement of the PNA-peptides within the DNA nanocage could be assessed using energy transfer between fluorescent dyes attached to the peptides. Chapter 4 builds on this work, by investigating the effect of PNA assembly with much longer polypeptides. Because SPPS can be challenging for long PNA-peptide sequences, a protocol was developed to conjugate PNA to biologically expressed proteins. Two different proteins were chosen to investigate the effect the DNA nanocage would have on protein function, as well as the protein orientation within the DNA nanocage based on its surface charge. Chapter 5 describes modifications to the initial DNA nanocage design to accommodate a PNA strand binding to an alternate binding site, which would facilitate peptide arrangements that better mimic the scaffold provided by the protein photosystem II (PSII), for assembling an artificial catalytic center.

Chapter 6 provides a prospective of using DNA nanocages to build an artificial catalytic center from photosystem II, which is called the oxygen-evolving complex (OEC). A cyclic PNA-peptide design using gamma modified PNA strands is also discussed, which would further stabilize the attached peptides when assembled in the DNA nanocage. The detailed design principles contained in this chapter provide a logical direction for the scientific community to continue and build on this work, which could have a meaningful impact on engineering practical and economical water oxidation catalysts. Chapter 7 summarizes the findings from Chapters 3, 4 and 5 and gives an outlook for using DNA nanoscaffolds to assemble polypeptides using PNA linkers, which mimic pro-

tein active sites for investigating protein structure and function, and concludes with an outlook for creating DNA-PNA-peptide complexes with novel functions. The methods developed in this dissertation lay the foundation for building functional DNA-PNA-protein complexes and assembling peptide-based artificial catalytic centers.

CHAPTER 2 – Materials and Methods

A major goal throughout the work described in this dissertation was to develop a toolkit that would facilitate gentle and controlled peptide and protein assembly within 3D DNA nanostructures using PNA linkers. Methods developed early on were used repeatedly and expanded upon in subsequent work. Due to the significant overlap of the methods used in different parts of this work, all the methods used in this dissertation were compiled into a single all-encompassing section described here. In this way, the methods can be placed in a logical arrangement that should facilitate others that wish to begin working in this exciting field.

Some typical results are included for many of the methods to illustrate the type of performance that was achieved. Some of these figures also contain supporting discussion to clarify complex interpretation of the results. The results included in this section are mostly figures that support some of the main results discussed in more detail in Chapters 3–5, thus those chapters will refer back to this section for these supporting figures. The following is an outline of the general method sections included here, organized by topic:

1. **PNA preparative methods**: PNA-peptide synthesis, purification and functionalization; γ-PNA functionalization and purification.
2. **Protein preparative methods**: Cytochrome c preparation; Azurin expression and purification.
3. **PNA-protein conjugation and purification**: PNA-protein conjugation, purification and characterization; ssDNA toehold mediated PNA and PNA-protein purification.

4. **DNA-PNA-polypeptide complex preparation**: DNA nanocage preparation; PNA-polypeptide assembly with the DNA nanocage.
5. **DNA-PNA-polypeptide characterization**: DNA-PNA-polypeptide thermal stability and fluorescence characterization; Cytochrome c activity and redox potential; Circular dichroism of cytochrome c.
6. **Modeling**: Molecular structures, models and chemical structures.

PNA-peptide synthesis, purification and functionalization

PNA coupling

Fluorenylmethyloxycarbonyl (Fmoc) protected PNA monomers were purchased from PolyOrg (Leominister, MA, USA) and Fmoc protected amino acids from Merck (DE). The PNA domains were synthesized using a CEM Liberty Automated Microwave Peptide Synthesizer (Figure 1-6B) following methods similar to those reported by Fabani and coworkers.[83] Deprotection of and coupling to PNA monomers and the first glycine amino acid were performed as follows. Deprotection was performed at room temperature in 20% piperidine in dimethylformamide (DMF) initially for 60 seconds and a second time for 12 minutes after washing with DMF at 25 °C. The UV absorbance of the cleaved Fmoc at 301 nm was monitored in the waste stream using the CEM UV monitoring accessory, as shown in Figure 2-1. Coupling was performed once for 30 minutes at 75 °C using microwave heating in the ratio of monomer/activator/base of 1.0/1.0/4.0 in 4 fold excess to resin, where the monomer was dissolved in DMF, the activator was 0.2 M 2-(1H-7-Azabenzotriazol-1-yl)--1,1,3,3-tetramethyl uronium hexafluorophosphate methanaminium (HATU) in DMF, and the base was a mixture of 0.4 M diisopropyletheyla-

mine (DIEA) and 0.4 M lutidine in N-methylpyrrolidone (NMP). Initially we tried using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) as reported by Fabani and coworkers,[83] but we observed some PyBOP related impurities as reported by Pritz and coworkers[84] that were not observed with HATU. Following the recommendation of several PNA protocols[67,83] all DMF used in PNA synthesis was anhydrous (Sigma).

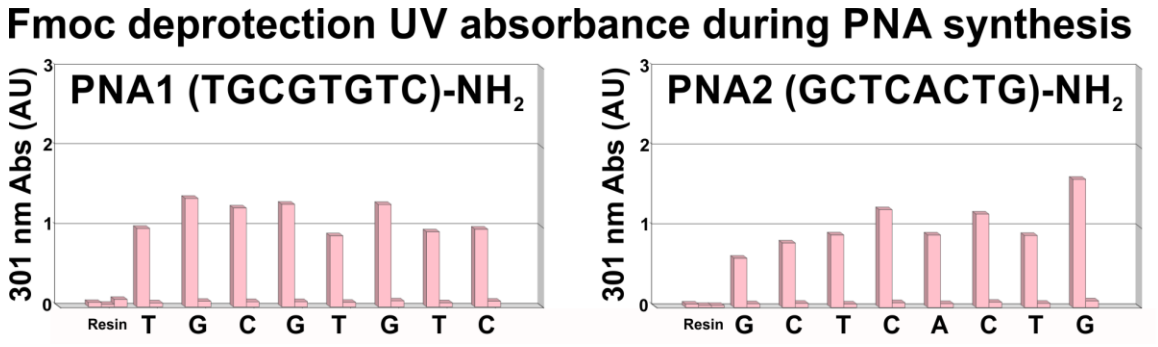


Figure 2-1. UV absorbance of Fmoc deprotection during PNA synthesis. In order to quantify the amount of Fmoc protecting group removed from the N terminal PNA base during each deprotection step, the UV absorbance at 301 nm of the waste line was measured using the CEM UV monitoring accessory. About 4–7% of the total absorbance for each PNA base was detected after the second step, underscoring the necessity for the second deprotection step.

Amino acid coupling

After verifying the PNA-Gly synthesis by RP-HPLC and MALDI, the remaining amino acids (Pro-Gly) were added using the synthesizer similar to the method reported by Fabani and coworkers[83] and summarized as follows. Deprotection was performed in 20% piperidine in DMF at 75 °C using microwave heating initially for 30 seconds and a second time for 3 minutes after washing with DMF. Coupling was performed once for 10 minutes at 75 °C using microwave heating in the ratio of monomer/activator/base of 1.0/1.0/4.0 in 4 fold excess to resin, where the activator was 0.2 M O-Benzotriazole-

N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU) in DMF and the base was 0.8 M DIEA in NMP. Because of the problems we found with PyBOP during PNA synthesis, we opted to use the more conventional peptide synthesis activator HBTU.

Fluorescent dye coupling

5(6)carboxyfluorescein (FAM) was coupled to PNA1-GPG construct with the synthesizer using a method similar to that reported by Kowalczyk and coworkers[85] summarized as follows. Deprotection was performed in 20% piperidine in DMF at 75 °C using microwave heating initially for 30 seconds and a second time for 3 minutes after washing with DMF. Coupling was performed once for 30 minutes at 75 °C using microwave heating in the ratio of monomer/activator/base of 1.0/1.0/4.0 in 4 fold excess to resin. The resin was washed with 20% piperidine twice with microwave heating to 75 °C for 3 minutes to remove polymerized dye as originally reported by Fischer and coworkers[86] using a method similar to that reported by Kowalczyk and coworkers.[85]

5(6)carboxytetramethylrhodamine (TMR) was coupled to PNA2-GPG similar to the method described by Kaihatsu and coworkers[68] as follows. Fmoc was deprotected by shaking with 20% piperidine in DMF 3 times for 5 minutes at room temperature. TMR was coupled in the ratio of monomer/activator/base of 1.0/1.0/4.0 in 4 fold excess to resin, where the activator was 0.2 M HBTU in DMF and the base was 0.8 M DIEA, by shaking at room temperature for 6–8 hours. The resin was washed with DMF followed by dichloromethane (DCM) and dried by lyophilization.

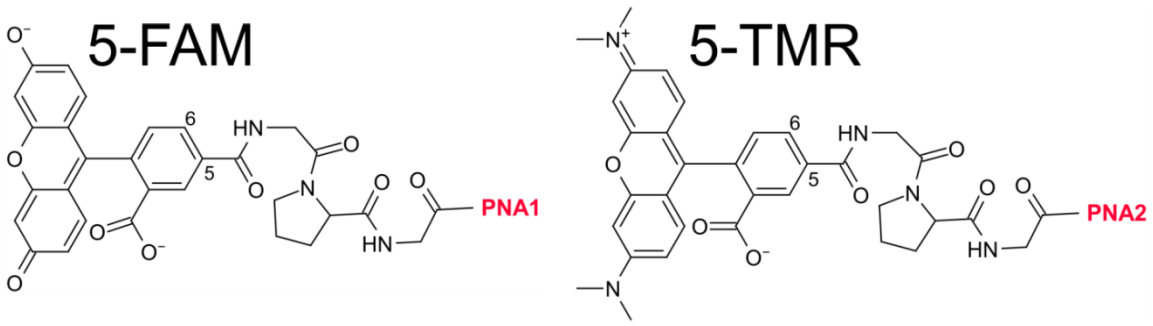


Figure 2-2. Chemical structures of FAM and TMR. Chemical structures of the dianion 5-carboxyfluorescein (left) and neutral zwitterion 5-tetramethylrhodamine (right) conjugated to the PNA-peptide via an amide linkage. The less expensive 5(6) dye isomer mixture was used in this study, although in some cases the dye isomers could be separated after conjugation to the PNA-peptide.

PNA-peptide acetylation

N terminally acetylated PNA-peptides were prepared as fluorescence controls by first deprotecting 3 times for 5 minutes at room temperature. First PNA1-GPG-NH_2 was acetylated in 0.5 M Acetic Anhydride, 0.125 M DIEA and 0.015 M Hydroxybenzotriazole (HOBt) by shaking for 3 hours at room temperature. After cleavage the RP-HPLC chromatogram showed several species which were identified by MALDI to be the product with from one up to four acetylations, and demonstrates the sensitivity of PNA to over-acetylation. Fortunately, the desired mono-acetylated product could be separated from the multi-acetylated impurities and collected by RP-HPLC. PNA2-GPG-NH_2 was acetylated using more mild conditions as published in a previous PNA capping protocol,[83] where the PNA-GPG-NH_2 was acetylated for 5 minutes at room temperature in DMF with 5% acetic anhydride and 5% lutidine. The resin was washed with DMF followed by DCM and dried by lyophilization. Only the mono-acetylated form of PNA2-GPG-Ac was observed.

Cleavage

All PNA-peptide products were cleaved from the resin by shaking for 2–3 hours in trifluoroacetic acid (TFA)/anisole/triisopropylsilane (TIS)/water in the ratios of 90.0/5.0/2.5/2.5%, precipitated in cold ether and lyophilized.

PNA-peptide purification and mass identification by MALDI

All PNA-peptide products were purified by RP-HPLC on a Waters system using an Agilent C18 Prep column (Polaris 5, C18-A 150 x 21.2mm). The purity was confirmed by RP-HPLC using an Agilent C18 Analytical column (Varian Polaris 5 C18-A 250 x 4.6mm) to be greater than 95% for all products except PNA2-GPG-TMR, which showed a purity of 90%. MALDI and RP-HPLC chromatograms are shown for all products in Figure 2-3 and Figure 2-4, respectively. The solvents used during RP-HPLC were: A) 99.9% water + 0.1% TFA, and B) 95% acetonitrile + 4.9% water + 0.1% TFA. The product masses were identified by MALDI mass spectrometry on a Voyager DE-STR MALDI-TOF (Applied Biosystems) in reflector mode using Sinapinic Acid as the matrix.[87]

The purified PNA-peptides were lyophilized and suspended in TAE.Mg buffer (40 mM Tris, 20 mM Acetate, 2 mM Ethylenediaminetetraacetic acid (EDTA), 12.5 mM Magnesium Acetate, pH 8.3). If the sample was not well dissolved it was heated and sonicated for 10–20 minutes, and if necessary, mixed with 5–30% acetonitrile (by volume). The extinction coefficients at the peak absorption for each PNA sequence were calculated online at http://biophysics.idtdna.com/UVSpectrum.html using Integrated DNA Technologies (IDT, Coralville, IA) published algorithm.[88] The calculated extinction

coefficients of PNA1 $\varepsilon_{258} = 69{,}932\ M^{-1}cm^{-1}$ and PNA2 $\varepsilon_{260} = 73{,}000\ M^{-1}cm^{-1}$ were used to measure the concentration of each dissolved PNA sequence. The extinction coefficients at 260 nm of FAM ($\varepsilon_{260} = 13{,}700\ M^{-1}cm^{-1}$) and TMR ($\varepsilon_{260} = 29{,}100\ M^{-1}cm^{-1}$) were used to calculate the PNA extinction coefficients for dye labeled samples and were assumed to be the same as those provided by the manufacturer (IDT) for ssDNA labeled with the same dyes.

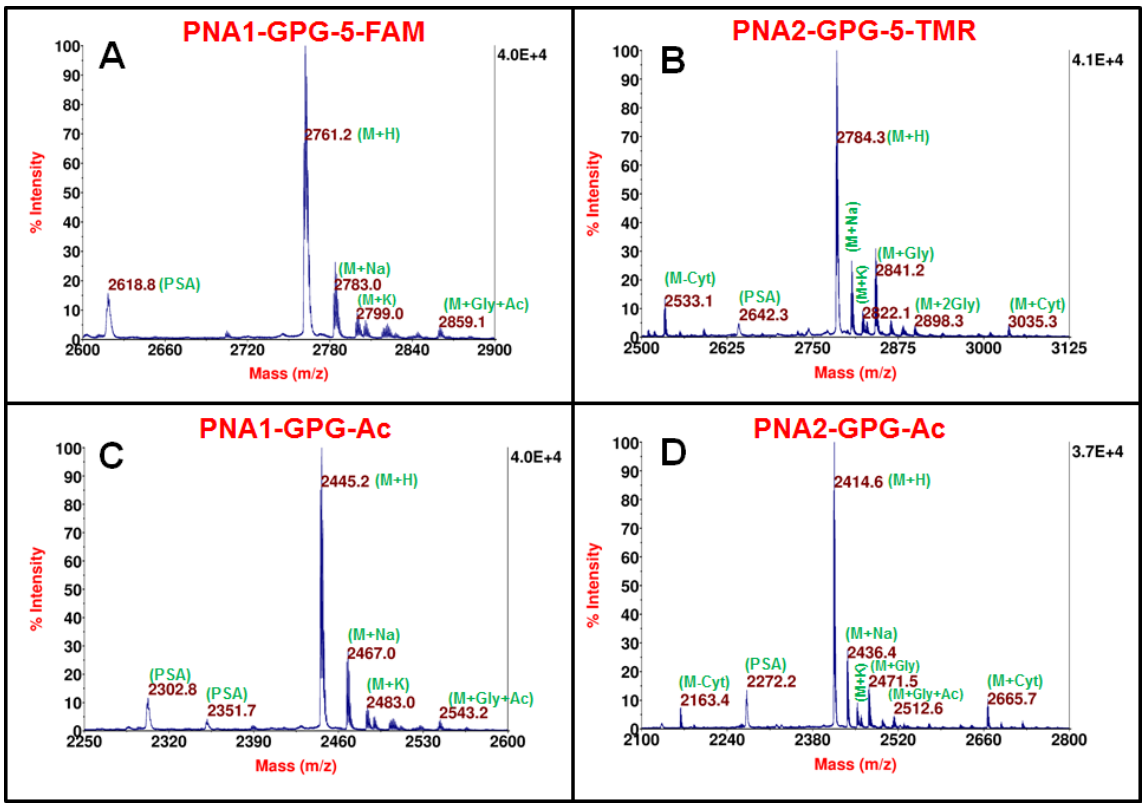


Figure 2-3. MALDI mass spectrum of the synthesized PNA-peptides. MALDI mass spectrum of PNA1-GPG-5-FAM [A], PNA2-GPG-5-TMR [B], PNA1-GPG-Ac (Acetylated) [C] and PNA2-GPG-Ac [D] after initial purification taken in reflector mode using Sinapinic Acid as the matrix. [A] The measured masses of 2761.2, 2784.3, 2445.2, 2414.6 Daltons (Da) agree within +0.2, +0.2, -1.8 and 0.6 Da of the expected mass of 2761.0, 2784.1, 2447.0, 2414.0 Da for PNA-peptides A, B, C and D, respectively. The spectra also show the corresponding sodium (M+Na) and potassium (M+K) adducts, as well as a minor impurity with a single glycine insertion (M+Gly) as well as a glycine insertion with an extra acetylation (M+Gly+Ac). PNA2-GPG-Ac and 5-TMR also show a minor impurity with a cytosine insertion (M+Cyt) and deletion (M-Cyt) as well as a dou-

ble glycine insertion (M+2Gly) for PNA2-GPG-5-TMR. The broad PSA peak is an artifact from post source fragmentation.

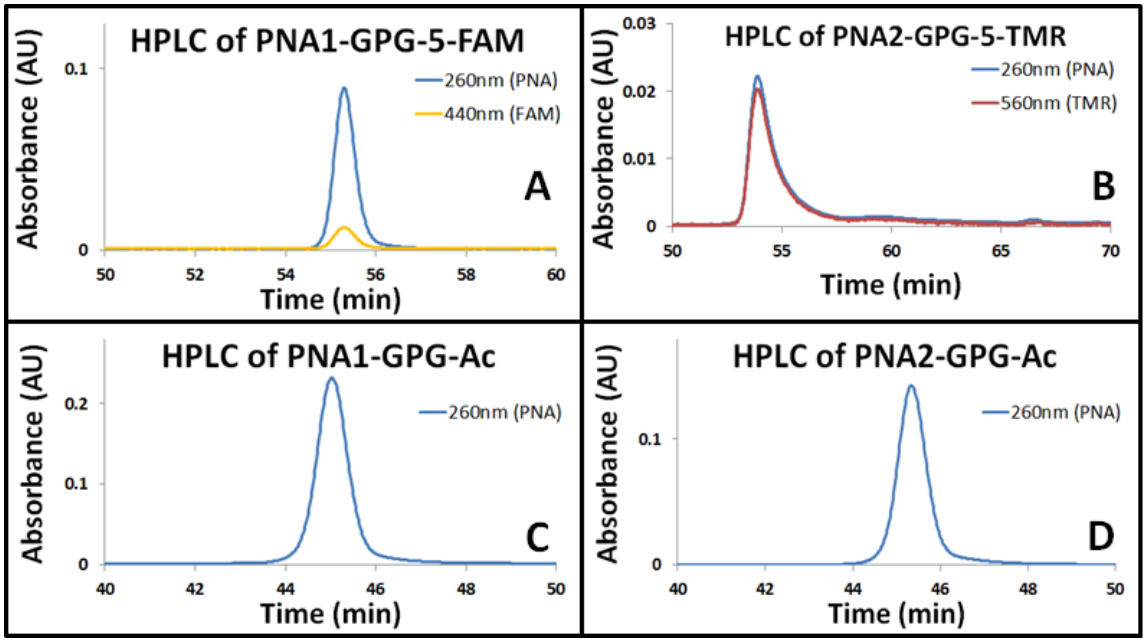


Figure 2-4. RP-HPLC chromatograms of the synthesized PNA-peptides. RP-HPLC chromatograms demonstrating the purity after initial purification of [A] PNA1-GPG-5-FAM (>95%), [B] PNA2-GPG-5-TMR (~90%), [C] PNA1-GPG-Ac (>95%) and [D] PNA2-GPG-Ac (>95%). The following stepwise gradients were used to mix solvent α (99.9% water, 0.1% TFA) and solvent β (95% acetonitrile, 4.9% water, 0.1% TFA) at a flow rate of 1ml/min: [A, B] 1%/min (0-10% β), 0.5%/min (10-20% β), 0.25%/min (20-35% β); [C, D] 1%/min (0-5% β), 0.5%/min (5-10% β), 0.25%/min (10-25% β). The nucleobase absorption at 260nm was monitored to detect the elution of the product. FAM and TMR absorption at 440nm and 560nm were monitored to confirm the presence of the dye label for [A] and [B], respectively. The absorption of FAM in [A] is substantially reduced at low pH[89] in its anionic (ε_{453} = 29,000 $M^{-1}cm^{-1}$) and neutral forms (ε_{434} = 11,000 $M^{-1}cm^{-1}$).

PNA maleimide functionalization

Sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (Sulfo-SMCC, Bioworld), an amine reactive cross-linker, was dissolved to 10 mM in 50 mM phosphate buffer at pH 7.2. PNA-Gly-NH_2 was dissolved in 0.1 M phosphate buffer at pH 7.2 containing 0.15 M NaCl and diluted with acetonitrile to 10–20% total by volume.

15x molar excess Sulfo-SMCC was mixed with the PNA for 1 hour at room temperature (final PNA concentration ~500 µM). The excess SMCC and unreacted PNA were removed using RP-HPLC, as shown in Figure 2-4, and then lyophilized. The matrix

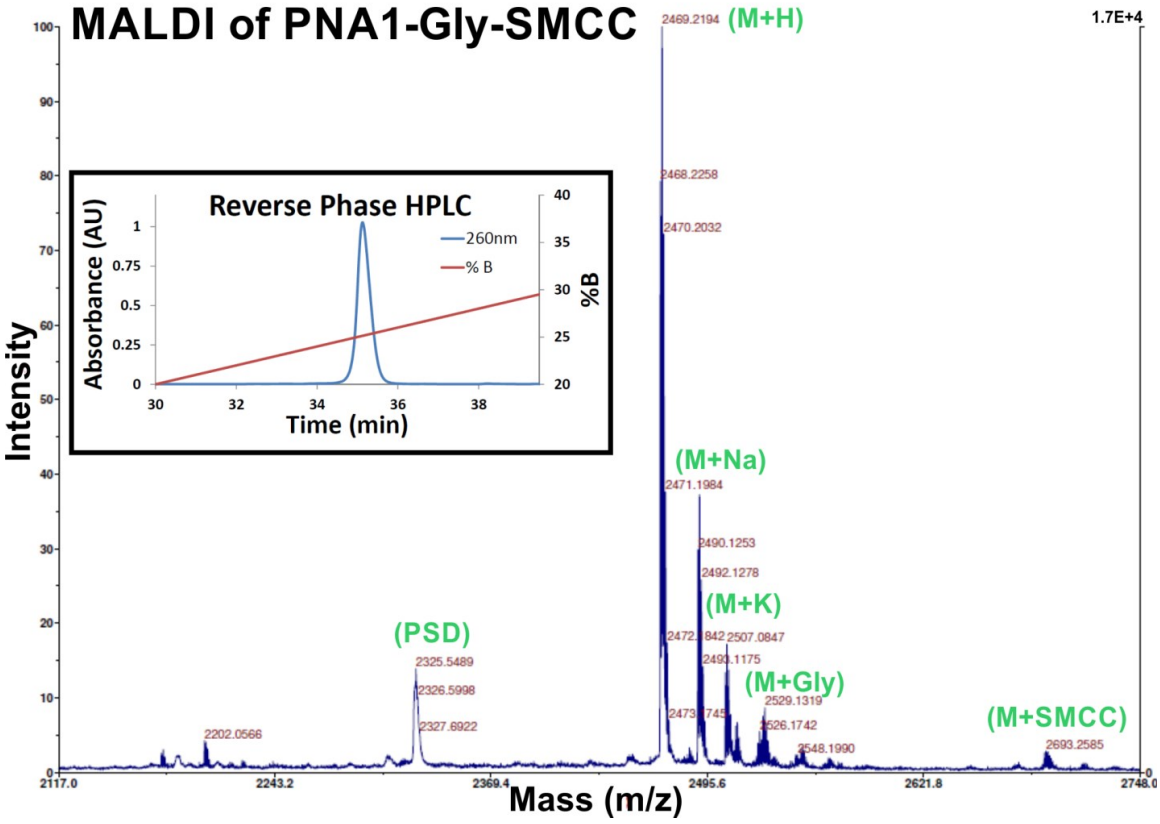


Figure 2-5. MALDI-MS spectrum and RP-HPLC of PNA-Gly-SMCC. MALDI-MS spectrum of PNA-Gly-SMCC after initial purification. The measurement was performed in reflector mode using sinapinic acid as the matrix. The measured mass of 2468.2 Daltons (Da) agrees within +0.2 Da of the expected mass of 2468.0. The spectra also show the corresponding sodium (M+Na) and potassium (M+K) adducts, as well as a minor impurity with a single glycine insertion (M+Gly) as well as a second SMCC attached (M+SMCC). The broad peak is an artifact from post source decay (PSD). The inset shows the RP-HPLC chromatogram: one main peak elutes from the column showing negligible impurities after the first purification step (purity ~100%). The following stepwise gradients were used to mix solvent α (99.9% water, 0.1% TFA) and solvent β (95% acetonitrile, 4.9% water, 0.1% TFA) at a flow rate of 1 ml/min: 1%/min (0–30% β). The nucleobase absorption at 260 nm was monitored to detect the elution of the product.

assisted laser desorption ionization mass spectrometry (MALDI-MS) spectrum is also shown in Figure 2-4.

γ-PNA functionalization and purification

γ-PNA Cy5 labeling

A scheme describing the Cy5 labeling of PNA using copper-free click chemistry is shown in Figure 5-3. γ-PNA1 and γ-PNA3 were purchased from PNA Bio (Panagene distributor) as a lyophilized powder and resuspended in nanopure water. The sequences of γ-PNA1 and γ-PNA3 are provided in Table 5-1. γ-PNA1 was heated to 60 ºC for 15 minutes to completely dissolve at ~500 μM. Cyanine 5 (Cy5) functionalized with the strained alkyne dibenzocyclooctyl (DBCO), referred to hereafter as Cy5-DBCO, was purchased from ClickChemistryTools (Scottsdale, AZ, USA) and dissolved in methanol to ~10 μM. 0.1 M sodium acetate buffer was prepared and the pH was adjusted to 4.0 using 0.1 M acetic acid. The buffer was mixed 1:3 (v/v) with methanol and the pH was adjusted once again to 4.0 by mixing acetic acid 1:10 (v/v) with buffer. The final buffer contained about 30 mM sodium acetate in 60% methanol with a pH of 4.0. In separate centrifuge tubes, 2x molar excess of Cy5-DBCO was added to γ-PNA1 and γ-PNA3 and mixed 1:8 (v/v) with the sodium acetate / methanol buffer (~50% methanol). The cocktail was reacted overnight at room temperature. In the morning, the pH was adjusted to 7.0 using sodium hydroxide and mixed 1:16 (v/v) with 0.5 M TCEP to reduce any γ-PNA disulfide dimers formed at the gamma position and incubated for 30 minutes.

γ-PNA Cy5 purification

The cocktail was mixed 2:1 with 80% glycerol and loaded on a 10% denaturing polyacrylamide gel electrophoresis (PAGE) gel (~10 μL / lane). The gel was run at 300 V for 973 Vhrs and visualized without removing it from the glass plates using a UVP Bio-Spectrum (Upland, CA, USA) with a 630 nm excitation filter (UVP, 38-0359-03) and a 695 nm emission filter (UVP, 38-0339-01) as shown in Figure 5-4. The gel was run for an additional 237 Vhrs (1210 total Vhrs) to allow the unreacted Cy5 to run off the gel so the product bands could be more easily visualized for excision. The desired product bands were cut and eluted with water in Costar® Spin-X® centrifuge tubes (Corning). The gel blocks were soaked from 1.25 hours to several days (115hrs), as described in Table 5-3, and spun down at 14k x g for 7-15 minutes to remove the eluted product from the gel blocks. The elution was performed 3–4 times, until the fluorescence of the gel blocks approached background levels. The progress of the elution was monitored by imaging the Cy5 fluorescence of the tubes for 5 seconds. The fluorescence intensity of a 2400 square pixel area of the gel block section of the brightest tube (i.e., containing the most product) was integrated and plotted as a function of the number of elutions. The data was fit using Origin version 9.1 with the following exponential function: $y(n) = y_0 + A * e^{-R_0 * n}$, where y_0 is the lower decay limit, A is the pre-exponential factor, R_0 is the rate of decay and n is the number of elutions. The fitting parameter values for γ-PNA1-Cy5 and γ-PNA3-Cy5 are shown in Table 5-4. The fraction of the product remaining in the gel blocks after each elution is plotted in Figure 5-4, listed in Table 5-3 and was calculated as follows: $Fract_{Gel\ Blocks}\ \%\ (n) = \frac{I_n - y_0}{A}$, where I_n is the integrated intensity of the gel blocks after the nth elution. The fraction of product eluted from the gel blocks after each

elution is listed in Table 5-3 and was calculated as one minus the fraction remaining in the gel blocks. The eluted fractions were pooled together and desalted by RP-HPLC on a Waters system using an Agilent C18 Prep column (Polaris 5, C18-A 150 x 21.2mm), lyophilized and reconstituted in water with 10-20% acetonitrile. The other bands with much lower fluorescence in the gel shown in Figure 5-4, were desalted using C18 spin columns (Pierce) per the manufacturer's instructions and found to be impurities by MALDI.

Cytochrome c preparation

A scheme outlining the preparation of unlabeled and labeled PNA-cytochrome c conjugates is shown in Figure 4-4.

Fluorescent labeling of cytochrome c with TMR

Fluorescent labeling of cytochrome c on amine residues was performed before the introduction of thiols on the surface. Equine heart cytochrome c (Sigma) was dissolved to about 400 μM in 0.1 M phosphate buffer at pH 7.2 containing 0.15 M NaCl and 1 mM Ethylenediaminetetraacetic acid (EDTA). Carboxytetramethylrhodamine (TMR) N-Hydroxysuccinimide ester (Pierce) was dissolved in DMF to 24 mM and then incubated in 3x molar excess with cytochrome c for 1 hour (~5% DMF (v/v)). Excess dye was removed using a NAP-5 column (GE Healthcare) equilibrated in the same buffer, and the sample was concentrated using 3 kDa MWCO centrifugal filters (EMD Millipore) to about 1 mM. The fluorescence spectrum is shown in Figure 2-19.

Thiolation of cytochrome c using SPDP

Our initial attempts to conjugate PNA to a naturally occurring surface cysteine on yeast cytochrome c were unsuccessful, and may reflect the poor accessibility of this residue.[90] We then resorted to using non-specific crosslinking techniques as previously reported[50] by first introducing thiols to surface amines of equine heart cytochrome c using the SPDP cross-linker. N-Succinimidyl 3-(2-pyridyldithio)-propionate (SPDP, Pierce), a heterobifunctional cross-linker, was dissolved to 80 mM in dimethyl sulfoxide (DMSO). For labeled constructs TMR labeled cytochrome c was concentrated to 1 mM and used directly. For unlabeled constructs, equine heart cytochrome c (Sigma) was dissolved to 5–10 mg/mL (400–800 μM) in 0.1 M phosphate buffer at pH 7.2 containing 0.15 M NaCl and 1 mM EDTA. 1.6x molar excess SPDP was mixed with either the TMR labeled or unlabeled cytochrome c at room temperature for 1 hour. Excess SPDP was removed using a NAP-5 column equilibrated in 0.1 M HEPES buffer at pH 7.2 containing 5 mM EDTA and which was previously degassed by vacuum. The number of thiols introduced per protein was quantified by the increase in absorbance at 343 nm from the release of pyridine-2-thione after reduction of SPDP by tris(2-carboxyethyl)phosphine (TCEP) (Figure 2-6). On average 1.1–1.4 and 2.6 thiols per protein were introduced when 1.6x and 2.3x molar excess of SPDP were used, respectively. This agrees with molecular weight distribution observed in the MALDI-MS spectrum (Figure 2-6).

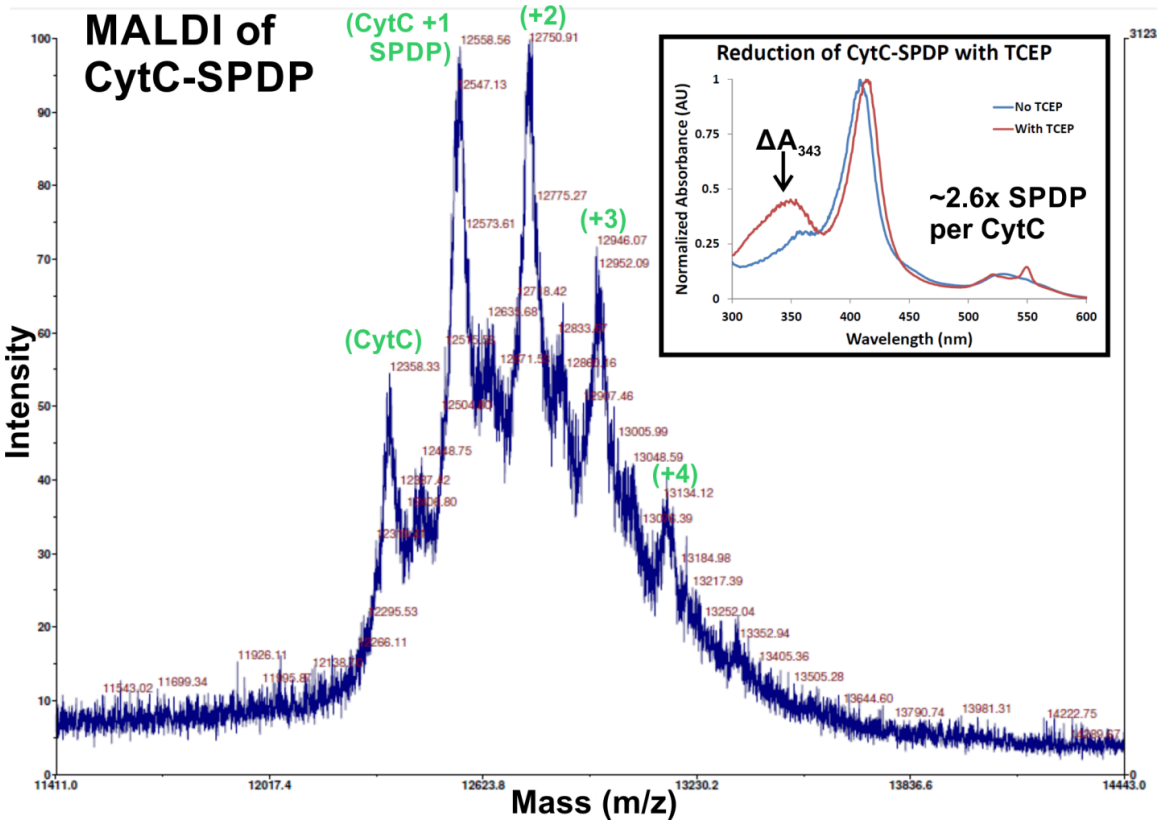


Figure 2-6. Quantification of SPDP thiolation of cytochrome c by MALDI-MS and UV-Vis. MALDI-MS spectrum of cytochrome c labeled with 2.3x molar excess of SPDP. The measurement was performed in linear mode using sinapinic acid as the matrix. The measured masses of 12358, 12558, 12751, 12946, 13134 Daltons (Da) agrees within 0, 0, -4, -9, and 20 Da of the expected masses for cytochrome c conjugated with zero (12358), one (12557), two (12756), three (12955) and four (13154) SPDP molecules (MW 199.3 Da). The inset shows the increase in absorbance at 343 nm when pyridine 2-thione is released after introducing TCEP to expose ~2.6 free surface thiols per cytochrome c.

Azurin expression and purification

The mutant version N42C of azurin (*Pseudomonas aeruginosa*) was expressed in *E. coli* and purified by using a slightly modified version of a previously published method.[91] Cells from *E. coli* strain JM109 were transformed using a pUC-derived plasmid (pGKN42C) containing the azurin gene preceded by a signal peptide for periplasmic translocation. After culturing, cells were harvested and resuspended in a solution of 20%

(w/v) sucrose in 30 mM Tris/HCl pH 8.0 containing 10 mM EDTA and 5 mM dithiothreitol (DTT) for 20 minutes at room temperature. Subsequently, the solution was centrifuged at 8000 rpm for 15 minutes and the supernatant was collected (sucrose fraction). The cells were resuspended in milliQ water containing 10 mM EDTA and 5 mM DTT at 4 °C, stirred for 20 minutes and centrifuged at 8000 rpm for 15 minutes. The supernatant was collected and added to the sucrose fraction, while the pellet was discarded. In *E. coli*, azurin is normally expressed in its apo-form, and the Zn containing form is isolated as a by-product of the expression.[92] A stepwise precipitation step was performed by lowering the pH of the solution to pH 4 by adding acetic acid. The precipitated proteins were removed by centrifugation (8000 rpm, 20 minutes). The resulting clarified solution was loaded on a home-packed CM sepharose Fast Flow (Amersham Biosciences) column (V_{tot} ~ 120 ml) equilibrated with 50 mM ammonium acetate buffer, in presence of 5 mM DTT. Elution was performed at 5 ml/min by using a pH gradient from pH 4 to pH 6.9 in 10 column volumes (CV). Fractions containing azurin were collected and, after buffer exchange, loaded onto a home-packed DEAE sepharose Fast Flow (Amersham Biosciences) column (V_{tot} ~ 80 ml) equilibrated with 5 mM Tris/HCl buffer, pH 8.5, containing 5 mM DTT. The protein was eluted using a salt gradient from 0 to 50 mM of NaCl in 10 CV. Two main peaks eluted from the column, one being the apo, and the other the Zn form of the protein, respectively. The Zn N42C azurin fractions were pooled together, concentrated and the buffer was exchanged to HEPES 20 mM, pH 7, by Amicon ultrafiltration. The protein was subsequently flash frozen in liquid nitrogen and stored at –80 °C.

All the chromatographic steps were performed on an Äkta Purifier system (GE Healthcare). The purification process was monitored by checking the purity of the protein

after each chromatographic step on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and by means of UV/Vis spectroscopy (Cary 50 spectrophotometer, Varian Inc., Agilent Technologies, USA). Isoelectric focusing gel electrophoresis was used to discriminate between apo and Zn containing azurin fractions after the DEAE column separation. The final product appeared on an SDS-PAGE gel as a single band with apparent mass of ~14 kDa.

PNA-protein conjugation, purification and characterization

A scheme outlining the preparation of unlabeled and labeled PNA-protein conjugates is shown in Figure 4-4.

PNA-cytochrome c conjugation

PNA-SMCC was dissolved to ~750 μM in 0.1 M HEPES buffer at pH 7.2, degassed by vacuum, containing 5 mM EDTA and diluted with acetonitrile to 20% (v/v) to help dissolve the PNA. Tris(2-carboxyethyl)phosphine (TCEP) was dissolved to 0.5 M and titrated to pH 7. Thiolated cytochrome c (CytC-SPDP) was reduced with a 20x molar excess TCEP for 30–60 minutes to release pyridine-2-thione and expose the introduced surface thiols. 1.0x to 1.8x molar excess of reduced CytC-SPDP was mixed with roughly equal volume PNA-SMCC to maintain 10% acetonitrile (v/v) in the cocktail to keep the PNA soluble, and was incubated at room temperature or 37 ^{0}C overnight. Figure 2-7 shows the MALDI-MS spectrum, which indicates the presence of the PNA-CytC conjugate. Alternate conjugation protocols used 10% methanol in place of the acetonitrile or just buffer (i.e., no organic solvents), which did not have any noticeable impact on the

conjugation efficiency at low concentration (<100 μM) , as shown in Figure 2-8. However, 10% acetonitrile did allow the PNA to remain soluble at much higher concentrations (>500 μM), which, in combination with high CytC concentrations (>500 μM), did help to improve the conjugation yield to around ~25% in subsequent conjugation preparations.

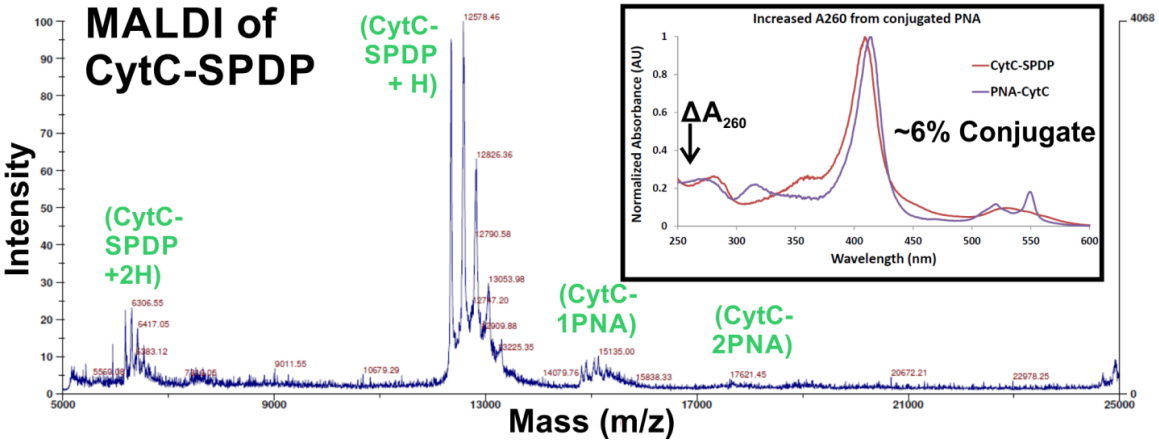


Figure 2-7. MALDI-MS spectrum of SPDP thiolated cytochrome c conjugated to PNA-SMCC. MALDI-MS spectrum of SPDP thiolated cytochrome c conjugated to PNA-SMCC. The measurement was performed in linear mode using sinapinic acid as the matrix. Unreacted SPDP thiolated cytochrome c is detected as a series of peaks centered at 12578 Da and the corresponding double ionization peaks centered at 6307 Da. The peaks centered at 15135 Da correspond to the thiolated cytochrome c (12578 Da) conjugated to a single PNA-SMCC (2468 Da) with an unreacted surface thiol (89 Da). A weak peak is also detected for a thiolated cytochrome c conjugated to two PNAs (17621 Da). The inset shows the increase in absorbance at 260 nm from native cytochrome c (CytC) and the cocktail mixture of CytC and PNA-CytC after removing unconjugated PNA, resulting in ~6% conjugation yield. For comparison, the spectrum of PNA-CytC is shown after purification by toehold mediated strand exchange.

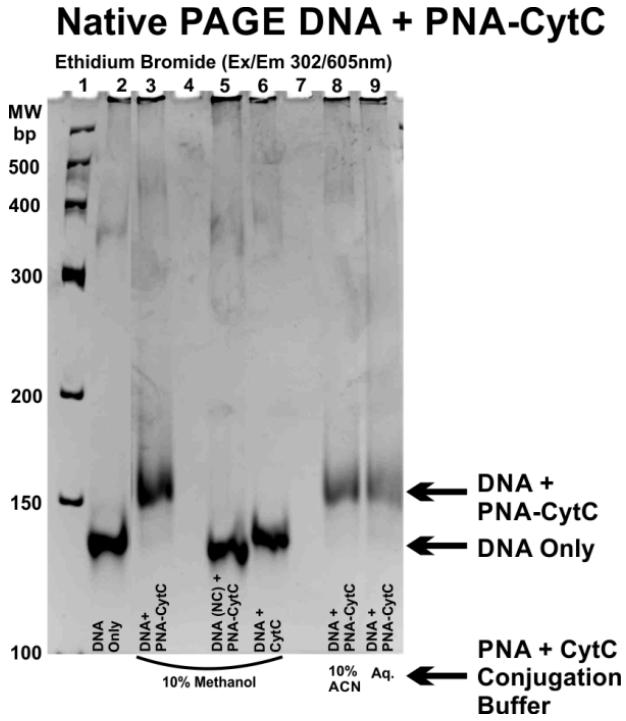


Figure 2-8. Native PAGE (7%) at 4 °C of the DNA nanocage incubated with PNA-CytC. PNA-SMCC was conjugated with thiolated cytochrome c in a buffer containing 10% methanol (lanes 3, 5, 6), 10% (lane 8) acetonitrile (ACN) or just aqueous buffer (lane 9). After removing unreacted PNA-SMCC by SE-FPLC, was incubated in 10x molar excess with the DNA nanocage (lanes 3, 8, 9) or the DNA nanocage with a binding sequence non-complementary (NC) to the PNA sequence (lane 5). The DNA nanocage was also incubated with 10x excess of free cytochrome c (lane 6). Only DNA nanocages with the correct PNA binding sequence and incubated with the PNA-CytC conjugate show the slower moving band containing cytochrome c. There was no noticeable difference in conjugation yield when the aqueous buffer contained 10% methanol or 10% acetonitrile by volume.

Aminoethylmaleimide (AEM) was dissolved to 50 mM in 0.1 M HEPES buffer at pH 7.2 containing 5 mM EDTA. 25x molar excess TCEP was added to the PNA-CytC cocktail (~250 μM) for 30 minutes to ensure all unreacted thiols were in the reduced form. 100x molar excess AEM was mixed with the PNA-CytC cocktail for 30 minutes to quench any free surface thiols and to restore the positive charge of the native surface lysines originally displaced by SPDP. Unreacted AEM and PNA were removed by Size

Exclusion Fast Protein Liquid Chromatography (SE-FPLC) with a Superdex 75 column (column volume 24 mL) using an Amersham Pharmacia (GE Healthcare) ÄKTA Purifier with a flow rate of 0.2 mL/min in 1x TAE.Mg buffer or, 0.1 M HEPES buffer pH 7.0 buffer containing 0.15 M NaCl, both or 10 mM sodium phosphate buffer pH 7.0 buffer containing 12 mM magnesium acetate and 0.15 M NaCl; each was degassed to remove oxygen. The low flow rate (0.2 mL/min) and moderate amount of salt (0.15 M NaCl) significantly improved separation of the unreacted PNA from the combined peak of the protein and PNA-protein. After subsequent toehold mediated ssDNA strand purification as described in the PNA-protein purification by toehold-mediated DNA strand displacement section, the purity of the PNA-CytC conjugates were analyzed by 15% SDS-PAGE with a 4% stacking layer with anode buffer (0.1 M Tris pH 8.8) and cathode buffer (0.1 M Tris, 0.1 M Tricine, 0.1% SDS, pH ~8.2). The gel was run at constant 70 mA for about 1 hour. The gel was subsequently silver stained as described in the PNA-protein assembly section. The concentration of PNA-cytochrome c conjugate was determined by deconvoluting the heme absorbance from the PNA absorbance at 260 nm using the calculated extinction coefficient for the PNA described in the PNA synthesis methods and the published ε_{410} of 106,000 $M^{-1}cm^{-1}$ for cytochrome c from the literature[93] along with the measured absorbance scan of the free cytochrome c.

A NAP-5 column was unable to adequately remove the unreacted PNA from the PNA-CytC, therefore a SE-FPLC chromatographic step was used to remove the unreacted AEM and PNA. Although SE-FPLC was not able to separate the PNA-CytC conjugate from the free protein, hybridization of the PNA-CytC conjugate to the DNA nanocage (MW ~95 kDa) can be easily separated from the free protein (MW 12.5 kDa) using SE-

FPLC. When the final product, which contained a mixture of free cytochrome c and PNA-CytC conjugate, was incubated with the DNA nanocage, PAGE in Figure 4-8 did not show any bands with the characteristic gel shift observed upon binding PNA alone, indicating that all unconjugated PNA was successfully removed from the sample.

PNA-azurin conjugation, purification and characterization

The N42C azurin mutant protein, prepared as described above, was exchanged into 0.1 M phosphate buffer pH 7.2 containing 0.15 M NaCl, concentrated to about 1.2 mM using 3 kDa MWCO centrifugal filters and reduced with a 10x molar excess of TCEP for 2–3 hours to reduce any intermolecular disulfide dimers. PNA-SMCC was dissolved in 0.1 M phosphate buffer at pH 7.2 containing 0.15 M NaCl. A 2.7x molar ratio of PNA-SMCC was added to the azurin and diluted with additional buffer to a final PNA concentration of about 100 μM and incubated at 37 ^{0}C overnight and analyzed by MALDI-MS, as shown in Figure 2-9. Unreacted PNA was removed using SE-FPLC with a HiLoad Prep grade Superdex 200 16/60 size exclusion column (GE Healthcare, 120 mL column volume) on an Amersham Pharmacia (GE Healthcare) ÄKTA Purifier controlled by Unicorn 5.11 software with a flow rate of 1 mL/min. To assess the purity of the PNA-azurin conjugate, NuPAGE 4–12% gradient SDS–PAGE gels (Invitrogen) were run at 180 V for 50 minutes and stained with Coomassie blue for 1 hour and subsequently destained overnight in water. The resulting gel image is shown in Figure 2-9. The ratio of PNA-azurin conjugate to unreacted azurin was determined to be 38% by deconvoluting the PNA and azurin absorbance at 260, 280 and 292 nm, which is in agreement with the SDS-PAGE band intensities. The PNA extinction coefficients were determined using the absorption

spectra of PNA alone and the extinction coefficient at 260 nm of 69,932 $M^{-1}cm^{-1}$ calculated in the PNA synthesis methods section. The azurin extinction coefficients were determined using the absorption spectra of azurin alone and the published extinction coefficient of 9,800 $M^{-1}cm^{-1}$ at 280 nm.[94] The extinction coefficient of PNA-azurin was ε_{260} = 76,792 $M^{-1}cm^{-1}$, and the mixture of PNA-azurin with unreacted azurin was ε_{260} = 87,974 $M^{-1}cm^{-1}$ which was used to determine the concentration of the PNA-azurin.

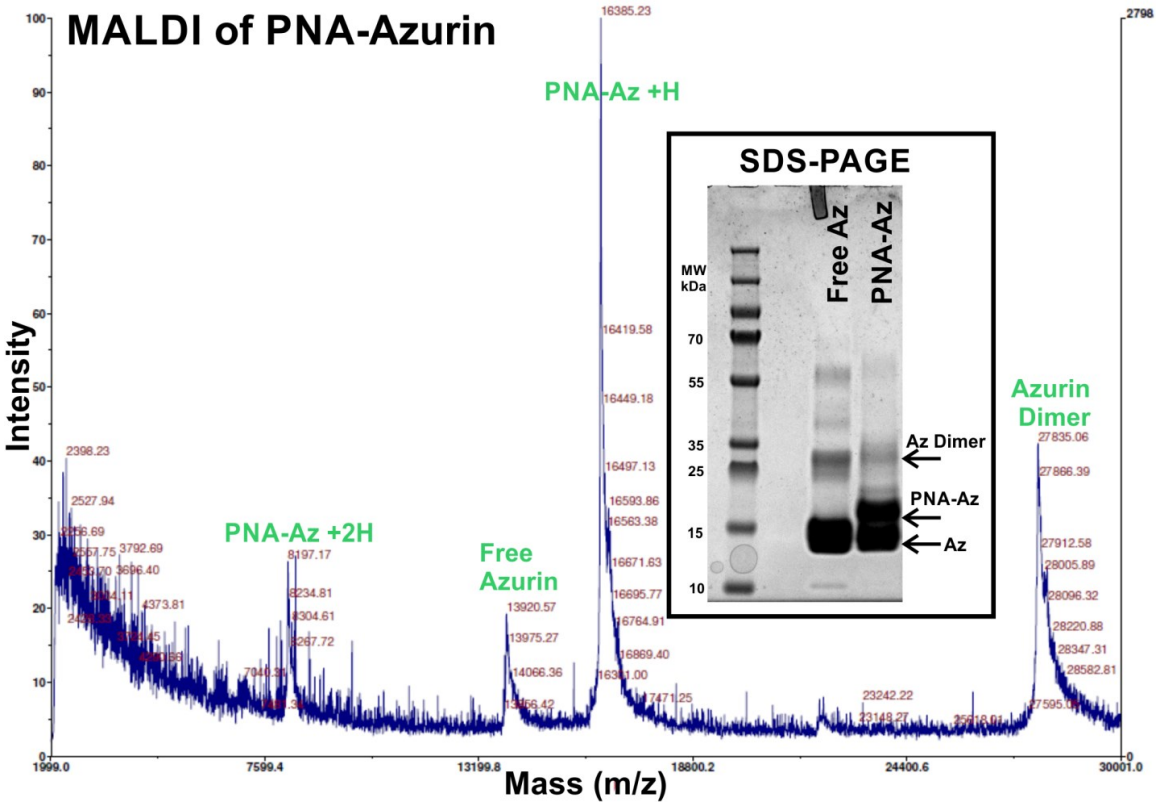


Figure 2-9. MALDI-MS spectrum and SDS-PAGE of azurin conjugated to PNA-SMCC. MALDI-MS spectrum of azurin conjugated to PNA-SMCC. The measurement was performed in linear mode using sinapinic acid as the matrix. The measured masses of 16385, 8197, 13920 and 27835 Dalton (Da) agree with the expected masses of PNA-azurin (M+H = 16388, M+2H = 8197), unreacted azurin (13.9 kDa) and an azurin dimer (27.8 kDa). The inset shows a 4–12% gradient SDS-PAGE of the mixture of free azurin and PNA-azurin conjugate.

During SE-FPLC, the PNA-azurin conjugate was not well separated from the unreacted azurin, but was separated from the unreacted PNA. When the PNA-azurin sample was incubated with the DNA nanocage, a faint band was visible on the gel with the characteristic shift of free PNA as shown in Figure 4-8, indicating a small amount of free PNA was not removed during the SE-FPLC run.

TMR labeling of PNA-azurin

After removal of unreacted PNA, some surface amines on the PNA-azurin conjugate were dye labeled. Carboxytetramethylrhodamine (TMR) N-Hydroxysuccinimide ester (Pierce) was dissolved in DMF to 10 mM and then incubated in 5x molar excess with PNA-Az in 0.1 M phosphate buffer pH 7.2 containing 0.15 M NaCl and 1 mM EDTA for 1 hour (~1.5% DMF (v/v)). Unreacted dye was removed using a NAP-5 column equilibrated in the same buffer and the labeled protein was subsequently concentrated to about 1 mM. The fluorescence spectrum is shown in Figure 2-19.

ssDNA toehold mediated PNA and PNA-protein purification

A scheme describing toehold mediated PNA and PNA-protein purification is shown in Figure 4-5.

Toehold DNA design

Typically PNA sequences have a greater binding affinity and thermal stability than their corresponding DNA sequences.[95] This property of PNA has been used to invade dsDNA in nanomechanical DNA devices.[96] To ensure the complementary toehold

DNA sequence had a significant thermodynamic driving force to displace the PNA sequence, an 8nt toehold domain was chosen that provided an increase in dissociation temperature of 27 ºC of the cDNA-toehold duplex over the PNA-toehold duplex. The 8nt length also should provide the most rapid strand displacement kinetics.[97]

PNA-FAM purification

Increasing molar excess (0.5x to 5x) of FAM labeled 16nt toehold DNA was incubated with TMR labeled PNA at 37 ^{0}C for 5 min and then cooled to 4 ^{0}C for 5 min. Additional aliquots of PNA-TMR were incubated at 37 ^{0}C with 2x excess of FAM labeled 16nt toehold DNA and then subsequently incubated with increasing amounts of fully complementary DNA (1x to 20x cDNA) based on the original amount of PNA-TMR. All aliquots were analyzed by native PAGE (7%) as shown in Figure 2-10.

PNA purification by toehold mediated DNA strand displacement was analyzed by native PAGE (7%) as shown in Figure 2-10. The ethidium bromide stain on the left shows the slowest moving DNA-FAM + PNA-TMR duplex band (~32 base pair), which has maximum intensity after just stoichiometric amounts of DNA-FAM is added and disappears after just after 2x molar excess cDNA is added. Free DNA-FAM begins to accumulate after 1x molar excess is added with respect to PNA-TMR. The middle band (16 base pair) shows the DNA-FAM + cDNA duplex form as the PNA-TMR is displaced. The TMR fluorescence image on the top right, shows the unbound and displaced PNA-TMR band migrates quite slowly due to the partial negative charge on the TMR dye. Some PNA-TMR does not bind to the DNA-FAM strand, possibly because it has an incompatible isomeric form, while a small amount dissociates during the gel run. The FAM

fluorescence image on the bottom right shows the presence of excess DNA-FAM after the PNA-TMR is saturated, as well as the duplex formed with DNA-FAM + cDNA. The two right most lanes show that PNA-TMR can also be liberated from DNA-FAM by incubating with cDNA at room temperature.

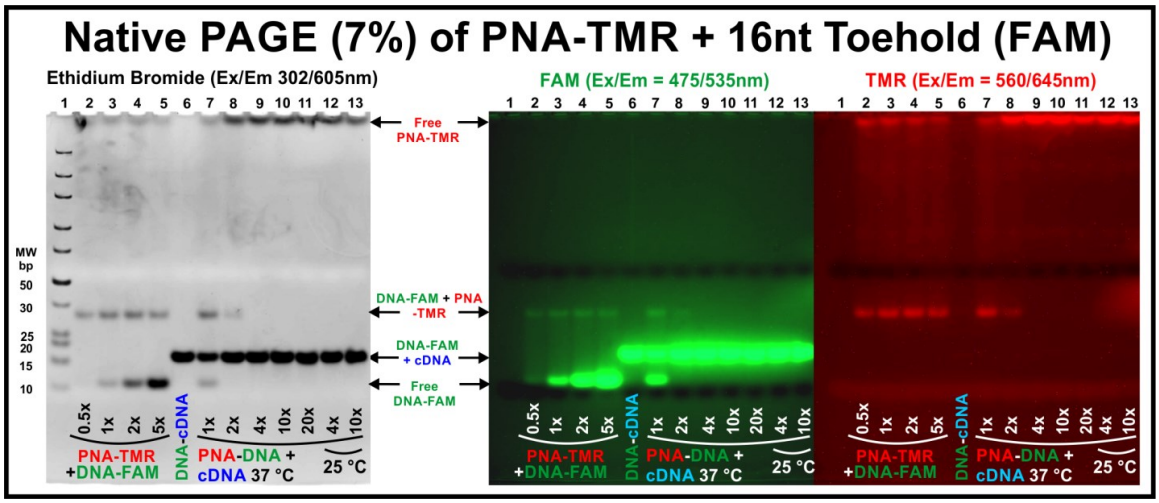


Figure 2-10. Native PAGE (7%) at 4 ^{0}C of toehold mediated ssDNA purification of PNA. Ethidium bromide stain [Left], FAM fluorescence [Center] and TMR fluorescence [Right]. Lanes 2–5: 2x PNA-TMR incubated with increasing amounts (0.5x to 5x) of FAM labeled 16nt toehold DNA (DNA-FAM) at 37 ^{0}C. Lane 6: 2x DNA-FAM incubated with 2x fully complementary DNA (cDNA) at 37 ^{0}C. Lanes 7–13: 2x PNA-TMR was incubated with 2x DNA-FAM at 37 ^{0}C cooled to 4 ^{0}C and then incubated with increasing amounts (1x to 20x) of cDNA at 37 ^{0}C (Lanes 7–11) or 25 ^{0}C (Lanes 12–13) to displace the PNA-TMR. Note: While the ethidium bromide image shows the ethidium bromide fluorescence of the DNA ladder in lane 1, the signal from the PNA-TMR and DNA-FAM bands are due to the native fluorescence emission of TMR and FAM at 605 nm due to direct excitation at 302 nm. The unlabeled cDNA band is not visualized by ethidium bromide, likely due to its short length.

2x molar excess 16nt toehold DNA was incubated at 37 ^{0}C for 15 min with TMR labeled PNA (PNA-TMR), cooled to 4 ^{0}C for several hours, and then purified using IEX-FPLC with the column cooled on ice, as shown in Figure 2-11. All products were retained on the column. The DNA-PNA-TMR duplex elutes prior to the excess ssDNA peak with good separation, as shown by the TMR absorption in the collected fraction

(Figure 2-11, inset). Some PNA-TMR also elutes with the ssDNA peak. Both PNA-TMR peaks can be recovered by incubating with the fully complementary cDNA strand and subsequent IEX-FPLC purification.

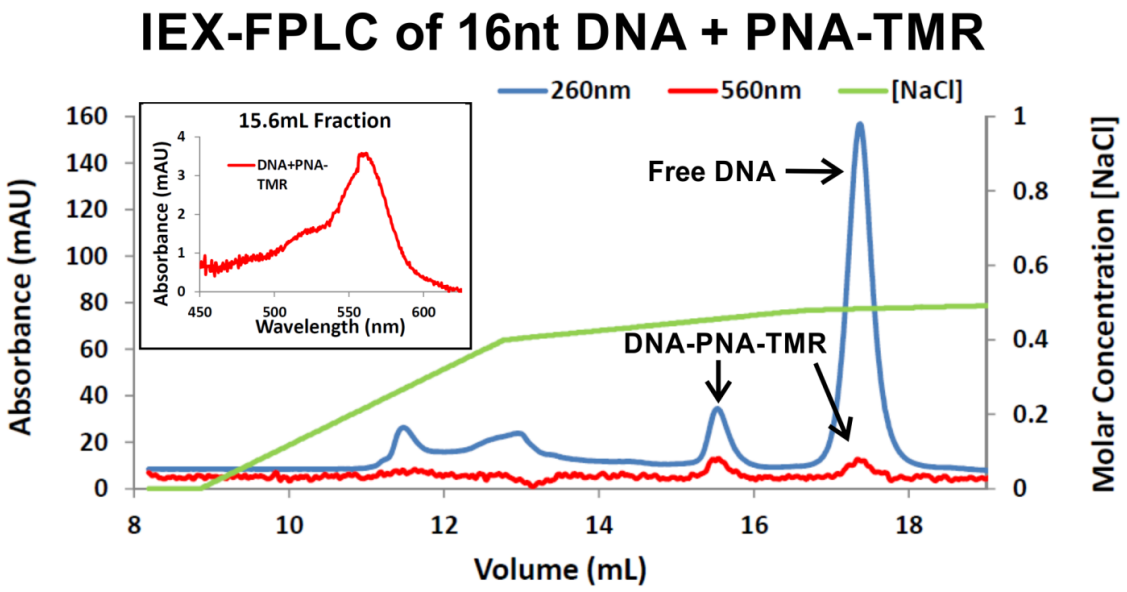


Figure 2-11. IEX-FPLC of toehold mediated ssDNA purification of PNA-TMR. IEX-FPLC chromatogram of PNA-TMR incubated with 2x excess 16nt toehold DNA with the column on ice. The following stepwise gradient was used to mix solvent α (TAE.Mg pH 9.0) and solvent β (TAE.Mg pH 9.0 + 2 M NaCl): 5%/min (0–20% β), 1%/min (20–24% β), 0.25%/min (24–28% β). The nucleobase absorption at 260 nm and the TMR absorption at 560 nm were monitored. The inset shows the TMR absorption of the first collected DNA-PNA-TMR fraction at 15.6 mL.

PNA-CytC purification

A 16nt or 66nt 'toehold' ssDNA strand was denatured by heating to 60 ^{0}C for 5 minutes. 2x molar excess of either toehold ssDNA was added to the crude mixture of PNA-CytC conjugate and free cytochrome c and incubated at room temperature or 37 ^{0}C for at least 20 minutes and then cooled on ice for at least 20 minutes. The DNA-PNA-protein complexes were purified from the free protein using SE-FPLC with a Superdex 75 column (GE Healthcare), with the column and buffers chilled on ice, as shown in Fig-

ure 4. The fractions containing the DNA-PNA-protein complexes were incubated with 2x molar excess of a ssDNA fully complementary to the toehold (cDNA) at room temperature for at least 15 minutes and then cooled on ice for at least 5 minutes. The cDNA and toehold DNA were removed from the PNA-protein using ion exchange FPLC (IEX-FPLC) using the gradient described in Figure 4-7. The sequences of the strands used in this method are described in the Figure 4-6.

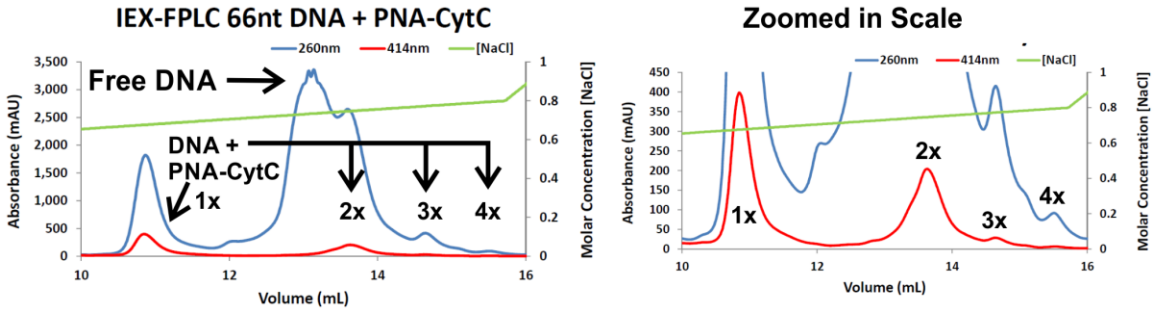


Figure 2-12. IEX-FPLC of toehold mediated ssDNA purification of PNA-CytC. IEX-FPLC chromatogram of PNA-CytC incubated with 2x excess 66nt toehold DNA. The 1x peak indicates the elution of CytC labeled with a single PNA and hybridized to a single toehold DNA strand. 2x, 3x and 4x indicate 2, 3 or 4 DNA-PNAs per protein, which are more visible on the right chromatogram with a zoomed in scale. The following stepwise gradient was used to mix solvent α (TAE.Mg pH 9.0) and solvent β (TAE.Mg pH 9.0 + 2 M NaCl): 2.5%/min (15–40% β), 15%/min (40–100% β. The nucleobase absorption at 260 nm and the CytC Soret absorption at 414 nm were monitored.

Although we initially found that PNA-CytC when hybridized to the 16nt toehold strand would not bind to the IEX-FPLC anion exchange column, we later found that PNA-CytC when hybridized to the larger and more stable 66nt toehold strand did stick to the anion exchange column, as shown in Figure 2-12. The column was also kept at room temperature, indicating the DNA-PNA-CytC complex is stable on the anion exchange column. The chromatogram shows several peaks with both significant A_{414} and A_{260}, indicating the presence of PNA-CytC hybridized to the 66nt toehold DNA strand in each

peak. The first peak with significant A_{414} has the shortest retention time and the highest A_{414}/A_{260} ratio, indicating the elution of cytochrome c labeled with a single PNA strand and hybridized to a single 66nt toehold strand. The second peak, which elutes just after the unhybridized toehold DNA strands (i.e., free DNA), has a longer retention time and a lower A_{414}/A_{260} ratio, indicating the elution of cytochrome c labeled with two PNA strands and hybridized to two toehold DNA strands. Similarly two other peaks are observed with longer retention times and lower A_{414}/A_{260} ratios, indicating a small amount of cytochrome c has three or four PNAs attached with their corresponding toehold DNA strands. IEX-FPLC provides much better separation of each PNA-CytC species compared to SE-FPLC, as shown in Figure 4-7, and does not require cooling the column. Thus IEX-FPLC is a superior separation technique and is more convenient than purification by means of SE-FPLC.

DNA nanocage preparation

ssDNA purification

Single stranded DNA of the sequences defined in Figure 3-2 were purchased from IDT (Integrated DNA Technologies) as a lyophilized powder, except for dye labeled strands which were purchased HPLC purified. IDT introduced FAM to strand S2 during the synthesis using a FAM modified thymine base and introduced TMR after synthesis by coupling a N-Hydroxysuccinimide ester functionalized TMR to an amine modified thymine base. Each strand was suspended in nanopure water to a concentration of 0.5 OD/μL and purified by denaturing PAGE using an SE 600 Ruby dual cooled standard vertical unit powered by a Fisher FB1000 supply in a buffer of 1xTBE (89 mM Tris, 89

mM Boric Acid, 2 mM EDTA, pH 8.5). The gels were heated to 35 °C using a Fisher Scientific Isotemp 3006S temperature controlled water bath circulating through the SE 600 Ruby lower buffer chamber. A 9% gel (8.3 M urea, 9% 19:1 Acrylamide:Bisacrylamide polymerized with ammonium persulfate (APS) and catalyzed by Tetramethylethylenediamine (TEMED) in 1xTBE) was used for 63nt and 66nt strands and an 8% gel was used for 55nt strands. The bands of the desired sequence were identified by staining with a 1x Ethidium Bromide (EtBr) solution (0.5 μg/mL) for 5–10 minutes, destaining in distilled water for 5–10 minutes and visualized briefly on a UV Transilluminator with 302 nm light. The bands were cut and shaken overnight in elution buffer (500 mM ammonium acetate, 10 mM magnesium acetate, and 2 mM EDTA) at room temperature. The tracking dyes and EtBr stain were removed by washing with butanol followed by cold ethanol precipitation and centrifugation. The DNA pellet was dried with a vacuum centrifuge and suspended in nanopure water. The concentration was determined using the measured absorbance at 260 nm and the extinction coefficients provided by the manufacturer.

DNA nanocage assembly

Four 63nt DNA strands were designed with mutually complementary sequences such that a DNA tetrahedron forms through hydrogen bonding between all base pairs, as previously reported.[65] A temperature gradient was used to allow the most energetically favorable state to form before lower temperatures are reached that can stabilize less favorable conformations. Using stoichiometric strand ratios and low concentrations, we ob-

served high assembly yields (~90%) with only minor aggregation as detected by Size Exclusion Fast Protein Liquid Chromatography (SE-FPLC).

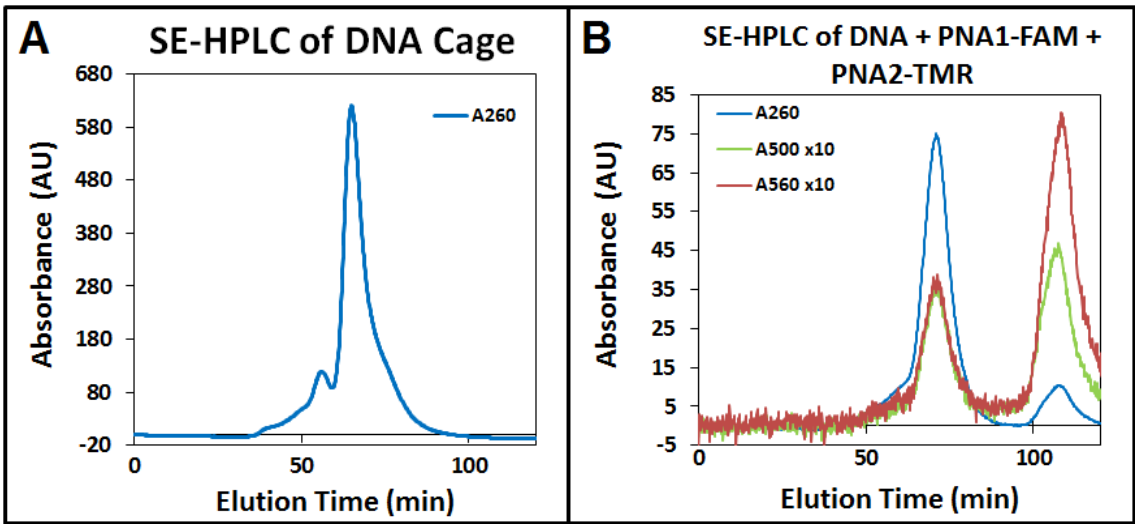


Figure 2-13. SE-FPLC chromatograms of the DNA nanocage with and without PNA-peptides. SE-FPLC chromatotrograms of [A] the DNA nanocage alone at room temperature and [B] the DNA nanocage with both dye labeled PNA-peptides (DNA + PNA1-FAM + PNA2-TMR) added in 5x molar excess at 4 °C. The absorbance of DNA and PNA at 260 nm were recorded for both chromatograms and for the [B] chromatogram the absorbance of FAM (500 nm) and TMR (560 nm) were also recorded. A_{500} and A_{560} are shown 10x larger than the actual recorded values. Both chromatograms were recorded at a flow rate of 1 mL/min. [A] Shows small amounts of dimers and higher order oligomer aggregates, followed by the main peak (yield ~90%) containing the monomeric DNA nanocage. [B] The first peak in the chromatogram represents the DNA-PNA-peptide complex and shows strong dye absorption from the bound labeled PNA-peptides. The second peak consists of the excess dye labeled PNA-peptides removed during purification.

Unless otherwise noted, 2.25 nmol of each DNA strand was added to 45 mL (50 nM) of TAE.Mg buffer, in 50 mL centrifuge tubes. The tubes were heated to 90 °C and then cooled on ice. The DNA nanocage samples were concentrated with 15 mL Amicon Ultra 30 kDa MWCO centrifugal filters (EMD Millipore) and purified by SE-FPLC with an Amersham Biosciences (GE Healthcare) HiLoad Prep grade Superdex 200 column using an Amersham Pharmacia (GE Healthcare) ÄKTA Purifier controlled by Unicorn

5.11 software with a flow rate of 1 mL/min. A representative SE-FPLC chromatogram is shown in Figure 2-9. The chromatogram shows a small peak of dimers and higher order oligomer aggregates, followed by the main peak (yield ~90%) containing the monomeric DNA nanocage. Alternatively, the annealed DNA nanocages were concentrated to 0.25–10 mL and purified by IEX-FPLC using a Mono-Q 5/50 GL (1 mL column volume) anion exchange column (GE Healthcare) on an Amersham Pharmacia (GE Healthcare) ÄKTA Purifier controlled by Unicorn 5.11 software with a flow rate of 1 mL/min. The following stepwise gradient was used to mix solvent α (TAE.Mg pH 9.0) and solvent β (TAE.Mg pH 9.0 + 2 M NaCl): 3.75%/min (20–35% β), 0.25%/min (35–40% β), 1%/min (40–50% β). The DNA nanocages were concentrated again to 5–15 μM. A representative chromatogram is shown in Figure 2-14A.

The IEX-FPLC chromatogram of the annealed DNA nanocage is shown in Figure 2-14. The monomeric DNA nanocage elutes as the first main peak, which is composed of 3–4 overlaid peaks, and is followed by a series of small peaks containing aggregated DNA nanocages. While heterogeneity is observed for the main DNA nanocage peak(s), all of these peaks pooled together run as a single band during native PAGE, as shown in lane 2 in Figure 2-14C. This heterogeneity in elution could be due to weak associations between different oligomeric complexes of DNA nanocages that have a small effect on the IEX retention time, which are observed during shallow gradients (0.25% or 5 mM NaCl / min), but that easily dissociate back to a monomeric form during gel migration. Although, agarose gel and electroelution is the current standard for purifying large DNA nanostructures,[98] given the excellent separation of the DNA tetrahedron from higher oli-

gomeric aggregates using IEX-FPLC with a shallow salt gradient, it may also be possible to purify moderate to large DNA nanostructures in the same fashion.

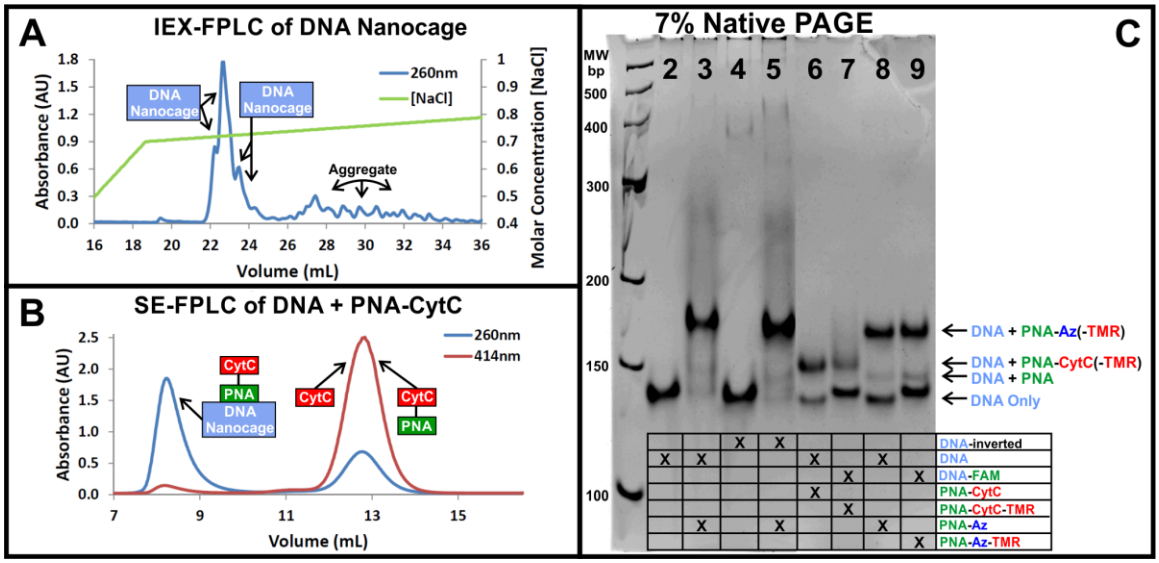


Figure 2-14. DNA nanocage assembly and PNA incubation. [A] Ion Exchange FPLC chromatogram of the annealed DNA nanocage. The following gradient was used to mix solvent α (TAE.Mg pH 9.0) and solvent β (TAE.Mg pH 9.0 + 2 M NaCl): 3.75%/min (20–35% β), 0.25%/min (35–40% β), 1%/min (40–50% β). [B] Size Exclusion FPLC chromatogram of the DNA nanocage incubated with 3x excess PNA-CytC (mixture with free CytC). The column and buffers were chilled on ice during the run. [C] Native PAGE (7%) at 4 ^{0}C showing the fraction of PNA-protein hybridized to the DNA nanocage after SE-FPLC purification, as indicated in parenthesis for each lane as follows: Lane 6 = DNA + PNA-CytC (~75%), Lane 7 = DNA-FAM + PNA-CytC-TMR (~30%), Lane 8 = DNA + PNA-Az-TMR (~60%), Lane 9 = DNA-FAM + PNA-Az-TMR (~50%). This gel also compares the mobility shift induced from PNA-azurin binding (without purification): Lane 2 and 3 have the standard DNA nanocage without and with PNA-azurin, respectively, Lane 4 and 5 have the inverted DNA nanocage without and with PNA-azurin, respectively, where the PNA binding domain is moved across the helix, rotating the orientation of protein attachment by about 180 degrees.

The DNA absorbance at 258 nm was measured on a Beckman Coulter DU640 spectrophotometer using a Hellman TrayCell fiber optic ultra-micro cell with 0.2mm and 1mm path length caps. The molar extinction coefficient of the DNA nanocage was calculated by adding up the extinction coefficients of each of the 6 edges of double stranded

DNA in the nanocage (A = 321,775, B = 322,314, C = 323,636, D = 319,211, E = 316,521, F = 320,058 $M^{-1}cm^{-1}$, where A–F are the edge labels given by the original publisher of this structure[65] and used in Figure 3-2, along with the extinction coefficient of 12 single stranded adenosine bases (151,128 $M^{-1}cm^{-1}$) at the vertices. These extinction coefficients were calculated online at http://biophysics.idtdna.com/UVSpectrum.html using IDT's published algorithm,[88] which is based on a nearest-neighbor model. This algorithm also accounts for the sequence dependent hypochromic absorbance reduction observed in dsDNA from the absorbance of its constituent ssDNA. The following extinction coefficients ($M^{-1}cm^{-1}$) were calculated:

A) Fully assembled DNA nanocage with all bases on the edges hybridized, ε_{258}= 2,074,643,

B) DNA nanocage from A, with 8nt removed for binding PNA1, ε_{258}= 2,074,181,

C) DNA nanocage from A, with 8nt removed for binding PNA2, ε_{258}= 2,074,894,

D) DNA nanocage from A, with 16nt removed for binding both PNA1 and PNA2, ε_{258}= 1,977,651.

PNA-polypeptide assembly with the DNA nanocage

PNA-peptide incubation with the DNA nanocage

The desired amount of preassembled and purified DNA nanocage, typically 20–1000 pmol, was diluted in 1x TAE.Mg buffer to 2 μM concentration, mixed with varying molar ratios of PNA-peptide, incubated at room temperature for 5–10 minutes and then chilled to 4 °C using a thermocycler. For analytical studies, the samples were analyzed

directly using native PAGE at 4 °C, as shown in Figure 2-15. For large scale preparations the sample was purified at 4 °C by SE-FPLC as described in the DNA Nanocage Assembly section. A representative chromatogram is shown in Figure 2-13. The main peak has a large absorbance at 260 nm (A_{260}) from the DNA and PNA nucleobases as well as absorbance from the PNA1-peptide FAM label at 500 nm (A_{500}) and PNA2-peptide TMR label at 560 nm (A_{560}). The A_{500} and A_{560} trace intensities are about 20 times less than A_{260}, which is expected for quantitative PNA-peptide binding to the DNA nanocage based on the extinction coefficients ε_{258}= 2,117,443, ε_{500} = 96,840 ($\varepsilon_{258}/\varepsilon_{500}$ = 21.9), ε_{560} = 91,000 $M^{-1}cm^{-1}$ ($\varepsilon_{258}/\varepsilon_{560}$ = 23.3), as detailed in the PNA-peptide hybridization efficiency section. The chromatogram in Figure 2-13 also shows a second well separated peak at a longer elution time from the unhybridized PNA-peptides. Alternatively, because of the near quantitative yield of the PNA-peptide incubation, for small scale preparations the excess PNA-peptide was removed using a 0.5 mL Amicon Ultra 30 kDa MWCO centrifugal filters (EMD Millipore) by spinning 3–4 times, where after the first 2–3 spins the sample were mixed with 0.5 mL of cold (4 °C) TAE.Mg buffer to ensure all free PNA-peptide was removed. Both PNA-peptides could also be assembled in a stepwise fashion. After each PNA-peptide was added to the DNA nanocage the excess PNA-peptide was also removed using centrifugal filters. The process was repeated for adding the second PNA-peptide. The same end product was achieved regardless of whether the PNA-peptides were assembled simultaneously or in a stepwise fashion.

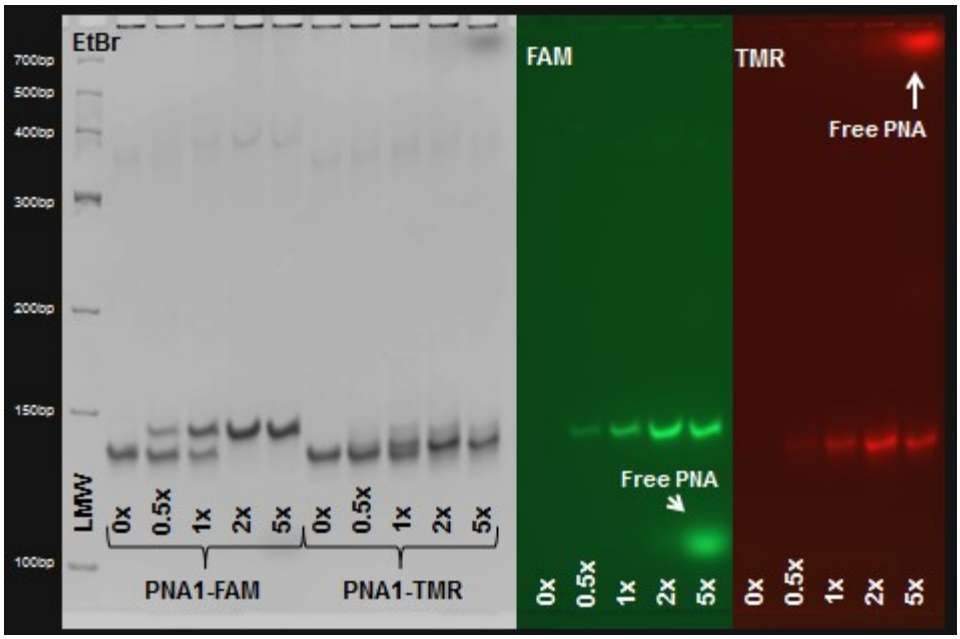


Figure 2-15. Native PAGE of the DNA nanocage incubated with PNA1-GPG at 4 °C. 7% Native PAGE at 4 °C of the DNA nanocage incubated with the indicated molar ratio of PNA1-GPG labeled with FAM and with TMR. [Left] Gel image (EtBr) showing the EtBr fluorescence (Ex. 302 nm, Em. 605 nm) after staining. [Middle] Gel image (FAM) taken of the native FAM fluorescence (Ex. 475 nm, Em. 535 nm) prior to EtBr staining. [Right] Gel image (TMR) taken of the native TMR fluorescence (Ex. 560 nm, Em. 645 nm) prior to EtBr staining. For more information refer to Figure 3-4.

PNA-protein incubation with the DNA nanocage

The complexes listed in Figure 4-3 were formed as follows. For gel hybridization experiments, 0.5x to 5x (PNA-CytC) or 0x to 8x (PNA-Az) PNA-protein conjugate was incubated for 20–30 minutes at room temperature with the preassembled DNA nanocage diluted to 2 μM in TAE.Mg buffer with 1 M NaCl to prevent non-specific binding of the protein to the DNA nanocage, then cooled to 4 ^{0}C for another 15–20 minutes, then loaded directly and analyzed by native PAGE at 4 ^{0}C. For larger scale preparations, 2x to 5x of each PNA-protein conjugate was used instead with 0.5 to 1 M NaCl, and excess PNA-protein was removed by gel filtration using an Amersham Biosciences (GE Healthcare)

Superdex 75 column cooled with ice, using an Amersham Pharmacia (GE Healthcare) ÄKTA Purifier with a flow rate of 1 mL/min of degassed 1x TAE.Mg buffer also cooled on ice. The collected fractions were concentrated to the desired amount using Amicon Ultra 30 kDa MWCO centrifuge filters (EMD Millipore). A representative chromatogram is shown in Figure 2-14B.

Dye labeled and unlabeled DNA-PNA-protein and DNA-PNA complexes were analyzed by native PAGE (7%) with the gels cooled to 4 ^{0}C and imaged on a UVP Bio-Spectrum, as described previously.[79] Native dye fluorescence was imaged first by exciting at either 475 nm (FAM) or 560 nm (TMR) and imaging at 535 nm (FAM) or 645 nm (TMR). Next, the gels were stained with ethidium bromide for 5–10 minutes and destained in nanopure water for 5–10 minutes, then excited at 302 nm and imaged at 605 nm. The gels were washed in nanopure water for an additional 10 minutes then fixed in the gel by rocking for 1 hour in 40 mL of 12.5% (v/v) glutaraldehyde in water. Excess glutaraldehyde was removed by rinsing the gel several times with deionized water followed by soaking twice for 15 minutes in deionized water. The gels were stained in the dark for 1 hour by rocking in 40–80 mL of 1% (w/v) silver nitrate in water. Excess silver nitrate was removed by rinsing the gel 3–4 times followed by soaking the gel in water for 5 minutes and then subsequently for 15–20 minutes. The gel was transferred to 10–20 mL of developer solution (0.25% (v/v) formaldehyde, 6.25% (w/v) sodium bicarbonate in water) where it was gently agitated until the bands of interest were nicely visible and then immediately transferred to 20–30 mL of fixing solution (10% (v/v) glycerol, 10% (v/v) acetic acid in water). The gel was then imaged on the UVP BioSpectrum via white light transillumination from the Manual BioLite. Optimal contrast was achieved using a 535

nm emission filter and converting the image to black and white. Color was added to the image using CorelDRAW version 16.0 (Corel corporation) to produce a yellow background with brown bands as typically shown with silver stain. Silver staining has proven to be an excellent method for identifying protein containing gel bands of small DNA nanostructures,[99] but it may not be practical with large DNA origami nanostructures due to the background DNA staining from such large (>4 MDa) scaffolds.

PNA-peptide hybridization efficiency

In order to determine the degree of PNA-peptide hybridization to the DNA nanocage, the ratio of the PNA and DNA nucleobase absorption at 258 nm to that of the PNA dye label (FAM = A_{500}/A_{258}, TMR = A_{560}/A_{258}) were normalized to the calculated ratio of their respective extinction coefficients (FAM = $\varepsilon_{500} / \varepsilon_{258}$, TMR = $\varepsilon_{560} / \varepsilon_{258}$. The spectra were measured at room temperature (DNA + PNA1-TMR) or with the sample kept on ice for measuring the absorption spectrum (DNA + PNA1-FAM) to ensure stable hybridization in the presence of electrostatic repulsion of FAM^{2-} away from the DNA nanocage. The extinction coefficients of the dyes were provided by the manufacturer, Integrated DNA Technologies (FAM: ε_{495} = 75,000 and ε_{260} = 13,700 $M^{-1}cm^{-1}$, TMR: ε_{559} = 91,000 and ε_{260} = 29,100 $M^{-1}cm^{-1}$), where the observed absorption peaks were slightly red shifted to 500 nm and 560 nm, respectively, after conjugation to PNA or DNA. The extinction coefficient for the DNA and PNA at 258 nm in the complex was ε_{258}= 2,074,643 $M^{-1}cm^{-1}$, and was calculated as described in the DNA Nanocage Assembly section. For the fully assembled complex with both dye labeled PNA-peptides, the extinction

coefficient of TMR at 500nm was determined experimentally to be 21,840 $M^{-1}cm^{-1}$ from the ratio of the absorption maxima at 560 nm (A_{560}), where the extinction coefficient is known, to the absorption at 500 nm (A_{500}), where $A_{500} = 0.24 \times A_{560}$. Adding up all the constituent extinction coefficients for the fully assembled complex with both labeled PNA-peptides yields: ε_{258}= 2,117,443, ε_{500} = 96,840 ($\varepsilon_{258}/\varepsilon_{500}$ = 21.9), ε_{560} = 91,000 $M^{-1}cm^{-1}$ ($\varepsilon_{258}/\varepsilon_{560}$ = 23.3).

The absorption spectra were measured as described in the DNA Nanocage Assembly section. The purified samples were analyzed by native PAGE. A 7% gel was typically used (7% 19:1 Acrylamide:Bisacrylamide polymerized with APS and catalyzed by TEMED in 1xTAE.Mg). Samples were mixed with 10x native tracking dye (0.2% Bromophenol Blue, 0.2% Xylene cyanol, 50% glycerol in 1x TAE.Mg) to a final ratio of 1:1. The gels were typically run at a constant 12.5 V/cm (200 V/16 cm) with the gels cooled to the specified temperature using a Fisher Scientific Isotemp 3006S temperature controlled water bath circulating through the SE 600 Ruby lower buffer chamber. Gels were visualized on a UVP BioSpectrum (Upland, CA, USA) equipped with a BioLite Manual transilluminator with a 150 W Halogen light source and Ascent 4000C 4.2 M pixel color CCD camera. FAM was visualized using a 475 nm excitation filter (UVP, 38-0340-03) and a 535 nm emission filter (UVP, 38-0340-01). TMR was visualized using a 560 nm excitation filter (UVP, 38-0341-03) and a 645 nm emission filter (UVP, 38-0341-01). DNA was visualized by EtBr staining and imaged at 302 nm using a Benchtop 2UV Transilluminator (UVP, 95-0495-01) and a 605 nm emission filter (UVP, 38-0220-01). The PNA-peptide hybridization efficiency to the DNA nanocage, shown in Figure 3-4,

was determined by integrating the FAM and TMR fluorescence intensity of each band of the unstained complex scaled by the amount of the complex in that lane, as determined by the Ethidium Bromide (EtBr) fluorescence intensity of the stained complex, and then normalized to the maximum FAM or TMR fluorescence.

PNA-polypeptide hybridization kinetics

The peak fluorescence (521 nm) of the FAM labeled DNA nanocage was measured using a Nanolog Fluorometer (Horiba Jobin Yvon, 450 W Xenon continuous wave lamp, R928 photomultiplier tube, 1200 grooves/mm excitation and emission gratings) with a 3 mm quartz cuvette (Starna Cells, 16.45F-Q-3/Z15), by exciting the donor (FAM) at 450 nm in order to minimize the amount of direct acceptor (TMR) excitation. The fluorescence intensity was integrated for 0.5 seconds and measured once per second with excitation and emission slit widths of 1.25 nm. A computer controlled NESLAB RTE 7 circulating, refrigerated water bath (Thermo Scientific) was used to maintain a 25 °C or 11 °C sample temperature through the FluorEssence v3.5 software interface. 15 pmol of DNA-FAM nanocage was diluted into 75 μL of TAE.Mg buffer and the fluorescence peak was collected for one minute. Data collection was paused for about 15 seconds while 5x molar excess (75 pmol) of PNA1-GPG-TMR (1 μL of 76 μM), 4x TMR labeled PNA-azurin or 5x TMR labeled PNA-CytC was added and mixed. Data collection was resumed until 15 total minutes of data was collected. The FAM fluorescence was quenched by the TMR label as well as the native heme absorbance in the case of cytochrome c. Similarly, DNA-FAM was monitored at 11 ^{0}C as 2x PNA-CytC (unlabeled) was added. The data was processed by normalizing the minimum and maximum fluores-

cence values to zero and one, respectively, and then inverted $(-1 * (x_i - 1))$ to show the fraction of PNA bound as a function of time, as shown in Figure 2-16 and Figure 3-6. The binding kinetics was fit with an exponential function using OriginPro version 9.1.0 with the following equation: $y = y_0 - A * e^{-R_0 * t}$, where y_0 is the upper limit once binding terminates, A is the exponential prefactor and R_0 is the rate of PNA-protein hybridization. The fitting parameters are shown in Table 3-1 for PNA1-GPG-TMR and Table 2-1 for the PNA-proteins.

TMR labeled PNA-proteins were added to separate FAM labeled DNA nanocages (DNA-FAM) to investigate the hybridization kinetics based on changes in Förster resonance energy transfer (FRET), as the fluorescence from the excited donor FAM becomes quenched by the acceptor TMR during PNA-protein hybridization. The peak FAM fluorescence of DNA-FAM was monitored as either 5x molar excess PNA-CytC-TMR or 4x molar excess of PNA-Az-TMR were separately added and mixed at 25 ºC in the presence of 1 M NaCl to prevent non-specific binding. DNA-FAM fluorescence was rapidly quenched upon TMR labeled PNA-protein binding as shown in Figure 2-16A and Figure 2-16C. The raw fluorescence data was normalized and inverted to show the fraction of PNA-protein bound to the DNA nanocage as a function of time, as shown in Figure 2-16B and Figure 2-16D. PNA-Az-TMR and PNA-CytC-TMR were found to reach 99% hybridization within 2 minutes and within 4 minutes, respectively, as shown in Figure 2-16. These rates are slightly faster than the 5–10 minutes previously observed for PNA-peptides.[79] The positively charged PNA-CytC may take longer to bind because of stronger electrostatic interaction with the DNA nanocage in conformations that slow the PNA from locating its binding site. 2x molar excess of unlabeled PNA-CytC was also incubat-

ed with DNA-FAM as a control for another fluorescence experiment, but it was found that the heme absorption also quenched the FAM fluorescence such that the binding kinetics could be monitored. PNA-CytC was found to bind to the DNA nanocage at 11 ºC within 16 minutes, as shown in Figure 2-16E and Figure 2-16F.

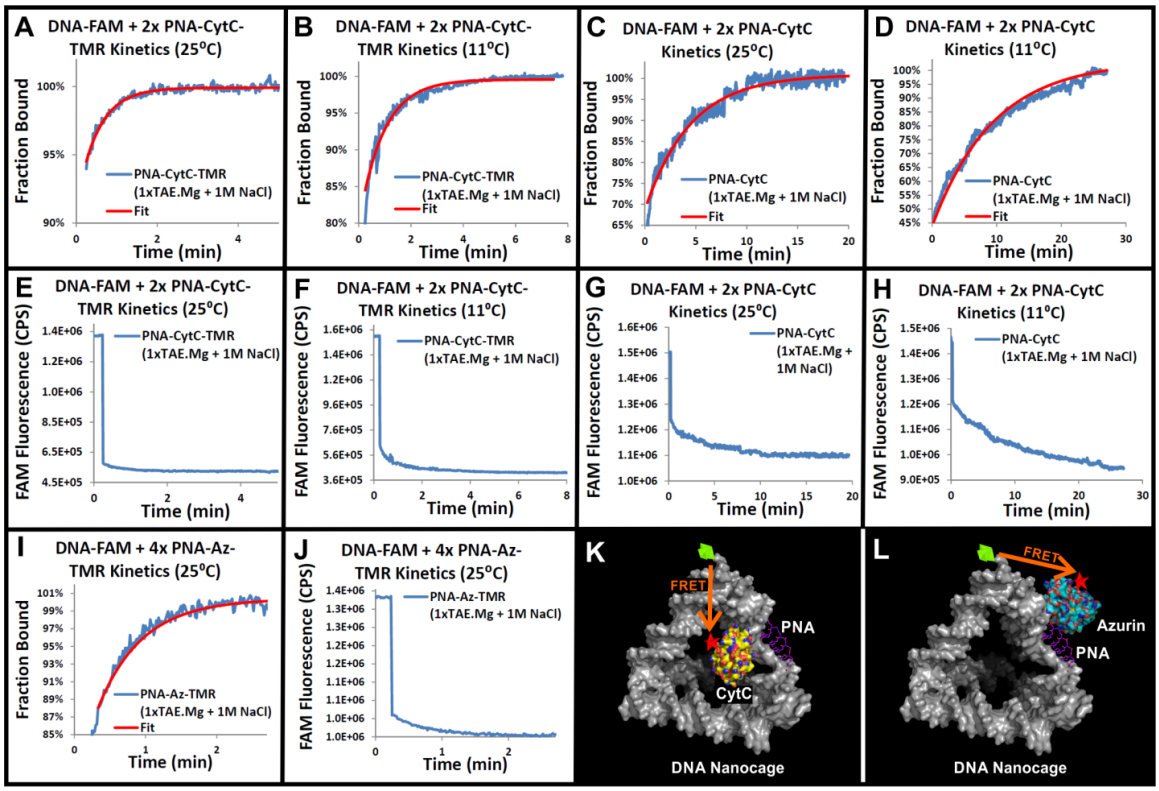


Figure 2-16. Time course PNA-protein hybridization with the DNA nanocage. Molecular models of the FAM labeled DNA nanocage with TMR labeled CytC [K] and Azurin [L]. 2x to 4x excess PNA-protein was mixed for ~15 seconds (no data collected) and then the FAM fluorescence at 520 nm was monitored over time for PNA-CytC-TMR at 25°C [A, E] and PNA-CytC-TMR at 11°C [B, F], PNA-CytC (unlabeled) at 25°C [C, G] and PNA-CytC (unlabeled) at 11°C [D, H] and PNA-Az-TMR at 25°C [I, J]. Plot of the peak FAM fluorescence at 521 nm as a function of time after the PNA-protein was added [E, F, G, H, J]. The peak FAM fluorescence was normalized and inverted to show the fraction of PNA-protein bound and plotted as a function of time [A, B, C, D, I]. The exponential fit was produced as described in the methods section and the fitting parameters are shown in Table 2-1.

Table 2-1. Fitting parameters and time to reach 99% hybridization for PNA-protein binding kinetics with DNA nanocage. Table of parameters obtained by fitting the hybridization kinetics data shown in Figure 2-16 to an exponential function. The equation is described in the PNA-protein assembly with the DNA nanocage methods section, and was used to determine the time for each PNA-protein to reach 99% hybridization.

Fitting Parameter	[25 °C] DNA-FAM + PNA-CytC-TMR	[25 °C] DNA-FAM + PNA-Az-TMR	[11 °C] DNA-FAM + PNA-CytC
y_0	100.7 ± 0.05%	100.4 ± 0.2%	100.1 ± 0.05%
A (counts)	(209.4 ± 1.2) x 10^{-3}	(211.6 ± 6.2) x 10^{-3}	(280.5 ± 0.4) x 10^{-3}
R_0 (counts/sec)	(9.9 ± 0.1) x 10^{-3}	(26.8 ± 1.2) x 10^{-3}	(3.29 ± 0.02) x 10^{-3}
R^2	0.9941	0.9685	0.9982
Time to 99% Hybridization	252 sec (4.2 min)	101 sec (1.7 min)	977 sec (16.3 min)

DNA-PNA-polypeptide thermal stability and fluorescence characterization

Sample preparation

For thermal stability measurements, the following constructs were assembled by incubating 2–5x molar excess of PNA-polypeptide with the DNA nanocage and purified by SE-FPLC at 4 °C, as described in the DNA Nanocage Assembly section: 1) DNA-FAM + PNA1-TMR, 2) DNA-TMR + PNA1-FAM, 3) DNA-FAM + PNA-CytC-TMR, 4) DNA-FAM + PNA-Az-TMR, and controls 5) DNA-FAM, and 6) DNA + PNA1-FAM. For TCSPC measurements the following complexes were assembled by incubation and kept at 4 °C during purification and thereafter: 1) DNA + PNA1-FAM + PNA2-TMR, a donor only control 2) DNA + PNA1-FAM + PNA2-Ac (acetylated), and acceptor only controls 3) DNA + PNA1-Ac + PNA2-TMR, 4) DNA + PNA-CytC-TMR, 5) DNA + PNA-Az-TMR. The named constructs are all defined in Figure 3-3.

Steady state fluorescence spectra

The same setup was used as described in the PNA hybridization kinetics section. We excited FAM at 440 nm for thermal dissociation experiments for PNA-peptides and 450 nm for PNA-proteins. However, the laser intensity for TCSPC experiments was much stronger at 450 nm so we chose to excite FAM at 450 nm for both TCSPC experiments and steady state FRET experiments, which allowed for direct comparison of energy transfer efficiency. Furthermore, after more careful analysis of the TMR and FAM absorption spectra, we noticed that TMR had a slightly lower absorption at 450 nm than 440 nm and FAM had higher absorption, so the 450 nm excitation wavelength is preferred. Donor and acceptor emission spectra were collected in 1 nm increments, integrating the fluorescence intensity for 0.5 seconds per data point, with excitation and emission slit widths as specified: 1.25 nm for steady state FRET measurements and 2.5 nm for thermal dissociation experiments. Fluorescence measurements were preceded by measuring the absorption on a JASCO V-670 spectrophotometer in order to normalize the fluorescence. The peak emission was normalized to the concentration of DNA nanocage from the absorbance at 260 nm for each construct. For thermal dissociation experiments, a water bath was used to control the sample temperature as described in the PNA hybridization kinetics section. The temperature was increased from 11 °C to 85 °C and back to 11°C in 0.5°C steps. Each temperature was equilibrated ±0.25 °C for 30 seconds before the fluorescence spectrum was acquired. All spectra were corrected for systematic responses of the gratings, monochromator components and detector using the correction file (Mcorr_R928P.spc) provided by the manufacturer (Horiba Jobin Yvon). Although both complexes shown in Figure 3-7 were measured at 210 nM concentration, the higher

fluorescence signal observed for DNA-TMR + PNA1-FAM is likely because of the higher FAM quantum yield when attached to the PNA over when attached directly to the DNA nanocage.

Dissociation temperature (T_D) curves

Two donor only controls were prepared to quantify the change in donor (FAM) fluorescence with temperature. A control was prepared with the donor coupled directly to the DNA nanocage (DNA-FAM) as well as a control with the donor coupled to the PNA1-peptide and hybridized to the DNA nanocage (DNA + PNA1-FAM). The steady state fluorescence emission spectra were monitored as a function of temperature as described in the Steady State Fluorescence Spectra section. The fractional change from the maximum donor fluorescence at each temperature was converted into a correction factor using the following formula: $CF(\lambda) = 1 + \frac{I_{11°C} - I_\lambda}{I_{11°C}}$, where $CF(\lambda)$ is the wavelength dependent correction factor, $I_{11°C}$ is the maximum donor fluorescence of the control which occurs at the coldest temperature of the experiment (11 °C) and I_λ is the donor fluorescence of the control at each wavelength, as shown in Figure 2-17.

The maximum donor fluorescence of the scan at each temperature was corrected by the corresponding control (i.e., DNA-FAM + PNA1-TMR, DNA-FAM + PNA-CytC-TMR and DNA-FAM + PNA-Az-TMR were corrected by DNA-FAM; DNA-TMR + PNA1-FAM was corrected by DNA + PNA1-FAM). The corrected data were normalized to the maximum peak donor fluorescence and fit to a sigmoidal dose response curve using Origin version 8.6 (PNA-peptide complexes) or 9.1 (PNA-protein complexes), with

the following equation: $y = A_1 + (\frac{A_2 - A_1}{1+10^{(LOG(x_0)-x)p}})$, where A_1 and A_2 are the lower and upper asymptotes, respectively, x is the normalized, corrected fluorescence at a given temperature, LOG(x_0) is the base 10 logarithm at the inflection point of the curve, x_0 , which is also the dissociation temperature (T_D), and p is the variable Hill slope at x_0. The fit assumes a two state model where the associated complex [DNA + PNA1] dissociates directly to its constituent components [PNA1] and [DNA] without any intermediate states.

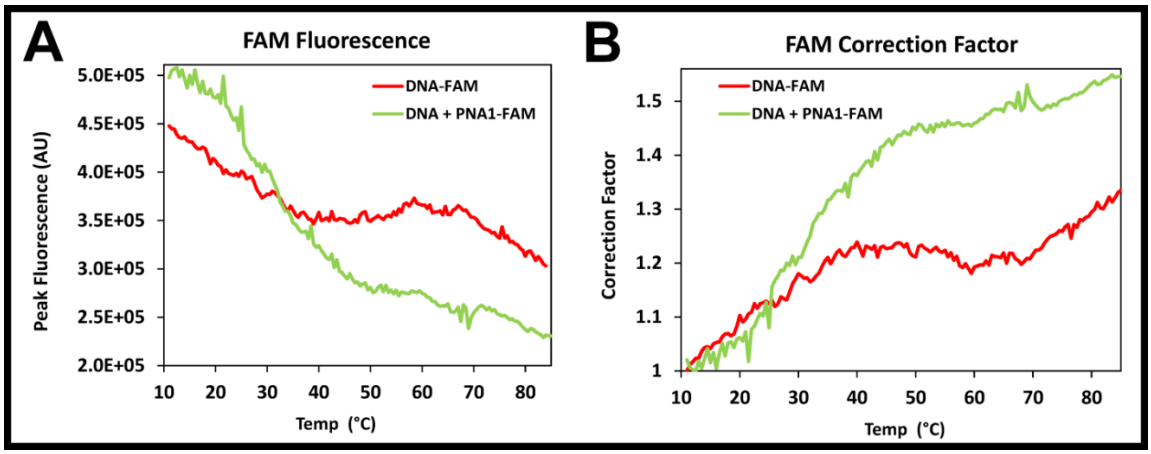


Figure 2-17. Temperature dependence of FAM Fluorescence. [A] Temperature dependent peak fluorescence of FAM labeled on the DNA nanocage (DNA-FAM) and labeled on the PNA-peptide and hybridized in the unlabeled DNA nanocage (DNA + PNA1-FAM). [B] Correction factor calculated from the fractional change in FAM fluorescence for each construct from [A] with respect to the maximum fluorescence of each scan.

The predicted dissociation temperatures were calculated using the method described by Giesen, Nielsen and coworkers[95] with the following equation: $T_{D\,pred} = c_0 + c_1 * T_{D\,DNA} + c_2 * f_{pyr} + c_3 * L$, where $c_0, c_1, c_2,$ and c_3 are constants with the values of 20.79, 0.83, -26.13 and 0.44, respectively, $T_{D\,DNA}$ is the predicted dissociation temperature of a DNA duplex of the same sequence calculated online at

http://biophysics.idtdna.com using the methods described by Owczarzy and coworkers[100] based on the nearest neighbor model, f_{pyr} is the fraction of the sequence with pyrimidine (C or T) bases, and L is the length of the sequence.

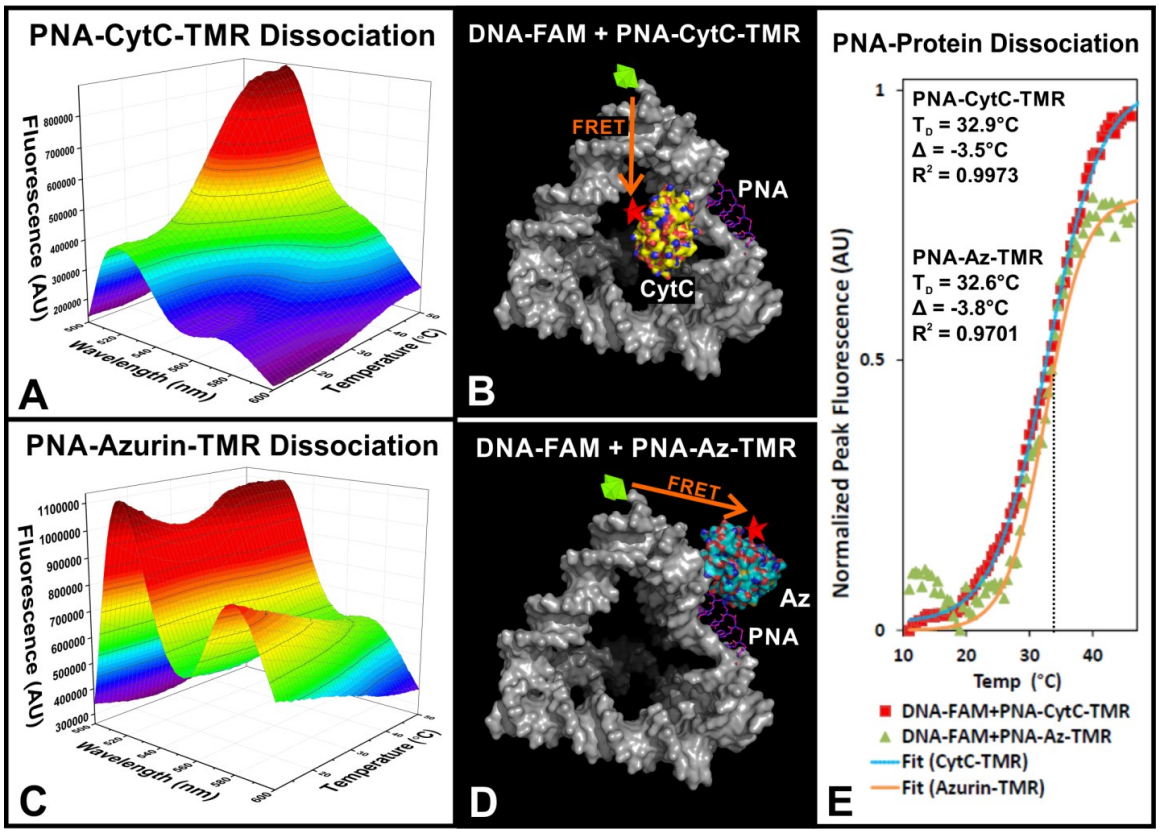


Figure 2-18. Fluorescence spectra of the TMR labeled PNA-proteins bound to the FAM labeled DNA nanocage as a function of temperature. The FAM labeled DNA nanocage (DNA-FAM) containing either TMR labeled PNA-CytC (A) or PNA-azurin (C) was excited at 450 nm and the emission was monitored as a function of wavelength as the temperature was increased from 11 ^{0}C to 85 ^{0}C. The molecular models of the constructs used to measure the spectra in from A and C are shown in B and D. The peak donor fluorescence of each construct is plotted versus temperature and fitted with a sigmoidal dose response curve (E), where the inflection point corresponds to the dissociation temperature (T_D) indicated on the plot. The difference (Δ) from a theoretical value (36.6 °C) is also shown, which is greater than the value predicted for a similar DNA sequence (34.5 °C).

Donor quantum yield

The quantum yield (Q_Y,Φ) of the donor (FAM) was calculated by measuring its peak fluorescence intensity when conjugated to the PNA-peptide and hybridized to the DNA nanocage (DNA + PNA1-FAM and DNA + PNA1-FAM + PNA2-Ac) and when conjugated directly to the DNA nanocage (DNA-FAM), and comparing these values to the fluorescence intensity of the dye standard Sodium Fluorescein in 0.1 M NaOH, which has Φ_R = 0.925 at 22 °C reported by Magde and coworkers.[101] The Q_Y was calculated using the following equation: $\Phi = \Phi_R \frac{\frac{I}{A}}{\frac{I_R}{A_R}} \frac{\eta^2}{{\eta_R}^2}$, where the fluorescence intensity of the sample (I) and reference (I_R) are normalized by their absorbances (A and A_R, respectively) at the excitation wavelength and scaled by the ratio of the square of the two index of refraction (η *and* η_R, respectively). For quantum yields measured at low concentration, the excitation and emission slit widths were opened up to 2.5nm to increase the fluorescent signal. However, at these low concentrations the dye absorbance could not be determined reliably. In this case the fluorescence intensities of the donor complex and the reference were normalized by the estimated sample concentration based on the dilution factor from a known initial sample concentration.

Calculating the absorption spectra and extinction coefficients for the overlap integral of TMR labeled PNA-proteins

The overlap between the donor (FAM) fluorescence emission and acceptor (TMR) absorption is used to determine the Förster radius. We determined the spectra for the conjugate when hybridized to the DNA nanocage because of the dependence of the

TMR absorption spectrum its environment. The absorption of the FAM labeled DNA nanocage (DNA-FAM) was subtracted from that of the DNA-FAM nanocage hybridized to either TMR labeled protein, as shown in Figure 2-19. The determined absorption spectra for DNA + PNA-CytC-TMR and DNA + PNA-Az-TMR were used along with the fluorescence emission of the DNA-FAM nanocage to determine the overlap integral in subsequent Förster distance calculations. The extinction coefficient of PNA-Az-TMR and PNA-CytC-TMR at 555 nm in separate DNA-FAM nanocages was calculated to be 32,000 $M^{-1}cm^{-1}$ and 71,638 $M^{-1}cm^{-1}$ as follows: $\varepsilon_{555} = \varepsilon_{258} * \frac{A_{555}}{A_{258}}$ where ε_{258} is the extinction of the nucleobases in the DNA nanocage and PNA sequence of 2,074,643 $M^{-1}cm^{-1}$ as calculated previously[79] and A_{258} and A_{555} are the observed absorbance values for the complex. Absorption spectra were measured on a Beckman Coulter DU640 spectrophotometer using a Hellman TrayCell fiber optic ultra-micro cell with 0.2 mm or 1 mm path length caps.

The TMR absorbance spectra when attached to either protein showed an increase in the absorption peak at 518 nm, compared to the free dye, as shown in Figure 2-19. This phenomenon is commonly observed when TMR self-associates, such as when conjugated to proteins[102] or in solution at high concentration.[103] The higher energy peak at 518 nm for PNA-CytC-TMR indicates that it may have a higher degree of labeling than for PNA-Az-TMR. However, because of the convolution between the TMR absorption spectral shift and amplitude due to multiple dye labels per protein, we were unable to quantify the extent of TMR labeling on each protein. Figure 2-19C and Figure 2-19D show the spectra of PNA-CytC-TMR and PNA-Az-TMR within a mixture with CytC-TMR and Az-TMR, respectively, prior to and after hybridization with the DNA nanocage. The 518 nm peak is

significantly reduced upon binding to the DNA nanocage. This may be partially explained by the lower effective concentration of the dye labeled PNA-protein bound to the DNA nanocage, making it much less likely for the dye labeled proteins to self-associate.

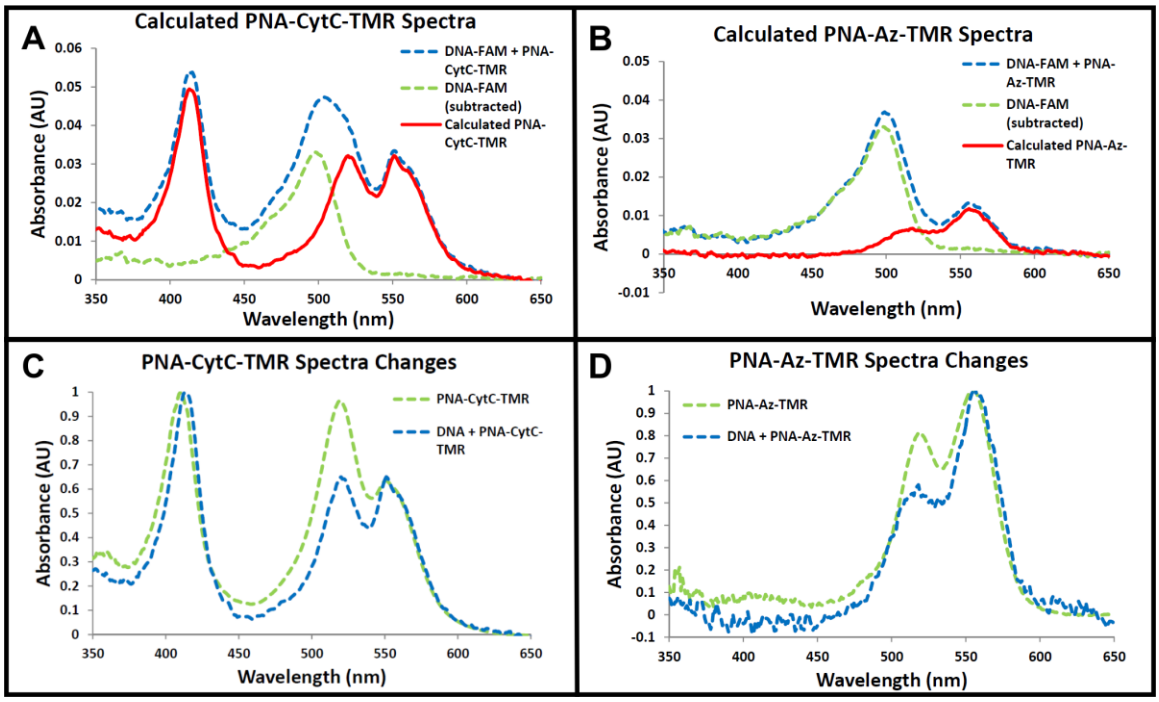


Figure 2-19. Calculated dye labeled PNA-protein absorption spectra and spectral changes upon hybridization to the DNA nanocage. Absorption spectra of the FAM labeled DNA nanocage with TMR labeled PNA-protein (blue), the FAM labeled DNA nanocage alone (green) and the absorption spectra of the TMR labeled protein (red) calculated by subtracting the green curve from the blue curve for PNA-CytC-TMR [A] or PNA-Az-TMR [B]. Changes to the TMR absorption spectra upon biding to the DNA nanocage for PNA-CytC-TMR [C] or PNA-Az-TMR [D]. Prior to binding to the DNA nanocage, the samples of PNA-CytC-TMR and PNA-Az-TMR also contain TMR labeled protein without PNA; the effect this has on the observed changes in spectra is discussed below.

The reduction in the 518 nm peak is more significant for PNA-CytC-TMR than for PNA-Az-TMR. Because cytochrome c was labeled with TMR prior to conjugation with PNA, the PNAs may conjugate preferentially to cytochrome c's with fewer TMR labels, which would lead to a larger reduction in the 518 nm peak compared to the PNA-Az-TMR, which was labeled with TMR after PNA conjugation. The slight difference in the Soret

peak of 411 nm and 413 nm between PNA-CytC-TMR and DNA-PNA-CytC-TMR is within the expected range for fully oxidized (410 nm) and fully reduced (414 nm) sta468.

Fluorescence lifetime

Time Correlated Single Photon Counting (TCSPC) spectra were measured using a 20 MHz Fianium FemtoPower1060, Supercontinuum Laser Source with the output set to 450 nm or 500 nm to excite the donor or acceptor, respectively, at a 10 nm spectral bandwidth using a combination of Acousto-Optic Tunable Filter (Fianium Ltd., Eugene, OR) and an interference filter centered at the excitation wavelength. The fluorescent signal was collected at a right angle to the excitation beam and detected using a double-grating monochromator (Jobin-Yvon, Gemini-180) and a microchannel plate photomultiplier tube (Hamamatsu R3809U-50). Data was acquired with a single photon counting card (Becker-Hickl, SPC-830). Lifetime data were collected through a polarizer set at the magic angle (54.7°). The instrument response function (IRF) was measured by sample scattering at the excitation wavelength and had a full width at half-maximum (FWHM) of 105–115 ps for data collected for the DAS and 60–118 ps for data collected for anisotropy. The sample was cooled to 11 °C using a refrigerated water bath to maintain quantitative PNA-peptide binding to the DNA nanocage during the measurement without water condensing on the fluorescence cuvette. All decay kinetics used to construct the DAS were collected for the same amount of time at each wavelength. The home-built software ASUFIT version 3.1 (PNA-peptide complexes) or 3.2 (PNA-protein complexes) running on MATLAB version 7.0.4 was used to fit the decay kinetics using a multi-exponential model for lifetime data or for DAS, all recorded wavelengths were fit simultaneously and

deconvoluted by global analysis. The lifetimes shown in Figure 2-21 are an average of the components identified in Table 2-2. The fluorescence kinetics data as a function of decay time and wavelength were fit simultaneously using the following multi-exponential decay model, $F(\lambda, t) = \sum_i A_i(\lambda) * \exp(-\frac{t}{\tau_i(\lambda)})$, where F is the fluorescence intensity decay kinetics at the measured wavelength, $\tau_i(\lambda)$ is the calculated lifetime of the identified decay component and $A_i(\lambda)$ is the calculated amplitude of decay at the measured wavelength.

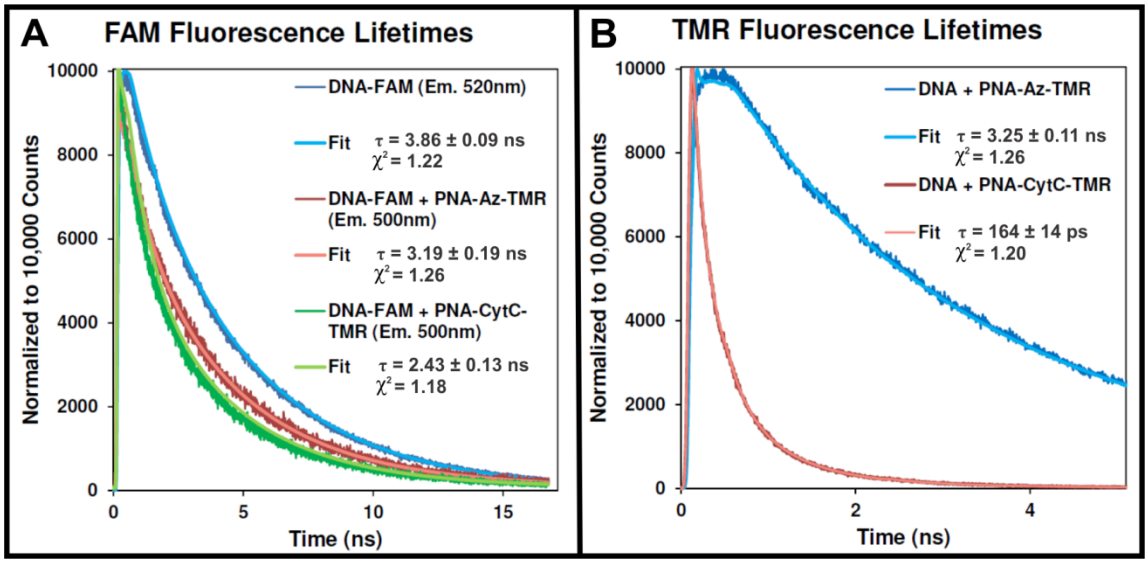


Figure 2-20. Donor (FAM) fluorescence decay kinetics attached to DNA nanocage with and without the acceptor (TMR) labeled PNA-proteins hybridized to the DNA nanocage. Fluorescence decay kinetics measured by TCSPC at 11°C of [A] the donor (FAM) when excited at 450 nm attached to the DNA nanocage alone (blue) and when hybridized to the acceptor (TMR) labeled PNA-azurin (red) or PNA-cytochrome c (green) as well as [B] the acceptor (TMR) when excited at 500 nm attached to the either PNA-CytC or PNA-azurin and hybridized to the unlabeled DNA nanocage.

The lifetimes shown in Figure 3-8 (A) are an average of the components identified in Table 3-2 as calculated using the equation $\tau(\lambda) = \frac{\sum_i A_i(\lambda) * \tau_i(\lambda))}{\sum_i A_i(\lambda)}$, where the amplitude $A_i(\lambda)$ is the percentage of each component of a given lifetime $\tau_i(\lambda)$. The energy

transfer efficiency was calculated from the percent change in donor lifetime in the presence of the acceptor using the following equation: $E = 1 - \frac{\tau_{DA}}{\tau_D}$, where E is the energy transfer efficiency, τ_D is the lifetime of the donor alone and τ_{DA} is the lifetime of the donor in the presence of the acceptor.

Table 2-2. Lifetime components and calculated average lifetimes of the donor and acceptor fluorophores. Each "%" column shows the relative amount of each lifetime component. The lifetime of the donor (FAM) in the presence of the acceptor (TMR) was measured in the donor emission region (500 nm) when the donor was excited at 450 nm. The lifetime of the acceptor (TMR) was measured at 600 nm by exciting the acceptor at 500 nm. DNA + PNA-CytC-TMR complex required four fitting parameters due to the higher degree of TMR labeling on cytochrome c and correspondingly more complex decay pathways compared to azurin.

FAM Construct	τ_1	% of τ_1	τ_2	% of τ_2	Average Lifetime
DNA-FAM alone	4.49 ± 0.06 ns	81.5%	1.08 ± 0.24 ns	18.5%	3.86 ± 0.09 ns
DNA-FAM + PNA-Az-TMR	4.10 ± 0.14 ns	71.5%	0.9 ± 0.3 ns	28.5%	3.19 ± 0.19 ns
DNA-FAM + PNA-CytC-TMR	3.75 ± 0.12 ns	55.4%	0.80 ± 0.14 ns	44.6%	2.43 ± 0.13 ns

TMR Construct	τ_1	% of τ_1	τ_2	% of τ_2	τ_3	% of τ_3	τ_4	% of τ_4	Average Lifetime
DNA + PNA-Az-TMR	3.82 ± 0.07 ns	74.8	1.57 ± 0.21 ns	25.2	-	-	-	-	3.25 ± 0.11 ns
DNA + PNA-CytC-TMR	32 ± 7 ps	63.1	192 ± 16 ps	23.5	624 ± 29 ps	12.3	2.05 ± 0.2 ns	1.1	164 ± 14 ps

Fluorescence decay kinetics were measured by exciting the DNA-FAM at 450 nm and measuring the fluorescence decay at 500 nm where it should be dominated by the FAM kinetics. The decay was fit by two exponentials globally over the wavelength region from 500 to 700 nm. A 1 ps decay component was necessary in some of the fits due to a slight time offset between the measured decay kinetics and the measured instrument

response function (IRF). This is well below the resolution limit of the instrument and therefore was not included for calculating the lifetimes. The lifetimes of each construct were calculated by weighting the lifetime from each exponential by the amplitude percentage required to produce the fit, as shown in Table 2-2. The DNA-FAM lifetime of 3.86 ± 0.09 ns is decreased to 3.19 ± 0.19 ns in the presence of PNA-Az-TMR and decreased to 2.43 ± 0.13 ns in the presence of the PNA-CytC-TMR due to energy transfer. The change in donor lifetime was used to calculate the energy transfer efficiency of 17 ± 5% for PNA-Az-TMR and 37 ± 3% for PNA-CytC-TMR, as described previously,[79] which are both in agreement with the results obtained using steady state fluorescence. TCSPC was also used to measure the fluorescence decay kinetics of the TMR labeled proteins attached to the unlabeled DNA nanocage, which are plotted in Figure 2-20B. The average lifetime of PNA-Az-TMR in the DNA nanocage was 3.25 ± 0.11 ns, which is in agreement with the lifetime of monomeric TMR.[104] The average lifetime of PNA-CytC-TMR in the DNA nanocage was significantly reduced to 164 ± 14 ps, which is likely because of the increased absorbance of the non-fluorescent energy band at 518 nm.[102]

Decay kinetics were also measured at various wavelengths between 500–640 nm that span the FAM emission spectrum and fit using global analysis with two exponentials. The amplitudes are plotted as a function of wavelength as the decay associated spectra (DAS) shown in Figure 2-21. Just as with the lifetime fitting, a 1 ps exponential term was used to compensate for the IRF time delay offset artifact, but was not included during subsequent analysis. DNA-FAM + PNA-Az-TMR had a 4.3 ± 0.1 ns component, which is from both the native FAM decay and TMR decay after energy transfer. The DAS fit also contains a 1.1 ± 0.4 ns component that shows reduced amplitude (increased fluores-

cence) in the TMR absorption region at 590 nm compared to the 4.3 ns component, indicating energy transfer from DNA-FAM to PNA-Az-TMR. DNA-FAM + PNA-CytC-TMR also has a 4.0 ± 0.1 ns component for the native FAM decay as well as 1.0 ± 0.15

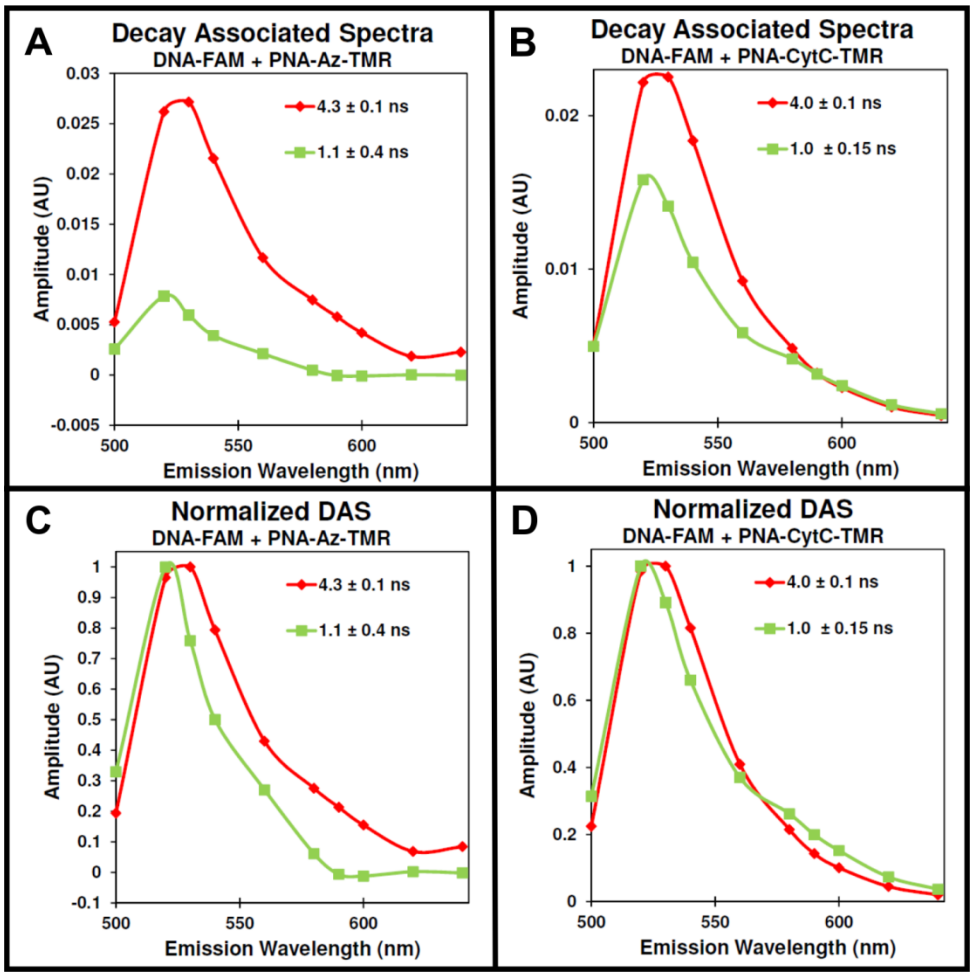


Figure 2-21. Decay associated spectra (DAS) of the FAM labeled DNA nanocage hybridized to either the TMR labeled PNA-azurin or PNA-cytochrome c. Fluorescence decay spectra were collected at several wavelengths and fit simultaneously to two components using global analysis. The relative amplitudes of each component from the analysis are plotted as a function of wavelength in the Decay Associated Spectra (DAS) for the FAM labeled DNA nanocage hybridized to either the TMR labeled PNA-azurin (A) or PNA-cytochrome c (B). Each component of the DAS from A and B were normalized and plotted in C and D, respectively, to highlight the reduced amplitude from energy transfer at 590 nm for PNA-Az-TMR (C) and at 540 nm for PNA-CytC-TMR (D).

ns component with reduced amplitude at 540 nm indicating energy transfer from DNA-FAM to PNA-CytC-TMR. The shift of the reduced amplitude domain from 590 to 540 nm indicates energy transfer to PNA-CytC-TMR is almost exclusively due to the 518 nm absorption band, which is expected because of the more significant overlap with the DNA-FAM emission. However, because the 518 nm TMR band is non-fluorescent,[102] there is much less amplitude decrease (increased fluorescence) due to energy transfer from FAM to TMR between 580–600 nm as compared to the same band in the 4.3 ns component for PNA-Az-TMR.

Fluorescence anisotropy

Fluorescence decay measurements were performed as described in the Fluorescence Lifetime section with the laser tuned to excite either the donor (FAM) at 450 nm or the acceptor (TMR) at 500 nm. We measured DNA nanocage complexes containing PNA at 11 °C to stabilize the PNA-DNA hybridization and measured at 22 °C for DNA nanocage complexes without PNA. DNA dye labeling was performed by the manufacturer (IDT) through a thymine base internally modified with amino hexane linker, which is located at the vertex adjacent to the PNA1 binding gap in the fully assembled DNA nanocage. The anisotropy (r) of each dye was calculated as a function of decay time (t) by comparing the emission with polarization both parallel (VV, vertical excitation and emission) and perpendicular (VH, vertical excitation and horizontal emission) relative to the polarization of the excitation using the following equation: $r(t) = \frac{I_{VV}(t) - G(t) * I_{VH}(t)}{I_{VV}(t) + 2G(t) * I_{VH}(t)}$, where $I_{VV}(t)$ and $I_{VH}(t)$ are the emission spectra collected as a function of time parallel and perpendicular to the excitation polarization, respectively, and G(t) is the "G factor"

as a function of time as defined by: $G(t) = \frac{I_{HV}(t)}{I_{HH}(t)}$, where I_{HV} and I_{HH} are the emission spectra collected as a function of time perpendicular and parallel to the excitation polarization, respectively, when the excitation polarization is rotated 90 degrees from what was used to measure $I_{VV}(t)$ and $I_{VH}(t)$. The "G factor" for our measurement setup was 1.5 at 520 nm emission and 1.3 at 600 nm emission. The anisotropy is plotted in Figure 8 along with the central moving average at each data point ± 0.02 ns (9 data points) for data before 0.3 ns and ± 0.12 ns (49 data points) for data after 0.3 ns.

Förster radius, energy transfer efficiency and distance calculations

The Förster radii of each complex were calculated using the following equation $R_0{}^6(cm^6) = 8.79\ x\ 10^{-25}(\kappa^2 n^{-4} Q_D J(\lambda))$, where κ is the orientation factor and is assumed to be ⅔ for a randomly oriented dye, n is the index of refraction of the aqueous 1x TAE.Mg buffer, which is assumed to be the same as water (n =1.333), Q_D is the quantum yield of the donor, as calculated for each complex in the Donor Quantum Yield section, and $J(\lambda)$ is the overlap integral ($M^{-1}cm^3$), which is defined as follows: $J(\lambda) = \int_0^\infty F_D(\lambda)\, \varepsilon_A(\lambda)\lambda^4 d\lambda$, where F_D is the donor fluorescence at each wavelength, with the total donor fluorescence overlap region normalized to unity, ε_A is the extinction coefficient ($M^{-1}cm^{-1}$) of the acceptor at each wavelength, which was calculated based on the extinction coefficient for the peak absorbance of TMR at 560 nm (91,000 $M^{-1}cm^{-1}$) as provided by the manufacturer IDT, and λ is the wavelength in centimeters. The overlap integral was calculated using numerical approximation by summing the product of the integrand from the start of the fluorescence scan at 470 nm to 650 nm, where the acceptor

absorption is negligible. The Förster distance for each complex was calculated using the following equation: $r = R_0 \sqrt[6]{\frac{1}{E} - 1}$, where R_0 is the calculated Förster radius, and E is the energy transfer efficiency, which is calculated for each complex as follows: $E = 1 - \frac{F_{DA}}{F_D}$, where F_{DA} and F_D are the peak donor fluorescence intensities with and without the acceptor, respectively, normalized by the absorption of each complex at the excitation wavelength, or normalized by the sample concentration for very dilute samples.

Förster distance uncertainty calculations

The anisotropy data was used to put limits on the orientation factor κ based on a method developed by Lakowicz and coworkers[13] using the following equation: ${\kappa_{min}}^2 = \frac{2}{3}\left[1 - \frac{(d_D + d_A)}{2}\right]$ and ${\kappa_{max}}^2 = \frac{2}{3}\left[1 + d_D + d_A + 3\, d_D d_A\right]$ and $d_i = \sqrt{\frac{r_i}{r_0}}$, where r_i is the fraction of depolarization from the local fluorophore motion and r_0 is fraction of depolarization from the rotational diffusion of the DNA nanocage. Because the DNA nanocage is so large (radius ~3 nm) we assume that the depolarization from rotational diffusion during the lifetime of our fluorophores (~4 ns) is negligible. Thus we can use our measured anisotropies for r_i (FAM = 0.12, TMR = 0.24) and the maximum possible aniosotropy of 0.4 for r_0 to calculate the depolarization factors (FAM = 0.55, TMR = 0.78) and limits on the orientation factor (κ_{min} = 0.2, κ_{max} = 1.7). The anisotropy of the TMR labeled PNA-proteins could not be reliably determined due to the effects of multiple dye labels on some proteins. Instead, the anisotropy of the TMR labeled DNA nanocage was used for the acceptor in the DNA-PNA-protein complexes, because the rotation of the DNA

nanocage is the slowest moving component of the system and therefore estimates the highest possible anisotropy and largest possible uncertainty in the measured distances. For each of the distances calculated using the Förster equation and shown in Figure 3-10, the orientation factor limits were used to calculate the upper (r_{max}) and lower (r_{min}) distance limits from the following equations: $r_{min} = r_{meas} * \sqrt[6]{[\frac{\kappa_{min}}{\frac{2}{3}}]}$ and $r_{max} = r_{meas} * \sqrt[6]{[\frac{\kappa_{max}}{\frac{2}{3}}]}$, where r_{meas} are the measured distances from the Förster equation. The uncertainties reported in Figure 3-10 were the average difference between each measured distance and its corresponding upper and lower limits.

Molecular structures, models and chemical structures

Molecular structures of DNA nanocage, PNA, cytochrome c and azurin

The molecular structures of the 8mer PNA,[105] cytochrome c,[106] azurin[107] and the photosystem II active site[108] shown in Figure 2-18, Figure 4-1, Figure 4-4, Figure 4-5, Figure 4-10, Figure 4-12 and Figure 6-5 were taken from the protein data bank and rendered using PyMOL[109] version 0.99 or 1.1. The molecular model of the DNA tetrahedron shown in Figure 2-18, Figure 4-10, Figure 4-12 and Figure 6-5 was assembled in Coot[110] version 0.7.1 and then exported as a PDB file that was then used as a model to generate the structural figures shown in Figure 9. PDB accession codes for the PNA, cytochrome c, azurin and photosystem II active site protein models are 3MBS, 1HRC, 1JVO and 2AXT, respectively.

PNA-protein chemical linker structures

The structures of the chemical linkers used to connect the PNA to both proteins shown in Figure 9 as well as the chemical structure of luminol in Figure 1 were drawn using Marvin Sketch version 6.0.0.[111] The minimum and maximum length of the PNA-protein linker indicated in Figure 9 was calculated as twice the minimum and maximum projected radii, respectively, as predicted using the geometry calculator in Marvin Sketch.

Cytochrome c activity and redox potential

Luminol oxidation activity

Preassembled DNA nanocages were incubated with either 3x excess PNA-CytC, 5x molar excess of free cytochrome c or 5x molar excess of PNA-FAM for 30 minutes at room temperature and cooled to 4 ^{0}C for 30 additional minutes. Excess PNA-CytC was removed from the DNA-PNA-CytC sample by SE-FPLC as described in the PNA-protein assembly methods section. A 15% native PAGE (resolving gel) was cast leaving 1–2 mm below the bottom edge of the comb. After polymerization a 4% native PAGE (stacking gel) was poured on top of the resolving gel. 100 pmol of each sample was loaded on the gel (cooled to 4 ^{0}C) and run for 2–3 hours at 200 V until the DNA nanocage complexes had entered the resolving gel; the progress of the gel was monitored using the FAM fluorescence from the DNA-PNA-FAM construct in lane 4. The gel was trimmed to fit in a small plastic container, soaked for 5 minutes in 7–10 mL of SuperSignal West Femto enhanced luminol chemluminescence solution (Thermo Pierce), washed briefly for 10 sec and then imaged using a UVP BioSpectrum (Upland, CA) equipped with an Ascent 4000C 4.2 million pixel color CCD camera. The image was captured every 20 minutes

for 2 hours using Dynamic Integration mode. A 32 square pixel area in the band of interest, free of any detector noise spots, was integrated to determine the mean area density for the DNA + PNA-CytC sample and the DNA + free cytochrome c control band. The mean area density of the same size, nearby spot without any stain on ethidium bromide image was subtracted to remove the background. The background corrected average mean density for each band was plotted as a function of integration time, as shown in Figure 6C. The DNA + PNA-CytC data was fit using OriginPro version 9.1.0 with the following exponential function: $y(t) = y_0 - A * e^{-R_0 * t}$where y_0 is the upper limit of 56,160 counts, A is the pre-exponential factor of 56,034 counts, R_0 is the rate of decay of 0.011 counts per second and t is the time in seconds. The gel was subsequently washed for 10 minutes, stained with ethidium bromide for 5 minutes, destained for 5 minutes in nanopure water and visualized on a UVP Biospectrum as described previously.[79]

Redox potential titration with ferri/ferrocyanide

The method described by Craig and Nichols [112] was modified to work with smaller volumes and lower concentrations. The DNA-PNA-CytC complex was prepared as described in the DNA nanocage assembly section, by incubating the DNA nanocage with 3x molar excess of PNA-CytC. Free cytochrome c and excess PNA-CytC and were removed by SE-FPLC at 4 °C into 10 mM phosphate buffer at pH 7.0 containing 15 mM magnesium acetate. Equine heart cytochrome c (Sigma) was dissolved in the same buffer and used as a control in a separate experiment. Stock concentrations of 1.875 mM ferricyanide, 100 mM ferrocyanide and 100 mM dithionite were prepared just prior to use in 10 mM phosphate buffer at pH 7.0 containing 15 mM magnesium acetate and degassed

by vacuum to remove oxygen. DNA-PNA-CytC was diluted until the cytochrome c concentration was about 10 μM and degassed by vacuum to remove oxygen. Then three separate 13.5 μL aliquots were diluted with 1.5 μL of the ferricyanide, ferrocyanide or dithionite (1:10 dilution) stocks solutions to maintain the same DNA-PNA-CytC concentration for each aliquot. The same process was used for the equine heart cytochrome c as a control. Absorbance scans of each sample were measured at room temperature (free cytochrome c) or 6 ^{0}C (DNA-PNA-CytC and free cytochrome c) using an ultra-micro B0631081 10 μL, 10 mm cuvette (Perkin Elmer) on a Lamda XLS+ spectrophotometer (Perkin Elmer) to determine the heme absorbance at 550 nm of the fully reduced cytochrome c in dithionite and fully oxidized cytochrome c in ferricyanide, as shown in Figure 2-22. The cuvette reservoir volume was increased to accommodate the ~30 μL of total volume added during the experiment by cutting off about 1 cm of the ends of two 1000 μL pipet tips and inserting them into the entry and exit holes in the cuvette. Next, a stepwise oxidation of the sample was performed by adding six separate 375 pmol (1 μL) aliquots of the sample in ferricyanide directly into the cuvette containing the sample in ferrocyanide and mixed with a pipet, followed by 3 separate 750 pmol (2 μL) aliquots, where the heme absorbance was measured after each step. For the DNA-PNA-CytC complex, which showed a lower redox potential, the concentration of ferricyanide added was reduced by 4x, by using half of the stock ferricyanide concentration (1.875 mM) and adding aliquots of half the volume of sample in ferricyanide (95.75 pmol or 0.5 μL and 187.5 pmol or 1 μL). A UV-Vis scan of the 10 mM solution of ferrocyanide used during the experiment was found to contain a small fraction (0.163%) already oxidized to the

ferricyanide form. This resulted in a baseline ferricyanide concentration of 16.3 μM before aliquots of sample in ferricyanide were added.

The collected data was analyzed as follows. The absorbance at 550 nm for each step was normalized to the change in absorbance from the fully reduced state (sample in dithionite normalized to 1.0) to the fully oxidized state (sample in ferricyanide normalized to 0.0) to determine the fraction of reduced (ferroCytC) and oxidized (ferriCytC) cytochrome c using the following equation: $Fract_{ferroCytC} = \frac{A_{sample,550} - A_{ferriCytC,550}}{A_{ferroCytC,550} - A_{ferriCytC,550}}$ and $Fract_{ferriCytC} = 1 - Fract_{ferroCytC}$, where $A_{ferriCytC,550}$ is the cytochrome c absorbance at 550 nm in the fully oxidized state (i.e., in ferricyanide), and $A_{ferroCytC,550}$ is the cytochrome c absorbance at 550 nm in the fully reduced state (i.e., in dithionite) and $A_{sample,550}$ is the cytochrome c absorbance in ferrocyanide after each aliquot of sample containing ferricyanide was added. Due to small variations in concentration between samples, the $A_{ferroCytC,550}$ value in the fully reduced state was determined from the measured $A_{ferriCytC,550}$ value as follows: $A_{ferroCytC,550} = \frac{A_{ferriCytC,550}}{22.5\%}$, where 22.5% was the average measured fraction of absorbance in the fully oxidized state with respect to the fully reduced state. For the DNA-PNA-CytC complex, which showed a lower redox potential, the sample became fully oxidized during the titration, so the value of $A_{ferroCytC,550}$ at the end of the redox titration was used for $A_{ferriCytC,550}$. Data points with an $A_{sample,550} - A_{ferriCytC,550}$ less than 1 mAU were not included because they were within the noise of the spectrophotometer (±0.3 mAU), which were generally data collected with a $Fract_{ferroCytC}$. < 1%.

For each step, $\log\frac{Fract_{ferroCytC}}{Fract_{ferriCytC}}$ was plotted against $\log\frac{Fract_{ferroCN}}{Fract_{ferriCN}}$ along with a linear fit of the data, as shown in Figure 7. The following equation was used to fit each dataset using Microsoft excel 2010: $y = mx - b$,where m is slope, and b is the y intercept. Only datasets with linear fits with $R^2 > 0.8$ were included. The x intercept of the fit was used with the Nernst equation to calculate the redox potential of cytochrome c as follows: $E^0_{CytC} = E^0_{FerriCN} - \frac{RT}{zF}x_{int}$, where $E^0_{FerriCN}$ is the redox potential of ferricyanide of 430 mV,[113] R is the ideal gas constant of 8.314 $\frac{J}{mol*k}$, T is the temperature of the experiment in kelvin, z is one for the number of electrons involved in the oxidation-reduction reaction, F is Faraday's constant of 96485 $\frac{C}{mol}$, and x_{int} is the x intercept obtained from the fit. Error bars were calculated as the standard error obtained 2 measurements.

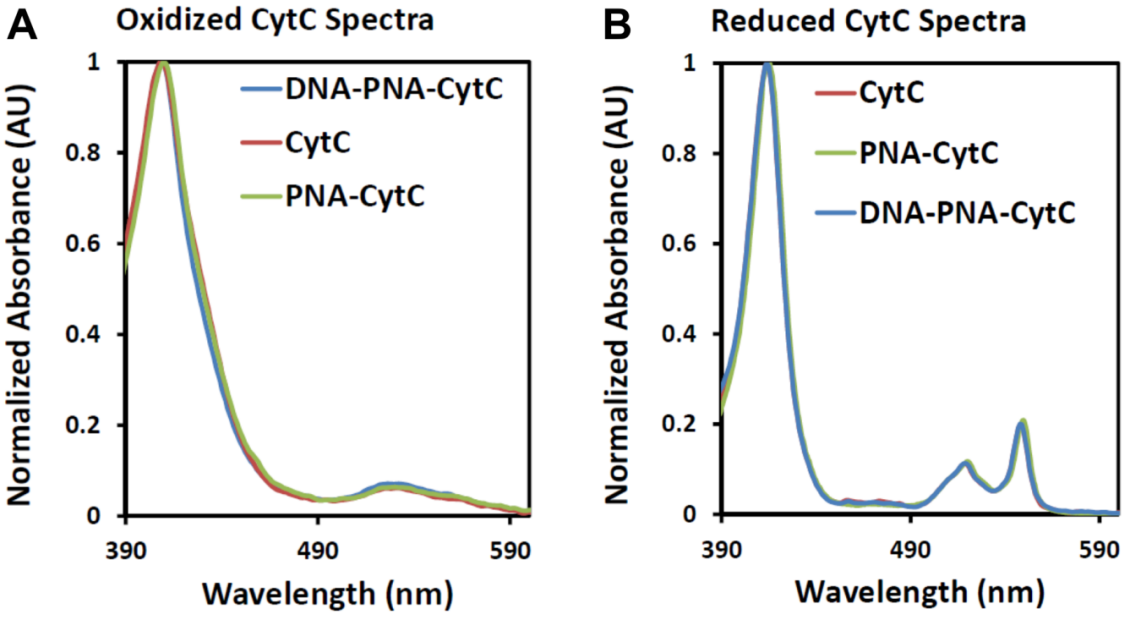


Figure 2-22. Visible spectra of cytochrome c, PNA-CytC and DNA-PNA-CytC. Visible spectra of cytochrome c (red), PNA-CytC (green) and DNA-PNA-CytC (blue) when oxidized with ferricyanide (A) or reduced with dithionite (B).

Circular dichroism of cytochrome c

Circular dichoism spectra were acquired using a JASCO J-710 spectropolarimeter in the UV region (190–250 nm) and Soret band (400–430 nm). All data were collected at 11 ^{0}C using the Peltier cooling accessory to ensure complete binding of the DNA-PNA-CytC complex, and the remaining samples for consistency. UV measurements were made in a Far-UV quartz cuvette with a 1 mm path length (Starna Cells 21-Q-1). Due to low sample concentration (~10 μM), Soret measurements were made in a Far-UV quartz cuvette with a 1 mm path length (Starna Cells 16.100-Q-10/Z8.5). The following scan parameters were used for the UV and Soret, respectively: Rate (50 nm/min, 20 nm/min), pitch (0.5 nm, 0.2 nm), bandwidth (1 nm, 2 nm), response time (4 sec). Each sample was mixed with 5% (v/v) 375 μM ferricyanide to oxidize the heme active site prior to data collection. The following samples were exchanged into 10 mM potassium phosphate buffer pH 7.0 and measured at the Soret and UV: CytC, PNA-CytC, buffer with ferricyanide. DNA-PNA-CytC was exchanged into sodium phosphate buffer pH 7.0 containing 10 mM magnesium acetate to stabilize the DNA nanocage and measured at the Soret band only, because the high absorbance from the DNA nanocage in the UV masked the features of the protein. A scan of the buffer alone containing ferricyanide was collected and used as a baseline to subtract from the data. All data for samples and controls were collected as an average of 10 scans. After baseline subtraction, the molar ellipticity [θ] of the UV and Soret data was calculated[114] as follows: $[\theta] = \frac{100*\theta}{C*l}$, where θ is the ellipticity in degrees recorded by the instrument, *C* is the molar sample concentration and *l* is the path length in centimeters. The results are shown in Figure 2-23.

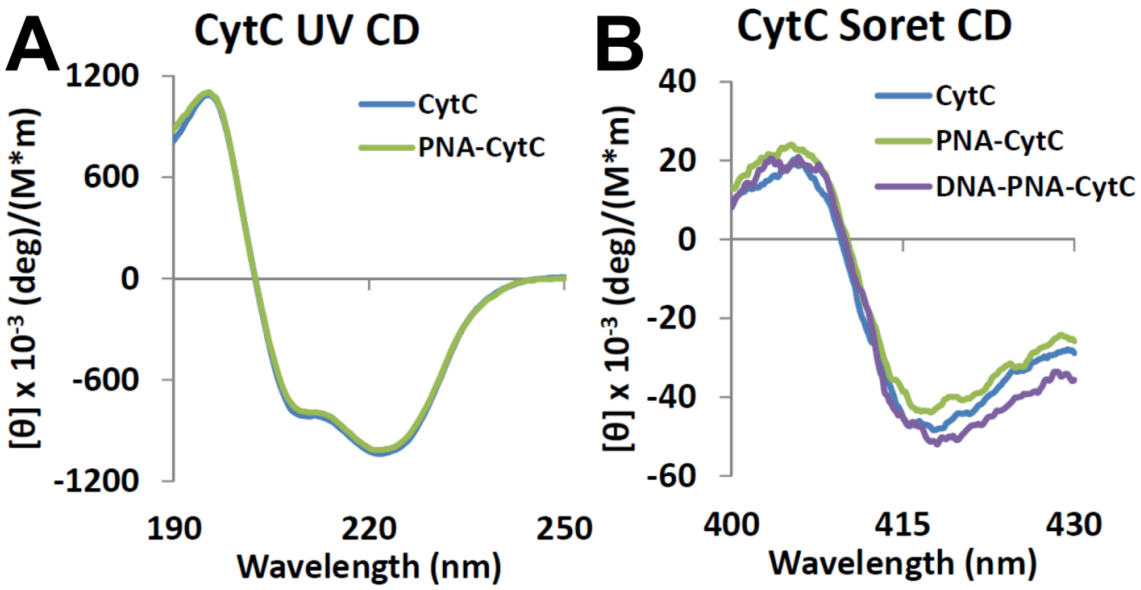


Figure 2-23. Circular dichroism of cytochrome c, PNA-CytC and DNA-PNA-CytC. UV [A] and Soret band [B] CD spectra of cytochrome c (CytC), when conjugated to PNA (PNA-CytC) and when hybridized into the DNA nanocage (DNA-PNA-CytC). Each spectrum is an average of 10 scans. Note: The large absorbance from the DNA nanocage in the UV prevented CD spectra for DNA-PNA-CytC in the UV region.

CHAPTER 3 – PNA-Peptide Assembly in a 3D DNA nanocage at Room Temperature

Adapted with permission from Flory, J. D.; Shinde, S.; Lin, S.; Liu, Y.; Yan, H.; Ghirlanda, G.; Fromme, P. *J. Am. Chem. Soc.* **2013**, *135*, 6985–6993.[79]

Abstract

Proteins and peptides fold into dynamic structures that access a broad functional landscape, however, designing artificial polypeptide systems is still a great challenge. Conversely, DNA engineering is now routinely used to build a wide variety of 2D and 3D nanostructures from hybridization based rules, and their functional diversity can be significantly expanded through site specific incorporation of the appropriate guest molecules. Here we demonstrate a new approach to rationally design 3D nucleic acid–amino acid complexes using Peptide Nucleic Acid (PNA) to assemble peptides inside a 3D DNA nanocage. The PNA-peptides were found to bind to the preassembled DNA nanocage in 5–10 minutes at room temperature and assembly could be performed in a stepwise fashion. Biophysical characterization of the DNA-PNA-peptide complex was performed using gel electrophoresis as well as steady state and time resolved fluorescence spectroscopy. Based on these results we have developed a model for the arrangement of the PNA-peptides inside the DNA nanocage. This work demonstrates a flexible new approach to leverage rationally designed nucleic acid (DNA-PNA) nanoscaffolds to guide polypeptide engineering.

Introduction

The field of structural DNA nanotechnology has transformed our thinking about the purpose and possibilities of canonical base pairing biopolymers, which can have a much broader structural landscape than what has evolved in natural organisms. The sequences of multiple DNA strands have been designed and assembled into a wide variety of interesting nanostructures, such as a 2D crystal lattice,[9] 3D tetrahedron,[65] 3D crystal,[20] octahedron,[115] various fully addressable 2D nanoscaffolds[21] and complex 3D shapes.[25] We refer the reader to several excellent reviews of the progress made in this dynamic field and challenges that remain.[13–15]

Because the functional diversity of unmodified DNA is limited, much effort has been focused on introducing functional groups. DNA scaffolds offer a flexible, biocompatible workbench to explore the wide variety of functions performed by proteins in nature, such as catalysis, scaffolding, molecular transport and molecular recognition among others. For example, DNA scaffolds have been used to organize proteins on 2D arrays,[8] to study antibody recognition of peptide epitopes[39] and to investigate the spatial dependence of enzyme cascades.[48] Remarkably nearly all of the biomolecules found in the biosphere are synthesized through unique metabolic pathways of reactions catalyzed by proteins. Some large proteins, such as the 350 kDa protein photosystem II (PSII), perform multiple functions including light absorption, photochemical conversion, electron transfer and catalytic water oxidation.[3] While light absorption and most of the electron transfer reaction take place in the membrane intrinsic part of PSII, the catalytic site for water oxidation is located in the membrane extrinsic part of PSII suggesting that a water soluble version of the water splitting complex is possible. The crystal structure of PSII[3] shows

that coordinating sphere around the metal cluster in the active site is only a small part of the protein, but that the amino acid ligands come from multiple protein subunits. We hypothesize that the scaffolding function of a protein could be replaced by strategically attaching peptides inside the well-defined cavity of a 3D nucleic acid nanoscaffold, which matches the size of the first and second coordination sphere of the catalytic center (<10 nm).

Different bio-orthogonal linking chemistries are required to control peptide assembly within a highly interconnected nucleic acid-polypeptide complex. Erben, Goodman and Turberfield coupled the protein cytochrome c to one of the strands of the same DNA tetrahedron used in our design and annealed the DNA-protein complex in one step.[50] While this strategy allows the same bio-orthogonal linking chemistry to be used on multiple DNA strands, it requires that all components be assembled at once and under conditions that favor proper DNA nanocage assembly, namely low concentration (sub micromolar) and high temperature annealing (optimally 90 °C ramped to room temperature). An alternative linking strategy that allows more control over the assembly process is to use nucleobase hybridization, where the nucleobase sequence is used to provide orthogonal attachment. This approach is routinely used to functionalize large DNA nanostructures, such as DNA origami,[29] because functionalization can be done after the DNA nanostructure is formed. To allow for sufficient thermal stability at room temperature, typical DNA hybridization linker sequences are at least 15–20 nucleotides (nt) long, which form 4.9–6.5nm long dsDNA helices upon hybridization. However, such long linker sequences put significant constraints on how the linker can be used in small DNA nanostructures (radius <4 nm). The linker length could be shortened by increasing the GC

content of the sequence, however this would limit the sequence variability needed to attach multiple peptides to the DNA nanostructure.

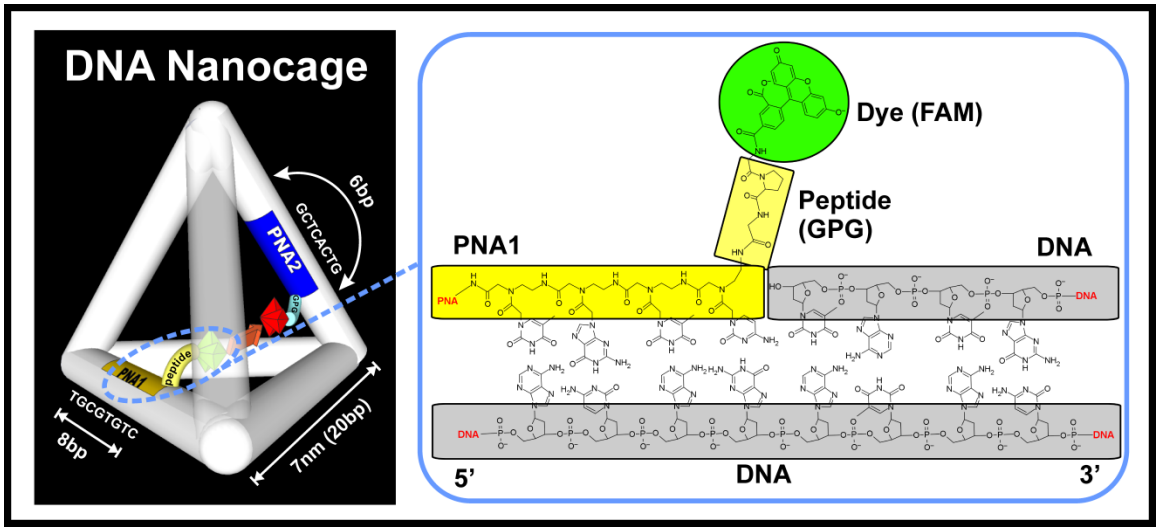


Figure 3-1. Schematic of the designed dye labeled DNA-PNA-peptide complex. [Left Figure] Cartoon of the proposed design for assembling two fluorescently labeled peptides into a DNA nanocage each using a PNA linker with the indicated sequences. The arrow indicates the direction of energy transfer between the fluorescein (FAM) donor and the tetramethylrhodamine (TMR) acceptor dyes. [Right Figure] Close up showing the chemical structures of the N terminal (5') end of the dye labeled PNA1-peptide, the 3' end of the adjacent DNA strand and the complementary DNA strand that drives the assembly of this peptide into the DNA nanocage. Chemical structures were drawn using Marvin Sketch.[111]

Our approach, as shown in the schematic in Figure 3-1 is to connect peptides to the DNA scaffold using short (8nt) linker sequences of the more thermostable Peptide Nucleic Acid (PNA) originally developed by Nielsen and coworkers.[66] PNA synthesis is compatible with peptide synthesis, so the polypeptides can be synthesized directly with the linking PNA sequences using established protocols.[83] There is a great interest in using PNA for gene targeting[116] and sensing[117,118] applications because of PNAs hypersensitivity to sequence mismatches,[119] its ability to invade double stranded DNA sequences[66] and its resistance to enzymatic degradation.[73] PNA has also been investigated as a nanoscaf-

fold. First it was incorporated into a two dimensional DNA array, resulting in only a slight unwinding of the helix,[78] and more recently used to arrange peptide ligands along a linear strand of DNA to assemble several protein subunits.[47]

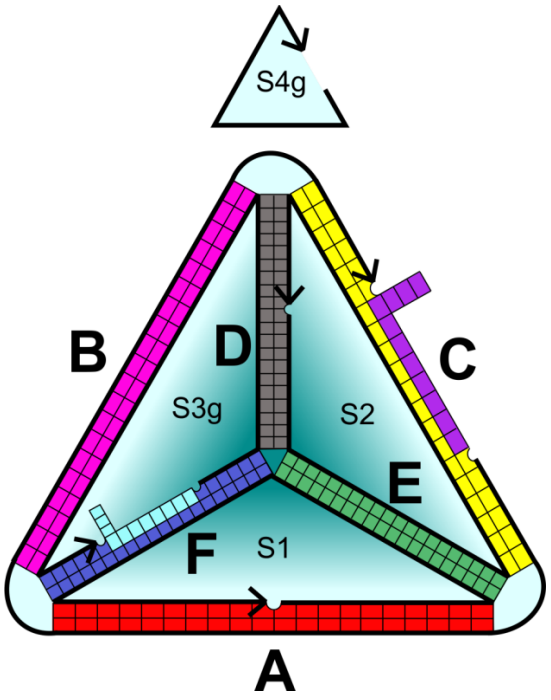
Here we report results for a design strategy using PNA to assemble peptides inside a 3D DNA nanostructure. Two fluorescently labeled PNA-peptides were synthesized using microwave assisted solid phase synthesis. The PNA-peptides were assembled into the DNA nanostructure in 5–10 minutes at room temperature. The peptides can be either assembled together or in a specified order. Steady state fluorescence was used to monitor Förster Resonance Energy Transfer (FRET) between a single dye labeled PNA-peptide bound to the DNA nanocage with a complementary dye label, in order to determine the temperature of PNA dissociation out of the DNA nanocage. Time resolved fluorescence spectroscopy was used to identify different interactions and decay pathways experienced by each fluorescent dye in the DNA-PNA-peptide complex. Finally, FRET theory was used to calculate the distances between the two PNA-peptides bound to the DNA nanocage and build a model of the designed complex. In the future this design principle could lead to the assembly of enzyme active centers inside DNA nanocages.

Results and Discussion

DNA-PNA-peptide complex design

Our design is based on a DNA tetrahedron assembled out of four 63nt ssDNA sequences originally designed by Goodman, Turberfield and coworkers.[65] The sequences of each strand are unique and fully addressable for incorporating guest molecules through hybridization based linkers such as PNA (Figure 3-2). The authors were able to identify

the location of the major and minor grooves and approximate base orientations through a series of DNA linking[65] and protein encapsulation experiments,[50] along with electron density from cryo-electron microscopy.[64] We inferred from their study that base positions around the 8th and 18th bases point inward and around the 3rd and 14th base point outward when counting from the vertex in the 3' direction. We introduced an 8nt single stranded gap in the center of opposite 20 base pair (bp) edges of the DNA nanocage to bind two PNA strands each containing a short prototype peptide (Gly-Pro-Gly) protruding from the N terminus (5') of the PNA sequence into the DNA nanocage at the 7th base position (Figure 3-1). Because of the structural symmetry of the DNA nanocage, the same strategy can be applied to any edge in the DNA nanocage, thereby allowing a number of different peptides to be assembled at sequence specific locations. The 8nt length of each PNA binding domain was chosen to provide sufficient thermal stability of PNA binding to the DNA nanocage with minimal disruption of the DNA nanocage structure in absence of the PNA-peptide. Standard dsDNA of B-DNA form has a helical pitch of 10.5bp/turn[120,121] whereas PNA-DNA duplexes have been observed to adopt a helical pitch from 13.0bp/turn[122] to 15.6bp/turn[78]. For the 8bp PNA-DNA domain in our design this leads to an underwinding of the helix by 53 to 90 degrees, respectively, along the edges that bind PNA. However, the original DNA nanocage design[65] included unhybridized adenosine nucleotides connecting each edge of the DNA nanocage, which should be able to counter rotate and neutralize this effect. Throughout this report we use the nomenclature defined in Figure 3-2 and Figure 3-3 when we describe the various strands and constructs used in each experiment. Briefly, PNA1 (TGCGTGTC) and PNA2 (GCTCACTG) are both PNA sequences written from C to N (3' to 5'), abbreviated as PNA1 and PNA2, respectively,



S1 - AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCACCCGCCATAGTAGACGTATCACC

S1P AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCCGAGTGACTAGTAGACGTATCACC

S1P-FAM AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCCGAGTGACTAGTAG/iFluorT/CGTATCACC

S2 - CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGAGGGTCCAATACCGACGATTACAG

S2P CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGACGCACAGATACCGACGATTACAG

S2P-FAM CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGACGCACAGATACCG/iFluorT/CGATTACAG

S2P-TMR CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGACGCACAGATACCG/i6-TAMN/CGATTACAG

S3g GATAAAACGTGTAGCAAGCTGTAATCGACGGGAAGAGCATGCCCATCCACTACTA

PNA2-peptide 5-TMR-Gly-Pro-Gly-GTCACTCG

S4g CGCATGACTCAACTGCCTGGTGATACGAGGATGGGCATGCTCTTCCCGACGGTAT

PNA1-peptide 5-FAM-Gly-Pro-Gly-CTGTGCGT

Figure 3-2. DNA and PNA sequences and schematic showing their arrangement in the complex. [Top Figure] Schematic of the PNA and DNA sequence layout, where the PNA1-peptide is attached to Edge C and the PNA2-peptide is attached across the DNA nanocage to Edge F. The strand nicks are indicated and the arrow is pointing in the 3' direction. Strand S4g is redrawn above for clarity. This schematic was redrawn from [65] with changes to incorporate the PNA-peptides. [Bottom Figure] DNA, PNA and peptide sequences with the color code matching the schematic in the top figure. The DNA and PNA sequences are written in the 5' to 3' direction. The peptides sequences are written using three letter amino acid codes from N to C. Dye labeled DNA strands are defined using the nomenclature for ordering from IDT, where the dye names are enclosed by "/" marks and are attached to a modified thymine base at the vertex.

DNA + PNA1-FAM = S1, S2P, S3, S4g, PNA1-GPG-FAM

DNA + PNA1-TMR = S1, S2P, S3, S4g, PNA1-GPG-TMR

DNA-TMR + PNA1-FAM = S1, S2P-TMR, S3, S4g, PNA1-GPG-FAM

DNA-FAM = S1, S2P-FAM, S3, S4g

DNA-FAM + PNA1-TMR = S1, S2P-FAM, S3, S4g, PNA1-GPG-TMR

DNA + PNA1-FAM + PNA2-TMR = S1P, S2P, S3g, S4g, PNA1-GPG-FAM, PNA2-GPG-TMR

DNA + PNA1-FAM + PNA2-Ac = S1P, S2P, S3g, S4g, PNA1-GPG-FAM, PNA2-GPG-Acetylated

DNA + PNA1-Ac + PNA2-TMR = S1P, S2P, S3g, S4g, PNA1-GPG-Ac, PNA2-GPG-TMR

DNA + PNA1-Ac + PNA2-Ac = S1P, S2P, S3g, S4g, PNA1-GPG-Ac, PNA2-GPG-Acetylated

DNA-FAM-TMR = S1P-FAM, S2P-TMR, S3g, S4g (No PNA)

Figure 3-3. Complex nomenclature with list of constituent DNA and PNA-peptide strand names.

GPG (Gly-Pro-Gly) is the peptide sequence and FAM (Fluorescein) and TMR (Tetra-methylrhodamine) are the dye labels.

The PNA1 sequence was slightly changed from the original DNA nanocage design[65] and the complementary S2 sequence was modified accordingly on the yellow edge to S2P (changes shaded in grey). The PNA2 sequence was also updated from the original DNA nanocage design in order to eliminate a stretch of 5 of 6 guanine bases (three consecutive) which proved to be more difficult to synthesize. To accommodate binding of the modified PNA2 sequence, the complementary S1 sequence was modified on the yellow edge to S1P (changes shaded in grey). S4 and S3 were also modified to S4g and S3g, respectively, by trimming back 8nt from the ends to create an 8nt gap, exposing an 8nt

ssDNA domain on the complementary strands, S2P and S1P, respectively, for binding PNA1 and PNA2, respectively.

PNA-peptide synthesis

We synthesized two dye labeled PNA-peptides, PNA1-GPG-FAM and PNA2-GPG-TMR, using microwave assisted solid phase synthesis following established protocols[67,83,84] with some modifications detailed in the Materials and Methods section called PNA-peptide synthesis, purification and functionalization. Unlabeled versions of each PNA-peptide were also prepared to serve as controls for fluorescence experiments by acetylating the N terminus. The PNA-peptide constructs were purified using Reverse-Phase High Pressure Liquid Chromatography (RP-HPLC) and identified using Matrix Assisted Laser Desorption Ionization (MALDI) mass spectrometry. MALDI and RP-HPLC chromatograms are shown for all products in Figure 2-3 and Figure 2-4, respectively.

Peptide Nucleic Acid (PNA) was originally designed by Nielsen and coworkers to link nucleobases together via amide bonds with the proper spacing in its aminoethylglycine backbone to hybridize with DNA.[119] The unique PNA backbone provides it with many an ability to invade double stranded DNA sequences,[66] resistance to enzymatic degradation[73] as well as compatibility with peptide synthesis. PNA monomers are now commercially available with Fmoc protection so they can be sequentially assembled following protocols compatible with milder solid phase peptide synthesis (SPPS) conditions.[67,83]

Although substantial similarity exists between the coupling mechanisms of consecutive PNAs and amino acids, PNA requires considerably more care during synthesis.

PNA coupling is less efficient than amino acid coupling and thus requires additional time or heat for efficient coupling. PNA also contains a carbonyl linkage to the nucleobase 'side chain' that is subject to nucleophilic attack by its free N terminus after deprotection, which can lead to nucleobase migration and isomerization.[123] We followed a recently published protocol for microwave assisted PNA synthesis, which emphasized the importance of turning off microwave heating during deprotection and capping steps to minimize side reactions.[83] Our initial synthesis of PNA1 showed an impurity associated with the recommended PyBOP activator, which we found was reported elsewhere,[84] so we opted to use HATU and anhydrous DMF as reported in many earlier PNA synthesis protocols.[67,84]

Microwave assisted solid phase synthesis was used to first synthesize the PNA domains with an N terminal glycine residue, using conditions optimized for PNA. During PNA2-Gly synthesis, the cytosine solution partially precipitated which led to an impurity with a deleted cytosine and another with an inserted cytosine as identified by MALDI in Figure 2-3. RP-HPLC showed these impurities were minor, as shown in Figure 2-4. The same synthesizer was used to add the remaining amino acids for the peptide using conditions optimized for peptides. Initially a capping step was performed after each PNA coupling to facilitate purification, but we found that without capping the yield was much improved and the removal of impurities for these short sequences was straightforward. Finally, the PNA-peptide conjugates were fluorescently labeled on resin with either 5(6)carboxyfluorescein (FAM) or 5(6)tetramethylrhodamine (TMR), or the free amino terminus was acetylated as a control for fluorescence studies. The 5(6) dye isomer mixtures were used due to substantially reduced cost with negligible spectral differences.

MALDI of the PNA-peptides identified a minor impurity with an extra glycine (+57 Da) in the peptide sequence, as shown in Figure 2-3. Each product was purified using a C18 prep column. The product was identified by MALDI and the purity verified by RP-HPLC using a C18 analytical column, as shown in Figure 2-3 and Figure 2-4, respectively.

PNA-peptide assembly with the DNA nanocage

PNA-DNA duplexes are typically formed from single stranded DNA and PNA by heating both strands to around 90 °C to remove any secondary structure or sample aggregation, followed by cooling to room temperature or below.[124] Alternatively, short PNA sequences have been shown to bind to their target DNA sequence after incubating for 5–10 minutes at room temperature.[75] This incubation strategy is especially important for hybridizing PNA-peptides to 3D DNA nanostructures, because assembling peptides inside a preformed DNA nanocage avoids the high temperature annealing step, which could denature the polypeptide structure. Furthermore the incubation strategy allows for higher PNA-peptide concentrations to drive more rapid assembly and lets the PNA-peptides be assembled in a desired order.

In order to determine the PNA-peptide hybridization efficiency, the DNA nanocage was prepared (without any PNA) by annealing and purified by Size Exclusion Fast Protein Liquid Chromatography (SE-FPLC), as shown in Figure 2-13 (A). The preformed DNA nanocage was incubated for 5–10 minutes at room temperature with increasing molar excess (0x to 5x) of PNA1-GPG labeled with FAM and separately with PNA1-GPG labeled with TMR. The samples were cooled to 4 °C and were directly analyzed using native polyacrylamide gel electrophoresis (PAGE) at 4 °C. The bands of in-

terest are shown in Figure 3-4 and the complete gel is shown in Figure 2-15. Both constructs exhibit a band shift upon PNA binding which could be due in part to the DNA nanocage becoming less spherical and more tetrahedral in shape with a slightly larger diameter after PNA binding. PNA1-GPG-FAM shows the most dramatic shift, which could be further enhanced from the FAM^{2-} dianion being repelled away from the negatively charged DNA nanocage, forcing the attached peptide to adopt an orientation pointing outward and further retarding migration of the complex through the gel. The chemical structures of FAM and TMR are shown in Figure 2-2.

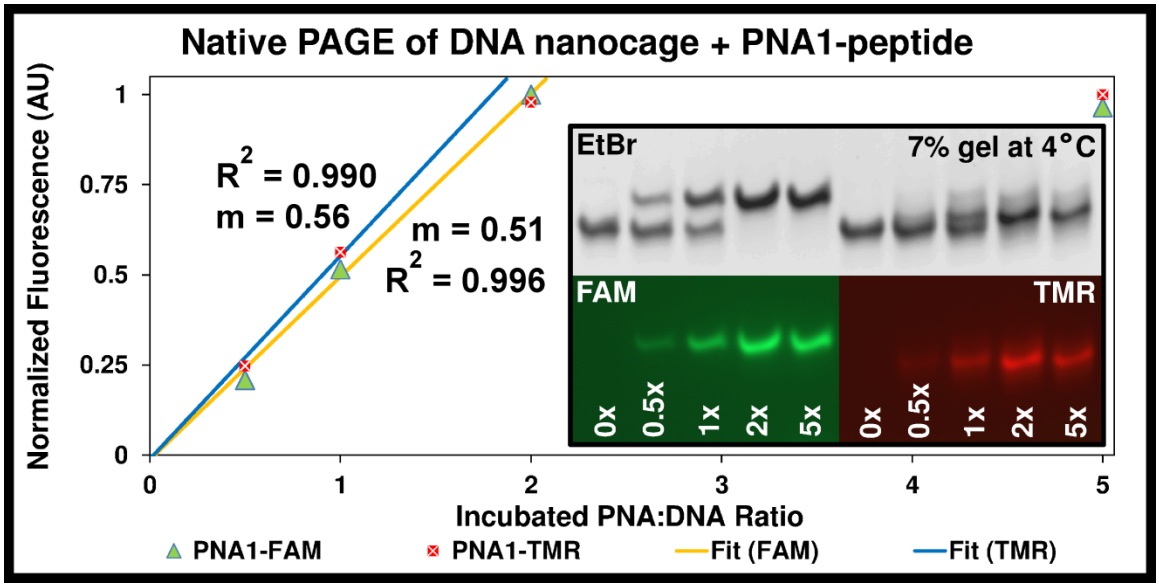


Figure 3-4. Native PAGE of the DNA nanocage hybridized to increasing molar excess of the PNA1-peptide labeled with the fluorescent dyes FAM and TMR. The DNA nanocage was incubated at room temperature for 5–10 minutes with increasing molar excess of PNA1-peptide labeled with TMR and with FAM. The upper image shows EtBr fluorescence (Ex. 302 nm, Em. 605 nm) after staining while the lower images were taken of the native dye fluorescence (FAM Ex. 475 nm, Em. 535 nm; TMR Ex. 560 nm, Em. 645 nm) prior to EtBr staining. The normalized integrated fluorescence intensities of the gel bands are shown plotted versus the incubated PNA to DNA ratio, where the initial hybridization fit a linear model. The slope m represents the fraction of PNA that binds to the DNA nanocage.

The gel in Figure 3-4 also shows that PNA1-GPG-TMR induces some structural heterogeneity upon binding. In addition to a distinct band showing TMR fluorescence, some minor band broadening is apparent that extends from the initial position of DNA nanocage without any PNA up to the PNA1-GPG-FAM band, suggesting the TMR peptide stabilizes a range of slightly different DNA nanocage conformations. Figure 3-4 also shows the ratio of PNA bound to the DNA nanocage calculated from the integrated band intensities, which increases linearly with the added molar excess until saturation is achieved by two times excess. The PNA1-GPG hybridization was further confirmed using absorption spectroscopy by comparing the absorption of the dye labeled PNA to that of the DNA nanocage after removing the unhybridized PNA. The PNA-peptide hybridization increases linearly until PNA-peptide saturation occurs just after two times excess for PNA1-GPG-TMR (extrapolated to 2.4x) and for PNA1-GPG-FAM (extrapolated to 2.5x), as shown in Figure 3-5.

The need for excess PNA-peptide to achieve quantitative binding to the DNA nanocage is likely because a fraction of the PNA-peptides irreversibly self-aggregate before binding to the DNA nanocage. Some PNA sequences show mild aggregation at concentrations as low as 1 μM and are often modified to improve solubility.[80] However, once bound to the DNA nanocage the PNA-peptides are quite stable at low temperatures, even once excess PNA-peptides are removed. To ensure quantitative PNA-peptide binding to the DNA nanocage, five times molar excess was used to prepare the constructs for all subsequent experiments. Samples were stored at -20 °C or 4 °C and experiments were carried out between 4 °C and 11 °C to prevent PNA-peptide dissociation.

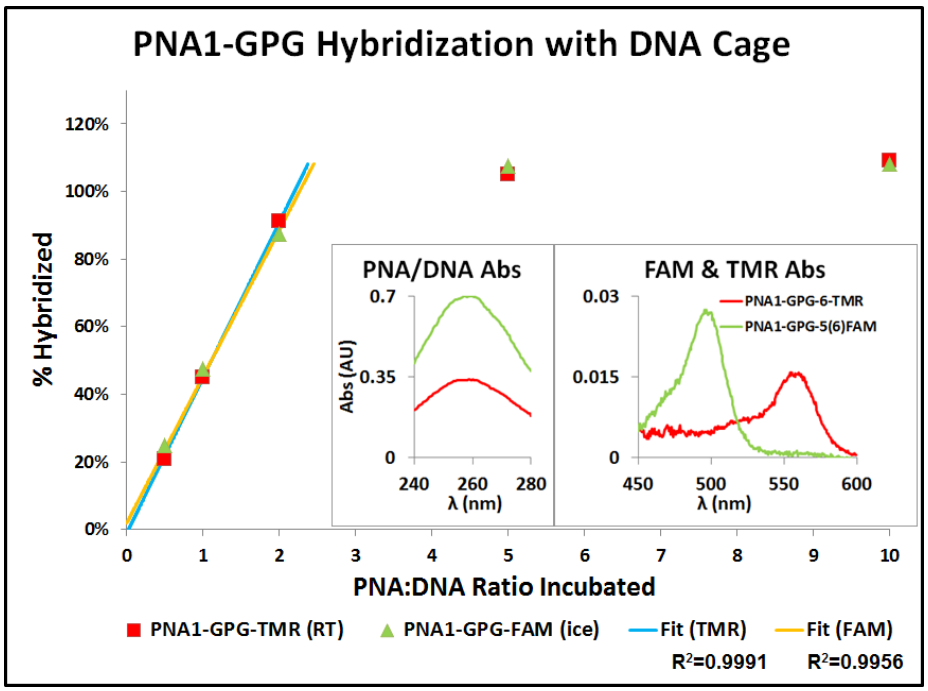


Figure 3-5. Absorbance of the labeled PNA1-peptide hybridized to the DNA nanocage. Graph showing the amount of PNA-peptide hybridized to the DNA nanocage plotted against the molar excess of PNA-peptide incubated with the DNA nanocage. The hybridization efficiency was calculated as $(A_{dye}/A_{258}) / (\varepsilon_{dye} / \varepsilon_{258})$, for each dye as described in the Materials and Methods section called PNA-peptide hybridization efficiency. The fitting parameters for the linear model are as follows: Fit (TMR): slope = 0.46, intercept = -0.01; Fit (FAM): slope = 0.43, intercept = 0.02, where the slope indicates the fraction of PNA-peptide that binds to the DNA nanocage. The inset shows the measured absorbance spectra of the DNA nanocage with PNA1-GPG-FAM (green) and the DNA nanocage with PNA1-GPG-TMR (red). The difference in absolute absorption between the samples was due to differences in concentrations (PNA1-GPG-FAM = 3–4 μM, PNA1-GPG-TMR = 2 μM).

A time course study was performed in order to determine the kinetics of PNA hybridization to the 3D DNA nanocage. Fluorescently labeled DNA nanocages were prepared using one strand labeled by the manufacturer (Integrated DNA Technologies) with a dye complementary to that on the PNA-peptide and located at the vertex adjacent to the PNA-peptide's N terminus. The fluorescence intensity of the FAM labeled DNA nanocage (DNA-FAM) was monitored as a function of time after mixing in five times molar excess PNA1-GPG-TMR (PNA1-TMR) with the DNA nanocage at 25 °C, as

shown in Figure 3-6 (A) The data from Figure 3-6 (A) was converted to the fraction of PNA-peptide bound as described in the methods section and plotted in Figure 3-6 (B). 86% of the PNA bound while mixing the PNA during the first 15 seconds, 94% after 1 minute and 99% after 6.8 minutes. Such rapid hybridization of short PNA sequences to DNA is in agreement with another kinetic study of PNA-DNA duplex formation.[75]

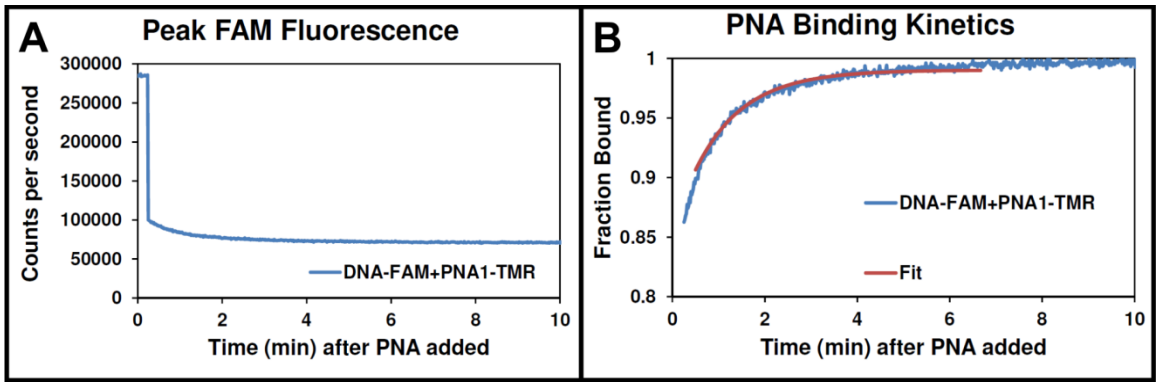


Figure 3-6. Time course PNA1-peptide hybridization with the DNA nanocage. 5x excess PNA1-TMR was added to the DNA nanocage labeled with FAM at 25 °C. [A] Plot of the peak FAM fluorescence at 521 nm as a function of time after the PNA1-TMR was added. [B] The data from [A] was normalized and inverted to show the fraction of PNA1-TMR bound as a function of time. The exponential fit was produced as described in the methods section and the fitting parameters are shown in Table 3-1. No data was collected for the first 15 seconds after PNA1-TMR was added and mixed with a pipet.

DNA nanocages were also prepared with both dye labeled PNA-peptides for fluorescence characterization. Five times molar excess of the each PNA-peptide was incubated at room temperature for 5–10 minutes and with the preformed DNA nanocage then chilled on ice for 1 hour. The unhybridized PNA-peptides were removed by SE-FPLC with the column cooled to 4 °C. The chromatogram of the fully assembled DNA nanocage with both dye labeled PNA-peptides (DNA + PNA1-FAM + PNA2-TMR) is shown in Figure 2-13 (B). The PNA-peptides could also be assembled in stepwise fashion by first forming the DNA nanocage with one PNA-peptide, removing excess PNA-

peptide using 30k MWCO centrifugal filters, adding the second PNA-peptide and removing excess PNA-peptide in the same fashion. This strategy produced the same end product as when both PNA-peptides were assembled simultaneously (data not shown).

Table 3-1. Fitting parameters and time to reach 99% hybridization for PNA1-GPG-TMR binding kinetics with DNA nanocage. Table of parameters obtained by fitting the hybridization kinetics data shown in Figure 3-6 to an exponential function. The equation is described in the methods section, and was used to determine the time for each PNA1-GPG-TMR to reach 99% hybridization.

Fitting Parameter	DNA-FAM + PNA1-GPG-TMR
y_0	99.02 ± 0.06%
A (counts)	(134.2 ± 0.3) x 10^{-3}
R_0 (counts/sec)	(15.7 ± 0.5) x 10^{-3}
R^2	0.9792
Time to 99% Hybridization	408 sec (6.8 min)

DNA-PNA-peptide thermal stability

We quantified the thermal stability of PNA1-GPG when labeled with each dye and bound to the DNA nanocage by incubating PNA1-GPG-TMR and PNA1-GPG-FAM with DNA nanocages labeled with the complementary dye at the vertex adjacent to the PNA-peptide (e.g., DNA-FAM and DNA-TMR, respectively), as shown in Figure 3-7 C and D. Figure 3-7 A and B show the fluoresce spectra taken of the complex DNA-FAM + PNA1-TMR and DNA-TMR + PNA1-FAM, respectively, as defined in Figure 3-3, while gradually heated from 11 °C to 85 °C. Initially for both complexes, the donor (FAM) fluorescence is significantly quenched because FRET occurs when the complex is fully assembled with the donor and acceptor in close proximity. As the temperature increases and the PNA-peptide dissociates from the DNA nanocage, the FRET signal becomes negligi-

ble as the separation between the donor and acceptor becomes large causing the donor fluorescence to increase significantly.

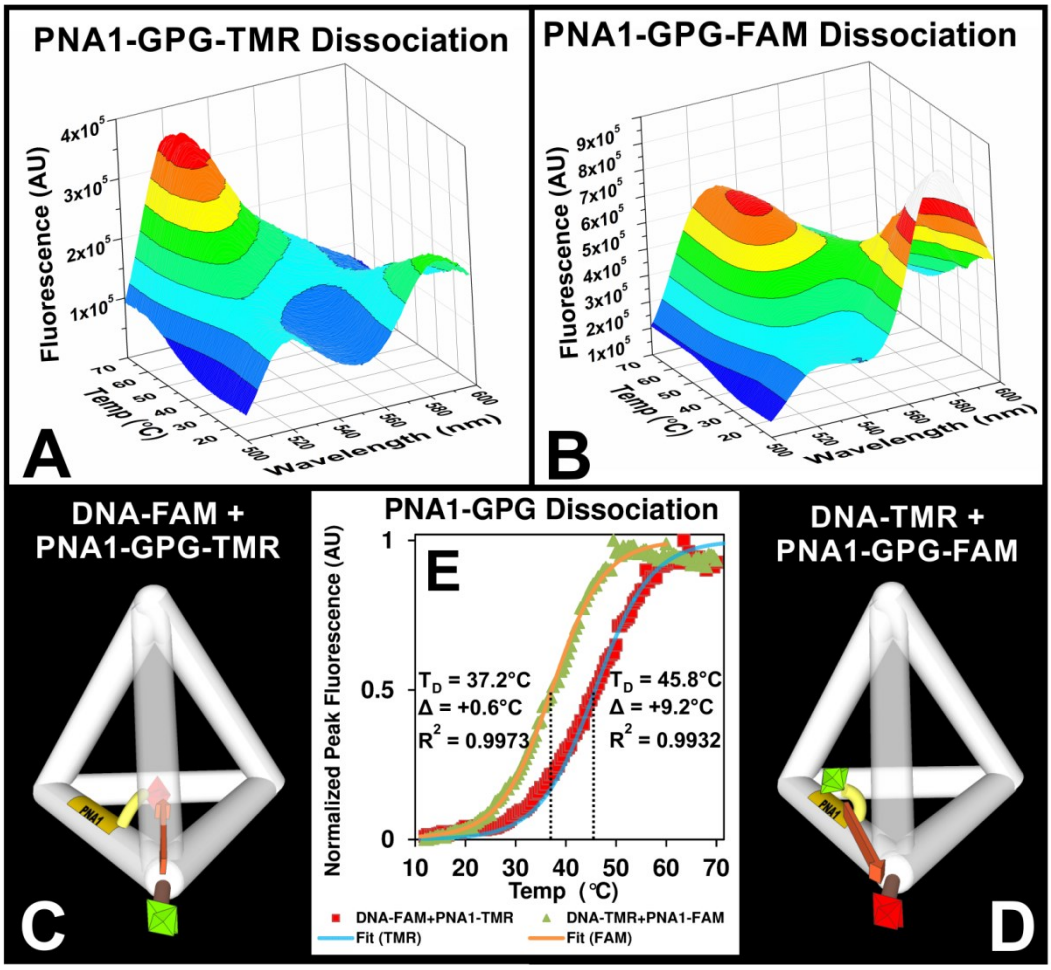


Figure 3-7. Fluorescence spectra of the TMR and FAM labeled PNA1-peptide bound to the FAM and TMR labeled DNA nanocage as a function of temperature. [A] PNA1-GPG labeled with TMR bound to the DNA nanocage labeled with FAM (DNA-FAM+PNA1-TMR), as shown in schematic "C" of the bottom figure. [B] PNA1-GPG labeled with FAM bound to the DNA nanocage labeled with TMR (DNA-TMR+PNA1-FAM), as shown in schematic "D" of the bottom figure. [E] The peak donor fluorescence is plotted versus temperature and fitted with a sigmoidal dose response curve, where the inflection point corresponds to the T_D indicated on the plot. The difference (Δ) from a theoretical value (36.6 °C) is also shown, which is greater than that predicted for a similar DNA sequence (34.5 °C).

The change in peak donor fluorescence with temperature in the presence of the acceptor was corrected for the change in fluorescence of the donor alone with tempera-

ture (Figure 2-17) and was fitted with a sigmoidal dose response curve (Figure 3-7 E), where the inflection point of the fitted curve indicates the dissociation temperature (T_D) of the PNA-peptide from the DNA nanocage. We use the term T_D instead of melting temperature (Tm) to avoid confusion from the melting of the DNA nanocage itself, which occurs at temperatures (~60 °C) above the dissociation of the PNA-peptides from the DNA nanocage.[125] We then used a method described by Giesen, Nielson and coworkers[95] to predict the T_D of a PNA-DNA duplexes based on the T_D of a DNA-DNA duplex with the same sequence, the length of the sequence and the fraction of pyrimidine bases also shown in Figure 3-7 (E). One possible interpretation of the higher T_D for PNA1-GPG-TMR might be that TMR interacts with guanine residues through π stacking[104] which could significantly stabilize PNA binding to the DNA nanocage.

While the stability of a DNA-PNA duplex depends on GC content, Sen and Nielsen showed that PNA strands have a much stronger increase in binding enthalpy and free energy with increasing purine content in the PNA strand than a DNA strand with the same sequence, possibly because of structural changes that were too subtle for them to detect by circular dichroism.[126] They demonstrated that a 10mer PNA sequence with 80% purine content increased the T_D to a complementary DNA sequence over a similar DNA sequence by 28.8 °C from 36.2 °C to 65.0 °C, whereas a 10mer PNA sequence with 80% pyrimidine content only increased the T_D by 2.8 °C to 39.0 °C. For applications that require higher PNA-peptide binding stability the purine content of our sequences could be increased further.

Fluorescence characterization: lifetime, anisotropy and energy transfer

We used Time Correlated Single Photon Counting (TCSPC) to measure the time resolved fluorescence decay of the dye labels to investigate the degree of interaction between both dye labeled PNA-peptides assembled into the DNA nanocage. Unlabeled PNA-peptides were prepared with N terminal acetylation (Ac) to serve as controls for our fluorescence experiments. Fluorescence decay kinetics of the donor alone (DNA + PNA1-FAM + PNA2-Ac) and the acceptor alone (DNA + PNA1-Ac + PNA2-TMR) were measured at their corresponding fluorescence maxima, and the fluorescence decay kinetics of the donor in the presence of the acceptor (DNA + PNA1-FAM + PNA2-TMR) was measured in the donor only emission region as shown in Figure 3-8 (A). The sequences and strand names of these complexes are defined in Figure 3-2 and Figure 3-3, respectively. Exponential fitting of each fluorescence decay required either two or three parameters, which are shown in Table 3-2. The long lived component is from the natural decay of each fluorophore. The short lived component is likely from fluorophore interactions with the DNA nanocage, such as quenching by nearby guanine nucleobases.[127] The medium length component identified for the donor in the presence of the acceptor is likely from energy transfer from donor to acceptor. For each decay the fitted components were averaged to facilitate comparing with the steady state fluorescence data discussed later. The calculated average lifetime of the donor alone is 3.9 ± 0.1 ns, which is in good agreement with another study,[128] as is the calculated average lifetime of the acceptor alone of 3.5 ± 0.1 ns.[104] The average donor lifetime is reduced to 2.5 ± 0.4 ns in the presence of the acceptor due to energy transfer from donor to acceptor, which corresponds to an energy transfer efficiency of 36%.

Table 3-2. Lifetime components and calculated average lifetimes of the donor and acceptor. Ex/Em (nm) indicates the excitation and emission wavelengths and each "%" column shows the relative amount of each lifetime component (τ). The lifetime of the donor (FAM) in the presence of the acceptor (TMR) was measured in the donor emission region (500 nm).

Name	Ex/Em (nm)	τ_1	% of τ_1	τ_2	% of τ_2	τ_3	% of τ_3	Average Lifetime
Donor (FAM) alone	450/520	4.20 ± 0.04ns	91.0 %	1.27 ± 0.4ns	9.0%	-	-	3.9 ± 0.1ns
Acceptor (TMR) alone	500/600	3.79 ± 0.06ns	84.3 %	1.88 ± 0.3ns	15.7 %	-	-	3.5 ± 0.1ns
Donor with Acceptor	450/500	4.3 ± 0.7ns	21.2 %	2.5 ± 0.3ns	56.4 %	0.7 ± 0.25ns	22.4 %	2.5 ± 0.4ns

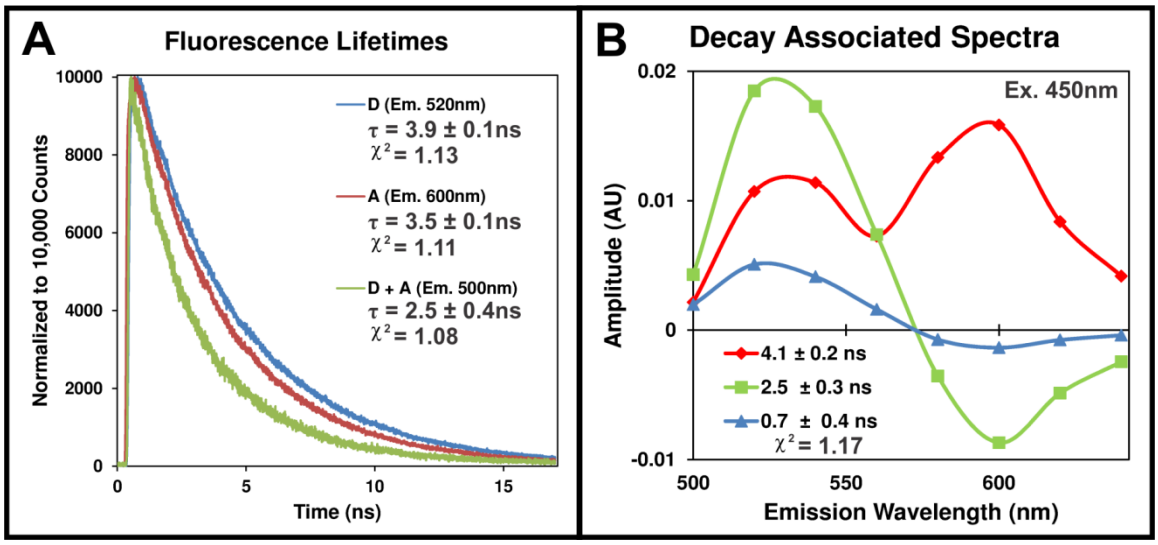


Figure 3-8. Fluorescence decay of the DNA nanocage with two fluorescently labeled PNA-peptides. [A] Plot of the fluorescence decay kinetics measured by TCSPC at 11 °C of the DNA nanocage with both PNA-peptides labeled with the donor (FAM) only (blue), the acceptor (TMR) only (red) and labeled with both the donor and acceptor and detected in the donor only emission region (green). [B] Fluorescence decay spectra were collected at several wavelengths and fit simultaneously to three components using global analysis. The relative amplitudes of each component from the analysis are plotted as a function of wavelength in the Decay Associated Spectra (DAS).

To gain insight into the different decay pathways of the donor excited state energy, we collected fluorescence decay kinetics data at several wavelengths spanning the donor and acceptor emission spectra. The kinetics curves at all recorded wavelengths were fit simultaneously to three components of the donor-acceptor system using global analysis. The relative amplitudes of each component from the analysis are plotted as a function of wavelength in the Decay Associated Spectra (DAS) as shown in Figure 3-8 (B). The most dominant component of donor decay has a lifetime of 2.5 ± 0.3 ns and shows a positive peak in the donor emission wavelength region (fluorescence decay) and a similarly sized negative peak (increase in fluorescence) in the acceptor region, which is characteristic of energy transfer from the donor to the acceptor. The second major component has a lifetime of 4.1 ± 0.2 ns, which has a positive peak in both the donor and acceptor region. This component indicates that there is a subpopulation of the donor that does not undergo energy transfer, but instead decays with its natural lifetime, possibly from an orthogonal orientation with respect to the acceptor. Global analysis is unable to resolve the differences in this component from the natural decay of the acceptor and thus provides an average lifetime of both components. The remaining component has a lifetime of 0.7 ± 0.4 ns with a positive peak in the donor region and small negative peak in the acceptor region suggesting that a second faster pathway for energy transfer exists. The DAS shows energy transfer between both dye labeled PNA peptides, indicative of proper complex assembly, and suggests that on average the dye labeled peptides adopt two distinct conformational states.

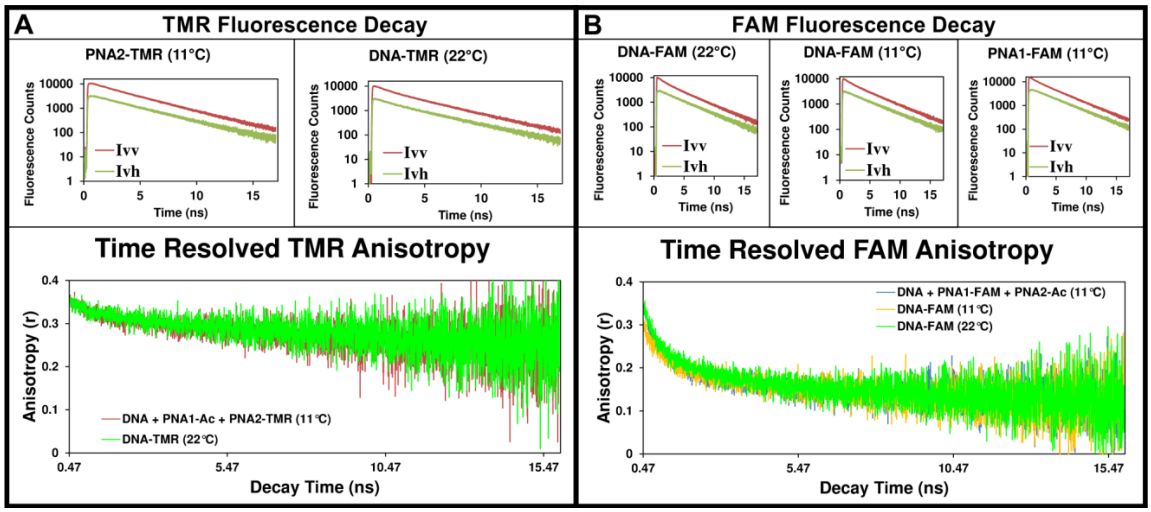


Figure 3-9. Fluorescence Decay and Anisotropy of FAM and TMR dyes attached to a complex of a PNA-peptide inside the DNA nanocage. [Top of A and B] Plot of the fluorescence decay of TMR at 600nm when excited at 500nm (A) and FAM at 520 nm when excited at 450 nm (B). The decay was measured by TCSPC at the indicted temperature both parallel (Ivv) and perpendicular (Ivh) to the excitation polarization, with the dye attached to the PNA and hybridized to the DNA nanocage or attached directly to the DNA nanocage. [Bottom Figure of A and B] The anisotropy was calculated for each data point of the decay, as described in the Materials and Methods section called Fluorescence anisotropy, and plotted for each complex.

TCSPC was also used to evaluate how freely the dye labels can rotate within the DNA-PNA-Peptide complex by probing the rate of depolarization of their emission when excited by polarized light as a function of time. The complex with the acceptor alone (DNA + PNA1-Ac + PNA2-TMR) was excited by vertically polarized light at 500 nm and the TMR emission was monitored at 600nm with polarization both parallel (VV, vertical excitation and emission) and perpendicular (VH, vertical excitation and horizontal emission) relative to the polarization of the excitation, as shown in Figure 3-9 (A). The difference in the emission intensity of VV and VH was used to determine the anisotropy using the equations described in the Materials and Methods section called Fluorescence anisotropy. The calculated anisotropy is plotted as a function of time in Figure 3-9 (A), which shows a steady decrease in anisotropy from 0.36 initially to 0.24 after 17 ns. The

maximum possible anisotropy value is 0.4 for a perfectly rigid complex with no depolarization, whereas the minimum value of zero indicates a rapidly rotating dye molecule that has fully depolarized. This measured value of 0.24 is in excellent agreement with a value (0.24) reported for the steady state anisotropy of a TMR labeled oligonucleotide, due to restricted dye mobility from π stacking with nearby guanine nucleobases.[104] The same method was used to determine the anisotropy of the DNA nanocage directly labeled with TMR (DNA-TMR). Interestingly, the calculated anisotropy of DNA-TMR decays in a nearly identical fashion over time to a value of 0.24 after 17 ns, as shown in Figure 3-9 (A). Evidently the TMR dye interacts equally with guanine when intercalated in the DNA helix along the edge or when stacked to one of the terminal guanines at the end of the three helices that converge at the vertex, as identified in the DNA sequences in Figure 3-2.

The complex with the donor alone (DNA + PNA1-FAM + PNA2-Ac) was measured in the same way as for TMR, however, the sample was excited at 450 nm and the FAM emission was monitored at 520 nm. The calculated FAM anisotropy decays much faster to a value of 0.12 after 17 ns, as shown in Figure 3-9 (B). This value is slightly higher than the steady state values reported in the literature[104] for FAM labeled oligonucleotides, which may be due to the 3D conformation of the DNA nanocage used in our study further restricting the motion of the dye. The calculated anisotropy of DNA-FAM at 11 °C and at 22 °C decay to values of 0.11 and 0.12, respectively, after 17 ns. Just as with TMR, the calculated anisotropy of FAM was found to decay in nearly identical fashion regardless of whether it was attached to the PNA-peptide or directly attached to the DNA nanocage.

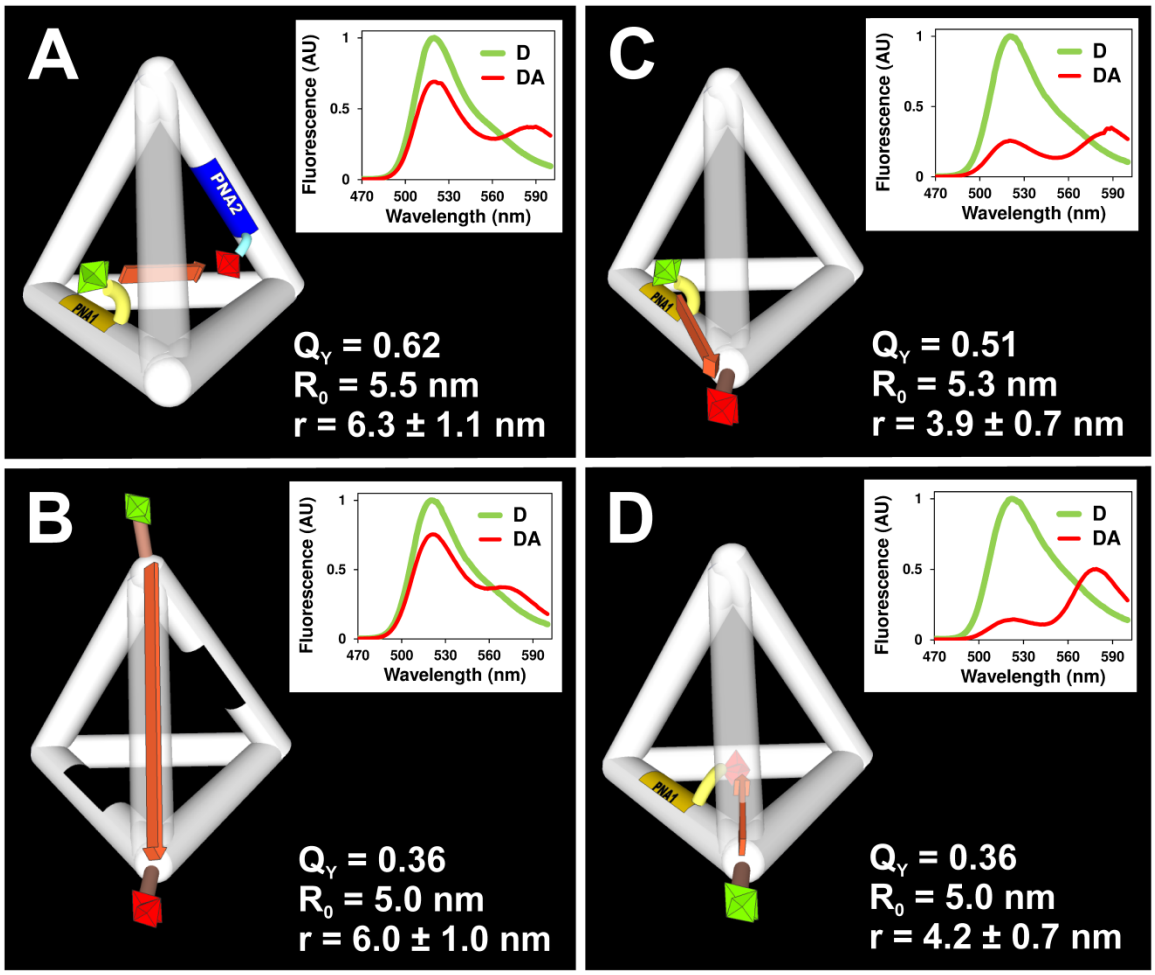


Figure 3-10. Förster distance measurements between FAM and TMR dyes attached to a complex of a DNA nanocage with 0, 1 or 2 PNA-peptides inside. A schematic is shown of each of the following complexes: [A] DNA + PNA1-FAM + PNA2-TMR, [B] DNA-TMR-FAM (no PNA), [C] DNA-TMR + PNA1-FAM, [D] DNA-FAM + PNA1-TMR. Next to each schematic is the fluorescence emission spectra (Ex. 450 nm) of the donor only (D) and of the donor with the acceptor (DA) measured at 11 °C along with the calculated energy transfer efficiency. Below the fluorescence spectra is the determined quantum yield (Q_Y), calculated Förster radius (R_0) and Förster distance (r).

The steady state fluorescence spectra were measured at 11 °C for the following complexes, as shown in Figure 3-10: (A) the designed DNA-PNA-peptide complex (DNA + PNA1-FAM + PNA2-TMR), (B) the DNA nanocage labeled at two adjacent vertices but without any bound PNA (DNA-FAM-TMR), as well as the constructs pre-

pared to measure PNA1-GPG thermal stability of (C) DNA-TMR + PNA1-FAM and (D) DNA-FAM + PNA1-TMR For each construct, Figure 3-10 shows a cartoon schematic, the fluorescence spectra with the change in donor fluorescence (D) in the presence of the acceptor (DA), the determined quantum yield (Q_Y), and the calculated Förster radius (R_0) and Förster Distance (r). The Q_Y of the donor only complexes were determined by comparing the fluorescence intensity of each complex (without the acceptor) to that of a sodium fluorescein standard,[101] when normalized to concentration. The calculated Q_Y for the donor (FAM) attached to the PNA1-peptide and hybridized to the DNA nanocage were higher (0.62 and 0.51) than when directly attached to the DNA nanocage vertex (0.36). This range of Q_Y were in agreement with those reported in the literature for FAM attached to dsDNA.[128,129] The change in donor (FAM) fluorescence in the presence of the acceptor (TMR) was used to determine the FRET efficiency and together with the Q_Y was used to calculate the corresponding Förster distances using the Förster equations described in the Materials and Methods section called Förster radius, energy transfer efficiency and distance calculations. For the designed construct shown in Figure 3-10 (A), the FRET efficiency determined from steady state fluorescence (30%) is similar to the value (36%) obtained by the change in average donor lifetime observed using TCSPC.

When calculating the Förster distances we assumed the dyes were randomly oriented, with an orientation factor (κ^2) of ⅔. However, the anisotropy data indicate that both dyes interact with the DNA nanocage, which could bias particular orientations and introduce uncertainty into the calculated distances. Dale and coworkers developed a method to determine the limits of the orientation factor based on the anisotropy data.[130] We used a variation of this method developed by Lakowicz and coworkers[131] and de-

scribed in the Materials and Methods section called Förster radius, energy transfer efficiency and distance calculations, to determine the upper and lower limits of the orientation factor (κ^2) to be between $\kappa^2_{min} = 0.2$ and $\kappa^2_{max} = 1.7$, which translates to the distance uncertainties shown in Figure 3-10 for each construct.

The calculated Förster distances of the different complexes in Figure 3-10 were used in combination with gel electrophoresis, lifetime and anisotropy data to develop a model of the insertion and structural arrangement of the fluorescently labeled PNA-peptides inside the DNA nanocage. The Förster distances of 6.3 ± 1.1 nm measured between edges through the center of the DNA nanocage (Figure 3-10 A) and of 6.0 ± 1.0 nm measured along one of the DNA nanocage edges (Figure 3-10 B) are slightly longer and shorter, respectively, than the corresponding distances of 7 nm and 5–6 nm, respectively, of a tetrahedron with straight edges. These measurements suggest a conformation of the DNA nanocage where the edges are bowed outward giving the tetrahedron a more spherical rather than pyramid shape, similar to a previously reported cryo-electron microscopy study.[64] The FAM labeled PNA-peptide causes a pronounced native gel band shift of the DNA nanocage upon binding, which suggests the negatively charged FAM^{2-} is repelled away from the DNA nanocage slowing its migration through the gel. The relatively high TMR anisotropy data suggests that it strongly interacts with the DNA nanocage and stabilizes the bound PNA-peptide, resulting in a 9.2 °C increase its dissociation temperature. Gel electrophoresis data of the TMR labeled PNA-peptide shows a dominant conformation that does not significantly affect the complex migration in the gel, which suggests the TMR may intercalate with the DNA nucleobases or bind on the interior of the DNA nanocage. The gel also shows minor band broadening suggesting that

TMR may also non-specifically interact with the surrounding DNA helix in a range of conformations. This TMR labeled PNA-peptide conformational heterogeneity is likely the source of the two energy transfer pathways identified using TCSPC. The donor-acceptor distances for the dominant energy pathway (2.5 ns, 36% FRET efficiency) was calculated to be 6.0 ± 1.0 nm, while the faster pathway (0.7 ns, 82% FRET efficiency) was calculated to be 4.3 ± 0.7 nm. The longer distance suggests a peptide conformation with the TMR intercalating with nearby DNA bases, while the shorter distance suggests a peptide conformation with the TMR extended into the center of the DNA nanocage toward the FAM labeled PNA-peptide.

Conclusion

We have demonstrated a rapid and flexible method using PNA linkers to assemble peptides inside a small 3D DNA nanocage. The PNA-peptides could be introduced quantitatively by incubating for 5–10 minutes at room temperature with the preassembled DNA nanocage, either in parallel or stepwise fashion, with as little as two times molar excess of PNA-peptide. The TMR dye label was found to significantly stabilize PNA-peptide binding to the DNA nanocage, increasing the T_D by 9.2 °C from a predicted value of 36.6 °C to 45.8 °C. Data from biophysical characterization using gel electrophoresis as well as steady state and time resolved fluorescence spectroscopy allowed us to develop a model for the arrangement of the PNA-peptides inside the DNA nanocage that is in agreement with the intended design.

Our design strategy can be further expanded to introduce four PNA binding gaps by shortening the length of the remaining two DNA strands (e.g., S1 and S2). In order to

access the remaining two edges of the DNA nanocage two of the existing four DNA strands must each be divided into a 13bp strand and a 34bp strand, which may require further optimization of the annealing conditions for proper DNA nanocage assembly. The gaps can also be translated to other edges contacted by the same strand, or shifted along the same edge to control the orientation of the N terminally attached peptide. Additional peptides can be introduced at the C terminus of the PNA during the initial synthesis or along the PNA backbone using orthogonal protecting groups.[82] A previous study estimated the DNA nanocage could accommodate a 60 kDa (~500 amino acid) globular protein,[50] but could just as easily be filled by a series of peptides assembled together using multiple PNAs. This work illustrates how a 3D DNA scaffold can be used to assemble polypeptides through rapid and controlled hybridization with PNA linkers. The flexibility of our method offers a biomimetic route to rebuild protein active sites and further expand the structure and functional landscape of polypeptide engineering.

CHAPTER 4 – Low Temperature Assembly of Functional 3D DNA-PNA-Protein Complexes

Adapted with permission from Flory, J. D.; Simmons, C. R.; Lin, S.; Johnson, T.; Andreoni, A.; Zook, J.; Ghirlanda, G.; Liu, Y.; Yan, H.; Fromme, P. Manuscript submitted for publication to *J. Am. Chem. Soc.* **2014**.[132]

Abstract

Proteins have evolved to carry out nearly all the work required of living organisms within complex inter and intracellular environments. However, systematically investigating the range of interactions experienced by a protein that influence its function remains challenging. DNA nanostructures are emerging as a convenient method to arrange a broad range of guest molecules. However, flexible methods are needed for arranging proteins in more biologically relevant 3D geometries under mild conditions that preserve protein function. Here we demonstrate how peptide nucleic acid (PNA) can be used to control the assembly of cytochrome c (12.5 kDa, pI 10.5) and azurin (13.9 kDa, pI 5.7) proteins into separate 3D DNA nanocages, in a process that maintains protein function. Toehold-mediated DNA strand displacement is introduced as a method to purify PNA-protein conjugates. The PNA-proteins were assembled within 2 minutes at room temperature, within 4 minutes at 11 ^{0}C, and hybridize with even greater efficiency than PNA conjugated to a short peptide. Gel electrophoresis, steady state and time resolved fluorescence spectroscopy were used to investigate the effect of protein surface charge on its interaction with the negatively charged DNA nanocage. These data were used to generate a model of the DNA-PNA-protein complexes that show the negatively charged azurin protein repelled away from the DNA nanocage while the positively charged cytochrome c

protein remains within and closely interacts with the DNA nanocage. When conjugated to PNA and incorporated into the DNA nanocage, the cytochrome c secondary structure and catalytic activity were maintained, and its redox potential was reduced modestly by 20 mV possibly due to neutralization of some positive surface charges. This work demonstrates a flexible new approach for using 3D nucleic acid (PNA-DNA) nanostructures to control the assembly of functional proteins, and facilitates further investigation of protein interactions as well as engineering more elaborate 3D protein complexes.

Introduction

Proteins evolved an enormous diversity of function that is carried out through massively complex pathways of interactions with other biomolecules, which are being compiled into streamlined databases of possible protein-protein interactions and networks.[55,56] However, many of these interactions are dependent on the method used to discover them or are non-functional,[57] so new methods are needed to validate and further characterize the specific properties of each interaction.[58] DNA nanostructures offer an intriguing platform to study the function of proteins through controlled manipulation of the protein orientation and proximity to other molecules of interest. The field of DNA nanotechnology has been extensively reviewed,[15,133] including specialized reviews focusing on protein functionalization.[134,135] Methods for manipulating proteins on 1D and 2D DNA nanostructures have been utilized for selectively arranging multiple enzymes,[42] investigating DNA binding with aptamers[44] and zinc finger proteins,[45] assembling complementary protein subunits[46] or protein ligands,[40,47] and for investigating the spatial dependence of enzyme cascades.[48,49] 3D DNA nanostructures have been used to encapsu-

late[50] and arrange proteins[51–53] as well as to organize membrane proteins for NMR structure determination.[54] Because natural protein interactions occur in complex 3D arrangements there is a need to develop methods for spatially and temporally controlling protein interactions in 3D, which may not be accessible in nanostructures of lower dimension. Furthermore, some proteins contain specific DNA binding domains[61] and many others are likely influenced by the presence of a negatively charged DNA scaffold, so a better understanding of the electrostatic effects of DNA nanostructures on proteins is needed to establish guidelines for using DNA scaffolds to study proteins. Previous work conjugating protein to DNA has largely focused on introducing proteins onto large 2D DNA lattices[8] or DNA origami nanostructures.[136] Small DNA nanostructures facilitate confining proteins to more biologically relevant interaction distances and frequencies, and can also be more easily incorporated into other materials and complexes.[137] Furthermore, small DNA nanostructures (<100 kDa) allow higher molar concentration to be achieved, which enables a broader range of protein assays, such as spectroelectrochemistry,[59] and structural analytical techniques, such as circular dichroism,[60] not practical with large DNA origami nanostructures (> 4 MDa).[21]

Previously, a DNA tetrahedron was used to encapsulate the protein cytochrome c.[50] In that work the authors directly conjugated the protein to one of the four ssDNAs required for proper assembly of the DNA tetrahedron, then assembled the DNA-protein complex in a single high temperature annealing step. It is unclear how the protein function was affected by the preparative methods and whether this method could be extended to a wider range of proteins. While it is possible to assemble DNA nanostructures at lower temperature, this generally reduces the assembly yield and complicates purification.

Furthermore, if multiple proteins were assembled on the same scaffold in one step, the timing of the interactions could not be controlled. Nucleobase hybridization coupling strategies can overcome this limitation, by first assembling the DNA nanostructure with a binding site for a guest molecule conjugated to a ssDNA, and then incorporating the guest molecule by hybridization under mild conditions.[29,41] However, even oligonucleotides of moderate length can substantially reduce the efficiency of incorporating proteins in close proximity.[41] Peptide nucleic acid (PNA)[66] is an artificial nucleic acid with enhanced binding affinity for DNA that shows great promise for applications in gene targeting,[116] sensing[118,138] and as a nanomaterial.[78,139,140] We previously showed that 8 nucleotide (nt) strands of PNA can be used to rapidly and efficiently assemble short, fluorescently labeled peptides at room temperature into an existing DNA tetrahedron design[65] modified to include a single-stranded PNA binding domain for each peptide.[79] It was unclear if the effectiveness of this strategy could be extended to much longer polypeptides and if it would preserve function.

In this report we conjugate a PNA linker to two small proteins, cytochrome c (12.5 kDa, pI 10.5) and azurin (14 kDa, pI 5.7), and assemble each conjugate into separate 3D DNA nanocages via nucleobase hybridization with the attached PNA linker, as shown in Figure 4-1. These proteins allowed us to investigate the feasibility of using PNA to assemble longer, functional polypeptides, as well as compare the effect of each protein's net surface charge on its interaction with the negatively charged 3D DNA nanocage. First we synthesized an 8nt PNA sequence and conjugated it to a surface thiol on either protein. Toehold-mediated DNA strand displacement was developed to purify the PNA-protein conjugates from unreacted protein. Gel electrophoresis was used to de-

termine the PNA-protein binding efficiency. Förster resonance energy transfer between the tetramethylrhodamine (TMR) labeled PNA-protein and the fluorescein (FAM) labeled DNA nanocage (DNA-FAM) was used to monitor the PNA-protein hybridization kinetics, thermal dissociation, and separation distances. Time resolved spectroscopy was used to analyze the interaction of the TMR labeled PNA-proteins with the DNA-FAM nanocage. Gel electrophoresis along with steady-state and time-resolved fluorescence data were used to construct a model of each DNA-PNA-protein complex. Finally, the function of cytochrome c, when conjugated to PNA and hybridized inside the DNA nanocage, was characterized by measuring the activity of luminol oxidation and by determining its redox potential using spectroelectrochemistry.

Results and Discussion

DNA-PNA-protein complex design

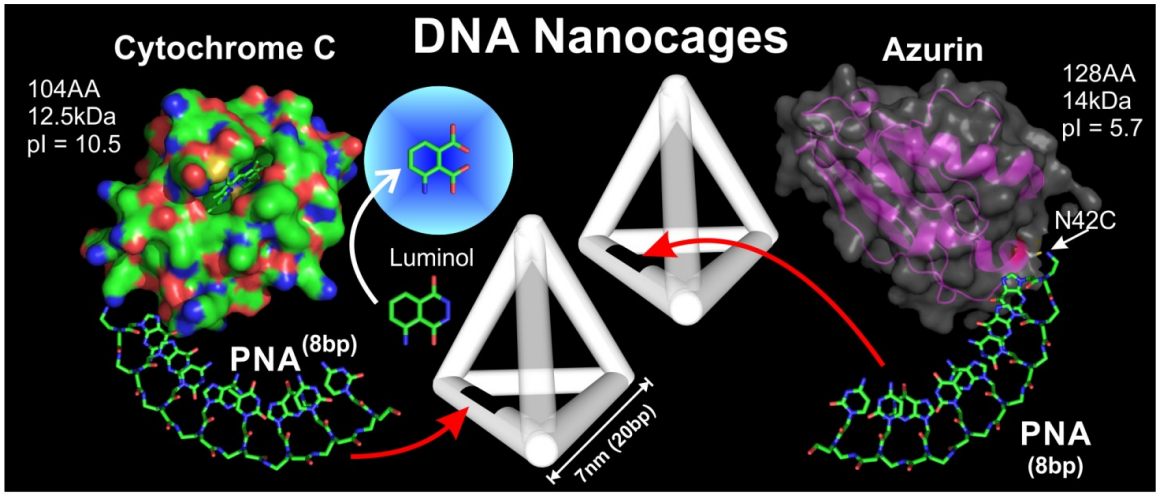
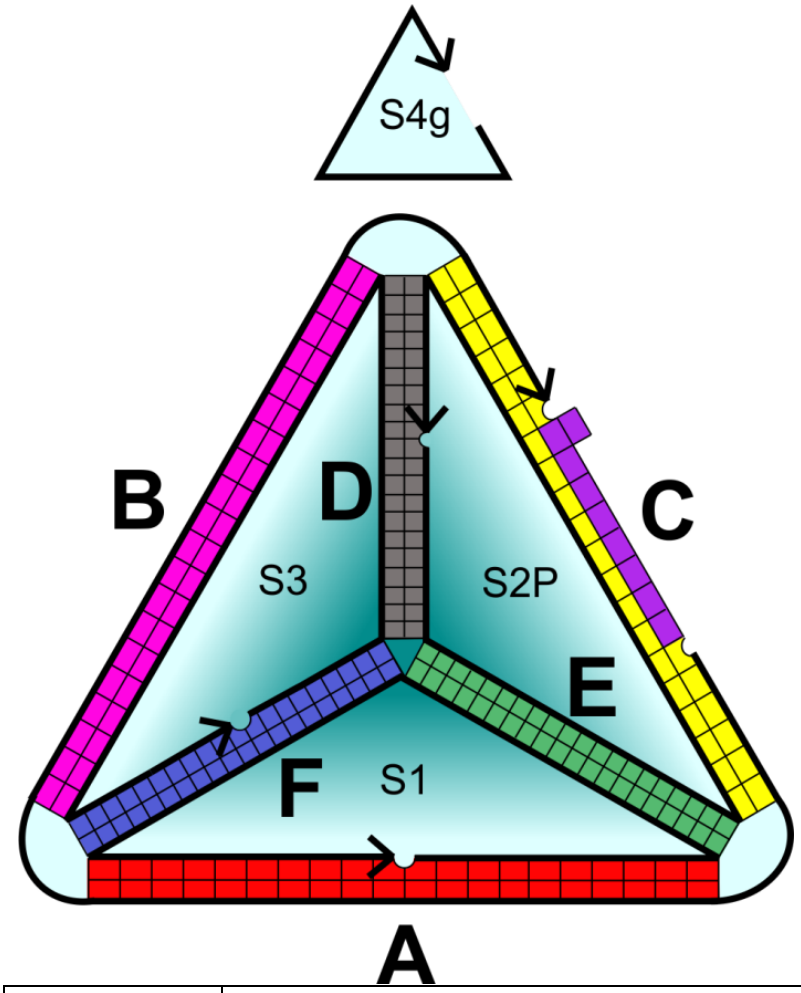


Figure 4-1. Proposed design to separately assemble two proteins into a DNA nanocage using a PNA linker.



Strand Name	Sequence
S1	AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCACCCGCCA-TAGTAGACGTATCACC
S2P	CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGACGCACAGA-TACCGACGATTACAG
S2P-FAM	CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGACGCACAGA-TACCG/iFluorT/CGATTACAG
S2Pg	ATACCGACGATTACAGCTTGCTACACGATTCAGACTTAG-GAATGTTCGACATGCG
S3	GGTGATAAAACGTGTAGCAAGCTGTAATCGACGGGAAGAG-CATGCCCATCCACTACTATGGCG
S4g	CGCATGACTCAACTGCCTGGTGATACGAGGATGGG-CATGCTCTTCCCGACGGTAT
S4P	CTCTTCCCGACGGTATACGCACAGCGCATGACTCAACTGCCTGGTGA-TACGAGGATGGGCATG
PNA-FAM	FAM-Gly-CTGTGCGT
PNA-TMR	TMR-Gly-CTGTGCGT
PNA-Az	Azurin-SMCC-Gly-CTGTGCGT
PNA-Az-TMR	TMR-Azurin-SMCC-Gly-CTGTGCGT
PNA-CytC	Cytochrome c-SPDP-SMCC-Gly-CTGTGCGT
PNA-CytC-TMR	TMR-Cytochrome c-SPDP-SMCC-Gly-CTGTGCGT

Figure 4-2. DNA and PNA sequences and schematic showing their arrangement in the complex. [Top] Schematic of the PNA and DNA sequence layout, where the PNA-protein is attached to Edge C. The strand nicks are indicated with the arrow pointing in the 3' direction. Strand S4g is redrawn above for clarity. This schematic was redrawn from Flory et al.[79] with changes to incorporate the PNA-proteins. [Bottom] DNA and PNA sequences with the color code matching the schematic in the top figure. Variations of strand S2 are included with a FAM label (S2-FAM) and with a PNA binding sequence (S2P) for an 'inverted' DNA nanocage design where the PNA binding domain is moved across the helix, rotating the orientation of protein attachment by about 180 degrees on edge C. The DNA and PNA sequences are written in the 5' to 3' direction. The PNA sequence has an N-terminal (5') glycine followed by the SMCC cross-linker for attaching to the surface thiol on azurin (N42C mutant) or cytochrome c (introduced with SPDP). Dye labeled DNA strands are defined using the nomenclature for ordering from IDT, where the dye names are enclosed by "/" marks and are attached to a modified thymine base at the vertex.

Previously we demonstrated the PNA-driven assembly of two fluorescently labeled peptides within a DNA nanocage.[79] Briefly, we introduced modifications to an existing DNA tetrahedron design by introducing two 8nt single stranded domains on opposite edges of the DNA nanocage for binding complementary PNA strands with short peptides on the N terminus oriented toward each other through the center of the DNA nanocage. The design allowed the DNA nanocage to be preassembled, purified and then subsequently populated with one or two PNA-peptides under mild conditions. Here, we extend this strategy to introduce two different metalloproteins into separate DNA nanocages, as shown in Figure 4-1, at the PNA binding site on one of the edges, as shown in the sequence schematic in Figure 4-2. Throughout this report we use the nomenclature defined in Figure 4-2, Figure 4-3 and Figure 4-6 when we describe the various strands and constructs used in each experiment.

Complex Name	Strands Required to Assemble Construct
DNA (nanocage)	S1, S2P, S3, S4g
DNA-FAM	S1, S2P-FAM, S3, S4g
DNAi (inverted nanocage)	S1, S2Pg, S3, S4P
DNA + PNA-Az	S1, S2P, S3, S4g, PNA-Az
DNA + PNA-Az-TMR	S1, S2P, S3, S4g, PNA-Az-TMR
DNA-FAM + PNA-Az-TMR	S1, S2P-FAM, S3, S4g, PNA-Az-TMR
DNA + PNA-CytC	S1, S2P, S3, S4g, PNA-CytC
DNA + PNA-CytC-TMR	S1, S2P, S3, S4g, PNA-CytC-TMR
DNA-FAM + PNA-CytC-TMR	S1, S2P-FAM, S3, S4g, PNA-CytC-TMR

Figure 4-3. Complex nomenclature with list of required DNA and PNA-protein strands for assembly. List of names of complexes referred to in this chapter with the corresponding strands and conjugate names required to assemble the complex.

PNA-protein conjugation, purification and fluorescent labeling

The schemes used to construct conjugates of PNA with either dye labeled or un-labeled azurin or cytochrome c are shown in Figure 4-4. Detailed methods for the conjugation, purification and fluorescent labeling and further discussion can be found in the Materials and Methods section called Fluorescent labeling of cytochrome c with TMR and section called TMR labeling of PNA-azurin. Briefly, the PNA sequence TGCGTGTC-Gly, written from C to N terminus (3′ to 5′) was prepared by microwave-assisted solid phase synthesis as described previously,[79,83] and the free N terminus was labeled with an SMCC cross-linker. RP-HPLC and MALDI of PNA-SMCC is shown in Figure 2-5. PNA-SMCC was conjugated to a single surface cysteine (N42C) containing azurin mutant (~40% yield).[107] The chemical structure of the linker connecting the PNA to azurin is shown in Figure 4-10B. The SPDP cross-linker was used to introduce 1.1–2.6 thiols per cytochrome c, conjugated to PNA-SMCC (yield 5–10%) and quenched using

aminoethylmalemide (AEM) to restore positive surface charges lost during the thiolation of surface amines with SPDP. The chemical structure of the linker connecting the PNA to cytochrome c is shown in Figure 4-10A. Some of the conjugates were dye labeled with TMR for subsequent fluorescent studies.

Further purification of the conjugates was not essential since the remaining free protein could be removed by size exclusion fast protein liquid chromatography (SE-FPLC) after the PNA-protein (15-16 kDa) hybridizes with the much larger DNA nanocage (80 kDa). However, characterization of the PNA-CytC conjugates within a mixture of free protein was challenging due to the low amount of conjugate (5–10%) and the significant absorbance of cytochrome c at 260 nm (ε_{260} = 20,140 $M^{-1}cm^{-1}$) compared to PNA (ε_{260} = 69,932 $M^{-1}cm^{-1}$), as shown in Figure 2-7. Due to the neutral charge and small size of the PNA linker (2.4 kDa) as compared to the protein (12.5 kDa), charge and molecular weight based separation techniques were not effective to separate the PNA-CytC conjugates from unreacted cytochrome c. This motivated our pursuit to develop a new method of purifying PNA-protein conjugates using toehold-mediated DNA strand displacement, as described in the next section.

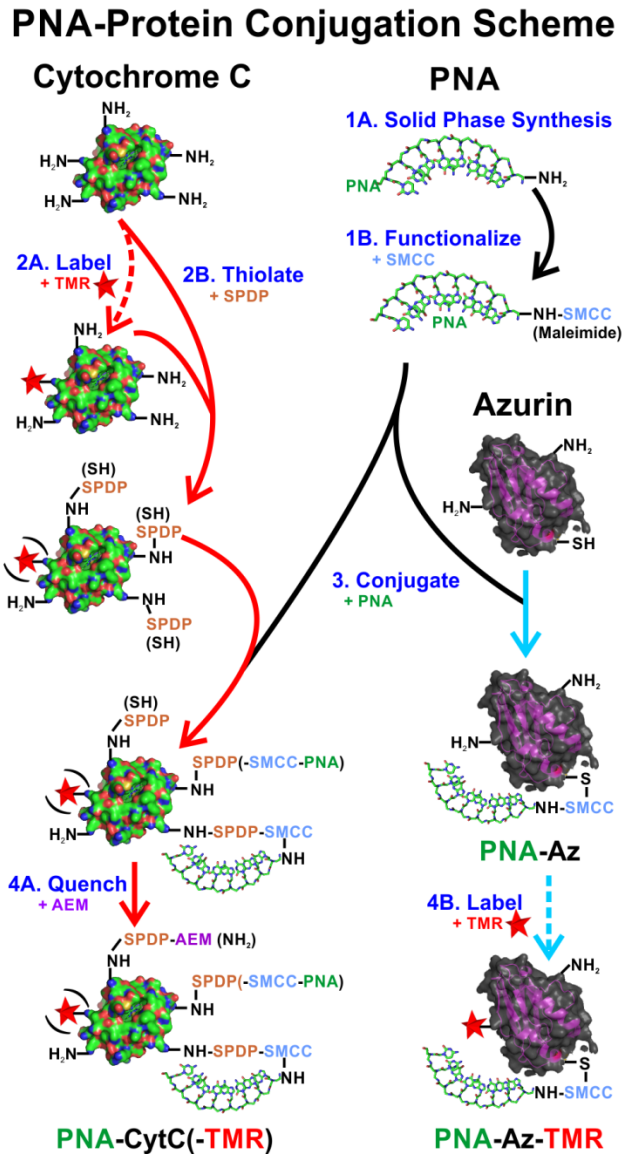


Figure 4-4. Scheme used to produce the following dye labeled and unlabeled PNA-protein conjugates: PNA-CytC, PNA-CytC-TMR, PNA-Az, PNA-Az-TMR. 1A) Solid Phase Synthesis of 8nt PNA sequence with N terminal glycine, 1B) Functionalization of PNA with maleimide on N terminus using SMCC cross-linker. 2A) Labeling of cytochrome c surface lysines with TMR dye, 2B) Functionalization of cytochrome c surface lysines with thiols using SPDP cross-linker, 3) Conjugation of PNA-SMCC to surface thiol(s) on cytochrome c and azurin, 4B) Quenching of unreacted surface thiols on cytochrome c with AEM to restore positive charge, 4B) Labeling azurin surface lysines with TMR dye.

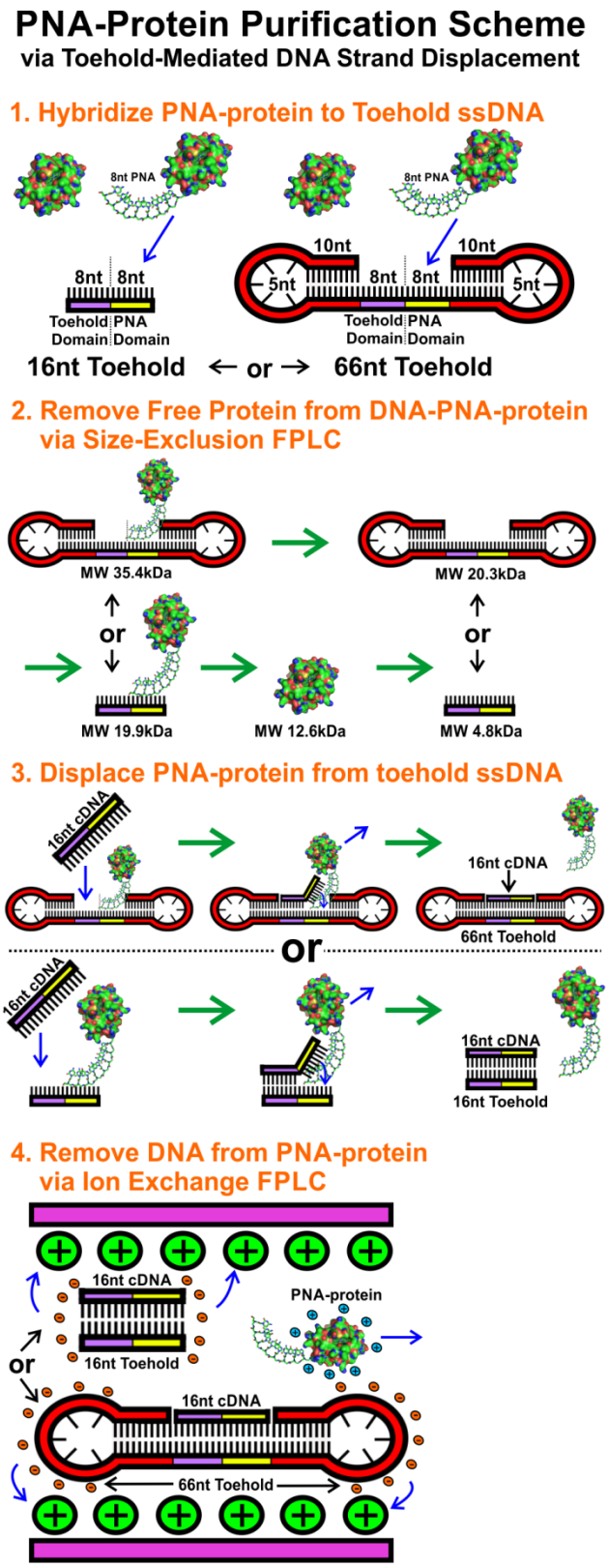
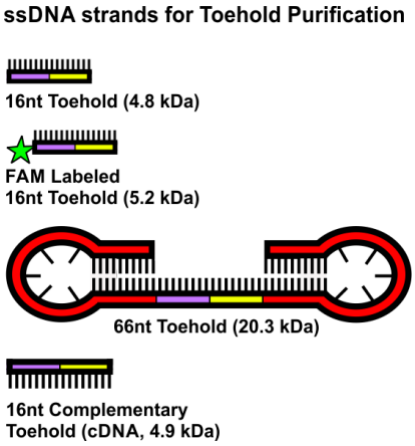


Figure 4-5. Scheme for purifying PNA-protein using toehold-mediated DNA strand displacement. Details about the procedure are described in the text.

PNA-protein purification by toehold-mediated DNA strand displacement

In order to further isolate and characterize the PNA-CytC conjugate from the free cytochrome c, we sought to leverage PNAs ability to hybridize to DNA. Toehold-mediated DNA strand displacement is an established technique that uses a partially complementary DNA strand to initially bind a target, but leaves a ssDNA 'toehold' domain available for a fully complementary strand to bind and release the target, due to the more favorable binding energies, as reviewed by Zhang and Seelig.[141] Here we adapted this technique to purify PNA-protein conjugates, as shown in Figure 4-5. To our knowledge this is the first application of this technique for purification purposes. In step 1, the crude mixture of PNA-CytC and free cytochrome c was incubated with either a 16nt or 66nt toehold ssDNA. Each toehold strand contains an 8nt domain (shown in yellow) fully complementary to the 8nt PNA sequence conjugated to the protein, and an 8nt toehold domain (shown in purple) that remains unhybridized. In step 2, the additional mass provided by either the 16nt (4.8 kDa) or 66nt (20.3 kDa) toehold in the DNA-PNA-protein complex allows SE-FPLC to be used to separate it from the free protein. In step 3, a ssDNA fully complementary to the toehold (cDNA) is incubated with the DNA-PNA-protein complex. The cDNA initially binds to the free part of the toehold DNA domain and then it completely displaces the PNA-protein. In step 4, the cDNA and toehold DNA are removed from the PNA-protein using ion exchange FPLC (IEX-FPLC). The sequences of the strands used in this method are described in Figure 4-6.



Strand Name	Sequence
16nt Toehold	AGCTCCAAACGCACAG
FAM Labled 16nt Toehold	/56-FAM/AGCTCCAAACGCACAG
66nt Toehold	CGAACATTCCAAAAAGGAATGTTCGAGCTCCAAAC-GCACAGCGATTACAG-CAAAAAGCTGTAATCG
16nt Complementary Toehold (cDNA)	CTGTGCGTTTGGAGCT

Figure 4-6. DNA sequences for toehold mediated PNA purification. DNA sequences used for purifying the PNA or PNA-protein conjugates by toehold mediated strand displacement. The sequences are written in the 5' to 3' direction. The 16nt toehold strand is comprised of an 8nt domain (yellow), which is fully complementary to PNA, an 8nt domain (purple) that provides a site for the complementary toehold strand (cDNA) to bind and displace the PNA or PNA-protein. A larger molecular weight 66nt toehold was also used, which was designed with self-complementary 10nt stem s (green and black) and 5nt loops (grey) flanking each binding domain.

Initially we suspected that the additional charge from the toehold DNA strand would make charge based separation techniques straightforward to isolate the DNA-PNA-protein complexes. Both IEX-FPLC and native PAGE were successful for separating fluorescently labeled PNA strands, as shown in Figure 2-10 and Figure 2-11, respectively, which is discussed further in the Materials and Methods section called ssDNA toehold mediated PNA and PNA-protein purification. Subsequent mixing of the DNA-PNA complex with the fully complementary (cDNA) strand liberated the PNA. When the same method was applied to PNA-CytC, however, the DNA-PNA-CytC hybridization was disrupted during purification. Two major factors may contribute to the dissociation of the toehold DNA from the PNA-CytC: 1) structural changes to the DNA as it binds to the anion exchange column may weaken its binding to PNA-CytC, and 2) the dipole

moment between the positively charged PNA-CytC and negatively charged DNA may increase the rate of dissociation in the presence of a large electric field during PAGE. To circumvent this issue, we performed SE-FPLC on the DNA-PNA-CytC mixture. Figure 4-7A shows the SE-FPLC chromatogram with the nucleobase absorbance at 260 nm and the heme absorbance of cytochrome c at 414 nm. Peaks 1 and 2 show significant A_{260} and A_{414}, which we assign to cytochrome c labeled with one (19.7 kDa) and two (26.9 kDa) PNA strands, respectively (masses include complementary toehold DNA strands). We assign the third peak to the free cytochrome c (12.5 kDa), because of the large A_{414} and modest A_{260}. The fourth peak, showing A_{260} but negligible A_{414}, is attributed to unhybridized 16nt toehold DNA strand (4.8 kDa).

In order to improve the separation of the PNA-CytC conjugates, a 66nt (20.3 kDa) toehold strand was designed to induce an even larger mass shift. Figure 4-7B shows the SE-FPLC chromatogram of PNA-CytC incubated with the 66nt toehold DNA. Peaks 1 and 2 show cytochrome c with one and two PNAs attached, respectively. These peaks show superior separation from the free cytochrome c peak to what was achieved with the 16nt toehold strand. Because of its larger size, the 66nt toehold DNA strand partially overlaps with the PNA-conjugate peaks, but which is removed during the subsequent purification step (vide infra). An additional peak is observed just prior to the free cytochrome c peak (peak 0), which could be PNA-CytC conjugate that dissociated from its complementary toehold DNA during the run.

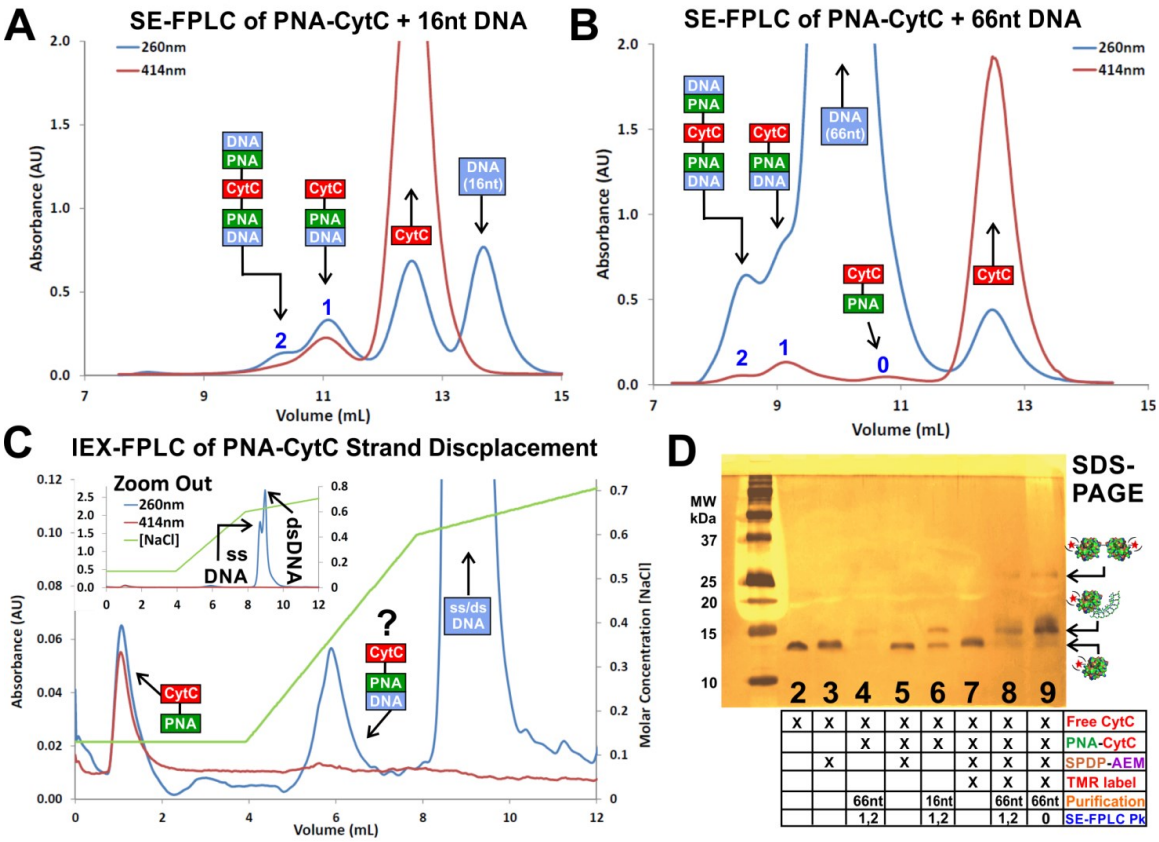


Figure 4-7. Results for PNA-protein conjugate purification by toehold-mediated strand displacement. Size Exclusion (SE) FPLC chromatogram of PNA-CytC (15.1 kDa, ε_{260} = 90k $M^{-1}cm^{-1}$, ε_{414} = 106k $M^{-1}cm^{-1}$) incubated with a complementary ssDNA of 16nt (4.8 kDa, ε_{260} = 157k $M^{-1}cm^{-1}$) [A] and 66nt (20.3 kDa, ε_{260} = 661k $M^{-1}cm^{-1}$) [B]. The nucleobase absorption at 260 nm and heme absorption at 414 nm were monitored to detect the elution of the product. [C] Ion Exchange (IEX) FPLC chromatogram showing the toehold and fully complementary ssDNA and corresponding dsDNA retained on the anion exchange column while the cationic PNA-CytC conjugate flows through. [D] Silver stain of a 4% stacking and 15% resolving SDS-PAGE of the SE-FPLC peaks of the PNA-CytC conjugates after purification using different complementary ssDNAs (16nt or 66nt). Lanes 2 and 3 are free cytochrome c and SPDP-AEM modified cytochrome c, respectively (refer to Figure 4-4 for SPDP-AEM complex); lanes 5 and 7 are the crude PNA-CytC and PNA-CytC-TMR mixtures with free cytochrome c; lane 4, 8 and 9 show conjugate purified using the 66nt toehold DNA; Lanes 4 and 8 show the overlapping peaks 1 and 2 combined for PNA-CytC and PNA-CytC-TMR, respectively, and lane 9 contains peak 0 from PNA-CytC-TMR.

All PNA-CytC conjugate peaks were collected and then incubated with fully complementary DNA (cDNA) strands to liberate the PNA-CytC conjugate. An IEX-

FPLC step was performed to separate the PNA-CytC from the DNA toehold. Figure 4-7C shows a typical IEX-FPLC chromatogram for the separation: the positively charged PNA-CytC conjugate does not bind to the resin and elutes in the void volume fraction, while the negatively charged ssDNA and dsDNA are retained on the column and are eluted when an increasing salt gradient is applied. The unknown peak eluting before the free DNA may be a small amount of conjugate still bound to the toehold DNA strand that was not completely displaced by the cDNA. Figure 4-7D shows an SDS-PAGE gel comparing unlabeled and TMR labeled PNA-CytC conjugates after toehold-mediated purification. On the gel, lanes 2 and 3 were loaded with free cytochrome c and SPDP-AEM modified cytochrome c, respectively (refer to Figure 4-4 for SPDP-AEM constructs). Both lanes show a single band at the expected molecular weight (~13 kDa). Lanes 5 and 7 were loaded with PNA-CytC and PNA-CytC-TMR, respectively. Both samples were purified by SE-FPLC to remove unreacted PNA, but still contain free cytochrome c. The low fraction of conjugate (~6%) estimated from the UV-Vis spectrum was not observed on the gel using silver stain. After toehold-mediated purification of the crude mixtures, the PNA-CytC peak from IEX-FPLC was concentrated and loaded in lane 6. A clear enrichment of the PNA-CytC conjugate is observed when using the 16nt toehold strand. The product obtained using a 66nt toehold DNA to purify PNA-CytC(-TMR) was loaded on lane 4, 8 and 9. A larger enrichment than what was obtained with the 16nt strand is observed both in the case of PNA-CytC (lane 4) and PNA-CytC-TMR (lane 8 and 9) conjugates.

PNA-protein assembly with the DNA nanocages

DNA nanocages with an 8nt single stranded domain for binding the PNA-protein conjugates were prepared by thermal annealing as described previously.[79] The DNA nanocage was purified using IEX-FPLC (Figure 2-14A), as described and discussed further in the Materials and Methods section called DNA nanocage assembly. In separate experiments, the preformed DNA nanocage was incubated at room temperature with increasing molar excess of PNA-CytC, PNA-azurin or with 5x molar excess of a fluorescently labeled PNA (PNA-FAM).[79] The sample was cooled to 4 °C and then directly analyzed using native PAGE at 4 °C. The results are shown in Figure 4-8: all the constructs exhibit a band shift upon PNA hybridization to the DNA nanocage. Both PNA-proteins achieve near quantitative binding after stoichiometric amounts are added. This is a noticeable improvement from our previous study of PNA- peptides,[79] which , due to self-aggregation,[80] required two-fold PNA excess for quantitative binding. This improvement in the hybridization efficiency may be explained by an enhanced solubility of the PNA after conjugation to the protein.

The smallest gel shift, with respect to the empty DNA nanocage, was observed for the binding of PNA-FAM, followed by a slightly larger shift induced from the binding of PNA-CytC, and then a more substantial shift caused from the binding of PNA-azurin. The binding of PNA itself induces a shift as observed previously,[79] likely because the DNA nanocage becomes more rigid as the single stranded binding domain for PNA becomes double stranded. The slight additional shift observed for PNA-CytC is likely due to a lower overall negative charge of the DNA-PNA-CytC nanocage assembly after binding the positively charged cytochrome c (pI ~10.5). The substantial shift induced by

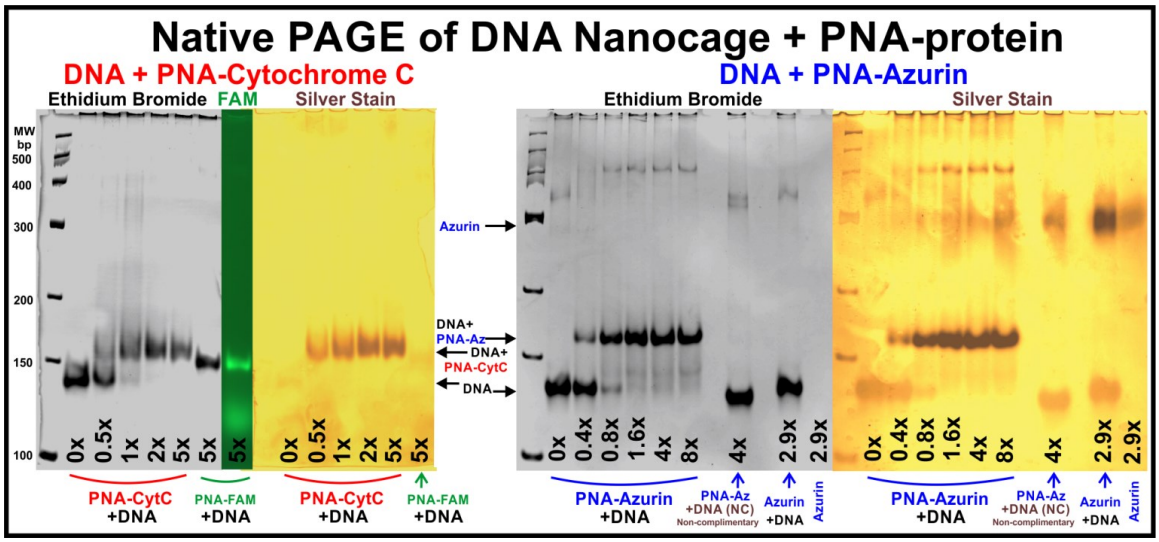


Figure 4-8. Gel showing the binding of PNA-cytochrome c, PNA-azurin or PNA-FAM to the DNA nanocage. [Left] Native PAGE (7%), run at 4 ^{0}C, of the DNA nanocage incubated with increasing molar excess (0x to 5x) of PNA-CytC or 5x PNA-FAM. From left to right: ethidium bromide staining of all lanes, FAM fluorescence (Ex. 475 nm, Em 535 nm) for the lane loaded with PNA-FAM, and silver staining of all lanes. [Right] Native PAGE (7%) of the DNA nanocage incubated with increasing molar excess (0x to 8x) of PNA-azurin, 4x PNA-azurin with a DNA nanocage with a non-complementary binding sequence, 2.9x free azurin with the DNA nanocage and 2.9x free azurin alone. From left to right: ethidium bromide and silver staining for all lanes.

PNA-azurin, suggests that the negatively charged protein is repelled away from the DNA nanocage; this increases the hydrodynamic radius of the assembly, therefore causing a slower migration in the gel, as compared with the empty DNA nanocage. A substantial gel shift was also observed for the same DNA nanocage attached to a slightly larger but also negatively charged green fluorescent protein (28 kDa, pI 5.7)[52] as well as streptavidin.[53] An alternative 'inverted' DNA nanocage design was assembled that moved the PNA binding domain across the helix, rotating the orientation of the point of protein attachment by about 180 degrees (sequences described in Figure 4-2). After incubating PNA-Az-TMR with the original DNA nanocage design as well as the inverted DNA nanocage design (DNAi), native PAGE showed no difference in the mobility of the com-

plexes, as shown in Figure 2-14C lanes 2–5. A possible explanation for this is that, regardless of the orientation provided by the PNA attached to the DNA nanocage, the glycine-SMCC linker provides sufficient flexibility for the protein to adopt its most thermodynamically favored orientation, based on the surrounding environment.

The kinetics of PNA-protein hybridization into the DNA nanocage was investigated at room temperature by monitoring quenching of the FAM labeled DNA nanocage fluorescence emission as TMR labeled PNA-proteins were added, as discussed in the Materials and Methods section called PNA-polypeptide hybridization kinetics. PNA-Az-TMR and PNA-CytC-TMR were found to reach 99% hybridization within 2 minutes, as shown in Figure 2-16. These rates are faster than the 5–10 minutes previously observed for TMR labeled PNA-peptides.[79] PNA-CytC-TMR was also found to hybridize to the DNA nanocage within 4 minutes at 11 ^{0}C. Since proteins are normally kept frozen or chilled to extend their useful lifetime, hybridization at low temperature may be ideal to study temperature sensitive proteins or protein complexes. The cytochrome c heme absorbance was also found to quench FAM fluorescence upon binding to the DNA nanocage. Unlabeled PNA-CytC hybridized within at 13 minutes at 25 ^{0}C and within 25 minutes at 11 ^{0}C, suggesting that the TMR dye label may facilitate binding of PNA-CytC. TMR is known to interact with DNA nucleobases, which may help to keep the protein close to the DNA nanocage to facilitate PNA hybridization.[104] The dissociation temperature (T_D) of the TMR labeled PNA-proteins bound to the FAM labeled DNA nanocage was determined by monitoring FRET of the complex while the temperature was slowly increased from 11 ^{0}C to 85 ^{0}C, as shown in Figure 2-18. The measured T_D of both PNA-proteins were 3–4 ^{0}C lower than the theoretical value for the PNA sequence alone.

This indicates that the presence of either protein destabilizes the PNA-DNA duplex to the same degree.

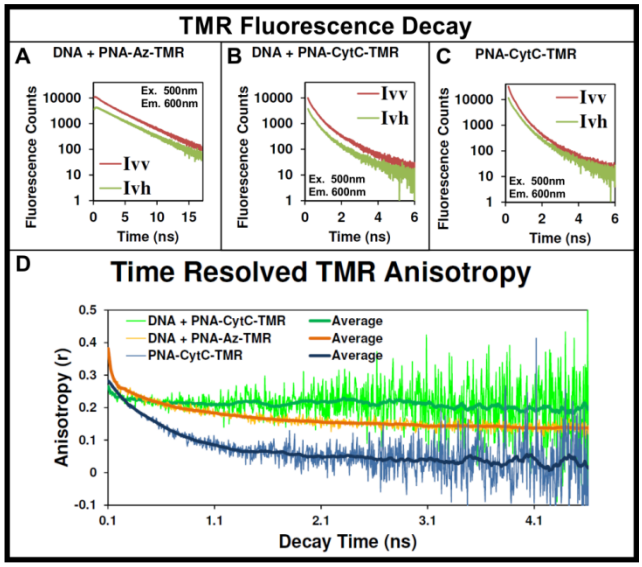


Figure 4-9. Fluorescence decay and anisotropy of two TMR dye labeled proteins hybridized to a DNA nanocage with PNA. (A, B, C) Plot of the parallel (Ivv) and perpendicular (Ivh) TMR fluorescence decay of PNA-Az-TMR hybridized to the DNA nanocage (A), and of PNA-CytC-TMR when hybridized to the DNA nanocage (B) and when free in solution (C). Panel (D): The anisotropy for each data point of the decays, shown in panels A, B, and C, was calculated as described in the Materials and Methods section called Fluorescence anisotropy, and the results are plotted here. The central moving average for each dataset is reported in the same plot for clarity.

Fluorescence characterization: anisotropy, lifetime and energy transfer

The arrangement of both proteins in the DNA nanocage environment was further investigated using fluorescence spectroscopy. We used time-correlated single photon counting (TCSPC) to evaluate how freely the dye labels can rotate when attached to the PNA-protein and when hybridized into the DNA nanocage by probing the rate of depolarization of their emission when excited by polarized light. Each complex was excited by vertically polarized light at 500 nm and the emission of TMR was monitored at 600 nm with polarization both parallel (VV, vertical excitation and emission) and perpendicular

(VH, vertical excitation and horizontal emission) relative to the polarization of the excitation, as shown in Figure 4-9 (A, B and C). The difference in the emission intensity of VV and VH was used to calculate the anisotropy decay kinetics using the equations described in the Materials and Methods section called Fluorescence anisotropy. The anisotropy decay and corresponding central moving average are plotted for each complex as a function of time in Figure 4-9 (D). The maximum possible anisotropy value is 0.4 for a perfectly rigid complex with no depolarization, whereas the minimum value of zero indicates a rapidly rotating dye molecule that has fully depolarized. The anisotropy for TMR decays with three distinct components, a <0.2 ns fast component due to energy transfer between proteins with multiple TMR labels, a 0.2–2 ns intermediate component due to the rotation of the protein and a >2 ns slow component due to the rotation of the DNA nanocage. Interpretation of the anisotropy decay of the fast components and absolute anisotropy values of the slow components are complex due to the differences in TMR labeling of both proteins and is discussed in the next paragraph. The intermediate decay component (0.2–2 ns) due to protein rotation is observed for PNA-Az-TMR in the DNA nanocage and for PNA-CytC-TMR in solution but not for PNA-CytC-TMR in the DNA nanocage. This suggests that cytochrome c remains within and closely interacts with the DNA nanocage, which significantly restricts its rotation, while azurin is repelled away from the DNA nanocage and rotates quite freely.

Interpretation of the anisotropy decay of TMR attached to both protein shown in Figure 4-9 at short (<0.2 ns) and long (>2 ns) times is complex because of the differences in labeling between both proteins. The fast initial decay for PNA-CytC-TMR from 0.4 to ~0.25 during the first 0.1 ns is not resolved during the measurement, but is observed both

when PNA-CytC-TMR is in solution and when bound to the DNA nanocage using PNA. Such rapid decay is likely because of multiple TMR dyes conjugated to each protein which rapidly transfer energy between each other and thus accelerate the rate of depolarization.[142] The initial decay for PNA-Az-TMR is slightly slower (~0.2 ns) than for PNA-CytC-TMR (~0.1 ns), which is likely because of a lower degree of TMR labeling on azurin. This is consistent with the observed changes in the TMR absorbance spectra for PNA-CytC-TMR and PNA-Az-TMR. After 2 ns the anisotropy levels out for each complex, where the PNA-CytC-TMR in solution depolarizes completely (0.04) but both PNA-CytC-TMR and PNA-Az-TMR hybridized to the DNA nanocage maintain substantial anisotropy due to the slow rotation of the DNA nanocage. The observed differences in the residual anisotropy of PNA-Az-TMR (0.14) and PNA-CytC-TMR (0.20) cannot be directly compared due to the differences in TMR labeling of both proteins.[142] However, the intermediate anisotropy decay component shows meaningful differences between the both proteins based on the DNA nanocage environment as previously discussed.

Additional TCSPC measurements were performed to obtain information on the position of each of the proteins inside the DNA nanocage after hybridization. The fluorescence emission decay of FAM on an empty DNA-FAM nanocage was compared to the decay of the same label after either PNA-CytC-TMR or PNA-Az-TMR was hybridized to the DNA-FAM nanocage. The results of these experiments are shown in Figure 2-20. From the time dependent intensity decay it is clear that the lifetime of FAM is reduced upon binding of either of the two TMR-labeled PNA-protein conjugates. The average lifetimes determined from the fitting the data in Figure 2-20 with exponential functions are presented in Table 2-2. The change in FAM lifetime in the presence of each TMR-

labeled PNA-protein was used to determine the energy transfer efficiency of 17 ± 5% for PNA-Az-TMR and 37 ± 3% for PNA-CytC-TMR as described in the Materials and Methods section called Fluorescence lifetime. The fluorescence decay kinetics of a DNA

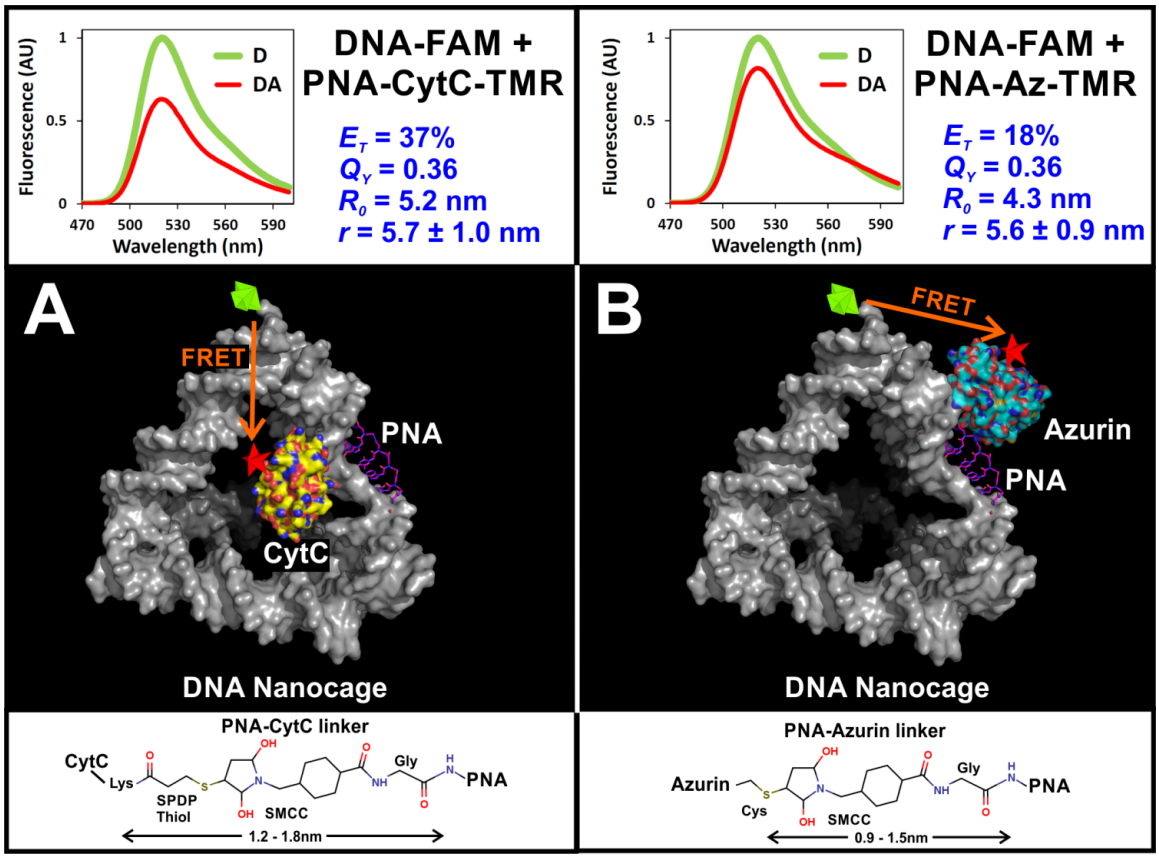


Figure 4-10. Steady State Fluorescence Spectra, Molecular Model of the DNA-PNA-protein complexes and linker chemical structure. [Top] Fluorescence emission spectra (Ex. 450 nm) of the donor only (D) and of the donor with the acceptor (DA) measured at 11 °C of the DNA-FAM nanocage with PNA-CytC-TMR (A) or PNA-Az-TMR (B). Each spectra also shows the calculated energy transfer efficiency (E_T), the determined quantum yield (Q_Y), calculated Förster radius (R_0), and Förster distance (r). [Middle] Molecular models of the dye labeled DNA nanocage hybridized to either a dye labeled PNA-cytochrome c conjugate (A) or PNA-azurin conjugate (B), with an arrow indicating the direction of energy transfer. The chemical structure of the linkers used to conjugate PNA to each protein is shown below each model.

nanocage labeled with FAM at the vertex adjacent to the TMR labeled PNA-protein was also measured at various wavelengths between 500-640 nm that span the FAM emission

spectrum. The data were fit using a global analysis algorithm and plotted as a function of wavelength in the Decay Associated Spectra (DAS) shown in Figure 2-21. Both the DAS and fluorescence decay indicate that energy transfer occurs from FAM to TMR in the DNA-PNA-protein nanostructures. These data are furthered discussed in the Materials and Methods section called Fluorescence lifetime.

Together with TCSPC, steady-state fluorescence spectroscopy was used to measure the distance between the TMR labeled PNA-proteins and the DNA nanocage vertex labeled with FAM. The steady state fluorescence spectra of DNA-FAM + PNA-CytC-TMR and DNA-FAM + PNA-Az-TMR were recorded at 11 °C, and are shown in Figure 4-10 panel (A), and (B), respectively. Figure 4-10 reports the following for each construct: the fluorescence spectra of the donor (D), the fluorescence spectra of the donor in the presence of the acceptor (DA), the energy transfer efficiency (E_T), the quantum yield of FAM bound to DNA (Q_Y), the Förster radius of the FAM-TMR FRET pair (R_0), and the Förster distance in the construct (r). The values of E_T, Q_Y, R_0, and r were calculated as described by Flory et al.[79] The energy transfer efficiency values calculated from the intensity change of the DNA-FAM fluorescence in the presence of either PNA-CytC-TMR or PNA-Az-TMR were 37% and 18%, respectively. These values agree well with the 37 ± 3% and 17 ± 5% calculated using the change in DNA-FAM lifetime as measured by TCSPC. The Förster radii for the FRET pairs DNA-FAM with PNA-CytC-TMR-FAM or PNA-Az-TMR were calculated to be 5.2 nm and 4.3 nm, respectively. The larger Förster radius for PNA-CytC-TMR was due to the greater overlap of its absorption spectrum with the DNA-FAM emission spectrum. The details of the calculations, as well as discussion about the spectral data used to obtain these values are presented in the Materials and

Methods section called Förster radius, energy transfer efficiency and distance calculations and section called Calculating the absorption spectra and extinction coefficients for the overlap integral of TMR labeled PNA-proteins, respectively. The Förster distances separating the FAM label on the DNA nanocage from the PNA-CytC-TMR or PNA-Az-TMR were found to be 5.7 ±1.0 nm and 5.6 ± 0.9 nm, respectively.

With the constraint that both proteins have approximately the same separation distance from the adjacent DNA nanocage vertex, we developed a model of the DNA-PNA-protein complexes that account for differences observed during PAGE and time resolved anisotropy, as shown in Figure 4-10. The significantly larger PAGE gel shift and faster anisotropy decay kinetics observed for TMR labeled PNA-azurin attached to the DNA nanocage suggest it is repelled away from the DNA nanocage creating a larger hydrodynamic radius for the complex and allows the protein to rotate freely. PNA-CytC induces a gel shift similar to that observed for the PNA alone, suggesting that the protein is confined within the DNA nanocage. The TMR anisotropy decay due to PNA-CytC rotation in solution is absent when hybridized into the DNA nanocage suggesting that PNA-CytC is tightly interacting with the DNA nanocage, and thus restricts its rotation. The chemical linkers used to connect PNA to each protein are shown in Figure 4-10 along with the projected length. Our model suggests that despite the same PNA binding site in the DNA nanocage, these chemical linkers are not rigid enough to hold two similarly sized proteins at the same location. Instead, the flexibility of the linker allows each protein to find a favorable orientation based on its surface electrostatic interactions with the negatively charged DNA nanocage. To improve control over the protein in the future, it may be advantageous to operate near the isoelectric point of the protein, increase the salt concentra-

tion, add a second specific conjugation site to further constrain the orientation of the protein, or include a complementary protein in an adjacent part of the DNA nanocage that can form a stable protein-protein complex.

Cytochrome c activity

One of the most challenging aspects of working with proteins is to identify relevant *in situ* methods to assess the integrity of their structural and functional properties. Cytochrome c is a redox active heme containing protein that functions as an electron carrier in the mitochondrion. Like other heme containing proteins, such as horseradish peroxidase, it has the ability to oxidize luminol in the presence of hydrogen peroxide to produce chemiluminescence. Furthermore, the heme absorbance at 550 nm changes significantly upon oxidation and reduction, thereby facilitating the spectroscopic assessment of its oxidation state. Wild type azurin is a copper containing redox active protein, however, the mutant used in this study contained redox inactive zinc in the metal center.

The activity of cytochrome c in the DNA-PNA-CytC complex was analyzed directly in a native polyacrylamide gel using the luminol assay, where the gel was imaged for 2 hours to monitor the progression of the reaction, as described in the Materials and Methods section called Luminol oxidation activity. Figure 4-11A and Figure 4-11B show the ethidium bromide stain and luminol chemiluminescence, respectively, of the DNA nanocage alone (lane 1), the DNA nanonocage with PNA-CytC (lane 2), the DNA nanocage incubated with 5x free cytochrome c (lane 3) and separately incubated with 5x PNA-FAM (lane 4). All DNA constructs migrate uniformly, with a noticeable shift for the DNA-PNA-CytC containing band from the empty DNA nanocages. Only the DNA-

PNA-CytC complex shows chemiluminescence. The intensity of the DNA-PNA-CytC (lane 2) and DNA + free cytochrome c (lane 3) bands in each image was integrated and plotted as a function of integration time, as shown in Figure 4-11C. Cytochrome c demonstrates substrate diffusion limited kinetic behavior as it catalyzes the oxidation of luminol in the DNA nanocage, which is in agreement with what was observed in solution.[143]

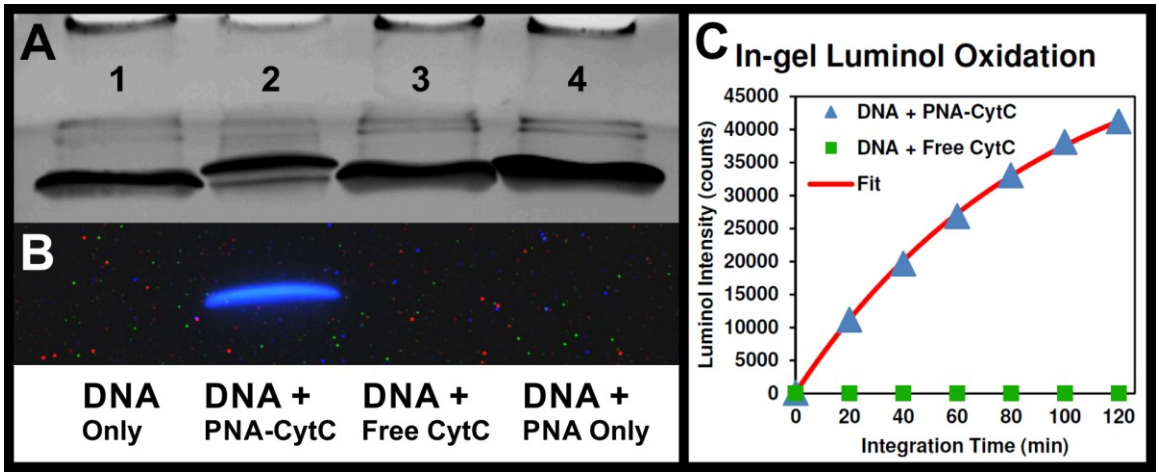


Figure 4-11. In-gel activity of PNA-cytochrome c bound to the DNA nanocage. [A] Ethidium bromide stain of a 4% stacking and 15% resolving native PAGE of 100 pmols of DNA nanocage alone (lane 1), DNA + PNA-CytC (lane 2), the DNA nanocage incubated with 5x free cytochrome c (lane 3) and separately incubated with 5x PNA-FAM (lane 4). [B] Chemiluminescence image of the gel from [A] due to luminol oxidation by PNA-CytC in the DNA nanocage. [C] Integrated chemiluminescence band intensities of DNA-PNA-CytC and DNA + Free CytC taken every 20 minutes.

Cytochrome c redox potential

The redox potential of cytochrome c was also investigated inside the DNA nanocage by monitoring the ratio of oxidized and reduced cytochrome c based on the absorbance change at 550 nm. The potential of the solution was adjusted using the ferricyanide-ferrocyanide redox couple, as described in the Materials and Methods section called Redox potential titration with ferri/ferrocyanide. A schematic of the reaction is shown in

Figure 4-12A. The fraction of ferricytochrome c (oxidized) is plotted as a function of the heme absorption spectra in Figure 4-12B. The log of the ratio of ferrocytochrome c to ferricytochrome c is plotted against the log of the ratio of ferrocyanide to ferricyanide (which set the solution potential) as shown in Figure 4-12C. The *x* intercept of a linear fit

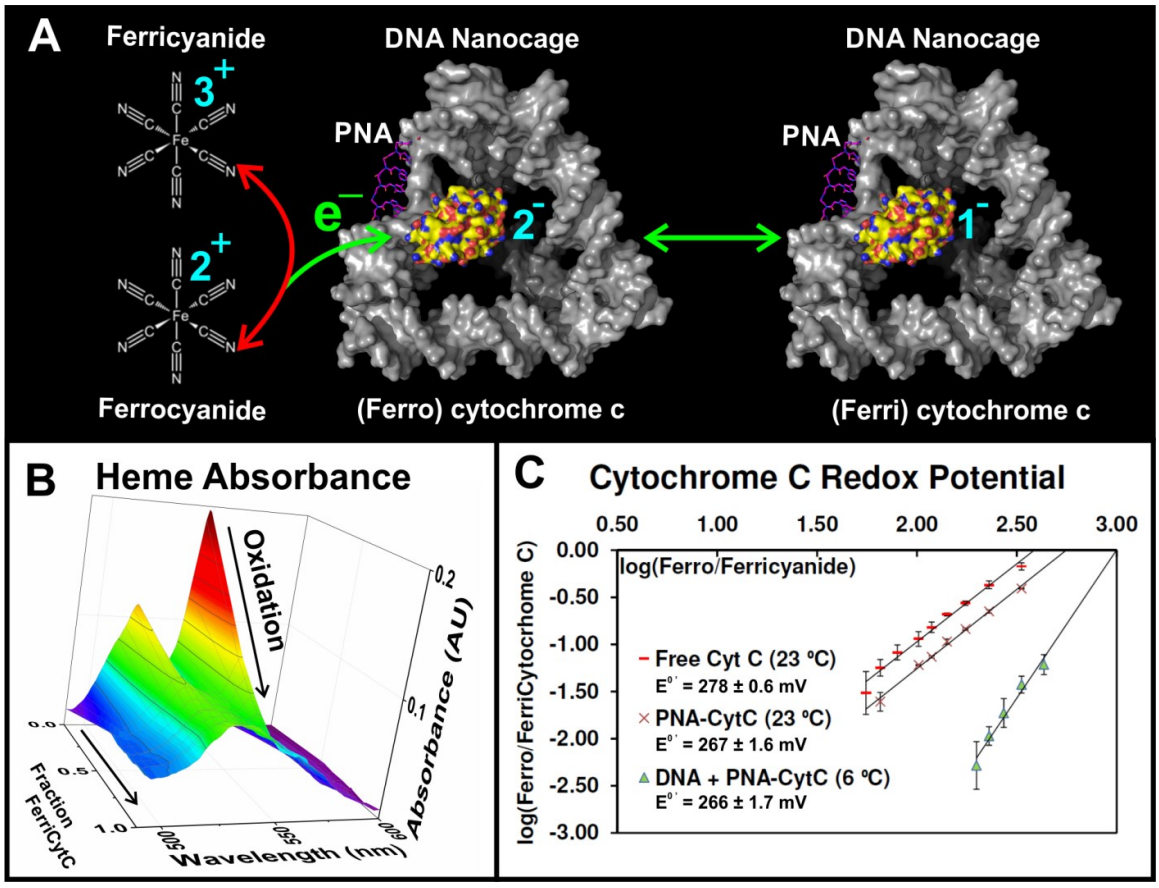


Figure 4-12. Spectroelectrochemical results of PNA-cytochrome when bound to the DNA nanocage. [A] Schematic of the oxidation of ferrocytochrome to ferricytochrome c inside the DNA nanocage by ferricyanide. [B] Absorption spectra of the DNA-PNA-CytC complex as a function of the fraction of ferricytochrome c (oxidized). [C] Log of the ratio of ferrocytochrome c to ferricytochrome c plotted against the log of the ratio of ferrocyanide to ferricyanide for free cytochrome c (red) and DNA-PNA-CytC (green). Also shown is a linear fit of each dataset and the calculated redox potential values.

of this plot determines the potential at equilibrium between ferri and ferrocytochrome c, which can be used to determine the redox potential using the Nernst equation. The calculated potential for free horse heart cytochrome c was 278 ± 0.6 mV at 23 ^{0}C, which is

about 10 mV higher than reported elsewhere,[144] and may be due to a systematic offset in our experimental setup. The calculated redox potential for PNA-CytC in solution was found to be 267 ± 1.6 mV at 23 ^{0}C. The experiment for the DNA-PNA-CytC complex was carried out at 6 ^{0}C to ensure the PNA-CytC stayed hybridized to the DNA nanocage. The redox potential of cytochrome c is known to be inversely dependent on temperature (–0.45 mV / ^{0}C), which has been explained due to the higher ordering of water at low temperature stabilizing the cytochrome c conformation in the reduced state.[145,146] The calculated redox potential for PNA-CytC when incorporated into the DNA nanocage was found to be 266 ± 1.7 mV at 6 ^{0}C with an equivalent potential of 258 ± 1.7 mV at 23 ^{0}C.

The 10 mV negative shift in the cytochrome c redox potential observed after conjugation to PNA and additional 10 mV shift observed once PNA-CytC is incorporated into the DNA nanocage may both be due to a less positive net charge of the protein, which helps to stabilize the net positive charge of the heme in the oxidized state. The positive charges from some surface lysine residues are displaced during PNA conjugation. Furthermore, close interaction of cytochrome c with the negatively charged DNA nanocage may neutralize some positive surface charges on the protein. This shows a similar trend to that reported for ferrocene conjugated to ssDNA upon hybridization with a complementary ssDNA.[147,148] The DNA microenvironment was also shown to effect the activity of a β-lactamase enzyme attached to a giant λ DNA (48.5 kbp) strand when the DNA was toggled between an unfolded and compact state.[149] This suggests that DNA nanostructures may be able to modulate the activity of proteins through a variety of mechanisms. Together, the activity and redox potential data indicate that the preparation

of the PNA-CytC conjugates as well as the incorporation of the conjugates in a DNA nanocage the maintain functionality of the protein.

Cytochrome c secondary structure

Circular dichroism (CD) of cytochrome c was used to determine the effect of conjugating PNA and hybridizating into the DNA nanocage on the secondary structure of the protein, as described in the Materials and Methods section called Circular dichroism of cytochrome c. The UV CD spectrum of cytochrome c and PNA-CytC (Figure 2-23A) both show similar negative-in-sign peaks at 222 nm and 208 nm and a positive-in-sign peak at 195 nm in agreement with the literature,[150] suggesting that the overall secondary structure of cytochrome c is unaffected by PNA conjugation. CD of the Soret band has also been used as a structural probe of the bond strength of the axial Met(80) ligand to the heme iron in the oxidized state of cytochrome c.[151] A pronounced negative-in-sign peak at 417 nm occurs in the oxidized state when Met(80) is in close proximity to the heme iron, but which disappears upon denaturation. The Soret CD spectrum of cytochrome c, PNA-CytC and DNA-PNA-CytC (Figure 2-23B) all show a negative-in-sign peak at 417 nm suggesting cytochrome c retains its native conformation around the heme active site after conjugation with PNA and incorporation into the DNA nanocage. This is further supported by nearly identical heme absorption spectra in both the oxidized and reduced states, as shown in Figure 2-22.

Conclusion

We have demonstrated gentle methods to conjugate, purify and assemble protein into a 3D DNA nanocage using a PNA linker. The results obtained with cytochrome c suggest that the methodology preserves protein functional activity. Toehold-mediated DNA strand displacement is introduced as an effective method for purifying PNA and enriching PNA-protein conjugates. Each of the PNA-protein conjugates was assembled into separate DNA nanocages within 2 minutes at room temperature, and within 4 minutes at 11 ^{0}C. The presence of protein increased the PNA solubility, allowing the PNA-protein conjugate to bind to the DNA nanocage in near stoichiometric quantities, with faster binding kinetics and with increased hybridization efficiency than the same PNA sequence conjugated to a short, neutral peptide.[79] The effect of the surface charge of two different proteins on their interactions with a DNA nanocage was also investigated. Gel electrophoresis, steady state and time-resolved fluorescence spectroscopy data were used to propose a model of the DNA-PNA-protein complexes. Our model suggests that the negatively charged azurin is repelled away from the DNA nanocage, whereas the positively charged cytochrome c remains within and closely interacts with the DNA nanocage. When incorporated into the DNA nanocage, cytochrome c maintained catalytic activity and its redox potential was reduced by 20 mV possibly due to neutralization of positive surface charges during PNA conjugation and from close interaction with the DNA nanocage.

This report describes a flexible approach to studying protein function and interactions within a 3D nucleic acid framework. The design could be further expanded to connect up to 4-6 polypeptides (up to 60 kDa)[50] on adjacent edges of the DNA nanocage by

introducing the appropriate PNA binding sequences. The ability to systematically introduce different proteins into defined locations within such a three dimensional structure may facilitate investigations of protein-protein interactions. The maturation of the DNA nanotechnology field has greatly simplified the process of designing a wide variety of DNA nanostructures. Our method of incorporating proteins into 3D DNA nanostructures by means of a PNA linker facilitates exploiting the enormous diversity of DNA nanostructures for investigating and engineering functional 3D protein-nucleic acid complexes.

CHAPTER 5 – Purification and assembly of Cy5 labeled γ-PNAs into a 3D DNA nanocage

Adapted with permission from Flory, J. D.; Johnson, T.; Simmons, C. R.; Ghirlanda, G.; Fromme, P. Manuscript submitted for publication to *Artificial DNA* **2014**.[152]

Abstract

PNA is hybrid molecule ideally suited for bridging the functional landscape of polypeptides with the structural diversity that can be engineered with DNA nanostructures. However, PNA can be more challenging to work with in aqueous solvents due to its hydrophobic nature. A 3D DNA nanocage was designed with binding sites for two fluorescently labeled PNA strands in close proximity. Denaturing polyacrylamide gel electrophoresis (PAGE) is introduced as an efficient method for purifying charged, dye-labeled PNA conjugates from large excesses of unreacted dye and unreacted, neutral PNA. Elution from the gel in water was monitored by fluorescence and found to be more efficient for the more soluble PNA strand. Native PAGE shows that both PNA strands hybridize to their intended binding sites within the DNA nanocage. The dye label increased the solubility of the PNA and the efficiency at which it hybridizes to the DNA nanocage.

Introduction

DNA nanostructures are an intriguing platform for manipulating molecules and materials on the nanometer scale,[15,133,153] including proteins and peptides.[134,135] DNA nanostructure assembly is governed mostly by Watson-Crick base pairing[11] and therefore is more straightforward to engineer than proteins or peptides.[17,18] Chemical strategies are

routinely used for conjugating nucleic acids to peptides or proteins in order to enhance the functionality of unmodified DNA. However, the number of efficient bioorthogonal coupling strategies can become limiting for creating more highly interconnected peptide complexes that could be used to mimic protein catalytic sites. Alternatively, nucleobase hybridization can be used to leverage a few robust chemical coupling methods for conjugating a number of polypeptides, each to a unique nucleic acid strand, which hybridize to specific binding sites within the DNA scaffold.

PNA[66] (peptide nucleic acid) is an intriguing synthetic polymer for bridging the diverse functional landscape of polypeptides with the straight-forward construction of nucleic acid nanostructures. Like peptides, PNAs are linked by amide bonds and can be synthesized using similar solid phase protocols,[67] thus allowing for seamless synthesis of PNA-peptide conjugates.[68] Applications have been recently reviewed[71] including a review focused on chemical modifications to PNA.[72] Its synthetic backbone makes it highly resistant to nucleases and proteases[73] and useful for targeting DNA[70] and RNA,[74] as well as using it as a molecular probe[75] and biosensor.[76,77] PNA has also been assembled within DNA scaffolds to investigate applications in bionanotechnology.[47,78,79] Previously we showed the versatility of using PNA linkers for assembling peptides[79] and proteins[132] into a 3D DNA nanocage. The assembly process was rapid and could be performed at or below room temperature, providing mild conditions that preserved the function of the attached polypeptides. The previous PNA sequence design provided modest thermal stability and could be functionalized at the N terminus. Substitutions to the gamma position on the PNA backbone, as shown in Figure 1-6, have been shown to increase the solubility and thermal stability of the PNA strand,[80] as well as provide a handle to introduce func-

tional groups without affecting its binding to a complementary strand.[81] Functionalizing PNA at the gamma position would also allow a peptide to be connected anywhere along the backbone or even at multiple locations.[47,82]

Here we report a 3D DNA nanocage design for binding two PNAs fluorescently labeled with Cy5. The PNAs were labeled using strain-promoted, copper-free click chemistry and purified using denaturing PAGE. Native PAGE was used to characterize the DNA nanocage when populated with either of two Cy5 labeled PNAs.

Results and Discussion

DNA-PNA complex design

Previously we incorporated two fluorescently labeled peptides[79] or proteins[132] within a 3D DNA nanocage using PNA linkers. Briefly, we modified an existing DNA tetrahedron design by introducing 8nt single stranded domains complementary to a PNA strand functionalized on the N terminus with a short peptide or small protein oriented toward the center of the DNA nanocage. This design allowed the DNA nanocage to be preassembled, purified and then subsequently populated with one or two PNA-polypeptides under mild conditions. Here we introduce a third 8nt PNA binding site (PNA3), to be used in conjunction with the PNA1 binding site for introducing guest molecules, as shown in Figure 5-1. The PNA3 binding site is in closer proximity to the PNA1 binding site (3 nm) than a similar PNA binding site on edge F used in the previous design (5–6 nm),[79] which is better suited for assembling protein active sites. Although the PNA sequences in our previous design were stable at room temperature[79], we sought to increase the thermal stability to make a more robust artificial OEC complex without significantly

changing the DNA nanocage design (i.e., the length of PNA binding sites). PNA sequences generally bind with higher thermal stability to a complementary DNA strand than a similar DNA sequence. The neutral charge of the PNA backbone provides some added stability, but sequences rich in purines can provided significantly higher thermal stabilities over the same sequence of DNA.[126] We updated the PNA1 and PNA3 binding sequences in order to double the purine content of PNA sequences from 37.5% (3 of 8) to

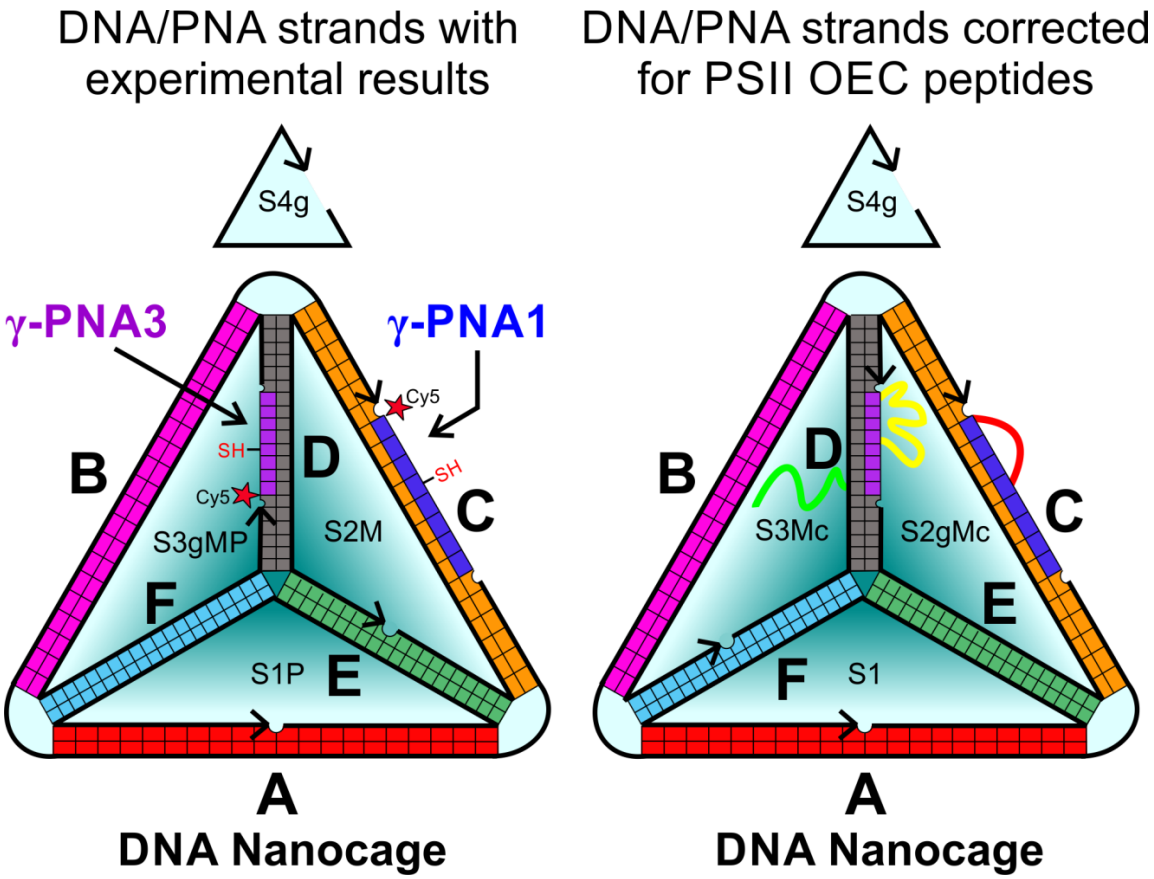


Figure 5-1. Schematics showing the DNA and PNA sequence arrangement in the complexes. γ-PNA1 and γ-PNA3 are labeled with Cy5, thiol functionalized at the gamma position of 4th PNA base, and attached to edge C and D respectively. This schematic was redrawn from[65] with changes to incorporate the gamma modified PNAs. The ssDNA strand nicks are indicated with the arrow pointing in the 3' direction. Strand S4g is redrawn above for clarity. [Left] DNA and γ-PNA strands used for the experiments described in the text. [Right] Sequences corrected to orient the peptide attached to γ-PNA3 toward the peptide attached to γ-PNA1 to more closely mimic the PSII active site, as shown in Figure 6-4.

Table 5-1. DNA, PNA and peptide strand names and sequences. DNA, PNA and peptide sequences with the color code matching the schematics in Figure 5-1. The DNA and PNA sequences are written in the 5' to 3' direction. The γ-PNA1 binding sequence is highlighted in blue and the γ-PNA3 binding sequence is highlighted in grey. N_3 indicates the PNA N terminal azide functionalization, and Lys(C3SH) is a thiol introduced to a lysine residue at the gamma position of the 4th PNA base. The peptides sequences are written using one letter amino acid codes from N to C termini. For each peptide, the N terminus is either free (NH_2) or maleimide functionalized, and the C terminus is either free (OH) or amidated ($CONH_2$). Ligands for the Mn_4O_5Ca cluster from the first coordination sphere of the OEC are enclosed in curly brackets {}. Dye labeled and thiolated DNA strands are defined using the nomenclature for ordering from IDT, where the dye names and thiol functionalization are enclosed by "/" marks and are attached to a modified thymine base at the BCD vertex for the dyes or to the 7th base from the DEF vertex for the thiol.

Strand Name	Sequence
S1	AGGCAGTTGAGAC-GAACATTCCTAAGTCTGAAATTTATCACCCGCCATAGTAGAC-GTATCACC
S1P	AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCCGAG-TGACTAGTAGACGTATCACC
S2M	GAATGTTCGACATGCGCCATCGCTATACCGA CGATTATCGCCGTTTACACGATTCAGACTTAG
S2M-FAM	GAATGTTCGACATGCGCCATCGCTATACCG/iFluorT/ CGATTATCGCCGTTTACACGATTCAGACTTAG
S2M-TMR	GAATGTTCGACATGCGCCATCGCTATACCG/i6-TAMN/ CGATTATCGCCGTTTACACGATTCAGACTTAG
S2M-Cy3	GAATGTTCGACATGCGCCATCGCTATACCG/iCy3/ CGATTATCGCCGTTTACACGATTCAGACTTAG
S2gMc	TACACGATTCAGACTTAGGAATGTTCGACATGCGCCATCGCTATAC-CGACGATTA
S3	GGTGATAAAACGTGTAGCAAGCTGTAATCGACGGGAAGAG-CATGCCCATCCACTACTATGGCG
S3MP	GGCGATAATCGACGGGAAGAGCATGCCCATCCACTACTAGTCAC-TCGGATAAAACGTGTAAAC
S3gMP	TAATCGACGGGAAGAGCATGCCCATCCACTACTAGTCACTCGGA-TAAAACGTGTA
S3Mc	CGGGTGATAAAACGTGTATCGCCGTTTAATCGA CGGGAAGAGCATGCCCATCCACTACTATGG
S3Mc-SH	CGGGTGATAAAACGTGTA/iThioMC6-D/CGCCGTTTAATCGA CGGGAAGAGCATGCCCATCCACTACTATGG
S4M	TGGCGCATGACTCAACTGCCTGGTGATACGAGGATGGG-CATGCTCTTCCCGACGGTATAGCGA
S4g	CGCATGACTCAACTGCCTGGTGATACGAGGATGGG-CATGCTCTTCCCGACGGTAT
γ-PNA1	N_3-AGCG(Lys-C3SH)ATGG-$CONH_2$
Red Peptide	Maleimide-GEIIFGG{E}TMRFWDK-$CONH_2$
γ-PNA3	N_3-AACG(Lys-C3SH)GCGA-$CONH_2$
Yellow Peptide	NH_2-AVFLI{Y}PIGQGSFS{D}GMPLGISGTFNFMIVFQA{E}HNILMC-$CONH_2$
Green Peptide	Maleimide-ADIINRANLGMEVM{HE}RNAHNFPL{D}L{A}-OH

Table 5-2. Complex nomenclature with list of constituent DNA and γ-PNA(-peptide) strand names. In the DNA-pep and DNA-PNA-aOEC complexes, strand S3Mc-SH is conjugated to the green peptide.

Complex Name	Strands Required to Assemble Construct
DNA1γ	S1, S2M, S3, S4g
DNA1γ-FAM	S1, S2M-FAM, S3, S4g
DNA1γ-TMR	S1, S2M-TMR, S3, S4g
DNA3γ	S1P, S2M, S3gMP, S4M
DNA3γ-FAM	S1P, S2M-FAM, S3gMP, S4M
DNA3γ-TMR	S1P, S2M-TMR, S3gMP, S4M
DNA3γ-Cy3	S1P, S2M-Cy3, S3gMP, S4M
DNA1γ + γ-PNA1	S1, S2M, S3, S4g, γ-PNA1
DNA1γ + γ-PNA1-Cy5	S1, S2M, S3, S4g, γ-PNA1-Cy5
DNA3γ + γ-PNA3	S1P, S2M, S3gMP, S4M, γ-PNA3
DNA3γ + γ-PNA3-Cy5	S1P, S2M, S3gMP, S4M, γ-PNA3-Cy5
DNA-pep	S1, S2gMc, S3Mc-SH(-pep), S4g
DNA-PNA-aOEC	S1, S2gMc, S3Mc-SH(-pep), S4g, γ-PNA1-pep, γ-PNA3-pep

75% (6 of 8), which is predicted to increase the dissociation temperature from about 37 ^{0}C to 49 ^{0}C and 53 ^{0}C, respectively.[95] Throughout this report we use the nomenclature defined in Figure 5-1, Table 5-1 and Table 5-2 when we describe the various strands and constructs used in each experiment.

Fluorescently labeled and unlabeled DNA nanocage assembly

Unlabeled DNA nanocages, for binding either γ-PNA1 (DNA1γ) or γ-PNA3 (DNA3γ), were assembled and purified by SE-FPLC, as previously described.[79] DNA nanocages were also assembled that were fluorescently labeled on the vertex shared between γ-PNA1 and γ-PNA3 with fluorescein (DNA1γ-FAM, DNA3γ-FAM) or tetramethylrhodamine (DNA1γ-TMR, DNA3γ-TMR) to characterize the γ-PNA binding ki-

netics and thermal stability and using Förster resonance energy transfer (FRET) with a γ-PNA labeled with tetramethylrhodamine (TMR) or fluorescein (FAM), respectively, as previously described.[79] The left panel of Figure 5-2 shows the native PAGE of the each of the fluorescently labeled and unlabeled DNA nanocages.

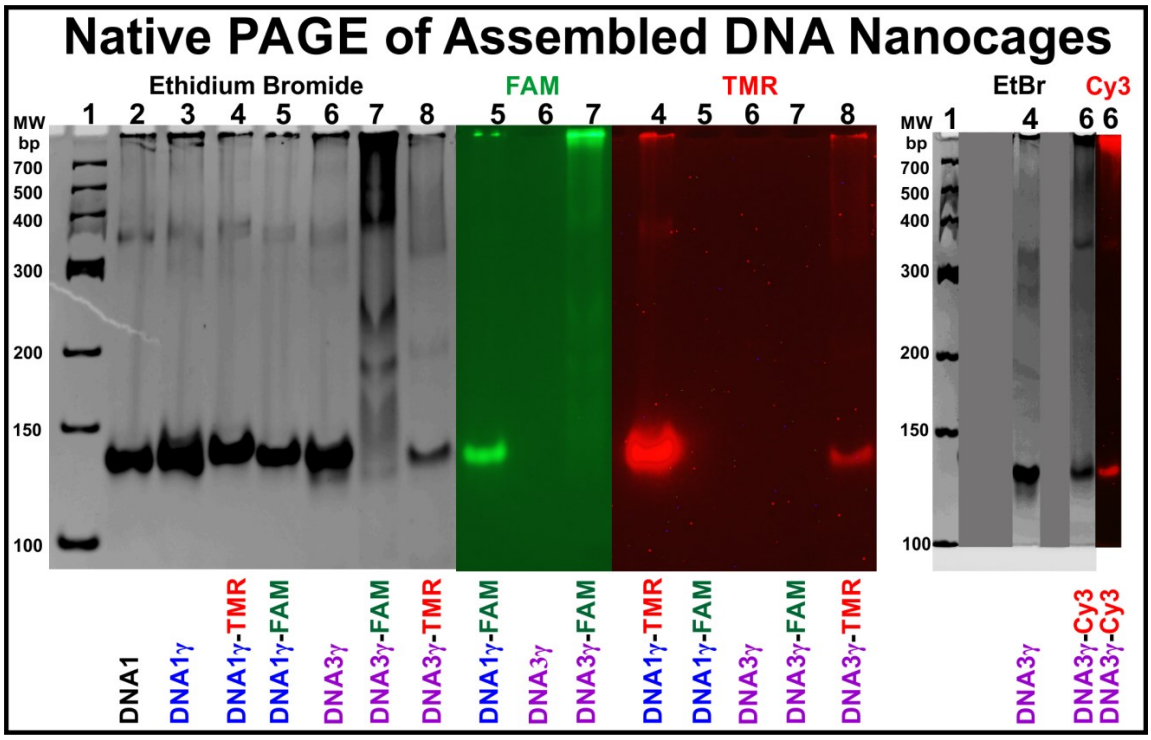


Figure 5-2. Native PAGE (7%) of labeled and unlabeled DNA nanocages for binding two γ-PNAs. [Lane 2] DNA1 is the unlabeled DNA nanocage with a binding site for a standard PNA1 sequence reported previously[79], as a control. DNA1γ is the same DNA nanocage with the binding sequence updated for γ-PNA1, shown unlabeled [Lane 3], labeled with TMR [Lane 4] or with FAM [Lane 5]. DNA3γ has the binding site moved to an adjacent edge with a sequence for binding γ-PNA3, shown unlabeled [Lane 6], labeled with FAM [Lane 7] or with TMR [Lane 8]. The ethidium bromide fluorescence (left panel), FAM fluorescence (middle panel) and TMR fluorescence (right panel) are shown. [Right] The DNA3γ nanocage unlabeled [Lane 4] and labeled with Cy3 [Lane 6]. The ethidium bromide (EtBr) fluorescence (left panel) and Cy3 fluorescence of lane 6 (right panel) are shown.

With exception of DNA3γ-FAM, all constructs assembled correctly and show a pronounced band that migrates at the same rate as a positive control (DNA1) described previously.[79] DNA3γ-FAM showed significant aggregation during assembly without any

dominant product. DNA3γ-TMR also showed more aggregation than was observed in the unlabeled DNA3γ construct or than in any of the DNA1γ constructs. This suggests that the assembly of DNA3γ may be more sensitive to dye functionalization at the vertex. To circumvent this issue, we tried assembling DNA3γ labeled with Cy3, which can be used for FRET studies by labeling the PNA with the complementary dye Cy5. The right panel of Figure 5-2, shows native PAGE of the DNA3γ labeled with Cy3, which assembled correctly and shows a defined band with the expected migration. Both gels in Figure 5-2 also show the native fluorescence of FAM, TMR or Cy3 for the assembled complexes.

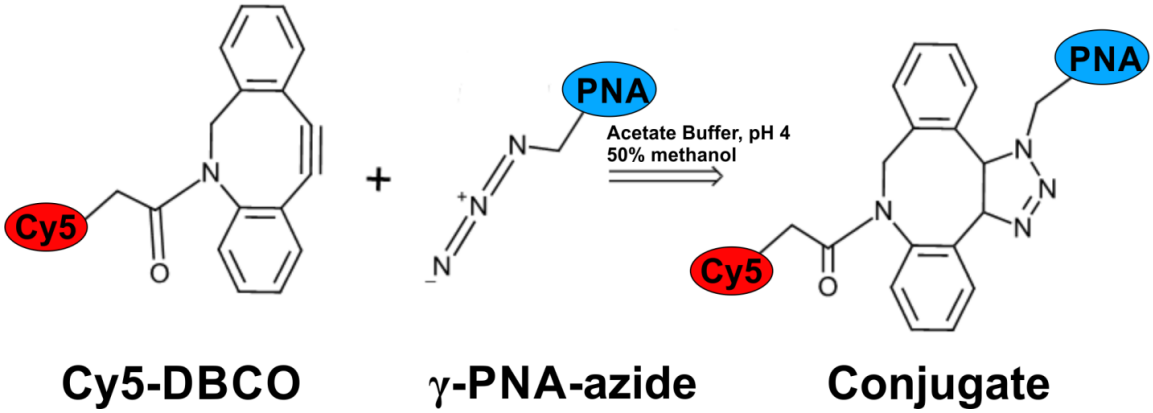


Figure 5-3. Scheme for labeling γ-PNAs with Cy5-DBCO using copper-free click chemistry. Cy5 functionalized with DBCO was conjugated to γ-PNA functionalized with an N terminal azide, as described in the methods section.

Fluorescent labeling and purification of γ-PNA

Cy5 functionalized with the strained alkyne dibenzocyclooctyl (DBCO) was used to label each γ-PNA at the N terminal azide position, as shown in Figure 5-3 and described in the methods section. Strained alkynes can be efficiently conjugated to azides without adding any copper catalyst,[154,155] which could interfere with assembling metallopeptide complexes and can be difficult to completely remove.[155,156] We were unable to separate the unreacted Cy5 from the γ-PNA-Cy5 conjugate using reverse-phase high

pressure liquid chromatography (RP-HPLC). We had similar problems when trying remove unreacted fluorescein (FAM) from γ-PNA-FAM, which suggests the dye labeled γ-PNA sequences may strongly interact with the unreacted dyes[157,158] or they have similar levels of hydrophobicity. Denaturing PAGE has been shown as an alternative to RP-HPLC for purifying unmodified PNA under acidic conditions (pH 3) with single-base resolution.[159] We adapted this technique to work under slightly alkaline conditions (pH 8.5) to take advantage of the difference in mobility in an electric field of unlabeled (neutral) PNA and negatively charged γ-PNA-Cy5 conjugates, as well as the significant difference in size between the unreacted Cy5 (1 kDa) and the γ-PNA-Cy5 conjugates (~3.5 kDa). Figure 5-4 shows the denaturing PAGE gel after the reaction of Cy5 with γ-PNA3. The unreacted Cy5 migrates significantly faster than the γ-PNA-Cy5 conjugate. matrix assisted laser desorption ionization mass spectrometry (MALDI-MS) of the other weakly fluorescent bands were found to contain impurities.

The bands for the γ-PNA-Cy5 conjugates were cut from the gel and eluted with nanopure water up to four times to extract the product, as detailed in Table 5-3. The fluorescent signal remaining in the gel blocks after each elution step, for both γ-PNA3-Cy5 and γ-PNA1-Cy5, was integrated and fit with an exponential function, as shown in Figure 5-4, and described in the methods section. Both samples are efficiently eluted after 3–4 elution steps. Despite a lower percent gel with larger pores, γ-PNA1-Cy5 elutes with slightly lower efficiency, which may be due to its reduced solubility compared to γ-PNA3-Cy5. Although the times allowed for elution during each step were significant different (1.25–115 hours), it appears that equilibrium was still achieved after each elution step, which suggests that the elution could be performed in much less time (i.e., 4 elu-

tions over 6 hours). The eluted product was desalted using RP-HPLC. The expected molecular weight of each collected product was detected by MALDI-MS, as shown in Figure 5-5 for γ-PNA1-Cy5 and Figure 5-6 for γ-PNA3-Cy5. Both samples show an impurity +70 Da from the expected molecular weights. However, native PAGE of the γ-PNA1-Cy5 shows a single species, which suggests the impurity is minor. Native PAGE analysis of γ-PNA3-Cy5 shows the expected band as well as a slower moving band, which may be dimerization of γ-PNA3-Cy5 via an intramolecular disulfide between thiols at the gamma position. The dimer may be the result of insufficient disulfide cleavage by tris(2-carboxyethyl)phosphine (TCEP) prior to PAGE analysis.

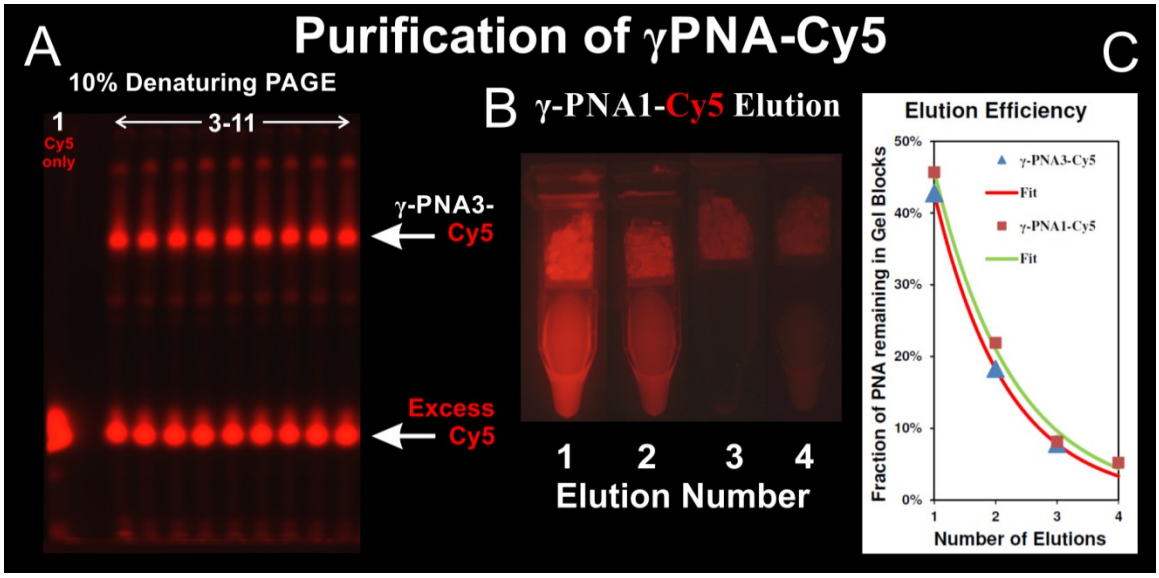


Figure 5-4. Purification of Cy5 labeled γ-PNAs using denaturing PAGE and time course elution from gel blocks. [A] Cy5 fluorescence image of a 10% denaturing PAGE gel containing a control with Cy5 only (Lane 1) and the γ-PNA3 and Cy5 reaction cocktail (Lanes 3-11). [B] Cy5 fluorescence image of the of γ-PNA1-Cy5 bands cut from a 7.5% gel, similar to that shown in [A], when eluted with water and centrifuged in Spin-X™ tubes up to 4 times to remove the product eluted from the gel blocks. [C] The fluorescence intensity of the gel blocks after each elution in [B] was used to determine the fraction of PNA remaining in the gel blocks for γ-PNA1-Cy5 and γ-PNA3-Cy5. Each curve was fit with an exponential decay function as described in the methods section. The parameters for the fitting functions are shown in Table 5-4.

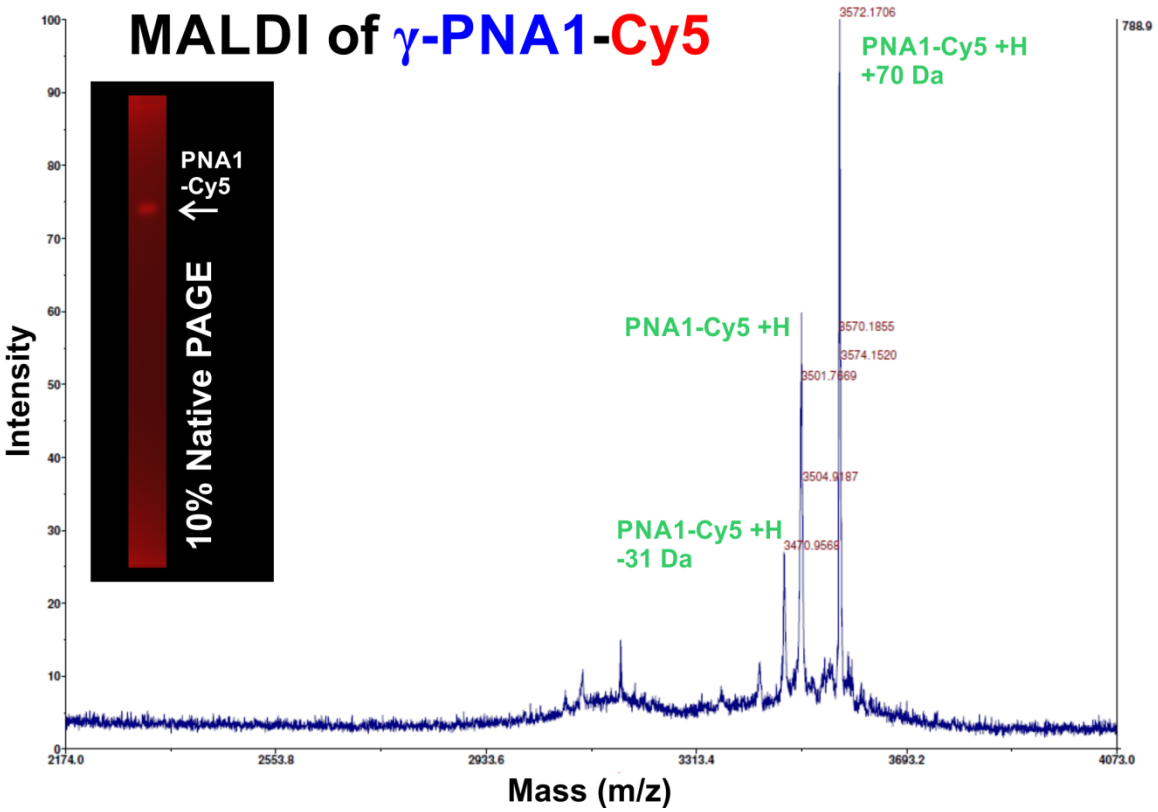


Figure 5-5. Characterization of purified Cy5 labeled γ-PNA1 by MALDI-MS and native PAGE. MALDI-MS spectrum of the purified γ-PNA1-Cy5 product. The measurement was performed in reflector mode using sinapinic acid as the matrix. The measured mass of 3,500.2 Dalton (Da) agrees with the expected mass for γ-PNA1-Cy5 (M+H) within –0.1 Da. The spectrum also shows two impurities 3,570 Da (M+H+70) and 3,469 Da (M+H–31). The inset shows a 10% native PAGE of the purified γ-PNA1-Cy5.

The expected molecular weight of each collected product was detected by MALDI-MS, as shown in Figure 5-5 for γ-PNA1-Cy5 and Figure 5-6 for γ-PNA3-Cy5. Both samples show an impurity +70 Da from the expected molecular weights. However, native PAGE of the γ-PNA1-Cy5 shows a single species, which suggests the impurity is minor. Native PAGE analysis of γ-PNA3-Cy5 shows the expected band as well as a slower moving band, which may be dimerization of γ-PNA3-Cy5 via an intramolecular disulfide between thiols at the gamma position. The dimer may be the result of insufficient disulfide cleavage by TCEP prior to PAGE analysis.

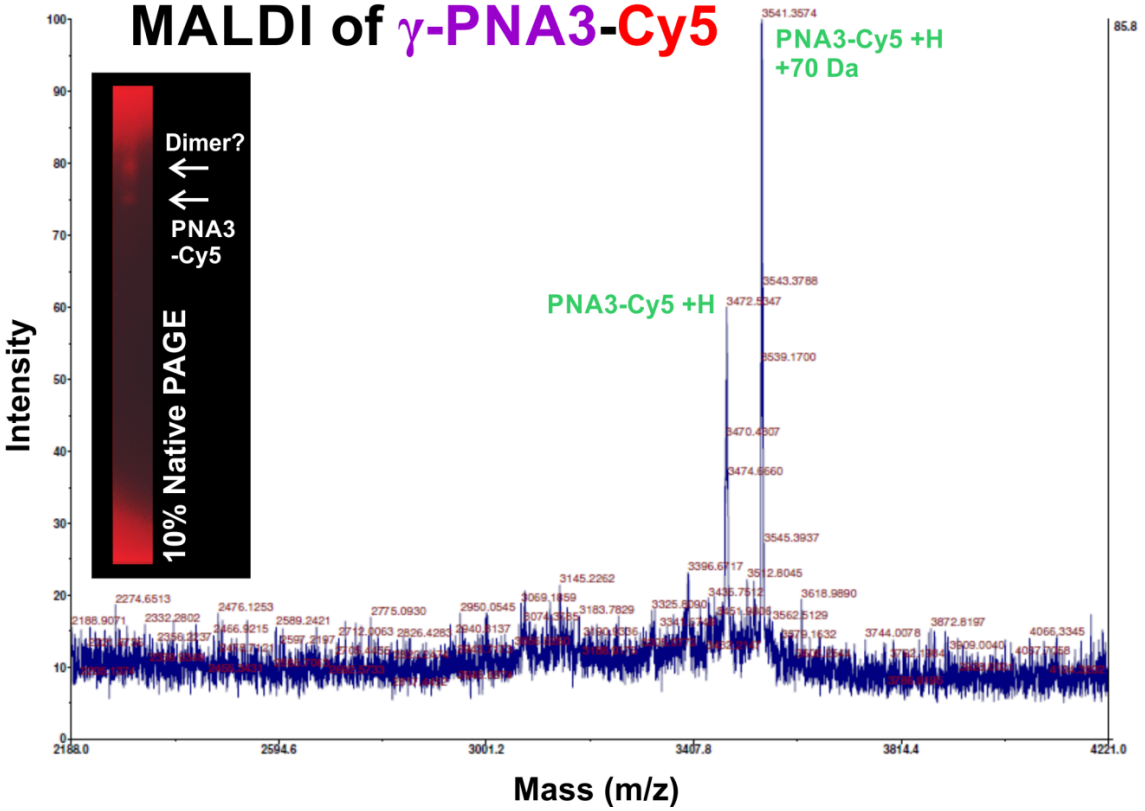


Figure 5-6. Characterization of purified Cy5 labeled γ-PNA3 by MALDI-MS and Native PAGE. MALDI-MS spectrum of the purified γ-PNA3-Cy5. The measurement was performed in reflector mode using sinapinic acid as the matrix. The measured mass of 3,470.5 Dalton (Da) agrees with the expected mass for γ-PNA3-Cy5 (M+H) within +1.3 Da. The spectrum also shows an impurity with 3,539 Da (M+H+70). The inset shows a 10% native PAGE of the purified γ-PNA3-Cy5.

Table 5-3. Progress of γ-PNA-Cy5 elution from gel blocks. For each elution, the table shows the number of hours allowed for each product to elute out of the gel blocks (Elution Time) prior to centrifugation and analysis, the fraction of product remaining in the gel blocks determined from the Cy5 fluorescence of the gel blocks, as well as the corresponding fraction of product eluted.

Strand	Elution Number	Elution Time (hrs)	Fraction in Gel Blocks	Fraction Eluted
γ-PNA1-Cy5 (7.5% gel)	1	44	46%	54%
	2	26	22%	78%
	3	68	8%	92%
	4	1.25	5%	95%
γ-PNA3-Cy5 (10% gel)	1	115	43%	57%
	2	6	18%	82%
	3	18	8%	92%

Table 5-4. Fitting parameters for of γ-PNA1-Cy5 and γ-PNA3-Cy5 elution. The fraction of product remaining in the gel blocks after each elution was fit to an exponential function with the parameters listed here, as described in the methods section. Because only three data points were collected for γ-PNA3-Cy5, error bars for each parameter and a goodness-of-fit value (R^2) could not be determined (ND).

Strand	y_0	A	R_0	R^2
γ-PNA1-Cy5	892 ± 898	22,074 ± 2,725	0.7798 ± 0.177	0.989
γ-PNA3-Cy5	4,382 ± ND	30,868 ± ND	0.8498 ± ND	ND

Assembly of the DNA nanocages with fluorescently labeled γ-PNAs

DNA nanocages, for binding either γ-PNA1 (DNA1γ) or γ-PNA3 (DNA3γ), were assembled and purified by ion-exchange fast protein liquid chromatography (IEX-FPLC), and incubated with increasing molar excess (0–3x or 10x) of the corresponding γ-PNA sequence at room temperature, as described previously.[79] The assembled complexes were analyzed directly by native PAGE at 4 ^{0}C, as shown in Figure 5-7. A small gel shift is observed for the DNA nanocage upon binding each γ-PNA, with respect to the empty DNA nanocage, and is attributed to the single stranded γ-PNA binding domain becoming double stranded and more rigid, as observed previously.[79] A 2-5x excess of unlabeled γ-PNA is required to achieve quantitative binding to the DNA nanocage. This is similar to the 2x excess of PNA-peptide required for quantitative binding to the same DNA nanocage.[79] Some PNA sequences show mild aggregation at concentrations as low as 1 μM,[80] which could prevent binding of some PNA to the DNA nanocage. The larger excess required for γ-PNA1 may be due to its reduced solubility compared to γ-PNA3. However, when γ-PNA1 is labeled with the negatively charged Cy5, only 2x excess is required for quantitative binding to the DNA nanocage, as shown in the right panel of Figure 5-7, suggesting the Cy5 label may improve the solubility of γ-PNA1. Figure 5-7

also shows the Cy5 fluorescence of the DNA nanocages containing the Cy5 labeled γ-PNAs, as well as a more diffuse band containing some unhybridized γ-PNA-Cy5. These results indicate the proper assembly of the γ-PNAs into the DNA nanocage before and after Cy5 labeling.

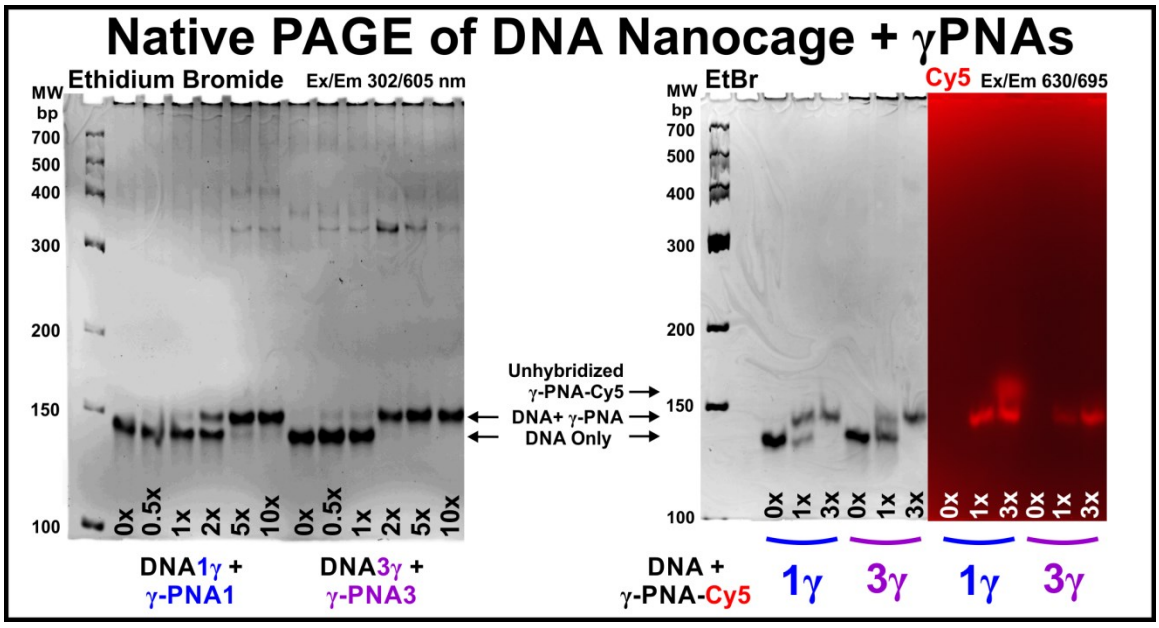


Figure 5-7. Gel showing the binding of the unlabeled and labeled γ-PNAs to the DNA nanocage. [Left] Ethidium bromide image of the unlabeled DNA1γ and DNA3γ nanocages with increasing amount of unlabeled γ-PNA1 and γ-PNA3, respectively. [Right] The unlabeled DNA1γ and DNA3γ nanocages with increasing amount of Cy5 labeled γ-PNA1 and γ-PNA3, respectively. The ethidium bromide (EtBr) fluorescence (left panel) and Cy5 fluorescence (right panel) are shown, with the corresponding excitation and emission wavelengths (Ex/Em).

Conclusion

We introduced denaturing PAGE as a method to efficiently purify PNA fluorescently labeled with Cy5. The fluorescence of the product was used to monitor its elution from the gel blocks, which allows the method to be extended to other charged molecules that may not be so easily detected. The Cy5 labeled γ-PNAs were assembled into DNA nanocages under benign conditions that were previously shown to preserve polypeptide

function.[132] A perspective application of using the DNA nanocage design for assembling cyclic PNA-peptides is described, which offer a promising approach for constructing artificial catalytic centers.

CHAPTER 6 – Prospective: Building a peptide-based artificial Oxygen-Evolving Complex within a 3D DNA nanocage using γ-PNA linkers

Abstract

While photosynthetic organisms offer a promising route to generate sustainable transportation fuels using solar energy, their low efficiency in converting sunlight to fuel presents a significant challenge for economic viability. A major limitation for high efficiency is photosystem II (PSII). The PSII reaction center-pigment-protein complexes catalyze the oxidation of water to produce all the oxygen in the Earth's atmosphere and enable conversion of solar energy to bioenergy for the entire biosphere. Under full sunlight the native PSII complexes are damaged every 30 minutes due to the accompanying chlorophyll-based photochemistry, resulting in the loss of 60% of the captured photons per day to heat.[160] The site of water oxidation is called the oxygen-evolving complex (OEC) and one could hypothesize that it may function independently of the PSII complex when reconstructed in a more stable environment. Three polypeptide domains from PSII are responsible for coordinating the metal cluster that catalyzes the water oxidation reaction. This chapter discusses a potential design concept for assembling these three OEC polypeptides from PSII into a DNA nanocage using γ-PNA linkers to form an artificial OEC. This artificial OEC has the potential to leverage the high catalytic activity of PSII for oxidizing water, but, by integrating it within an artificial system, may avoid the damage that limits the natural OEC within PSII.

Discussion

Natural Photosynthesis

The evolution of photosystem II (PSII) 2.5 billion years ago dramatically changed the makeup of life on earth by enabling photoautotrophic growth from abundant sources of water and sunlight. The overall chemical reaction catalyzed by PSII is: $2H_2O$ + light $\rightarrow O_2 + 4H^+ + 4e^-$. PSII provides the electrons to the photosynthetic electron transport chain shown in Figure 6-1. The light-driven electron transfer reactions in photosystem I (PSI) and PSII lead to the formation of an electrochemical gradient that drives synthesis of the high energy products ATP and NADPH, which are finally used to reduce atmospheric CO_2 into carbohydrates to produce all of biomass on earth. The synthesis of all complex organic compounds used for energy storage and other critical life processes depend on photosynthesis. As molecular oxygen, which is the byproduct from water splitting by PSII, began to significantly accumulate in the atmosphere about 500 million years ago,[161] it caused a large extinction of organisms that could not adapt to the presence of this powerful oxidant, but at the same time provided a large driving force to evolve complex cellular metabolisms and higher orders of life, such as ourselves.[161]

The light-driven photosynthetic electron transport chain, shown in Figure 6-1, is localized within the thylakoid membranes of the chloroplast, which separate the lumen from the stroma, and has been reviewed by Renger.[162] PSII uses light energy to catalyze the oxidation of water, in four sequential light-driven charge separation steps, where one proton and one electron are extracted in each step from two substrate water molecules. The protons are released into the thylakoid lumen and contribute to the electrochemical gradient used by ATP synthase to create ATP, which is the energy currency of the cell.

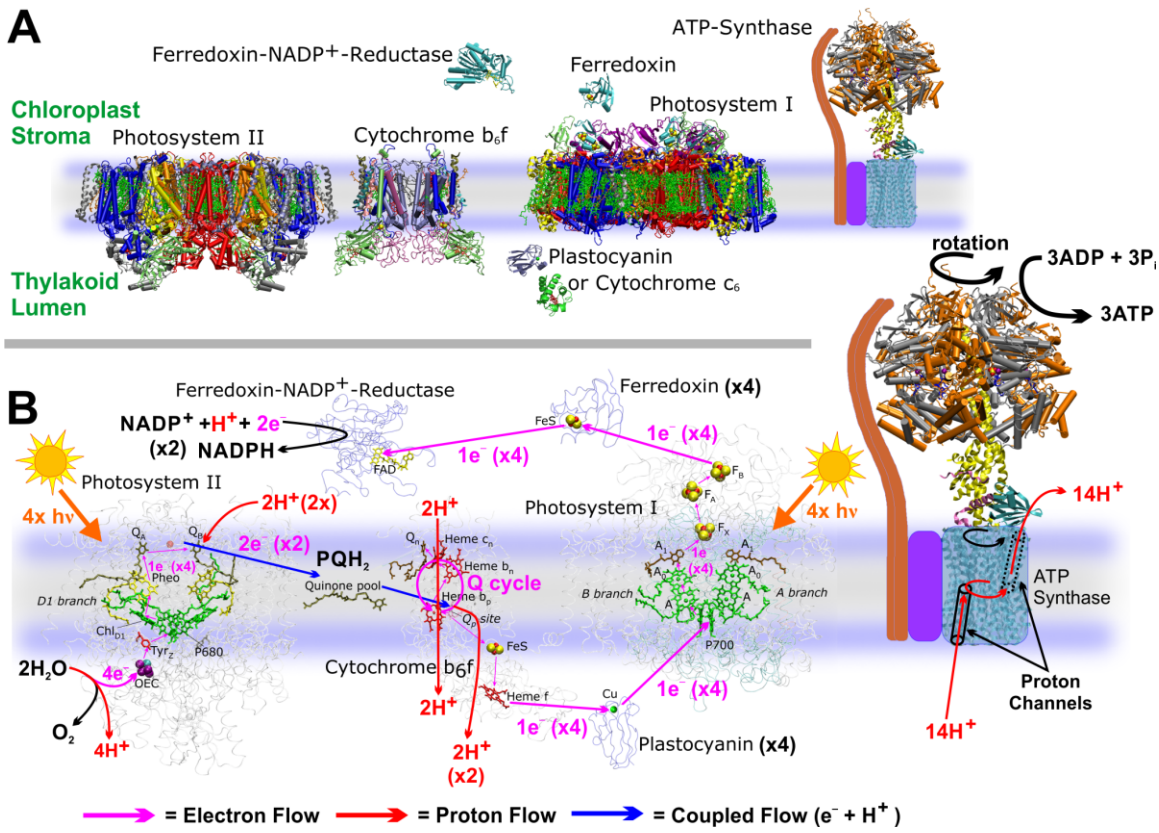


Figure 6-1. Photosynthetic electron transport chain. [A] Structure of the major protein complexes involved in the light reactions of the photosynthetic electron transport chain. [B] Flow of electrons and protons within the photosynthetic electron transport chain. Figure adapted with permission from Fromme et al.,[163] © Wiley 2008.

Light energy form the sun is captured by external and PSII internal antenna complexes and channeled to the center of the PSII complex to what is called the primary donor P680. When P680 is excited to $P680^*$, charge separation takes place and an electron is transported across the membrane to the terminal electron acceptor a plastoquinone (PQ) molecule in the Q_B binding pocket, which is located at the stromal acceptor side of the membrane. After two charge separation events, PQ^{--} takes up two protons from the stromal side and leaves the binding pocket as PQH_2 (i.e., plastoquinol). At the lumenal donor site, $P680^+$ is reduced to P680 by electrons extracted from water molecules bound to the

Mn_4O_5Ca cluster of the oxygen evolving complex (OEC) in PSII via a redox active tyrosine Tyr_Z.

Plastoquinol migrates within the thylakoid membrane to the cytochrome b_6f complex. The b_6f complex releases the two protons into the lumen, and pumps one additional proton per two electrons across the thylakoid membrane in the so-called Q-cycle, thereby further enhancing the electrochemical gradient for ATP production. The cytochrome b_6f complex transfers one electron to the soluble electron carrier protein plastocyanin, which is located in the lumen. Once the reduced plastocyanin diffuses into the lumen, an oxidized plastocyanin docks to the cytochrome b_6f complex to receive the second electron that originated from the bound plastoquinol. Reduced plastocyanin transfers its electron to photosystem I (PSI), which uses light energy to further increase the reducing potential of the electron in order to reduce ferredoxin on the stromal side. The net result of the PSI electron transport chain is to induce transmembrane charge separation from plastocyanin on the lumenal side to ferredoxin (fd) on the stromal side of the membrane. Ferredoxin-NADP reductase (FNR), then accepts two highly reducing electrons from two reduced ferredoxin molecules (fd^{red}) to catalyze the following redox reaction: $2fd^{red} + NADP^+ + H^+ \rightarrow NADPH + 2fd^{ox}$. The net result of the "light reactions" of photosynthesis is to provide highly reducing electrons to make NADPH and to produce an electrochemical gradient across the membrane, which is used by the ATP-synthase to synthesize ATP from ADP and inorganic phosphate (Pi). ATP and NADPH are then used in the "dark reactions" via the Calvin cycle to reduce CO_2 in to carbohydrates, which has been reviewed by Shively et al[164] and Raines.[165]

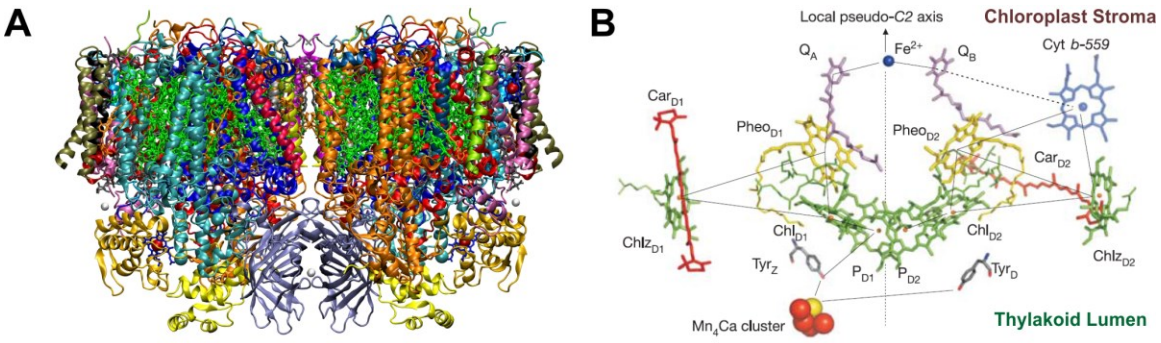


Figure 6-2. Photosystem II structure and electron transport chain. [A] Crystal structure of PSII solved by Umena et al[3] (PDB accession code 3ARC). Figure made from 3ARC by Ingo Grotjohann. [B] Electron transport chain in PSII (PDB accession code 2AXT).[108] Figure [B] reproduced with permission from Loll et al.,[108] © NPG 2005.

Photosystem II: Nature's solar-driven water oxidation catalyst

Light-induced water oxidation by PSII is a remarkable function that is highly desirable for designing artificial catalysts to economically produce fuels from solar energy. The structure of PSII is shown in Figure 6-2A, and its function has been recently reviewed by Vinyard, Ananyev and Dismukes.[166] The protein complex is a dimer, where each monomer consists of 19 protein subunits to which more than 50 cofactors (e.g., chlorophylls, carotenoids, lipids, quinones and the metal complex of the OEC) are bound. The antenna system consists of chlorophyll pigments (shown in green) that funnel the excitation energy to the reaction center, where the PSII cofactors of the electron transport chain are located. When the excitation energy reaches P680 (special chlorophyll molecules) the excited state $P680^*$ is formed, which catalyzes the charge separation event. $P680^*$ serves as the primary electron donor by sending its excited electron up the electron transport chain and then becoming the primary oxidant $P680^+$, which has the most oxidizing potential in any biological system of 1.1 V [$E^{0'}$ vs normal hydrogen electrode (NHE)][167] The PSII electron transport chain is shown in Figure 6-2B, where $P680^+$ is de-

noted as P_{D1}. The electron is quickly funneled away up the left branch through a thermodynamically favorable potential gradient, first to a pheophytin ($Pheo_{D1}$) molecule, then to a tightly bound plastoquinone (Q_A), and finally to a transiently bound plastoquinone (Q_B). After accepting two electrons and picking up two protons from the stroma, plastoquinol (PQH_2) diffuses into the membrane and is replaced by an oxidized plastoquinone to accept another two electrons and protons. The plastoquinols transport their electrons to cytochrome b_6f and on down the photosynthetic electron transport chain. The highly oxidizing $P680^+$ is reduced again by extraction of electrons from two water molecules bound to the Mn_4O_5Ca cluster contained within a domain of PSII called the oxygen-evolving complex (OEC). The electron transfer between $P680^+$ and the OEC is mediated through a special tyrosine amino acid (Tyr_Z). In each charge separation step one electron and one proton are extracted from the two water molecules by the OEC. After four photons are absorbed, and four oxidizing equivalents are stored by the Mn_4O_5Ca cluster in the OEC, molecular oxygen is released. While many models exist for this process[168–171] the exact mechanism of the water oxidation is still unknown.

Soon after dawn, PSII is able to capture and convert enough light energy to completely reduce the PQ pool to PQH_2, such that an average of 60% of the captured light energy each day is wasted as heat.[160] With the PQH_2 pool reduced, electrons from the intermediate states on the reducing side (Q_A^- or Q_B^-) recombine back to $P680^+$ form a triplet state ($^3P680$). $^3P680$ is highly energetic and can significantly damage the protein near the reaction center. $^3P680$ can also react with molecular oxygen ($^3O_2$) via triplet-triplet annihilation to form singlet oxygen ($^1O_2$), which can significantly damage the surrounding protein matrix.[172] After an average of 30 minutes in full sunlight, the damage to the

PSII reaction center accumulates to the point where PSII is inhibited (i.e., photoinhibition) and must be repaired, which involves a complex process including disassembly of large parts of the PSII complex. The repair process involves proteolytic degradation and re-synthesis of the D1 protein that binds most of the cofactors of the PSII electron transport chain, followed by a complex reassembly process that includes light-driven reassembly of the OEC.[173,174] While charge recombination also occurs in artificial systems, such as semiconductor photovoltaics (PV), which results in energy lost as heat, it does not damage the materials as it does in PSII.

Artificial Photosynthesis

Artificial systems can also be used to create simple fuels from sunlight and water.[175] Figure 6-3 shows a PV device, which uses a semiconductor to absorb light and create a potential that can be used to drive a platinum catalyst to electrolyze water and reduce protons to H_2. Artificial systems have some advantages over biological systems. Silicon based PV absorbs photons further into the infrared (700–1100 nm) than photosynthetic pigments, allowing for capture of about twice the number of photons.[176] PV is also about 4-fold more efficient converting solar energy into current than photosynthesis is to produce glucose.[176] Other rare earth metals, such as iridium[177] and ruthenium[178] are also capable of catalyzing water oxidation and with impressive longevity (>10k and >8k turnovers, respectively) and turnover frequencies (0.09 and 300 s^{-1}, respectively), but are not an economical or sustainable solution to create fuels from solar energy.[179] A number of other homogenous[180] and heterogenous[181,182] water oxidation catalysts have also been constructed, but which have not been able to match both the fast turnover frequency (60–

400 s^{-1}) and longevity (100–700k turnovers) of PSII.[183,184] Our current understanding of PSII has helped to determine design principles for constructing solar-driven artificial water splitting systems,[180] however, these efforts could be significantly enhanced by elucidating the catalytic mechanism of PSII.

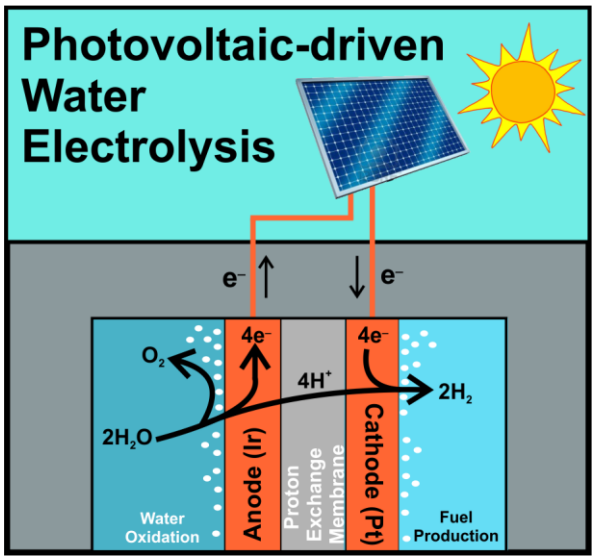


Figure 6-3. Artificial system to create fuel from sunlight and water. A photovoltaic device is used to drive water electrolysis to generate H_2 as a fuel from water.[185]

Potential design of an artificial oxygen-evolving complex

Solar-driven water oxidation by PSII is accomplished through three distinct functions:

1. Light absorption and energy transfer from the large antenna complex to a reaction center (P680).
2. Conversion of light energy into electrical potential through charge separation in the reaction center.
3. Storage of four oxidizing equivalents in a Mn_4O_5Ca cluster, which ultimately catalyzes the oxidation water: $2H_2O + \text{light} \rightarrow O_2 + 4H^+ + 4e^-$.

PSII is highly conserved among diverse species of photosynthetic organisms and this suggests that further optimization, either through natural evolution or mutagenesis, to improve the efficiency of this complex protein is unlikely.[161] From an engineering perspective, photosynthesis might be improved by decoupling each of its three functions, such that they could be controlled independently, and then recombined in an optimal manner so as to achieve a superior outcome. For example, photodamage of PSII occurs because of a lack of electron acceptors, which leads to charge recombination, the formation of ^{3}P680, $^1O_2$ and damage to the PSII protein around the reaction center, as previously described. Photodamage could be prevented if the pool of electron acceptors could be significantly expanded by directly coupling an artificial light-driven water splitting complex to a transparent electrode. While light absorption and most of the electron transfer reaction take place in the membrane intrinsic part of PSII, the catalytic site for water oxidation is located in the membrane extrinsic part of PSII suggesting that a water soluble version of the oxygen-evolving complex (OEC) could be produced. To this end, we propose a design strategy to decouple the water oxidation function from the light reactions in PSII by assembling three peptide domains from the OEC, which provide the ligand framework for its Mn_4O_5Ca active site, within an artificial scaffold. This design approach could enable further investigation into the PSII catalytic mechanism, as well as provide insights into how to optimize the catalyst for engineering into more practical and economical solar to fuel devices.

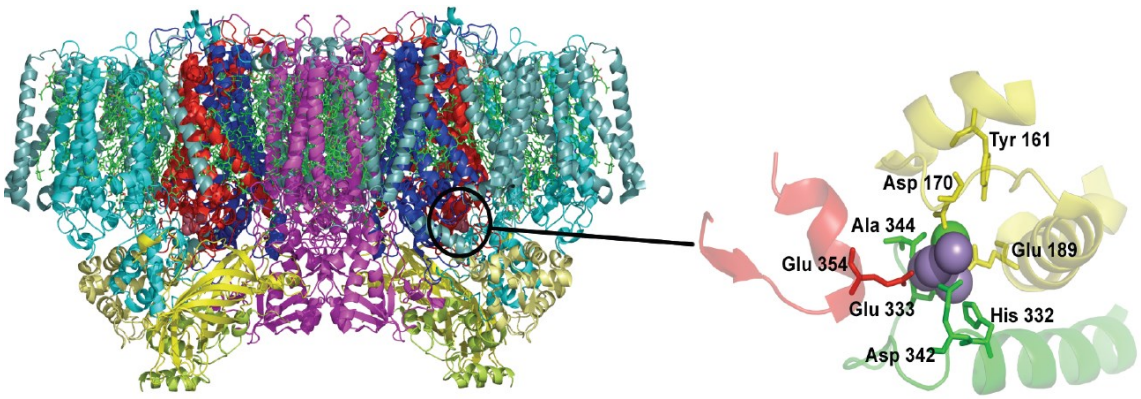


Figure 6-4. Peptides identified from PSII for assembling an artificial OEC. Crystal structure of photosystem II (PDB accession code 2AXT)[108] and close up view showing the peptides coordinating the Mn_4O_5Ca cluster for the artificial OEC. These peptides include the CD loop of PsbA (Ala156–His195) depicted in yellow, the C-terminus of PsbA (Ala318–Ala344) in green, and the short loop of PsbC (Gly347–Asp360) in red. Figure made from 2AXT by Ingo Grotjohann.

From the PSII crystal structure,[108] the essential ligands to coordinate the Mn_4O_5Ca cluster in the active site are found in two adjacent proteins of PSII (PsbA and PsbC). Although the ligands are not in close sequence proximity, they could be provided using three peptide segments (referred to as the OEC peptides), as shown in Figure 6-4. The sequences of each peptide are described in Table 5-1, which also denotes the amino acids directly coordinating the Mn_4O_5Ca cluster. A more recent, higher resolution crystal structure[3] confirmed the importance of the ligands provided by these three peptides and suggests Asp61 from PsbA may provide an additional ligand via a water molecule. Because it is not clear which residues in the secondary and tertiary coordination sphere are critical to perform the catalytic cycle, a logical first step is to assemble the basic peptide framework of the OEC before considering more complex and potentially superfluous peptide domains. The modular DNA-PNA scaffold design described in Chapters 3–5, provides flexibility for investigating various combinations of different peptide sequences, as well

as introducing additional peptide domains from PSII to further investigate their roles in the catalytic cycle.

By encapsulating the OEC peptides inside a 3D DNA nanostructure, it may be possible to perform high resolution structural analysis of the OEC during the catalytic cycle, to provide new insights into the mechanism. Previous x-ray analysis of 3D DNA nanocage crystals did not provide sufficient resolution to solve the structure.[186] However, incorporating the artificial OEC into the same DNA nanocage in the future could aid structure determination by stabilizing the DNA nanocage and providing asymmetry during crystal formation, as well as reducing the solvent content and increasing the electron density of the crystals during x-ray analysis. 3D DNA nanostructures are also conducive to electron transport[187] and have been incorporated within conductive metal oxide films.[137] Although DNA is a stable molecule,[188,189] the purine nucleobases can become oxidized at potentials as low as 1.0–1.3 V vs NHE at pH 7,[190,191] which is close to the thermodynamic redox potential for water oxidation of 0.82 V vs NHE at pH 7. Fortunately, PSII operates efficiently with a low overpotential[192] suggesting that an artificial OEC, which would be able to closely mimic the catalytic center of PSII, could be driven with potentials low enough to prevent damage to the DNA nanocage.

To further constrain the OEC peptides in an orientation that mimics the scaffold provided by PSII, the termini of each OEC peptide can be anchored to the DNA nanocage, with exception of the C terminal end of the green peptide, which provides a direct ligand to the Mn_4O_5Ca cluster, as shown in Figure 6-5. The OEC peptides could be oriented in the DNA nanocage in such a way that would allow for a single PNA strand to be used to connect both ends of either the red or yellow peptide at the N terminus and at

the gamma position of the 4th PNA base, to form cyclic PNA-peptides. The DNA nanocage would, in an ideal arrangement, place each peptide within 0.6 nm of its relative position in the PSII crystal structure. The strands used to form this model of the complex (DNA-PNA-aOEC) are shown in Table 5-1.

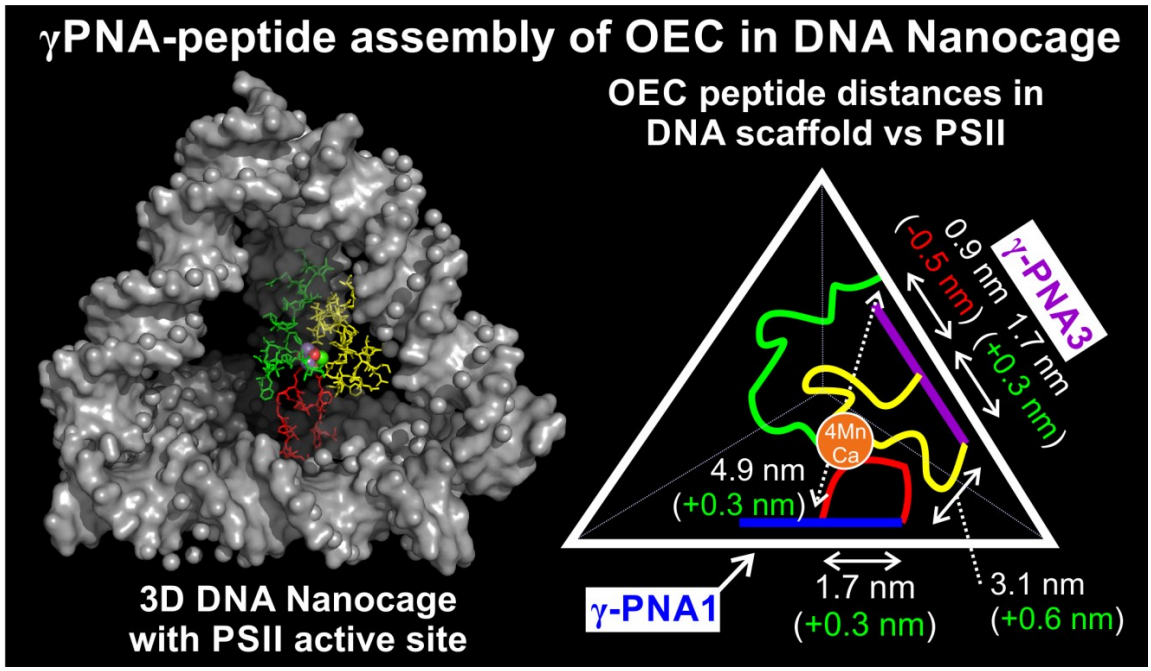


Figure 6-5. Model and schematic of assembling the PSII active site within a DNA nanocage using γ-PNA linkers. [Left] A molecular model of the photosystem II (PSII) oxygen-evolving complex (OEC) inside the DNA nanocage. [Right] Schematic of the DNA nanocage arranging three peptides from the OEC using γ-PNA linkers. The red and yellow peptides are cyclized with γ-PNA1 and γ-PNA3, respectively. The green peptide is attached directly to the DNA nanocage. The distances between the peptide points of attachment in the DNA nanocage model are labeled. The small differences in distances compared to the PSII crystal structure are shown in parenthesis, where positive values indicate the peptides are further apart in the DNA nanocage than in the protein.

Furthermore, by forming cyclic PNA-peptides, the same two orthogonal linking chemistries could be used to form any number of cyclic PNA-peptides, which can each be assembled at specific locations within the DNA nanocage based on the unique PNA sequence. While it would be ideal to also attach the green peptide to a third PNA sequence in order to maintain the most flexibility when assembling the OEC peptides and

constructing the artificial OEC, the described DNA nanocage design is best suited for no more than one 8nt PNA binding site per edge of the DNA nanocage. Because other near-by edges in the nanocage would anchor the peptide too far (>2 nm) from the red and yel-low peptides, the design presented here directly attaches the green peptide to the DNA nanocage by chemical crosslinking.

Potential strategy to form cyclic γ-PNA-peptide conjugates

A potential bioorthogonal conjugation strategy is now discussed, using thiol-maleimide and alkyne-azide chemistry, as a conceivable strategy to introduce the OEC peptides into the DNA nanocage, as shown in Figure 6-6. The γ-PNA1 and γ-PNA3 se-quences were designed to form cyclic PNA-peptides with the red and yellow peptides from the PSII active site, respectively. The following γ-PNA sequences were purchased (PNA Bio, Thousand Oaks, CA, USA) with the desired functionalization (written from N to C terminus): γ-PNA1 {N_3-AGCG(Lys-C3SH)ATGG}, γ-PNA3 {N_3-AACG(Lys-C3SH)GCGA}, where N_3 is an N terminal azide functionalization, and Lys-C3SH is a thiol with a 3 carbon linker conjugated to a lysine residue at the gamma position of the 4^{th} PNA base from the N terminus. The chemical structure of PNA showing the gamma sub-stitution position on the backbone is presented in Figure 1-6.

The peptide sequences are provided in Table 5-1. Because no cysteine or lysine residues are present in any of the peptide sequences identified from PSII, they can be added to the termini of each sequence as low cost, orthogonal handles, which can be functionalized in solution after peptide synthesis. The 15 amino acid (AA) red peptide

was designed with N terminal maleimide functionalization for conjugating to the thiol at the gamma position of γ-PNA1. A C terminal lysine residue was added to the red peptide

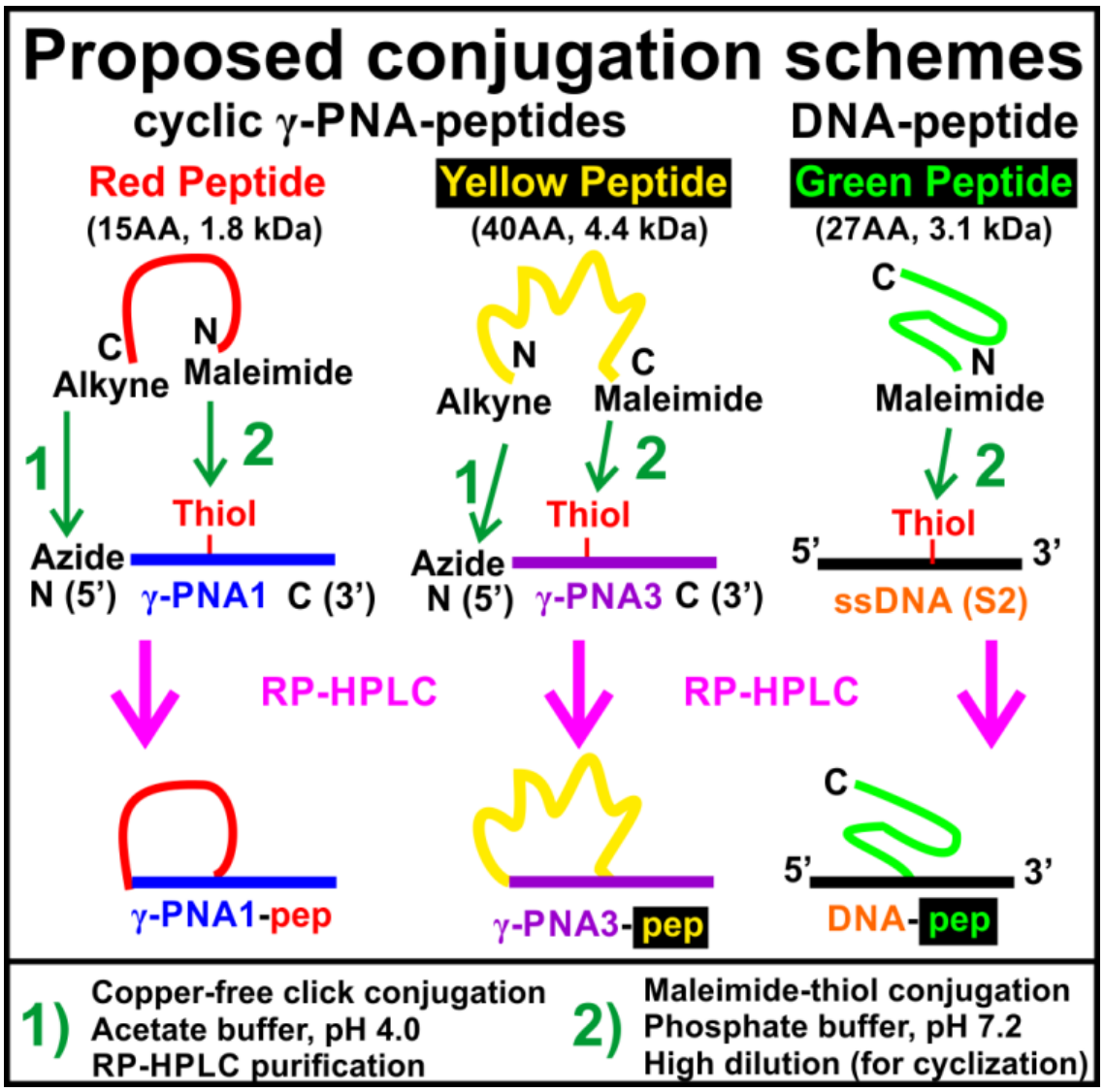


Figure 6-6. Proposed conjugation scheme to form cyclic γ-PNA-peptides and DNA-peptides for assembling the artificial OEC. Description provided in the text.

sequence as a handle to introduce alkyne functionalization for conjugating to the N terminal azide of γ-PNA1. The 40AA yellow peptide was designed with a free amino terminus for introducing alkyne functionalization for conjugating to the N terminal azide of γ-PNA3. A C terminal cysteine residue was added to the yellow peptide sequence as a han-

dle to introduce maleimide functionalization for conjugating to the thiol at the gamma position of γ-PNA3. The maleimide may be introduced with bis(maleimido)ethane, by using a large excess of to prevent cross-linking . The 27AA green peptide was designed with N terminal maleimide functionalization for conjugating to thiolated ssDNA. In order to complete the peptide functionalization, the strained alkyne dibenzocyclooctyl (DBCO) functionalized with N-hydroxysuccinimide (ClickChemistryTools, Scottsdale, AZ, USA) may be used to introduce the necessary alkyne to the C terminal lysine on the red peptide and the N terminus of the yellow peptide. Strained alkynes, such as DBCO, can be efficiently conjugated to azides without adding any copper catalyst,[154,155] which can be difficult to completely remove[155,156] and could compete with manganese binding during assembly of the artificial OEC.

The proposed conjugation scheme to form the cyclic γ-PNA-peptides is shown in Figure 6-6. In step 1, using the same approach for labeling the PNAs with Cy5, copper-free click chemistry is used to conjugate the C and N terminus of the red and yellow peptides functionalized with DBCO to the N terminal azide on γ-PNA1 or γ-PNA3, respectively. Because strained alkynes can sometimes have mild cross-reactivity with other nucleophiles,[193] the click reaction is carried out under mildly acidic conditions of pH 4 to ensure the thiol remains protonated and cannot react with DBCO. In step 2, the PNA-peptides are resuspended in phosphate buffer at pH 7.2 in the presence of TCEP to perform the cyclization reaction between the N and C terminal maleimide on the red and yellow peptide to γ-PNA1 and γ-PNA3, respectively. The reaction is performed at high dilution (nanomolar) to favor the formation of intramolecular cyclization over the competing intermolecular reaction leading to polymerization.[194] The N terminal maleimide of the

green peptide is directly conjugated to the thiol of ssDNA in the same fashion, but at high concentration (micro to millimolar), since there is no possibility of forming polymers. It is worth noting that by performing the click reaction first, TCEP can be used to reduce any oxidized thiols prior to maleimide conjugation, which would otherwise rapidly reduce the γ-PNA N terminal azides to primary amines.[195]

Synthesis of the OEC peptides used in the conjugation scheme proposed above might be challenging. I requested several companies (Pierce and Genscript) to synthesize the OEC peptides, but they were unable to successfully synthesize and purify all three peptides due to their hydrophobicity. Genscript was able to synthesize the more water soluble green peptide (27AA) with >95% purity and the shorter red peptide (15AA) with ~80% purity, but was unable to synthesize and purify the longer yellow peptide (40AA). RP-HPLC purification of the more hydrophobic peptides might be improved in the future using small amounts of methanol or ethanol mixed with water in the stationary phase to prevent peptide aggregation on the column, followed by eluting the peptide with isopropanol instead of acetonitrile.[196] In order to improve synthesis yields, the yellow peptide solubility could be improved by shortening the sequence or inserting charged residues.[197] Alternatively, the yellow peptide could be expressed in a biological system with a soluble fusion protein.[198]

Potential strategy to assemble and operate the artificial OEC

The DNA nanocage design used throughout this dissertation provides flexibility to control the timing and sequence of OEC peptide assembly using PNA linkers, allowing many assembly conditions to be investigated. The green peptide could be conjugated to

one of the ssDNA sequences (S3Mc-SH) and then assembled along with the DNA nanocage or conjugated to the preassembled DNA nanocage. The red and yellow peptides are each conjugated to separate PNAs and can be assembled concurrently or in a particular sequence by incubation each with the preformed DNA nanocage, with or without the green peptide. A stepwise process to assemble the OEC may also help to reduce the complexity of its formation, by first assembling mono or dimanganese complexes, such as those found in some bacterial reaction centers[199] and superoxide dismutase (mono-Mn) or in manganese catalase (di-Mn). Manganese catalase is hypothesized to be a precursor in the evolution of the OEC, which may provide a clue as to how to assemble more complex manganese containing catalytic centers.[200] Manganese superoxide dismutase also has a mild redox potential of 0.3–0.4 V vs NHE, which is easily tolerated by the DNA nanocage.[201] Formation of the Mn_4O_5Ca cluster can be performed after, or concurrent with, the OEC peptide assembly, and may be accomplished through the addition of calcium chloride and manganese acetate followed by, or concurrent with, the (electro)chemical oxidation of the manganese to the catalytically active manganese oxidation states (i.e., III, IV or V).[202]

The thermodynamic potential of water oxidation (0.82 V vs NHE) provides a 0.2–0.5 V window to drive the reaction without potential damage to the DNA nanocage through oxidation of the guanine nucleobases. Although this is a relatively small overhead, native PSII operates with low overpotential. *De novo* designed water oxidation catalysts also suffer from many limitations, such as low turnover rates, low stability and high cost.[203] The approach described here has the advantage that the OEC from PSII is known to function as a highly active and robust catalyst. If the activity of the assembled

artificial OEC is found to be low, the flexibility of the described methodology to assemble the DNA-PNA-polypeptide complexes may allow the assembled catalyst to be optimized further, perhaps through molecular evolution. By integrating the artificial OEC into nanopourous electrodes,[137] the artificial OEC could be driven directly, without a reaction center, and thus would not suffer the same damage that occurs in the native PSII reaction center, thus extending the longevity of the artificial OEC and improving its economic viability.

Conclusion

PSII is a remarkable water oxidation catalyst that currently exceeds the performance and stability of most synthetic catalysts, even those using cost-prohibitive rare earth metals (e.g., Pt, Ir, Ru). However, when integrated in the photosynthetic electron transport chain, PSII can quickly saturate the pool of electron acceptors, such that toxic radicals are generated during charge recombination, which cause irreversible damage to the catalyst. In this chapter we presented and discussed a novel approach to reassemble the oxygen-evolving complex, which encompasses the PSII catalytic site of water oxidation, inside a stable DNA nanocage. By integrating the OEC into synthetic systems it may prevent charge recombination and extend its longevity. In this way, it may be possible to leverage the performance of this highly active catalyst without being constrained by the constant damage that occurs within photosynthetic organisms.

Building an artificial OEC from PSII presents many challenges, such as synthesizing and purifying several hydrophobic peptides, assembling an active site containing five metal sites of two different metals (Mn_4O_5Ca) and must be driven using fairly high redox

potentials (>0.82 V vs NHE). The prospective design strategy proposed here provides flexibility to control the timing and sequence of OEC peptide assembly, allowing many assembly conditions to be investigated, and may provide further insight into the catalytic mechanism of one of nature's most remarkable proteins. Furthermore, by mimicking one of the best known water oxidation catalyst, which is made of earth abundant materials, the artificial OEC could provide a recipe to design economically viable solar to fuel catalysts.

CHAPTER 7 – Summary and Outlook

Summary

This dissertation describes a gentle and flexible methodology for using short (8nt) PNA linkers to rapidly assemble polypeptides within a 3D DNA nanocage, as a proof of concept for constructing artificial catalytic centers. This is the first time that PNA has been used to assemble polypeptides into 3D DNA nanostructures. Toehold-mediated DNA strand displacement is introduced as an effective method for purifying PNA and PNA-protein conjugates. Denaturing PAGE was introduced as an efficient method to purify PNA fluorescently labeled with Cy5. The PNA-polypeptides could be introduced quantitatively by incubating at room temperature with the preassembled DNA nanocage for 2 minutes or 5–10 minutes for the PNA-proteins or PNA-peptides, respectively. The PNA-proteins could also be assembled within 4 minutes at 11 ºC. The presence of protein increased the PNA solubility, allowing the PNA-protein conjugate to bind to the DNA nanocage in near stoichiometric quantities, with faster binding kinetics and with increased hybridization efficiency than the same PNA sequence conjugated to a short, neutral peptide.[79] A TMR dye label was found to significantly stabilize PNA-peptide binding to the DNA nanocage, increasing the dissociation temperature by 9.2 °C from a predicted value of 36.6 °C to 45.8 °C.

The effect of the surface charge of two different proteins on their interactions with a DNA nanocage was also investigated. Data from biophysical characterization using gel electrophoresis as well as steady state and time resolved fluorescence spectroscopy allowed us to develop a model for the arrangement of the PNA-polypeptides inside the DNA nanocage that is in agreement with the intended design. Our model suggests that

the negatively charged azurin is repelled away from the DNA nanocage, whereas the positively charged cytochrome c remains within and closely interacts with the DNA nanocage. The negatively charged DNA nanocage environment also reduces the redox potential of cytochrome c by 10 mV, but still remains catalytically active when incorporated into the DNA nanocage. A methodology for using cyclic PNA-peptides to construct an artificial oxygen-evolving complex is also described. This approach could also be extended to assemble artificial catalytic centers from one of the many other diverse enzymes found in nature, in order to investigate their function, and to improve their performance and economic viability.

Outlook

Proteins use a number of metals and cofactors to carry out a wide range of interesting chemistry, many of which have far fewer constraints than the OEC described in Chapter 6, and may be useful targets for assembling peptide based catalytic centers within DNA nanocages. Heme containing proteins, such as cytochrome c and horseradish peroxidase, have much simpler active sites, made up of iron coordinated within a porphyrin macrocycle, which might be more easily incorporated into a DNA nanocage. Iron is also commonly complexed by inorganic and cysteine sulfurs to form iron-sulfur clusters, such as those found in ferredoxins and hydrogenases, and, due to their spontaneous assembly *in vitro*, might be more easily incorporated into peptide assemblies stabilized by a DNA nanocage.[204] Giovanna Ghirlanda's group has *de novo* designed and synthesized several remarkable water soluble, peptide-based analogs of heme[205] and iron-sulfur cluster containing proteins,[206,207] which could be used as building blocks for assembling higher order

redox active complexes within DNA nanostructures to promote catalytic cascades, to engineer more elaborate artificial electron transfer pathways,[206] or to enhance solar-driven fuel production.[207] Together, the redox potentials of both heme and iron-sulfur cluster proteins span a remarkable range of over 1000 mV, which could be tuned for operating within the DNA nanocage.[208] Cupredoxins are a class of copper containing redox proteins that are water soluble and have relatively simple active sites and modest redox potentials, which are also good candidates to be assembled within DNA-PNA-peptide complexes.[209]

The 3D DNA nanocage design could also be modified to accommodate the remarkable diversity of enzymes and active sites found in nature. The DNA nanocage could be further expanded to connect up to 4-6 polypeptides (up to 60 kDa)[50] on adjacent edges of the DNA nanocage, by introducing the appropriate PNA binding sequences. The ability to systematically introduce different proteins or peptides into defined locations within such a three dimensional structure may also facilitate investigating protein-protein interactions. The maturation of the DNA nanotechnology field has greatly simplified the process of designing a wide variety of DNA nanostructures. Incorporating proteins and polypeptides into 3D DNA nanostructures by means of a PNA linker facilitates exploiting the enormous diversity of DNA nanostructures for investigating and engineering functional 3D protein-nucleic acid complexes. The flexibility of the methods described here offer a biomimetic route to rebuild protein active sites and further expand the structure and functional landscape of polypeptide engineering. This has the potential to revolutionize the process of designing and building robust catalysts by leveraging the recipes contained within nature's remarkable molecular machines, and also providing a flexible and

controlled artificial environment that might even improve them further towards commercial viability.

REFERENCES

(1) Protein structure levels http://en.wikipedia.org/wiki/File:Main_protein_structure_levels_en.svg (accessed Apr 2, 2014).

(2) Kantharaj, G. R. Plant cell structure http://preuniversity.grkraj.org/html/1_CELL_STRUCTURE.htm (accessed Apr 2, 2014).

(3) Umena, Y.; Kawakami, K.; Shen, J.-R.; Kamiya, N. *Nature* **2011**, *473*, 55–60.

(4) Watson, J.; Crick, F. *Nature* **1953**, *171*, 737–738.

(5) Holliday, R. *Genet. Res.* **1964**, *5*, 282–304.

(6) Holliday Junction http://en.wikipedia.org/wiki/File:Holliday_junction_coloured.png (accessed Mar 30, 2014).

(7) Lin, C.; Liu, Y.; Rinker, S.; Yan, H. *ChemPhysChem* **2006**, *7*, 1641–1647.

(8) Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H. *Science* **2003**, *301*, 1882–1884.

(9) Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, *394*, 539–544.

(10) He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. *Nature* **2008**, *452*, 198–201.

(11) Seeman, N. C. *J. Theor. Biol.* **1982**, *99*, 237–247.

(12) Kallenbach, N.; Ma, R.; Seeman, N. *Nature* **1983**, *305*, 829–831.

(13) Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F. *Science* **2008**, *321*, 1795–1799.

(14) Seeman, N. C. *Annu. Rev. Biochem.* **2010**, *79*, 65–87.

(15) Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. *Nat. Nanotechnol.* **2011**, *6*, 763–772.

(16) Andersen, E.; Nielsen, M. *Nat. Protoc.* **2011**, *1*.

(17) Sawyer, N.; Speltz, E. B.; Regan, L. *Biochem. Soc. Trans.* **2013**, *41*, 1131–1136.

(18) Hollfelder, F.; Lutz, S. *Curr. Opin. Struct. Biol.* **2013**, *23*, 569–570.

(19) LaBean, T.; Yan, H.; Kopatsch, J. *J. Am. Chem. Soc.* **2000**, *122*, 1848–1860.

(20) Zheng, J.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R.; Constantinou, P. E.; Ginell, S. L.; Mao, C.; Seeman, N. C. *Nature* **2009**, *461*, 74–77.

(21) Rothemund, P. W. K. *Nature* **2006**, *440*, 297–302.

(22) Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L. P.; Pedersen, J. S.; Birkedal, V.; Besenbacher, F.; Gothelf, K. V; Kjems, J. *Nature* **2009**, *459*, 73–76.

(23) Han, D.; Pal, S.; Nangreave, J.; Deng, Z.; Liu, Y.; Yan, H. *Science* **2011**, *332*, 342–346.

(24) Han, D.; Pal, S.; Yang, Y.; Jiang, S.; Nangreave, J.; Liu, Y.; Yan, H. *Science* **2013**, *339*, 1412–1415.

(25) Douglas, S. M.; Dietz, H.; Liedl, T.; Högberg, B.; Graf, F.; Shih, W. M. *Nature* **2009**, *459*, 414–418.

(26) Wang, W.; Wan, W.; Zhou, H.-H.; Niu, S.; Li, A. D. Q. *J. Am. Chem. Soc.* **2003**, *125*, 5248–5249.

(27) Ding, K.; Alemdaroglu, F. E.; Börsch, M.; Berger, R.; Herrmann, A. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 1172–1175.

(28) Sharma, J.; Chhabra, R.; Cheng, A.; Brownell, J. *Science* **2009**, *323*, 112–116.

(29) Pal, S.; Deng, Z.; Ding, B.; Yan, H.; Liu, Y. *Angew. Chemie Int. Ed.* **2010**, *122*, 2760–2764.

(30) Pal, S.; Varghese, R.; Deng, Z. *Angew. ...* **2011**, *123*, 4262–4265.

(31) Tanaka, K.; Clever, G. H.; Takezawa, Y.; Yamada, Y.; Kaul, C.; Shionoya, M.; Carell, T. *Nat. Nanotechnol.* **2006**, *1*, 190–194.

(32) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shionoya, M. *Science* **2003**, *299*, 1212.

(33) Ke, Y.; Lindsay, S.; Chang, Y.; Liu, Y.; Yan, H. *Science* **2008**, *319*, 180–183.

(34) Yurke, B.; Turberfield, a J.; Mills, a P.; Simmel, F. C.; Neumann, J. L. *Nature* **2000**, *406*, 605–608.

(35) Feng, L.; Park, S. H.; Reif, J. H.; Yan, H. *Angew. Chemie* **2003**, *115*, 4478–4482.

(36) Yin, P.; Yan, H.; Daniell, X. G.; Turberfield, A. J.; Reif, J. H. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 4906–4911.

(37) Lin, C.; Wang, X.; Liu, Y. *J. Am. ...* **2007**, *129*, 14475–14481.

(38) Lin, C.; Rinker, S.; Wang, X.; Liu, Y.; Seeman, N. C.; Yan, H. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17626–17631.

(39) Williams, B. a R.; Lund, K.; Liu, Y.; Yan, H.; Chaput, J. C. *Angew. Chemie* **2007**, *119*, 3111–3114.

(40) Eberhard, H.; Diezmann, F.; Seitz, O. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 4146–4150.

(41) Fu, J.; Liu, M.; Liu, Y. *J. Am. Chem. Soc.* **2012**, *22*, 5516–5519.

(42) Numajiri, K.; Yamazaki, T.; Kimura, M.; Kuzuya, A.; Komiyama, M. *J. Am. Chem. Soc.* **2010**, *132*, 9937–9939.

(43) Chhabra, R.; Sharma, J.; Ke, Y.; Liu, Y.; Rinker, S.; Lindsay, S.; Yan, H. *J. Am. Chem. Soc.* **2007**, *129*, 10304–10305.

(44) Zhou, C.; Yang, Z.; Liu, D. *J. Am. Chem. Soc.* **2012**, *134*, 1416–1418.

(45) Nakata, E.; Liew, F. F.; Uwatoko, C.; Kiyonaka, S.; Mori, Y.; Katsuda, Y.; Endo, M.; Sugiyama, H.; Morii, T. *Angew. Chemie* **2012**, *124*, 2471–2474.

(46) Erkelenz, M.; Kuo, C.-H.; Niemeyer, C. M. *J. Am. Chem. Soc.* **2011**, *133*, 16111–16118.

(47) Englund, E. A.; Wang, D.; Fujigaki, H.; Sakai, H.; Micklitsch, C. M.; Ghirlando, R.; Martin-Manso, G.; Pendrak, M. L.; Roberts, D. D.; Durell, S. R.; Appella, D. H. *Nat. Commun.* **2012**, *3*, 614.

(48) Wilner, O. I.; Weizmann, Y.; Gill, R.; Lioubashevski, O.; Freeman, R.; Willner, I. *Nat. Nanotechnol.* **2009**, *4*, 249–254.

(49) Fu, Y.; Zeng, D.; Chao, J.; Jin, Y.; Zhang, Z.; Liu, H.; Li, D.; Ma, H.; Huang, Q.; Gothelf, K. V.; Fan, C. *J. Am. Chem. Soc.* **2013**, *135*, 696–702.

(50) Erben, C. M.; Goodman, R. P.; Turberfield, A. J. *Angew. Chemie Int. Ed.* **2006**, *118*, 7574–7577.

(51) Zhang, C.; Tian, C.; Guo, F.; Liu, Z.; Jiang, W.; Mao, C. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 3382–3385.

(52) Duckworth, B. P.; Chen, Y.; Wollack, J. W.; Sham, Y.; Mueller, J. D.; Taton, T. A.; Distefano, M. D. *Angew. Chemie* **2007**, *119*, 8975–8978.

(53) Wong, N. Y.; Zhang, C.; Tan, L. H.; Lu, Y. *Small* **2011**, *7*, 1427–1430.

(54) Douglas, S.; Chou, J.; Shih, W. *Proc. Natl. Acad. Sci.* **2007**, *104*, 6644–6648.

(55) Franceschini, A.; Szklarczyk, D.; Frankild, S.; Kuhn, M.; Simonovic, M.; Roth, A.; Lin, J.; Minguez, P.; Bork, P.; von Mering, C.; Jensen, L. J. *Nucleic Acids Res.* **2013**, *41*, D808–15.

(56) Chatr-Aryamontri, A.; Breitkreutz, B.-J.; Heinicke, S.; Boucher, L.; Winter, A.; Stark, C.; Nixon, J.; Ramage, L.; Kolas, N.; O'Donnell, L.; Reguly, T.; Breitkreutz, A.; Sellam, A.; Chen, D.; Chang, C.; Rust, J.; Livstone, M.; Oughtred, R.; Dolinski, K.; Tyers, M. *Nucleic Acids Res.* **2013**, *41*, D816–23.

(57) Wodak, S. J.; Vlasblom, J.; Turinsky, A. L.; Pu, S. *Curr. Opin. Struct. Biol.* **2013**, *23*, 941–953.

(58) Mosca, R.; Pons, T.; Céol, A.; Valencia, A.; Aloy, P. *Curr. Opin. Struct. Biol.* **2013**, *23*, 929–940.

(59) Myer, Y.; Saturno, A.; Verma, B.; Pande, A. *J. Biol. Chem.* **1979**, *254*, 11202–11207.

(60) Greenfield, N. *Nat. Protoc.* **2007**, *1*, 2876–2890.

(61) Wolfe, S.; Nekludova, L.; Pabo, C. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 183–212.

(62) Bai, X.-C.; Martin, T. G.; Scheres, S. H. W.; Dietz, H. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20012–20017.

(63) Severcan, I.; Geary, C.; Chworos, A.; Voss, N.; Jacovetty, E.; Jaeger, L. *Nat. Chem.* **2010**, *2*, 772–779.

(64) Kato, T.; Goodman, R. P.; Erben, C. M.; Turberfield, A. J.; Namba, K. *Nano Lett.* **2009**, *9*, 2747–2750.

(65) Goodman, R. P.; Schaap, I. A. T.; Tardin, C. F.; Erben, C. M.; Berry, R. M.; Schmidt, C. F.; Turberfield, A. J. *Science* **2005**, *310*, 1661–1665.

(66) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–1500.

(67) Braasch, D. A.; Nulf, C. J.; Corey, D. R. In *Current protocols in nucleic acid chemistry*; John Wiley & Sons: New York, NY, 2002; pp. 4.11.1–4.11.18.

(68) Kaihatsu, K.; Corey, D. R. *Methods Mol. Biol.* **2004**, *283*, 207–216.

(69) CEM Corporation. CEM Liberty http://www.cem.com/content-cat686.html (accessed Mar 31, 2014).

(70) Lansdorp, P. M.; Verwoerd, N. P.; van de Rijke, F. M.; Dragowska, V.; Little, M. T.; Dirks, R. W.; Raap, a K.; Tanke, H. J. *Hum. Mol. Genet.* **1996**, *5*, 685–691.

(71) Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Manicardi, A.; Marchelli, R. *Curr. Top. Med. Chem.* **2011**, *11*, 1535–1554.

(72) Rozners, E. *J. Nucleic Acids* **2012**, *2012*, 518162.

(73) Demidov, V. V; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchard, O.; Sönnichsen, S. H.; Nielsen, P. E. *Biochem. Pharmacol.* **1994**, *48*, 1310–1313.

(74) Braasch, D.; Corey, D. *Biochemistry* **2002**, *41*.

(75) Kuhn, H.; Demidov, V. V; Coull, J. M.; Fiandaca, M. J.; Gildea, B. D.; Frank-Kamenetskii, M. D. *J. Am. Chem. Soc.* **2002**, *124*, 1097–1103.

(76) Corradini, R.; Feriotto, G.; Sforza, S.; Marchelli, R.; Gambari, R. *J. Mol. Recognit.* **2004**, *17*, 76–84.

(77) Wang, J.; Palecek, E.; Nielsen, P. *J. Am. Chem. Soc.* **1996**, *7863*, 7667–7670.

(78) Lukeman, P. S.; Mittal, A. C.; Seeman, N. C. *Chem. Commun.* **2004**, 1694–1695.

(79) Flory, J. D.; Shinde, S.; Lin, S.; Liu, Y.; Yan, H.; Ghirlanda, G.; Fromme, P. *J. Am. Chem. Soc.* **2013**, *135*, 6985–6993.

(80) Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. *J. Org. Chem.* **2011**, *76*, 5614–5627.

(81) Englund, E. a; Appella, D. H. *Org. Lett.* **2005**, *7*, 3465–3467.

(82) Englund, E. A.; Appella, D. H. *Angew. Chemie Int. Ed.* **2007**, *46*, 1414–1418.

(83) Fabani, M. M.; Abreu-Goodger, C.; Williams, D.; Lyons, P. a; Torres, A. G.; Smith, K. G. C.; Enright, A. J.; Gait, M. J.; Vigorito, E. *Nucleic Acids Res.* **2010**, *38*, 4466–4475.

(84) Pritz, S.; Yvonne, W.; Klemm, C.; Bienert, M. *Tetrahedron Lett.* **2006**, *47*, 5893–5896.

(85) Kowalczyk, R.; Harris, P. W. R.; Dunbar, R. P.; Brimble, M. A. *Synthesis (Stuttg).* **2009**, *13*, 2210–2222.

(86) Fischer, R.; Mader, O.; Jung, G.; Brock, R. *Bioconjug. Chem.* **2003**, *14*, 653–660.

(87) Butler, J. M.; Jiang-Baucom, P.; Huang, M.; Belgrader, P.; Girard, J. *Anal. Chem.* **1996**, *68*, 3283–3287.

(88) Tataurov, A. V; You, Y.; Owczarzy, R. *Biophys. Chem.* **2008**, *133*, 66–70.

(89) Sjöback, R.; Nygren, J.; Kubista, M. *Spectrochim. Acta Part A* **1995**, *51*, L7–L21.

(90) Louie, G.; Hutcheon, W.; Brayer, G. *J. Mol. Biol.* **1988**, *199*, 295–314.

(91) Van Amsterdam, I. M.; Ubbink, M.; Jeuken, L. J.; Verbeet, M. P.; Einsle, O.; Messerschmidt, A.; Canters, G. W. *Chem. – A Eur. J.* **2001**, *7*, 2398–2406.

(92) Van de Kamp, M.; Hali, F. C.; Rosato, N.; Agro, a F.; Canters, G. W. *Biochim. Biophys. Acta* **1990**, *1019*, 283–292.

(93) Margoliash, E.; Frohwirt, N.; Wiener, E. *Biochem. J.* **1959**, *71*, 559–570.

(94) Kolczak, U., C.; Dennison, A.; Messerschmidt; Canters, G. W. *Handbook of metalloproteins*; Messerschmidt, A.; Huber, R.; T. Poulos; Wieghardt, K., Eds.; John Wiley & Sons: Chichester, 2001; pp. 1170–1194.

(95) Giesen, U.; Kleider, W.; Berding, C.; Geiger, A.; Ørum, H.; Nielsen, P. E. *Nucleic Acids Res.* **1998**, *26*, 5004–5006.

(96) Ackermann, D.; Famulok, M. *Nucleic Acids Res.* **2013**, *41*, 4729–4739.

(97) Zhang, D. Y.; Winfree, E. *J. Am. Chem. Soc.* **2009**, *131*, 17303–17314.

(98) Bellot, G.; McClintock, M. a; Lin, C.; Shih, W. M. *Nat. Methods* **2011**, *8*, 192–194.

(99) Liu, M.; Fu, J.; Hejesen, C.; Yang, Y.; Woodbury, N. W.; Gothelf, K.; Liu, Y.; Yan, H. *Nat. Commun.* **2013**, *4*, 2127.

(100) Owczarzy, R.; You, Y.; Groth, C.; Tataurov, A. *Biochemistry* **2011**, *50*, 9353–9367.

(101) Magde, D.; Wong, R.; Seybold, P. G. *Photochem. Photobiol.* **2002**, *75*, 327–334.

(102) Meadows, D. L.; Shafer, J. S.; Schultz, J. S. *J. Immunol. Methods* **1991**, *143*, 263–272.

(103) Hamman, B. D.; Oleinikov, A. V; Jokhadze, G. G.; Bochkariov, D. E.; Traut, R. R.; Jameson, D. M. *J. Biol. Chem.* **1996**, *271*, 7568–7573.

(104) Wang, L.; Gaigalas, A. K.; Blasic, J.; Holden, M. J. *Spectrochim. Acta Part A* **2004**, *60*, 2741–2750.

(105) Yeh, J. I.; Pohl, E.; Truan, D.; He, W.; Sheldrick, G. M.; Du, S.; Achim, C. *Chemistry* **2010**, *16*, 11867–11875.

(106) Bushnell, G. W.; Louie, G. V; Brayer, G. D. *J. Mol. Biol.* **1990**, *214*, 585–595.

(107) Van Amsterdam, I. M. C.; Ubbink, M.; Einsle, O.; Messerschmidt, A.; Merli, A.; Cavazzini, D.; Rossi, G. L.; Canters, G. W. *Nat. Struct. Biol.* **2001**, *9*, 48–52.

(108) Loll, B.; Kern, J.; Saenger, W.; Zouni, A.; Biesiadka, J. *Nature* **2005**, *438*, 1040–1044.

(109) DeLano, W. L. (2002) The PYMOL molecular graphics system, DeLano Scientific, San Carlos, CA.

(110) Emsley, P.; Cowtan, K. *Acta Crystallogr. D. Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(111) Marvin Sketch. “Marvin Sketch was used for drawing, displaying, and characterizing chemical structures, MarvinSketch 5.11.3, 2012, ChemAxon (http://www.chemaxon.com).”

(112) Craig, D. B.; Nichols, E. R. *J. Chem. Educ.* **2006**, *83*, 1325.

(113) Dutton, P. L. *Methods Enzymol.* **1978**, *54*, 411–435.

(114) Sreerama, N.; Woody, R. W. *Methods Enzymol.* **2004**, *383*, 318–351.

(115) Shih, W.; Quispe, J.; Joyce, G. *Nature* **2004**, *427*, 618–621.

(116) Onyshchenko, M. I.; Gaynutdinov, T. I.; Englund, E. A.; Appella, D. H.; Neumann, R. D.; Panyutin, I. G. *Nucleic Acids Res.* **2011**, *39*, 7114–7123.

(117) Gaylord, B. S.; Massie, M. R.; Feinstein, S. C.; Bazan, G. C. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 34–39.

(118) Pokorski, J. K.; Nam, J. M.; Vega, R. A.; Mirkin, C. A.; Appella, D. H. *Chem. Commun.* **2005**, 2101–2103.

(119) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordcn, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566–568.

(120) Wang, J. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 200–203.

(121) Rhodes, D.; Klug, A. *Nature* **1980**, *286*, 573–578.

(122) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3*, 410–413.

(123) Eriksson, M.; Christensen, L.; Schmidt, J.; Haaima, G.; Orgel, L.; Nielsen, P. E. *Nouv. J. Chim.* **1998**, *22*, 1055–1059.

(124) Lesignoli, E.; Germini, A.; Corradini, R.; Sforza, S.; Galavema, G.; Dossena, A.; Marchelli, R. *J. Chromatogr. A* **2001**, *922*, 177–185.

(125) Özhalici-Ünal, H.; Armitage, B. A. *ACS Nano* **2009**, *3*, 425–433.

(126) Sen, A.; Nielsen, P. E. *Biophys. J.* **2006**, *90*, 1329–1337.

(127) Kaji, T.; Ito, S.; Iwai, S.; Miyasaka, H. *J. Phys. Chem. B* **2009**, *113*, 13917–13925.

(128) Delgadillo, R. F.; Parkhurst, L. J. *Photochem. Photobiol.* **2010**, *86*, 261–272.

(129) Noble, J. E.; Wang, L.; Cole, K. D.; Gaigalas, A. K. *Biophys. Chem.* **2005**, *113*, 255–263.

(130) Dale, R. E.; Eisinger, J.; Blumberg, W. E. *Biophys. J.* **1979**, *26*, 161–193.

(131) Lakowicz, J. R.; Gryczynski, I.; Cheung, H. C.; Wang, C.-K.; Johnson, M. L.; Joshi, N. *Biochemistry* **1988**, *27*, 9149–9160.

(132) Flory, J. D.; Simmons, C. R.; Lin, S.; Johnson, T.; Andreoni, A.; Zook, J.; Ghirlanda, G.; Liu, Y.; Yan, H.; Fromme, P. *Manuscript submitted for publication* **2014**.

(133) Linko, V.; Dietz, H. *Curr. Opin. Biotechnol.* **2013**, *24*, 555–561.

(134) Saccà, B.; Niemeyer, C. M. *Chem. Soc. Rev.* **2011**, *40*, 5910–5921.

(135) Diezmann, F.; Seitz, O. *Chem. Soc. Rev.* **2011**, *40*, 5789–5801.

(136) Saccà, B.; Meyer, R.; Erkelenz, M.; Kiko, K.; Arndt, A.; Schroeder, H.; Rabe, K. S.; Niemeyer, C. M. *Angew. Chemie Int. Ed.* **2010**, *49*, 9378–9383.

(137) Simmons, C. R.; Schmitt, D.; Wei, X.; Han, D.; Volosin, A. M.; Ladd, D. M.; Seo, D.-K.; Liu, Y.; Yan, H. *ACS Nano* **2011**, *5*, 6060–6068.

(138) Liu, J.; Cao, Z.; Lu, Y. *Chem. Rev.* **2009**, *109*, 1948–1998.

(139) Lytton-Jean, A. K. R.; Gibbs-Davis, J. M.; Long, H.; Schatz, G. C.; Mirkin, C. a.; Nguyen, S. T. *Adv. Mater.* **2009**, *21*, 706–709.

(140) Gaillard, C.; Girard, H. a.; Falck, C.; Paget, V.; Simic, V.; Ugolin, N.; Bergonzo, P.; Chevillard, S.; Arnault, J. C. *RSC Adv.* **2014**, *4*, 3566.

(141) Zhang, D. Y.; Seelig, G. *Nat. Chem.* **2011**, *3*, 103–113.

(142) Vogel, S. S.; Thaler, C.; Blank, P. S.; Koushik, S. V. In *FLIM Microscopy in Biology and Medicine*; Periasamy, A.; Clegg, R. M., Eds.; Chapman and Hall/CRC: London, UK, 2009; Vol. 1076, pp. 245–288.

(143) Radi, R.; Thomson, L.; Rubbo, H.; Prodanov, E. *Arch. Biochem. Biophys.* **1991**, *288*, 112–117.

(144) Wallace, C. J.; Proudfoot, a E. *Biochem. J.* **1987**, *245*, 773–779.

(145) Anderson, C.; Halsall, H. *Biochem. Biophys. Res. Commun.* **1977**, *76*, 339–344.

(146) Taniguchi, I.; Iseki, M.; Eto, T. *Bioelectrochemistry Bioenerg.* **1984**, *174*, 373–383.

(147) Hansen, M. N.; Farjami, E.; Kristiansen, M.; Clima, L.; Pedersen, S. U.; Daasbjerg, K.; Ferapontova, E. E.; Gothelf, K. V. *J. Org. Chem.* **2010**, *75*, 2474–2481.

(148) Ihara, T.; Maruo, Y.; Takenaka, S.; Takagi, M. *Nucleic Acids Res.* **1996**, *24*, 4273–4280.

(149) Rudiuk, S.; Venancio-Marques, A.; Baigl, D. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 12694–12698.

(150) Myer, Y. *J. Biol. Chem.* **1968**, *243*, 2115–2122.

(151) Santucci, R.; Ascoli, F. *J. Inorg. Biochem.* **1997**, *68*, 211–214.

(152) Flory, J. D.; Johnson, T.; Simmons, C. R.; Ghirlanda, G.; Fromme, P. *Manuscript submitted for publication* **2014**.

(153) Seeman, N. C. *Nano Lett.* **2010**, 1971–1978.

(154) Baskin, J. M.; Prescher, J. a; Laughlin, S. T.; Agard, N. J.; Chang, P. V; Miller, I. a; Lo, A.; Codelli, J. a; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16793–16797.

(155) Campbell-Verduyn, L. S.; Mirfeizi, L.; Schoonen, A. K.; Dierckx, R. a; Elsinga, P. H.; Feringa, B. L. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 11117–11120.

(156) Becer, C. R.; Hoogenboom, R.; Schubert, U. S. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 4900–4908.

(157) Armitage, B. In *DNA binders and related subjects*; Waring, M. J.; Chaires, J. B., Eds.; Springer: Berlin Heidelberg, 2005; Vol. 253, pp. 55–76.

(158) Smith, J.; Olson, D.; Armitage, B. *J. Am. Chem. Soc.* **1999**, *121*, 2686–2695.

(159) Dodd, D. W.; Hudson, R. H. E. *Electrophoresis* **2007**, *28*, 3884–3889.

(160) Melis, A. *Plant Sci.* **2009**, *177*, 272–280.

(161) Williamson, A.; Conlan, B.; Hillier, W.; Wydrzynski, T. *Photosynth. Res.* **2011**, *107*, 71–86.

(162) Renger, G. *Curr. Sci.* **2010**, *98*, 1305–1319.

(163) Fromme, P.; Grotjohann, I. *Photosynthetic protein complexes: a structural approach*; Fromme, P., Ed.; John Wiley & Sons: Weinheim, 2008.

(164) Shively, J. M.; van Keulen, G.; Meijer, W. G. *Annu. Rev. Microbiol.* **1998**, *52*, 191–230.

(165) Raines, C. a. *Photosynth. Res.* **2003**, *75*, 1–10.

(166) Vinyard, D. J.; Ananyev, G. M.; Dismukes, G. C. *Annu. Rev. Biochem.* **2013**, *82*, 577–606.

(167) Klimov, V. V.; Allakhverdiev, S. I.; Demeter, S.; Krasnovskii, A. A. *Dokl. Akad. Nauk SSSR* **1979**, *249*, 227–230.

(168) Cox, N.; Messinger, J. *Biochim. Biophys. Acta* **2013**, *1827*, 1020–1030.

(169) Luber, S.; Rivalta, I.; Umena, Y.; Kawakami, K.; Shen, J.-R.; Kamiya, N.; Brudvig, G. W.; Batista, V. S. *Biochemistry* **2011**, *50*, 6308–6311.

(170) Pal, R.; Negre, C. F. a; Vogt, L.; Pokhrel, R.; Ertem, M. Z.; Brudvig, G. W.; Batista, V. S. *Biochemistry* **2013**, *52*, 7703–7706.

(171) Siegbahn, P. E. M. *Biochim. Biophys. Acta - Bioenerg.* **2013**, *1827*, 1003–1019.

(172) Adir, N.; Zer, H.; Shochat, S.; Ohad, I. *Photosynth. Res.* **2003**, *76*, 343–370.

(173) Becker, K.; Cormann, K. U.; Nowaczyk, M. M. *J. Photochem. Photobiol. B.* **2011**, *104*, 204–211.

(174) Dasgupta, J.; Ananyev, G. M.; Dismukes, G. C. *Coord. Chem. Rev.* **2008**, *252*, 347–360.

(175) Walter, M. G.; Warren, E. L.; McKone, J. R.; Boettcher, S. W.; Mi, Q.; Santori, E. a; Lewis, N. S. *Chem. Rev.* **2010**, *110*, 6446–6473.

(176) Blankenship, R. E.; Tiede, D. M.; Barber, J.; Brudvig, G. W.; Fleming, G.; Ghirardi, M.; Gunner, M. R.; Junge, W.; Kramer, D. M.; Melis, a.; Moore, T. a.; Moser, C. C.; Nocera, D. G.; Nozik, a. J.; Ort, D. R.; Parson, W. W.; Prince, R. C.; Sayre, R. T. *Science* **2011**, *332*, 805–809.

(177) Lalrempuia, R.; McDaniel, N. D.; Müller-Bunz, H.; Bernhard, S.; Albrecht, M. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 9765–9768.

(178) Duan, L.; Bozoglian, F.; Mandal, S.; Stewart, B.; Privalov, T.; Llobet, A.; Sun, L. *Nat. Chem.* **2012**, *4*, 418–423.

(179) Gust, D.; Moore, T. a.; Moore, A. L. *Faraday Discuss.* **2012**, *155*, 9.

(180) Wasylenko, D. J.; Palmer, R. D.; Berlinguette, C. P. *Chem. Commun. (Camb).* **2013**, *49*, 218–227.

(181) Kanan, M. W.; Nocera, D. G. *Science* **2008**, *321*, 1072–1075.

(182) Harriman, A.; Pickering, I. J.; Thomas, J. M.; Christensen, P. a. *J. Chem. Soc. Faraday Trans. 1 Phys. Chem. Condens. Phases* **1988**, *84*, 2795.

(183) Lee, W. J.; Whitmarsh, J. *Plant Physiol.* **1989**, *89*, 932–940.

(184) Dismukes, G. C.; Brimblecombe, R.; Felton, G. a N.; Pryadun, R. S.; Sheats, J. E.; Spiccia, L.; Swiegers, G. F. *Acc. Chem. Res.* **2009**, *42*, 1935–1943.

(185) Moore, T. a. Photosynthetic water oxidation versus photovoltaic water electrolysis http://solarfuel.clas.asu.edu/photosynthetic-water-oxidation-versus-photovoltaic-water-electrolysis (accessed Mar 29, 2014).

(186) Rendek, K. N.; Fromme, R.; Grotjohann, I.; Fromme, P. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **2013**, *69*, 141–146.

(187) Lu, N.; Pei, H.; Ge, Z.; Simmons, C. R.; Yan, H.; Fan, C. *J. Am. Chem. Soc.* **2012**, *134*, 13148–13151.

(188) Haouchar, D.; Haile, J.; McDowell, M. C.; Murray, D. C.; White, N. E.; Allcock, R. J. N.; Phillips, M. J.; Prideaux, G. J.; Bunce, M. *Quat. Sci. Rev.* **2014**, *84*, 56–64.

(189) Conway, J. W.; McLaughlin, C. K.; Castor, K. J.; Sleiman, H. *Chem. Commun. (Camb).* **2013**, *49*, 1172–1174.

(190) Xie, H.; Yang, D.; Heller, A.; Gao, Z. *Biophys. J.* **2007**, *92*, L70–2.

(191) Steenken, S.; Jovanovic, S. *J. Am. Chem. ...* **1997**, *7863*, 617–618.

(192) McEvoy, J. P.; Brudvig, G. W. *Chem. Rev.* **2006**, *106*, 4455–4483.

(193) Debets, M. F.; van Hest, J. C. M.; Rutjes, F. P. J. T. *Org. Biomol. Chem.* **2013**, *11*, 6439–6455.

(194) White, C. J.; Yudin, A. K. *Nat. Chem.* **2011**, *3*, 509–524.

(195) Chen, Y.; Kamlet, A. S.; Steinman, J. B.; Liu, D. R. *Nat. Chem.* **2011**, *3*, 146–153.

(196) Gustafsson, M.; Curstedt, T.; Jornvall, H.; Johansson, J. *Biochem. J.* **1997**, *806*, 799–806.

(197) Cunningham, F.; Deber, C. M. *Methods* **2007**, *41*, 370–380.

(198) Tulumello, D. V; Johnson, R. M.; Isupov, I.; Deber, C. M. *Biopolymers* **2012**, *98*, 546–556.

(199) Allen, J.; Olson, T.; Oyala, P. *Proc. Natl. Acad. Sci.* **2012**, *109*, 2314–2318.

(200) Blankenship, R. E.; Hartman, H. *Trends Biochem. Sci.* **1998**, *23*, 94–97.

(201) Olson, T. L.; Williams, J. C.; Allen, J. P. *Biochim. Biophys. Acta* **2013**, *1827*, 914–922.

(202) Renger, G. *Biochim. Biophys. Acta* **2012**.

(203) Limburg, B.; Bouwman, E.; Bonnet, S. *Coord. Chem. Rev.* **2012**, *256*, 1451–1467.

(204) Malkin, R.; Rabinowitz, J. C. *Biochem. Biophys. Res. Commun.* **1966**, *23*, 822–827.

(205) Ghirlanda, G.; Osyczka, A.; Liu, W.; Antolovich, M.; Smith, K. M.; Dutton, P. L.; Wand, a J.; DeGrado, W. F. *J. Am. Chem. Soc.* **2004**, *126*, 8141–8147.

(206) Roy, A.; Sarrou, I.; Vaughn, M. *Biochemistry* **2013**, *52*, 7586–7594.

(207) Roy, A.; Madden, C.; Ghirlanda, G. *Chem. Commun. (Camb).* **2012**, *48*, 9816–9818.

(208) Bott, A. *Curr. Sep.* **1999**, *18*, 47–54.

(209) Dennison, C. *Coord. Chem. Rev.* **2005**, *249*, 3025–3054.

APPENDIX A

PERMISION TO REPRODUCE SELECTED FIGURES

American Association for the Advancement of Science (AAAS)

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

Get the printable license.

License Number	3363750839792
License date	Apr 07, 2014
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires
Licensed content author	Hao Yan, Sung Ha Park, Gleb Finkelstein, John H. Reif, Thomas H. LaBean
Licensed content date	Sep 26, 2003
Volume number	301
Issue number	5641
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Order reference number	Figure 1C (bottom left)
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size(pages)	220
Total	**0.00 USD**

4/7/2014 RightsLink- Your Account

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

Get the printable license.

License Number	3363780089189
License date	Apr 07, 2014
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	Rapid Chiral Assembly of Rigid DNA Building Blocks for Molecular Nanofabrication
Licensed content author	R. P. Goodman, I. A. T. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, A. J. Turberfield
Licensed content date	Dec 9, 2005
Volume number	310
Issue number	5754
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Order reference number	Figure 1
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size(pages)	220
Total	**0.00 USD**

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

Get the printable license.

License Number	3363761490155
License date	Apr 07, 2014
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	DNA Origami with Complex Curvatures in Three-Dimensional Space
Licensed content author	Dongran Han, Suchetan Pal, Jeanette Nangreave, Zhengtao Deng, Yan Liu, Hao Yan
Licensed content date	Apr 15, 2011
Volume number	332
Issue number	6027
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Order reference number	Figure 3
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size(pages)	220
Total	**0.00 USD**

4/7/2014 RightsLink- Your Account

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

Get the printable license.

License Number	3363770257525
License date	Apr 07, 2014
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	DNA Gridiron Nanostructures Based on Four-Arm Junctions
Licensed content author	Dongran Han, Suchetan Pal, Yang Yang, Shuoxing Jiang, Jeanette Nangreave, Yan Liu, Hao Yan
Licensed content date	Mar 22, 2013
Volume number	339
Issue number	6126
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	2
Order reference number	Figure 1 and 3
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size(pages)	220
Total	**0.00 USD**

American chemical society (ACS)

Nature Publishing Group (NPG)

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

Get the printable license.

License Number	3363760462710
License date	Apr 07, 2014
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Design and self-assembly of two-dimensional DNA crystals
Licensed content author	Erik Winfree, Furong Liu, Lisa A. Wenzler and Nadrian C. Seeman
Licensed content date	Aug 6, 1998
Type of Use	reuse in a dissertation / thesis
Volume number	394
Issue number	6693
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Figures	Figure 1 and Figure 3
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size (number of pages)	220
Total	**0.00 USD**

4/7/2014 RightsLink- Your Account

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

Get the printable license.

License Number	3363751313289
License date	Apr 07, 2014
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Towards complete cofactor arrangement in the 3.0[thinsp]A resolution structure of photosystem II
Licensed content author	Bernhard Loll, Jan Kern, Wolfram Saenger, Athina Zouni and Jacek Biesiadka
Licensed content date	Dec 14, 2005
Type of Use	reuse in a dissertation / thesis
Volume number	438
Issue number	7070
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 3 (A)
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size (number of pages)	220
Total	**0.00 USD**

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

Get the printable license.

License Number	3363760994606
License date	Apr 07, 2014
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra
Licensed content author	Yu He,Tao Ye,Min Su,Chuan Zhang,Alexander E. Ribbe et al.
Licensed content date	Mar 13, 2008
Type of Use	reuse in a dissertation / thesis
Volume number	452
Issue number	7184
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	4
High-res required	no
Figures	Figures 1, 2, 3, 4
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size (number of pages)	220
Total	**0.00 USD**

4/7/2014 RightsLink- Your Account

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

Get the printable license.

License Number	3363760207166
License date	Apr 07, 2014
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Self-assembly of a nanoscale DNA box with a controllable lid
Licensed content author	Ebbe S. Andersen, Mingdong Dong, Morten M. Nielsen, Kasper Jahn, Ramesh Subramani et al.
Licensed content date	May 7, 2009
Type of Use	reuse in a dissertation / thesis
Volume number	459
Issue number	7243
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
High-res required	no
Figures	Figure 1 and Figure 3
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size (number of pages)	220
Total	**0.00 USD**

License Details

This is a License Agreement between Justin Flory ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

Get the printable license.

License Number	3363780769399
License date	Apr 07, 2014
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Communications
Licensed content title	Programmable multivalent display of receptor ligands using peptide nucleic acid nanoscaffolds
Licensed content author	Ethan A. Englund, Deyun Wang, Hidetsugu Fujigaki, Hiroyasu Sakai, Christopher M. Micklitsch, Rodolfo Ghirlando
Licensed content date	Jan 10, 2012
Type of Use	reuse in a dissertation / thesis
Volume number	3
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
High-res required	no
Figures	Figure 2 and 4
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size (number of pages)	220
Total	**0.00 USD**

Oxford University Press (Oxford)

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

Get the printable license.

License Number	3363780967245
License date	Apr 07, 2014
Licensed content publisher	Oxford University Press
Licensed content publication	Human Molecular Genetics
Licensed content title	Heterogeneity in Telomere Length of Human Chromosomes:
Licensed content author	Peter M. Lansdorp, Nico P. Verwoerd, Frans M. van de Rijke, Visia Dragowska, Marie-Térèse Little, Roeland W. Dirks, Anton K. Raap, Hans J. Tanke
Licensed content date	05/01/1996
Volume number	5
Issue number	5
Type of Use	Thesis/Dissertation
Requestor type	Academic/Educational institute
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No
Author of this OUP article	No
Order reference number	Figure 1
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Estimated size(pages)	220
Publisher VAT ID	GB 125 5067 30
Total	**0.00 USD**

John Wiley and Sons (Wiley)

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and John Wiley and Sons ("John Wiley and Sons"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

Get the printable license.

License Number	3363780238884
License date	Apr 07, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	Angewandte Chemie
Licensed content title	Single-Molecule Protein Encapsulation in a Rigid DNA Cage
Licensed copyright line	Copyright © 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
Licensed content author	Christoph M. Erben,Russell P. Goodman,Andrew J. Turberfield
Licensed content date	Nov 6, 2006
Start page	7574
End page	7577
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1
Will you be translating?	No
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Expected size (number of pages)	220
Total	**0.00 USD**

4/7/2014 RightsLink- Your Account

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and John Wiley and Sons ("John Wiley and Sons"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

Get the printable license.

License Number	3363750315368
License date	Apr 07, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	ChemPhysChem
Licensed content title	DNA Tile Based Self-Assembly: Building Complex Nanoarchitectures
Licensed copyright line	Copyright © 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
Licensed content author	Chenxiang Lin,Yan Liu,Sherri Rinker,Hao Yan
Licensed content date	Jul 10, 2006
Start page	1641
End page	1647
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1
Will you be translating?	No
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Expected size (number of pages)	220
Total	**0.00 USD**

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and John Wiley and Sons ("John Wiley and Sons"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

Get the printable license.

License Number	3363770733983
License date	Apr 07, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	Angewandte Chemie
Licensed content title	Self-Assembled Peptide Nanoarrays: An Approach to Studying Protein–Protein Interactions
Licensed copyright line	Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
Licensed content author	Berea A. R. Williams,Kyle Lund,Yan Liu,Hao Yan,John C. Chaput
Licensed content date	Mar 16, 2007
Start page	3111
End page	3114
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 3
Will you be translating?	No
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Expected size (number of pages)	220
Total	**0.00 USD**

4/7/2014 RightsLink- Your Account

Copyright Clearance Center RightsLink®

My Orders My Library My Profile Welcome justin.flory@asu.edu Log out | Help

My Orders > Orders > All Orders

License Details

This is a License Agreement between Justin Flory ("You") and John Wiley and Sons ("John Wiley and Sons"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

Get the printable license.

License Number	3363760657541
License date	Apr 07, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	WileyoBooks
Licensed content title	Overview of Photosynthesis
Book title	Photosynthetic Protein Complexes: A Structural Approach
Licensed copyright line	Copyright © 2008 Wiley-VCH Verlag GmbH & Co. KGaA
Licensed content author	Petra Fromme,Ingo Grotjohann
Licensed content date	Sep 5, 2008
Start page	1
End page	22
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	From Chapter 1, section 1.2 "The Protein Complexes of Oxygenic Photosynthesis"
Will you be translating?	No
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Expected size (number of pages)	220
Total	**0.00 USD**

License Details

This is a License Agreement between Justin Flory ("You") and John Wiley and Sons ("John Wiley and Sons"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

Get the printable license.

License Number	3363770926574
License date	Apr 07, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	Angewandte Chemie
Licensed content title	DNA as a Molecular Ruler: Interrogation of a Tandem SH2 Domain with Self-Assembled, Bivalent DNA–Peptide Complexes
Licensed copyright line	Copyright © 2011 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
Licensed content author	Hendrik Eberhard,Franziska Diezmann,Oliver Seitz
Licensed content date	Mar 31, 2011
Start page	4232
End page	4236
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1
Will you be translating?	No
Title of your thesis / dissertation	PNA-Polypeptide Assembly in a 3D DNA Nanocage for Building Artificial Catalytic Centers
Expected completion date	May 2014
Expected size (number of pages)	220
Total	**0.00 USD**

APPENDIX B

CO-AUTHOR APPROVAL

I verify that the following co-authors have approved of my use of our co-authored work, either published or submitted for publication, for use in my dissertation:

Alessio Andreoni (Arizona State University)

Petra Fromme (Arizona State University)

Giovanna Ghirlanda (Arizona State University)

Trey Johnson (Arizona State University)

Su Lin (Arizona State University)

Yan Liu (Arizona State University)

Sandip Shinde (Arizona State University)

Chad Simmons (Arizona State University)

Hao Yan (Arizona State University)

James Zook (Arizona State University)

BIOGRAPHICAL SKETCH

Justin Flory received his B.S. in physics at Sonoma State University in 2002, where he also developed a fascination with the sophisticated molecular machinery of biological systems. Mr. Flory first applied his academic training by working as an engineer for Symmetricom until 2007 and later working on implantable blood sugar monitoring systems for Abbott Diabetes Care until 2009. During these experiences Mr. Flory realized his fascination with applying our understanding of living organisms to solve important technological challenges, such as mitigating climate change while meeting our future energy demand. This inspired Mr. Flory to join the Biological Design program at Arizona State University, where he researched how to build an artificial device to convert water and sunlight into fuel, based on the natural process of photosynthesis. He worked for Dr. Petra Fromme in the Center for Bio-Inspired Solar Fuel Production, sponsored by the US Department of Energy. Upon receiving his doctorate in May 2014, he plans to apply his research experience in artificial photosynthesis to improve solar energy technologies.