Designing Recognition Elements based on DNA Scaffolds

by

Berea Williams

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

.

ARIZONA STATE UNIVERSITY

May 2010

UMI Number: 3410557

All rights reserved

INFORMATION TO ALL USERS The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3410557 Copyright 2010 by ProQuest LLC. All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC 789 East Eisenhower Parkway P.O. Box 1346 Ann Arbor, MI 48106-1346

Designing Recognition Elements based on DNA Scaffolds

by

Berea Williams

has been approved

April 2010

Graduate Supervisory Committee:

John Chaput, Chair Hao Yan Yan Liu

ACCEPTED BY THE GRADUATE COLLEGE

ABSTRACT

Following the discovery of deoxyribonucleic acid (DNA) and its role as the genetic carrier of life, researchers have investigated the potential use of DNA as a molecular tool in chemical biology. Along this line, my Ph.D. thesis examines the use of DNA oligonucleotides and peptide-oligonucleotide conjugates (POCs) as material to capture and organize biological macromolecules in nanotechnology and molecular biology.

Using POC molecules, we developed a new method of positioning proteins at addressable locations on DNA nanostructures. Described in chapter 2 is the design of a DNA nanostructure that displayed single-stranded DNA probes that were complimentary in sequence to an oligonucleotide conjugated to the c-myc peptide. When the DNA array was assembled, the c-myc POCs were organized on the array surface and remained accessible for recognition by the c-myc antibody. This research was the first demonstration of peptide nanoarrays using POCs to specifically organize proteins on DNA nanostructures.

In chapter 4, we created bivalent reagents (termed synbodies) using POC molecules composed of complimentary strands of DNA independently conjugated to two peptides that exhibited non-overlapping binding to a target protein. When the POCs are hybridized, bivalent reagents are created that display two peptides on a double-stranded DNA scaffold with varying spatial separation between the peptides. We have observed that the synbodies with the optimal peptide separation distance exhibit ~1000-fold enhancement in binding affinity to the target protein in comparison to the peptides that comprise the bivalent reagents.

Alternatively, chapter 3 describes the use of *in vitro* evolution to select for single strand DNA molecules (termed aptamers) with discrete tertiary structures that are capable of specifically binding the histone H4 protein with a acetyl modification. When compared to a commercial antibody, the DNA aptamer binds with similar affinity, but much stronger specificity when probed with different modified and unmodified protein targets. This research demonstrated the ability of DNA aptamers to specifically recognize small chemical modifications among protein targets.

Together, these projects demonstrate the utility of DNA oligonucleotides and POC molecules as chemical reagents to precisely organize biological materials on the nanoscale and produce high affinity protein binding agents.

ACKNOWLEDGMENTS

There have been many people in the past five years that have helped me complete my goal of attaining a Ph.D. Most of all, I must thank my Ph.D. advisor, Dr. John Chaput. Without Dr. Chaput's support and encouragement, I would not have accomplished half of the achievements I did. Additionally, I really appreciated Dr. Chaput's understanding of my need to participate in "extracurricular" activities that consumed some of my time that should have been spent conducting research.

I was extremely fortunate to have excellent mentors that challenged and championed me during my Ph.D. work; Dr. Hao Yan, Dr. Yan Liu, Dr. Neal Woodbury, Dr. Chris Diehnelt, Dr. Stuart Lindsay, Dr. Stephen Johnston, Dr. Ira Bennett, and Dr. Jameson Wetmore. I would also like to thank all my friends and colleagues that supported me through this experience; Jo Woods, Carole Flores, Dr. Paul Belcher, Dr. Matt Greving, Jinglin Fu, Dr. Sherri Rinker, Dr. Elizabeth McCullum, and the entire Chaput lab (Megan, Dr. Wellensiek., Rui, Hanyang, Bing, Josh, Brian P., Richard, Jinglei, Linda, Dr. Nelsson, Karen, Doris, Jenn, YingWei, Mike, Brett, and Nora).

In addition, I am extremely thankful for the support of my family; Adam, June, Bruce, Lillie, Justin, Jason, Mable, Shirley, Uncle Tony, Ruth, Eamon, Linda, and Larry. Their unconditionally love made me proud of myself no matter what I accomplished.

Finally, I would like to thank the National Science Foundation for providing me with the IGERT fellowship, which allowed me to conduct my research without worrying about financial support and taught me how to put scientific research in a societal context.

v

TABLE OF CONTENTS

LIST OF TABL	ES x
LIST OF FIGU	XESxi
CHAPTER	
1 INT	IRODUCTION 1
1	.1 Structural DNA Nanotechnology
1	.2 Nucleic Acid Aptamers
1	.3 Synthetic Protein Affinity Reagents
1	.4 Linking Strategies for Multivalent Affinity Reagents
	1.4.1 Small Molecule Linkers24
	1.4.2 Polymer-Based Linkers28
	1.4.3 Dendrimer Scaffolds
	1.4.4 Cyclic Core Scaffolds
	1.4.5 Amino Acid Linkers46
	1.4.6 Deoxyribonucleic Acid Linkers55
1	.5 References
2 SEI	LF-ASSEMBLED PEPTIDE NANOARRAYS: A NEW APPROACH TO
	STUDYING PROTEIN-PROTEIN INTERACTIONS 79
2	2.1 Contributions
2	2.2 Abstract
2	2.3 Introduction
2	2.4 Results

Page

	2.5 Conclusions
	2.6 Experimental Design
	2.7 References
3	EVOLUTION OF A HISTONE H4-K16 ACETYL-SPECIFIC DNA
	APTAMER
	3.1 Contributions
	3.2 Abstract
	3.3 Introduction
	3.4 Results 102
	3.5 Conclusions 111
	3.6 Experimental Design 113
	3.7 References 119
4	CREATING PROTEIN AFFINITY REAGENTS BY COMBINING
	PEPTIDE LIGANDS ON SYNTHETIC DNA SCAFFOLDS 121
	4.1 Contributions
	4.2 Abstract 123
	4.3 Introduction
	4.4 Results 128
	4.5 Conclusions
	4.6 Experimental Design155
	4.7 References

5	SYNTHESIS OF PEPTIDE-OLIGONUCLEOTIDE CONJUGATES USING
	A HETEROBIFUNCTIONAL CROSSLINKER 172
	5.1 Abstract 173
	5.2 Introduction 174
	5.3 Basic Protocol 1 178
	5.4 Basic Protocol 2
	5.5 Basic Protocol 3194
	5.6 Basic Protocol 4
	5.7 Basic Protocol 5
	5.8 Commentary
	5.9 References
6	CONCLUDING REMARKS 221
	6.1 Abstract
	6.2 Conclusions 223
	6.3 Key Observations 226
	6.4 Future Direction
	6.5 References
REFEREN	CES 233
APPENDIX	ζ 249
	COPYRIGHT PERMISSION
,	The American Association of the Advancement of Science 250
	American Chemical Society

Page

Elsvier	260
Nature Publishing Group	264
International Union of Crystallography	266
Wiley-VCH Verlag GmbH & Co. KGaA	266

LIST	OF	TABLES

Table	Page
1.1	Flexible Acyclic Linkers
1.2	Dendrimer and Cyclic Scaffolds
1.3	Amino Acid Linkers
1.4	Deoxyribosenucleic Acid Linkers
1.5	Strengths and Weaknesses of Various Linker Strategies
2.1	DNA sequences used to assemble the ABCD array
3.1	Dissociation Constants and Specificity Values of Aptamer 4.20 and H4-
	K16Ac Antibody by SPR 112
3.2	Round 4 DNA Sequences 116
4.1	Gal80 Peptide Fluorescent Values
4.2	Transferrin Peptide Fluorescent Values
4.3	Transferrin-Peptide Crosslinking Analysis 145
4.4	Gal80 Synbody Oligonucleotide Sequences
4.5	Transferrin Synbody Oligonucleotide Sequences
5.1	Urea-PAGE Mixture
5.2	Migration of Running Dye in Relation to Oligonucleotide Length in
	Nucleotides 186
5.3	Molar Extinction Coefficient for each Nucleobase 192
5.4	Native-PAGE Mixture
5.5	Gradient of Mobile Phase B for RP-HPLC Purification of Crude POC
	Material 207

LIST OF FIGURES

Figure	Page
1.1	Deoxyribonucleic Acid Nucleobases 2
1.2	Double Stranded B-form DNA Duplex 4
1.3	Specific Positioning of Streptavidin on a DNA Array
1.4	Functionalization of DNA Nanoarrays with a Thrombin Aptamer
1.5	Structure of Succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate
1.6	Antibody Functionalized DNA Nanoarrays 11
1.7	Crystal Structure of a RNA Aptamer 13
1.8	Chemical Structure of Theophylline and Caffeine
1.9	Crystal Structure of a Nucleosome 17
1.10	Monoclonal Antibody Production
1.11	Structure Activity Relationship by Nuclear Magnetic Resonance
1.12	Combinatorial Target-Guided Ligand Assembly 27
1. 13	Polymer-Linked Ligand Dimers (PLDs) 29
1.14	PEG Spacer Linked Bivalent Aptamers
1.15	Dendrimer Display
1. 16	Multiple Antigen Peptide (MAP)
1. 17	Shiga-Like Toxin I Inhibitor, STARFISH 40
1. 18	Dendritic β–Cyclodextrin Derivatives
1. 19	Multivalent Homotrimer CD40L Mimicks 44
1.20	Engineered Antibody Conformations 47

,

•

.

1.21	Avidity Multimers: Avimers
1.22	Multivalent Peptide Ligands to Human Melanocortin Receptor 4 53
1.23	Oligonucleotide-Linked Organic Fragments
1.24	DNA Linked Bivalent Thrombin Inhibitor 59
1.25	Self-Assembling Bivalent Aptamers Using a DNA Tile
1.26	Combinatorial Examination of Linker and Ligands (CELL)
2.1	Self-assembled 2D Peptide Nanostructure
2.2	Non-denaturing Gel Electrophoresis Assay Showing Efficient Hybridization
	of the DNA-tagged Peptide Fusion to the DNA Array
2.3	AFM Imaging of the Peptide Nanoarray
3.1	Selection Strategy Used to Identify H4-K16Ac-specific Aptamers
3.2	Capillary Electrophoresis Separation of Histone Aptamers 105
3.2 3.3	Capillary Electrophoresis Separation of Histone Aptamers 105 Secondary Structure of Histone H4-K16Ac Aptamers using the DNA
3.2 3.3	Capillary Electrophoresis Separation of Histone Aptamers
3.23.33.4	Capillary Electrophoresis Separation of Histone Aptamers
3.23.33.43.5	Capillary Electrophoresis Separation of Histone Aptamers
 3.2 3.3 3.4 3.5 4.1 	Capillary Electrophoresis Separation of Histone Aptamers
 3.2 3.3 3.4 3.5 4.1 4.2 	Capillary Electrophoresis Separation of Histone Aptamers
 3.2 3.3 3.4 3.5 4.1 4.2 4.3 	Capillary Electrophoresis Separation of Histone Aptamers
 3.2 3.3 3.4 3.5 4.1 4.2 4.3 4.4 	Capillary Electrophoresis Separation of Histone Aptamers
 3.2 3.3 3.4 3.5 4.1 4.2 4.3 4.4 	Capillary Electrophoresis Separation of Histone Aptamers

Figure

4.6	Relative Binding Affinity for Homo-Peptide Gal80 Synbody
	Combinations 138
4.7	Characterization of the Gal80 Synbody 141
4.8	Transferrin Ligand Discovery Using Peptide Microarrays 142
4.9	Generation of a Transferrin Synbody 146
4.10	Binding Affinity of SC-6 to Human Transferrin 148
4.11	Specificity Assay of Gal80 Synbody 149
4.12	Control Pulldown Assay for Non-specific Binding to Endogenous E. coli
	Proteins 150
4.13	SPR Analysis of Gal80 Peptides and Synbody Construct 13 167
5.1	Example of Heterobifunctional Crosslinking Reagents 176
5.2	Amine-Modified Oligonucleotide Conjugation 177
5.3	Amine-modified Deoxynucleotide Phosphoramidites 179
5.4	Conjugation of SMCC to Amine Modified Deoxynucleotide189
5.5	Conjugation of SMCC-Oligonucleotide to Cysteine-Containing Polypeptides
5.6	Example chromatograms of the reverse phase-HPLC analysis 198
5.7	MALDI-TOF mass spectroscopy analysis 205
5.8	Analysis of POCs by native-PAGE mobility shift assay

CHAPTER 1

INTRODUCTION

From the discovery of the double helix to the ability to sequence the human genome, our understanding and application of deoxyribonucleic acid (DNA) has changed dramatically over the last 50 years (1). This relatively simple polymer, made of adenosine (A), thymidine (T), cytidine (C), and guanodine (G) nucleotides (Figure 1.1), has the *in vivo* role of encoding the genetic instructions necessary for a cell to survive and replicate. In the laboratory, researchers have exploited the use of this polymer and its relatively simple hybridization principles by transforming DNA into a powerful *in vitro* molecular tool. Beyond the storage of genetic information, DNA has been used as a scaffold for complex molecular structures, high affinity protein binding reagents, fluorescent signal reporters, nano-mechanical devices, and conductive nano-wire networks (1-3).

The introduction of this dissertation will highlight the use of DNA as a molecular tool the areas of structural DNA nanotechnology and *in vitro* evolution of functional nucleic acid sequences, while chapters 2 and 3 of this thesis will focus on recent work our laboratory has contributed to the advancement of these fields. The knowledge gained from these projects was then applied to the development of bivalent affinity reagents, and the full report of using double stranded DNA as a linking strategy is given in chapter 4. The research presented in chapters 2 and 4 rely on the use of peptide-oligonucleotide conjugate molecules and a detailed protocol of their synthesis is described in chapter 5.





Figure 1.1. Deoxyribonucleic Acid Nucleotides. Purine bases, adenine and guanine, pair with pyrimidine bases, thymine and cytosine, respectively. Nucleobases are covalently linked to a deoxyribose ring to form nucleosides. The addition of a phosphate group at the 5' carbon position of the deoxyribose forms the nucleotides. Joining of the 5' phosphate group of one nucleotide to the 3' hydroxyl group of another nucleotide forms the nucleic acid polymer DNA.

1.1 Structural DNA Nanotechnology

In nature, DNA is most often found in the double stranded form where the nucleobases of two single strands of DNA pair together in an anti-parallel fashion. The intermolecular base pairing follows the rules estabilished by James Watson and Francis Crick, such that adenine pairs with thymine and cytosine pairs with guanine. This base pairing allows single stranded DNA to remain hybridized as a double helix, maintaining precise nanoscale dimensions (Figure 1.2). Using the fundamental Watson and Crick base pairing principles, Ned Seeman conceived the idea of designing specific DNA sequences to self-assemble into branched nanostructures, such as a 4-arm junction (4). Additionally, Dr. Seeman theorized that single stranded overhangs, or "sticky ends", could be combined with branched DNA molecules to enable individual structures to interact with one another. He proposed that the building blocks could be engineered to have complementary sticky ends such that the individual units could self-hybridize into 2dimensional arrays composed purely of DNA (4). The rational design of DNA sequences that self-assemble into desired structures is the basis of structural DNA nanotechnology. Since Dr. Seeman's original proposal in 1982, DNA has been used as a building block to create a diverse set of complex DNA arrays composed of elaborate 3-dimensional structures, chemical modifications, and even controllable moving parts (1).

Beyond the creation of several types of basic building blocks, DNA nanostructures have been used to arrange biological material with nanometer (nm) precision. The predictable assembly of DNA structures allows for the precise placement



Figure 1.2. Double Stranded B-form DNA Duplex. Two single strands of DNA with complementary sequences will base pair in an anti-parallel fashion resulting in a double stranded helix composed of 10 base pairs per turn with a width of 2 nanometers and a length of 3.4 nanometers.

of short segments of single stranded DNA (ssDNA) that act as probes for the nanoscale organization of other materials. These probes can be used to recruit complementary nucleic acid sequences that contain chemical modifications or exhibit a particular function. The use of ssDNA probes strands to recruit external materials has led to the functionalization of DNA arrays with small organic molecules, inorganic particles, functional nucleic acid sequences and proteins (5).

The programmable nanometer arrangement of the protein streptavidin was originally demonstrated using a DNA nanoarray composed of four distinct 4-arm (4 x 4) branched DNA tiles (6). Each DNA tile consisted of unique sticky end sequences that were designed to hybridize with other 4 x 4 DNA tiles to generate a large repeating DNA lattice. The shape and size of the resulting scaffold was determined by the design of the sticky end sequences (Figure 1.3). Specific 4 x 4 DNA tiles were designed to contain a ssDNA probe that was complementary in sequence to a biotin modified DNA strand. When the strands were self-assembled, the biotinylated strands were displayed on the corresponding DNA tiles at discrete locations within the array. When the DNA lattices were subsequently incubated with streptavidin, which has strong binding affinity to biotin (K_d ~10⁻¹⁵ M), the protein bound the biotin molecules and was arranged on the DNA nanoarrays in an organized manner. This approach described the use of DNA nanoarrays to specifically organize small molecules and ultimately organize biological material.

While this method demonstrates an effective way of arranging streptavidin on DNA nanoarrays using the biotin/streptavidin interaction, its' application is limited by the number of small molecules that are easily conjugated to DNA. Alternatively,



Figure 1.3. Patterning of Streptavidin on a DNA Array. A-F) Directed by the sticky end design of each 4 x 4 DNA tile, tiles containing biotinylated DNA specifically position streptavidin at discrete locations on the array with nanometer precision. Reprinted (adapted in part or whole) with permission from Nano Letters (6). Copyright 2005 American Chemical Society.

incorporating single strands of DNA or RNA that are capable of binding a target ligand (termed aptamers) into DNA nanoarrays can greatly extend the functionality of nanostructures with a wider range of biological material. One of the first demonstrations of this approach integrated a thrombin binding DNA aptamer into a DNA triple crossover tile (Figure 1.4) (7). The DNA tile was designed with precise spatial separation of each thrombin aptamer and when the array was incubated with thrombin, the protein was specifically positioned with 17 nanometer interparticle spacing. Because of the wide range of aptamer/target combinations that exist (8, 9), the application of aptamers to structural DNA nanotechnology greatly expanded the repertoire of molecules that can be patterened using DNA nanoarrays.

In collaboration with Dr. Hao Yan and Dr. Yan Liu, our laboratory further expanded the addressability of DNA arrays for proteins using peptide-oligonucleotide conjugates (POCs). There are many chemical methods to covalently link DNA oligonucleotides to amino acid polymers, including in-line, solid phase synthesis and fragment conjugation. The solid phase synthesis approach utilizes step-wise peptide and DNA synthesis, and protecting strategies that are compatible with both synthesis steps. Alternatively, fragment conjugation covalently links a peptide to a DNA sequence in solution post-synthesis and does not require special equipment (*10*). A detailed fragment conjugation protocol is described in chapter 5, where a cysteine containing polypeptide is linked to an amine-modified DNA sequence using the hetero-bifunctional crosslinking reagent, succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) (Figure 1.5).



Figure 1.4. Functionalization of DNA Nanoarrays with a Thrombin Aptamer. A) The thrombin DNA aptamer was designed into a DNA tile, B) self-assembled into a three-helix bundle, and C-D) incubated with thrombin. The thrombin protein was specifically positioned at the locations of the aptamer (red) and not of that of a random sequence (black). Reprinted with permission Angewandte Chemie International Edition (7). Copyright 2005 Gesellschaft Deutscher Chemiker.



succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate

Figure 1.5. Structure of Succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate. This small molecule, also known as SMCC, is a bifunctional crosslinking reagent. The maleimide portion of the molecule specifically reacts with thiols and the Nhydroxysuccinimide reacts with primary amine moieties. This strategy was used to create POCs for the functionalization of an ABCD array using the c-myc peptide and its cognate antibody (termed nanodisplay). Chapter 2 provides a detailed account of this work. In brief, the DNA portion of the POC molecule was designed to be complementary in sequence to a ssDNA probe on the D-tile of an ABCD array (*11*). When the POC molecules were hybridized to the array, the c-myc polypeptides were specifically positioned on the D-tile of the array and were capable of binding and organizing the anti-c-myc antibody (Figure 1.6). The resulting arrays positioned the antibody in repeating lines spaced by 64 nanometers. This research demonstrated the use of POCs to functionalize DNA arrays with biological material and the further expansion of DNA nanotechnology.



Figure 1.6. Antibody Functionalized DNA Nanoarrays. A D-tile was designed to display peptide-oligonucleotide conjugate molecules and together with tiles A-C, self-assemble into an ABCD array. The assembled array presented the peptide in repeating lines spaced by 64 nanometers. When incubated with its cognate antibody, the antibody was also displayed in repeating lines spaced every 64 nanometers. Reprinted with permission from Angewandte Chemie International Edition (*11*). Copyright 2007 Gesellschaft Deutscher Chemiker.

1.2 Nucleic Acid Aptamers

DNA is found naturally in the double stranded form through the hybridization of two complementary single stranded sequences. However, following similar base pairing rules, a single strand of DNA can self-hybridize intramolecularly to form a stably folded tertiary structure that can discrete interactions with ligands, similar to the way proteins interact. It is this folded structure that allows some single stranded nucleic acid sequences to adopt an induced fit with ligands such as proteins, peptides, and small molecules (Figure 1.7).

The term aptamer was coined by Andrew Ellington and Jack Szostak to describe a single stranded DNA or RNA molecule that folds into a 3-dimensional structure capable of binding a target molecule (12). The selection of aptamers from a pool of nucleic acid sequences was independently developed by two laboratories in 1990 (12, 13). It was shown that functional RNA molecules could be identified in vitro from a large pool of unique sequences when a selective pressure (binding a target molecule) was applied to the library. The sequences that functioned under the selective conditions would be enriched by polymerase chain reaction (PCR) and were subjected to another round of selective pressure. After multiple partitioning and enrichment steps, the nucleic acid sequences that survived the selection would have the desired function determined by the selective pressure. This process was termed Systematic Evolution of Ligands by Exponential Enrichment (SELEX) by Craig Tuerk and Larry Gold (13).



A

Figure 1.7. Crystal Structure of an RNA Aptamer. A) Cartoon ribbon and B) stick representation of the tertiary structure of a malachite green binding RNA aptamer developed through SELEX (PDB: 1F1T) (14). The RNA structure is stabilized through hydrogen bond interactions, while some nucleotide bases are flipped outside the core structure in order to increase interactions with ligands, similar to the way a protein folds and functions.

Two characteristics are analyzed when determining the quality of an aptamer; affinity and specificity. The affinity is defined as the strength of a binding interaction between two molecules and is measured in terms of dissociation constant (K_d). The dissociation constant for a given interaction is the concentration of the ligand at which 50% of the analyte is bound. For example, antibodies typically have low nanomolar (nM) K_d values for an antigen, therefore, the antibody is at low nanomolar concentration when 50% of the antigen is bound in solution. While affinity is an excellent measure of the quality of a molecular recognition event, specificity of the interaction is also a critical characteristic. Specificity is the measure of binding fidelity between a given molecule and the binding partner compared to the binding to molecules other than the target. This can be calculated by taking the ratio of the off-target K_d divided by the on-target K_d , the larger the ratio is the more selective the binding agent. It is very important that an aptamer binds the target molecule with high affinity, but does not bind other molecules with high affinity as well.

To date, the most specific aptamer ever identified was discovered by the Polisky laboratory and is capable of discriminating between two molecules that differ by a single methyl group (15). The target for the RNA aptamer selection was theophylline (Figure 1.8A), a bronchodialator for treatment of asthma and bronchitis. After five rounds of SELEX for theophylline binding, three rounds of "counter-SELEX" was employed where the library of theophylline binding aptamers were subjected to a caffeine (Figure 1.8B) elution. The molecules that bound caffeine were discarded, while those retained on the theophylline column were amplified. The best RNA aptamer obtained from the SELEX





theophylline

Figure 1.8. Chemical Structure of Theophylline and Caffeine. One of the most specific aptamers ever discovered is able to discriminate A) theophylline from B) caffeine, which differs by a single methyl group (tip of arrow).

process had high affinity to theophylline ($K_d = 600$ nM) and greater than 10,000 fold specificity against caffeine. This research demonstrated the ability of an aptamer to discriminate between molecules with subtle differences as small as a single methyl group.

Small modifications such as methyl, phosphoryl, and acetyl moieties can have a dramatic effect in biological processing. For example, genetic DNA is tightly packed inside cells by wrapping itself around four dimerized histone proteins (H2a, H2b, H3, and H4) to form nucleosomes (Figure 1.9). The processing of DNA is directly controlled by the addition and removal of small molecules, called post-translational modifications (PTMs), on the N-terminus of histone proteins. PTMs consist of over 20 small molecules and modulate most eukaryotic genomes (*16*). PTMs on histone proteins create a specific message to other proteins on how to process the DNA, and the pattern of PTMs is called the epigenetic code.

The most abundant modification on histone proteins is acetylation, which is the addition of an acetyl moiety (COCH₃) onto lysine amino acid side chains (*16*). The addition of an acetyl residue (molecular weight of 43 g/mol) to a histone protein may seem insignificant, however it actually has a dramatic biological affect. Recent research has shown the acetylation of histone H4 protein at lysine residue position 16 affects chromatin structure and protein interactions (*17*). The presence of the acetyl modification promotes the accessibility of proteins and cofactors to the genomic DNA, therefore increasing transcription. However, the same acetyl modification on a different histone protein or a different lysine residue on the same histone protein may have a different influence. Being able to distinguish histone proteins with specific PTMs on specific



Figure 1.9. Crystal Structure of a Nucleosome. The nucleosome (A) front view and B) side view) is made of 8 histone proteins, H2A (blue), H2B (red), H3 (orange), and H4 (green), and 146 base pairs of DNA (PDB:1EQZ) (18).

amino acid residues will provide much needed insight into the role of PTMs on DNA processing and in genetic disease states.

Since the initial development of *in vitro* evolution of nucleic acids, hundreds of aptamers have been evolved to bind a wide range of target molecules such as peptides, proteins, small molecules, and even whole cells (*19*). To date, however, only two aptamers have been evolved to a particular PTM, a farnysyl binding RNA aptamer and a histone H4-K16 acetyl (AcH4-K16) binding DNA aptamer (*20, 21*). The latter was discovered in collaboration between our laboratory and Dr. Stuart Lindsay's laboratory. Chapter 3 describes the identification of a novel aptamer to the acetyl PTM on histone H4 at position K16. The resulting aptamer has high affinity and high specificity for AcH4-K16 over H4 without the PTM and H4 with the acetyl at another lysine position. This research is a first step in the generation of high quality aptamers to proteins with PTMs that may help deconvolute the epigenetic code that controls the processing of DNA.

1.3 Synthetic Protein Affinity Reagents

The ability to detect human diseases at an early stage requires advanced technology to detect small, yet significant protein concentrations corresponding to considerable biological changes. Protein biomarkers are a direct indicator of the presence of a disease and their relative abundance can act as a human health profile. Currently, the amount of protein biomarkers present in the blood is determined by a variety of analytical methods, such as enzyme-linked immunosorbant assay (ELISA), pathology, histology, mass spectroscopy, and protein capture array technology (*22*).

One similarity between most biomaker detection techniques is the use of antibodies as the protein affinity reagent. Due to their characteristically high affinity and specificity to their corresponding target, antibodies are the "gold standard" for protein detection. Antibodies used for research and diagnostic purposes and are developed through the inoculation of mammals with a target antigen, followed by a long incubation period and a second round of injections (Figure 1.10). During this process, polyclonal antibodies are expressed by B-lymphocytes located in the spleen and are removed from the animal for purification. Select B-lymphocytes expressing antibodies with desired characteristics are fused with immortalized myeloma cells to create hybridoma cells. The hybridomas are then screened for activity, and the active antibodies are isolated as monoclonal species. This process can take around 30 weeks and cost ~\$4,000 (Antagene, <u>http://www.antageneinc.com/monoclonal_antibody.html</u>). In addition to their expensive and long development time, antibodies are very sensitive to their environment leading to 90% denaturation when subject to passive adsorption on a solid support (23).



Figure 1.10. Monoclonal Antibody Production. An animal is first immunized with an antigen to initiate the immune system to produce antibodies against the antigen. The spleen is then removed from the animal and B-lymphocytes are collected. The B-lymphocytes are fused with myeloma cells to produce hybridoma cancer cells. These cells are screened for antibody function and partitioned based on performance, resulting in monoclonal entities. Reprinted with permission from Access Excellence @ the National Health Museum, copyright 2007 (24).

Due to the limitations of traditional antibodies, a variety of alternative protein capture agents have been developed, such as immunoglobulin fragments, proteins, aptamers, peptides, and small molecules. Immunoglobulin fragments correspond to the variable regions of the antigen recognition fragments (Fv) of an antibody. The Fv region has a conserved structure composed of β -sheets and loops structures, however, the length and amino acid assignment in this region is hypervariable. The Fv region is structurally more stable then a complete immunoglobulin, and has been the focus of single chain antibody phage display libraries (25). Fv's have been evolved to have similar binding characteristics as antibodies, but within a shorter time frame and lower cost.

Similarly, proteins with stable tertiary structure that can withstand numerous amino acid substitutions or sequence insertions can serve as scaffolds to create protein libraries (26). Affibodies are an example of a protein scaffold, made up of the Z-domain of protein A with 13 randomized residues on 2 α -helices. Affibodies have been created to selectively bind to Taq Polymerase and Her2/neu with affinities of ~50 nM (27).

Alternatively, peptides and small molecules are easily synthesized and modified, and remain stable on solid supports. However, peptides and small molecules typically have low micromolar (μ M) binding affinity and need to be enhanced in order to be useful for most biological assays (28). One approach is to combine two or more ligands together as multivalent affinity reagent. Tethering multiple noncompetitive ligands together as one molecule reduces the degree of motion at the binding site and reduces the entropy loss during binding. As a result, the energy of binding between the multivalent ligand molecule and the target receptor is decreased, resulting in a stronger binding interaction and a lower K_d value. This phenomenon is called the enthalpic chelate effect (29).
1.4 Linking Strategies for Multivalent Affinity Reagents

Many biological interactions involve multivalency, including antibodies. Each class of antibodies displays a different number of binding sites; IgD (two), IgG (two), IgE (two), IgA (two, four or six), and IgM (ten) (30). For example, IgM is a naturally occurring pentavalent antibody (5 antibodies linked together, each displaying two binding sites) that is produced by B-lymphocytes in order to eliminate pathogens in the early stages of infection. Although IgM antibodies are much lower in abundance (~10%) than other antibodies, they are much more effective at binding a given pathogen due to their pentavalency (30).

The increase in binding affinity and specificity gained through multivalency has been the base of inspiration for researchers when it comes to creating high affinity reagents. Whitesides and coworkers used the idea of multivalency to create one of the strongest receptor-ligand pairs that currently exists (~ 4 x 10^{-17} M) (*31*). The extremely strong binding affinity observed in this example is not a typical result for most multivalent reagents, because identifying the optimal ligands, linker length and tether composition needed to maintain the enthalpic chelate affect can be very difficult (*32*). Therefore, many linking strategies have been attempted using small molecules, polymers, dendrimers, cyclic scaffolds, amino acids, and DNA. In the following section, these linking strategies are described in detail and a few examples of the resulting multivalent reagents.

1.4.1 Small Molecule Linkers

One of the original examples of creating multivalent molecules was conducted by Fesik and coworkers, who utilized nuclear magnetic resonance (NMR) to investigate the structure-activity relationship (SAR) between a target protein and a library of small molecules (33). Discrete binding locations of small molecules on FK506 binding protein (FKBP) were determined using two dimensional ¹⁵N- heteronuclear single-quantum (¹⁵N-HSOC) coherence shift changes. Once a lead compound was identified and optimized for a specific binding site, a second screen of the remaining library was conducted in the presence or absence of the lead compound (Figure 1.11A). Again NMR was used to identify the approximate binding location of the second ligand using the ¹⁵N- and ¹Hamide chemical shift. Based on further structural analysis, the approximate distance between the two compounds was determined and linkers were designed to span the distance without having a steric clash with the protein. Five bivalent compounds were synthesized with carbon linkers of varying length ((CH_2)_n where n=3,4,5,6) (Figure 1.11B). Although the two monomeric compounds bound FKBP with dissociation constant values of 2 μ M and 100 μ M, all of the bivalent affinity reagents designed from these monomeric units had K_{d} values in the nanomolar range. The best bivalent FKBP reagent exhibited a binding affinity of 19 nM, more than a 5,000-fold improvement over the lowest affinity lead compound. Although the starting molecules only bound with micromolar affinity, this research deomonstrated that structural information can be used to design bivalent compounds that transform low affinity ligands into high affinity binding reagents.



Figure 1.11. Structure Activity Relationship by Nuclear Magnetic Resonance. A) Schematic of the SAR by NMR method. Lead compounds are identified from a screen and optimized for binding using NMR structural information. A second compound, which binds a neighboring binding site, is identified and linked to the first compound. B) A small library of bivalent reagents with varying linker lengths are synthesized and screened for binding affinity to FK506 binding protein. In this example, an ~100 fold improvement in binding affinity is observed for the bivalent compound 10 over the monomer 2. From Shuker, S. B., et al. (1996) Science 274, 1531-1534 (*33*). Reprinted with permission from AAAS.

Since the SAR by NMR publication in 1996, structure based drug design has rapidly increased the generation of bivalent ligands and the rate of drug discovery (34-36). However, due to the time consuming nature of NMR, alternative methods have been developed for producing high affinity bivalent ligands that do not utilize structural information. Ellman and coworkers have developed a combinatorial library and linker screening technique that allows for the most productive combination of ligands and linkages to be identified in 4 steps: 1) prepare a set of potential binding elements that contain the same chemical linking moiety; 2) screen the small library for all possible molecules that bind, even weakly, to the target; 3) prepare all possible combinations of bivalent elements; and 4) screen the bivalent library for high affinity compounds (Figure 1.12A) (37). Using this strategy, 305 potential tyrsosine kinase c-Src inhibitors were screened, all of which contained an O-methyl oxime chemical linking moiety. Of the 305 small molecules examined, 47 showed greater than 70% inhibition at 500 µM c-Src protein concentration. All possible combinations of the 37 lead molecules were generated by condensing the aldehyde precursors using 5 different length O',O'- diaminoalkanediol linkers. The bivalent library was screened using ELISA, and the highest binding bivalent ligand was further characterized with 8 different linkers (Figure 1.12B). The linker length and composition made as much as a 100-fold difference in binding affinity, with the best binder possessing a K_d value of 64 nM. The monomeric lead compounds, 90 and 273, had binding affinities of $\sim 40 \ \mu$ M, however the optimal linker between the two compounds transformed the entities into a bivalent affinity reagent with greater than 600-fold improvement in binding affinity.



Figure 1.12. Combinatorial Target-Guided Ligand Assembly. A) Cartoon representation of the target-guided ligand assembly method. Monomer binding elements synthesized with the same linkage moiety are screened against the target protein. Monomer units that bind the target are linked in a pair-wise fashion, covering all possible combinations. B) Bivalent binding elements with the highest affinity are optimized for linker length and composition. Reprinted with permission from Proceedings of the National Academy of Science USA (*37*). Copyright 2000 National Academy of Science USA.

1.4.2 Polymer-Based Linkers

The literature contains numerous examples of multivalent compounds that are linked together by small molecules and we direct the reader to previously published reviews that cover this topic in great detail (38-40). An extension of small molecule based scaffolding is the use of polymers composed of a varying number of repeating structural units to produce the optimal spatial distance between ligands needed to enhance their binding affinity. This strategy was initially demonstrated by creating potent bivalent ligands to cyclic-nucleotide-gate (CNG) channels, which contain four cyclic-nucleotide binding sites (41). Kramer and Karpen developed polymer-linked ligand dimers (PLD) by covalently linking cyclic guanosine monophosphate (cGMP) to flanking ends of bifunctional vinylsulphone-derivatized polyethylene glycol (PEG), creating a barbellshaped molecule (Figure 1.13A). PLDs were constructed containing different units (n) of PEG resulting in different length bivalent molecules and were assayed for relative binding affinity to CNG. It was hypothesized that some PLDs would be too short, while others would be too long to allow each cGMP molecule to bind the target receptor synergistically (Figure 1.13B). Only when the length of the PLD was "just right" would an increase in binding affinity be observed. A series of PLDs containing varying units of PEG with lengths spanning 15-123 Å were assayed to determine the optimal polymer length to activate two different types of CNG channels; rat olfactory (OLF) and bovine rod photoreceptor (RET) (Figure 1.13C). The optimal distance for a PLD to RET was 30 Å, while enhanced affinity was observed for OLF at a wider spread of distances, with the highest affinity at a polymer length of 39 Å. The best PLD for each channel resulted in



Figure 1.13. Polymer-Linked Ligand Dimers (PLDs). A) Cyclic GMP moieties are linked through varying length units (n) of ethylene glycol. B) Illustration of PLDs of different length binding to a cell receptor with four binding sites. If the linker is too short (left) or too long (right), then both ligands are not able to bind simultaneously. C) The optimal linker length of cGMP PLDs differs for each cell receptor target; olfactory (OLF) and rod photoreceptor (RET). Reprinted by permission from Macmillan Publishers Ltd: Nature (*41*), copyright 1998.

greater than 100-fold improvement in activation of the CNG channel over the monomeric cGMP molecule.

There are many examples of PEG tethered multivalent affinity reagents due to its ease of synthesis and commercial availability (42). Schaffer et al. prepared homo- and heterodimer affinity reagents to insulin receptor using phage display selected peptides linked through PEG (43). The bivalent peptides were joined through different orientations of the terminus, N to N, N to C, C to N, and C to C, and screened for insulin receptor binding. The heterodimers outperformed the homodimers in insulin receptor binding, with the best binding PEG linked bivalent reagent coupled C-terminus to N-terminus and a binding affinity of high picomolar (pM).

Although PEG is a very simple polymer unit, it can have a dramatic influence on the binding affinity of multivalent reagents. Kim et al. demonstrated the effect of spacer length on thrombin inhibition using two DNA aptamers (44). Two aptamers that bound thrombin at non-overlapping epitopes were linked together using triethylene glycol spacers, ((CH₂)₂-O)₃, of varying unit numbers to optimize binding to thrombin (Figure 1.14A-B). When clotting time was compared, the bivalent molecule with 8 spacer units (\sim 3-4 nm) was most effective with the longest clotting time when compared to the individual aptamers (15Apt and 27 Apt), replacement of 27Apt with a random sequence (Bi'8S), or molecules with spacer units of shorter or longer length (Figure 1.14C). Kinetic binding studies revealed a dissociation rate from thrombin of 1/50 the speed for the bivalent reagent when compared to the monomeric aptamers. Based on the relative rates of dissociation, it suggests the bivalent reagent has a K_d value of 8.7 nM, while the 15Apt and 27Apt have dissociation values of 450 nM and 700 pM, respectively.

•



Figure 1.14. PEG Spacer Linked Bivalent Aptamers. Illustration of A) monovalent and B) bivalent thrombin aptamers. The two aptamers are linked through a PEG spacer of different length. C) Comparison of thrombin clotting time caused by each inhibitor indicates the optimal linker for the bivalent reagent is with 8 triethylene glycol spacers. Reprinted with permission from Proceedings of the National Academy of Science USA (44). Copyright 2008 National Academy of Science USA.

Another example of bivalent thrombin aptamers was generated using very long PEG repeating units (hexaethylene glycol, spacer 18, $(CH_2)_2$ -O- $((CH_2)_2$ -O)₄- $(CH_2)_2$ -O-PO₃- $(CH_2)_2$) (45). Interestingly, they too found that linker length was directly related to thrombin binding affinity, with the shortest linker (~24 nm) giving the best affinity. Although the aptamer sequences and linker length was different than reported in the manuscript by Kim et al. (3-4 nm), the bivalent reagent still had an improved affinity of ~100-fold over the individual aptamers. However, since the affinity of the starting aptamers was much stronger (2.9 nM and 4.0 nM), the resulting bivalent reagent had a dissociation constant of 30 pM.

Although PEG is the most highly used polymer to link ligands for multivalent reagents, other strategies such as ring opening-metathesis polymerization (ROMP) have been very successful. Kiessling and coworkers developed the ROMP method to synthesize polymers of varying density of mannose residues for increased concanavalin A (Con A) receptor clustering (46). By varying the ratio of galactose (does not bind Con A) to mannose, the binding affinity of the multivalent reagent can be tightly controlled. Experiments showed a clear relationship between an increase in the density of mannose to an increase in Con A binding, up to the point of 70% mannose density. This research demonstrated the use of multivalent affinity reagents as a tool to study receptor cluster formation and the regulation of natural multivalent interactions. A summary of all the small molecule and polymer based linking strategies described in these units is given in Table 1.1.

Linker Strategies Ligand SAR by NMR Small M	d tvne				
SAR by NMR Small M	a cy pe	Ligand Discovery Method	Affinity	Number of Examples	Reference
	folecule	Structural Analysis	19 nM- 350 µM	100s	33, 36
Combinatorial Target- Small M Guided Assembly	Aolecule	Small Library Screen	60 nM- 250 nM	2	34, 37
Polymer-Linked Ligand Small M Dimer	folecule	Naturally Occurring	4 nM	1	41
PEG Linked Bivalent Pept Peptides	tide	Phage Display	20 pM	1	43
PEG Linked Aptamer DNA A	Aptamer	SELEX	1 nM	2	44, 45
Olefin Metathesis Small M	Iolecule	Naturally Occurring	92 µM - 290µM	100s	46, 47

1.4.3 Dendrimer Scaffolds

Dendrimers are large molecules composed of multiple polymers linked to a central core molecule. The polymeric units and central compound can be anything from sugar moieties to amino acids (49-51). Dendrimers are of great interest due to their versatility of application (52). A recent publication described the transformation of weak affinity collagen-binding peptides discovered through phage display into high affinity binding reagents using dendritic wedges (Figure 1.15A) (53). This process was cleverly coined dendrimer display. Two rounds of phage display panning were conducted to discover a new consensus sequence that differed from all previously reported collagen binding peptides. The peptide was combined with a pentavalent dendritic platform to display 5 copies of the collagen binding peptide (Figure 1.15B, where R_2 =NO₂, biotin or fluorescein, R_3 =peptide, and R_4 =H). While the monomer peptide bound collagen with a dissociation constant of 61 μ M, the dendrimer displaying 5 copies of the peptide bound with over 100-fold stronger affinity with K_d value of 550 nM (Figure 1.15C).

One of the most well known type of peptide based dendrimers is the multiple antigen peptide (MAP) (51). MAPs are usually composed of lysine residues branched through the α and ε amine moieties on the peptide backbone and side chain, respectively. A variety of MAP valencies can be easily engineered by altering the number of lysine residues (Figure 1.16A-C). Tam et al. generated a three lysine MAP (K₂K) displaying short arginine-leucine-tyrosine-arginine peptides for enhanced antimicrobial activity (Figure 1.16D) (54). The combination of α and ε connections resulted in a ε peptide backbone that spans 21 atoms. Although the tetrapeptide MAP did not have any antimicrobial activity, a dimeric tetrapeptide MAP, formed through the carboxy terminus and displaying 8 peptides, had antimicorbial activity against 4 gram-positive, 3 gramnegative and 3 fungi with low micromolar affinity. Additionally, the MAPs that displayed longer peptide sequences performed better than those with shorter sequences, suggesting that small peptides may get buried in the MAP core structure.



Figure 1.15. Dendrimer Display. A) Peptide ligands identified to bind collagen through 3 rounds of phage display are used to create a multivalent dendrimer. B) The pentavalent dendritic wedge used to display the peptide (\mathbf{R}) discovered from phage display. C) Binding analysis revealed the pentavalent molecule has a 100-fold improvement in affinity over the monovalent ligand. Reprinted (adapted in part or whole) with permission from Journal of the American Chemical Society (*53*). Copyright 2009 American Chemical Society.



Figure 1.16. Multiple Antigen Peptide (MAP). The α and ε amine of the amide backbone and side chain, respectively, of a lysine residue is used to branch the MAP compound. A-C) Depending on the number of lysine residues, difference branching densities will occur. D) Chemical structure of a K₂K Map bearing Arg-Leu-Tyr-Arg peptides. Reprinted with permission from European Journal of Biochemistry (*54*). Copyright 2007 John Wiley & Sons.

1.4.4 Cyclic Core Scaffolds

The dendrimeric scaffolds are well suited for the creation of multivalent affinity reagents, however the resulting molecules are relatively flat and linear compounds. Alternatively, ligands can be linked to a cyclic core scaffold, which can increase the 3-dimensional spatial separation. One example of this general approach uses a central β -galactose residue to connect 5 dimeric trisaccharide ligands (Figure 1.17A) (55). The resulting molecule was designed to target Shiga-like toxin (SLT), and contains 15 binding sites arranged symmetrically around the surface. A model of the multivalent molecule bound to the SLT suggests possible binding sites between the ligands and the protein (Figure 1.17B). The binding interaction positions the 5 arms of the multivalent reagent in a manner that resembles a starfish, therefore the molecule was appropriately termed STARFISH. The STARFISH molecule exhibited more than 1-million fold increase in inhibition of STL over the trisaccharide **R** in figure 1.17A.

Another heavily utilized sugar based cyclic scaffold is cyclodextrin (CD) (56). A cyclic dextrin compound baring mannose moieties was designed to bind Con A, as previously targeted by the ROMP method (57). β -CD scaffolds were synthesized displaying different densities of the glycoligand (Figure 1.18A). As suspected, an increase in valency showed an increase in binding affinity to Con A, with the strongest affinity CD molecule possessing the most mannose sugars (Figure 1.18B). The β -CD design has many advantages including high water solubility and the ability to dimerize into a cage like structure for drug delivery (Figure 1.18C).



Figure 1.17. Shiga-Like Toxin I Inhibitor, STARFISH. A) The core of the scaffold is composed of β -galactose (3) and is linked to 5 dimeric trisaccarides (**R**), creating a pentavalent compound. B) A model of the resulting five-ray cluster resembles a starfish and is named accordingly, STARFISH. The STARFISH has greater than 1 million fold higher inhibitory activity over the univalent ligands. Reprinted by permission from Macmillan Publishers Ltd: Nature (*55*), copyright 2000.



Figure 1.18. Dendritic β -Cyclodextrin Derivatives. A) β -cyclodextrin structures can be made with varying number of dextrin moieties and saccharide substitutions to make linear and branched multivalent reagents. B) Binding inhibition of concanavalin A demonstrates the affect of valency on binding with the lowest IC₅₀ values obtained with the highest valency compounds. C) A dimer of cyclodextrin-glycodendrimer conjugates can be used as a drug delivery system, with the cyclodextrin dimer as the drug carrier and the glycodendrimer as the cell-targeting reagent. Reprinted (adapted in part or whole) with permission from Journal of the American Chemical Society (*57*). Copyright 2004 American Chemical Society. Cyclic molecular core scaffolds used to create multivalent reagents have been designed as simple as modified benzene (58) and adamantine (59) rings to as complex as acetylated pentacyclen (60). Fan et al. used the acetylated pentacyclen ring to present D-galactose at a 5-fold symmetry that matched the binding sites of heat labile enterotoxin from Escherichia coli. A small library of multivalent reagents was generated using this design with varying linker lengths between the pentacyclen and galactose ligands. The best pentavalent ligand exhibited an IC₅₀ value of 560 nM, 100,000-fold better than the monomeric galactose unit, but not as strong as the oligosaccharide portion of the natural heat labile enterotoxin, GM1-OS (IC₅₀= 10 nM).

Similar to small molecular linkages, not every cyclic scaffold is suitable for all target applications. When the same group, Fan and Hol (above), tried to target cholera toxin, multiple cyclic macromolecules were screened to identify the optimal scaffold necessary for a high affinity binding reagent (*61*). They designed four cyclic scaffolds, 3 decapeptides and 1 pentacyclen, each conjugated to 5 galactose-based ligands with varying linker lengths. For cholera toxin, the pentacyclen core was not the ideal scaffold. Instead the decapeptide consisting of repeating lysine- γ -aminobutyric acid units with a 2-unit linker length between the core and the galactose moiety displayed the best IC₅₀ value of low micromolar. This multivalent reagent again had over 100,000-fold enhancement in affinity over the monomeric galactose unit.

Similarly, Trouche et al. needed to systematically examine multiple cyclic scaffold types and tether lengths in order to find the optimal CD40 multivalent binding reagent (62). A homotrimer reagent was developed based on the natural ligand to CD40

(CD40L) (Figure 1.19A). The multivalent reagent was composed of a cyclic scaffold and three CD40 ligands spaced by a linker. Nine different scaffold types were first screened with simple lysine-glycine-tyrosine-tyrosine-amino hexanoic acid sequences to identify the optimal scaffold (Figure 1.19B). One of the smaller, cyclic scaffolds (**B**) was determined to be most favorable, and was used to screen different linker lengths (Figure 1.19C). The linkers were composed of different carbon chain lengths and amino acid isomers. The best binding molecule as determined by surface plasmon resonance is **B-1** with a pentane carbon linker (Figure 1.19D). Furthermore, Trouche et al. tested different ligand conformations and found the optimal multivalent CD40L mimick (**B-14**) exhibited a dissociation constant of 2.8 nM, more than 40-fold over the **B-1** compound. A summary of this, and the above mentioned dendrimer linkers and cyclic scaffolds, is given in Table 1.2.



Figure 1.19. Multivalent Homotrimer CD40L Mimicks. A) A trimeric compound was designed based on the X-ray structure of CD40L homotrimer (PDB: 1ALY). The timeric compound consists of a core, three linkers and three CD40 binding motifs. B) Multiple core structures were assayed for optimal binding affinity. C) Core structure **B** indicated the best scaffold for CD40 binding and the linker length was optimized for this scaffold. D) Compound **B-14** resulted in the highest percentage of apoptosis with a K_d of 2.8 nM. Reprinted (adapted in part or whole) with permission from Journal of the American Chemical Society (*62*). Copyright 2007 American Chemical Society.

er Strategies	Ligand type	Ligand Discovery	Affinity	Number of	Reference
		Method		Examples	
rimer Display	Peptide	Phage Display	550 nM	1	53
: Lysine MAP	Peptide	Naturally Occurring and Phage Dismlay	48 µM- 3.9 mM	100s	54, 63
ve Dimeric	Sugar Moiety	Naturally Occurring	1.2 µM	1	55
haride Ligands)				
odendrimer- clodextrin	Sugar Moiety	Naturally Occurring	30 µM	1	57
ed Pentacyclen Ring	Sugar Moiety	Naturally Occurring	560-900 nM	2	60, 64
ltiple Cyclic	Sugar Moiety	Naturally Occurring	1 µM	1	61
romolecules					
ltiple Cyclic	Peptide	Small Library Screen	2.8 nM	1	62
romolecules					

1.4.5 Amino Acid Linkers

Of all the tethering molecules used to combine ligands as multivalent compounds, amino acid linkers are the most common. More specifically, the most common amino acid linker designs consist of a combination of glycine and serine residues due to their flexibility, limited structural interference, and protease resistance (65). When linking peptides, proteins or immunoglobulins, the incorporation of amino acid tethers can be as simple as adding a few residues during solid phase peptide synthesis or by engineering the DNA sequence to contain the linker sequence before expression.

In the last decade, linking naturally monomeric immunoglobulins (such as IgD, IgE, and IgG) into multimeric units has gained much attention. Antibodies are well known as the "gold standard" for protein affinity reagents with high levels of specificity and affinity. However, antibodies are limited in their use due to structural instability, mutational changes over time, and the use of animals for their development. To overcome these limitations, researchers have begun using only the target binding region (TBR) of the antibody displayed on the variable region (Fv) of the heavy chain and light chain of the antibody (Figure 1.20A). By joining the Fv together, single chain antibody fragments (scFv) are created with more stability and facile bacterial expression, which eliminates the need for animal immunization. Furthermore, linking multiple scFvs together can generate homovalent or heterovalent reagents termed diabodies (two fragments), triabodies (three fragments), or tetrabodies (four fragments) (Figure 1.20B). The homovalent versions have shown to enhance affinity over the monomeric unit,



Figure 1.20. Engineered Antibody Comformations. A) A traditional IgG antibody consists of 2 heavy chains and 2 light chains each of which contain variable antigen binding regions. Different fragments of each chain can be used as monomeric or multimeric untis to make affinity reagents. B) Depending on the length of the amino acid linker connecting the fragments, bivalent (diabodies), trivalent (triabodies), and tetravalent (tetrabodies) antibody fragments can be created. Reprinted from Biomolecular Engineering, 18, Kortt, A. A., Dolezal, O., Power, B. E., and Hudson, P. J. Author(s), Dimeric and trimeric antibodies: high avidity scFvs for cancer targeting, 95-108, Copyright 2001 with permission from Elsevier (66).

while the heterovalent reagents have the ability to bind two different targets in a bispecific manner (66, 67).

For these types of molecules, the linker length, orientation and composition plays a very important role in determining the valency of the multivalent scFv (68). A linker consisting of 3 to 12 residues results in two scFv molecules associated together in a diabody protein (69). If the linker is reduced to less than 3 residues, it forces the scFv to associate into a triabody, tetrabody, or multimers depending on linker orientation and composition (Figure 1.20B).

One of the initial examples of linking multiple antibody TBRs together to create a bispecific molecule was conducted by Neri et al. and was described as chelating recombinant antibodies (CRAbs) (70). This work describes the combination of two antibody fragments that recognize non-overlapping sites of the target, hen egg lysozyme (HEL). The TBRs of two anti-HEL antibodies were conjugated through short 15 amino acid hydrophobic linkers between the heavy and light chains and 18 amino acid linkers between the two light chains. The engineered construct was expressed in Escherichia coli and assayed for binding affinity to HEL. The CRAb construct bound HEL with greater than 20-fold enhancement over the individual fragments.

Another version of the amino acid linking strategy has been demonstrated by tethering multiple exon shuffled A-domains together that bind non-overlapping sites on a



^aAvimer sequences provided in **Supplementary Data** online, ^bELISA binding assay. ^cAlphaquest competition assay. ^dA549 soft-agar colony formation assays and cell-scatter assays. ^eBiacore (SPR) binding assay. ^tWhole PBMC fraction proliferation assays with MLR, anti-CD3 or staphylococcal enterotoxin 8 stimulation. ^BIL-6 stimulation of TF-1 cell proliferation. ^bNot determined.

Figure 1.21. Avidity Multimers: Avimers. A) Using exon shuffling and phage display, multiple protein binding domains to a given target can be linked together to enhance the affinity. B) The binding affinities of the multidomain avimers can rival that of traditional antibodies with picomolar or lower IC₅₀ values. Reprinted by permission from Macmillan Publishers Ltd: Nature Biotechnology (71), copyright 2005.

single target (Figure 1.21A) (71). Out of the ~35 amino acids that comprise the A domain, 28 residues are not conserved, which is similar to the 25-30 amino acid hypervariable region of an antibody. The optimized A domain units are spaced by ~5 amino acids and termed avimers. Five avimer molecules were independently generated to 5 unique target proteins and all exhibited sub-nanomolar affinities (Figure 1.21B). The strongest affinity avimer contained 3 A-domains and was developed to interluekin 6. This avimer exhibited an IC₅₀ value of high femtomolar to low picomolar and showed strong specificity against a related protein, interleukin 1.

Similarly, two or more polypeptides can be linked together using a short string of amino acids (72, 73). However, determining the optimal amino acid linker and ligand pair may be a time consuming process if done as individual steps. To streamline this process, Kodadek and coworkers developed a method of identifying both the linker and ligand components simultaneously in a single assay (74). This strategy was independently demonstrated on two protein targets, Mdm2 and ubiquitin. A library of short polypeptoid and polypeptide polymers containing ~7 amino acids were generated by split pool synthesis on densely packed beads and capped with previously identified ligands to Mdm2 and ubiquitin, respectively. The library was screened against the respective targets and polymers presented on the beads that bound each target were sequenced by Edman degradation. The resulting bivalent peptoid reagent bound Mdm2 with over 4,000-fold affinity enhancement compared to the monovalent ligands, while the bivalent peptide bound ubiquitin with only 3-fold affinity improvement in solution. It is hypothesized that the bivalent peptide bound ubiquitin tightly when immobilized on a surface because it

allowed adjacent ligands to bind collaboratively in an avidity-based binding event while the Mdm2 reagent bound in a truly bivalent, synergistic fashion.

Additionally, a combination of linker types may be necessary in order to identify the optimal spatial distance between ligands. Handl et al. used a combination of amino acid and PEG linkers (termed PEGO in this paper) to optimize a bivalent peptide ligand for binding to the human Melanocortin Receptor 4 (hMC4R) (75). Bivalent reagents were design to bind the hMC4R homodimer by linking two copies of a truncated heptapeptide version of the naturally occurring melanocyte stimulating hormone (MSH). Sixteen different compounds were synthesized with different combinations of PEGO and prolineglycine units (Figure 1.22A). On average, the bivalent reagents that contained both PEGO and amino acid linkers performed only slightly better than those with only amino acid units (Figure 1.22B). The reagent with the best binding affinity (11 nM) to hMC4R contained 2 units of (proline-glycine)₃, with or without the addition of PEGO (Figure 1.22C).

Beyond simple amino acid linkages, Nunn and coworkers developed high affinity peptide (HAP) binders to vascular endothelial growth factor receptor (VEGF-2) and c-Met by combining biotinylated ligands using avidin as the scaffold (76). Peptide ligands discovered from phage display panning were synthesized with a biotin moiety and incubated with avidin to create homo- and heterovalent tetrameric affinity reagents. For both protein targets, the heterovalent affinity reagents resulted in stronger binding over the homovalent agents. The highest affinity hetero-tetravalent HAP binder generated to VEGF-2 exhibited a K_d of 500 pM, greater than 10-fold improvement over the

homovalent reagents and more than 40-fold over the individual peptides. The heterodimer c-Met HAP binder exhibited a dissociation constant of 800 pM, which is more than a 1,000-fold enhancement in binding affinity over the weakest affinity peptide used to generate the HAP. For an overview of this linking strategy and the previously mentioned amino acid tethers, see Table 1.3



Figure 1.22. Multivalent Peptide Ligands to Human Melanocortin Receptor 4. A) A small library of homobivalent α -melanocyte stimulating hormone peptides were linked together using a combination of proline-glycine and PEGO linkers. B) The bivalent ligands with PEGO inserts (\blacktriangle) did not perform much better than those with only proline-glycine linkers (\blacksquare). C) The highest affinity ligands contained 2 (proline-glycine)₃ linker, with or without PEGO inserts, the resulting molecules bound with low nanomolar affinity. Reprinted (adapted in part or whole) with permission from Bioconjugate Chemistry (75). Copyright 2007 American Chemical Society.

Linker Strategies	Ligand type	Ligand Discovery Method	Affinity	Number of Examples	Reference
Diabodies, Triabodies,	Antibody	Animal Immunization	1 nM-133nM	100s	67, 77
Teterabodies Chelating Recombinant	Fragments Antibody	Animal Immunization	1 nM	2	70, 78
Antibodies	Fragments	and Phage Display			
Avimer Exon Shuffled A-	A-Domain	Exon Shuffling and	100-200 pM	5	IL
Domains	Protein	Phage Display			
Peptide	Peptide or	Microarray Screen	1.3-12 µM	2	74
Extension	Peptoid				
PEG and Amino Acid	Peptides	Naturally Occurring	11 nM		75
Linked Peptide Ligands	I				
Avidin Linked Peptides	Peptides	Phage Display	500-800 pM	2	76

1.4.6 Deoxyribonucleic Acid Linkers

The commercial availability of standard and modified synthetic deoxyribonucleic acid (DNA) sequences for relatively inexpensive (0.40 per base) has made DNA a logical linker molecule for the development of multivalent affinity reagents. Spatial distances can be easily tailored when using DNA due to its predictable hybridization through Watson and Crick base pairing, and geometry of 3.4 Å distance between base pairs and width of ~2 nm for the dsDNA molecule. An additional feature that makes DNA exceptionally useful as a tethering molecule is its flexible nature when single stranded but extreme rigidity when double stranded, with a persistence length of at least 50 nm (79).

Similar to the bispecific diabodies, linking two different antibodies together using dsDNA can produce bispecific reagents with dual binding function. Chaudri et al. independently linked complimentary strands of DNA to anti-human IgG and anti-mouse IgG antibodies (79). When assayed, the bivalent reagent exhibited the ability to bind both subtypes of antibodies, while the individual monomeric unit could only detect their respective subtype. Additionally, this DNA scaffold design is able to bring two cell types together in close proximity by first incubating each conjugated antibody with their respective cell type, then combining the two and allowing for the hybridization between the complimentary strands of DNA.

Using modified DNA, Sprinz et al. developed a novel screening method of identifying non-competing ligands through oligonucleotide-linked organic fragments (80). In this procedure, each molecule in a small library is linked to a unique sequence of

modified DNA that is partially unique and partially complimentary to the other sequences in the library allowing hybridization into bivalent elements (Figure 1.23A-B). These reagents are then screened against the target of interest and those that bind with high affinity are recovered (Figure 1.23C-D). The species in the bound fraction are identified using specific primers to the unique region of the DNA during PCR (Figure 1.23F). A single example of this method was performed against streptavidin using 6 oligonucleotide-linked organic fragments. Only the oligonucleotides conjugated to biotin were observed in the recovered fraction and amplified by PCR. A model of the bivalent biotin-DNA complex bound to streptavidin is depicted in Figure 1.23G.

Baird and coworkers investigated the role of dsDNA rigidity and separation distance of bivalent dinitrophenyl (DNP) molecules in the regulation of IgE-Fc ϵ RI induced signal transduction (*81*). Toward this aim, DNP was modified on the 5' end of varying lengths of complimentary strands of DNA. The small library of bivalent DNP molecules linked through dsDNA were titrated against anti-DNP IgE incubated with Fc ϵ RI free in solution or bound to cells. No difference was observed between the bivalent and monovalent molecules when the assay was conducted in solution, however when conducted against cells, the bivalent reagents showed 2-fold enhancement in affinity. A surprising result was observed when the molecules were tested for their ability to trigger cell degranulation. A length dependent trend was observed, with the shortest linkers outperforming the longer linkers. Only when a flexible region (O-(CH₂)₃-OPO₂)₅ was added to the longest linker, did the bivalent reagent have any effect on cell degranulation.



Figure 1.23. Oligonucleotide-Linked Organic Fragments. A) General strategy of identifying bivalent organic fragments through an oligonucleotide-linkage. B) Illustration of the bidentate biotin molecule binding two different binding sites on streptavidin (PDB code for streptavidin: ISWG, NDB for duplex DNA: BDL084). Reprinted from Bioorganic and Medicinal Chemistry Letters, 15, Sprinz, K. I., Tagore, D. M., and Hamilton, A. D., Self-assembly of bivalent protein-binding agents based on oligonucleotide-linked organic fragments, 3908-3911, Copyright 2005, with permission from Elsevier (*80*).

Furthermore, the Baird group designed trivalent DNP molecules using a "Y" shaped dsDNA linker to further study the length dependence of tyrosine phosphorylation of Fc ϵ RI (82). Similar to the bivalent system, the shorter dsDNA linkers (~5 nm) activated the cellular cascade more effectively than the longer scaffolds (~12 nm). These dsDNA scaffold systems provide researchers a systematical approach to study receptor-mediated signaling systems.

Another logical use of the DNA linking system is to connect two DNA aptamers, which can easily be synthesized as a single string of nucleic acids. Mayer and coworkers demonstrated this method by linking two non-overlapping thrombin binding DNA aptamers with two different length poly deoxyadenosine (dA) linkers, 5 and 15 dA (Figure 1.24) (83). When tested for thrombin binding, both constructs showed ~2-fold enhanced binding affinity over the individual aptamers. However, when the constructs were assayed for thrombin inhibition of peptide cleavage, the bivalent reagent linked by 15dA inhibited with greater than 10-fold activity over the monomeric aptamer. In further studies, this design was successfully used as a caging molecule with controllable release for thrombin inhibition (84).

Rinker et al. took this approach one step further by using a DNA nanostructure as a scaffold for a bivalent DNA aptamer against thrombin (*85*). In this work, multi-helix DNA tiles (4- and 5-helix bundles) were designed to spatially separate two DNA aptamers that bound non-overlapping sites on thrombin (Figure 1.25A). The DNA aptamers were placed at different positions at the end of the 4- or 5-helix bundle creating 4 distances between the aptamers; 2, 3.5, 5.3, and 6.9 nm. The range efficiently spanned


Figure 1.24. DNA Linked Bivalent Thrombin Inhibitor. A) Crystal structure of thrombin with exosite I (red) and exosite II (blue) highlighted (PDB code: 1THR). B) Bivalent aptamer design with a poly adenosine linker (orange). C) Multiple bivalent constructs were generated with different lengths of adenosine linkers (*83*). Reprinted with permission from ChemBioChem. Copyright 2007 John Wiley & Sons.



Figure 1.25. Self-Assembling Bivalent Aptamers Using a DNA Tile. A) A schematic illustrating the DNA tile design to specifically space the two aptamers by different distances. B) Gel mobility shift assay results indicate the optimal distance for thrombin binding to the bivalent aptamer-tile is 5.3 nanometers. The grey and striped bars represent 4-helix and 5-helix tiles, respectively. Reprinted by permission from Macmillan Publishers Ltd: Nature Nanotechnology (*85*), copyright 2008.

the 4.1 x 4.1 nm dimensional of thrombin. Native gel mobility shift assay was conducted on each construct, and the tile with aptamers separated by 5.3 nm bound the highest percentage of thrombin (Figure 1.25B). It was estimated that the 5.3 nm separated bivalent aptamer tile had \sim 10-fold binding enhancement over the individual aptamers. This result is very consistent with the affinity enhancement obtained by Mayer and coworkers.

In addition to these uses of DNA as a linking tool, our laboratory has developed a completely novel method of creating bivalent affinity reagents using a synthetic DNA scaffold through a combinatorial examination of linkers and ligands (CELL) (86). In this method, non-overlapping ligands to a given protein target are covalently linked at different positions within complementary strands of DNA (Figure 1.26A), resulting in a small library of bivalent reagents, termed synbodies, with different combinations of ligands and spatial separations between ligand pairs (Figure 1.26B). This bivalent library is then screened for relative binding affinity, and the synthetic antibody with the best binding response is further characterized (Figure 1.26C). An example of a synbody developed through the CELL process is shown in figure 1.26D. Two peptides, BP1 and BP2, were combined on the DNA scaffold with the optimal orientation and spatial separation creating a synbody to Gal80 with ~1,000-fold improvement in binding affinity over the individual ligands. Multiple examples of the CELL process have been demonstrated and complete experiment detail and results are given in chapter 4. Additionally, Table 1.4 compares all DNA linking strategies presented in this section.



Figure 1.26. Combinatorial Examination of Linker and Ligands (CELL). A) Bivalent affinity reagents, termed synbodies, are created by systematically spacing peptide ligands on a dsDNA scaffold. B) The optimal peptide ligand pair and separation distance is screened for binding affinity. C) A small library of synbodies are screened for affinity against Gal80 protein. The optimal peptide orientation and separation distance was observed for forward SC-13. D) The affinity of the synbody to Gal80 is ~1,000-fold improved over the individual peptide ligands. Reprinted with permission from Journal of the American Chemical Society (*86*). Copyright 2009 American Chemical Society.

Ligand type Ligand Discovery Attinity Number of Examples Reference od Small Molecule Small Library Screen Not Determined 1 80 oed Dintitrophenyl Naturally Occurring 1.5-3 nM 2 81, 82 oed DNA Aptamer SELEX 4.3 nM 1 83 DNA Aptamer SELEX ND 1 85 Peptide Microarray Screen 3-87 nM 2 86
cdSmall MoleculeSmall Library ScreenNot Determined180cdDintitrophenylNaturally Occurring1.5-3 nM281, 82cdDNA AptamerSELEX4.3 nM183DNA AptamerSELEXND185PeptideMicroarray Screen3-87 nM286
cedDintitrophenylNaturally Occurring1.5-3 nM281, 82cedDNA AptamerSELEX4.3 nM183DNA AptamerSELEXND185PeptideMicroarray Screen3-87 nM286Is
ced DNA Aptamer SELEX 4.3 nM 1 83 DNA Aptamer SELEX ND 1 85 Peptide Microarray Screen 3-87 nM 2 86
DNA Aptamer SELEX ND 1 85 Peptide Microarray Screen 3-87 nM 2 86
Peptide Microarray Screen 3-87 nM 2 86 rs

Table 1.4. Deoxyribosenucleic Acid Linkers

Currently, the discovery of new proteins is outpacing the development rate of protein affinity reagents. This bottleneck has created a need for high-throughput methods of creating affinity reagents to the proteome. One such approach to fill this need is to link multiple copies of weak affinity ligands together, creating a single multivalent reagent with high binding affinity needed to characterize proteins. Reviewed in this chapter are just a few of the many strategies that exist to generate multivalent reagents. Each strategy has been successful in creating at least one protein affinity reagent, however, they may not be amendable to other protein targets. Because of the strengths and weakness of each strategy (summarized in Table 1.5), multiple strategies may need to be employed in order to create the ideal multivalent reagent for a given target protein and application.

	Weaknesses	Requires 3D protein structure	Requires extensive organic synthesis Limited to small proteins	Requires organic synthesis Requires extensive organic synthesis	Requires minor organic synthesis	Limited to known aptamers	Not a systematic strategy Requires organic synthesis Limited control on composition and density
f Various Linker Strategies	Strengths	Structural method	Systematic strategy	No 3D structure needed Potentially high-throughput Systematic strategy	No 3D structure needed Systematic strategy	No 3D structure needed Systematic strategy Uses commercial reagents	No 3D structure needed
Table 1.5. Strengths and Weaknesses o	Linker Strategies	SAR by NMR		Combinatorial Target-Guided Assembly	Polymer-Linked Ligand Dimers	PEG Linked Aptamers	Olefin Metathesis

65

.

Dendrimer Display No 3D structure needed Requires phag Not syster Not syster Not syster Not syster Five Dimeric Trisaccharide Structural method Requires all prot Five Dimeric Trisaccharide Structural method Requires stressive on Ijveodendrimer-Cyclodextrin No 3D structure needed Requires extensive on Systematic strategy Requires extensive on Systematic strategy Acylated Pentacyclen Ring Structural method Requires 3D prot Acylated Pentacyclen Ring No 3D structure needed Requires 3D prot Acylated Pentacyclen Ring Structural method Requires attensive on Intiple Cyclic Macromolecules No 3D structure needed Requires attensive on Systematic strategy Diabodies, Triabodies, No 3D structure needed Requires animal in	e 1.5. Continued Linker Strategies	Strengths	Weaknesses
Dimeric Trisaccharide Structural method Requires minor org Ligands Structural method Requires 3D prot Usiands Designed composition and density Requires extensive or endrimer-Cyclodextrin No 3D structure needed Requires extensive or Systematic strategy Requires extensive or Systematic strategy Designed composition and density Requires 3D prot Red Structural method Requires 3D prot Red Structural method Requires animal in Cyclic Macromolecules No 3D structure needed Requires extensive or Systematic strategy Designed composition and density Requires extensive or Red Structure needed Requires extensive or Cyclic Macromolecules No 3D structure needed Requires animal in Dodies, Triabodies, No 3D structure needed Requires animal in	endrimer Display	No 3D structure needed	Requires phage display Not systematic
Dimeric TrisaccharideStructural methodRequires 3D protLigandsDesigned composition and densityRequires extensive ondendrimer-CyclodextrinNo 3D structure neededRequires extensive onSystematic strategySystematic strategyRequires extensive ondendrimer-Cyclone RingStructural methodRequires extensive onsted Pentacyclen RingStructural methodRequires attensive oncyclic MacromoleculesNo 3D structure neededRequires extensive onsodies, Triabodies,No 3D structure neededRequires minor orgTeterahodiesNo 3D structure neededRequires animal inTeterahodiesNo 3D structure neededRequires animal in			Requires minor organic synthesis
dendrimer-CyclodextrinNo 3D structure neededRequires extensive on Systematic strategySystematic strategySystematic strategyDesigned composition and densityRequires 3D protated Pentacyclen RingStructural methodRequires 3D protc Cyclic MacromoleculesNo 3D structure neededRequires minor orgSystematic strategyNo 3D structure neededRequires minor orgabodies, Triabodies, Triabodies, Triabodies, Triabodies, Triabodies, TotalNo 3D structure neededRequires animal in	Dimeric Trisaccharide Ligands	Structural method Designed composition and density	Requires 3D protein structure Requires extensive organic synthesis
ated Pentacyclen Ring Structural method Requires 3D prote Requires extensive of Requires extensive of Requires extensive of Systematic strategy Systematic strategy No 3D structure needed Requires animal in Teterahodies, Triabodies, No 3D structure needed Requires animal in Teterahodies	dendrimer-Cyclodextrin	No 3D structure needed Systematic strategy Designed composition and density	Requires extensive organic synthesis
e Cyclic Macromolecules No 3D structure needed Requires minor org Systematic strategy abodies, Triabodies, No 3D structure needed Requires animal in Teterahodies	ated Pentacyclen Ring	Structural method	Requires 3D protein structure Requires extensive organic synthesis
abodies, Triabodies, No 3D structure needed Requires animal in Teterahodies	e Cyclic Macromolecules	No 3D structure needed Systematic strategy	Requires minor organic synthesis
Systematic strategy Generated i	abodies, Triabodies, Teterabodies	No 3D structure needed Systematic strategy	Requires animal immunization Generated <i>in vitro</i>

.

66

	Weaknesses	Requires animal immunization Generated <i>in vitro</i>	Requires phage display Generated <i>in vitro</i>	Requires a peptide synthesizer Requires a lead ligand	Required specialty solid-phase synthesis	Limited control on spacial orientation and linker length	Requires small molecules compatible with DNA conjugation Limited control on linker length
	Strengths	No 3D structure needed	No 3D structure needed, systematic strategy	No 3D structure needed Uses commercial reagents Systematic strategy	No 3D structure needed	No 3D structure needed Uses commercial reagents	Potentially high-throughput No 3D structure needed Systematic strategy Uses commercial reagents
1 adie 1.3. Continuea	Linker Strategies	Chelating Recombinant Antibodies	Avimer Exon Shuffled A-Domains	Peptide Extension	PEG and Amino Acid Linked Peptide Ligands	Avidin Linked Peptides	Oligonucleotide Linked Organic Fragments

Table 1.5. Conti

67

Linker Strategies	Strengths	Weaknesses
DNP-dsDNA	No 3D structure needed	Limited to DNP binding proteins
	Uses commercial reagents	Limited control on linker length
		Not systematic
Bivalent Aptamers Spaced by	No 3D structure needed	Limited to known aptamers
ssDNA	Uses commercial reagents	Not systematic
DNA Array Based Bivalent	No 3D structure needed	Limited to known aptamers
Aptamer	Uses commercial reagents	
	Systematic strategy	
Combinatorial Examination of	Uses commercial reagents,	Requires eptitope mapping
Linkers and Ligands	No 3D structure needed	
	Potentially high-throughput	
	Systematic strategy	

Table 1.5. Continued

1.5 References

- 1. Endo, M., and Sugiyama, H. (2009) Chemical approaches to DNA nanotechnology, *ChemBioChem 10*, 2420-2443.
- 2. Carell, T., Behrens, C., and Gierlich, J. (2003) Electrontransfer through DNA and metal-containing DNA, *Org. Biomol. Chem.* 1, 2221-2228.
- 3. Cheng, A. K. H., Sen, D., and Yu, H.-Z. (2009) Design and testing of aptamerbased electrochemical biosensors for proteins and small molecules, *Bioelectrochem.* 77, 1-12.
- 4. Seeman, N. C. (1982) Nucleic acid junctions and lattices, J. Theor. Biol. 99, 237-247.
- 5. Aldaye, F. A., Palmer, A. L., and Sleiman, H. F. (2008) Assembling materials with DNA as the guide, *Science 321*, 1795-1799.
- 6. Park, S. H., Yin, P., Liu, Y., Reif, J. H., LaBean, T. H., and Yan, H. (2005) Programmable DNA self-assemblies for nanoscale organization of ligands and proteins, *Nano Lett.* 5, 729-733.
- 7. Liu, Y., Lin, C. X., Li, H. Y., and Yan, H. (2005) Aptamer-directed self-assembly of protein arrays on a DNA nanostructure, *Angew. Chem. Int. Ed.* 44, 4333-4338.
- 8. Bunka, D. H. J., and Stockley, P. G. (2006) Aptamers come of age- at last, *Nat. Rev. Micro.* 4, 588-596.
- 9. Lee, J. F., Hesselberth, J. R., Meyers, L. A., and Ellington, A. D. (2004) Aptamer Database, *Nucleic Acids Res.* 32, D95-100.
- 10. Singh, Y., Spinelli, N., and Defrancq, E. (2008) Chemical strategies for oligonucleotide-conjugates synthesis, *Curr. Org. Chem.* 12, 263-290.
- 11. Williams, B. A. R., Lund, K., Liu, Y., Yan, H., and Chaput, J. C. (2007) Selfassembled peptide nanoarrays: An approach to studying protein-protein interactions, *Angew. Chem. Int. Ed.* 46, 3051-3054.

- 12. Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands, *Nature 346*, 818-822.
- 13. Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science 249*, 505-510.
- 14. Baugh, C., Grate, D. r., and Wilson, C. (2000) 2.8 Å crystal structure of the malachite green aptamer, *J. Mol. Biol.* 301, 117-128.
- 15. Jenison, R. D., Gill, S. C., Pardi, A., and Polisky, B. (1994) High-resolution molecular discrimination by RNA, *Science 263*, 1425-1429.
- 16. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications, *Nat. Biotechnol. 21*, 255-261.
- 17. Shrogen-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 acetyation controls chromatin structure and protein interactions, *Science 311*, 844-847.
- 18. Harp, J. M., Hanson, B. L., Timm, D. E., and Bunick, G. J. (2000) Asymmetries in the nucleosome core particle at 2.5 A resolution, *Acta Crystall. Section D 56*, 1513-1534.
- 19. Mairal, T., Ozalp, V. C., Sanchez, P. L., Mir, M., Katakis, I., and O'Sullivan, C. K. (2008) Aptamers: molecular tools for analytical applications, *Anal. Bioanal. Chem.* 390, 989-1007.
- 20. Gilbert, B. A., Sha, M., Wathen, S. T., and Rando, R. R. (1997) RNA aptamers that specifically bind to a K Ras-derived farnesylated peptide, *Bio. Med. Chem. 5*, 1115-1122.
- 21. Williams, B. A. R., Lin, L., Lindsay, S. M., and Chaput, J. C. (2009) Evolution of a histone H4-K16 acetyl-specific DNA aptamer, *J. Am. Chem. Soc. 131*, 6330-6331.

- 22. Hu, S., Loo, J. A., and Wong, D. T. (2006) Human body fluid proteome analysis, *Proteomics 6*, 6326-6353.
- 23. Butler, J. E., Ni, L., Brown, W. R., Joshi, K. S., Chang, J., Rosenberg, B., and Voss, E. W. (1993) The immunochemistry of sandwich ELISAs-VI. Greater than 90% of monoclonal and 75% of polyclonal antifluorescyl capture antibodies (CAbs) are denatured by passive adsorption, *Mol. Immunol.* 30, 1165-1175.
- 24. Access Excellence @ the National Health Museum, http://www.accessexcellence.org/RC/VL/GG/monoclonal.php
- Alamargo, J. C., Quintero-Hernandez, V., Ortiz-Leon, M., Velandia, A., Smith, S. L., and Becerril, B. (2006) Design and validation of a synthetic V_H repertoire with tailored diversity for protein recognition, *J. Mol. Rec.* 19, 413-422.
- 26. Binz, H. K., Amstutz, P., and Pluckthun, A. (2005) Engineering novel binding proteins from nonimmunoglobulin domains, *Nat. Biotechnol. Rev.* 23, 1257-1268.
- 27. Hosse, R. J., Rothe, A., and Power, B. E. (2006) A new generation of protein display scaffolds for molecular recognition, *Protein Science 15*, 14-27.
- 28. Reddy, M. M., Bachhawat-Sikder, K., and Kodadek, T. (2004) Transformation of low-affinity lead compounds into high-affinity protein capture agents, *Chem. Biol. 11*, 1127-1137.
- 29. Searle, M. S., Sharman, G. J., Groves, P., Benhamu, B., Beauregard, D. A., Westwell, M. S., Dancer, R. J., Maguire, A. J., Try, A. C., and Williams, D. H. (1996) Enthalpic (electrostatic) contribution to the chelate effect: a correlation between ligand binding constant and a specific hydrogen bond strength in complexes of glycopeptide antibiotics with cell wall analogues, J. Am. Chem. Soc. 1, 2781-2786.
- 30. Mammen, M., Choi, S.-K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design in use of multivalent ligands and inhibitors, *Angewandte Chemie International Edition* 37, 2754-2794.

- 31. Rao, J., Lahiri, J., Isaacs, L., Weis, R. M., and Whitesides, G. M. (1998) A Trivalent System from Vancomycin?D-Ala-D-Ala with Higher Affinity Than Avidin?Biotin, *Science 280*, 708-711.
- 32. Shewmake, T. A., Solis, F. J., Gillies, R. J., and Caplan, M. R. (2008) Effects of linker length and flexibility on multivalent targeting, *Biomacromolecules* 9, 3057-3064.
- 33. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR, *Science 274*, 1531-1534.
- 34. Zhou, J. Z. (2008) Structure-directed combinatorial library design, *Current Opinion in Chemical Biology 12*, 379-385.
- 35. Sridhar, J., Wei, Z.-L., Nowak, I., Lewin, N. E., Ayres, J. A., Pearce, L. V., Blumberg, P. M., and Kozikowski, A. P. (2003) New bivalent PKC ligands linked by a carbon spacer: Enhancement in binding affinity, *Journal of Medicinal Chemistry 46*, 4196-4204.
- 36. Hajduk, P. J. (2006) SAR by NMR: Putting the pieces together, *Mol. Interv.* 6, 267-272.
- 37. Maly, D. J., Choong, I. C., and Ellman, J. S. (2000) Combinatorial target-guided ligand assembly: Identification of potent subtype-selective c-Src inhibitors, *Proc. Natl. Acad. Sci. USA* 97, 2419-2424.
- 38. Brase, S., Dahmen, S., and Lormann, M. E. P. (2003) Multifunctional linkers as an efficient tool for the synthesis of diverse small molecule libraries: The triazene anchors, in *Combinatorial Chemistry*, *Pt B*, pp 127-150, Academic Press Inc, San Diego.
- 39. Reyes, S. J., and Burgess, K. (2006) Heterovalent selectivity and the combinatorial advantage, *Chemical Society Reviews* 35, 416-423.
- 40. Wright, D., and Usher, L. (2001) Multivalent binding in the design of bioactive compounds, *Current Organic Chemistry 5*, 1107.

- 41. Kramer, R. H., and Karpen, J. W. (1998) Spanning binding sites on allosteric proteins with polymer-linked ligand dimers, *Nature 395*, 710-713.
- 42. Susuki, Y., Hojo, M., Okazaki, I., Kamata, H., Sasaki, M., Maeda, M., Nomizu, M., Yamamoto, Y., Nakagawa, S., Mayaumi, T., and Kawasaki, K. (2002) Preparation and biological activities of a bivalent poly(ethylene glycol) hybrid containing an active site and its synergistic site of fibronectin, *Chemical and Pharmaceutical Bulletin 50*, 1229-1232.
- Schaffer, L., Brissette, R. E., Spetzler, J. C., Pillutla, R. C., Ostergaard, S., Lennick, M., Brandt, J., Fletcher, P. W., Danielsen, D. M., Hsiao, K.-C., Anderson, A. S., Dedova, O., Ribel, U., Hoeg-Jensen, T., Hansen, P. H., Blume, A. J., Markussen, J., and Goldstein, N. I. (2003) Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks, *Proc. Natl. Acad. Sci. USA* 8, 4435-4439.
- 44. Kim, Y., Cao, Z., and Tan, W. (2008) Molecular assembly for high-performance bivalent nucleic acid inhibitor, *Proc. Natl. Acad. Sci. USA 105*, 5664-5669.
- 45. Tian, L., and Heyduk, T. (2008) Bivalent ligands with long nanometer-scale flexible linkers, *Biochemistry* 48, 264-275.
- 46. Cairo, C. W., Gestwicki, J. E., Kanai, M., and Kiessling, L. L. (2002) Control of multivalent interactions by binding epitope density, *J. Am. Chem. Soc.* 124, 1615-1619.
- 47. Kehoe, J. W., Maly, D. J., Verdugo, D. E., Armstrong, J. I., Cook, B. N., Ouyang, Y.-B., Moore, K. L., Ellman, J. A., and Bertozzi, C. R. (2002) Tyrosylprotein sulfotransferase inhibitors generated by combinatorial target-Guided ligand assembly, *Bioorganic & Medicinal Chemistry Letters 12*, 329-332.
- 48. Smith, D., Pentzer, E. B., and Nguyen, S. T. (2007) Bioactive and therapeutic ROMP polymers, *Polymer Rev.* 47, 419-459.
- 49. Roy, R. (1996) Syntheses and some applications of chemically defined multivalent glycoconjugates, *Current Opinion in Structural Biology* 6, 692-702.

- 50. Schluter, A. D., and Rabe, J. P. (2000) Dendronized polymers: synthesis, characterization, assembly at interfaces, and manipulation, *Angew. Chem. Int. Ed. 39*, 864-883.
- Crespo, L., Sancilmens, G., Pons, M., Giralt, E., Royo, M., and Albericio, F. (2005) Peptide and amide bond-contianing dendrimers, *Chemisty Reviews 105*, 1663-1381.
- 52. Kim, Y., and Zimmerman, S. C. (1998) Applications of dendrimers in bio-organic chemistry, *Current Opinion in Chemical Biology 2*, 733-742.
- 53. Helms, B. A., Reulen, S. W. A., Nijhuis, S., Graaf-Heuvelmans, P. T. H. M. d., Merkx, M., and Meijer, E. W. (2009) High-affinity peptide-based collagen targeting using synthetic phage mimics: From phage display to dendrimer display, J. Am. Chem. Soc. 131, 11683-11685.
- 54. Tam, J. P., Lu, Y.-A., and Yang, J.-L. (2002) Antimicrobial dendrimeric peptides, *European Journal of Biochemistry 269*, 923-932.
- 55. Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., Read, R. J., and Bundle, D. R. (2000) Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands, *Nature 403*, 669-672.
- 56. Chen, Y., and Liu, Y. (2010) Cyclodextrin-based bioactive supramolecular assemblies, *Chem. Soc. Rev.* 39, 495-505.
- 57. Benito, J. M., Gomez-Garcia, M., Ortiz Mellet, C., Baussanne, I., Defaye, J., and Garcia Fernandez, J. M. (2004) Optimizing saccharide-directed molecular delivery to biological receptors: design, synthesis, and biological evaluation of glycodendrimer, B-cyclodextrin conjugates, *J. Am. Chem. Soc. 126*, 10355-10363.
- 58. Perez-Balderas, F., Morales-Sanfrutos, J., Hernandez-Mateo, F., Isac-Garcla, J., and Santoyo-Gonzalez, F. (2009) Click multivalent domogeneous neoglycoconjugates - Synthesis and evaluation of their binding affinities, *European Journal of Organic Chemistry 15*, 2441-2453.

- 59. Nasr, K., Pannier, N., Frangioni, J. V., and Maison, W. (2008) Rigid multivalent scaffolds based on adamantane, *The Journal of Organic Chemistry* 73, 1056-1060.
- 60. Fan, E., Zhang, Z., Minke, W. E., Hou, Z., Verlinde, C. L. M. J., and Hol, W. G. J. (2000) High-Affinity pentavalent ligands of escherichia coli heat-labile enterotoxin by modular structure-based design, *J. Am. Chem. Soc. 122*, 2663-2664.
- 61. Zhang, Z., Liu, J., Verlinde, C. L. M. J., Hol, W. G. J., and Fan, E. (2004) Large cyclic peptides as cores of multivalent ligands: Application to inhibitors of receptor binding by cholera toxin, *The Journal of Organic Chemistry* 69, 7737-7740.
- 62. Trouche, N., Wieckowski, S. b., Sun, W., Chaloin, O., Hoebeke, J., Fournel, S., and Guichard, G. (2007) Small multivalent architectures mimicking homotrimers of the TNF superfamily member CD40L: Delineating the relationship between structure and effector function, *J. Am. Chem. Soc.* 129, 13480-13492.
- 63. Niederhafner, P., Scaronebestlk, J., Je, J., zcaron, and ek. (2005) Peptide dendrimers, *Journal of Peptide Science 11*, 757-788.
- 64. Merritt, E. A., Zhang, Z., Pickens, J. C., Ahn, M., Hol, W. G., and Fan, E. (2002) Characterization and cyrstal structure of a high-affinity pentavalent receptorbinidng inhibitor for cholera toxin and E. coli heat labile enterotoxin, *J. Am. Chem. Soc. 124*, 8818-8824.
- 65. Turner, D. J., Ritter, M. A., and George, A. J. T. (1997) Importance of the linker in expression of single-chain Fv antibody fragments: optimisation of peptide sequence using phage display technology, *Journal of Immunological Methods* 205, 43-54.
- 66. Kortt, A. A., Dolezal, O., Power, B. E., and Hudson, P. J. (2001) Dimeric and trimeric antibodies: high avidity scFvs for cancer targeting, *Biomolecular Engineering 18*, 95-108.
- 67. Todorovska, A., Roovers, R. C., Dolezal, O., Kortt, A. A., Hoogenboom, H. R., and Hudson, P. J. (2001) Design and application of diabodies, triabodies and tetrabodies for cancer targeting, *Journal of Immunological Methods 248*, 47-66.

- 68. Volkel, T., Korn, T., Bach, M., Muller, R., and Kontermann, F. R. (2001) Optimized linker sequences for the expression of monomeric and dimeric bispecific single-chain diabodies, *Protein Engineering 14*, 815-823.
- 69. Atwell, J. L., Breheney, K. A., Lawrence, L. J., McCoy, A. J., Kortt, A. A., and Hudson, P. J. (1999) scFv multimer of the anti-neuramindase antibody NC10: length of the linker between VH and VL domains dictate precisely the transition between diabodies and triabodies, *Protein Engineering 12*, 597-604.
- 70. Neri, D., Momo, M., Prosero, T., and Winter, G. (1995) High-affinity antigen bidning by chelating recombinant antibodies (CRAbs), *Journal of Molecular Biology 246*, 367-373.
- 71. Silverman, J., Liu, Q., Lu, Q., Bakker, A., To, W., Duguay, A., Alba, B. M., Smith, R., Rivas, A., Li, P., Le, H., Whitehorn, E., Moore, K. W., Swimmer, C., Perlroth, V., Vogt, M., Kolkman, J., and Stemmer, W. P. C. (2005) Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains, *Nature Biotechnology 23*, 1556-1561.
- Kazantsev, A., Walker, H. A., Slepko, N., Bear, J. E., Preisinger, E., Steffan, J. S., Zhu, Y.-Z., Gertler, F. B., Housman, D. E., Marsh, J. L., and Thompson, L. M. (2002) A bivalent Huntingtin binding peptide suppresses polyglutamine aggregation and pathogenesis in Drosophila, *Nature Genetics 30*, 367-376.
- 73. Cowburn, D., Zheng, J., Xu, Q., and Barany, G. (1995) Enhanced affinities and specificities of consolidated ligands for the Src homology (SH)3 and SH2 domains of abelson protein-tyrosine kinase, *The Journal of Biological Chemistry* 270, 26738-26741.
- 74. Reddy, M. M., Bachhawat-Sikder, K., and Kodadek, T. (2004) Transformation of low-affinity lead compounds into high-affinity protein capture agents, *Proc. Natl. Acad. Sci. USA 11*, 1127-1137.
- 75. Handl, H. L., Sankaranarayanan, R., Josan, J. S., Vagner, J., Mash, E. A., Gillies, R. J., and Hruby, V. J. (2007) Synthesis and evaluation of bivalent NDP--MSH(7) peptide ligands for binding to the human melanocortin receptor 4 (hMC4R), *Bioconjugate Chemistry 18*, 1101-1109.

- 76. Shrivastava, A., Wronski, M. A., Sato, A. K., Dransfield, D. T., Sexton, D., Bogdan, N., Pillai, R., Nanjappan, P., Song, B., Marinelli, E., DeOliveira, D., Luneau, C., Devlin, M., Muruganandam, A., Abujoub, A., Connelly, G., Wu, Q. L., Conley, G., Chang, Q., Tweedle, M. F., Ladner, R. C., Swenson, R. E., and Nunn, A. D. (2005) A distinct strategy to generate high-affinity peptide binders to receptor tyrosine kinases, *Protein Engineering, Design, & Selection 18*, 417-424.
- 77. Nielsen, U. B., Adams, G. P., Weiner, L. M., and Marks, J. D. (2000) Targeting of Bivalent Anti-ErbB2 Diabody Antibody Fragments to Tumor Cells Is Independent of the Intrinsic Antibody Affinity, *Cancer Res 60*, 6434-6440.
- 78. Wright, M. J., and Deonarain, M. P. (2007) Phage display of chelating recombinant antibody libraries, *Molecular Immunology* 44, 2860-2869.
- 79. Toth, K., Sauermann, V., and Langowski, J. (1998) DNA curvature in solution measured by fluorescence resonance energy transfer, *Biochemistry* 37, 8173-8179.
- 80. Sprinz, K. I., Tagore, D. M., and Hamilton, A. D. (2005) Self-assembly of bivalent protein-binding agents based on oligonucleotide-linked organic fragments, *Bioorganic & Medicinal Chemistry Letters* 15, 3908-3911.
- 81. Paar, J. M., Harris, N. T., Holowka, D., and Baird, B. (2002) Bivalent ligands with rigid double-stranded DNA spacers reveal structural constraints on signaling by Fc{epsilon}RI, *J Immunol 169*, 856-864.
- 82. Sil, D., Lee, J. B., Luo, D., Holowka, D., and Baird, B. (2007) Trivalent ligands with rigid DNA spacers reveal structural requirements for IgE receptor signaling in RBL mst cells, 2, 674-684.
- 83. Muller, J., Wulffen, B., Potzsch, B., and Mayer, G. (2007) Multidomain targeting generates a high-affinity thrombin-inhibiting bivalent aptamer, *ChemBioChem 8*, 2223-2226.
- 84. Mayer, G., Muller, J., Mack, T., Freitag, D. F., Hover, T., Potzsch, B., and Heckel, A. (2009) Differential regulation of protein subdomain activity with caged bivalent ligands, *ChemBioChem 10*, 654-657.

- 85. Rinker, S., Ke, Y., Liu, Y., Chhabra, R., and Yan, H. (2008) Self-assembled DNA nanostructures for distance dependent multivalent ligand-protein binding, *Nat. Nanotechnol.* 3, 418-422.
- 86. Williams, B. A. R., Diehnelt, C. W., Belcher, P., Greving, M., Woodbury, N. W., Johnston, S. A., and Chaput, J. C. (2009) Creating protein affinity reagents by combining peptide ligands on synthetic DNA scaffolds, *J. Am. Chem. Soc. 131*, 17233-17241.

CHAPTER 2

SELF-ASSEMBLED PEPTIDE NANOARRAYS: A NEW APPROACH TO STUDYING PROTEIN-PROTEIN INTERACTIONS

Berea A.R. Williams, Kyle Lund, Yan Liu, Hao Yan, and John C. Chaput

PUBLISHED: Angewandte Chemie International Edition (2007) 119: 3111 – 3114. Reprinted with permission from Angewandte Chemie International Edition. Copyright 2007 Gesellschaft Deutscher Chemiker.

2.1 Contributions

The following chapter describes a collaborative research project between the laboratories of John Chaput and Hao Yan that sought to develop a ABCD-tile DNA nanoarray with functional c-myc peptides displayed at precise locations capable of binding and positioning the cognate anti-c-myc antibody. Kyle Lund ordered and annealed the strands that assembled into the ABCD tile DNA nanoarray used in this research. Berea Williams designed and synthesized the c-myc peptide-oligonucleotide conjugate (POC) that annealed to the D-tile of the DNA array. Ms. Williams purified and assayed the function of the c-myc POC using polyacrylamide gel mobility shift assay to verify the full construction of the POC and that the recognition of the anti-c-myc antibody to the c-myc peptide. Ms. Williams and Mr. Lund both contributed to the atomic force microscopy characterization of the ABCD-tile DNA nanoarray with and without the POC and the anti-c-myc antibody. Ms. Williams and Dr. Chaput wrote the manuscript, with helpful comments from Mr. Lund, Dr. Yan, and Dr. Liu. The result of this work was published in Angewandte Chemie International Edition.

2.2 Abstract

Structural DNA nanotechnology provides a unique approach to constructing micrometer-scale objects with nanometer-scale features. Despite recent advances, the ability to create patterned surfaces with more than one molecular component remains a challenging problem. Here we describe a possible strategy for producing high-density peptide arrays capable of displaying many different amino acid sequences at well-defined and addressable locations on the same DNA nanostructure. This strategy relies on the genetic information encoded in the nucleic acid portion of a DNA-tagged peptide to position the amino acid sequence at a predetermined location on the array. We demonstrate the feasibility of this approach by monitoring the formation and capture of the myc peptide epitope on a repeating ABCD tiled array by atomic force microscopy. We suggest that this strategy could be used to construct high-density arrays containing many different proteins at spatially defined positions.

2.3 Introduction

Directed molecular assembly of DNA nanostructures based on Watson-Crick base pairing provides a unique approach to constructing micrometer-scale objects with nanometer-scale features (1). Early progress in this area showed that DNA could be used to construct a variety of geometrically diverse objects, including cubes (2), knots (3), truncated octahedron (4), and Borromean rings (5). More recently, DNA has been used as a building block material to produce large patterned arrays of repeating periodicity. Twodimensional (2D) arrays composed of thousands to hundreds of thousands of DNA tiles have been made using positional information encoded in the sticky ends of nonoverlapping DNA to direct the association of one or more individual tiles toward the formation of large lattice structures (6, 7). Although the design of complex DNA nanostructures is of fundamental importance, the intrinsic value of DNA as a building block material lies in its ability to organize other molecules with nanometer-scale spacing (8). This unique property helps distinguish protein nanoarrays from conventional microarrays, since proteins deposited onto glass slides are spatially restricted from interacting with one another due to the micrometer-scale distances between spots. By contrast, protein nanoarrays are assembled in solution on DNA scaffolds that position individual proteins within a few nanometers of each other. The advantages of solution phase chemistry and site-specific positioning make DNA an attractive candidate for developing high-density protein arrays.

Indeed, many examples now show the general utility of DNA as a molecular scaffold for assembling proteins and metallic particles onto array surfaces (9). The most

common method used to functionalize DNA arrays involves Au-sulfur or biotinstreptavidin attachment chemistry (9-11). While these strategies are useful for creating patterned surfaces with one type of molecular component, the design of more divergent surfaces containing many different types of molecular components, each at well-defined and addressable locations on the same array surface, remains a challenging problem. As a possible solution, we have developed a general method to produce high-density peptide arrays that relies on the addressable information encoded in the nucleic acid portion of a DNA-tagged peptide to position specific amino acid sequences at predetermined locations on the DNA array via in situ hybridization to complementary DNA capture probes. We call this approach nanodisplay, since the individual peptides are displayed on the surface of a DNA nanostructure. Given the large number of peptide-nucleic acid combinations that can be constructed, and the mild conditions under which these arrays form, it should be possible to use this technology as a universal platform for generating program-driven peptide nanoarrays. The peptide nanoarrays produced by nanodisplay could, in theory, be used to study a wide-range of protein-protein and protein-inorganic interactions at the nanometer-scale level. Here we demonstrate the feasibility of this approach by capturing the myc-epitope peptide displayed on a 2D DNA array with antimyc mouse antibody 9E10. Using gel electrophoresis and atomic force microscopy (AFM), we were able to validate each step in the assembly process and demonstrated that the myc peptide epitope remains functional when displayed on the DNA array.

2.4 Results

We began by assembling a synthetic DNA scaffold from a set of 22 synthetic oligonucleotides designed to adopt one of four double-crossover (DX) motifs (Fig. 2.1, tiles A-D) (6). The array design, called an ABCD tile array, is a common motif in structural DNA nanotechnology. We modified the D tile to contain a DNA capture probe that functions as a tag for positioning DNA-peptide fusions at a specific addressable locations on the DNA array (9). The DX tile is an ideal building block material for the construction of DNA nanostructures due to the rigid nature of the double-crossover motif, and the intrinsic nanometer-scale dimensions of the tile: ~4 nm x 16 nm (width x length) and ~2 nm in thickness. When imaged by atomic force microscopy (AFM) the ABCD tile array is predicted to give a series of regularly spaced topographical features that indicate the location of the probes on the array surface.

The myc peptide fusion was constructed using standard peptide coupling chemistry to covalently link the myc-epitope peptide to a 5'-amino modified DNA strand (12). This strategy, originally designed to improve the binding affinity of antisense oligonucleotides, functions by first attaching a bivalent coupling reagent, 4- (maleimidomethyl)-1-cyclohexane carboxylic acid N-hydroxysuccinimide ester (SMCC), to the amine-modified oligonucleotide, purifying away the unreacted SMCC, and coupling the SMCC-modified oligonucleotide to a cysteine residue on the peptide. The DNA portion of the DNA-tagged peptide is complementary in sequence to the capture probe on the D tile.



Figure 2.1. Self-assembled 2D Peptide Nanostructure. DX tiles A-D are preassembled in solution from a set of 22 DNA strands. The tiles then further assemble into 2D arrays with the D tile displaying a unique capture probe for in situ hybridization to the DNApeptide conjugate. The peptide conjugated to DNA retains its conformation after hybridization to the DNA nanoarray allowing antibody recognition. Reprinted with permission from Angewandte Chemie International Edition. Copyright 2007 Gesellschaft Deutscher Chemiker.

The myc peptide array (Fig. 2.1) was assembled from four DX tiles using sticky end cohesion. The four tiles were formed separately by heating the strands to 94 °C and slowly cooling to 30 °C. The tiles were then combined with the myc peptide fusion and further annealed by slowly cooling from 40 °C to 24 °C. During the second annealing step, the four DX tiles assemble into a periodic 2D array, and the DNA-tagged peptides hybridize to the DNA capture probes on the D tile. We found that simply heating all of the strands to 94 °C and cooling slowly to 24 °C also led to efficient array formation.

Correct hybridization between the DNA capture probe and the myc peptide fusion was confirmed by gel mobility shift assays performed under native conditions. As illustrated in Figure 2.2A, the band corresponding to the DNA capture probe shifts to slower mobility when incubated with the complementary DNA tag or DNA-tagged peptide fusion (lane 1 versus lanes 2 and 3). The change in electrophoretic mobility is consistent with the formation of a double-stranded DNA helix. To demonstrate that the DNA-tagged peptide fusion remains functional when hybridized to the DNA capture probe, we performed a second gel shift assay (Fig. 2.2B) using the anti-myc antibody to capture the myc peptide fusion on the D tile. Myc peptides displayed on D tiles were incubated with increasing concentrations of anti-myc mouse antibody. Complete capture of the myc peptide fusion by the antibody was visualized by the disappearance of the band corresponding to the D tile displaying the myc peptide epitope. These two experiments indicate that the myc peptide fusion anneals to the DNA capture probe and remains accessible to the anti-myc antibody when localized on the DNA tile.

86



Figure 2.2. Non-denaturing Gel Electrophoresis Assay Showing Efficient Hybridization of the DNA-tagged Peptide Fusion to the DNA Array. A) Native gel mobility shift assay was used to examine the hybridization efficiency of the myc peptide fusion to the DNA capture probe. Lane 1, single-stranded DNA capture probe; lane 2, SMCC conjugated DNA tag annealed to the DNA capture probe; and lane 3, DNA tagged peptide fusion annealed to the DNA capture probe; B) Native gel mobility shift assay to demonstrate immunoprecipitation of the myc epitope displayed on the D tile by the myc antibody 9E10. Lane 1, D tile alone; lane 2, myc peptide fusion annealed to the D tile; and lanes 3-7, binding of the myc peptide nanoarray (18 pmole) with increasing concentrations (8, 11, 67, 90, and 120 pmole) of anti-myc antibody. Reprinted with permission from Angewandte Chemie International Edition. Copyright 2007 Gesellschaft Deutscher Chemiker.

AFM images were collected on arrays assembled in the absence and presence of the myc peptide fusion to verify that the DNA array formed properly in the presence of the myc peptide. In each case, a 2 μ L aliquot of the final array solution was deposited onto a freshly cleaved mica surface, washed with buffer, and imaged by AFM. Figures 2.3A through 2.3C provide a schematic illustration of the three different states of the DNA surface. The AFM images in Figures 2.3D-F show complete array formation, as indicated by a series of parallel lines separated by ~64 nm, which is the approximate distance separating the D tile repeats. The lines on the array are due to changes in the surface height when the AFM probe-tip comes in contact with the DNA capture probe. Comparison of the line height profiles for arrays formed in the absence and presence of the myc peptide fusion (Figures 2.3G versus 2.3H) indicate that arrays formed in the presence of the myc fusion give sharper lines with greater height (~ 0.5 nm versus ~ 1.5 nm, respectively). We attribute the change in height to the increased rigidity of the DNA probe when annealed to the myc peptide fusion, and the increased bulk volume due to the presence of the peptide (Fig. 2.3A, B).

We collected AFM images of the myc peptide array following a brief incubation with the myc antibody. The resulting array shows a series of parallel lines that signify the location of the myc-antibody bound to the myc-peptide epitope annealed to the ABCD tile array. This is evident by the observation that the lines on the array appear much larger following incubation with the anti-myc antibody, than images collected in the absence of the myc antibody (Fig. 2.3E versus 2.3F).



Figure 2.3. AFM Imaging of the Peptide Nanoarray. A-C) Schematic illustration showing the DNA capture probe on the DNA surface, annealed to the myc peptide fusion, and immunocaptured by the anti-myc antibody. D-F) AFM images were collected for the array before hybridization of the myc peptide fusion, after hybridization of the myc peptide, and following incubation with the myc antibody. G-I) Height profiles were determined for the array, the array displaying the myc peptide epitope, and the array with the myc antibody bound to the myc epitope. Reprinted with permission from Angewandte Chemie International Edition. Copyright 2007 Gesellschaft Deutscher Chemiker.

Comparative height profiles taken before and after the addition of the myc antibody (Fig. 2.3H and 2.3I) show that the line height increases from \sim 1.5 nm to \sim 3.0 nm Moreover, almost no antibody binding was observed anywhere else on the array, other than where the myc peptide was annealed to when the antibody is bound to the peptide array. The increased line height is consistent with the larger size of myc antibody (\sim 65 kDa). This experiment confirms that peptides displayed on DNA nanostructures remain accessible to exogenously added proteins. Although we cannot rule-out the possibility that some non-specific interactions between the peptide and the DNA surface may occur at low, perhaps undetectable levels, results from these experiments indicate that such interactions do not impede antibody binding.

2.5 Conclusion

In summary, we present a novel approach for constructing tailor-made peptide and protein nanoarrays with addressable surface features. The construction of highdensity peptide arrays with nanometer-scale features represents a new strategy for studying protein-protein and protein-inorganic interactions for nanobiotechnology and nanoelectronic applications. A major advantage of nanodisplay over conventional microarray systems is the ability to detect substrate binding in solution at the singlemolecule or near single-molecule level. We suggest that this technology in combination with more diverse DNA surface architectures, such as origami assembly, may lead to self-assembled nanoelectronics and nanobiochips capable of monitoring protein pathways at local concentrations reminiscent of cellular pathways (*13, 14*).

2.6 Experimental Design

Myc Peptide Fusion. The myc peptide was conjugated to a synthetic DNA 20-mer as described previously (12). A 5'-amine modified DNA oligonucleotide (5'-ACC AGC TGT GCA GGC CTC GC-3') purchased from Integrative DNA Technologies (Coralville, IA) was conjugated to the bifunctional linker 4-(maleimidomethyl)-1-cyclohexane carboxylic acid N-hydroxysuccinimide ester (SMCC, Sigma Aldrich) by combining 1 mL of SMCC (1 mg/mL) in acetonitrile with 2 mL of DNA (195 nmol) in 0.1 M KHPO₄ buffer (pH 7.8). Following a 3 h incubation at room temperature, a second portion (100 μ L) of SMCC (10 mg/ml) was added and the reaction was allowed to continue overnight at room temperature. The SMCC conjugated DNA was purified by reverse phase HPLC, the sample was collected, and the conjugation was verified by MALDI-MS (calculated-6477.3 Da, observed- 6465.87 Da). The sample was exchanged into water by size exclusion chromatography on a Nap-5 column (Amersham Bioscience). To construct the peptide fusion molecule, the myc epitope peptide (NH2 -EQKLISEEDLC-COOH, 20 nmol) was incubated with the SMCC-conjugated DNA (10 nmol) in 1 mL of 0.1 M KHPO₄ buffer (pH 7.8) for 3 h at room temperature. HPLC analysis indicated 85% product conversion. The myc peptide-DNA conjugation was verified by MALDI-MS (calculated- 7783.75 Da, observed- 7774.89 Da).

Gel Electrophoretic Mobility Shift Assay. Hybridization of myc peptide fusion to the DNA capture probe was verified by 10% native PAGE. The DNA capture probe (26 pmol) was incubated with the myc peptide fusion (25 pmol) in 1 M NaCl and 0.1 M NaHPO₄ buffer (pH 7.0) in an eppendorf tube by heating to 90 °C for 5 min and cooling on ice for 30 min. Samples (9 μ L) were mixed with 1 μ L of 10X native loading dye (0.2% bromophenol blue, 0.2% xylene cyanole, 50% glycerol, 1x TAE-Mg²⁺) and analyzed by 10% native polyacrylamide gel electrophoresis. The resulting gel was visualized by staining with Stains-All (Sigma Aldrich).

Immunoprecipitation of the D-tile by the Myc Antibody 9E10. The D tile containing the myc peptide fusion (18 pmol) was incubated with increasing amounts (0, 8, 11, 18, 24, and 32 pmol) of anti-myc antibody 9E10 (Delta Biolabs) at room temperature for 30 min. Samples (9 μ L) were mixed with 1 μ L of 10X native loading dye (0.2% bromophenol blue, 0.2% xylene cyanole, 50% glycerol, 1x TAE-Mg²⁺) and analyzed by 5% native polyacrylamide gel electrophoresis. The resulting gel was visualized by staining with Stains-All (Sigma Aldrich).

ABCD Array Assembly. The synthetic DNA oligonucleotides used to construct the DX array (Table 2.1) have been reported previously (13), with the exception of two additional strands that make up the DNA capture probe. Tiles A-D were annealed separately at a concentration of 1 μ M in 1x TAE-Mg²⁺ (40 mM Tris base, 20 mM acetic acid, 2 mM EDTA, and 12.5 mM magnesium acetate) by heating to 94 °C and cooling to 30 °C at a rate of 0.8 °C per minutes. The D tile, displaying the myc peptide fusion, was annealed in the presence of the DNA-tagged peptide (120 pmoles). The ABCD array was formed by combining a small aliquot (15 μ L) of each tile in a single reaction tube, heating to 40 °C, and slowly cooling to room temperature overnight. Table 2.1. DNA sequences used to assemble the ABCD array.

A-Tile	A1: 5'-CACAGCGGTAGCGTGGACTAG-3'
	A2: 5'-GACTGCCGACTGGTGCTCACCGTAGTTGCTGG-3'
	A3: 5'-GCAGTACGTGTGGCACAACGGCATGACATACACCGATACGAT-3'
	A4: 5'-ACTATGCTAGTCCTGTATGTCATGCCGTTGTGCCTGAGCACCAGTCG-3'
	A5: 5'-CTGACGCCAGCAACTACGGACACGTACTGCATCGTATCGGACGCTACC-3'
B-Tile	B1: 5'-CGTCAGGCTGCTGTGGTCGTGC-3'
	B2: 5'-GCCATCCGTCGATACGGCACCATGATGCACG-3'
	B3: 5'-CGCTACCGTGCATCATGGACTAACCAGTGACCGCATCGGACAGCAGC-3'
	B4: 5'-GCAGTCGCACGACCTGGCGTTGTACTACGCAATCCTGCCGTATGCACG-3'
	B5: 5'-CACTGGTTAGTGGATTGCGTAGTACAACGCCACCGATGCGGT-3'
C-Tile	C1: 5'-GTAGCGCCGTTAGTGGATGTC-3'
	C2: 5'-AGTTCAGTGTCAATGCTCACCGATTCAACCAG-3'
	C3: 5'-TGTCGTATCGTGGCTGTGTAATCATAGCGGCACCAACTGGCA-3'
	C4: 5'-GATGGCGACATCCTGCCGCTATGATTACACAGCCTGAGCATTGACAC-3'
	C5: 5'-TCGTTGCTGGTTGAATCGGACGATACGACATGCCAGTTGGACTAACGG-3'
D-Tile	D1: 5'-CAACGACGAATCGTGGCTGCCG-3'
	D2: 5'-CATAGTCGTAGTGTCATCACCAGTTGTATCG-3'
	D3: 5'-CGCATGGGATAGCGACCAACCGTTACACCGATGCGGTACATA-3'
	D4: 5'-GCTGTGCGATACAACTGGCATGCGTATGTGTGACGGCGAGGCCT
	GCACAGCTGGT-3'
	D5: 5'-CGTCACACCGCATCGGACGATTGC-3'
	D6: 5'-TGAACTCGGCAGCCTGTAACGCTGGCGCGAGGCCTGCACAGCTGGT-3'
	D7: 5'-GCCAGCGGTTGGTCGCTATCCTGATGACACTACG-3'

^aA-, B-, and C-tiles consist of 5 DNA strands. The D tile is assembled from seven DNA strands. Strands D4 and D6 contain the unpaired site designed to capture the DNA-tagged peptide fusion. Reprinted with permission from Angewandte Chemie International Edition. Copyright 2007 Gesellschaft Deutscher Chemiker.
Atomic Force Microscopy. The AFM imaging was performed under 1xTAE/Mg in a fluid cell on a PicoPlus AFM instrument (Molecular Imaging) in AAC mode (Tapping mode). The tip used in these experiments was the thinner and shorter cantilever of the NP-S tips (Veeco Inc.). All samples (2 μ L of a 2 μ M ABCD array solution) were spotted on freshly cleaved pieces of mica (Ted Pella, Inc.), allowed to dry for 2 min., and incubated in 200 μ L of 1 x TAE-Mg²⁺. The images with the mouse monoclonal c-myc antibody binding the myc epitope were prepared in a similar manner. The DNA array was allowed to bind to the mica surface as before and 2 μ L of a 100 μ g/mL antibody solution was added to the mica and incubated for 2 minutes. The array was imaged as described above.

2.7 References

- 1. Seeman, N. C. (1982) Nucleic acid junctions and lattices, *J. Theor. Biol.* 99, 237-247.
- 2. Chen, J., and Seeman, N. C. (1991) The Synthesis From Dna Of A Molecule With The Connectivity Of A Cube, *Abst. Pap. Am. Chem. Soc. 202*, 162-BIOL.
- 3. Mueller, J. E., Du, S. M., and Seeman, N. C. (1991) Design And Synthesis Of A Knot From Single-Stranded-Dna, *J. Am. Chem. Soc. 113*, 6306-6308.
- 4. Zhang, Y. W., and Seeman, N. C. (1994) A Truncated Octahedron Built From Dna Using Solid-Support Methodology, *Biophys. J.* 66, A24-A24.
- 5. Mao, C., Sun, W., and Seeman, N. C. (1997) Assembly of Borromean rings from DNA, *Nature 386*, 137-138.
- 6. Winfree, E., Liu, F. R., Wenzler, L. A., and Seeman, N. C. (1998) Design and self-assembly of two-dimensional DNA crystals, *Nature 394*, 539-544.
- 7. Yan, H., Park, S. H., Finkelstein, G., Reif, J. H., and LaBean, T. H. (2003) DNAtemplated self-assembly of protein arrays and highly conductive nanowires, *Science 301*, 1882-1884.
- 8. Lund, K., Liu, Y., Wang, H., Lindsay, S., Yan, H. (2005) Self-assembly of a finite size addressable DNA Nanoarrays, In preparation.
- 9. Le, J. D., Pinto, Y., Seeman, N. C., Musier-Forsyth, K., Taton, T. A., and Kiehl, R. A. (2004) DNA-templated self-assembly of metallic nanocomponent arrays on a surface, *Nano Letters* 4, 2343-2347.
- 10. Li, H. Y., Park, S. H., Reif, J. H., LaBean, T. H., and Yan, H. (2004) DNAtemplated self-assembly of protein and nanoparticle linear arrays, *J. Am. Chem. Soc. 126*, 418-419.

.

- 11. Xiao, S. J., Liu, F. R., Rosen, A. E., Hainfeld, J. F., Seeman, N. C., Musier-Forsyth, K., and Kiehl, R. A. (2002) Selfassembly of metallic nanoparticle arrays by DNA scaffolding, *J. Nanopart. Res.* 4, 313-317.
- 12. Harrison, J. G., and Balasubramanian, S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates, *Nucleic Acids Res. 26*, 3136-3145.
- 13. Yan, H., LaBean, T. H., Feng, L. P., and Reif, J. H. (2003) Directed nucleation assembly of DNA tile complexes for barcode-patterned lattices, *Proc. Nat. Acad. Sci. U. S. A. 100*, 8103-8108.
- 14. Rothemund, P. W. K. (2006) Folding DNA to create nanoscale shapes and patterns, *Nature 440*, 297-302.

CHAPTER 3

EVOLUTION OF A HISTONE H4-K16 ACETYL-SPECIFIC DNA APTAMER

Berea A. R. Williams, Liyun Lin, Stuart M. Lindsay, and John C. Chaput

PUBLISHED: Journal of the American Chemical Society (2009) 131: 6330-6331.

Reprinted with permission from Journal of the American Chemical Society.

Copyright 2009 American Chemical Society.

3.1 Contributions

The following chapter describes a collaborative research project between the laboratories of John Chaput and Stuart Lindsay that sought to develop a DNA aptamer with high affinity and specificity to histone H4 protein with an acetyl post-translational modification located on lysine residue 16 (H4K16-Ac). Liyun Lin conducted de novo evolution starting from a naïve DNA library using capillary electrophoresis as the partitioning technique to enrich aptamers that bound H4K16-Ac. Dr. Lin cloned and sequenced 20 DNA aptamers after round 4 of the selection. Berea Williams categorized the 20 aptamers based on the complexity of their predicted secondary structure of using mFold. Ms. Williams chose three aptamers to represent simple, medium, and complex structures. Dr. Lin tested the binding affinity of the 3 aptamers to H4K16-Ac by affinity capillary electrophoresis. Ms. Williams characterized the affinity and specificity of the 3 aptamers against the target H4K16-Ac and 2 off targets, H4 and H4K8-Ac using surface plasmon resonance (SPR). Similarly, Ms. Williams tested the affinity and specificity of a commercially available anti-H4K16-Ac antibody against H4K16-Ac, H4, and H4K8-Ac. Ms. Williams and Dr. Chaput wrote the manuscript, with helpful comments from Dr. Lin and Dr. Lindsay. The result of this work was published in the Journal of the American Chemical Society.

3.2 Abstract

We report the in vitro selection of DNA aptamers that bind to histone H4 proteins acetylated at lysine 16. The best aptamer identified in this selection binds to the target protein with a K_d of 21 nM, and discriminates against both the non-acetylated protein and histone H4 proteins acetylated at lysine 8. Comparative binding assays performed with a chip-quality antibody reveal that this aptamer binds to the acetylated histone target with similar affinity to a commercial antibody, but shows significantly greater specificity (15-fold versus 2,400-fold) for the target molecule. This result demonstrates that aptamers that are both modification and location specific can be generated to bind specific protein post-translational modifications.

3.3 Introduction

The development of high quality affinity reagents to human proteins represents a major challenge in basic and applied biomedicine. Many large-scale biological assays rely on the use of antibodies to interrogate the nature and function of the human proteome (1, 2). Unfortunately, only a small portion of human proteins have antibodies that are available for use in routine molecular and cellular biology assays (3). Even less common are antibodies with high affinity and specificity to specific post-translational modifications (PTMs) and caution is often urged when using antibodies to detect modified proteins in biological samples (4, 5). This shortfall has created a tremendous need for new molecular tools that maintain many of the recognition properties of antibodies, but overcome or avoid some of their limitations (5).

Aptamers, pieces of single-stranded DNA or RNA that fold into threedimensional structures with binding sites that are complementary in shape and charge to target antigens, have received much attention as possible alternatives to traditional antibodies (6, 7). Because these molecules can be produced in vitro by test-tube evolution methods, their recognition and binding properties can be tailored to specific target antigens. Indeed, aptamers have now been created to bind virtually any target including ions, small molecules, drugs, peptides, proteins, and even whole cells (8, 9). Despite these advances, very few aptamers have been identified that bind specific protein PTMs. In fact, only one literature-reported aptamer exists that binds a PTM and this aptamer shows only a 10-fold preference against the unmodified target (10, 11).

3.4 Results

In this report, we address the question of whether aptamers can be created that bind subtle PTMs and distinguish their site of occurrence in a protein sequence. We chose histone H4 acetylated at lysine 16 (H4-K16Ac) as our target due to the importance of this modification in regulating gene activation and silencing (12). Because the K4-H16Ac modification is located on the N-terminal tail, which is a region of the protein that remains unfolded and accessible to chromatin modifying enzymes when assembled into nucleosomes, we used a 15-mer peptide containing residues Gly6 to Lys20 to represent this portion of the protein (13). The use of synthetic peptides as target molecules helped to simplify the selection, as pure histone H4-K16Ac protein is not readily available.

The initial library contained $\sim 10^{14}$ distinct single-stranded DNA molecules with 48 random nucleotide positions flanked on both sides with constant primer-binding sites for PCR. The selection strategy (Figure 3.1) involved a negative selection step to remove molecules that bind to the unmodified histone H4 (H4-K16) tail sequence, followed by a positive selection step to enrich for molecules that bind to the desired H4-K16Ac target. For each round of selection, the single-stranded DNA pool was passed through an affinity matrix displaying the H4 peptide sequence to selectively remove molecules that bind to the non-acetylated H4 tail. Those molecules that remained in the pool were incubated with the H4-K16Ac peptide and functional aptamers were separated from the DNA pool by injecting the mixture onto a neutral coated capillary. Five injections were made per round of selection, and $\sim 10^{11}$ unique DNA sequences were sampled in the starting pool.



selection included a negative followed by a positive selection step to ensure high specificity binding. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

We chose the capillary electrophoresis (CE) protocol for the positive selection step because this technique enables solution-based separation of bound aptamers from the unbound pool (14-16). We felt that this approach would help reduce the occurrence of non-specific binders, which is sometimes a problem with traditional bead-based selections. CE-based selections have the added benefit of ultrahigh partitioning, which enables the discovery of high affinity aptamers in a minimum number of selection rounds (14-16). Electrophoresis was performed using an electric field of 30 kV in a 57 cm long capillary with an inner diameter of 50 mm. Under these conditions, the unbound DNA migrated faster than the DNA-peptide complex, which enabled us to collect the bound DNA in a separate vial by applying pressure to the column after the unbound DNA passed into a waste vial. The DNA from each round of selection was amplified by PCR and made single-stranded by denaturing the DNA product on streptavidin-coated agarose beads. After 4 rounds of in vitro selection and amplification, a second peak became visible in the CE chromatogram (Figure 3.2), indicating that the pool had become enriched in molecules with high affinity to the H4-K16Ac histone tail.

DNA molecules obtained from round 4 of the selection were cloned and sequenced. Analysis of these sequences revealed no one dominant sequence or class of related sequences, but rather a large number of unrelated sequences. We, and others, have reported similar results (14-16), which suggests that the high partition coefficient and minimum number of selection rounds associated with CE-SELEX helps facilitate the discovery of many high affinity aptamers. To categorize the different aptamers,



Figure 3.2. Capillary Electrophoresis Separation of Histone Aptamers. The DNA pool (0.1 μM) from round 1 (pink trace) and round 4 (blue trace) was incubated with histone H4-K16Ac peptide (0.1 nM) and separated by capillary electrophoresis. The unbound DNA migrates as a single peak around 7.5 minutes and the DNA bound to the Histone H4-K16Ac peptide migrates at 8 minutes. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.



Figure 3.3. Secondary Structure of Histone H4-K16Ac Aptamers using the DNA version of mFold. Clones were categorized into simple, intermediate, and complex structures based on their stem loop, stem loop with bulges, and multiple stem loop structures, respectively. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

all of the sequences were organized into different groups (Figure 3.3) based on the complexity of their predicted secondary structure. mFold was used to identify the lowest energy motif for each aptamer. Group I contained aptamers whose structures were dominated by a single stem-loop. Group II contained aptamers with an internal buldge-loop. Group III consisted of aptamers with multiple stem-loop structures. From this set, we chose clones 4.9, 4.3, and 4.20 to represent aptamers with simple, intermediate, and complex secondary structures (Figure 3.4), respectively.

We screened these three aptamers for binding using affinity capillary electrophoresis (ACE) to measure the dissociation constant (K_d) of each aptamer to the H4-K16Ac target. Clone 4.20 had the highest binding affinity with a K_d of 47 ± 24 nM, while clones 4.9 and 4.3 bind the target with K_d 's of 83 ± 75 and 140 ± 30 nM, respectively. As a validation method, we examined the affinity and specificity of clone 4.20 using a more sensitive Biacore T-100 surface plasmon resonance (SPR) instrument. A biotinylated version of clone 4.20 was immobilized onto a streptavidin-coated biosensor chip and the binding response to each target was measured at different concentrations. The maximum binding response taken from each sensogram was plotted versus peptide concentration and the data was fit to a standard hyperbolic curve (Figure 3.5).

Based on the notion that aptamers with more elaborate secondary structures contain more information content than aptamers with simpler structural motifs (17), we chose clone 4.20 to examine the specificity for the H4-K16Ac tail sequence. The binding affinity constant of clone 4.20 to H4-K16Ac was compared with the affinity of this



Figure 3.4. Predicted Secondary Structure of H4-K16Ac Binding Aptamers. Clone 4.9, 4.3, and 4.20 represent aptamers with simple, intermediate, and complex secondary structures, respectively. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.



Figure 3.5. Affinity plots for Anti-H4-K16Ac Antibody and Clone 4.20 Aptamer. A-C) SPR affinity fits for the monoclonal anti-H4-K16Ac antibody against A) H4 and B) H4-K16Ac. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

aptamer to the unmodified variant H4-K16. Analysis of the affinity plots revealed that clone 4.20 binds the H4-K16Ac target with a K_d of 21 nM and discriminates against the unmodified H4-K16 tail by a dramatic 2,400-fold (Table 3.1), corresponding to a DDG of binding of 4.6 kcal/mol. To examine whether clone 4.20 is specific to the acetyl modification at position 16 in the tail, a second binding assay was performed using a peptide with an acetyl group at lysine position 8 (H4-K8Ac). SPR analysis showed that clone 4.20 is >2,400-fold more selective against an acetylated lysine residue at position 8 in the H4 tail sequence, thereby demonstrating that this aptamer is both modification and location specific.

For comparison purposes, the same binding assays were performed using a standard chip quality antibody raised to bind the H4-K16Ac target. This antibody, which is advertised as a highly specific affinity reagent, binds to the desired H4-K16Ac target with a K_d of 6 nM, but shows only 12- and 15-fold specificity against the H4-K16 and H4-K8Ac off-target sequences (Table 3.1). This result highlights a common problem among many commercial antibodies, which is their limited ability to distinguish close PTMs in biological assays (*18*).

3.5 Conclusion

In summary, we describe the creation of a highly specific DNA aptamer to an important histone PTM. The generality of this approach coupled with the low number of selection steps, provides a facile method for generating high quality protein affinity reagents. We suggest that molecules with similar properties could be made that bind a wide range of PTMs with high affinity and specificity.

.

	Clone 4.20		H4-K16Ac Antibody	
Peptide	K _d (nM)	Specificity ¹	K _d (nM)	Specificity ¹
H4-K16Ac	21 ± 11	-	5.9 ± 3.5	-
H4	50,000 ²	2400	87 ± 25	12
H4-K8Ac	$> 50,000^2$	> 2400	94 ± 32	16

Table 3.1. Dissociation Constants and Specificity Values of Aptamer 4.20 and H4-K16Ac Antibody by SPR.

¹Specificity is defined as K_d (off-target)/ K_d (on-target).

²Experimental error ~10%

.

3.6 Experimental Design

Oligonucleotide Sequences.

DNA Library (Keck Facility at Yale University)

5'-GGCGGCGATGAGGATGAC-N48-ACCACTGCGTGACTGCCC-3

N=1:1:1:1 mixture of A:T:G:C

Forward PCR Primer (IDT, Coralville, IA)

5'-FAM-GGCGGCGATGAGGATGAC-3'

FAM= 6-carboxyfluorescein

Reverse PCR Primer (IDT, Coralville, IA)

5'-Bio-GGGCAGTCACGCAGTGGT-3'

Bio= Biotin

Peptide Sequences.

Histone H4 peptide (ASU Protein Facility)

GGKGLGKGGAKRHRK-Bio

Bio= Biotin

Histone H4-K16Ac peptide (Protein Core Laboratory, Baylor College of Medicine)

GGKGLGKGGAK(Ac)RHRK

Ac= acetyl group

Histone H4-K8Ac peptide (Protein Core Laboratory, Baylor College of Medicine)

GGK(Ac)GLGKGGAKRHRK

Ac= acetyl group

(5'-GGCGGCGATGAGGATGAC-(N48)-Reagents. The DNA library ACCACTGCGTGACTGCCC-3) was purchased from the Keck Facility (Yale University). The random region was synthesized using a biased ratio of nucleotide phosphoramidites to give a 1:1:1:1 mixture of A: T: G: C in the oligonucleotide product. The PCR primers were purchased from Integrated DNA Technologies (Coralville, IA). The forward and reverse primers were synthesized with a 6-carboxyfluorescein (6-FAM) and biotin modifications on their 5'-ends, respectively. The H4 peptide tail sequences were synthesized at the ASU Protein Facility, and the acetylated H4 peptide at lysine position 16 (H4-K16Ac) and acetylated H4 peptide at lysine position 8 (H4-K8Ac) were made at the Protein Chemistry Core Laboratory at Baylor College of Medicine. The monoclonal anti-acetyl-Histone H4 (K16) antibody was purchased from Upstate Biotechnology Inc (cat. number: 07-329).

In Vitro Selection. For each round of selection, the DNA pool was made singlestranded by immobilizing the dsDNA onto streptavidin-coated agarose beads (Millipore) and denaturing the strands with 0.15 M NaOH. The sense strand was collected in the flow-through, neutralized, and folded by heat denaturing the DNA at 95 °C for 5 min. and cooling on ice in the presence of selection buffer (3 mM KCl, 10 mM Na2HPO4, 2 mM KH2PO4, 10 mM NaCl, and 5 mM MgCl2, pH 7.4). The DNA pool (0.5 μ M) was incubated in selection buffer (200 μ L) with the H4 peptide (0.5 μ M) containing a Cterminal biotin residue for 1 hour at 24 °C, and passed over streptavidin-coated agarose beads to remove molecules that bind to the non-acetylated target. The unbound DNA was collected, purified by ethanol precipitation and used as input for the positive selection step. DNA molecules that remained in the pool were refolded and the DNA pool (0.1 μ M) was incubated with histone H4-K16Ac peptide (0.1 nM) in selection buffer for 1 hour at 24 °C. The peptide-bound DNA fraction was separated using capillary electrophoresis (see below), recovered in a vial, and amplified by PCR. After four rounds of selection, the DNA was cloned into a Topo TA plasmid (Invitrogen, CA), transformed into Escherichia coli Top10 cells and sequenced at the ASU Sequencing Facility (Table 3.2).

Capillary Electrophoresis Selection. All capillary electrophoresis (CE) separations were performed on a P/ACE 2100 Capillary Electrophoresis system (Beckman Coulter, Inc., Fullerton, CA). Prior to use, the separation capillary (polyacrylamide coated, 50 μ m I.D., 360 μ m O.D., total length = 57 cm, length to detector = 50 cm) was rinsed with ultrapure water for 10 min and equilibrated with selection buffer. All experiments were conducted at 20 °C. A portion of the ssDNA-peptide mixture (~10 nL) was injected onto the separation capillary using pressure injections (1 s, 0.1 psi) and 30 kV was applied across the capillary for 2 min. Laser-induced fluorescence (LIF) was used to monitor the separation of the 6-FAM labeled DNA. Excitation was generated using the 488 nm line of an Ar⁺ laser (Beckman Coulter) and emission was collected at 520 nm. Two major peaks were observed corresponding to bound and unbound fractions of the DNA (Figure 3.2). The bound fraction was collected and amplified using PCR. Five CE separations were performed for every round of selection to increase the number of sequence copies in the assay.

Table 3.2. Round 4 DNA Sequences.

Clone		D	NA Sequence (5'-	-3')	
4.1	AGGAG	GGAACAGC	ACTGAAGCAC	CATAGGGACC	AGAACTA
4.2	GCTGTC	GGGTAGCC	GTTAGAAAG	AGGTTGGAGT	ACACGGC
4.3	GTACGC	CAGGTTAAA	ATCCCAAATG	GTCCGAGGGT	TGCGGCA
4.4	TGGTCG	GGCCAAA	AGTAAGGGGC	GAGGGTCGGG	GATGGGC
4.5	CGAGG	CAGCGCAG	CAGGCGCAAC	GATGAGGGGG	GGTATGG
4.6	GGAGG	CCAGTCAA	GAGGGTGATC	GCACTGGGGG	GTTGGAT
4.7	TCGCTG	TGCAGTCT	CAATGAAGA	FGGAGGTTAA	TAGGGTG
4.8	GGCCGC	GTGGTTA	CGGCAAAAGC	GGGGTCGTGT	CGGAAAT
4.9	GTCTAA	GTAAACTO	GTGGGAAGGC	GCTATGTTTTC	CTGTCCGC
4.10	AGGGGG	GTAGCCCA	GGTCAGATGT	GTGGTGCATA	GTGCGGG
4.11	GGGGGG	GAGGGAGG	GAAGGCACG	FCGCTTTGTAC	CTGTCGA
4.12	GGGGGA	ATACCCGA	GTCAGGGGAT	TGAGTGGTAT	TAGAGA
4.13	GAAGAG	GTGGTGGG	AATCCGCCCC	GTGGACAGGC	GGAGAGA
4.14	CACGGG	GAGAGAC	AAGTGTGCTA	CTGACGAGG	GTCCCGCT
4.15	CGAGTA	CGGCATG	AGGCCATTGG	GCGGCAAGTA	ACCCCGA
4.16	CGGTAC	GAACTACTI	TAACCCAGGG	CGCGCGGAGT	AGGAAA
4.17	TCAATT	GGAGGGG	AGGGACGCAC	CAGGAGCGTA	ACTAGCG
4.18	GAGGGG	GAGGTGGC	GGAAGGGTGA	AGAGGTACAT	TTTTGGC
4.19	AACTGO	GCAGCCAA	GCGTCACGCA	CCAGGAAAG 7	GGGCGTT
4.20	AGACG	FAAGTTAA	TTGGACTTGG	ICGTGTGCGG(CACAGCG
	1	10	20	30	40

Dissociation Constants (K_d) of the Aptamers. Affinity capillary electrophoresis (ACE) was used to obtain K_d values for the different aptamers as previously described (1). Folded clones 4.3, 4.9, and 4.20 were incubated in selection buffer with increasing concentrations of peptide and injected onto the CE instrument. Relative LIF of bound and unbound fractions were recorded for each sample and the peak height of the free DNA was used to generate K_d values by plotting the concentration versus peak height and fit to a 1:1 binding model using the program Origin (Table 3.1).

Surface Plasmon Resonance. Affinity constants for clone 4.20 aptamer were determined using surface plasmon resonance (SPR) (T100 instrument from Biacore). Streptavidin was immobilized on a CM5 chip using standard NHS/EDC immobilization chemistry. Optimal streptavidin (0.2 µg/ml) immobilization occurred in10mM NaOAc buffer, pH 5.0, resulting in 2500 response units. 5' biotinylated clone 20 was annealed in the presence of 1x TAE Mg2+ buffer (20 mM Tris, 20 mM Acetic acid, 2mM EDTA, and 12.5 mM Magnesium acetate, pH 8.0). The annealed aptamer was immobilized on the streptavidin-coated chip yielding over 1,000 response units. The anti-acetyl-Histone H4 antibody was immobilized through standard NHS/EDC immobilization chemistry. PBS tween (10 mM NaH2PO4, 138 mM NaCl, 2.7 mM KCl, 0.05% Tween, pH7.4) buffer was used as the running and dilution buffer in the binding assays at a flow rate of 30 µL/min. Each assay consisted of a 60 second contact period and a 260 second dissociation time, allowing the response units to return to baseline. Any nonspecific binding to the reference cell was subtracted from the sample flow cell response. All sensograms were double referenced using buffer injections. A concentration series of H4K16Ac, H4-K8Ac and H4 peptide ligands were assayed, between 1-500 nM and 0.5-100 μ M, respectively. K_d values were determined from equilibrium binding responses using Biacore evaluation software to fit the curves using 1:1 binding (Figure 3.5).

.

,

3.7 References

- 1. Hober, S., and Uhlen, M. (2008) Human protein atlas and the use of microarray technologies, *Curr. Opin. Biotech.* 19, 30-35.
- 2. Uhlen, M. (2007) Mapping the human proteome using antibodies, *Mol. Cell. Proteomics* 6.8, 1455-1456.
- 3. Blow, N. (2007) The generation game, *Nature 447*, 741-744.
- 4. Kazanecki, C. C., Kowalski, A. J., Ding, T., Rittling, S. R., and Denhardt, D. T. (2007) Characterization of anti-osteopontin monoclonal antibodies: binding sensitivity to post-translational modifications, *J. Cell. Biochem.* 102, 925-935.
- Taussig, M. J., Stoevesandt, O., Borrebaeck, C. A. K., Bradbury, A., Cahill, D., Cambillau, C., de Daruvar, A., Dubel, S., Eichler, J., Frank, R., Gibson, T., Gloriam, D., Gold, L., Herberg, F., Hermjakob, H., Hoheisel, J., Joos, T., Kallioniemi, O., Koegll, M., Konthur, Z., Korn, B., Kremmer, E., Krobitsch, S., Landegren, U., van der Maarel, S., McCafferty, J., Muyldermans, S., Nygren, P., Palcy, S., Pluckthun, A., Polic, B., Przybylski, M., Saviranta, P., Sawyer, A., Sherman, D., Skerra, A., Templin, M., Ueffing, M., and Uhlen, M. (2007) ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome, *Nature Methods 4*, 13-17.
- 6. Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands, *Nature 346*, 818-822.
- 7. Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science 249*, 505-510.
- 8. Wilson, D. S., and Szostak, J. W. (1999) In vitro selection of functional nucleic acids, *Annu. Rev. Biochem.* 68, 611-647.
- 9. Bunka, D. H. J., and Stockley, P. G. (2006) Aptamers come of age- at last, *Nature* 4, 588-596.

- 10. Gilbert, B. A., Sha, M., Wathen, S. T., and Rando, R. R. (1997) RNA aptamers that specifically bind to a K Ras-derived farnesylated peptide, *Bio. Med. Chem. 5*, 1115-1122.
- 11. Vaish, N. K., Dong, F., Andrews, L., Schweppe, R. E., Ahn, N. G., Blatt, L., and Seiwert, S. D. (2002) Monitoring post-translational modification of proteins with allosteric ribozymes, *Nat. Biotech.* 20, 810-815.
- Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions, *Science 311*, 844-847.
- 13. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature 389*, 251-260.
- 14. Mendonsa, S. D., and Bowser, M. T. (2004) In vitro evolution of functional DNA using capillary electrophoresis, *J. Am. Chem. Soc.* 126, 20-21.
- 15. Mosing, R. K., Mendonsa, S. D., and Bowser, M. T. (2005) Capillary Electrophoresis-SELEX Selection of Aptamers with Affinity for HIV-1 Reverse Transcriptase, *Anal. Chem.* 77, 6107-6112.
- 16. Lin, L., Hom, D., Lindsay, S. M., and Chaput, J. C. (2007) In vitro selection of histone H4 aptamers for recognition imaging microscopy, *J. Am. Chem. Soc. 129*, 14568-14569.
- 17. Carothers, J. M., Oestreich, S. C., and Szostak, J. W. (2006) Aptamers selected for higher-affinity binding are not more specific for the target ligand, *J. Am. Chem. Soc. 128*, 7929-7937.
- 18. Zichi, D., Eaton, B., Singer, B., and Gold, L. (2008) Proteomics and diagnostics: let's get specific, again, *Curr. Opin. Chem. Biol.* 12, 78-85.

CHAPTER 4

CREATING PROTEIN AFFINITY REAGENTS BY COMBINING PEPTIDE LIGANDS ON SYNTHETIC DNA SCAFFOLDS

Berea A. R. Williams, Chris W. Diehnelt, Paul Belcher, Matthew Greving, Neal W. Woodbury, Stephen A. Johnston, and John C. Chaput

PUBLISHED: Journal of the American Chemical Society (2009) 131: 17233-17241. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

4.1 Contributions

The following chapter describes a collaborative research project between the laboratories of John Chaput, Neal Woodbury, and Stephen Johnston to develop a novel method of generating high quality synthetic affinity reagents by a combinatorial examination of linkers and ligands (CELL). Dr. Johnston developed the idea of identifying peptide ligands from a peptide microarray that bound non-overlapping sites of a given protein target. Dr. Chaput conceived the idea of using dsDNA as a linking system to spatially separate two peptide ligands for optimal binding. Dr. Belcher and Dr. Diehnelt conducted the peptide microarray fingerprinting assays and characterized the affinity of the peptide ligands using SPR. Dr. Greving performed the peptide-protein crosslinking experiments to map the peptide binding sites to the protein crystal structure. Berea Williams designed the dsDNA scaffolding and SMCC linking system to create all POC molecules. Ms. Williams synthesized and purified the POC molecules using an optimized method further described in chapter 5. The POC molecules were hybridized into synbody molecules and characterized using polyacrylamide gel mobility shift assays by Ms. Williams. Dr. Diehnelt and Ms. Williams conducted the distance screen for the Gal80 synbodies, while Ms. Williams conducted the distance screen for the transferrin synbodies independently. Ms. Williams tested the affinity and specificity of all the synbody molecules using SPR, ELISA, fluorescent anisotropy, and pull down assays. Ms. Williams and Dr. Chaput wrote the manuscript using helpful input from Dr. Woodbury, Dr. Greving, Dr. Belcher, Dr. Diehnelt and Dr. Johnston. The result of this work was published in the Journal of the American Chemical Society.

4.2 Abstract

A full understanding of the proteome will require ligands to all of the proteins encoded by genomes. While antibodies represent the principle affinity reagents used to bind proteins, their limitations have created a need for new ligands to large numbers of proteins. Here we propose a general concept to obtain protein affinity reagents that avoids animal immunization and iterative selection steps. Central to this process is the idea that small peptide libraries contain sequences that will bind to independent regions on a protein surface, and that these ligands can be combined on synthetic scaffolds to create high affinity bivalent reagents. To demonstrate the feasibility of this approach, an array of 4,000 unique 12-mer peptides was screened to identify sequences that bind to nonoverlapping sites on the yeast regulatory protein Gal80. Individual peptide ligands were screened at different distances using a novel DNA linking strategy to identify the optimal peptide pair and peptide pair separation distance required to transform two weaker ligands into a single high affinity protein capture reagent. A synthetic antibody or synbody was created with 5 nM affinity to Gal80 that functions in conventional ELISA and pull-down assays. We validated our synthetic antibody approach by creating a second synbody to human transferrin. In both cases, we observed an increase in binding affinity of ~1000-fold ($\Delta\Delta G = ~4.1$ kcal/mol) between the individual peptides and final bivalent synbody construct.

4.3 Introduction

In the post-genomic era, there exists a tremendous need for protein affinity reagents that can be used to explore the complexity and function of the proteome (1-3). Although traditional antibodies are commonly used for this purpose, only a limited number of human proteins have antibodies that are available for use in standard cellular and molecular biology assays (4,5). This observation is not always evident, as a disproportionate number of antibodies have been raised to a relatively small number of targets (1). Antibodies are further limited by their slow production time, high cost, and poor stability. These problems have prompted researchers to develop synthetic affinity reagents that function with antibody-like properties, but avoid many of the problems associated with traditional animal immunization and hybridoma technologies (δ) . Artificial antibodies currently being developed for this purpose include immunoglobulin domains (scFv, Fab, and Fv), a wide range of alternative protein scaffolds, nucleic acid aptamers, and some small molecule ligands (7-14). While these protein affinity reagents are often easier to construct and engineer than traditional antibodies, the process of their discovery remains labor intensive and often requires iterative rounds of in vitro selection and amplification. Thus, new methods are needed to chemically synthesize protein affinity reagents on scales that are amenable to high throughput production (15).

The main barrier to the development of synthetic antibodies has been the absence of effective methods for generating protein affinity reagents with high affinity to their target proteins. Most small molecule ligands isolated from combinatorial libraries have binding dissociation constants (K_d) in the micromolar range, while typical commercial monoclonal antibodies bind their targets with low nanomolar affinity. One solution to this problem is to create multivalent binding agents by combining two or more moderate affinity (1-10 μ M) ligands on a synthetic tether or polymer (*16*). Transitioning this approach to a general discovery platform requires developing methods to rapidly identify synthetic ligands to protein targets, and a simple and robust system to link these ligands into multivalent affinity reagents (*17*).

Here we sought to develop a scalable method for generating high quality synthetic antibodies that we call synbodies. In contrast to many other synthetic antibody strategies, the central concept developed here (Figure 4.1) is that i) small libraries of short, presumably unstructured, polypeptides would have sufficient diversity to contain members that bind to independent sites on a protein target; ii) these peptides would be flexible enough to allow them to be linked into bivalent reagents; iii) engineerable materials such as DNA could be used to spatially constrain two or more peptides at different distances and orientations; and iv) the construct would result in a synbody with higher affinity than the individual peptide ligands alone. To test this concept, we identified peptides from a microarray that bind to independent sites on the yeast regulatory protein Gal80. We then used a novel DNA linking strategy that we refer to as combinatorial examination of ligands and linkers (CELL) to determine the optimal distance and strand orientation required to transform two weaker affinity ligands into a single high affinity reagent. The synbody produced by this process has an affinity of 5 nM, and detects Gal80 in ELISA and pull-down assays. We validated our approach by generating a second synbody to human transferrin, a common blood plasma protein. In

both cases, we observed an increase in the binding affinity of ~ 1000 -fold between the individual peptides and final synbody construct.



Figure 4.1. Illustration of the Synbody Concept. Synbodies are created by linking two peptides (I, II) that bind the target independently. In this case the linker (III) is double-stranded DNA. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

4.4 Results

Ligand Discovery by Peptide Microarray. We used custom peptide microarrays to identify peptides that bound to independent sites on the surface of the yeast regulatory protein Gal80. Microarrays were synthesized with ~4,000 unique 12-mer polypeptide sequences. The peptide sequence at each position in the array was generated randomly, selecting one of the eight amino acids, Gly, Thr, Gln, Lys, Ser, Trp, Leu, and Arg, for each residue in the sequence. Based on experience with peptide phage library selections (18), all peptides were designed to contain at least one tryptophan residue in their sequence. Gal80 protein labeled with a fluorescent tag was incubated in the presence and absence of the synthetic Gal4 activation domain peptide (Gal4 AD, residues 847-881) (19), which binds the Gal80 repressor site with high affinity, and both samples were deposited onto separate peptide microarrays. Analysis was performed to distinguish peptides that bind the Gal80 repressor site from peptides that bind elsewhere on the protein surface (Figure 4.2A-B). The two array images revealed that 1,644 peptides bind the Gal80 protein with fluorescence intensity greater than 2- fold above the background fluorescence of the microarray. Of these peptides, 957 sequences were blocked by the Gal4 AD peptide, suggesting that this group of peptides binds the repressor site of Gal80.



Figure 4.2. Protein Ligand Discovery Using Peptide Microarrays. Gal80 binding peptides were identified from a peptide microarray containing ~4,000 unique features. Fluorescently labeled Gal80 protein was incubated with the peptide microarray in the absence A) and presence B) of a known ligand to the Gal80 repressor site. Higher overall fluorescence observed when Gal80 was incubated in the absence of the synthetic ligand indicates that many peptides on the microarray bind the repressor site. C-D), scatter plot analysis (non-normalized) was used to identify peptides with high affinity to nonoverlapping sites on Gal80 and low affinity to α_1 -antitrypsin and transferrin. Black circles indicate peptide sequences that were identified with a high fluorescent ratio of Gal80 to α_1 -antitrypsin and transferrin. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

Peptides blocked by Gal4 AD were termed activation domain (AD) peptides, while peptides that were not blocked by the AD peptide were termed Gal80 binding peptides (BP).

To distinguish peptides with specific affinity to Gal80 from peptides that nonspecifically bind to any protein surface, a second set of microarray assays were performed using the human serum proteins, α_1 -antitrypsin and transferrin, to represent two common blood proteins. Array analysis revealed that 28% and 52% of the selected peptides showed significant binding to α_1 -antitrypsin and transferrin, respectively, demonstrating that this collection of Gal80 binding peptides had high non-specific binding. The remaining peptides exhibited varying degrees of affinity for Gal80, and a ranking of the fluorescent intensity for Gal80 relative to α_1 -antitrypsin and transferrin allowed a subset of Gal80 binding peptides to be identified (Figure 4.2C-D) that showed high fluorescence for Gal80 and low fluorescence for α_1 -antitrypsin and transferrin (Table 4.1). In this collection, four peptides were identified that bound the Gal80 repressor site (peptides AD1-4), and six peptides that recognized a different region or regions on the Gal80 protein surface (peptides BP1-6). Each of the ten candidate peptides was synthesized, and assayed on a Biacore surface plasmon resonance (SPR) Flexchip for affinity to recombinant Gal80 protein. The peptides BP1 and AD1 were found to have the highest relative binding to Gal80 (Figure 4.3A).


Figure 4.3. Analysis of the Gal80 Binding Peptides. A) The relative binding affinity of each peptide was determined by SPR. BP1 and AD1 were identified as the two peptides that bound Gal80 with the highest response. B) Separate, non-overlapping binding sites were validated by mapping BP1 and AD1 to the surface of the Gal80 protein. Gal80 regions cross-linked to BP1 and AD1 are circled on the crystal structure (PDB ID: 3BTV) in green and purple, respectively. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

Name	Sequence (N-C)	Gal80	Gal80	Gal80/Gal80-	AAT	Transferrin
			Blocked	Blocked		
BP1	EGEWTEGKLSLR	236.55	11.86	19.95	0.75	1.09
BP2	WSKSRLLSGLSR	125.49	14.21	8.83	0.77	1.33
BP3	SERWESTSELRR	186.77	10.35	18.05	0.63	1.20
BP4	LWLETREGSLTR	180.19	10.22	17.63	0.58	1.52
BP5	GWSERGGSRTRL	143.39	17.81	8.05	0.69	1.32
BP6	WLESSGSTSLRR	141.74	16.50	8.59	1.08	1.03
AD1	GTEKGTSGWLKT	106.85	1.64	65.04	1.05	2.24
AD2	SGWGSSELRKER	21.11	0.68	30.85	0.50	0.72
AD3	TLWSETRTKRER	82.73	1.94	42.65	0.58	0.86
AD4	KKWSKLTEKLRT	56.28	1.94	28.99	0.77	1.78

Table 4.1. Gal80 Peptide Fluorescent Values¹

¹Fluorescent values are the average of Cy3 and Cy5 fluorescent intensities per peptide spot divided by the background fluorescent intensity of the array. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society. Mapping Peptide Binding to Gal80 protein. To determine whether BP1 and AD1 do indeed recognize separate, non-overlapping sites on the Gal80 protein surface, peptide mapping was used to determine the binding location of each peptide. Similar to traditional protein-protein interface mapping (20), BP1 and AD1 were separately crosslinked to the Gal80 protein using a bifunctional crosslinking reagent. The protein-peptide complex was digested with trypsin and analyzed by MALDI-TOF mass spectrometry. The resulting mass spectra (Figure 4.4) showed that Gal80 amino acid residues 384-420 crosslink to residues 1-4 of peptide AD1, and Gal80 amino acid residues 1-8 crosslink to residues 9-12 of peptide BP1. Mapping the peptide binding sites to the surface of the Xray crystal structure of Gal80 (Figure 4.3B) demonstrated that the two peptides recognized different sites on the Gal80 surface, with AD1 binding the Gal4 AD-binding region of Gal80 (21).

Synbody Design. Because BP1 and AD1 could potentially bind Gal80 in many different conformations, it was unclear what distance and which type of linker would be required to transform the two peptides into a high affinity protein capture reagent. While previous approaches to this problem have relied on structural knowledge and extensive chemical synthesis (22, 23), we sought to develop a general strategy that could be performed on any water-soluble protein without prior structural information. We hypothesized that unstructured peptides might be less sensitive to linker restrictions as was the case in previous constructs (22). For this purpose, we developed an assay that we refer to as combinatorial examination of ligands and linkers (CELL), which combines the multiplex capability of SPR with the nanoscale precision of DNA self-assembly to

,#



Figure 4.4. MALDI-TOF Analysis of Gal80 Crosslinked to Each Peptide Ligand after Trypsin Digestion. A) Mass spectrum of digested Gal80 protein with crosslinker (top) and Gal80 protein crosslinked to AD1 peptide (bottom). The peak that appears in the bottom but not the top spectra (indicated by arrow) corresponds to amino acid residues 1-4 of AD1 crosslinked to residues 384-420 of Gal80 (expected = 4871.5, observed = 4871.7). B) Mass spectrum of Gal80 protein with crosslinker (top) and Gal80 protein crosslinked to BP1 peptide (bottom). The peak that appears in the bottom but not the top spectra (indicated by arrow) corresponds to amino acid residues 1-8 of BP1 crosslinked to residues 9-12 of Gal80 (expected = 2074.9, observed = 2074.8). Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.



Figure 4.5. Combinatorial Examination of Ligands and Linkers. A) A model of the DNA backbone showing the spatial separation between individual peptide positions. Modified nucleotide positions (red) indicate the locations where peptides were conjugated to the DNA. B) Combinatorial analysis of bivalent DNA-peptide fusion molecules was performed in a single step using a Biacore Flexchip that measured 400 independent binding interactions. C) The combinatorial peptide pair and peptide pair distance assay was used to screen BP1 and AD1 in two orientations at six different base pair distances on the DNA scaffold. The inset displays the binding response of the forward synbodies at nucleotide positions 9 and 13. The synbody with the highest relative response to Gal80 was determined to be the combination with BP1 and AD1 at positions 1' and 13, respectively. Approximate linear distances in nanometers are given in parenthesis. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

simultaneously search multiple peptide pairs and peptide pair distances in a single binding assay (24). The molecular design (Figure 4.5A) with DNA as a linker relies on DNA-peptide conjugates that self-assemble into bivalent DNA-peptide fusion molecules with two peptides positioned at different distances along the DNA backbone. DNA is a logical choice for a synthetic scaffold as DNA adopts a helical structure that is predictable, rigid over short distances, and easy to engineer. Using standard amine coupling chemistry (25), the template strand was conjugated to the C-terminus of either BP1 or AD1 (Scheme 4.1). This strand was then annealed to a complementary strand that contained either the BP1 or AD1 peptide conjugated at nucleotide positions 9, 13, 15, 17, 24, 26, and 28, thereby creating a small combinatorial library of bivalent fusion molecules separated by distances of 3-9 nanometers.

SPR analysis on a Flexchip (Figure 4.5B) allowed all possible homo- and hetero-BP1 and AD1 peptide pairs to be analyzed in a single experiment. The set of bivalent DNA-peptide fusion molecules were immobilized on the Flexchip, and relative Gal80 binding was determined by flowing Gal80 over the surface. Bivalent complexes constructed of hetero-peptide pairs (Figure 4.5C) showed higher overall binding to Gal80, although the homo-peptide pairs (Figure 4.6) did show substantial binding to Gal80, possibly because Gal80 is a dimeric protein in solution. Of the hetero-pairs tested, a clear trend emerged in which BP1 was favored on the template strand and AD1 on the complementary strand. One reason for this might be that the bivalent affinity reagent has a chiral preference due to the helicity of natural B-form DNA. The synbody with the



Scheme 4.1. Peptide Conjugation to ssDNA. The amine modified DNA oligonucleotide is conjugated to succinimidyl-4-[N-maleimidomethyl]cyclohexane-1-carboxylate (SMCC) through a reaction of the primary amine on the DNA with the NHS ester of SMCC. The maleimide portion of the DNA-SMCC is then conjugated to the peptide using a cysteine residue at the C-terminus of the peptide. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.



Figure 4.6. Relative Binding Affinity for Homo-Peptide Gal80 Synbody Combinations. Synbodies with homo-peptide pairs were spaced at different distances on the DNA scaffold and assayed for affinity to Gal80. Homo-BP1 synbody distance 24 was not assayed. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

highest binding response for Gal80 occurred at a distance of ~4.3 nm with BP1 at position 1' on the template strand and AD1 at position 13 on the complementary strand. We refer to this complex as synbody construct-13 (SC-13). It is interesting that SC-13, which has almost 180° separation between the two peptides shows the highest binding response to Gal80, while the remaining constructs, including SC-9 (Figure 4.5C insert), trend toward lower binding response units. This suggests that multiple distances will yield synbodies with enhanced binding, since the inherent flexibility of the peptide is able to conform to the protein surface.

Synbody Affinity Measurements. The dissociation constant (K_d) of SC-13 was determined by SPR and this value was compared to the affinity of the individual peptide ligands. A summary of this data is given in Figure 4.7A. The individual peptides, either alone or coupled to double-stranded DNA, bound Gal80 with affinity constants of ~5 mM. No binding was observed for the DNA linker alone (K_d > 400 μ M) demonstrating that the DNA linker itself does not interact directly with the target protein. When both peptides were positioned on the DNA linker at the optimal orientation and spatial separation distance, the affinity of SC-13 for Gal80 increased 1000-fold to yield an equilibrium dissociation constant of 5 nM. The affinity of SC-13 and individual peptides was validated independently using an ELISA-type assay (Figure 4.7C) and by fluorescence anisotropy (Figure 4.7B). This dramatic change in binding affinity demonstrates that synbodies can be created with affinity constants similar to traditional antibodies.

Transferrin Ligand Discovery and Peptide Mapping. The CELL process was used to generate a second synbody to the human serum protein transferrin. Transferrin ligands were identified from a custom microarray consisting of 10,000 individual 20-mer peptides (Figure 4.8). We hypothesized that this larger array would cover sufficient sequence space to identify peptides with affinity to different sites on the transferrin protein without the need for a discrete transferrin blocking agent. In these experiments, Alexa-555 labeled transferrin was incubated with the 10,000-peptide microarray in competition with Alexa-647 labeled E. coli lysate. The bacterial lysate served as competitor to aid in the identification of peptides with high specificity to transferrin. The fluorescent ratio of Alexa-555 transferrin to Alexa-647 E. coli lysate was calculated for each peptide on the microarray. Ten peptides were identified with more than 5-fold specificity to transferrin (Table 4.2).

From the list of ten transferrin-binding peptides, peptides 23 (TRF23) and 26 (TRF26) were found to bind different sites on the transferrin surface. Each peptide was immobilized on beads, crosslinked to transferrin using formaldehyde, and digested by trypsin. The formaldehyde crosslinked fragments were reversed with heat and analyzed by MALDI-TOF mass spectrometry. The mass spectra data (Table 4.3) showed that TRF23 was crosslinked to transferrin amino acid residues 415-433, 582-599, and 664-679 and TRF26 was crosslinked to transferrin amino acid residues 435-447 and 448-470. The peptide binding sites were mapped to the surface of the X-ray crystal structure of transferrin (Figure 4.9A), which revealed that TRF23 and TRF26 bind different non-overlapping sites on the transferrin protein.



Figure 4.7. Characterization of the Gal80 Synbody. A) Dissociation constants for BP1 and AD1 were measured as individual peptides, individual peptides on the DNA scaffold, and as a bivalent synbody. The bivalent synbody improves the affinity of the peptides by 1000-fold to produce a synthetic antibody with a K_d of 5.6 nM. B) Solution phase binding affinity of SC-13 was determined by fluorescence anisotropy. Gal80 protein was titrated against fluorescein-labeled synbody and fluorescent anisotropy was measured by exciting at 480 nm and emitting at 525 nm. The average of three experiments resulted in a K_d for SC-13 of 3.0 ± 1.3 nM. C) The dissociation constant of the synbody was validated by ELISA and was determined to be 3.9 ± 0.3 nM. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.



Figure 4.8. Transferrin Ligand Discovery Using Peptide Microarrays. Custom peptide microarrays containing 10,000 unique 20-mer peptides were used to identify transferrin ligands. Alexa-555 labeled transferrin and Alexa-647 labeled E. coli lysate was incubated with the microarray. The slide was scanned for fluorescence at A) 565 and B) 665 nanometers, respectively. C) The overlay of fluorescence at 565 and 665 nanometers was used to identify peptides with affinity to human transferrin in the presence of competing total E. coli lysate. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

Transferrin Synbody Construction and Screen. Transferrin synbody constructs were generated using TRF23 and TRF26 conjugated to different nucleotide positions along the dsDNA scaffold. For the transferrin synbodies, we spatially separated the two peptides at every third position along the DNA scaffold (base pair position 3, 6, 9, 12, 15, 18, 21, 24, and 27) as shown in Figure 4.9B. The small library of transferrin synbodies spanned a distance of ~1.0-9.2 nanometers. The synbodies were screened against transferrin for relative binding response using a Biacore T100 SPR instrument (Figure 4.9C). The synbody with the highest binding response for transferrin occurred at a base pair 6 with a distance of ~2.0 nm and is referred to as transferrin synbody construct-6 (TRF SC-6).

Transferrin Synbody Affinity Measurements. The transferrin peptides and TRF SC-6 were assayed for affinity to transferrin using SPR. Transferrin was immobilized on the SPR chip surface and each ligand was passed over the surface while response units were measured. TRF23 and TRF26 had moderate affinity for transferrin with apparent K_d values of 17.4 μ M and 120 μ M, respectively. We noticed that the dissociation constants for the transferrin ligands are higher than what was previously observed for the Gal80 ligands. This difference could be due to a large entropic penalty associated with the longer peptides used in the transferrin study (12-mer vs 20-mer peptides). TRF SC-6 resulted in a K_d of 117 nM by SPR, which is a ~1,000-fold improvement in binding affinity over the weaker affinity peptide sequence. We validated the binding affinity of TRF SC-6 by fluorescence anisotropy, which produced a K_d of 86. 5 nM (Figure 4.9D), and demonstrated that TRF SC-6 functions in an ELISA like assay with a dissociation constant of 68.4 nM (Figure 4.10).

Table 4.2	Transferrin	Peptide	Fluorescent	Values ¹
1 auto 4.2	. Italistellill	repuide	Tuorescent	values

Name	Sequence (N-C)	Transferrin	Transferrin/E. coli
TRF19	KEDNPGYSSEQDYNKLDGSC	19954.5	7.6
TRF20	GQTQFAMHRFQQWYKIKGSC	16295.8	7.3
TRF21	QYHHFMNLKRQGRAQAYGSC	15007.3	7.0
TRF22	HAYKGPGDMRRFNHSGMGSC	6012.8	6.2
TRF23	FRGWAHIFFGPHVIYRGGSC	12277.8	5.7
TRF24	SVKPWRPLITGNRWLNSGSC	10104.3	5.7
TRF25	APYAPQQIHYWSTLGFKGSC	7701.0	5.6
TRF26	AHKVVPQRQIRHAYNRYGSC	22763.0	5.6
TRF27	LDPLFNTSIMVNWHRWMGSC	16298.8	5.4
TRF28	RFQLTQHYAQFWGHYTWGSC	7502.0	5.2

¹Fluorescent values are for Alexa-555 labeled transferrin and Alexa-657 labeled E. coli lysate per peptide spot. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

m 11 / A		• • •		A 11	1 *	. 1 •
Toble / 7	monat	OPPRING L	antida.	('roodlan	171100	A mo 11/010
1 200 24 3	TAUSI	CI I III-F	сынас	V TOSSIII	кшул	ALIAIVSIS
10010 1101	TINTOT	VIIII I		CI COULL		LIANNA J DIDA
				1	0	•

Peptide	Observed Mass (Da)	Transferrin Fragment	Amino Acid Fragment
TRF23	1798.9	KCSTSSLLEACTFRRP	664-679
TRF23	2003.9	APNHAVVTRKDKEACVHK	582-599
TRF23	2015.9	SDNCEDTPEAGYFAVAVVK	415-433
TRF26	1435.7	SASDLTWDNLKGK	435-447
TRF26	2518.3	KSCHTAVGRTAGWNIPMGLLYNK	448-470

Reprinted with permission from Journal of the American Chemical Society. Copyright

2009 American Chemical Society.



Figure 4.9. Generation of a Transferrin Synbody. A) The transferrin peptides, TRF23 and TRF26, were mapped using PyMOL to the X-ray crystal of transferrin (PDB ID: 2HAV) by protein crosslinking experiments shown in red and blue, respectively. B) A cartoon of the dsDNA scaffold used to spatially separate TRF23 and TRF26 peptides. TRF23 was conjugated to the 1' base pair position on the template strand (orange) and TRF26 was conjugated to one of every third base pair position on the complementary strand (green). C) The peptide distance assay was used to screen TRF23 and TRF26 at nine different base pair positions on the DNA scaffold. The synbody with the highest relative response to transferrin occurred when TRF26 was at the 1' position and TRF23 was at the 6 base pair position termed TRF SC-6. Approximate linear distances in nanometers are given in parenthesis and the cartoon above the graph indicates the peptide positions when looking down the dsDNA helix. D) The average of three fluorescence anisotropy experiments of TRF SC-6 resulted in a K_d of 86.5 ± 18.6 nM. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

Specificity Measurements. A standard pull-down experiment was performed to evaluate the specificity of Gal80 SC-13. This synbody was chosen because it exhibited 10-fold higher affinity for its desired target than TRF SC-6. The synbody was modified with biotin, immobilized on streptavidin coated magnetic beads, and the resin was incubated with total soluble E. coli lysate that contained 3% recombinant Gal80 protein. The beads were washed thoroughly with buffer and the protein that remained bound to the resin was eluted with SDS and analyzed by denaturing gel electrophoresis. A silverstained image of the resulting gel (Figure 4.11) indicated that the Gal80 synbody successfully enriched Gal80 protein from a crude mixture of E. coli proteins. Close inspection of the gel revealed the presence of four low intensity bands at 18, 26, 40, and 70 kDa, respectively. Whether these four proteins were enriched by the Gal80 synbody or the Gal80 protein itself is not yet known. One possibility is that one or more of these E. coli proteins bind Gal80 and are co-eluted in the pull-down assay. Comparison of lanes 4 and 5 reveal that three additional bands observed in the elution lane were due to nonspecific binding to the streptavidin coated magnetic beads. A control experiment involving E. coli lysate without any Gal80 protein present failed to detect any significant binding of SC-13 to endogenous E. coli proteins (Figure 4.12). These data demonstrate that the Gal80 synbody functions as a strong protein capture reagent capable of discriminating Gal80 from many other proteins present in a complex biological mixture.



Figure 4.10. Binding Affinity of SC-6 to Human Transferrin. The binding affinity of SC-6 to transferrin was determined to be 68.4 ± 28.1 nM using an ELISA-type assay. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.



Figure 4.11. Specificity Assay of Gal80 Synbody. A standard pull-down assay was used to evaluate the specificity of the Gal80 synbody. Immobilized SC-13 was able to pull Gal80 protein out of a complex mixture of E. coli lysate. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.



Figure 4.12. Control Pulldown Assay for Non-specific Binding to Endogenous E. coli Proteins. Background binding was analyzed using streptavidin coated magnetic (SA) beads (lane 3), SA beads containing the dsDNA scaffold (lane 6), and SC-13 (lane 9). Lanes 1, 4, and 7 contain the protein ladder; lanes 2, 5, and 8 contain crude E. coli lysate; lanes 3, 6, and 9 contain SDS elution from the beads. Each elution contained 100-fold greater volume than the crude lysate lane. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

4.5 Conclusions

In this work we describe a new approach to creating synthetic protein affinity reagents called synbodies. We first demonstrated that two peptides that bind different sites on the Gal80 protein could be isolated from a relatively small peptide library of ~4,000 unique 12-mers displayed on a microarray surface. The independent binding of the peptides was supported by a crosslinking assay. A novel DNA linking strategy was designed to spatially separate the two peptides at different distances and orientations. One peptide was linked to DNA at a constant position, while the second peptide was linked to different positions on a complementary strand. Annealing the two strands created 24 unique bivalent synbodies, which were simultaneously assayed in a single SPR experiment. The synbody with the highest binding response (SC-13) was further characterized and resulted in a $\sim 1,000$ fold improvement in affinity over the individual peptides. The affinity of the best synbody (5 nM) is comparable to a typical antibody affinity and functions in conventional ELISA and pull-down assays. A second set of experiments was performed on the human blood plasma protein transferrin, and these assays resulted in a similar improvement in binding affinity over the individual peptides. Together, these experiments show how a synthetic protein affinity reagent can be created without resorting to animal immunization methods or iterative rounds of in vitro selection and amplification.

During the course of our experiments, it was discovered that a large portion (~40%) of peptide sequences displayed on the microarray exhibited significant affinity to the target protein. This outcome stands in contrast to most directed evolution experiments

where selections often yield only one or a limited number of solutions, most of which bind the dominant epitope of a given target protein (12). One interpretation of this result is that polypeptide ligands are relatively common in protein sequence space, and therefore large combinatorial libraries are not necessary to find small sequences with relatively simple functions. Indeed, it was discovered that a simple screen could be used to identify ligands that recognized their cognate target with low micromolar binding affinity.

Step two of our synthetic antibody process involved developing a general approach that could be used to transform any two peptide ligands that showed affinity to separate, non-overlapping sites into a single high affinity protein capture reagent. This challenge required designing a strategy that could be broadly applied to a wide range of protein targets. While earlier work on the structure-activity-relationship of protein ligands almost always required some structural knowledge of the target protein (13, 22), we sought to create a strategy that functioned independent of any protein information. Our solution to this problem was to use DNA to systematically explore different peptide pairs and peptide pair separation distances in a single binding assay. DNA is an ideal building block material for this purpose as it allows for small subtle differences in the length to be explored in a systematic fashion. Coupling a combinatorial library of bivalent DNApeptide fusion molecules to the surface of an SPR Flexchip, and screening the different complexes for binding made it possible to rapidly search different peptide pairs and peptide pair distances for optimal binding to the target protein. Through this process, it was discovered that two modest affinity ligands could be transformed into a single high affinity protein-binding reagent. Characterization of the resulting molecule demonstrated that purely chemical methods enabled a synthetic antibody to be created that functioned as an effective antibody mimic.

One interesting phenomenon to come from this study was the observation that two non-competing ligands optimally spaced on a synthetic scaffold improved the binding affinity of the individual peptides by 1000-fold. In a perfect cooperative binding event, where both ligands recognize two independent sites, one might expect the binding affinity of a bivalent affinity reagent to be at or near the product of the affinities of the two peptides (16). In the current study, for example, that would have produced a synbody with an affinity for Gal80 of ~25 pM, which is 200-fold better than the affinity we observed in our binding assays. Determining whether the binding constants of bivalent affinity reagents scale linearly with the affinity of their individual ligands, and how the chemical composition of the linker impacts the net increase in binding affinity remain two very interesting questions. If for example, the affinity constant of bivalent synbodies do indeed scale linearly with the affinity of the individual ligands, then improving the quality, and possibly orientation of the ligands should lead to the creation of synbodies that are able to more closely approximate the cooperativity of a perfect bivalent binding event.

In conclusion, we describe a novel strategy that could be use to develop synthetic antibodies from available chemical building blocks without resorting to protein design, in vitro selection, or animal immunization. The simplicity of this technique suggests that this technology should be amenable to automation, which would make it possible to rapidly generate synbodies to larger numbers of protein targets. These molecules could then be used to investigate the complexity and function of the human proteome—a task currently limited by the availability of high quality antibodies.

4.6 Experimental Design

Protein Expression and Purification. The Gal80 protein was expressed and purified as previously described (21). In brief, a pET28a plasmid containing the Gal80 insert was expressed in E. coli BL21 cells at 17 °C for 16 hrs. Soluble protein was purified by Ni-NTA affinity chromatography, separated from the column by thrombin proteolysis, and purified a second time on a heparin affinity column. The resulting Gal80 protein was concentrated and analyzed by SDS-PAGE and MALDI-TOF mass spectrometry.

Gal80 Peptide Microarrays. Peptide microarrays were used to identify lead peptides for the Gal80 synbodies. Four custom peptide microarrays were synthesized by LC Sciences (Houston, TX) that each contained 3919 unique 12-mer peptides, synthesized from C- to N-terminus on an amino propylsilane slide spaced by a Ahx- β Ala- β Ala linker (26). Individual sequences were chosen from eight natural amino acids (Gly, Thr, Glu, Lys, Ser, Trp, Leu, Arg) using a random sequence generator. Gal80, transferrin, and α_1 -antitrypsin proteins were individually labeled with Cy3 and Cy5 fluorescent dyes and applied to the arrays in four combinations: (i) α_1 -antitrypsin-Cy3 + transferrin-Cy5; (ii) α_1 -antitrypsin-Cy5 + transferrin-Cy3; (iii) Gal80-Cy5 + Gal80-Cy3 (blocked with Gal4 activation domain (AD) peptide); and (iv) Gal80-Cy3 + Gal80-Cy5 (blocked with Gal4 AD peptide). Gal80 was analyzed in the absence and presence of the Gal4 synthetic AD peptide ligand (MDOTAYNAFGITTGMFNTTTMD DVYNYLFDDEDT) by pre-incubating the protein with the ligand (1.5 µM) prior to microarray analysis (19). In each case, the array surface was blocked with 1% BSA, 0.5%

non-fat milk, and PBS-Tween (0.05% Tween-20 in PBS buffer, pH 7.4), washed three times with PBS-Tween, and incubated with the labeled protein (1 μ M) in blocking buffer overnight at 4 °C. Fluorescent images of the array were acquired by scanning the array at Cy3 and Cy5 wavelengths (570 and 670 nm, respectively) using an Axon GenePix 400B Microarray Scanner (Molecular Devices).

Transferrin Peptide Microarrays. Transferrin peptide ligands were identified from a 10.000-feature microarray of 20-mer peptides. The microarray was generated in-house as described previously (27). In brief, random peptides were designed to contain any amino acid with the exception of cysteine, which was added to the C-terminus of each sequence. The peptides were synthesized by Alta Biosceinces Ltd (Birmingham, UK) at >70% purity and diluted in PBS with 25% DMS to 1 mg/ml. The peptides were spotted onto a sulfo-SMCC activated polylysine slide, quenched with mercaptohexanol and stored under argon at 4 °C. The peptide microarray was blocked with 1% BSA and E. coli lysate. The slide was washed 3x with TBST buffer (tris-saline buffer with 0.05% tween) followed by 3 washes with water. Human serum transferrin protein (Sigma) was labeled with Alexa-555 and E. coli lysate was labeled with Alexa-647. Alexa-555 labeled transferfin (1.0 µM) and Alexa-647 labeled E. coli lysate were incubated with the microarry for 3 h at 24 °C. The slide was washed 3x with TBST buffer followed by 3 water washes. The slides were scanned for fluorescence at 565 and 665 nm, respectively (Figure 4.8)

Gal80 Peptide Selection. Relative peptide binding values were calculated as the average of the Cy3 and Cy5 fluorescence intensity per peptide spot divided by the

background fluorescence intensity of the array. Fluorescent binding values were obtained for all 3919 peptides against Gal80, transferrin and α_1 -antitrypsin. Gal80 binding peptides were identified as the subset of sequences that showed a 20-fold preference for Gal80 over two common blood proteins (transferrin and α_1 -antitrypsin). Gal80 binding peptides whose fluorescence intensity dropped by 4-fold or more in the blocking assay with the Gal4 AD peptide were classified as ligands overlapping the Gal80 AD binding site. These sequences were labeled AD1-4 as they mimicked the binding of Gal4 AD peptide. Sequences whose fluorescence intensity was not altered by the presence of the synthetic ligand were labeled BP 1-6, as these peptides recognized non-exclusionary regions on the protein surface. Table 4.1 lists the fluorescent intensity values of BP1-6 and AD1-4.

Transferrin Peptide Selection. Fluorescent binding values were determined for all 10,000 peptides present on the microarray for human transferrin. Alexa-647 labeled E. coli lysate was used to identify peptides with non-specific binding affinity. The ratio of fluorescent intensities of transferrin to E. coli lysate (λ = 565 to 665) was calculated for each peptide spot. Ten peptides were identified that had >5-fold more fluorescence for transferrin over E. coli lysate (Table 4.2).

Peptide Characterization. Peptides BP1-6 and AD1-4 were synthesized with a Gly-Ser-Cys sequence at their C-terminus, purified by HPLC, and verified by MALDI-TOF mass spectrometry. The set of ten peptides were assayed in parallel for Gal80 binding by surface plasmon resonance using a Biacore FlexChip instrument. The surface was treated with 11-amino-1-undecanethiol hydrochloride (MUAM, Dojindo), and peptides were coupled to the array using the bifunctional coupling reagent

sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (sulfo-SMCC). Peptides were spotted at concentrations of 0.05, 0.1, 0.5, and 1.0 mM in PBS-Tween using the Spotarray 72 microarray (Perkin Elmer) with split-type pins (Telechem, SPM7 pins). Following spotting, the array was incubated in a humidity chamber for 1 h and allowed to air dry. The SPR chip was blocked for 3 cycles of 8 min each with 0.7 mg/mL BSA in PBS-Tween and 1 M mercaptohexanol. The matrix (20 x 18 immobilized peptides) was assayed for Gal80 binding by flowing the protein (1 mM) over the chip for 12 min with a flow rate of 1 mL/min and recording the protein dissociation for 10 min. Following each dissociation cycle, the surface was regenerated with a short pulse of 10 mM NaOH in 150 mM NaCl. All data was corrected by reference subtraction and analyzed using the Biacore software package.

Peptide Mapping to Gal80. The binding regions for BP1 and AD1 were determined by protein-protein interface mapping (20). In brief, Gal80 was separately incubated with BP1 and AD1 at stoichiometric concentrations for 1 h in PBS buffer (pH 7.0). A negative control containing Gal80 without the BP1 or AD1 was also performed. Following incubation, a 1:1 mixture of bifunctional crosslinking reagent (BS²G-d₀ and BS²G-d₄, Fisher Scientific) was added to all samples and the mixture was allowed to stand for an additional 45 min. Unreacted crosslinker and peptides were removed from the reaction mixture using a 10 kDa spin-filter with an ammonium bicarbonate (pH 8.5) buffer exchange. The crosslinked samples were digested overnight at 37 °C with trypsin. Digested fragments were separated from undigested protein with a 10 kDa spin-filter and the flow-through was evaporated to dryness. Dry samples were dissolved in a minimal

volume of 0.1% aqueous TFA, desalted using C-18 ZipTip (Millipore, Billerica MA) and analyzed by MALDI-TOF mass spectrometry (Figure 4.4A). An additional analysis using Sulfo-SBED (Thermo Fisher Scientific) photoreactive crosslinker was required for the BP1 peptide due to significant formation of an intramolecular cyclic peptide product with the BS²G-d₀:BS²G-d₄ crosslinker. Analysis of BP1 was performed as described above with the exception that the samples were irradiated at 365 nm for 8 hours, and digest fragments were separated from undigested protein using a 10 kDa spin-filter, and the flow-through was incubated with monomeric avidin beads. The avidin beads were washed extensively and captured fragments were eluted with 50% aqueous acetonitrile containing 0.1% TFA. The eluent was evaporated and redissolved in a small volume of 50% aqueous acetonitrile containing 0.1% TFA and analyzed by MALDI-TOF mass spectrometry (Figure 4.4B).

Peptide Mapping to Transferrin. Transferrin peptides, TRF23 and TRF26, were immobilized in separate spin-column vials (ThermoFisher Scientific) on 25 μ L of UltraLink Iodoacetyl Resin (ThermoFisher Scientific) via the C-terminal cysteine using the manufacturer's recommended protocol. A third spin-column vial containing UltraLink Resin quenched with β -mercaptoethanol was prepared as a negative control. 10 μ M transferrin prepared in 1X PBS buffer was incubated with the prepared resin for 60 min., after which 0.5% (v:v) formaldehyde was added to the samples for 15 min. Formaldehyde crosslinking was quenched with the addition of 100 mM Tris pH 8.5. The resin was washed 3 times with 100 mM Tris pH 8.5, then 3 times with 10 mM glycine pH 2.5, 3 times with nanopure H₂O and finally 3 times with 100 mM Tris pH 8.5. 300 nM proteomics grade Trypsin (Sigma-Aldrich) was prepared in 100 mM Tris pH 8.5 and incubated with the samples for 4 hours at 37 °C. After trypsin digestion, the resin was washed 3 times with 100 mM Tris pH 8.5, then 3 times with 10 mM glycine pH 2.5 and finally 3 times with nanopure H_2O then dried by centrifugation in the spin-column vials. Approximately 20 μ L of nanopure H₂O was added to the bottom of each spin-column vial, below (not in contact with) the resin bed and the vial was partially closed with a screw-cap thereby creating a humid environment inside the vial. Formaldehyde crosslinks were reversed with heat by placing the vials in an oven at 70 °C overnight. Following crosslink reversal, 25 µL of 75% nanopure H₂O 25% acetonitrile was added to the resin to dissolve free transferrin digested peptide fragments. This solution was spun to the bottom of the spin-column vial in a centrifuge then evaporated in a vacuum centrifuge leaving a faint white residue at the bottom. This residue was redissolved with 2 μ L 1:1 Acetonitrile:H2O containing 0.1% trifluoroacetic acid and saturated a-cyano-4hydroxycinammic acid matrix, then spotted on a MALDI-MS target plate and analyzed with MALDI-MS (Table 4.3).

Synbody Construction. The peptides were conjugated to synthetic DNA (Table 4.4 and 4.5) using standard amine coupling chemistry (25) (Scheme 4.1). The amine modified DNA oligonucleotide (Keck Facility, Yale University) was treated with 66 mL of succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) (1 mg/mL) in acetonitrile with 200 mL of DNA (20 nmol) in 0.1 M K₂HPO₄ buffer, pH 7.2. The mixture was incubated for 30 min at 24°C, ethanol precipitated, re-suspended in 200 mL of K₂HPO₄ buffer (0.1 M, pH 7.2), and incubated with 200 mL of peptide (100 nmol) for

3 h at 24°C. Conjugated DNA-peptide molecules were purified by denaturing PAGE, resuspended in H₂O, and quantified by UV absorbance. The template strand was conjugated to a second DNA molecule containing a 3'-biotin or thiol moiety by UV cross-linking. This was achieved by annealing the template-peptide conjugate (2 nmol) with a complementary DNA strand containing a 5'-psoralen and 3'-biotin or 3'-thiol (4 nmol) in cross-linking buffer (100 mM KCl, 1 mM spermidine, 200 mM Hepes, and 1 mM EDTA, pH 8.0), and irradiating the mixture with ultra violet light (366 nm) for 15 min. Cross-linked DNA-peptide fusions were again purified by denaturing PAGE, resuspended in H₂O, and quantified by UV absorbance. The disulfide bond on the crosslinked DNA was reduced prior to use by incubating the molecule in 10 mM tris(2carboxyethyl) phosphine hydrochloride (TCEP) for 30 min at 24°C. Bivalent DNApeptide fusion molecules were assembled by annealing the strands in 50 mM NaCl.

Gal80 Synbody Distance Screen. Bivalent DNA-peptide conjugates were printed onto an activated Flexchip surface using the same protocol described above for peptide printing. Following the immobilization step, the chip surface was blocked for 3 cycles of 8 min each with 0.7 mg/mL BSA in PBS-tween and 1 M mercaptohexanol. Two concentrations of Gal80 protein (1 mM and 100 nM) were prepared in running buffer and tested for binding to the 20 X 18 matrix of immobilized synthetic antibody constructs. Each protein solution was flowed across the chip surface for 12 min at 1 mL/min and protein dissociation was recorded for another 10 min under the same flow rate. At the end of each dissociation cycle, the surface was regenerated with a short pulse of 10 mM NaOH in 150 mM NaCl. All

Name	Sequence (5'-3')	Notes
Template	5CCGAAACAACCGCGAGAGGCACGCGCGTAGC	5= amino-modifier C6 dC
Fluorescein -Template	5CCGAAACAACCGCGAGAGGCACGCGCGTAGC3	5= amino-modifier C6 dC
		3= 6-Fluorescein
Variable Construct 13	GCTACGCGCGTGCCTCTCG5GGTTGTTTCGGG	5= amino-modifier C6 dC
Variable Construct 15	GCTACGCGCGTGCCTCT5GCGGTTGTTTCGGG	5= amino-modifier C6 dC
Variable Construct 17	GCTACGCGCGTGCCT5TCGCGGTTGTTTCGGG	5= amino-modifier C6 dC
Variable Construct 24	GCTACGCG5GTGCCTCTCGCGGTTGTTTCGGG	5= amino-modifier C6 dC
Variable Construct 26	GCTACG5GCGTGCCTCTCGCGGTTGTTTCGGG	5= amino-modifier C6 dC
Variable Construct 28	GCTA5GCGCGTGCCTCTCGCGGTTGTTTCGGG	5= amino-modifier C6 dC
Crosslinking Template	5CCGAAACAACCGCGAGAGGCACGCGCGTAGCC	5= amino-modifier C6 dC
	GTCACCGGCTAT	
Crosslinker-thiol	4TAGCCGGTGTGAAGTTTCTGCTAGTAATG6	4= psoralen
		6= thiol modifier C3
Crosslinker-biotin	4TAGCCGGTGTGAAGTTTCTGCTAGTAATG6	4= psoralen
		6= biotin

Table 4.4. Gal80 Synbody Oligonucleotide Sequences.

Reprinted with permission from Journal of the American Chemical Society. Copyright

2009 American Chemical Society.

Name	Sequence (5'-3')	Notes
Template	5CCGAAACAACCGCGAGAGGCACGCGCGTAGC	5= amino-modifier C6 dC
Fluorescein -Template	5CCGAAACAACCGCGAGAGGCACGCGCGTAGC3	5= amino-modifier C6 dC
		3=6-Fluorescein
Variable Construct 3	GCTACGCGCGTGCCTCTCGCGGTTGTTTC5GG	5= amino-modifier C6 dG
Variable Construct 6	GCTACGCGCGTGCCTCTCGCGGTTGT5TCGGG	5= amino-modifier C6 dT
Variable Construct 9	GCTACGCGCGTGCCTCTCGCGGT5GTTTCGGG	5= amino-modifier C6 dT
Variable Construct 12	GCTACGCGCGTGCCTCTCGC5GTTGTTTCGGG	5= amino-modifier C6 dG
Variable Construct 15	GCTACGCGCGTGCCTCT5GCGGTTGTTTCGGG	5= amino-modifier C6 dC
Variable Construct 18	GCTACGCGCGTGCC5CTCGCGGTTGTTTCGGG	5= amino-modifier C6 dT
Variable Construct 21	GCTACGCGCGT5CCTCTCGCGGTTGTTTCGGG	5= amino-modifier C6 dG
Variable Construct 24	GCTACGCG5GTGCCTCTCGCGGTTGTTTCGGG	5= amino-modifier C6 dC
Variable Construct 27	GCTAC5CGCGTGCCTCTCGCGGTTGTTTCGGG	5= amino-modifier C6 dG

Table 4.5. Transferrin Synbody Oligonucleotide Sequences.

Reprinted with permission from Journal of the American Chemical Society. Copyright

2009 American Chemical Society.

data were corrected by reference subtraction and analyzed using the Biacore software package.

Transferrin Synbody Distance Screen. Transferrin synbody constructs were assayed for relative binding using a Biacore T100 surface plasmon resonance instrument. Transferrin was immobilized on a CM5 chip using standard NHS-amine chemistry and each synbody construct (1 μ M) was passed over the transferrin chip in PBS-tween buffer with a flow rate of 30 mL/min. The binding response of each transferrin synbody construct was measured, with all sensograms double reference subtracted from buffer injections and the reference cell.

Affinity Determination by SPR. Binding affinities for the synbodies and individual components were determined using a Biacore T100 surface plasmon resonance instrument. Gal80 protein was immobilized on a CM5 chip using standard amine coupling chemistry, which resulted in 18,500 response units immobilized to the chip giving an R_{max} of 289 for the peptides and 4800 for the bivalent DNA-peptide conjugates. Binding assays conducted at lower immobilization levels (6,000 RU) produced similar K_d values (1.2 nM for the synbody). Transferrin protein was immobilized on a separate CM5 chip using standard amine coupling chemistry, yielding 6,962 response units immobilized. Individual binding assays were performed at multiple concentrations in standard PBS-tween buffer with a flow rate of 30 mL/min. Each assay consisted of a 100 sec contact time followed by a 300 sec dissociation time. All sensograms were double referenced using buffer injections and the reference cell to subtract nonspecific background binding. Solution binding affinity values were determined using the affinity

fits in the Biacore software package using a 1:1 binding model and represent the average of at least two independent trials. Example affinity plots for Gal80 SC-13 and peptides are shown in Figure 4.13.

Fluorescent Anisotropy. Fluorescein-labeled synbodies, constructed from a fluorescein-labeled oligonucleotide, were diluted into 1x HBS buffer containing 5 mM MgCl₂ to a final concentration that was 10-fold below the expected K_d. This solution was incubated with increasing concentrations of protein that spanned the K_d values. The complex was incubated in a 96-well black plate in a total volume of 100 mL at room temp for 1 hr. The plate was excited at 490 nm with vertically polarized light (I_{vv}), with vertical and horizontal components (I_{vh}) detected at 525 nm (SpectroMax plate reader, Spectro). Fluorescent anisotropy values (r) were calculated by the SpectroMax software using the equation $r = (I_{vv}-GI_{vh})/((I_v+2GI_{vh})$. A G-value of 1.2 was used and the data was subtracted from a buffer-only control (r_f) to give Δr values ($\Delta r = r-r_f$). Experiments were performed in triplicate and the final K_d was determined using GraphPad Prism to fit a hyperbolic curve to the data.

Affinity Determination by ELISA. Enzyme linked immunosorbant assays (ELISA) were conducted by incubating 200 ng of Gal80 protein in 0.1 M sodium bicarbonate, pH 9.8, in a nontreated Maxisorb NUNC 96-well plate overnight at 4°C in a humidifier. The solution was removed and replaced with 100 mL of blocking buffer (2% BSA in PBS, pH 7.4), which was incubated for 1 h at 37°C in the humidifier. The solution was removed and the plate was washed three times with PBS-tween and tapped dry. The biontylated Gal80 synbody, dsDNA, and peptides were added to the plate at varying concentrations

in PBS-tween. The ligands were incubated with the plate for 1 h at 37°C. The ligand solution was removed and the plate was washed three times with PBS-tween. Horseradish peroxidase conjugated streptavidin was diluted 1:1000 in 0.1% BSA in PBS-tween, and 50 mL was added to each well and incubated for 1 h at 37°C. The strepavidin solution was removed and the plate was washed three times with buffer PBS-tween. 50 mL of TMB (3,3', 5,5'-tetramethylbenzidine) was added to each well and the solution was incubated for 15 min at 24°C. 50 mL of 0.5 M HCl was added to stop the reaction and the plate was scanned immediately using a SpectroMax plate reader. The transferrin synbody was assayed for binding affinity in the same manner with the exception that the transferrin synbody was labeled with fluorescein, which enabled direct fluorescent measurements to be made by scanning the plate for fluorescence at 525 nm. These assays were conducted in triplicate, the data were then normalized by subtracting all fluorescent values from the no protein control, plotted, and fit using GraphPad Prism.

Pull-down Assay. Biotinylated Gal80 synbody was immobilized onto prewashed streptavidin coated magnetic beads (DynaBeads, Invitrogen) by incubating 150 pmoles synbody with 30 μ l bead slurry at room temperature for 30 min. Gal80 protein (80 pmoles) and E. coli lysate (10 mL of A₂₈₀-27.8) were incubated with the synbody beads for 30 min at 4 °C. The beads were washed three times with 30 μ l of 1x HBS buffer supplemented with 5 mM MgCl₂. Gal80 was eluted by incubating the synbody beads with 20 μ l protein loading buffer (50 mM Tris, pH 6.8, 10% sodium dodecyl sulfate, 1% β-mercaptoethanol, 50% glycerol, 0.05%


Figure 4.13. SPR Analysis of Gal80 Peptides and Synbody Construct 13. A-C) Representative affinity plots for BP1, AD1, and SC-13 taken from multiple SPR assays, respectively. Affinity plots were created using the 1:1 binding model with the Biacore evaluation software. Reprinted with permission from Journal of the American Chemical Society. Copyright 2009 American Chemical Society.

bromophenol blue). 0.1 μ l ladder (Novex sharp, Invitrogen), 0.5 μ l pure protein (1.6 μ M), 0.1 μ l spiked lysate, and 5 μ l elution were run on a 4-12% SDS-PAGE gel (NuPAGE, Invitrogen) for 45 min at 200V and imaged using SilverXpress silver staining kit (Invitrogen). Beads without the synbody were used as a control to evaluate non-specific binding to the resin. A second control was conducted using E. coli lysate not expressing Gal80 to assess the potential for the Gal80 synbody to bind other proteins present in E. coli lysate (Figure 4.12).

÷

4.7 References

- 1. Blow, N. (2007) The generation game, *Nature* 447, 741-744.
- 2. Hober, S., and Uhlen, M. (2008) Human protein atlas and the use of microarray technologies, *Curr. Opin. Biotechnol. 19*, 30-35.
- 3. Uhlen, M. (2007) Mapping the human proteome using antibodies, *Mol. Cell. Proteomics 6.8,* 1455-1456.
- 4. Chambers, R. S. (2005) High-throughput antibody production, Curr. Opin. Chem. Biol. 9, 46-50.
- 5. Zichi, D., Eaton, B., Singer, B., and Gold, L. (2008) Proteomics and diagnostics: let's get specific, again, *Curr. Opin. Chem. Biol.* 12, 78-85.
- Taussig, M. J., Stoevesandt, O., Borrebaeck, C., Bradbury, A., Cahill, D., Cambillau, C., de Daruvar, A., Dubel, S., Eichler, J., Frank, R., Gibson, T., Gloriam, D., Gold, L., Herberg, F., Hermjakob, H., Hoheisel, J., Joos, T., Kallioniemi, O., Koegll, M., Konthur, Z., Korn, B., Kremmer, E., Krobitsch, S., Landegren, U., van der Maarel, S., McCafferty, J., Muyldermans, S., Nygren, P., Palcy, S., Pluckthun, A., Polic, B., Przybylski, M., Saviranta, P., Sawyer, A., Sherman, D., Skerra, A., Templin, M., Ueffing, M., and Uhlen, M. (2007) ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome, *Nat. Methods 4*, 13-17.
- 7. Hoogenboom, H. R. (2005) Selecting and screening recombinant antibody libraries, *Nat. Biotechnol.* 23, 1105-1116.
- 8. Hudson, P. J., and Souriau, C. (2003) Engineered antibodies, *Nat. Med. 9*, 129-134.
- 9. Hey, T., Fiedler, E., Rudolph, R., and Fiedler, M. (2005) Artificial, non-antibody binding proteins for pharmeceutical and industrial applications, *Trends Biotechnol.* 23, 514-522.
- 10. Bradbury, A. R. M., and Marks, J. D. (2004) Antibodies from phage antibody libraries, *J. Immunol. Methods 290*, 29-49.

- 11. Hosse, R. J., Rothe, A., and Power, B. E. (2006) A new generation of protein display scaffolds for molecular recognition, *Protein Sci.* 15, 14-27.
- 12. Wilson, D. S., and Szostak, J. W. (1999) In vitro selection of functional nucleic acids., *Annu. Rev. Biochem.* 68, 611-647.
- 13. Erlanson, D. A., Wells, J. A., and Braisted, A. C. (2004) Tethering: fragmentbased drug discovery, *Annu. Rev. Biophys. Biomol. Struct.* 33, 199-223.
- 14. Williams, B. A. R., Lin, L., Lindsay, S. M., and Chaput, J. C. (2009) Evolution of a histone H4-K16 acetyl-specific DNA aptamer, *J. Am. Chem. Soc. 131*, 6330-6331.
- Kodadek, T., Reddy, M. M., Olivios, H. J., Bachhawat-Sikder, K., and Alluri, P. G. (2004) Synthetic molecules as antibody replacements, *Acc. Chem. Res.* 37, 711-718.
- 16. Mammen, M., Choi, S.-K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design in use of multivalent ligands and inhibitors, *Angew. Chem. Int. Ed.* 37, 2754-2794.
- 17. Reddy, M. M., Bachhawat-Sikder, K., and Kodadek, T. (2004) Transformation of low-affinity lead compounds into high-affinity protein capture agents, *Chem. Biol. 11*, 1127-1137.
- 18. Barry, M. A., Dower, W. J., and Johnston, S. A. (1996) Toward cell-targeting gene therapy vectors: Selection of cell-binding peptides from random peptide-presenting phage libraries, *Nat. Med.* 2, 229-305.
- 19. Johnston, S. A., Salmeron, J. M., and Dincher, S. S. (1987) Interaction of positive and negative regulatory proteins in the galactose regulon of yeast, *Cell 50*, 143-146.
- 20. Sinz, A., Kalkhof, S., and Ihling, C. (2005) Mapping protein interfaces by a trifunctional crosslinker combined with MALDI-TOF and ESI-FTICR mass spectrometry, J. Am. Soc. Mass Spectrom. 16, 1921-1931.

- 21. Kumar, P. R., Yu, Y., Sternglanz, R., Johnston, S. A., and Joshua-Tor, L. (2008) NADP regulates the yeast GAL induction system, *Science 319*, 1090-1092.
- 22. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR, *Science 274*, 1531-1534.
- 23. Udugamasooriya, D. G., Dineen, S. P., Brekken, R. A., and Kodadek, T. (2008) A peptoid "antibody surrogate" that antagonizes VEGF receptor 2 activity, *J. Am. Chem. Soc.* 130, 5744-5752.
- 24. Homola, J., Vaisocherova, H., Dostalek, J., and Piliarik, M. (2005) Multi-analyte surface plasmon resonance biosensing, *Methods* 37, 26-36.
- 25. Harrison, J. G., and Balasubramanian, S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates, *Nucleic Acids Res. 26*, 3136-3145.
- 26. Pelloise, J. P., Zhou, X., Srivannavit, O., Zhou, T., Gulari, E., and Gao, X. (2002) Individually addressable parallel peptide synthesis on microchips, *Nat. Biotechnol.* 20, 922-926.
- Boltz, K., Gonzalez-Moa, M. J., Stafford, P., Johnston, S. A., and Svarovsky, S. A. (2009) Peptide microarrays for carbohydrate recognition, *Analyst 134*, 650-652.

CHAPTER 5

SYNTHESIS OF PEPTIDE-OLIGONUCLEOTIDE CONJUGATES USING A HETEROBIFUNCTIONAL CROSSLINKER

Berea Williams and John Chaput

Submitted: Current Protocols in Nucleic Acid Research

.

5.1 Abstract

Peptide-oligonucleotide conjugates (POCs) are molecular chimeras composed of a nucleic acid moiety covalently attached to a polypeptide moiety. POCs have been used in numerous applications from therapeutics to nanotechnology, and most recently as combinatorial agents in the assembly of bivalent protein affinity reagents. This unit describes the synthesis and purification of POC molecules using the heterobifunctional crosslinking reagent succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC), which enables amine-modified oligonucleotides to become covalently linked to cysteine-modified polypeptides. This solution-based protocol consists of a two-step synthesis followed by a single purification step.

5.2 Introduction

Peptide-oligonucleotide conjugates (POCs) possess unique properties that can be exploited in a wide range of applications from nanotechnology to drug delivery and antisense technology (1). Several synthetic protocols have been published that describe different approaches for making POC molecules (2). Most of these strategies involve inline, solid-phase synthesis procedures in which the polypeptide and oligonucleotide are synthesized directly on a solid support (3). The in-line synthesis approach requires peptide and DNA synthesizers, along with protecting group strategies that are compatible with the solid-phase synthesis of both polymers.

Fragment conjugation represents a convenient alternative to in-line synthesis. The fragment conjugation strategy presented in this unit utilizes heterobifunctional crosslinking reagents, composed of an aliphatic chain containing a maleimide group on one side and an *N*-hydroxysuccinimide (NHS) on the other (Figure 5.1). This reagent allows cysteine-modified peptides to become covalently attached to amine-modified oligonucleotides. The fragment conjugation approach is highly versatile and can be used to create POCs that crosslink to any position in the peptide or nucleic acid sequence. Unlike the in-line synthesis approach, this technique is performed in solution using standard laboratory techniques without the need for specialized equipment and reagents.

The current unit describes the use of SMCC to covalently link amine-modified DNA to cysteine-containing peptides. The protocol was adapted from a previously published method (4), and requires the use of high purity (>95%) amine-modified DNA. A purification method using denaturing polyacrylamide gel electrophoresis (urea-PAGE)

is described in Basic Protocol 1. As shown in Figure 5.2, the primary amine modification on the DNA is reacted with the NHS ester moiety of the SMCC reagent (Basic Protocol 2) to attach the bifunctional molecule to the DNA. The SMCC-modified DNA is then conjugated to the polypeptide by a Michael-type addition of the sulfhydryl group of the cysteine residue to the maleimide moiety of the SMCC linker (Basic Protocol 3). Peptideoligonucleotide conjugates can be purified using native-PAGE (Basic Protocol 4) or high performance liquid chromatography (HPLC) (Basic Protocol 5).



succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate

SMCC



(N-e-maleimidocaproyloxy)succinimide ester

EMCS







Figure 5.1. Example of heterobifunctional crosslinking reagents that contain maleimide and succinimidyl ester moieties, all of which are compatible with the protocol described in this unit.



Figure 5.2. Amine-modified oligonucleotide conjugation to a cysteine-containing peptide using the heterobifunctional crosslinking reagent SMCC.

5.3 Basic Protocol 1

Preparation of Amine-Modified Oligodeoxynucleotide

Amine-modified oligonucleotides are commercially available and can be purchased from several vendors including the W.M. Keck facility at Yale University and Integrated DNA Technologies (IDT). Most amine-modified DNA is synthesized using one of the four non-standard, amine-modified phosphoramidites (Figure 5.3). For modifications that occur at the 3' terminus of the DNA strand, oligonucleotides can be synthesized from a universal linker or from a CPG column pre-charged with an amine moiety. The amine-modified DNA oligonucleotide must be purified prior to SMCC conjugation in order to remove any primary amines that would otherwise compete with the DNA for the crosslinking reagent. Many companies provide oligonucleotide purification, however this additional service can be cost prohibitive for many labs. Fortunately, amine-modified oligonucleotides are easy to purify by denaturing urea-PAGE, and a standard procedure is given here.



available from Glen Research that can be used as building blocks to construct amine-

D

A

modified deoxyoligonucleotides.

Materials

Amine-modified oligonucleotide

Bisacrylamide (Promega)

Acrylamide (Promega)

8 M urea (BDH)

50 mM ethylenediaminetetraacetic acid, pH 8.0 (EDTA, Pierce)

10X TBE buffer (1 M tris, 1 M boric acid, 10 mM EDTA, pH 8.3)

N',N',N',N', Tetraethylmethylenediamine (TEMED, Pierce)

10% (w/v) ammonium persulfate (APS, EMD Biosciences)

Running dye (0.05% (w/v) bromophenol blue and 0.05% (w/v) xylene cyanol in 1X TBE buffer)

200-proof ethanol (Sigma)

70% ethanol

3 M sodium acetate, pH 5.2 (Sigma)

Gel elution buffer (500 mM ammonium acetate, 10 mM magnesium acetate, 1 mM

EDTA, pH 8.3)

Stir bar

Stir plate

Gel plates (19.7 x 16 cm and 19.7 x 18.5 cm)

Electrophoresis apparatus

Spacers (1.5 mm thick)

Comb (1 well with 2 maker lanes)

Power supply

Heat block

Disposable tubes (1.5 mL)

Plastic transfer pipette (pulled capillary)

Beaker (100 mL)

Plastic syringe (50 mL)

Plastic wrap

UV-active thin layer chromatography (TLC) plate

Handheld UV lamp- 254 nm

Microcentrifuge

Ultrafiltration spin filter tubes (0.45 µm, Millipore)

Black permanent marker

Razor blade or scalpel

Spatula

Vortex

Prepare urea-PAGE gel

1. Prepare the gel plates, spacers, and comb using the manufacturer's recommended protocol or as described in the following reference (5).

Purification gels are 1.5 mm thick and often utilize a comb with a single well.

2. In a 100 mL beaker, combine the acrylamide, bisacrylamide, urea, 10X TBE buffer and water using the amounts listed in Table 5.1.

Oligonucleotides of ≤ 25 nucleotide bases are purified using a 20% polyacrylamide gel. Longer oligonucleotides may require a lower percentage of acrylamide. CAUTION: Acrylamide and bisacrylamide are hazardous. Use appropriate safety precautions and laboratory apparel.

- 3. Stir the mixture using a stir bar and stir plate until the solution is homogeneous. To speed up the process, heat the mixture at 50 °C. Allow the solution to cool before moving to step 4.
- Add 500 μL of 10% APS and 50 μL of TEMED to the 100 mL beaker and stir for 30 seconds.
- 5. Carefully pour the acylamide solution into the prepared gel plates to ensure that no leaks or air bubbles are present between the plates.
- 6. Insert the comb at the top of the gel and dislodge any remaining air bubbles by gently tapping on the glass.
- 7. Allow 30 minutes for the solution to polymerize between the gel plates.

Pre-run the urea-PAGE gel

- 8. Remove the bottom spacer and comb from the polymerized gel.
- 9. Rinse the gel plates with water to remove any polymerized acrylamide from the outside surface of the glass.

Acrylamide %	Acrylamide (g)	Bisacrylamide (g)	Urea (g)	10X TBE buffer (mL)	Deionized water (mL)	10% APS (μL)	TEMED (µL)
6	5.7	0.3	42	10	40	500	50
8	7.6	0.4	42	10	40	500	50
10	9.5	0.5	42	10	40	500	50
12	11.4	0.6	42	10	40	500	50
14	13.3	0.7	42	10	40	500	50
16	15.2	0.8	42	10	40	500	50
18	17.1	0.9	42	10	40	500	50
20	19.0	1.0	42	10	40	500	50

Table 5.1. Urea-PAGE Mixture

.

.

- 10. Place the gel plates in the gel electrophoresis apparatus and secure the plates as suggested by the manufacturer.
- 11. Fill the bottom reservoir of the gel apparatus with 1X TBE buffer. Remove any air bubbles trapped under the gel plates by tilting the apparatus to the side or using a syringe full of 1X TBE buffer to displace the bubbles.
- 12. Pour 1X TBE buffer into the top reservoir of the electrophoresis apparatus until the solution is ~3 cm above the gel.
- 13. Rinse the well of the gel with 1X TBE buffer to remove any residual polyacrylamide.
- 14. Connect the electrophoresis apparatus to the power supply and set the power supply to 20 constant watts for 30 minutes.

The ideal wattage for the gel will heat the glass plates so they are warm but not too hot to touch. Temperatures above ~ 70 °C can break the glass plates.

15. While the gel is pre-running, prepare the crude oligonucleotide.

Prepare the oligonucleotide

16. Resuspend the oligonucleotide in deionized water for a final concentration of 1 μmole/mL.

Add 1 mL of water to a 1 μ mole scale DNA synthesis, and 200 μ L of water to a 0.2 μ mole scale DNA synthesis. If the pellet does not dissolve immediately, heat the mixture at 37 °C until homogeneous.

17. Transfer 1/4 the volume of the oligonucleotide solution to a 1.5 mL tube.

- 18. Add 1/5 the volume of 8 M urea and 1/10 the volume of 50 mM EDTA to the oligonucleotide solution in step 17.
- 19. Heat the mixture to 90 °C for 5 minutes to denature the DNA oligonucleotide.

Load and run the urea-PAGE gel

- 20. Once the gel has finished pre-running, disconnect the gel apparatus from the power supply and rinse the wells of the urea-PAGE gel with 1X TBE to remove any residual urea.
- 21. Load the denatured oligonucleotide into the large well of the urea-PAGE gel using a pulled capillary of a plastic transfer pipette.
- 22. Load 5 μ L of the running dye into the small wells located on each side of the large well.

Running time will vary depending on the size of the oligonucleotide, but the two dyes can be used to approximate the location of the oligonucleotide in the gel according to the acrylamide percentage as shown in Table 5.2.

23. Reconnect the gel apparatus to the power supply and set the power at a constant20 watts.

Gel extraction

- 24. Carefully remove the gel from the glass plates using a spatula to peal the gel away from the glass.
- 25. Wrap the gel front and back in a single layer of plastic wrap.

Table 5.2. Migration of Running Dye in Relation to

Oligonucleotide Length in Nucleotides (nt)

Acrylamide %	Bromophenol Blue (nt)	Xylene Cyanol (nt)		
6	25	103		
8	18	75		
10	14	58		
12	11	48		
14	9	40		
16	8	35		
18	7	31		
20	6	28		

.

- 26. Place the gel on a UV-active TLC plate and image the gel by holding a handheld UV lamp turned on at 254 nm ~ 6 inches above the gel. The DNA will absorb the light and cast a shadow on the TLC plate.
- 27. Using a black permanent marker, trace the band corresponding to the oligonucleotide on the plastic wrap. Do this quickly, as the UV light will damage the DNA.
- 28. Turn the UV lamp off and cut the oligonucleotide band out of the gel using a clean razor or scalpel.
- 29. Cut the gel slice into 1-2 cm size pieces and transfer ~ 10 gel pieces to each 0.45 μ m spin filter tube.
- 30. Place the tubes in the -20 °C freezer for 10 minutes.
- 31. Remove the tubes from the freezer and further crush the gel pieces with a spatula, being careful not to break the membrane in the spin filter.
- 32. Add 200 μ L of gel elution buffer to each spin filter tube.
- 33. Let the mixture sit at room temperature for 4 hours.

Gentle agitation or rotation can increase the rate of diffusion.

- 34. Place the spin filter tubes in a microcentrifuge and spin for 5 minutes at 10,000 rpm.
- 35. Remove the filters from the tube and transfer the filtrate to a new 1.5 mL disposable tube.
- 36. Repeat steps 32-35 two more times.

As an alternative to the buffer diffusion method, electroelution can be used to recover oligonucleotide material from the gel pieces.

Ethanol precipitation

- 37. Dispense 200 µL of gel elution buffer into one or more 1.5 mL disposable tube.
- 38. Add 20 µL of 3 M sodium acetate and 800 µL of 200-proof ethanol to each tube.
- 39. Place each tube in the -80 °C freezer or on dry ice for 30 minutes.
- 40. Remove tubes from the freezer and centrifuge the mixture at 13,000 rpm for 30 minutes at 4 °C.
- 41. After centrifugation, a white pellet will be visible near the bottom of the tube.
- 42. Discard the supernatant and repeat steps 38-41 with 70% ethanol instead of 200proof ethanol.
- 43. Carefully discard the supernatant and allow the pellet to air dry.

5.4 Basic Protocol 2

Conjugation of SMCC to the Amine-Modified Oligonucleotide

The pure amine-modified oligonucleotide is conjugated to SMCC (Figure 5.4) through a coupling reaction between the primary amine on the DNA strand and the NHS ester moiety of SMCC. The rate of the reaction is dependent on the pH of the solution (6), and must be tightly controlled to avoid cross-reactivity of the amine with the maleimide moiety of the SMCC reagent (7). Optimal buffer conditions are pH 7.0 - 8.0 for both the initial coupling step between the amine-modified oligonucleotide and SMCC reagent, and the second coupling step between the peptide and SMCC-oligonucleotide.



Figure 5.4. Conjugation of SMCC to an amine-modified deoxyoligonucleotide.

Materials

Pure amine-modified oligonucleotide (Basic Protocol 1)

100 mM KH₂PO₄ buffer, pH 7.2 (BDH)

Succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC, Pierce)

Acetonitrile (Sigma)

200-proof ethanol (Sigma)

70% ethanol

3 M sodium acetate, pH 5.2 (Sigma)

Disposable tubes (1.5 mL)

Microcentrifuge

Vortex

Conjugate SMCC to the DNA oligonucleotide

1. Resuspend the purified oligonucleotide in 100 μ L of ultrapure water.

If you have multiple tubes, add a fraction of the volume of water to each tube, resuspend each pellet, and then combine the solution into one 1.5 mL tube.

2. Quantify the oligonucleotide using UV absorbance and Beer's Law. The extinction coefficient for the oligonucleotide sequence can be calculated based on the base composition.

The extinction coefficient of the oligonucleotide can be calculated by adding together the extinction coefficients of each nucleobase in the sequence, see Table 5.3. The concentration of the oligonucleotide in solution can be Table 5.3. Molar Extinction Coefficient for each

Nucleobase

Nucleobase	Extinction coefficient (ϵ) at 260 nm ¹
Т	8700 M ⁻¹ cm ⁻¹
dC	$7400 \text{ M}^{-1} \text{cm}^{-1}$
dG	$11500 \text{ M}^{-1} \text{cm}^{-1}$
dA	$15400 \text{ M}^{-1} \text{cm}^{-1}$

¹Determined at 25 °C and neutral pH conditions (Fasman, 1975).

calculated using Beer's law $(A_{254}=cl\varepsilon)$, where c is the molar concentration, l is the pathlength of the UV cell (typically 1 cm), and ε is the molar extinction coefficient of the oligonucleotide sequence.

- 3. Transfer 20 nmoles of the oligonucleotide from step 1 to a new 1.5 mL tube.
- 4. Add 134 μ L of 100 mM KH₂PO₄ buffer, pH 7.2 to the tube.
- 5. In a separate 1.5 mL tube, dissolve 1 mg of SMCC in 1 mL of acetonitrile for a final concentration of 3 mM.
- Transfer 67.0 μL of SMCC (200 nmoles) solution to the oligonucleotide solution prepared in step 4.
- 7. Vortex the solution vigorously for 5 seconds. Centrifuge the tube briefly.
- 8. Allow the reaction to take place at room temperature for 30 minutes.
- 9. Conduct an *Ethanol precipitation* on the SMCC-oligonucleotide conjugate material as described in Basic Protocol 1.

SMCC is soluble in ethanol while DNA is not. The majority of excess SMCC is separated from the SMCC-oligonucleotide conjugate when the ethanol fraction is removed. The SMCC-oligonucleotide conjugate is best stored as a dry solid at -20 °C.

5.5 Basic Protocol 3

Conjugation of SMCC-Oligonucleotide to a Cysteine-Containing Peptide

The sulfhydryl group of a cysteine residue is used to conjugate the peptide to the maleimide functional group of the SMCC-oligonucleotide intermediate (Figure 5.5). This fragment conjugation strategy is limited to peptides that contain one cysteine residue per sequence, preferably at the N- or C-terminus. When incorporating a cysteine residue at the peptide terminus, the addition of a flexible amino acid linker, such as glycine-serine dipeptide, between the peptide and terminal cysteine residue provides a convenient spacer that can reduce interference between the functional region of the polypeptide and the DNA polymer. It is critical to ensure that all reagents used in the synthesis and deprotection of the polypeptide are removed prior to conjugation to the SMCC-oligonucleotide. Many commercial companies provide HPLC purification services to obtain peptides with high purity (often 70% to >98%). For this protocol, it is suggested to use peptides with > 98% purity.



Figure 5.5. Conjugation of SMCC-oligonucleotide to a cysteine-containing polypeptide.

Materials

SMCC-oligonucleotide (Basic Protocol 2) 100 mM KH₂PO₄ buffer, pH 7.2 (BDH) Polypeptide (>98% purity) Siliconized disposable tubes (1.5 mL, Fisher)

Microcentrifuge

Vortex

Conjugate SMCC-oligonucleotide to Peptide

- Resuspend the peptide in water for a final concentration of 1 mg/mL solution.
 Depending on the amino acid composition of the polypeptide, different solvents may be necessary to dissolve the peptide. Solvents such as acetonitrile, methanol, acetic acid and dimethylsulfoxide (DMSO) are compatible dissolving agents. Do not use any solvents that contain sulfhydryl groups.
 - 2. Resuspend the SMCC-oligonucleotide conjugate in 200 μ L of 100 mM KH₂PO₄ buffer, pH 7.2 and transfer to a 1.5 mL siliconized tube.

Polypeptides can nonspecifically bind to standard polypropylene tubes, therefore siliconized tubes are used to reduce the peptide interaction with the tube.

- 3. Add 100 nmoles of peptide to the tube in step 2.
- 4. Vortex the solution for 5 seconds.

5. Centrifuge the tube briefly.

.

6. Let the reaction take place at room temperature for at least 3 hours.

5.6 Basic Protocol 4

Purification of Peptide-oligonucleotide conjugates using Native-PAGE

POCs can be purified using standard urea-PAGE (Basic Protocol 1), native-PAGE (Basic Protocol 4), or HPLC (Basic Protocol 5) purification. Urea-PAGE differs from native-PAGE purification by running the gel under non-denaturing conditions and stabilizing the gel temperature at 4 °C. Urea-PAGE purification can be used to purify the POC product, however we, along with others, have observed the degradation of some POC sequences at high temperatures caused by the urea-PAGE running conditions (8). For these POC sequences, native-PAGE or HPLC purification is used. HPLC is a relatively fast method for POC purification, however retention times of the starting reagents and POC product may vary based on the amino acid composition of the peptide. Native-PAGE provides a convenient alternative to HPLC purification and optimal resolution between the oligonucleotide and POC product is easy to obtain.

Materials

Crude POC material (Basic Protocol 3)

40% acrylamide/bisacrylamide 19:1 solution (BioRad)

10X TAE-Mg²⁺ (400 mM tris, 200 mM acetic acid, 20 mM EDTA, 125 mM magnesium acetate, pH 7.5)

N', N', N', N', Tetraethylmethylenediamine (TEMED, Pierce)

10% ammonium persulfate (APS, EMD Biosciences)

Glycerol (BDH)

Native running dye (0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol, 80%

(v/v) glycerol)

NAP-10 (GE Lifesciences)

Beaker (100 mL)

Siliconized disposable tubes (1.5 mL, Fisher)

Stir bar

Stir plate

Syringe (50 mL)

Microcentrifuge

Gel plates (19.7 x 16 cm and 19.7 x 18.5 cm)

Spacers (1.5 mm thick)

Comb (10 wells)

Electrophoresis apparatus

Power supply

SyberGold (Invitrogen, optional)

Prepare native-PAGE gel

- 1. Prepare the gel plates, spacers, and comb using the manufacturer's recommended protocol or as described in the following reference (5).
- Combine the 40% acrylamide/bisacrylamide solution, 10X TAE-Mg²⁺ buffer and water in a 100 mL beaker using the amounts listed in Table 5.4 for a given acrylamide percentage.

Gel %	40% Acrylamide/bisacrylamide 19:1	10X TAE-Mg ²⁺	Deionized	10% APS	TEMED
	(mL)	buffer (mL)	water (mL)	(µL)	(µL)
6	6	4	30	500	50
8	8	4	28	500	50
10	10	4	26	500	50
12	12	4	24	500	50
14	14	4	22	500	50
16	16	4	20	500	50
18	18	4	18	500	50
20	20	4	16	500	50

Table 5.4. Native-PAGE Mixture

For oligonucleotides with <25 bases, use a 20% acrylamide gel concentration.

- 3. Stir the mixture using a stir bar and stir plate until the solution is homogeneous.
- 4. Add 500 μ L of 10% APS and 50 μ L of TEMED and mix for 30 seconds.
- Carefully pour the solution into the prepared gel plates to ensure that no leaks or air bubbles are present between the plates.
- 6. Insert the 10-well comb at the top of the gel and dislodge any remaining air bubbles by gently tapping on the glass.
- 7. Allow 60 minutes for the solution to polymerize between the plates.

Pre-run the native-PAGE gel

- 8. Remove the bottom spacer and comb from the polymerized gel.
- Rinse the gel plates with water to remove any polymerized material on the surface of the glass.
- 10. Place the gel plates in the gel electrophoresis apparatus and secure the plates as suggested by the manufacturer.
- 11. Fill the bottom reservoir of the gel apparatus with 1X TAE-Mg²⁺ buffer. Remove any air bubbles trapped under the gel plates by tilting the apparatus to the side or using a syringe full of 1X TAE-Mg²⁺ buffer to displace the bubbles.
- 12. Pour 1X TAE-Mg²⁺ buffer in the top reservoir of the electrophoresis apparatus to \sim 3 cm above the top of the gel.
- 13. Rinse the wells of the gel with 1X TAE-Mg²⁺ buffer to remove any residual polyacrylamide.

14. Connect the electrophoresis apparatus to the power supply and set the power to a constant 200 volts for 30 minutes.

Native gel electrophoresis should be run using constant voltage to prevent the gel plates from heating. It is best to place the gel apparatus in a 4 $^{\circ}$ C room or connect the apparatus to a water bath with water cycling at 4 $^{\circ}$ C to keep the gel plates cool.

15. While the gel is pre-running, prepare the POC material by adding 1/10 the volume of glycerol and 1/10 the volume of 10X TAE-Mg²⁺ buffer to the crude POC solution in step 6 of Basic Protocol 3.

Run the native-PAGE gel for POC purification

- 16. After the pre-run, disconnect the gel apparatus from the power supply and rinse the 10 wells with 1X TAE-Mg²⁺ buffer.
- 17. Load 50 μ L of prepared POC material into each well and 5 μ L of native running dye in the two outermost wells flanking the POC material.
- 18. Reconnect the gel apparatus to the power supply and start the gel running at 200 volts at 4 °C. Running time will vary with gel percentage and length of the oligonucleotide.

Isolation of POC material

 Recover the POC material from the gel as described in *Gel extraction* of Basic Protocol 1.
If the POC material is prepared on a small scale or is not detectable by UV, use SyberGold staining methods as described by manufacturer. If the POC material is stained with SyberGold and eluted as described in **Gel extraction**, conduct an **Ethanol precipitation** after step 29 of **Isolation of POC material** to remove any excess SyberGold.

- 20. To remove excess salts from the POC solution, prepare a NAP-10 column by washing the column with 10 column volumes of water.
- 21. Combine the purified POC solution in a 1.5 mL siliconized tube.
- 22. Add water to the tube until the final volume of the POC solution is 1 mL.
- 23. Add the 1 mL of POC to the top of the NAP-10 column.
- 24. Label four 1.5 mL siliconized tubes as elutions #1-4.
- 25. Place tube #1 below the NAP-10 column for collection purposes.
- 26. Add 0.5 mL of water to the top of the NAP-10 column and collect the output 0.5 mL in the 1.5 mL tube from step 25.
- 27. Repeat steps 25-26 three more times with tubes #2-4.
- 28. Lyophilize the four tubes to dryness.

If you notice tube #4 contains a gel like material, repeat steps 22-28.

29. Resuspend the POC material in 100 μ L of water and quantify using UV absorbance as described in Basic Protocol 1.

The POC material is quantified based on the oligonucleotide sequence since the contribution of absorbance at 254 nm caused by the amino acids is considered negligible compared to the absorbance of the DNA nucleobases.

30. Calculate the overall synthesis yield by taking the ratio of product obtained in step 29 (in moles) to the starting amount of pure DNA from Basic Protocol 2 (20 nanomoles). *The average yield of POC material after synthesis and purification by native-PAGE is 10% to 30%.*

5.7 Basic Protocol 5

Purification of Peptide-Oligonucleotide Conjugates using Reverse Phase-HPLC

As an alternative to urea- and native-PAGE purification, reverse phase-HPLC (RP-HPLC) can be used to purify POCs. Described below is a RP-HPLC protocol for the purification of POC material that has been slightly modified from a previously published method (4). Initial analysis of the POC material is performed by injecting a small amount of the crude POC material onto the HPLC column and running the complete gradient method to obtain an analytical chromatogram trace. Migration of the POC material is monitored using absorbance at 254 nm, and retention times will vary depending on the sequence composition of both the peptide and DNA polymer. Additional injections of the oligonucleotide and SMCC-oligonucleotide intermediate may be needed to determine which peak corresponds to the POC. Once the peaks in the analytical chromatogram trace have been assigned, either an analytical or semi-preparative HPLC column (depending on the amount of POC material) is used to purify the remaining POC material.

Materials

Crude POC material (Basic Protocol 3)

HPLC mobile phase A: 0.1 M ammonium acetate, pH 7.0

HPLC mobile phase B: acetonitrile (HPLC grade, Sigma)

Nanopure water

High-performance liquid chromatography (HPLC) system with:

Injector (autosampler preferred), sample loop, and syringe (for manual injections) Binary pumping system UV/Vis detector with wavelength detection between 200 and 300 nm Analytical column: reverse phase column (i.e., Source 5RPC ST 4.6/150, Amersham) Semi-preparative column: reverse phase column (i.e., Source 15RPC ST 4.6/100, Amersham) Automatic fraction collector (optional) Lyophilizer Disposable siliconized tubes (1.5 mL, Fisher)

Prepare HPLC instrument and RP-HPLC analytical injection

- 1. Connect the analytical RP-HPLC column to the HPLC instrument as directed by the manufacturer's handbook.
- 2. Turn the HPLC instrument on and set the UV detector to 254 nm.
- 3. Turn on the binary pumps and begin pumping 100% mobile phase B through the column at a flow rate of 0.8 mL/min for 5 minutes.
- 4. Change the solvent from 100% mobile phase B to 95% mobile phase A over 1 minute with a continuous flow rate of 0.8 mL/min. Flow 95% mobile phase A through the column until a flat baseline is achieved.
- 5. Program the HPLC gradient system to start at 95% mobile phase A and increase the percentage of mobile phase B over time as shown in Table 5.5.

Table 5.5. Gradient of Mobile Phase B for RP-

HPLC Purification of Crude POC Material^a

Elapse time (min)	Percentage mobile phase B ^b
0	5
5	5
30	30
31	100
35	100
36	5
40	5

^aGradient conditions are based on a 0.8 mL/min flow rate using an analytical Amersham Source 5RPC ST 4.6/150 column at room temperature with a 40 minute injection cycle. Mobile phase A: 0.1 M ammonium acetate, pH 7.0. Mobile phase B: acetonitrile. ^bPercentage is at elapsed time.

- 6. Start the gradient program while tracking the absorbance at 254 nm. Continue to do full gradient cycles until the baseline is flat for the entire gradient program.
- Dilute 5 μL of POC material from Basic Protocol 3 in 20 μL of 0.1 M ammonium acetate, pH 7.0.
- 8. Inject 25 μ L of POC sample from step 7 into the sample column of the HPLC instrument and elute the POC material by starting the gradient program described in Table 5.

Purify crude POC material using analytical or semi-preparative RP-HPLC

- 9. Depending on the amount of POC material, use an analytical or semi-preparative column accordingly.
- 10. Prepare the remaining 95 μL POC material from Basic Protocol 3 in 405 μL of 0.1 M ammonium acetate, pH 7.0.
- 11. Repeat step 8 five times by injecting 100 μ L of POC material from step 10 and running the same gradient listed in Table 5.5.

Elute the material at a flow rate of 0.8 mL/min when using an analytical column or 3 mL/min for semi-preparative column. Collect the assigned POC fraction as determined in the analytical chromatogram using an automatic fraction collector or manually by observing the chromatogram in real time.

12. Exchange the purified POC material into water using a Nap-10 column as described in steps 21-27 of *Isolation of POC material*.

- 13. Lyophilize the fractions to dryness.
- 14. Resuspend the POC material in 100 μ L of water and quantify using UV absorbance as described in Basic Protocol 1.

The POC material is quantified based on the oligonucleotide sequence since the contribution of absorbance at 254 nm by the amino acids is considered negligible compared to the absorbance of the nucleobases.

15. Calculate the overall synthesis yield by taking the ratio of product obtained in step 14 (in moles) to the starting amount of pure DNA from Basic Protocol 2 (20 nanomoles). The average yield of POC material after synthesis and purification by RP-HPLC is 30% to 50%.

5.8 Commentary

Background Information

Several methods exist to conjugate polypeptides to synthetic DNA and RNA (11-13). The most commonly used method of peptide-oligonucleotide construction is in-line solid phase synthesis. Using this technique, the peptide and oligonucleotide strands are synthesized sequentially using automated synthesizers. Typically, the peptide is synthesized first on a solid support, and a modified N-terminal or the natural C-terminal hydroxyl group is used to create a phosphodiester linkage between the peptide and the first nucleotide (1, 12). The remaining oligonucleotide sequence is then elongated using standard phosphoramidite chemistry on a DNA synthesizer. Although the in-line solid phase synthesis method is straightforward, the strategy is limited to amino acid sequences with chemical protecting groups that are compatible with oligonucleotide synthesis and deprotection chemistries (9).

Fragment conjugation is an alternative strategy to the solid phase method of creating POC molecules. Since this approach uses independently synthesized and deprotected peptides and oligonucleotides, fragment conjugation overcomes the need for compatible protecting groups and allows for the conjugation of longer polymers containing any combination of amino acids and nucleotides. Also, because the peptide and oligonucleotide polymers are commercially available, the fragment conjugation method is applicable to most laboratories.

The fragment conjugation method presented here utilizes the heterobifunctional crosslinking reagent SMCC. This method was originally reported by Harrison and

Balasubramanian in 1997 to characterize the hybridization properties of POCs (4, 14). The functionality of the SMCC molecule allows for a wide range of oligonucleotide and peptide sequences to be conjugated. However, one constraint of this fragment conjugation method is that the peptide sequence must contain only one cysteine residue. One added benefit of this strategy is that it is compatible with other chemical modifications including biotin, 2,4-dinitrophenyl (DNP), psoralen, puromycin, and fluorescein labels.

Applications

The protein recognition and cellular uptake capabilities of polypeptides, along with the base pair recognition of oligonucleotides, allows for a wide range of application for POC molecules (15,16). POCs were originally used for antisense technology where viral or cellular gene expression is controlled by hybridizing a small DNA sequence to the mRNA region that encodes a gene of interest. POCs are well suited for this type of application because the peptide portion of the molecule is able to recognize a particular cell type and transport the oligonucleotide sequence across the cell membrane where the oligonucleotide can selectively hybridize with mRNA and inhibit translation.

Since the original use of POC molecules as antisense reagents, many new applications have emerged that extend the use of POCs into the field of bionanotechnology. One effort focuses on the advancement of structural DNA nanotechnology in which the nucleic acid moiety of the POC molecules are used to direct individual peptides to specific locations on the surface of a DNA nanostructure. In the first example, the c-myc peptide was conjugated to a short 20-nt single-stranded DNA molecule that was complimentary in sequence to a DNA probe presented on the surface of a two-dimensional DNA nanoarray (17). When hybridized, the tiled array produced a nanoscale pattern of evenly spaced parallel lines separated by 64 nanometers. When probed with the cognate anti-c-myc antibody, the antibody specifically recognized the c-myc peptide displayed on the surface of the nanoarray. This technology of displaying peptides on the surfaces of DNA nanostructures was termed nanodisplay. A second example of nanodisplay recently appeared when POCs displayed on DNA nanotubes were used to template-direct the nucleation of gold nanoparticles from soluble chemical precursors. Transmission electron microscopy images demonstrated the formation of gold nanoparticles of a discrete size evenly spaced on the surface of the DNA nanotube (18). Together, these two examples demonstrate the utility of POC molecules as chemical reagents that can be used to precisely organize biological and inorganic materials on the nanoscale.

More recently, POC molecules were used to create bivalent protein affinity reagents, termed synbodies (19). This chemical design relies on the principle of multivalency in which peptides with affinity to non-overlapping epitopes on a protein surface can be used to create bivalent affinity reagents that bind their target proteins with significantly higher affinity than the individual peptides alone. Synbodies are constructed by synthesizing a small library of POCs using two peptides that are independently conjugated to complimentary strands of synthetic DNA. When the POCs are hybridized together, the two peptides are displayed on the dsDNA scaffold with varying spatial separation and orientation depending on the location of the amine modification within the DNA sequence. By positioning the peptides at different positions along the DNA backbone it is possible to quickly determine the optimal distance required to achieve high affinity binding. The synbody with the optimal peptide separation and orientation has ~1,000-fold improvement in binding affinity to the target protein when compared to the individual peptides that make up the bivalent reagent. The creation of synbodies demonstrates a unique application of POC molecules to produce nanoscale, bivalent protein affinity reagents.

Critical Parameters and Troubleshooting

When constructing POCs using the fragment conjugation protocol there are several critical parameters that need to be addressed. Of principal importance is the need for high purity (>95%) oligonucleotides and peptides that are free of unwanted amine and thiol impurities commonly found in solid-phase synthesis reagents. It is also important to be aware of denaturants and buffers that contain molecules that can react with SMCC.

Second, alternative purification strategies must be employed when scaling the reaction to concentrations greater than 20 nmoles per tube. In the described method, ethanol precipitation is used to remove the excess SMCC, however, this separation method is not sufficient for large-scale conjugation reactions. Alternatively, it is advised to use HPLC or urea-PAGE to purify the SMCC-oligonucleotide product from the excess SMCC for large-scale conjugations. Example chromatograms of SMCC-oligonucleotide and crude POC material using RP-HPLC is shown in Figure 5.6.

Validation of peptide-oligonucleotide conjugation can be performed using two different methods: gel mobility shift assay or mass spectrometry (MS). Depending on the amino acid and nucleotide composition of the POC, it can be difficult to crystallize the POC material using a matrix when performing matrix-assisted laser desorption ionization (MALDI) MS due to the different matrixes needed for the peptide and oligonucleotide polymers. For optimal matrix crystallization, it is best to try all of the following matrixes when performing MALDI-MS analysis of POC material; 4-hydroxy-α-cyano-cinnamic acid. sinapinic acid, 3-hydroxypicolinic acid. and 2,4,6trihydroxyacetophenone/diammonium salt (1). An example of MALDI-MS analysis of crude and pure POC material using sinapinic acid is shown in Figure 5.7. Electrospray ionization (ESI) MS (20) and analytical native-PAGE are alternative methods of analyzing POC material that do not require matrix optimization. An example of native-PAGE mobility shift analysis of POC material is shown in Figure 5.8.



Figure 5.6. Reverse phase-HPLC analysis of A) the c-myc peptide (EQKLISEEDLC), B) crude POC material (EQKLISEEDLC-ACCAGCTGTGCAGGCCTCGC), and C) 1 nanomole of SMCC-oligonucleotide spiked into the POC mixture.



Figure 5.7. MALDI mass spectroscopy analysis of A) SMCC-oligonucleotide and B) purified POC material.



Figure 5.8. Native-PAGE mobility shift analysis of POC material. Retardation of mobility is observed between the ssDNA (ACCAGCTGTGCAGGCCTCGC) (lane1), SMCC-oligonucleotide (SMCC-ACCAGCTGTGCAGGCCTCGC) hybridized to its complimentary strand (lane 2) and POC (EQKLISEEDLC-ACCAGCTGTGCAGGCCTCGC) hybridized to its complimentary strand (lane 3), thereby confirming the presence of the polypeptide (*17*). Image reproduced with permission from Wiley-VCH Verlag GmbH & Co.

Anticipated Results

As with all synthetic steps, the yield of the final product depends on the purity of the starting materials, with higher purity peptides and DNA giving higher overall yields. Using the methods described above, a yield of 2-6 nmoles of pure POC material can be obtained from native-PAGE purification and 6-10 nmoles from RP-HPLC purification when conducting a 20 nanomole conjugation reaction. The yield may vary depending on the polypeptide sequence and the location of the amine-modified nucleobase in the DNA sequence. To optimize the yield of the reaction, vary the ratio of oligonucleotide to SMCC during the first conjugation step and the peptide to SMCC-oligonucleotide during the second conjugation step.

Time Consideration

The synthesis and purification of peptide-oligonucleotide conjugates using either the native-PAGE or HPLC purification methods can be performed in 2-3 days. Additional material can be generated in the same time frame by conducting multiple 20 nanomole reactions in parallel.

4

5.9 References

- 1. Tung, C.-H. (2000) Preparation and application of peptide-oligonucleotide conjugates, *Bioconjugate Chemistry 11*, 605-618.
- 2. Singh, Y., Spinelli, N., and Defrancq, E. (2008) Chemical strategies for oligonucleotide-conjugates synthesis, *Curr. Org. Chem.* 12, 263-290.
- 3. Grandas, A., Marchan, V., Debethune, L., and Pedroso, E. (2007) Stepwise solidphase synthesis of nucleopeptides, *Curr. Protoc. Nucleic Acid Chem.* 4.22.1-4.22.54.
- 4. Harrison, J. G., and Balasubramanian, S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates, *Nucleic Acids Res.* 26, 3136-3145.
- 5. Sharma, S., and Yadav, B. R. (2009) Denaturing urea-polyacrylamide gel electrophoresis (PAGE) based microsatellite analysis, <u>http://www.protocol-online.org</u>.
- 6. Tournier, E. J. M., Wallach, J., and Blond, P. (1998) Sulfosuccinimidyl 4-(N-maleimidomethyl)-1-cyclohexane carboxylate as a bifunctional immobilization agent. Optimization of the coupling conditions, *Analytical Chimica Acta 361*, 33-44.
- 7. Brewer, C. F., and Riehm, J. P. (1967) Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins, *Anal. Biochem. 18*, 248-255.
- 8. Portela, C., Albericio, F., Eritja, R., Castedo, L., and Mascarenas, J. L. (2007) Dsoligonucleotide-peptide conjugates featuring peptides from the leucine-zipper region of Fos as switchable receptors for the oncoprotein Jun, *ChemBioChem* 8, 1110-1114.
- 9. Venkatesan, N., and Kim, B. H. (2006) Peptide Conjugates of Oligonucleotides:,Äâ Synthesis and Applications, *Chemical Reviews 106*, 3712-3761.
- Moroder, H., Steger, J., Graber, D., Fauster, K., Trappl, K., Marquez, V., Polacek, N., Wilson, D. N., and Micura, R. (2009) Non-hydrolyzable RNA-peptide conjugates: A powerful advance in the synthesis of mimics for 3 prime-peptidyl tRNA termini, *Angew. Chem. 121*, 4116-4120.

- 11. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the in vitro seleciton of peptides and proteins, *Proc. Natl. Acad. Sci. USA 94*, 122297-112302.
- 12. Haralambidis, J., Duncan, L., Angus, K., and Tregear, G. W. (1990) The synthesis of polyamide-oligonucleotide conjugate molecules, *Nucleic Acids Res.* 18, 493-499.
- 13. Tung, C.-H., and Stein, S. (2000) Preparation and applications of peptideoligonucleotide conjugates, *Bioconj. Chem. 11*, 605-618.
- 14. Harrison, J. G., and Balasubramanian, S. (1997) A convenient synthetic route to oligonucleotide conjugates, *Bioorg. Med. Chem.Letters* 7, 1041-1046.
- Leonetti, J. P., Degols, G., and Lebleu, B. (1990) Biological activity of oligonucleotide-poly(L-lysine) conjugates: mechanism of cell uptake, *Bioconj. Chem.* 1, 149-153.
- 16. Soukchareun, S., Tregear, G. W., and Haralambidis, J. (1995) Preparation and characterization of antisense oligonucleotide-peptide hybrids containing viral fusion peptides, *Bioconj. Chem.* 6, 43-53.
- 17. Williams, B. A. R., Lund, K., Liu, Y., Yan, H., and Chaput, J. C. (2007) Selfassembled peptide nanoarrays: an approach to studying protein-protein interactions, *Angew. Chem. Int. Ed.* 46, 3051-3054.
- 18. Stearns, L. A., Chhabra, R., Sharma, J., Liu, Y., Petuskey, W. T., Yan, H., and Chaput, J. C. (2009) Template-directed nucleation and growth of inorganic nanoparticles on DNA scaffolds, *Angew. Chem. Int. Ed.* 48, 8494-8496.
- 19. Williams, B. A. R., Diehnelt, C. W., Belcher, P., Greving, M., Woodbury, N. W., Johnston, S. A., and Chaput, J. C. (2009) Creating protein affinity reagents by combining peptide ligands on synthetic DNA scaffolds, *J. Am. Chem. Soc. 131*, 17233-17241.
- 20. Tengvall, U., Auirola, S., Antopolsky, M., Azhaev, A., and Beigelman, L. (2003) Characterization of antisense oligonucleotide-peptide conjugates with negative ionization electrospray mass spectrometry and liquid chromatography-mass spectrometry, *J. Pharma. Biomed. Anal.* 32, 581-590.

CHAPTER 6

CONCLUDING REMARKS

6.1 Abstract

The work presented in the preceding chapters explores the use of DNA as a molecular tool for organizing and capturing protein targets. Each chapter describes different applications of DNA in nanotechnology and molecular biology. Chapter 3 demonstrates the capabilities of DNA to specifically capture proteins with small chemical additions termed post-translational modifications. The research described in chapter 2 utilizes peptide-oligonucleotide conjugates (POCs) to specifically position proteins with nanometer precision using DNA nanoarrays, while chapter 4 uses POC molecules to create high affinity bivalent protein capture reagents. The method to synthesize these POC molecules was a core component of my Ph.D. and is described in detail in chapter 5.

Although each project described in the preceding chapters resulted in successful publication, many critical parameters were noted that allowed these projects to reach fruition. The conclusions and some key observations from each research project are summarized here.

6.2 Conclusions

The research described in chapter 2 sought to expand the utility of traditional DNA nanoarrays by developing a universal method for displaying functional polypeptides at discrete locations (1). In this research, the c-myc peptide was conjugated to single-stranded DNA strand that was complimentary in sequence to a DNA probe strand on the D-tile of the ABCD nanoarray. When hybridized, the ABCD array was designed to display one c-myc peptide on every D-tile of the array. The atomic force microscopy (AFM) images revealed repeating parallel lines of c-myc peptide separated by 64 nanometers, correlating precisely with our design. When probed with the cognate anti-c-myc antibody, an increase in the height profile was observed where the c-myc peptides were positioned, suggesting the antibody bound the c-myc peptide and was specifically organized with nanometer precision. This research was the first demonstration of POCs to organize proteins with nanometer precision on DNA nanoarrays.

Chapter 3 describes the discovery and characterization of ssDNA as a binding reagent (aptamer) with high affinity and specificity to a histone protein with a common post-translational modification (PTM) (*3*). In this work, four rounds of *de novo* evolution from a naïve DNA library were employed against histone H4 peptide with an acetyl modification on lysine residue 16 (H4K16-Ac). Each round of selection consisted of a positive and negative selection step to enrich sequences with affinity for H4K16-Ac and not H4, respectively. The selection resulted in multiple, unique aptamer sequences that did not converge on to a single sequence. Three of the aptamer sequences were

determined to have low nanomolar binding affinity to H4K16-Ac by affinity capillary electrophoresis. This result suggests that a selection does not need to converge on a family of sequences to be successful. The highest affinity aptamer (4.20) and a commercially available anti-H4K16-Ac antibody were tested by surface plasmon resonance for affinity to H4K16-Ac and specificity over H4 and H4K8-Ac. The results showed the affinity of the aptamer and antibody were comparable, however, the specificity of the aptamer was >150 fold better in comparison to the commercial antibody.

The knowledge gained from the research conducted in chapter 2 led us to apply POC molecules to the development of high affinity, bivalent protein capture reagents. Chapter 4 describes a novel use of POC molecules through a process described as combinatorial examination of ligands and linkers (CELL). Through this process, a library of independently conjugated POC molecules are constructed to complementary strands of DNA using peptide ligands that bind non-overlapping sites on a given target protein. When hybridized, two POCs come together to make bivalent affinity reagents, termed synbodies, with the optimal pair of POCs generating a high affinity and specificity reagent similar to a traditional antibody. In chapter 4, low nanomolar affinity synbodies were developed to Gal80 and transferrin protein. The synbodies with the optimal peptide pair and peptide pair distance had ~1,000-fold increase in binding affinity over the individual ligands and functioned in standard ELISA and pull-down assays.

The relatively simple chemical method of linking peptides to DNA, termed POCs, used in making peptide functional DNA nanoarrays and synbody molecules is thoroughly described in chapter 5. Cystiene-containing peptides are covalently linked to aminemodified ssDNA using a heterobifunctional crosslinking molecule called succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC). The fragment conjugation method presented in chapter 5 uses a two-step reaction technique with a single purification step. Depending on the method of purification (native PAGE or HPLC), yield can range between 10% and 50%.

6.3 Key Observations

Each project presented in this dissertation resulted in successful advancements in its respective field, however key observations were noted during the process that may not have appeared in the resulting publication. As in chapter 2, although the research demonstrated the specific positioning of the anti-c-myc antibody with nanometer precision, it was observed that not every peptide displayed on the array bound an antibody. This could be caused by steric hindrance between neighboring peptides or an interaction between the antibody and the AFM tip that prevents the antibody from binding all the peptides. It is hypothesized that the latter of the two is correct due to observations of the antibody sticking to the AFM tip during measurements of the ABCD array when incubated with the antibody. Blocking the AFM tip with an agent such as bovine serum albumin (BSA) or dialyzing the antibody into a compatible buffer prior to incubation of the antibody with the array could reduce the nonspecific interaction.

Some difficulty was encountered when assaying the affinity and specificity of histone peptides with post-translational modifications as described in chapter 3. Due to the complexity of synthesizing a protein with an acetyl PTM, amino acid residues 6-20 of histone H4 with the PTM at K16 or K8 were used to represent the N-terminal tail of the protein. The synthesis of acetylated polypeptides is very costly, therefore, we were not able to add extra labels or tags to the peptide in order to assays the aptamers affinity using standard techniques such as ELISA, fluorescent anisotropy, or spin filter-binding assay. For this reason, we employed the label free technique of surface plasmon resonance to measure the affinity of the aptamer and antibody to the H4K16-Ac peptide.

In these experiments, we immobilized the biotinylated aptamer on a streptavidin coated CM5 chip. A dramatic difference in binding affinity was observed depending on the location of the biotin in the aptamer sequence. Three different aptamer constructs were tested at the 5' or 3' end of the sequences, with and without the 5' and 3' primer regions. The best binding affinity resulted when the aptamer was biotinylated at the 5' end of the core sequence without the primer regions. This result suggested that the tertiary structure of the aptamer was altered depending on the location of the biotin and the presence of the primer regions. To avoid this issue, it might be advantageous to label the peptide sequence with (even with the added cost) a tag instead of labeling the aptamer.

During the CELL process described in chapter 4, an interesting trend was observed; the synbody molecules had an ~1,000 fold affinity enhancement over the starting peptides. The peptides identified to bind non-overlapping sites on Gal80 had ~5 μ M affinities, while the resulting synbody molecule had an affinity of 5nM. The starting transferrin peptide ligands had affinities of 150 μ M and 50 μ M, while the resulting synbody had an affinity of ~100 nM, which is an ~1,000 affinity improvement over the average affinity of the two peptide ligands. This trend suggests that better affinity ligands will result in better affinity synbody molecules, therefore, developing methods of identifying better ligands will improve the CELL outcome.

In the protocol presented in chapter 5, the most common issues encountered when making POCs are caused by the presence of impurities in the starting oligonucleotide or polypeptide solutions. These impurities compete with the oligonucleotide or peptide in the reaction with SMCC resulting in undesired side products. It is essential to use purified oligonucleotide and polypeptide material with >95% purity when generating POCs. Also, scaling up the reaction from 20 nanomoles to larger amounts has not been successfully achieved using the method described in chapter 5. It is hypothesized that ethanol precipitation step that is used to separate the oligonucleotide-SMCC product from the unreacted SMCC is not efficient at scales larger than 20 nanomoles. Therefore, for larger scales synthesis of POC material, it is suggested to use HPLC to separate excess SMCC away from the oligonucleotide-SMCC product. Alternatively, multiple reactions of 20 nanomole each can be conducted simultaneously in separate tubes.

6.4 Future Direction

The results in chapter 2 demonstrate that peptides displayed on DNA nanoarrays remain functional and accessible to exogenously added proteins. This suggests that the method could be used to generate tailor-made protein nanoarrays by discretely positioning functional proteins in nanometer intervals utilizing DNA nanostructures as a scaffold. This method of organizing proteins with nanometer precision may be a useful tool for studying protein-protein interactions or enzymatic pathways at local concentrations similar to that of a cellular environment. Alternatively, displaying peptides capable of binding inorganic material on DNA nanoarrays could have application in nanoelectronics or DNA based computing (2). The next step in this field would be to demonstrate the addressability of the DNA nanoarrays by displaying different peptide sequences at discrete locations on the array and probing each one individually with its cognate antibody or binding partner.

The work demonstrated in chapter 3 shows that aptamers can be evolved to bind proteins with PTMs that are specific to both the modification and position of the modification in the protein sequence. It is continually acknowledged that highly specific affinity reagents against PTMs are needed to understand and characterize the role of PTMs in nature and disease states (4). In the near future, the field of medical diagnostics will surely include the detection of PTMs when reading the epigenetic code. Accurate detection of these modifications will only be possible if there are affinity reagents capable of discriminating between PTMs and the position of the modification. With the results presented in chapter 3, it is suggested that a library of aptamers capable of specifically detecting PTMs could be used to replace antibodies in disease-profiling medical diagnostics.

The CELL process described in chapter 4 is a powerful technique capable of generating artificial affinity reagents termed synbodies. This process can generate synbodies with similar binding characteristics to typical antibodies, but at a much lower development cost and in a shorter timeframe. Characteristics of the synbody molecules suggest that they could replace antibodies in certain applications such as ELISA, affinity purification, pull-down, and immunofluorescence assays. In particular, the use of synbodies in mildly denaturing conditions, such as surface applications, could be of great interest due to the inherent stability problem that encompasses antibodies. Further demonstration of the CELL process against more interesting protein targets, such as cancer biomarkers, and in a high-throughput manner is needed to fully realize the potential of the CELL technology. Additionally, rapid development of synbodies to many protein targets in a high-throughput method would greatly increase the utility of the CELL method. One approach to high-throughput identification of synbodies is to screen the bivalent reagents using crude POC molecules in an ELISA type assay. The highest binding reagent can then be recreated with high purity for further analysis.

Although great advancements were made to each of the respective fields, much more research is needed to translate these ideas to realistic applications. Each of the described research fields is fairly young in the scientific community and it may take many more years to realize the full impact of the contributions described in this dissertation. With the use of high-throughput technology, more examples or functionality could be applied to each field and advance the corresponding areas of research to a more useful position beyond what has been discussed. However incremental our research has been, this work clearly demonstrates the versatility of DNA as a molecular tool for nanotechnology and molecular biology.

6.5 References

- 1. Williams, B. A. R., Lund, K., Liu, Y., Yan, H., and Chaput, J. C. (2007) Selfassembled peptide nanoarrays: an approach to studying protein-protein interactions, *Angew. Chem. Int. Ed.* 46, 3051-3054.
- 2. Williams, B. A. R., Lin, L., Lindsay, S. M., and Chaput, J. C. (2009) Evolution of a histone H4-K16 acetyl-specific DNA aptamer, *J. Am. Chem. Soc. 131*, 6330-6331.
- 3. Stearns, L. A., Chhabra, R., Sharma, J., Liu, Y., Petuskey, W. T., Yan, H., and Chaput, J. C. (2009) Template-directed nucleation and growth of inorganic nanoparticles on DNA Scaffolds, *Angew. Chem. Int. Ed.* 48, 8494-8496.
- 4. Blow, N. (2007) The generation game, *Nature* 447, 741-744.

REFERENCES

- 1. Endo, M., and Sugiyama, H. (2009) Chemical approaches to DNA nanotechnology, *ChemBioChem 10*, 2420-2443.
- 2. Carell, T., Behrens, C., and Gierlich, J. (2003) Electrontransfer through DNA and metal-containing DNA, *Org. Biomol. Chem. 1*, 2221-2228.
- 3. Cheng, A. K. H., Sen, D., and Yu, H.-Z. (2009) Design and testing of aptamerbased electrochemical biosensors for proteins and small molecules, *Bioelectrochem.* 77, 1-12.
- 4. Seeman, N. C. (1982) Nucleic acid junctions and lattices, J. Theor. Biol. 99, 237-247.
- 5. Aldaye, F. A., Palmer, A. L., and Sleiman, H. F. (2008) Assembling materials with DNA as the guide, *Science 321*, 1795-1799.
- 6. Park, S. H., Yin, P., Liu, Y., Reif, J. H., LaBean, T. H., and Yan, H. (2005) Programmable DNA self-assemblies for nanoscale organization of ligands and proteins, *Nano Lett.* 5, 729-733.
- 7. Liu, Y., Lin, C. X., Li, H. Y., and Yan, H. (2005) Aptamer-directed self-assembly of protein arrays on a DNA nanostructure, *Angew. Chem. Int. Ed.* 44, 4333-4338.
- 8. Bunka, D. H. J., and Stockley, P. G. (2006) Aptamers come of age- at last, *Nat. Rev. Micro.* 4, 588-596.
- 9. Lee, J. F., Hesselberth, J. R., Meyers, L. A., and Ellington, A. D. (2004) Aptamer Database, *Nucleic Acids Res.* 32, D95-100.
- 10. Singh, Y., Spinelli, N., and Defrancq, E. (2008) Chemical strategies for oligonucleotide-conjugates synthesis, *Curr. Org. Chem.* 12, 263-290.
- 11. Williams, B. A. R., Lund, K., Liu, Y., Yan, H., and Chaput, J. C. (2007) Selfassembled peptide nanoarrays: An approach to studying protein-protein interactions, *Angew. Chem. Int. Ed.* 46, 3051-3054.

- 12. Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands, *Nature 346*, 818-822.
- 13. Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science 249*, 505-510.
- 14. Baugh, C., Grate, D. r., and Wilson, C. (2000) 2.8 Å crystal structure of the malachite green aptamer, *J. Mol. Biol.* 301, 117-128.
- 15. Jenison, R. D., Gill, S. C., Pardi, A., and Polisky, B. (1994) High-resolution molecular discrimination by RNA, *Science 263*, 1425-1429.
- 16. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications, *Nat. Biotechnol. 21*, 255-261.
- 17. Shrogen-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 acetyation controls chromatin structure and protein interactions, *Science 311*, 844-847.
- 18. Harp, J. M., Hanson, B. L., Timm, D. E., and Bunick, G. J. (2000) Asymmetries in the nucleosome core particle at 2.5 A resolution, *Acta Crystall. Section D 56*, 1513-1534.
- 19. Mairal, T., Ozalp, V. C., Sanchez, P. L., Mir, M., Katakis, I., and O'Sullivan, C. K. (2008) Aptamers: molecular tools for analytical applications, *Anal. Bioanal. Chem.* 390, 989-1007.
- 20. Gilbert, B. A., Sha, M., Wathen, S. T., and Rando, R. R. (1997) RNA aptamers that specifically bind to a K Ras-derived farnesylated peptide, *Bio. Med. Chem. 5*, 1115-1122.
- 21. Williams, B. A. R., Lin, L., Lindsay, S. M., and Chaput, J. C. (2009) Evolution of a histone H4-K16 acetyl-specific DNA aptamer, *J. Am. Chem. Soc. 131*, 6330-6331.

- 22. Hu, S., Loo, J. A., and Wong, D. T. (2006) Human body fluid proteome analysis, *Proteomics* 6, 6326-6353.
- Butler, J. E., Ni, L., Brown, W. R., Joshi, K. S., Chang, J., Rosenberg, B., and Voss, E. W. (1993) The immunochemistry of sandwich ELISAs-VI. Greater than 90% of monoclonal and 75% of polyclonal antifluorescyl capture antibodies (CAbs) are denatured by passive adsorption, *Mol. Immunol.* 30, 1165-1175.
- 24. Access Excellence @ the National Health Museum, http://www.accessexcellence.org/RC/VL/GG/monoclonal.php
- 25. Alamargo, J. C., Quintero-Hernandez, V., Ortiz-Leon, M., Velandia, A., Smith, S. L., and Becerril, B. (2006) Design and validation of a synthetic V_H repertoire with tailored diversity for protein recognition, *J. Mol. Rec.* 19, 413-422.
- 26. Binz, H. K., Amstutz, P., and Pluckthun, A. (2005) Engineering novel binding proteins from nonimmunoglobulin domains, *Nat. Biotechnol. Rev.* 23, 1257-1268.
- 27. Hosse, R. J., Rothe, A., and Power, B. E. (2006) A new generation of protein display scaffolds for molecular recognition, *Protein Science 15*, 14-27.
- 28. Reddy, M. M., Bachhawat-Sikder, K., and Kodadek, T. (2004) Transformation of low-affinity lead compounds into high-affinity protein capture agents, *Chem. Biol.* 11, 1127-1137.
- 29. Searle, M. S., Sharman, G. J., Groves, P., Benhamu, B., Beauregard, D. A., Westwell, M. S., Dancer, R. J., Maguire, A. J., Try, A. C., and Williams, D. H. (1996) Enthalpic (electrostatic) contribution to the chelate effect: a correlation between ligand binding constant and a specific hydrogen bond strength in complexes of glycopeptide antibiotics with cell wall analogues, J. Am. Chem. Soc. 1, 2781-2786.
- 30. Mammen, M., Choi, S.-K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design in use of multivalent ligands and inhibitors, *Angew. Chem. Int. Ed.* 37, 2754-2794.

- 31. Rao, J., Lahiri, J., Isaacs, L., Weis, R. M., and Whitesides, G. M. (1998) A trivalent system from vancomycin-D-Ala-D-Ala with higher affinity than avidinbiotin, *Science 280*, 708-711.
- 32. Shewmake, T. A., Solis, F. J., Gillies, R. J., and Caplan, M. R. (2008) Effects of linker length and flexibility on multivalent targeting, *Biomacromol. 9*, 3057-3064.
- 33. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR, *Science* 274, 1531-1534.
- 34. Zhou, J. Z. (2008) Structure-directed combinatorial library design, Curr. Opin. Chem. Biol. 12, 379-385.
- 35. Sridhar, J., Wei, Z.-L., Nowak, I., Lewin, N. E., Ayres, J. A., Pearce, L. V., Blumberg, P. M., and Kozikowski, A. P. (2003) New bivalent PKC ligands linked by a carbon spacer: Enhancement in binding affinity, *J. Med. Chem.* 46, 4196-4204.
- 36. Maly, D. J., Choong, I. C., and Ellman, J. S. (2000) Combinatorial target-guided ligand assembly: Identification of potent subtype-selective c-Src inhibitors, *Proc. Natl. Acad. Sci. USA* 97, 2419-2424.
- 37. Brase, S., Dahmen, S., and Lormann, M. E. P. (2003) Multifunctional linkers as an efficient tool for the synthesis of diverse small molecule libraries: The triazene anchors, *Combin. Chem.*, *Pt B*, pp 127-150, Academic Press Inc, San Diego.
- 38. Reyes, S. J., and Burgess, K. (2006) Heterovalent selectivity and the combinatorial advantage, *Chem. Soc. Rev.* 35, 416-423.
- 39. Wright, D., and Usher, L. (2001) Multivalent binding in the design of bioactive compounds, *Curr. Org. Chem. 5*, 1107.
- 40. Kramer, R. H., and Karpen, J. W. (1998) Spanning binding sites on allosteric proteins with polymer-linked ligand dimers, *Nature 395*, 710-713.
- 41. Susuki, Y., Hojo, M., Okazaki, I., Kamata, H., Sasaki, M., Maeda, M., Nomizu, M., Yamamoto, Y., Nakagawa, S., Mayaumi, T., and Kawasaki, K. (2002)

Preparation and biological activities of a bivalent poly(ethylene glycol) hybrid containing an active site and its synergistic site of fibronectin, *Chem. Pharmaceut. Bull.* 50, 1229-1232.

- Schaffer, L., Brissette, R. E., Spetzler, J. C., Pillutla, R. C., Ostergaard, S., Lennick, M., Brandt, J., Fletcher, P. W., Danielsen, D. M., Hsiao, K.-C., Anderson, A. S., Dedova, O., Ribel, U., Hoeg-Jensen, T., Hansen, P. H., Blume, A. J., Markussen, J., and Goldstein, N. I. (2003) Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks, *Proc. Natl. Acad. Sci. USA 8*, 4435-4439.
- 43. Kim, Y., Cao, Z., and Tan, W. (2008) Molecular assembly for high-performance bivalent nucleic acid inhibitor, *Proc. Natl. Acad. Sci. USA 105*, 5664-5669.
- 44. Tian, L., and Heyduk, T. (2008) Bivalent ligands with long nanometer-scale flexible linkers, *Biochem.* 48, 264-275.
- 45. Cairo, C. W., Gestwicki, J. E., Kanai, M., and Kiessling, L. L. (2002) Control of multivalent interactions by binding epitope density, *J. Am. Chem. Soc. 124*, 1615-1619.
- 46. Roy, R. (1996) Syntheses and some applications of chemically defined multivalent glycoconjugates, *Curr. Opin. Struct. Biol.* 6, 692-702.
- 47. Schluter, A. D., and Rabe, J. P. (2000) Dendronized polymers: synthesis, characterization, assembly at interfaces, and manipulation, *Angew. Chem. Int. Ed.* 39, 864-883.
- 48. Crespo, L., Sancilmens, G., Pons, M., Giralt, E., Royo, M., and Albericio, F. (2005) Peptide and amide bond-contianing dendrimers, *Chem. Rev.* 105, 1663-1381.
- 49. Kim, Y., and Zimmerman, S. C. (1998) Applications of dendrimers in bio-organic chemistry, *Curr. Opin. Chem. Biol.* 2, 733-742.
- 50. Helms, B. A., Reulen, S. W. A., Nijhuis, S., Graaf-Heuvelmans, P. T. H. M. d., Merkx, M., and Meijer, E. W. (2009) High-affinity peptide-based collagen

targeting using synthetic phage mimics: From phage display to dendrimer display, J. Am. Chem. Soc. 131, 11683-11685.

- 51. Tam, J. P., Lu, Y.-A., and Yang, J.-L. (2002) Antimicrobial dendrimeric peptides, *Europ. J. Biochem.* 269, 923-932.
- 52. Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., Read, R. J., and Bundle, D. R. (2000) Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands, *Nature 403*, 669-672.
- 53. Chen, Y., and Liu, Y. (2010) Cyclodextrin-based bioactive supramolecular assemblies, *Chem. Soc. Rev.* 39, 495-505.
- 54. Benito, J. M., Gomez-Garcia, M., Ortiz Mellet, C., Baussanne, I., Defaye, J., and Garcia Fernandez, J. M. (2004) Optimizing saccharide-directed molecular delivery to biological receptors: design, synthesis, and biological evaluation of glycodendrimer, B-cyclodextrin conjugates, J. Am. Chem. Soc. 126, 10355-10363.
- 55. Perez-Balderas, F., Morales-Sanfrutos, J., Hernandez-Mateo, F., Isac-Garcla, J., and Santoyo-Gonzalez, F. (2009) Click multivalent domogeneous neoglycoconjugates - Synthesis and evaluation of their binding affinities, *Europ.* J. Org. Chem. 15, 2441-2453.
- 56. Nasr, K., Pannier, N., Frangioni, J. V., and Maison, W. (2008) Rigid multivalent scaffolds based on adamantane, *J. Org. Chem.* 73, 1056-1060.
- 57. Fan, E., Zhang, Z., Minke, W. E., Hou, Z., Verlinde, C. L. M. J., and Hol, W. G. J. (2000) High-Affinity pentavalent ligands of escherichia coli heat-labile enterotoxin by modular structure-based design, *J. Am. Chem. Soc.* 122, 2663-2664.
- 58. Zhang, Z., Liu, J., Verlinde, C. L. M. J., Hol, W. G. J., and Fan, E. (2004) Large cyclic peptides as cores of multivalent ligands: Application to inhibitors of receptor binding by cholera toxin, *J. Org. Chem.* 69, 7737-7740.
- 59. Trouche, N., Wieckowski, S. b., Sun, W., Chaloin, O., Hoebeke, J., Fournel, S., and Guichard, G. (2007) Small multivalent architectures mimicking homotrimers
of the TNF superfamily member CD40L: Delineating the relationship between structure and effector function, J. Am. Chem. Soc. 129, 13480-13492.

- 60. Turner, D. J., Ritter, M. A., and George, A. J. T. (1997) Importance of the linker in expression of single-chain Fv antibody fragments: optimisation of peptide sequence using phage display technology, *J. Immunol. Meth.* 205, 43-54.
- 61. Todorovska, A., Roovers, R. C., Dolezal, O., Kortt, A. A., Hoogenboom, H. R., and Hudson, P. J. (2001) Design and application of diabodies, triabodies and tetrabodies for cancer targeting, *J. Immunol. Meth.* 248, 47-66.
- 62. Kortt, A. A., Dolezal, O., Power, B. E., and Hudson, P. J. (2001) Dimeric and trimeric antibodies: high avidity scFvs for cancer targeting, *Biomol. Eng. 18*, 95-108.
- 63. Volkel, T., Korn, T., Bach, M., Muller, R., and Kontermann, F. R. (2001) Optimized linker sequences for the expression of monomeric and dimeric bispecific single-chain diabodies, *Protein Eng.* 14, 815-823.
- 64. Atwell, J. L., Breheney, K. A., Lawrence, L. J., McCoy, A. J., Kortt, A. A., and Hudson, P. J. (1999) scFv multimer of the anti-neuramindase antibody NC10: length ot the linker between VH and VL domains dictate precisely the transition between diabodies and triabodies, *Protein Eng. 12*, 597-604.
- 65. Neri, D., Momo, M., Prosero, T., and Winter, G. (1995) High-affinity antigen bidning by chelating recombinant antibodies (CRAbs), *J. Mol. Biol. 246*, 367-373.
- 66. Silverman, J., Liu, Q., Lu, Q., Bakker, A., To, W., Duguay, A., Alba, B. M., Smith, R., Rivas, A., Li, P., Le, H., Whitehorn, E., Moore, K. W., Swimmer, C., Perlroth, V., Vogt, M., Kolkman, J., and Stemmer, W. P. C. (2005) Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains, *Nat. Biotechnol.* 23, 1556-1561.
- Kazantsev, A., Walker, H. A., Slepko, N., Bear, J. E., Preisinger, E., Steffan, J. S., Zhu, Y.-Z., Gertler, F. B., Housman, D. E., Marsh, J. L., and Thompson, L. M. (2002) A bivalent Huntingtin binding peptide suppresses polyglutamine aggregation and pathogenesis in Drosophila, *Nature Genetics* 30, 367-376.

- 68. Cowburn, D., Zheng, J., Xu, Q., and Barany, G. (1995) Enhanced affinities and specificities of consolidated ligands for the Src homology (SH)3 and SH2 domains of abelson protein-tyrosine kinase, J. Biol. Chem. 270, 26738-26741.
- 69. Reddy, M. M., Bachhawat-Sikder, K., and Kodadek, T. (2004) Transformation of low-affinity lead compounds into high-affinity protein capture agents, *Proc. Natl. Acad. Sci. USA 11*, 1127-1137.
- Handl, H. L., Sankaranarayanan, R., Josan, J. S., Vagner, J., Mash, E. A., Gillies, R. J., and Hruby, V. J. (2007) Synthesis and evaluation of bivalent NDP--MSH(7) peptide ligands for binding to the human melanocortin receptor 4 (hMC4R), *Bioconj. Chem. 18*, 1101-1109.
- 71. Shrivastava, A., Wronski, M. A., Sato, A. K., Dransfield, D. T., Sexton, D., Bogdan, N., Pillai, R., Nanjappan, P., Song, B., Marinelli, E., DeOliveira, D., Luneau, C., Devlin, M., Muruganandam, A., Abujoub, A., Connelly, G., Wu, Q. L., Conley, G., Chang, Q., Tweedle, M. F., Ladner, R. C., Swenson, R. E., and Nunn, A. D. (2005) A distinct strategy to generate high-affinity peptide binders to receptor tyrosine kinases, *Protein Eng. Des. Sel. 18*, 417-424.
- 72. Toth, K., Sauermann, V., and Langowski, J. (1998) DNA curvature in solution measured by fluorescence resonance energy transfer, *Biochem.* 37, 8173-8179.
- 73. Chaudri, Z. N., Bartlet-Jones, M., Panayotou, G., Klonisch, T., Roitt, I. M., Lund, T., and Delves, P. J. (1999) Dual specificity antibodies using a double-stranded oligonucleotide bridge, *Fed. Soc. Biochem. Mol. Biol.* 450, 23-26.
- 74. Sprinz, K. I., Tagore, D. M., and Hamilton, A. D. (2005) Self-assembly of bivalent protein-binding agents based on oligonucleotide-linked organic fragments, *Bioorg. Med. Chem. Letters* 15, 3908-3911.
- 75. Paar, J. M., Harris, N. T., Holowka, D., and Baird, B. (2002) Bivalent ligands with rigid double-stranded DNA spacers reveal structural constraints on signaling by FcεRI, *J. Immunol.* 169, 856-864.
- 76. Sil, D., Lee, J. B., Luo, D., Holowka, D., and Baird, B. (2007) Trivalent ligands with rigid DNA spacers reveal structural requirements for IgE receptor signaling in RBL mast cells, *ACS Chem. Biol.* 2, 674-684.

77. Muller, J., Wulffen, B., Potzsch, B., and Mayer, G. (2007) Multidomain targeting generates a high-affinity thrombin-inhibiting bivalent aptamer, *ChemBioChem* 8, 2223-2226.

•

- 78. Mayer, G., Muller, J., Mack, T., Freitag, D. F., Hover, T., Potzsch, B., and Heckel, A. (2009) Differential regulation of protein subdomain activity with caged bivalent ligands, *ChemBioChem 10*, 654-657.
- 79. Rinker, S., Ke, Y., Liu, Y., Chhabra, R., and Yan, H. (2008) Self-assembled DNA nanostructures for distance dependent multivalent ligand-protein binding, *Nat. Nanotechnol.* 3, 418-422.
- 80. Williams, B. A. R., Diehnelt, C. W., Belcher, P., Greving, M., Woodbury, N. W., Johnston, S. A., and Chaput, J. C. (2009) Creating protein affinity reagents by combining peptide ligands on synthetic DNA scaffolds, *J. Am. Chem. Soc. 131*, 17233-17241.
- 81. Seeman, N. C. (1982) Nucleic acid junctions and lattices, J. Theor. Biol. 99, 237-247.
- 82. Chen, J., and Seeman, N. C. (1991) The Synthesis From Dna Of A Molecule With The Connectivity Of A Cube, *Abst. Pap. Am. Chem. Soc. 202*, 162-BIOL.
- 83. Mueller, J. E., Du, S. M., and Seeman, N. C. (1991) Design And Synthesis Of A Knot From Single-Stranded-Dna, J. Am. Chem. Soc. 113, 6306-6308.
- 84. Zhang, Y. W., and Seeman, N. C. (1994) A Truncated Octahedron Built From Dna Using Solid-Support Methodology, *Biophys. J.* 66, A24-A24.
- 85. Mao, C., Sun, W., and Seeman, N. C. (1997) Assembly of Borromean rings from DNA, *Nature 386*, 137-138.
- 86. Winfree, E., Liu, F. R., Wenzler, L. A., and Seeman, N. C. (1998) Design and self-assembly of two-dimensional DNA crystals, *Nature 394*, 539-544.

- 87. Yan, H., Park, S. H., Finkelstein, G., Reif, J. H., and LaBean, T. H. (2003) DNAtemplated self-assembly of protein arrays and highly conductive nanowires, *Science 301*, 1882-1884.
- 88. Lund, K., Liu, Y., Wang, H., Lindsay, S., Yan, H. (2005) Self-assembly of a finite size addressable DNA Nanoarrays, In preparation.
- 89. Le, J. D., Pinto, Y., Seeman, N. C., Musier-Forsyth, K., Taton, T. A., and Kiehl, R. A. (2004) DNA-templated self-assembly of metallic nanocomponent arrays on a surface, *Nano Letters* 4, 2343-2347.
- 90. Li, H. Y., Park, S. H., Reif, J. H., LaBean, T. H., and Yan, H. (2004) DNAtemplated self-assembly of protein and nanoparticle linear arrays, *J. Am. Chem. Soc. 126*, 418-419.
- 91. Xiao, S. J., Liu, F. R., Rosen, A. E., Hainfeld, J. F., Seeman, N. C., Musier-Forsyth, K., and Kiehl, R. A. (2002) Selfassembly of metallic nanoparticle arrays by DNA scaffolding, *J. Nanopart. Res.* 4, 313-317.
- 92. Harrison, J. G., and Balasubramanian, S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates, *Nucleic Acids Res. 26*, 3136-3145.
- 93. Yan, H., LaBean, T. H., Feng, L. P., and Reif, J. H. (2003) Directed nucleation assembly of DNA tile complexes for barcode-patterned lattices, *Proc. Nat. Acad. Sci. U. S. A. 100*, 8103-8108.
- 94. Rothemund, P. W. K. (2006) Folding DNA to create nanoscale shapes and patterns, *Nature 440*, 297-302.
- 95. Hober, S., and Uhlen, M. (2008) Human protein atlas and the use of microarray technologies, *Curr. Opin. Biotech.* 19, 30-35.
- 96. Uhlen, M. (2007) Mapping the human proteome using antibodies, *Mol. Cell. Proteomics* 6.8, 1455-1456.
- 97. Blow, N. (2007) The generation game, *Nature 447*, 741-744.

- 98. Kazanecki, C. C., Kowalski, A. J., Ding, T., Rittling, S. R., and Denhardt, D. T. (2007) Characterization of anti-osteopontin monoclonal antibodies: binding sensitivity to post-translational modifications, *J. Cell. Biochem.* 102, 925-935.
- 99. Taussig, M. J., Stoevesandt, O., Borrebaeck, C. A. K., Bradbury, A., Cahill, D., Cambillau, C., de Daruvar, A., Dubel, S., Eichler, J., Frank, R., Gibson, T., Gloriam, D., Gold, L., Herberg, F., Hermjakob, H., Hoheisel, J., Joos, T., Kallioniemi, O., Koegll, M., Konthur, Z., Korn, B., Kremmer, E., Krobitsch, S., Landegren, U., van der Maarel, S., McCafferty, J., Muyldermans, S., Nygren, P., Palcy, S., Pluckthun, A., Polic, B., Przybylski, M., Saviranta, P., Sawyer, A., Sherman, D., Skerra, A., Templin, M., Ueffing, M., and Uhlen, M. (2007) ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome, *Nature Methods 4*, 13-17.
- 100. Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands, *Nature 346*, 818-822.
- 101. Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science 249*, 505-510.
- 102. Wilson, D. S., and Szostak, J. W. (1999) In vitro selection of functional nucleic acids, *Annu. Rev. Biochem.* 68, 611-647.
- 103. Bunka, D. H. J., and Stockley, P. G. (2006) Aptamers come of age- at last, *Nature* 4, 588-596.
- 104. Gilbert, B. A., Sha, M., Wathen, S. T., and Rando, R. R. (1997) RNA aptamers that specifically bind to a K Ras-derived farnesylated peptide, *Bio. Med. Chem. 5*, 1115-1122.
- 105. Vaish, N. K., Dong, F., Andrews, L., Schweppe, R. E., Ahn, N. G., Blatt, L., and Seiwert, S. D. (2002) Monitoring post-translational modification of proteins with allosteric ribozymes, *Nat. Biotech.* 20, 810-815.
- Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions, *Science 311*, 844-847.

- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature 389*, 251-260.
- 108. Mendonsa, S. D., and Bowser, M. T. (2004) In vitro evolution of functional DNA using capillary electrophoresis, J. Am. Chem. Soc. 126, 20-21.
- 109. Mosing, R. K., Mendonsa, S. D., and Bowser, M. T. (2005) Capillary Electrophoresis-SELEX Selection of Aptamers with Affinity for HIV-1 Reverse Transcriptase, *Anal. Chem.* 77, 6107-6112.
- 110. Lin, L., Hom, D., Lindsay, S. M., and Chaput, J. C. (2007) In vitro selection of histone H4 aptamers for recognition imaging microscopy, *J. Am. Chem. Soc. 129*, 14568-14569.
- 111. Carothers, J. M., Oestreich, S. C., and Szostak, J. W. (2006) Aptamers selected for higher-affinity binding are not more specific for the target ligand, *J. Am. Chem. Soc. 128*, 7929-7937.
- 112. Zichi, D., Eaton, B., Singer, B., and Gold, L. (2008) Proteomics and diagnostics: let's get specific, again, *Curr. Opin. Chem. Biol.* 12, 78-85.
- 113. Blow, N. (2007) The generation game, Nature 447, 741-744.
- 114. Hober, S., and Uhlen, M. (2008) Human protein atlas and the use of microarray technologies, *Curr. Opin. Biotechnol.* 19, 30-35.
- 115. Uhlen, M. (2007) Mapping the human proteome using antibodies, *Mol. Cell. Proteomics* 6.8, 1455-1456.
- 116. Chambers, R. S. (2005) High-throughput antibody production, Curr. Opin. Chem. Biol. 9, 46-50.
- 117. Zichi, D., Eaton, B., Singer, B., and Gold, L. (2008) Proteomics and diagnostics: let's get specific, again, *Curr. Opin. Chem. Biol.* 12, 78-85.

- 118. Taussig, M. J., Stoevesandt, O., Borrebaeck, C., Bradbury, A., Cahill, D., Cambillau, C., de Daruvar, A., Dubel, S., Eichler, J., Frank, R., Gibson, T., Gloriam, D., Gold, L., Herberg, F., Hermjakob, H., Hoheisel, J., Joos, T., Kallioniemi, O., Koegll, M., Konthur, Z., Korn, B., Kremmer, E., Krobitsch, S., Landegren, U., van der Maarel, S., McCafferty, J., Muyldermans, S., Nygren, P., Palcy, S., Pluckthun, A., Polic, B., Przybylski, M., Saviranta, P., Sawyer, A., Sherman, D., Skerra, A., Templin, M., Ueffing, M., and Uhlen, M. (2007) ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome, *Nat. Methods 4*, 13-17.
- 119. Hoogenboom, H. R. (2005) Selecting and screening recombinant antibody libraries, *Nat. Biotechnol.* 23, 1105-1116.
- 120. Hudson, P. J., and Souriau, C. (2003) Engineered antibodies, Nat. Med. 9, 129-134.
- 121. Hey, T., Fiedler, E., Rudolph, R., and Fiedler, M. (2005) Artificial, non-antibody binding proteins for pharmeceutical and industrial applications, *Trends Biotechnol.* 23, 514-522.
- 122. Bradbury, A. R. M., and Marks, J. D. (2004) Antibodies from phage antibody libraries, J. Immunol. Methods 290, 29-49.
- 123. Hosse, R. J., Rothe, A., and Power, B. E. (2006) A new generation of protein display scaffolds for molecular recognition, *Protein Sci.* 15, 14-27.
- 124. Wilson, D. S., and Szostak, J. W. (1999) In vitro selection of functional nucleic acids., *Annu. Rev. Biochem.* 68, 611-647.
- 125. Erlanson, D. A., Wells, J. A., and Braisted, A. C. (2004) Tethering: fragmentbased drug discovery, Annu. Rev. Biophys. Biomol. Struct. 33, 199-223.
- 126. Williams, B. A. R., Lin, L., Lindsay, S. M., and Chaput, J. C. (2009) Evolution of a histone H4-K16 acetyl-specific DNA aptamer, J. Am. Chem. Soc. 131, 6330-6331.

- Kodadek, T., Reddy, M. M., Olivios, H. J., Bachhawat-Sikder, K., and Alluri, P. G. (2004) Synthetic molecules as antibody replacements, *Acc. Chem. Res.* 37, 711-718.
- 128. Mammen, M., Choi, S.-K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design in use of multivalent ligands and inhibitors, *Angew. Chem. Int. Ed.* 37, 2754-2794.
- 129. Reddy, M. M., Bachhawat-Sikder, K., and Kodadek, T. (2004) Transformation of low-affinity lead compounds into high-affinity protein capture agents, *Chem. Biol.* 11, 1127-1137.
- 130. Barry, M. A., Dower, W. J., and Johnston, S. A. (1996) Toward cell-targeting gene therapy vectors: Selection of cell-binding peptides from random peptide-presenting phage libraries, *Nat. Med. 2*, 229-305.
- 131. Johnston, S. A., Salmeron, J. M., and Dincher, S. S. (1987) Interaction of positive and negative regulatory proteins in the galactose regulon of yeast, *Cell 50*, 143-146.
- 132. Sinz, A., Kalkhof, S., and Ihling, C. (2005) Mapping protein interfaces by a trifunctional crosslinker combined with MALDI-TOF and ESI-FTICR mass spectrometry, J. Am. Soc. Mass Spectrom. 16, 1921-1931.
- 133. Kumar, P. R., Yu, Y., Sternglanz, R., Johnston, S. A., and Joshua-Tor, L. (2008) NADP regulates the yeast GAL induction system, *Science 319*, 1090-1092.
- 134. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR, *Science 274*, 1531-1534.
- 135. Udugamasooriya, D. G., Dineen, S. P., Brekken, R. A., and Kodadek, T. (2008) A peptoid "antibody surrogate" that antagonizes VEGF receptor 2 activity, *J. Am. Chem. Soc.* 130, 5744-5752.
- 136. Homola, J., Vaisocherova, H., Dostalek, J., and Piliarik, M. (2005) Multi-analyte surface plasmon resonance biosensing, *Methods* 37, 26-36.

- 137. Harrison, J. G., and Balasubramanian, S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates, *Nucleic Acids Res. 26*, 3136-3145.
- 138. Pelloise, J. P., Zhou, X., Srivannavit, O., Zhou, T., Gulari, E., and Gao, X. (2002) Individually addressable parallel peptide synthesis on microchips, *Nat. Biotechnol.* 20, 922-926.
- Boltz, K., Gonzalez-Moa, M. J., Stafford, P., Johnston, S. A., and Svarovsky, S. A. (2009) Peptide microarrays for carbohydrate recognition, *Analyst 134*, 650-652.
- 140. Tung, C.-H. (2000) Preparation and application of peptide-oligonucleotide conjugates, *Bioconj. Chem. 11*, 605-618.
- 141. Singh, Y., Spinelli, N., and Defrancq, E. (2008) Chemical strategies for oligonucleotide-conjugates synthesis, *Curr. Org. Chem. 12*, 263-290.
- 142. Grandas, A., Marchan, V., Debethune, L., and Pedroso, E. (2007) Stepwise solidphase synthesis of nucleopeptides, *Curr. Prot. Nucleic Acid Chem.* Unit 4.22.1-4.22.54
- 143. Harrison, J. G., and Balasubramanian, S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates, *Nucleic Acids Res. 26*, 3136-3145.
- 144. Sharma, S., and Yadav, B. R. (2009) Denaturing urea-polyacrylamide gel electrophoresis (PAGE) based microsatellite analysis, <u>http://www.protocol-online.org</u>.
- 145. Tournier, E. J. M., Wallach, J., and Blond, P. (1998) Sulfosuccinimidyl 4-(N-maleimidomethyl)-1-cyclohexane carboxylate as a bifunctional immobilization agent. Optimization of the coupling conditions, *Anal. Chimica Acta 361*, 33-44.
- 146. Brewer, C. F., and Riehm, J. P. (1967) Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins, *Anal. Bioch.* 18, 248-255.

- 147. Hashida, S., Imagawa, M., Inoue, S., Ruan, K.-H., and Ishikawa, E. (1984) More useful maleimide compounds for the conjufation of Fab' to horseradish peroxidase through thiol Groups in the hinge, *J. App. Biochem.* 6, 56-63.
- 148. Tung, C.-H., and Stein, S. (2000) Preparation and applications of peptideoligonucleotide conjugates, *Bioconj. Chem. 11*, 605-618.
- 149. Williams, B. A. R., Lund, K., Liu, Y., Yan, H., and Chaput, J. C. (2007) Selfassembled peptide nanoarrays: an approach to studying protein-protein interactions, *Angew. Chem. Int. Ed.* 46, 3051-3054.
- 150. Stearns, L. A., Chhabra, R., Sharma, J., Liu, Y., Petuskey, W. T., Yan, H., and Chaput, J. C. (2009) Template-directed nucleation and growth of inorganic nanoparticles on DNA scaffolds, *Angew. Chem. Int. Ed.* 48, 8494-8496.
- 151. Williams, B. A. R., Diehnelt, C. W., Belcher, P., Greving, M., Woodbury, N. W., Johnston, S. A., and Chaput, J. C. (2009) Creating protein affinity reagents by combining peptide ligands on synthetic DNA scaffolds, *J. Am. Chem. Soc. 131*, 17233-17241.
- 152. Tengvall, U., Auirola, S., Antopolsky, M., Azhaev, A., and Beigelman, L. (2003) Characterization of antisense oligonucleotide-peptide conjugates with negative ionization electrospray mass spectrometry and liquid chromatography-mass spectrometry, J. Phamecuet. Biomed. Analysis 32, 581-590.
- 153. Williams, B. A. R., Lund, K., Liu, Y., Yan, H., and Chaput, J. C. (2007) Selfassembled peptide nanoarrays: an approach to studying protein-protein interactions, *Angew. Chem. Int. Ed.* 46, 3051-3054.
- 154. Williams, B. A. R., Lin, L., Lindsay, S. M., and Chaput, J. C. (2009) Evolution of a histone H4-K16 acetyl-specific DNA aptamer, J. Am. Chem. Soc. 131, 6330-6331.
- 155. Stearns, L. A., Chhabra, R., Sharma, J., Liu, Y., Petuskey, W. T., Yan, H., and Chaput, J. C. (2009) Template-directed nucleation and growth of inorganic nanoparticles on DNA Scaffolds, *Angew. Chem. Int. Ed.* 48, 8494-8496.
- 156. Blow, N. (2007) The generation game, *Nature* 447, 741-744.

APPENDIX

COPYRIGHT PERMISSION

•

The American Association for the Advancement of Science

License terms and conditions: Mar 12, 2010 This is a License Agreement between Berea A Williams ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") provided by Copyright Clearance Center ("CCC"). 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. All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form. License Number 2386810908972 License date Mar 12, 2010 Licensed content publisher The American Association for the Advancement of Science Licensed content publication Science Licensed content title Discovering High-Affinity Ligands for Proteins: SAR by NMR Licensed content author Suzanne B. Shuker, Philip J. Hajduk, Robert P. Meadows, Stephen W. Fesik Licensed content date Nov 29, 1996 Type of Use Thesis / Dissertation Requestor type Other Individual Format Print Portion Text Excerpt Number of pages requested 2 Order reference number ref33 Title of your thesis / dissertation Molecular Recognition Using DNA As A Scaffolding Tool Expected completion date Apr 2010 Estimated size (pages) 250 Total 0.00 USD

Terms and Conditions American Association for the Advancement of Science TERMS AND CONDITIONS Regarding your request, we are pleased to grant you non-exclusive, non-transferable permission, to republish the AAAS material identified above in your work identified above, subject to the terms and conditions herein. We must be contacted for permission for any uses other than those specifically identified in your request above. The following credit line must be printed along with the AAAS material: "From [Full Reference Citation]. Reprinted with permission from AAAS." All required credit lines and notices must be visible any time a user accesses any part of the AAAS material and must appear on any printed copies and authorized user might make. This permission does not apply to figures / photos / artwork or any other content or materials included in your work that are credited to non-AAAS sources. If the requested material is sourced to or references non-AAAS sources, you must obtain authorization from that source as well before using that material. You agree to hold harmless and indemnify AAAS against any claims arising from your use of any content in your work that is credited to non-AAAS sources. If the AAAS material covered by this permission was published in Science during the years 1974 - 1994, you must also obtain permission from the author, who may grant or withhold permission, and who may or may not charge a fee if permission is granted. See original article for author's address. This condition does not apply to news articles. The AAAS material may not be modified or altered except that figures and tables may be modified with permission from the author. Author permission for any such changes must be secured prior to your use. Whenever possible, we ask that electronic uses of the AAAS material permitted herein include a hyperlink to the original work on AAAS's website (hyperlink may be embedded in the reference citation). AAAS material reproduced in your work identified herein must not account for more than 30% of the

total contents of that work. AAAS must publish the full paper prior to use of any text. AAAS material must not imply any endorsement by the American Association for the Advancement of Science. This permission is not valid for the use of the AAAS and/or Science logos. AAAS makes no representations or warranties as to the accuracy of any information contained in the AAAS material covered by this permission, including any warranties of merchantability or fitness for a particular purpose. If permission fees for this use are waived, please note that AAAS reserves the right to charge for reproduction of this material in the future. Permission is not valid unless payment is received within sixty (60) days of the issuance of this permission. If payment is not received within this time period then all rights granted herein shall be revoked and this permission will be considered null and void. In the event of breach of any of the terms and conditions herein or any of CCC's Billing and Payment terms and conditions, all rights granted herein shall be revoked and this permission will be considered null and void. AAAS reserves the right to terminate this permission and all rights granted herein at its discretion, for any purpose, at any time. In the event that AAAS elects to terminate this permission, you will have no further right to publish, publicly perform, publicly display, distribute or otherwise use any matter in which the AAAS content had been included, and all fees paid hereunder shall be fully refunded to you. Notification of termination will be sent to the contact information as supplied by you during the request process and termination shall be immediate upon sending the notice. Neither AAAS nor CCC shall be liable for any costs, expenses, or damages you may incur as a result of the termination of this permission, beyond the refund noted above. This Permission may not be amended except by written document signed by both parties. The terms above are applicable to all permissions granted for the use of AAAS material. Below you will find additional conditions that apply to your particular type of use. FOR A THESIS OR DISSERTATION If you are using figure(s)/table(s), permission is granted for use in print and electronic versions of your dissertation or thesis. A full text article may be used in print versions only of a dissertation or thesis. Permission covers the distribution of your dissertation or thesis on demand by ProQuest / UMI, provided the AAAS material covered by this permission remains in situ. If you are an Original Author on the AAAS article being reproduced, please refer to your License to Publish for rules on reproducing your paper in a dissertation or thesis. FOR JOURNALS: Permission covers both print and electronic versions of your journal article, however the AAAS material may not be used in any manner other than within the context of your article. FOR BOOKS/TEXTBOOKS: If this license is to reuse figures/tables, then permission is granted for non-exclusive world rights in all languages in both print and electronic formats (electronic formats are defined below). If this license is to reuse a text excerpt or a full text article, then permission is granted for non-exclusive world rights in English only. You have the option of securing either print or electronic rights or both, but electronic rights are not automatically granted and do garner additional fees. Permission for translations of text excerpts or full text articles into other languages must be obtained separately. Licenses granted for use of AAAS material in electronic format books/textbooks are valid only in cases where the electronic version is equivalent to or substitutes for the print version of the book/textbook. The AAAS material reproduced as permitted herein must remain in situ

and must not be exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit), except in the case of permitted textbook companions as noted below. You must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher." If your book is an academic textbook, permission covers the following companions to your textbook, provided such companions are distributed only in conjunction with your textbook at no additional cost to the user: -Password-protected website - Instructor's image CD/DVD and/or PowerPoint resource -Student CD/DVD All companions must contain instructions to users that the AAAS material may be used for non-commercial, classroom purposes only. Any other uses require the prior written permission from AAAS. Rights also extend to copies/files of your Work (as described above) that you are required to provide for use by the visually and/or print disabled in compliance with state and federal laws. This permission only covers a single edition of your work as identified in your request. FOR NEWSLETTERS: Permission covers print and/or electronic versions, provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit) FOR ANNUAL REPORTS: Permission covers print and electronic versions provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit) FOR PROMOTIONAL/MARKETING USES: Permission covers the use of AAAS material in promotional or marketing pieces such as information packets, media kits, product slide kits, brochures, or flyers limited to a single print run. The AAAS Material may not be used in any manner which implies endorsement or promotion by the American Association for the Advancement of Science (AAAS) or Science of any product or service. AAAS does not permit the reproduction of its name, logo or text on promotional literature. If permission to use a full text article is permitted, The Science article covered by this permission must not be altered in any way. No additional printing may be set onto an article copy other than the copyright credit line required above. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to, the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto the article copies. Additionally, article copies must be a freestanding part of any information package (i.e. media kit) into which they are inserted. They may not be physically attached to anything, such as an advertising insert, or have anything attached to them, such as a sample product. Article copies must be easily removable from any kits or informational packages in which they are used. The only exception is that article copies may be inserted into three-ring binders. FOR CORPORATE INTERNAL USE: The AAAS material covered by this permission may

not be altered in any way. No additional printing may be set onto an article copy other than the required credit line. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto article copies. If you are making article copies, copies are restricted to the number indicated in your request and must be distributed only to internal employees for internal use. If you are using AAAS Material in Presentation Slides, the required credit line must be visible on the slide where the AAAS material will be reprinted If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher." Access to any such CD, DVD, Flash Drive or Web page must be restricted to your organization's employees only. FOR CME COURSE and SCIENTIFIC SOCIETY MEETINGS: Permission is restricted to the particular Course, Seminar, Conference, or Meeting indicated in your request. If this license covers a text excerpt or a Full Text Article, access to the reprinted AAAS material must be restricted to attendees of your event only (if you have been granted electronic rights for use of a full text article on your website, your website must be password protected, or access restricted so that only attendees can access the content on your site). If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher." FOR POLICY REPORTS: These rights are granted only to non-profit government agencies. Permission covers print and electronic versions of a report, provided the required credit line appears in both versions and provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately. FOR CLASSROOM PHOTOCOPIES: Permission covers distribution in print copy format only. Article copies must be freestanding and not part of a course pack. They may not be physically attached to anything or have anything attached to them. FOR COURSEPACKS OR COURSE WEBSITES: These rights cover use of the AAAS material in one class at one institution. Permission is valid only for a single semester after which the AAAS material must be removed from the Electronic Course website, unless new permission is obtained for an additional semester. If the material is to be distributed online, access must be restricted to students and instructors enrolled in that particular course by some means of password or access control. FOR WEBSITES: You must include the following notice in any electronic versions, either adjacent to the reprinted

AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher." Permissions for the use of Full Text articles on third party websites are granted on a case by case basis and only in cases where access to the AAAS Material is restricted by some means of password or access control. Alternately, an E-Print may be purchased through our reprints department (brocheleau@rockwaterinc.com). REGARDING FULL TEXT ARTICLE USE ON THE WORLD WIDE WEB IF YOU ARE AN 'ORIGINAL AUTHOR' OF A SCIENCE PAPER If you chose "Original Author" as the Requestor Type, you are warranting that you are one of authors listed on the License Agreement as a "Licensed content author" or that you are acting on that author's behalf to use the Licensed content in a new work that one of the authors listed on the License Agreement as a "Licensed content author" has written. Original Authors may post the 'Accepted Version' of their full text article on their personal or on their University website and not on any other website. The 'Accepted Version' is the version of the paper accepted for publication by AAAS including changes resulting from peer review but prior to AAAS's copy editing and production (in other words not the AAAS published version). FOR MOVIES / FILM / TELEVISION: Permission is granted to use, record, film, photograph, and/or tape the AAAS material in connection with your program/film and in any medium your program/film may be shown or heard, including but not limited to broadcast and cable television, radio, print, world wide web, and videocassette. The required credit line should run in the program/film's end credits. FOR MUSEUM EXHIBITIONS: Permission is granted to use the AAAS material as part of a single exhibition for the duration of that exhibit. Permission for use of the material in promotional materials for the exhibit must be cleared separately with AAAS (please contact us at permissions@aaas.org). FOR TRANSLATIONS: Translation rights apply only to the language identified in your request summary above. The following disclaimer must appear with your translation, on the first page of the article, after the credit line: "This translation is not an official translation by AAAS staff, nor is it endorsed by AAAS as accurate. In crucial matters, please refer to the official Englishlanguage version originally published by AAAS." FOR USE ON A COVER: Permission is granted to use the AAAS material on the cover of a journal issue, newsletter issue, book, textbook, or annual report in print and electronic formats provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately By using the AAAS Material identified in your request, you agree to abide by all the terms and conditions herein. Questions about these terms can be directed to the AAAS Permissions department permissions@aaas.org. v 2 Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required. If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number

RLNK10750200. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time. Make Payment To:Copyright Clearance CenterDept 001 P.O. Box 843006 Boston, MA 02284-3006 If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2386810908972 and noting the reason for cancellation. Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

American Chemical Society

License terms and conditions: Mar 11, 2010This is a License Agreement between Berea A Williams ("You") and American Chemical Society ("American Chemical Society") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by American Chemical Society, and the payment terms and conditions. *All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.* License Number 2337740461527 License Date Dec 28, 2009 Licensed content publisher American Chemical Society Licensed content publication Journal of the American Chemical Society Licensed content title Evolution of a Histone H4-K16 Acetyl-Specific DNA Aptamer Licensed content author Berea A. R. Williams et al. Licensed content date May 1, 2009 Volume number 131 Issue number 18 Type of Use Thesis/Dissertation Requestor type Not specified Format Print and Electronic Portion 50% or more of original article Author of this ACS article Yes Order reference number H4-Aptamer Title of the thesis / dissertation Molecular Recognition using DNA As a Scaffolding Tool Expected License Number 2337740697024 License Date Dec 28, 2009 Licensed content publisher American Chemical Society Licensed content publication Journal of the American Chemical Society Licensed content title Creating Protein Affinity Reagents by Combining Peptide Ligands on Synthetic DNA Scaffolds Licensed content author Berea A. R. Williams et al. Licensed content date Dec 1, 2009 Volume number 131 Issue number 47 Type of Use Thesis/Dissertation Requestor type Not specified Format Print and Electronic Portion 50% or more of original article Author of this ACS article Yes Order reference number synbody Title of the thesis / dissertation Molecular Recognition using DNA As a Scaffolding Tool License Number 2386800665858 License Date Mar 12, 2010 Licensed content publisher American Chemical Society Licensed content publication Nano Letters Licensed content title Programmable DNA Self-Assemblies for Nanoscale Organization of Ligands and Proteins Licensed content author Sung Ha Park et al. Licensed content date Apr 1, 2005 Volume number 5 Issue number 4 Type of Use Requestor Thesis/Dissertation type Not specified Format Print Portion Table/Figure/Chart Number of Table/Figure/Charts 1 Author of this ACS article No License Number 2386820393434 License Date Mar 12, 2010 Licensed content publisher American Chemical Society Licensed content publication Journal of the American Chemical Society Licensed content title High-Affinity Peptide-Based Collagen Targeting Using Synthetic Phage Mimics: From Phage Display to Dendrimer Display Licensed content author Brett A. Helms et al. Licensed content date Aug 1, 2009 Volume number 131 Issue number 33 Type of Use Thesis/Dissertation Requestor type Not specified Format Print Portion Table/Figure/Chart Number of Table/Figure/Charts 2 Author of this ACS article No License Number 2386820537183 License Date Mar 12, 2010 Licensed content publisher American Chemical Society Licensed content publication Journal of the American Chemical Society Licensed content title Optimizing Saccharide-Directed Molecular Delivery to Biological Receptors: Design, Synthesis, and Biological Evaluation of Glycodendrimer-Cyclodextrin Conjugates Licensed content author Juan M. Benito et al. Licensed content date Aug 1, 2004 Volume number 126 Issue number 33

Type of Use Thesis/Dissertation Requestor type Not specified Format Print Portion Table/Figure/Chart Number of Table/Figure/Charts 2 Author of this ACS article No License Number 238682080768 License Date Mar 12, 2010 Licensed content publisher American Chemical Society Licensed content publication Journal of the American Chemical Society Licensed content title Small Multivalent Architectures Mimicking Homotrimers of the TNF Superfamily Member CD40L: Delineating the Relationship between Structure and Effector Function Licensed content author Nathalie Trouche et al. Licensed content date Nov 1, 2007 Volume number 129 Issue number 44 Type of Use Thesis/Dissertation Requestor Not specified type Format Print Portion Table/Figure/Chart Number of Table/Figure/Charts 3 Author of this ACS article No License Number 2386830026005 License Date Mar 12, 2010 Licensed content publisher American Chemical Society Licensed content publication Bioconjugate Chemistry Licensed content title Synthesis and Evaluation of Bivalent NDP- α -MSH(7) Peptide Ligands for Binding to the Human Melanocortin Receptor 4 (hMC4R) Licensed content author Heather L. Handl et al. Licensed content date Jul 1, 2007 Volume number 18 Issue number 4 Type of Use Thesis/Dissertation Requestor type Not specified Format Print Portion Table/Figure/Chart Number of Table/Figure/Charts 2 Author of this ACS article No Title of the thesis / dissertation Molecular Recognition Using DNA As A Scaffolding Tool Expected completion date Apr 2010 Estimated size(pages) 250 Billing Type Invoice Billing Address ASU-Biodesign 1001 McAllister Tempe, AZ 85287-5201 United States Customer reference info Total 0.00 USD

ACS / RIGHTSLINK TERMS & CONDITIONS THESIS/DISSERTATION INTRODUCTION The publisher for this copyrighted material is the American Chemical Society. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at). LIMITED LICENSE Publisher hereby grants to you a non-exclusive license to use this material. Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process; any form of republication must be completed within 60 days from the date hereof (although copies prepared before then may be distributed thereafter). GEOGRAPHIC RIGHTS: SCOPE Licenses may be exercised anywhere in the world. RESERVATION OF RIGHTS Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions. PORTION RIGHTS STATEMENT: DISCLAIMER If you seek to reuse a portion from an ACS publication, it is your responsibility to examine each portion as published to determine whether a credit to, or copyright notice of, a third party owner was published adjacent to the item. You may only obtain permission via Rightslink to use material owned by ACS. Permission to use any material published in an ACS publication, journal, or article which is reprinted with permission of a third party must be obtained from the third party owner. ACS disclaims any responsibility for any use you make of items

owned by third parties without their permission. REVOCATION The American Chemical Society reserves the right to revoke a license for any reason, including but not limited to advertising and promotional uses of ACS content, third party usage, and incorrect figure source attribution. LICENSE CONTINGENT ON PAYMENT While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials. COPYRIGHT NOTICE: DISCLAIMER You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Reprinted with permission from REFERENCE CITATION. Copyright YEAR American Chemical Society." WARRANTIES: NONE Publisher makes no representations or warranties with respect to the licensed material. INDEMNITY You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license. NO TRANSFER OF LICENSE This license is personal to you or your publisher and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission. NO AMENDMENT EXCEPT IN WRITING This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf). OBJECTION TO CONTRARY TERMS Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control. JURISDICTION This license transaction shall be governed by and construed in accordance with the laws of the District of Columbia. You hereby agree to submit to the jurisdiction of the courts located in the District of Columbia for purposes of resolving any disputes that may arise in connection with this licensing transaction. THESES/DISSERTATION TERMS _Publishing implications of electronic publication of theses and dissertation material Students and their mentors should be aware that posting of theses and dissertation material on the Web prior to submission of material from that thesis or dissertation to an ACS journal may affect

publication in that journal. Whether Web posting is considered prior publication may be evaluated on a case-by-case basis by the journals editor. If an ACS journal editor considers Web posting to be prior publication, the paper will not be accepted for

publication in that journal. If you intend to submit your unpublished paper to ACS for publication, check with the appropriate editor prior to posting your manuscript electronically. If your paper has already been published by ACS and you want to include the text or portions of the text in your thesis/dissertation in *print or microfilm formats*, please print the ACS copyright credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [FULL REFERENCE CITATION.] Copyright [YEAR] American Chemical Society." Include appropriate information. *Submission to a Dissertation Distributor*: If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (*not the individual ACS* *article by itself*) for electronic dissemination through the distributor; ACSâ€TMs copyright credit line should be printed on the first page of the ACS paper. *Use on an Intranet:* The inclusion of your ACS unpublished or published manuscript is permitted in your thesis in print and microfilm formats. If ACS has published your paper you may include the manuscript in your thesis on an intranet that is not publicly available. Your ACS article cannot be posted electronically on a publicly available medium (i.e. one that is not password protected), such as but not limited to, electronic archives, Internet, library server, etc. The only material from your paper that can be posted on a public electronic medium is the article abstract, figures, and tables, and you may link to the articleâ€[™]s DOI or post the articleâ€[™]s author-directed URL link provided by ACS. This paragraph does not pertain to the dissertation distributor paragraph above. * Other conditions: * v1.1 *Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.* *If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10712411. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time. Make Payment To: Copyright Clearance Center Dept 001 P.O. Box 843006 Boston, MA 02284-3006 If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2337740461527 and noting the reason for cancellation. Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.*

Elsevier

License terms and conditions: Mar 12, 2010 This is a License Agreement between Berea A Williams ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions. All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form. Supplier Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB,UK Registered Company Number 1982084 Customer name Berea A Williams Customer address ASU-Biodesign Tempe, AZ 85287-5201 License Number 2386810197008 License date Mar 12, 2010 Licensed content publisher Elsevier Licensed content publication Journal of Molecular Biology Licensed content title 2.8 Å crystal structure of the malachite green aptamer Licensed content author Christopher Baugh. Dilârâ Grate, Charles Wilson Licensed content date 4 August 2000 Volume number 301 number Pages Thesis / Dissertation Portion Issue 1 12 Type of Use Figures/tables/illustrations Number of Figures/tables/illustrations 1 Format Print You are an author of the Elsevier article No Are you translating? No Order Reference Number Expected publication date Apr 2010 Elsevier VAT number GB 494 6272 12 Permissions price 0.00 USD Value added tax 0.0% 0.00 USD Total 0.00 US License number 2386830325581 License date Mar 12, 2010 Licensed content publisher Elsevier Licensed content publication Bioorganic & Medicinal Chemistry Letters Licensed content title Self-assembly of bivalent protein-binding agents based on oligonucleotide-linked organic fragments Licensed content author K. Ingrid Sprinz, Debarati M. Tagore, Andrew D. Hamilton Licensed content date 1 September 2005 Volume number 15 Issue number 17 Pages 4 Type of Use Thesis / Dissertation Portion Figures/tables/illustrations Number of Figures/tables/illustrations1Format Print You are an author of the Elsevier article No Are you translating? No Order Reference Number ref 73 Expected publication date Apr 2010 Elsevier VAT number GB 494 6272 12 Permissions price 0.00 USD Value added tax 0.0% 0.00 USD License Number 2387691369174 License date Mar 14, 2010 Licensed content publisher Elsevier Licensed content publication Biomolecular Engineering Licensed content title Dimeric and trimeric antibodies: high avidity scFvs for cancer targeting Licensed content author Alexander A. Kortt, Olan Dolezal, Barbara E. Power, Peter J. Hudson Licensed content date 15 October 2001 Volume number 18 Issue number 3 Pages 14 Type of Use Thesis / Dissertation Portion Figures/tables/illustrations Number of Figures/tables/illustrations 2 Format Print You are an author of the Elsevier article No Are you translating? No Order Reference Number ref 62 Expected publication date Apr 2010 Elsevier VAT number GB 494 6272 12 Permissions price 0.00 USD Value added tax 0.0% 0.00 USD

Terms and Conditions

INTRODUCTION 1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"),

at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com). GENERAL TERMS 2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated. 3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows: "Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier." 4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given. 5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com) 6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee. 7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions. 8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and If full payment is not received on a timely basis, then any license conditions. preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials. 9. Warranties: Publisher makes no representations or warranties with respect to the licensed material. 10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license. 11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission. 12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC

on publisher's behalf).13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control. 14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions. LIMITED LICENSE The following terms and conditions apply only to specific license types: 15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages. 16. Website: The following terms and conditions apply to electronic reserve and author websites: Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if: This license was made in connection with a course, This permission is granted for 1 year only. You may obtain a license for future website posting, All content posted to the web site must maintain the copyright information line on the bottom of each image. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com, and Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu. 17. Author website for journals with the following additional clauses: All content posted to the web site must maintain the copyright information line on the bottom of each image, and the permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version, A hypertext must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx, As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published

article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting. Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu. 18. Author website for books with the following additional clauses: Authors are permitted to place a brief summary of their work online only. A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com All content posted to the web site must maintain the copyright information line on the bottom of each image. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu. 19. Website (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. or for books to the Elsevier homepage at http://www.elsevier.com 20. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. 21. Other Conditions: None v1.6 Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required. If you would like to pay for this license now, please remit this license along with your payment made pavable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10750198. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time. Make Payment To: Copyright Clearance Center Dept 001 P.O. Box 843006 Boston, MA 02284-3006 If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2386810197008 and noting the reason for cancellation. Questions? customercare@copyright.com or $\pm 1-877-622-5543$ (toll free in the US) or $\pm 1-978-646-646$ 2777.

Nature Publishing Group

License terms and conditions: Mar 12, 2010 This is a License Agreement between Berea A Williams ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions. All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form. License Number 2386811387235 License date Mar 12, 2010 Licensed content publisher Nature Publishing Group Licensed content publication Nature Licensed content title Spanning binding sites on allosteric proteins with polymer-linked ligand dimmers Licensed content author Richard H. Kramer and Jeffrey W. Karpen Volume number 395 Issue number 6703 Pages pp710-713 Year of publication 1998 Portion used Figures / tables Number of figures / tables 2 License Number 2386820033734 License date Mar 12, 2010 Licensed content publisher Nature Publishing Group Licensed content publication Nature Licensed content title Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands Licensed content author Pavel I. Kitov, Joanna M. Sadowska, George Mulvey, Glen D. Armstrong, Hong Ling et al. Volume number 403 Issue number 6770 Pages pp669-672 Year of publication 2000 Portion used Figures / tables Number of figures / tables 2 License Number 2386821261632 License date Mar 12, 2010 Licensed content publisher Nature Publishing Group Licensed content publication Nature Biotechnology Licensed content title Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains Licensed content author Joshua Silverman, Qiang Lu, Alice Bakker, Wayne To Amy Duguay et al. Volume number 23 Issue number Pages pp1556-1561 Year of publication 2005 Portion used Figures / tables Number of figures / tables 2 License Number 2386821388272 License date Mar 12, 2010 Licensed content publisher Nature Publishing Group Licensed content publication Nature Nanotechnology Licensed content title Self-assembled DNA nanostructures for distance-dependent multivalent ligandprotein binding Licensed content author Sherri Rinker, Yonggang Ke, Yan Liu, Rahul Chhabra and Hao Yan Volume number 3 Issue number 7 Pages pp418-422 Year of publication 2008 Portion used Figures / tables Number of figures / tables 2 Requestor type Student Type of Use Thesis / Dissertation High-res requested No Billing Type Invoice Company Berea A Williams Billing Address ASU-Biodesign 1001 McAllister Tempe, AZ 85287-5201 United States Customer reference info Total 0.00 USD

Terms and Conditions Terms and Conditions for Permissions Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use, subject to the conditions below: 1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material. 2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version. Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run). NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission. 3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free). 4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage. 5. The credit line should read: Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication) For AOP papers, the credit line should read: Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX) 6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows: Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication) 7. Translations of 401 words up to a approval. whole article require NPG Please visit http://www.macmillanmedicalcommunications.com for more information. Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows: Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication) We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you. v1.1 Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required. If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK10750214. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time. Make Payment To: Copyright Clearance Center Dept 001 P.O. Box 843006 Boston, MA 02284-3006 If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2386821388272 and noting the reason for cancellation. Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

International Union of Crystallography

Permission is hereby granted, on behalf of the IUCr, for you to reproduce the material specified below, subject to the following conditions: 1. Reproduction is intended in a primary journal, secondary journal, CD-ROM, book or thesis. 2. The original article in which the material appeared is cited. 3. IUCr's copyright permission is indicated next to the Figure/Table in print. In electronic form, this acknowledgement must be visible at the same time as the Figure/Table, and must be hyperlinked to Crystallography Journals Online (http://journals.iucr.org/).

Wiley-VCH Verlag GmbH & Co. KGaA

Thank you for your request. We hereby grant permission for the requested use expected that due credit is given to the original source. For material published before 2007 additionally: Please note that the author's permission is also required. Please note that we only grant rights for a printed version, but not the rights for an electronic/ online/ web/ microfiche publication, but you are free to create a link to the article in question which is posted on our website (http://www3.interscience.wiley.com) If material appears within our work with credit to another source, authorisation from that source must be obtained. Credit must include the following components: - Books: Author(s)/ Editor(s) Name(s): Title of the Book. Page(s). Publication year. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. - Journals: Author(s) Name(s): Title of the Article. Name of the Journal. Publication year. Volume. Page(s). Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. With kind regards Bettina Loycke Senior Rights Manager Wiley-VCH Verlag GmbH & Co. KGaA Boschstr. 12 69469 Weinheim Germany Phone: +49 (0) 62 01- 606 - 280 Fax: +49(0) 62 01 - 606 - 332Email: rights@wiley-vch.de