ABSTRACT

Title of Document: CHARACTERIZATION OF BRANCHED HISTIDINE-LYSINE POLYPEPTIDES USED FOR NUCLEIC ACID DELIVERY AND THEIR COMPLEXES WITH DNA AND SMALL INTERFERING RNA.

Lucas James Tricoli, Doctor of Philosophy, 2013

Directed By: Associate Professor Jason D. Kahn, Department of Chemistry and Biochemistry

DNA and siRNA must be packaged for protection prior to transfection in vivo when they are administered to humans or other animals. Dr. A. J. Mixson's laboratory at the University of Maryland, Baltimore Medical School has developed a family of branched histidine-lysine (HK) peptides that confer improved transfection in mice compared to naked nucleic acid. The branched polymer denoted H3K4b has a superior ability to transfect siRNA in vitro compared to H3K(+G)4b. H3K(+G)4b, made by the addition of two glycines to each of the original H3K4b branches, is presumably a more flexible polymer, and it allows for better transfection of plasmid DNA than H3K4b. Biophysical characterization of the HK-DNA and HK-siRNA complexes is aimed at understanding how the structures of both peptides affect their biological activity. This characterization was performed using ethidium bromide exclusion from nucleic acid, DNase I plasmid degradation, dynamic light scattering,

circular dichroism, isothermal titration calorimetry and atomic force microscopy. Results from the characterization suggest that H3K4b forms condensed regions of packaged plasmid with some repeating accessibility of the DNA, compared to a more even coating of plasmid by H3K(+G)4b. Based on these results a "coating versus clumping" model was developed to relate the transfection efficiency of each peptide to its binding of plasmid DNA. A specific model for packaging of siRNA with these peptides was not developed, but we believe that characteristics that lead to effective transfection of plasmid are not key to siRNA delivery. A better understanding of characteristics important to peptide-nucleic acid complex formation may lead to the development of improved transfection agents.

CHARACTERIZATION OF BRANCHED HISTIDINE-LYSINE POLYPEPTIDES USED FOR NUCLEIC ACID DELIVERY AND THEIR COMPLEXES WITH DNA AND SMALL INTERFERING RNA.

By

Lucas James Tricoli

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Associate Professor Jason D. Kahn, Chair, Research Advisor
Associate Professor Douglas Julin
Assistant Professor Joonil Seog
Adjunct Professor Steven Rokita
Associate Professor Richard Stewart, Dean's Representative

Dedication

My thesis and research is dedicated to anyone who has looked into the night sky or at the beauty of the natural world around them and wondered "what all is there to know about life?" Carl Sagan once stated:

"For me it is far better to grasp the universe as it really is than to persist in delusion however satisfying and reassuring."

The pursuit of knowledge will constantly drive researchers to further mankind's ultimate understanding of the true nature of the universe.

Acknowledgements

I would like to first thank my Ph.D. advisor, Dr. Kahn for his guidance and the wisdom he has imparted to me during my graduate career. Dr. Kahn taught me the importance of critical and well-organized thinking and that nothing is outside the realm of understanding. I will always carry the importance of these scientific values with me in my pursuit for knowledge.

Dr. Mixson and his lab made possible the research I conducted for which I am very grateful.

I owe thanks to our group of collaborators on the "HK peptide" project, which include Dr. Mixson, Dr. Kahn, Dr. Seog, Justin (Szu-Ting) Chou, Adam Karcz, Amy Lee and Justin Hustedt. Our monthly, bi-weekly and weekly meetings provided invaluable insight onto the research conducted for this project.

My Ph.D. committee was supportive and showed me all the right avenues to take, making this a successful research project. They strived to test my knowledge and allowed me to keep prospective on the important aspects of a research project.

My lab mates, past and present deserve a huge thanks for their scientific help and personal support. My lab members, Dr. Haeusler, Dr. Gowestki, Dr. Goodson, Sarah Sucyan, Xiaoyu Wang and Jason Hustedt provided a stimulating lab environment, which proved enjoyable during my time here. I owe a special thanks to Dr. Haeusler, Dr. Gowetski and Jason Hustedt for their scientific feedback and contributions to my thesis research project. Members of the Rokita and Julin labs, who include Dr. Shadrick, Dr. McTamney, Dr. Servinski and Charles Mueller also provided much appreciated support.

I would like to thank everyone who allowed me the use of the instrumentation that made it possible to complete my research. I owe Dr. Sintim and his lab a large thanks for the use of his fluorometer on all the fluorescence experiments conducted. I greatly appreciate Dr. Bonenberger's help in scheduling time for use of the DLS.

My parents, Margaret Tricoli and Dr. James Tricoli and girlfriend, Anastasia Poulos have been a foundation, which I could rely upon for their tireless personal, professional and financial support. Without them, the realization of my dreams would not have been possible. My accomplishments and success are theirs, just as much as they are mine.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

1x TEN	A Buffer with 10 mM Tris, 1 mM EDTA, 100 mM
AFM	Atomic Force Microscopy
AO	Acridine Orange
CD	Circular Dichroism
DLS	Dynamic Light Scattering
DNA	Deoxyribose Nucleic Acid
EMSA	Electrophoretic Mobility Shift Assay
EtBr	Ethidium Bromide
H3K4b	Histidine-lysine Branched Peptide
H3K(+G)4b	Histidine-lysine Branched Peptide with Glycines
ITC	Isothermal Titration Calorimetry
I/Io	Fluorescence intensity divided by initial fluorescence intensity
N/P Ratio	Protonatable Nitrogen to Phosphate ratio
PAMAM	Poly(amidoamine)
PDI	Polydispersivity Index
PEI	Polyethyleneimine
PGA	Poly-L-Glutamic Acid
PLL	Poly-L-Lysine
PPL	Polyproline
RNAi	RNA Interference
siRNA	Small Interfering RNA

Chapter 1: Introduction

1.1 Biological Information Storage

The digital age has produced devices that transfer vast amounts of information. Mankind has created data storage devices, which nature has been working on and perfecting for billions of years[1; 2; 3]. The density of digital information stored per gram of nucleic acids is significantly greater than any storage device mankind has created to date[4; 5]. Starting with the proposed structure of B-form DNA by Watson and Crick[6], the rapid evolution of the biological sciences has uncovered a wealth of information about the intricacies of nucleic acids as data storage molecules. Once we can completely interpret the genetic code of an organism, a comprehensive understanding of its biological information exchange will be within reach.

Key to the ability of nucleic acids to act as data storage molecules is the central dogma of molecular biology[7]. The central dogma is essential to the propagation of information in an organism because it describes how cellular data propagates from DNA to RNA and finally to creation of protein[7]. Additions to the central dogma include DNA and RNA replication and also RNA reverse transcription back to DNA[7]. Proteins serve as the major intracellular workhorse, comprising everything from biological structures to cellular regulatory agents. The central dogma's elegant scheme is the foundation for many studies conducted in the biological sciences.

1.2 Nucleic Acids as the Key to Disease Treatment

Advances in technology and a new understanding of genetics over the past 40 years has made possible the ability to model and understand the biological

information storage system[8]. The ultimate objective is to understand biological information exchange adequately enough to improve quality of life. As the primary roles of nucleic acids and genetics in diseases become better understood, a way to correct the disease seems attainable[9; 10; 11]. Most diseases are categorized as a malfunctioning protein, most typically caused by a defect in gene expression or an infectious disease[12]. Treatment of diseases such as bacterial infections with antibiotics has proven to be extremely effective in many cases, but bacteria develop resistance to current antibiotic treatment regimens[13; 14 15]. There is no effective treatment for many genetic diseases, such as Huntington's disease[16]. In addition there are many treatment regimens, especially for diseases like cancer that prove to be marginally effective in some cases and highly detrimental to the patient[17]. The vast amount of information acquired by studying the central dogma and genetics has led to advances in our understanding of how diseases occur and allow us the ability to attack the main root cause of inherited and acquired disorders[8].

Nucleic acid delivery is one of many promising therapies to deal with these disorders. It has major appeal in possibly providing a permanent replacement or regulation of a defective gene so no further treatment need be applied[11]. Nucleic acid therapy has potential applications in treatment of hereditary, acquired or even infectious diseases[11]. There exist many ways to deliver a nucleic acid in order to effect treatment of a disease. With such a wide variety of applications and therapeutic potential, it is important to characterize what makes the best delivery vector for a specific application. The project I have conducted observed biophysical characteristics of two different nucleic acid binding peptides that have potential

therapeutic uses. This project's broader impact on the understanding of peptide's interaction with nucleic acids may have applications to effective methods of disease treatment. Eventually the knowledge gained here can be applied to the creation of more efficient methods of nucleic acid therapy.

1.3 Types of Nucleic Acid Payloads for Delivery

In the realm of therapeutic applications there are a few classes of nucleic acid payloads that can be delivered to effect therapy. Treatment falls into two major categories, gene delivery or RNA interference (RNAi). Gene delivery involves permanent or temporary replacement of a malfunctioning gene causing disease, with a functional one. Plasmids can be utilized as a delivery vehicle for genes[18]. Viruses have evolved to naturally inject their genetic material into a host cell, and they can also be engineered to deliver a therapeutic gene[19]. Currently, one of the most successful utilizations of gene therapy under development has been treatment of Severe Combined Immunodeficiency-X1 (X-linked SCID, also referred to as adenosine deaminase (ADA) deficiency), which is genetic disorder that causes problems with T and NK lymphocyte production[9]. A virus was used to successfully deliver a functional copy of the gene that causes this disease. Despite this success, the FDA has no approved gene delivery therapeutics for sale as of 2011[20].

The second treatment is through RNAi-mediated repression of over-expressed or aberrantly expressed genes. The major appeal of siRNA therapy compared to typical small molecule drugs is the ability to suppress the production of the faulty protein causing a disease before its translation. Plasmids can also be utilized to contain short hairpin RNA (shRNA), which are used to make siRNA precursors[21]. A

more direct approach is to manufacture and deliver the siRNA used for RNAi mediation of protein expression. The appeal of siRNA as a therapeutic compared to typical small molecule approaches is the ability to directly regulate a malfunctioning gene causing a disease, which has lead to over 30 clinical trials to date[22]. Unfortunately, the furthest these trials have reached is phase III for an siRNA therapeutic targeting age-related macular degeneration[23]. Many other treatments are still in phase I or II trials and pilot feasibility studies[23]. However, the future holds potential for treatments of diseases like Alzheimer's and Huntington's, in which over expression of proteins are the causes of the disease[16; 24; 25].

1.4 Key Characteristics of a Nucleic Acid Delivery Agent

Addressing what constitutes a good delivery vector is key to its creation. A myriad of factors that go into consideration of a good delivery vector can be boiled down into four critical aspects (Figure 1.1). First, the gene delivery vector has to effectively package and protect the nucleic acid (RNA or DNA) of interest during its passage through a host of biological media to its target[11]. Without protection, most of the nucleic acid will be destroyed before it reaches it target. Even in the event that some of the gene therapy agent does reach its target, without proper packaging and condensation, it is unlikely it will be able to enter the cell and be effective[26]. Second, the release of the therapeutic nucleic acid in the correct location is critical[11]. The therapeutic agent complex formed with the nucleic acid has to be robust and resilient to destruction. However, it also cannot be so robust as to never be degraded, inhibiting release of the nucleic acid to allow binding to its target. Third, the vector in question also has to be relatively inexpensive and easy to manufacture[11]. During

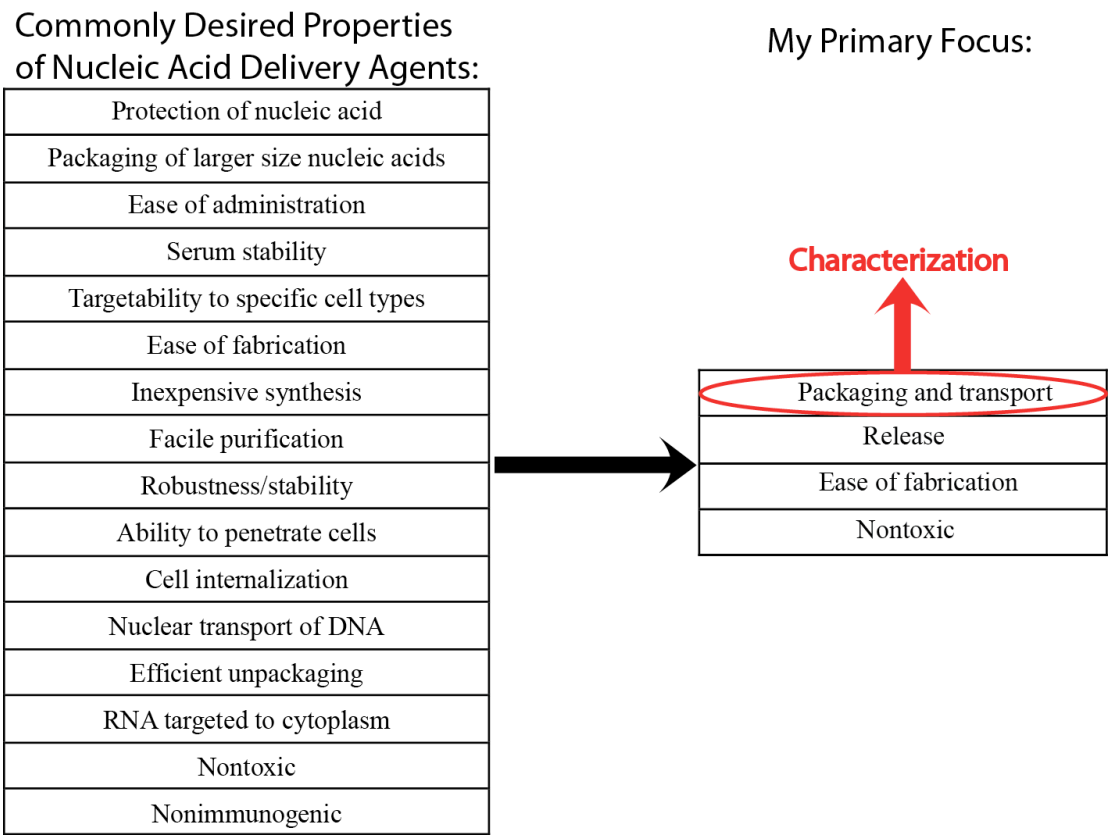


Figure 1.1 ***Nucleic Acid Delivery Agent Requirements.*** **A comprehensive list of requirements for an effective nucleic acid delivery agent, adapted from Lavigne[11]. These factors can be boiled down into four major aspects. The two factors that are critical to biochemically characterize are packaging and release of nucleic acids. A focus on creation of nucleic acid delivery agents that are more effective at packaging and release of their payload can be obtained through better insight into their function.**

this manufacturing process, it would also be positive to have a vector that is easy to chemically modify in the event it has to be adapted to a new set of conditions. The final key consideration for a good gene delivery agent is an inherent lack of toxicity to the host[11; 27]. If the vector is toxic, any chances for advancement to clinical trials are marginal, discouraging any future therapeutic use. Packaging and release are the two major characteristics of the four that can be biophysically characterized. Characterization of these two attributes of a gene delivery vector can yield a wealth of information about how they function.

1.5 Viral and Non-Viral Nucleic Acid Delivery Agents

In light of these key characteristics for a delivery agent, the most prevalent vectors that fit the criteria listed above fall into two major categories, viral and non-viral[11]. Viral vectors are a logical choice as a gene delivery agent due to the inherent ability of virus to deliver genetic material to the nucleus[8; 11]. Some of the first trials with gene therapy were accomplished in the early 1980's through the 1990's with

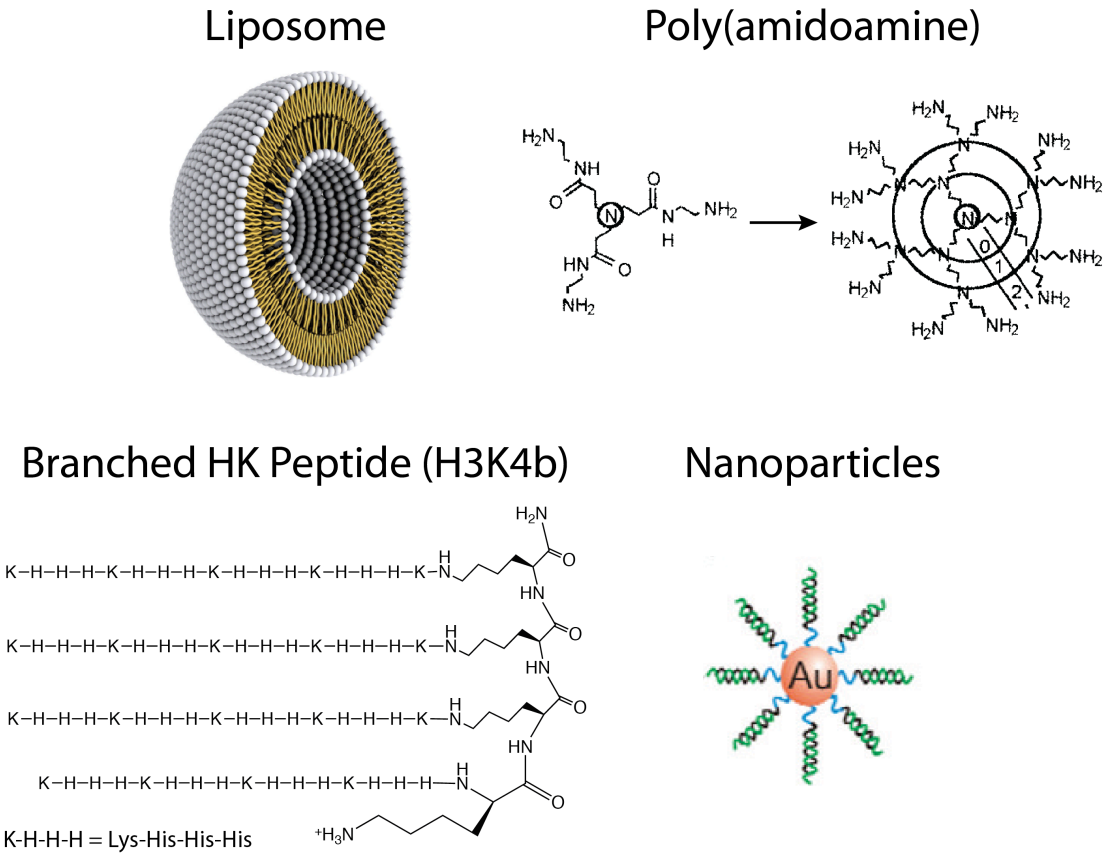


Figure 1.2 ***Non-viral Transfection Agents.*** **Images of four primary gene delivery vectors used currently, adapted from a couple of sources[11; 28; 29]. Liposomes form a spherical structure with a positively charge surface for where nucleic acids associate. Poly(amidoamine) (PAMAM) and branched HK peptides both have an extensive branched network with an abundance of positive charge to associate with nucleic acids. Gold nanoparticles, which have been made popular by Chad Mirkin, rely on functionalizing the surface with nucleic acids.**

virus delivery of recombinant DNA[8]. In 2000 the first fully successful clinical instance of gene therapy, dealing with treatment of SCID-X1, was accomplished[9]. For all the promise and success that viral vectors have demonstrated, there have been

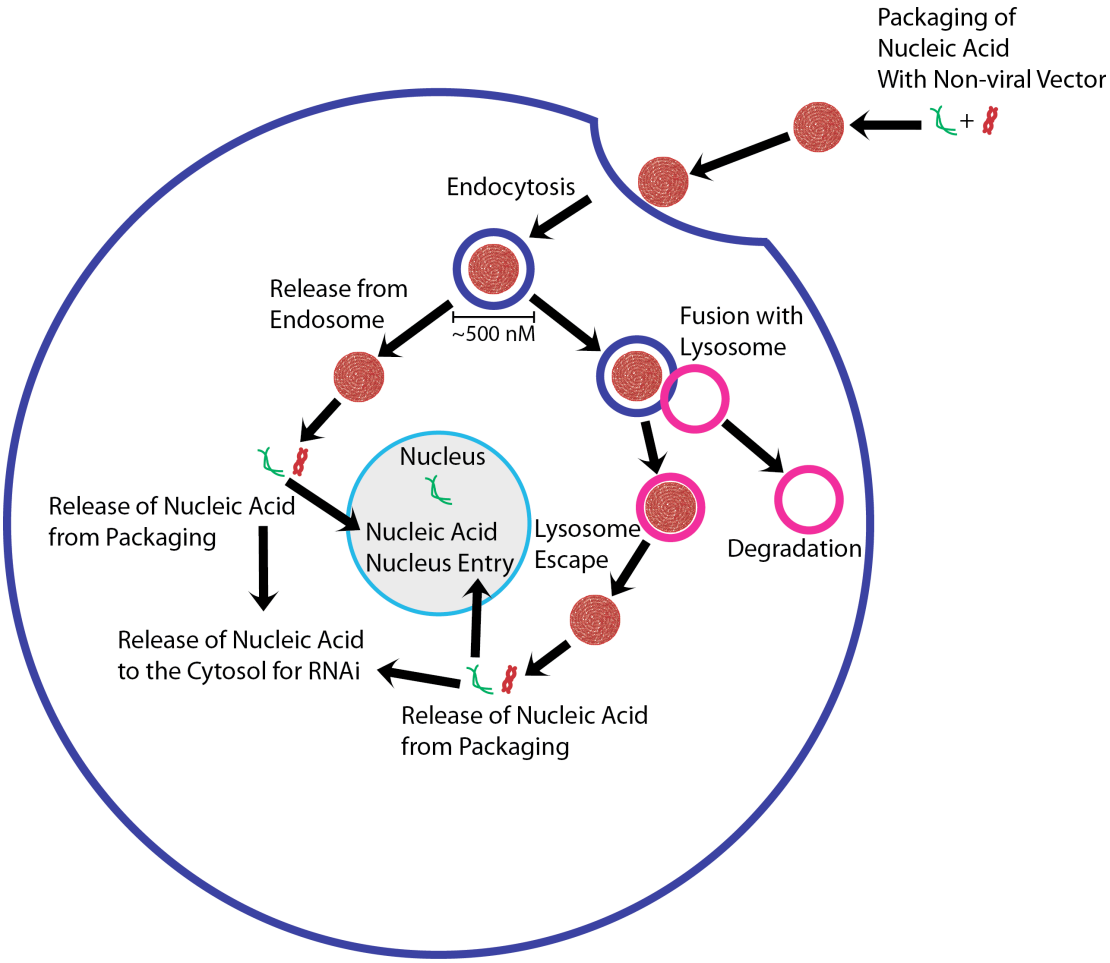


Figure 1.3 ***Cellular Fates of a Nucleic Acid Complex*****. An illustration of the possible fates of a packaged nucleic acid upon entry into a cell, adapted from Lavigne [11]. The nucleic acid (green) is packaged by a delivery agent and introduced to the cell. The packaged complex associates with the cell membrane where an invagination will form. This invagination will become an endosome containing the complex for cell internalization. The endosome is approximately 500 nm in diameter, which is very similar to the size of packaged complexes. The endosome containing a complex will be very densely packed. Once the endosome enters the cell the complex has one of two fates; it can be released or possibly fuse with a lysosome. If released, the nucleic acid has a reasonable probability of being unpackaged and delivered to the nucleus. If fused with a lysosome, the packaged nucleic acid could be degraded. If degradation does not occur then escape from the lysosome will occur and subsequent release to the nucleus.**

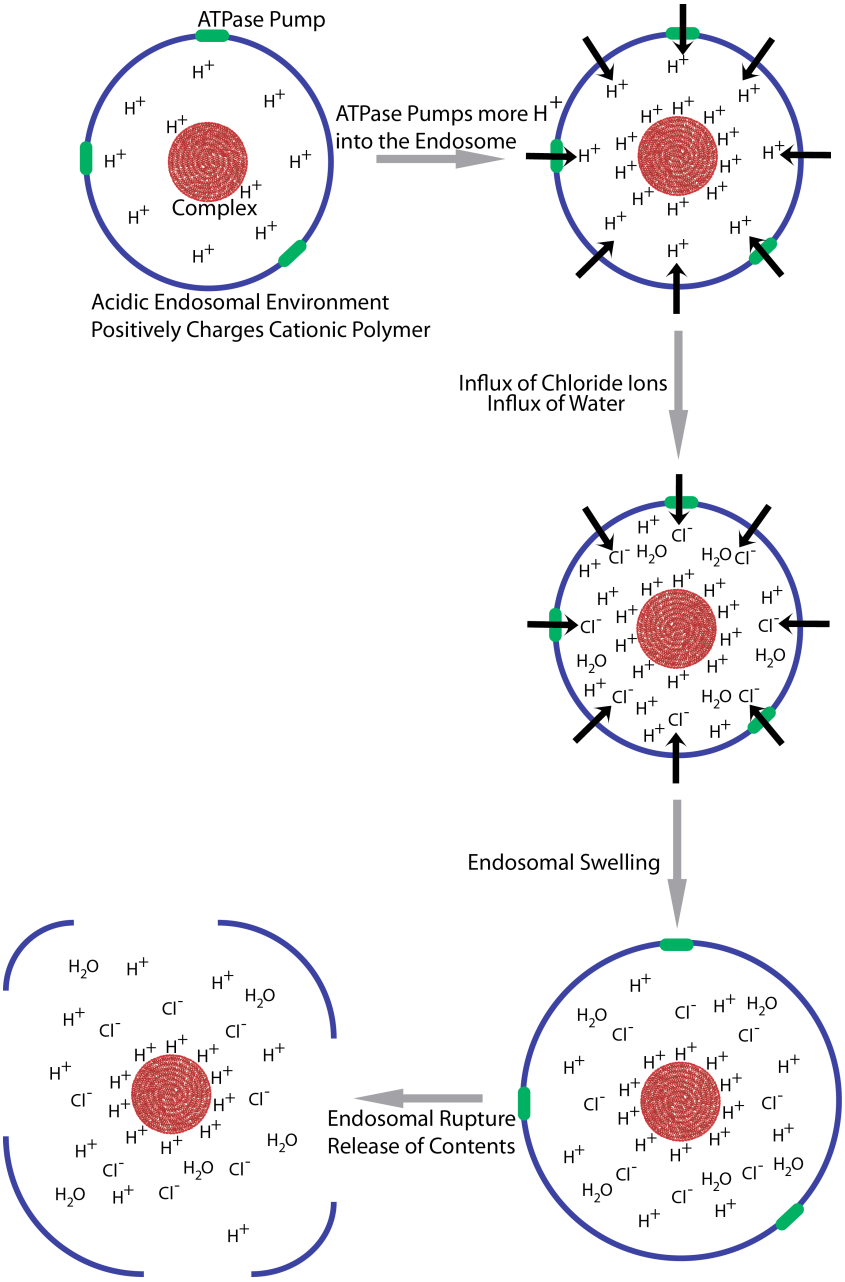
setbacks. These vectors are very costly to make, considering the extensive genetic modifications the virus has to undergo to be non-toxic to the host[11; 19]. Even after modification, there still is a possibility of toxicity by causing an immunogenic response in the host[10; 11; 27].

In light of these difficulties, non-viral methods of gene delivery have a greater

appeal as being easy to make, and they hold the promise to be just as effective as viral vectors[27]. A few of the most studied and used thus far are liposomes, nanoparticles, poly-amidoamine (PAMAM) and modified branched polymers (Figure 1.2)[29; 30; 31; 32; 33]. A general illustration for the fate of a delivery agent-nucleic acid complex is shown in Figure 1.3. The advantages of non-viral vectors include easy manufacturing making them significantly more cost effective to produce. During the manufacturing process they also retain great potential for modifications that will allow them to be adapted to the unique biological conditions that gene delivery presents[11]. Cationic lipids or liposomes are among some of the first non-viral agents used in successful protection and transfection of DNA[34]. Past formulations for liposomes consisted of a lipid bilayer shell in which the nucleic acid of interest is encapsulated. A very commonly used formulation today is lipofectamine, which forms a cationic particle where nucleic acids bind to the surface to form a delivery complex. Cationic lipid vesicles can stick to cellular surfaces in order to facilitate transfection[34].

Nanoparticles are a relatively new approach made popular by Chad Mirkin[29]. They involve functionalizing the outside of a particle with a dense population of nucleic acids such as RNA[29]. The thought is that the highly dense population of RNA aids in stabilizing the nucleic acids during their delivery through biological media[29].

All of these non-viral vectors have the ability to create condensed particles containing nucleic acid that are thought to be well protected during transport through the blood stream to its cellular target. In efforts to further the development of small cationic molecules for gene delivery came the creation of small poly(amidoamine) (PAMAM) and polyethyleneimine (PEI) molecules. Both of these molecules were



created around the same time, and showed an impressive ability in transfection of DNA, especially PAMAM[26; 35]. The complexes formed by these cationic molecules with nucleic acids are very different from liposome-nucleic acid complexes. It is believed that the highly positively charged nature of

Figure 1.4 ***Release of a Nucleic Acid Delivery Complex by the Proton Sponge Mechanism.*** **The amine cationic polymer accepts protons in the acidic endosome environment. This leads to a highly positively charged delivery complex. The ATPase pump in the endosomal membrane pump more protons into the endosome to maintain the acidic environment. Chloride counter ions flow in to counteract the high positive charge of the complex. Water also enters the endosome to counteract the increase concentration of ions. This sudden influx of ions and water causes destabilization and rupture of the endosome releasing its contents[36].**

these molecules leads to association with nucleic acid, forming a small aggregated complex of polymer and nucleic acid[26]. The highly branched nature and abundance

of basic amine groups in these molecules can provide a positively charged and extended surface area for binding and complexing with nucleic acids[35; 37]. Many of these cationic polymer are designed to have a pKa lower than biological pH, which leads to a high buffering capacity at near-neutral pH. The amine groups of these molecules can accept a large amount of protons in the endosome and become highly positively charged. A proton pump replaces the protons taken from the endosome by the peptide. The influx of positive charged protons leads to an influx of chloride counter ions and also water, causing swelling and destabilization of the endosome[36]. This mechanism is called the proton sponge hypothesis (Figure 1.4)[36]. Polypeptides have also been developed as potential transfection agents embodying many of the same successful aspects possessed by PEI and PAMAM. In addition to endosomal swelling, the polypeptide's branched nature is also an important characteristic which is believed to aid in condensation of plasmid DNA, contributing to their transfection ability.

Despite non-viral vectors' great potential for effective gene delivery, a serious set of difficulties still exists before widespread implementation can occur. Cationic lipids suffer from not being as effective as viral vectors at transfection[30]. PEI has proven to be effective in delivery of DNA and siRNA[26], but it is toxic to cells[11]. Modifications to both vectors have led to decreased toxicity of PEI variants and more efficient transfection by liposomes. However, without further modifications, lipsomes and PEI are difficult to consider as the only viable options for delivery vectors[11; 26]. This leaves us with PAMAM and modified polymers as another set of possible vectors. Both have a characteristic branched nature to them, which

presumably has contributed to their initial success as delivery agents. They also have limited cytotoxcicity in transfection trials accomplished thus far[33].

1.6 The Effectiveness of Polypeptides as Nucleic Acid Delivery Vehicles

Our collaborator Dr. A. James Mixson and his lab have worked extensively on the use of modified branched polymers. They have demonstrated great initial success with transfection and broad possibilities for modification during synthesis[33]. The credit for conception and success of these peptides goes to Dr. Mixson, and his lab has provided the in vivo experiments on the response of luciferase activity to nucleic acid delivery.

These peptides are synthesized from naturally occurring histidine and lysine, making them recognizable as a peptide substrate and degradable in serum. Two of

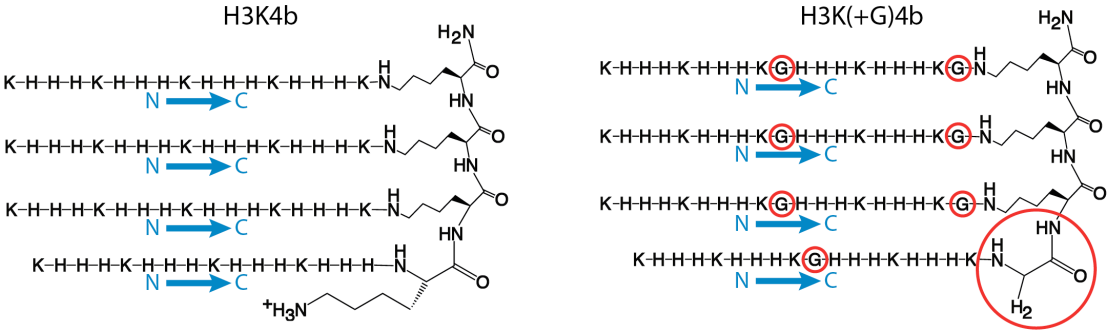


Figure 1.5 ***Structure of H3K4b and H3K(+G)4b.*** **A schematic of the two histidine lysine (HK) polypeptides, H3K4b and H3K(+G)4b. They each have a three-lysine core with four branches (4b) synthesized off of the three epsilon amino of each lysine and one N-terminal alpha amino group. Each branch is composed of a HHHK (H3K) repeated four times. Both have an approximate charge of +24 at pH 7.5 assuming a histidine side chain pKa of 6. The charge of the peptide can get up to +72 if all the terminal and side chain amines become protonated at low pH. The only difference between the peptides is the H3K(+G)4b has two glycines inserted at the start and middle of each branched region (circled in red), which may provide added flexibility to each branch.**

the peptides they have worked with, designated H3K4b and H3K(+G)4b, are shown in Figure 1.5. These two peptides are characterized by a three-lysine core which has four branches synthesized off of it, each containing a series of four HHHK repeats.

The only difference between these two peptides is a pair of glycines inserted at the beginning and the middle of each branch in H3K(+G)4b. Though a relatively small alteration in structure, these added glycines provide a significant difference in the expression of a transfected luciferase gene and also knockdown of endogenously expressed luciferase by siRNA.

In the first experiment a plasmid containing a gene for luciferase was complexed with H3K4b or H3K(+G)4b and transfected into a culture of MDA-MB-435 cells. Luciferase activity of the cultures was then measured after 24 hours of exposure to the transfection complex. As seen in Figure 1.6, H3K(+G)4b is a significantly more efficient transfection agent for plasmid DNA, as indicated by the

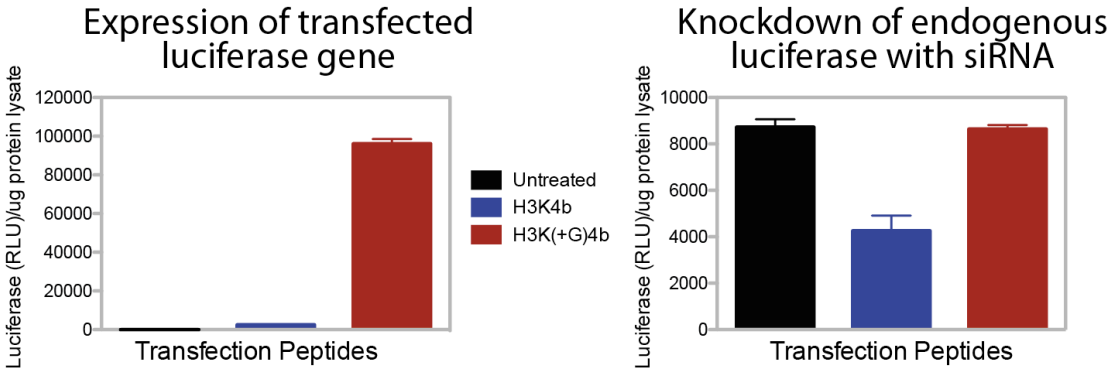


Figure 1.6 ***Expression and Knockdown of Luciferase Activity by Nucleic Acids Delivered by H3K4b and H3K(+G)4b*****. All transfection experiments were conducted and data provided by the Mixson lab (unpublished results). Markedly better expression with a luciferase expressing plasmid is seen with transfection using H3K(+G)4b rather than H3K4b. Knockdown of luciferase expression is 50% for H3K4b and insignificant for H3K(+G)4b.**

markedly higher expression of luciferase over that of H3K4b. The H3K4b peptide had hardly any ability to transfect of plasmid as indicated by the very low level of luciferase expression. The second experiment was designed to measure the ability of each peptide to deliver siRNA. This time MDA-MB-435 cells that stably expressed luciferase were exposed to an H3K4b-siRNA or H3K(+G)4b-siRNA complex in which the siRNA was designed to knock down luciferase expression. This

experiment demonstrated that RNA delivered by H3K4b inhibits luciferase expression by about 50% and therefore is a much better carrier of siRNA than H3K(+G)4b, which showed no ability to deliver of siRNA. These stark differences in transfection efficiency for DNA versus siRNA by the two peptides present a classic biochemistry question of how a slight structural difference can cause a dramatic functional difference.

1.7 Proposed Characterization of H3K4b and H3K(+G)4b-Nucleic Acid Complexes

The transfection data obtained by Dr. Mixson prompted us to ask what aspects of H3K4b and H3K(+G)4b interaction with siRNA and plasmid DNA cause the significant differences observed. It is clear what factors are desired of an ideal nucleic acid delivery vector, however the biochemical mechanisms that make a non-viral vector work are still poorly understood, therefore we would like to elucidate interactions of H3K4b and H3K(+G)4b with nucleic acids that contribute to transport, protection, cellular uptake and release. A clear understanding of these biochemical factors contributing to gene delivery agent interaction with nucleic acids can help to pave the way for creation of better agents.

Our objective is to correlate the in vitro data for transfection of each of these peptides with ethidium bromide exclusion, DNase I degradation, circular dichroism, isothermal titration calorimetry, dynamic light scattering and atomic force microscopy characterization to draw conclusions about what structural and biochemical interactions are causing these differences in transfection. We can then begin to understand the differences in the interaction of these two peptides with nucleic acids. If any differences in H3K4b and H3K(+G)4b's ability to bind nucleic

acids is discovered then we can hopefully draw conclusions about the biophysical control of transfection efficiency. This should provide a starting point for a list of important structural aspects that should be built into a gene delivery vector for selective uptake and delivery of siRNA or plasmid DNA.

Chapter 2: Nucleic Acid Coating by Histidine-Lysine Rich Branched Peptides

General Introduction:

H3K4b and H3K(+G)4b are cationic polypeptides used for delivery of nucleic acids. The binding of these two peptides to nucleic acids was assessed through dye displacement from siRNA and DNA and also DNase I probing of peptide-complexed DNA. Dye exclusion experiments demonstrate that H3K4b and H3K(+G)4b exclude ethidium brominde to different extents as the peptides are added to dye-DNA complexes. The dye exclusion is sensitive to local structure in peptide-DNA complexes, which reports on the uniformity with which each peptide coats nucleic acid. Exposure of the peptide-DNA complex to DNase I should reveal if the peptides form qualitatively different plasmid complexes. Exposure of DNA to solvent should show how densely packed each peptide is on the DNA and will report on differences in protection. Differences in the coverage of nucleic acid with H3K4b and H3K(+G)4b with nucleic acid will shed light on the intricate relationship between the structure of these two peptides and how this affects their functions as nucleic acid delivery agents.

2.1 Materials and Methods:

2.1.1 General Procedures and Materials Used

Materials Used:

Plasmid DNA was purified from *E. coli* cells using Qiagen Midiprep and Maxiprep kits. The DNA was a pRset derivative called prGCN4. Thanks goes to Dr. Dan Gowetski for allowing me use of his plasmid stock. DNA concentration was calculated by quantitation on a UV-visible spectrometer at a wavelength of 260 nm.

All 21 base pair β-gal and VEGF siRNA was provided by Dr. Mixson's laboratory. H3K4b and H3K(+G)4b peptides were synthesized and given to us by Dr. Mixson's laboratory. Peptides were HPLC purified and provided as a concentrated 1.5 mM stock solution. Concentrations of both H3K4b and H3K(+G)4b peptides were checked by UV-visible spectrometer quantitation at a wavelength of 214 nm. All water and buffers used were filtered before use with a Whatman Puradisc 25 mm sterile and endotoxin free 0.2 μm PES filter media syringe filter. DNase I was purchased from NEB. Ethidium bromide, acridine orange, and agarose were purchased from Sigma Chemicals.

Concentrations of peptide and nucleic acid used and the calculation of N/P ratio:

The N/P is the ratio of all the possibly charged nitrogen groups on the H3K4b and H3K(+G)4b peptides to negatively charged phosphate groups on the DNA. There are 20 lysines, 48 histidines and 4 terminal amines on each branch which yield a total of 72 possibly charged nitrogens on both H3K4b and H3K(+G)4b peptides. For the 3 kb prGCN4 plasmid there are 6000 phosphate groups. The above values were then multiplied by the concentration of each component used to come up with a concentration of nitrogen or phosphate groups for each biomolecule. The concentration of nitrogens was then divided by the phosphate concentration to yield all N/P ratios presented in this document. The key ratio to observe in all cases is that N/P = 3 is the ratio at which there is 1:1 charge neutralization of peptide to phosphate backbone of the DNA at pH 7.5. The charge of the peptide is estimated to be +24 at pH 7.5, which assumes that all of the lysines will be protonated. It is important to note the +24 charge is based on the assumption that none of the amino acid side chain

pKas will be altered by the surrounding environment. The assumption that the side chain pKas will not be affected by proximity to DNA or the surrounding environment may not be accurate, making the peptides' charge different from our estimation.

Siliconization:

Peptide sticking to tips and cuvettes proved to have a large impact on results. Therefore, all reactions conducted with peptide and complex formation used siliconized tips and tubes. All cuvettes also had to be siliconized to prevent sticking of peptide and complex to the sides. The procedure for siliconizing the fluorescence and UV-Visible spectrometer cuvettes is in Appendix 1.

2.1.2 Exclusion of Ethidium Bromide and Acridine Orange from Nucleic Acid

A Cary Eclipse fluorometer was used for the fluorescent ethidium bromide (EtBr) exclusion assessment. A siliconized 700 µL Hellma fluorescence cuvette was used to contain all analyzed solutions. A slit width of 20 nm was used and the samples were excited at a wavelength of 335 nm and fluorescence was observed from 450-650 nm, with a maximum around 607 nm. A total sample volume of 500 µL experiment consisted of a 1x TEN (10 mM Tris, 1 mM EDTA, 100 mM or 50 mM NaCl pH 7.5) buffering solution, 0.5 nM DNA (1.5 µM base pairs plasmid) and 5 µM EtBr. A 25 µL volume of a 10 nM DNA stock solution was used to get 0.5 nM of DNA in cuvette. Aliquots of 4 µL from a 1 µM peptide stock were titrated into the cuvette yielding 8 nM additions of H3K4b or H3K(+G)4b. After a total of 32 nM of peptide was added to the DNA, aliquots of 4 µL from a 4 µM peptide stock were titrated into the cuvette yielding 32 nM additions of H3K4b or H3K(+G)4b. After

each aliquot was added, a fluorescence spectrum was taken, and a decrease in fluorescence signal was observed with addition of peptide.

Excitation of fluorescent acridine orange was measured in the same way, except that excitation was at 415 nm and emission spectra were obtained from 500-700 nm. For siRNA experiments, 71 nM siRNA was used in the 500 µL cuvette to yield the same base pair concentration used in the DNA experiments. All fluorescence results are volume corrected by calculating the percent of EtBr-DNA fluorescence left after adding in each aliquot of peptide.

2.1.3 DNase I Degradation of Plasmid

DNase I was used to probe the protection of plasmid DNA-H3K4b and H3K(+G)4b complexes. All reactions and gels were conducted the same day to ensure the most reproducible results. To accomplish this 400 µL reactions with series of different peptide concentrations were prepared with a fixed DNA concentration. All reactions contained 1x TEN (10 mM Tris, 1 mM EDTA, 100 mM, pH 7.5) and 1x DNase I buffer (10 mM Tris-HCl, 2.5 mM $MgCl_2$, 10 mM $CaCl_2$, pH 7.6) in addition to the following DNA and peptide concentrations to yield a series of different charge ratios as in Table 2.1. The N/P describes the ratio of possibly positively charged Nitrogens on the peptide (N) to Phosphates on the backbone of the DNA (P).

N/P	**0**	**0.48**	**0.96**	**1.8**	**3.6**	**4.8**	**7.2**	**9.6**
[DNA] (nM)	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
[H3K4b] or [H3K(+G)4b] (nM)	0	20	40	75	150	200	300	400

Table 2.1 Table of Concentrations Used to Yield each N/P Ratio

After the appropriate amounts of peptide and DNA were combined, the reaction was allowed to incubate for about 45 minutes at room temperature. An 80 µL aliquot was

removed before the DNase I reaction was started to provide a zero time point. A 1:1000 dilution of the original DNase I 1,000 unit stock solution was created. A 50 µL aliquot of the 1:1000 dilution contains 0.0027 units of DNase I, which was then added to each reaction. One unit is defined as the amount of enzyme which will completely degrade 1 µg of pBR322 DNA in 10 minutes at 37 °C in DNase I Reaction Buffer (NEB website). Each reaction mixture was then incubated at 37 ^{o}C for 2, 5 and 10 minutes at the specified DNase I concentration. DNase I reactions for each N/P ratio tube were staggered 20 seconds apart to allow for a buffer region between time points when taking aliquots out of each reaction. A volume of 80 µL was removed at each time point and mixed into 50 µL of quenching solution (stock solution comprised of: 1x TEN, 25 mM EDTA, 0.24 mg Proteinase K) to stop the reaction immediately. Table 2.2 shows the start time points for DNase I addition and all the time points when an 80 µL aliquot was removed (blue), to provide a sample for every N/P ratio at 2, 5 and 10 minutes. This time course scheme was followed for the H3K4b, H3K4(+G)b and DNA control sample in the same exact manner.

N/P	**0.46**	**.90**	**1.7**	**3.4**	**4.6**	**6.8**	**9.1**
DNase I Added	0”	20”	40”	1’	6’20”	6’40”	7’
2’	2’	2’20”	2’40”	3’	8’20”	8’40”	9’
5’	5’	5’20”	5’40”	6’	11’20”	11’40”	12’
10’	10’	10’20”	10’40”	11’	16’20”	16’40”	17’

Table 2.2 Time Course for Incubation of Reaction Mixtures with DNase I

After all of the time course volumes were dispensed into the quenching reaction for each N/P ratio, all samples were allowed to incubate at 37 ^{o}C for at least 15 minutes. After this incubation period, 150 mM NaCl was added to every sample to aid in dissociation of the degraded peptide from the DNA. One volume of phenol/chloroform/isoamyl alcohol (25:24:1 ratio), pH 8, was added to every

reaction, inverted and finger tapped to mix (NOT VORTEXED; as this causes shearing of the plasmid) and spun at 6,000 x *g* for 14 minutes in a tabletop eppendorf 5415D centrifuge. It was spun at 6,000 x *g* because the smaller 600 µL low retention tubes tend to break apart at the higher stresses of 16,100 x *g*. The aqueous phase was then transferred to a fresh 1.5 mL tube (not another low retention tube, as the all the peptide is removed at this point). The remaining phenol:chloroform phase was then back extracted with an additional 100 µL of water to recover any remaining DNA. This mixture was finger tapped to mix and spun at 6,000 x *g* for 14 minutes. This aqueous phase was then added to the previously extracted phase and one volume of chloroform was added to the combined aqueous phase, mixed by finger tapping, and spun at 16,100 x *g* for 14 minutes. The aqueous phase was then transferred to a new tube. Two volumes of 4 ^{o}C 100% EtOH were then added to all samples, mixed, and then allowed to cool in the -80 ^{o}C freezer for 10 minutes. Samples were then spun at 16,100 xg at 4 ^{o}C for at least 20 minutes in a tabletop refrigerated Eppendorf centrifuge. The EtOH was then carefully pipetted off and 200 µL of 70% EtOH was added to the pellet (not visible). The samples were mixed then spun down again at 16,100 x *g* at 4 ^{o}C for at least 20 minutes. The EtOH was then again pipetted off and the samples were dried in a speed-vac for 15 minutes to remove any remaining EtOH. The samples were then dissolved in 10 µL of water with 2 µL of 6 x loading dye (2.5% Ficoll-400, 11 mM EDTA, 3.3 mM Tris-HCl, 0.017% SDS, 0.015% bromophenol blue, pH 8.0) and warmed for 15 minutes at 37 ^{o}C to facilitate redissolving. Samples were then loaded onto a 0.5% agarose gel. The gel was run at 40 V, 100 mA, (approximately 4 W) for 20 minutes in 1x TBE buffer to allow entry

of the DNA into the gel (this proved to give cleaner bands). Once the DNA entered the gel at lower voltage, the gel was set to 80 V, 100 mA, (approximately 8 W) for 2 hours. The gel was then placed into a dish with enough 1x TBE buffer to cover the top of the gel (about 300 mL) and 25 μL of 25 mM EtBr solution was added for staining. Staining proceeded for 15 minutes, and then the staining solution replaced with distilled water and the gel was allowed to destain for 10 minutes. The gel was then visualized on a UV transilluminator and a photograph taken with a digital camera gel imaging system.

2.1.4 Image J Quantitation of DNase I Degradation Gel

An Image J gel quantitation of the percentage of supercoiled DNA left after DNase I exposure was carried out as follows. All images were adjusted for color and contrast in Adobe Photoshop. Images were then imported into Image J and further contrast adjustment was made before the gels were quantitated. The gel image was then background subtracted by using the *subtract background* option under the *process* tab. A Box was then drawn around the first lane in the row, encompassing the supercoiled, nicked and linear bands. The first lane with the box drawn around it was selected by using the *select first lane* option under *analyze,* then *gel* tabs. Subsequent lanes in the row were then selected by moving the box over the next lane and then using the *select next lane* option under the *analyze*, then *gel* tabs. Once all lanes were selected in a row, the *plot lane* option was then used under the *analyze*, then *gel* tabs. A plot of all the peaks for each lane is then displayed. The gel analysis routine averages the results over each line within the box, providing the equivalent of volume integration over each band. A line was then drawn under each peak to isolate

the volume of the peak for analysis. After all peaks were isolated, the wand tool was used to measure the area of supercoiled, nicked and linear bands for each lane. The supercoiled, nicked and linear areas were added up to get the total area of plasmid DNA bands in each lane. The supercoiled area was then divided by the total area for each lane to come up with a percentage of supercoiled DNA left in each lane after exposure to DNase I.

2.2 Results:

2.2.1 EtBr Exclusion Resulting from H3K4b and H3K(+G)4b Binding of Plasmid

The Ethidium Brominde (EtBr) exclusion experiments were designed to obtain a binding constant for H3K4b and H3K(+G)4b binding to nucleic acid. We also wanted to observe any differences in the ability of both peptides to displace EtBr upon binding to nucleic acid and relate it to their transfection efficiency. Based on the in vivo data, we speculated that H3K(+G)4b would coat the DNA more completley and exclude more EtBr than H3K4b. From the fluorescence exclusion experiment we were not able to derive a binding constant for either peptide because they bind too tightly to plasmid, however, we were able to observe differences in EtBr displacement. Exclusion of EtBr from plasmid DNA with binding of H3K4b and H3K(+G)4b demonstrated a drop in fluorescence intensity as a function of increasing N/P ratios. Quenching in fluorescence intensity is primarily due to displacement of EtBr from intercalation sites between the base pairs of the plasmid[26]. EtBr displacement is believed to occur due to compaction of the plasmid during

complex formation. No interaction was observed between either H3K4b or

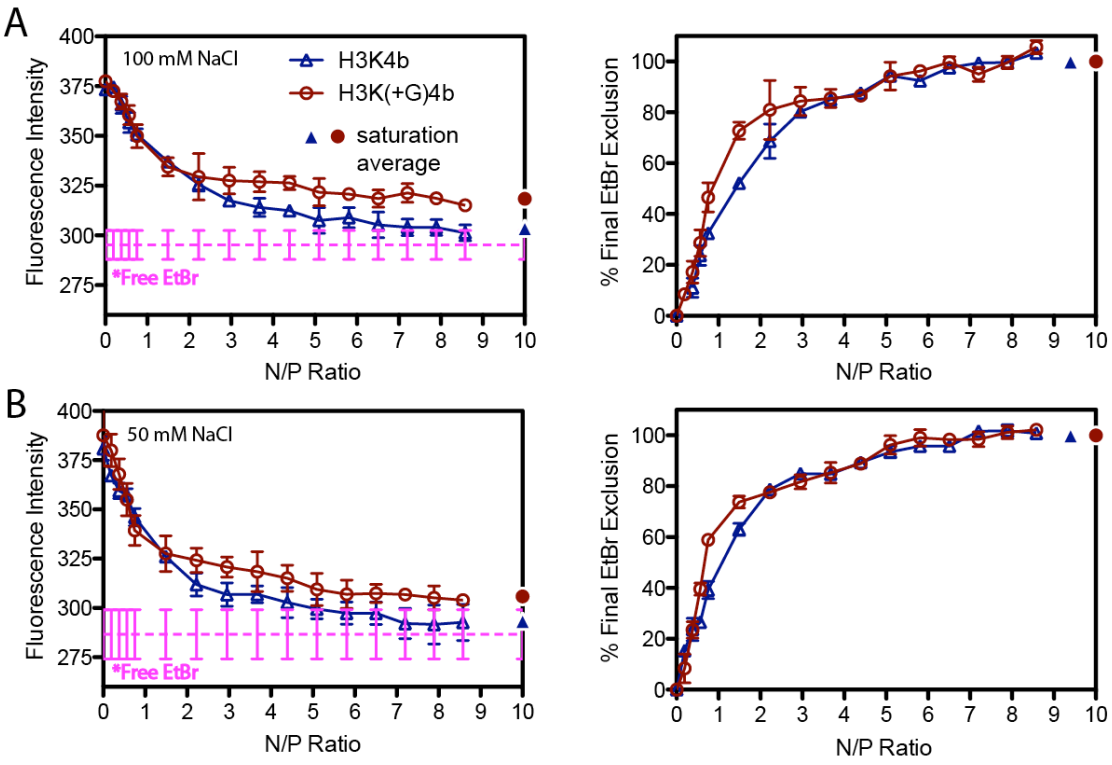


Figure 2.1 ***Exclusion of EtBr bound to DNA through H3K4b and H3K(+G)4b peptide binding.*** **EtBr bound DNA was titrated with increasing concentrations of peptide presented as the N/P ratio. All spectra have been corrected for concentration and fluorescence changes due to dilution. In the left panels, the fluorescence intensity is plotted vs. N/P ratio. A decrease in fluorescence is observed as peptide is added. The error bars are the standard deviation from an average of two trials for the 100 mM NaCl and three trials for the 50 mM NaCl EtBr exclusion experiments. The dashed pink line is the EtBr fluorescence alone (without DNA). Curves for H3K4b are presented in blue and those for H3K(+G)4b, in red. A and B are the experiments for two different NaCl concentrations, 100 mM and 50 mM. At both salt concentrations, H3K4b shows a shallower initial drop in fluorescence, but it reaches a greater final extent of EtBr exclusion than H3K(+G)4b. H3K(+G)4b demonstrates a sharper drop in fluorescence at a lower N/P ratio and plateaus at its full extent of EtBr exclusion at lower N/P ratios than H3K4b. The graphs to the right present the same data, normalized to the final extent of EtBr excluded, with 100% being the maximum extent of EtBr exclusion from the plasmid for each peptide. The saturation average is the average of the last four data points from each titration, which yields our value for 100% EtBr displacement. These data clearly demonstrate greater exclusion of EtBr upon H3K(+G)4b binding to DNA than H3K4b at lower N/P ratios (0-2) N/P. This indicates that H3K(+G)4b covers more base pairs of DNA at lower peptide ratios than H3K4b.**

H3K(+G)4b and EtBr alone (data not shown), indicating that the mechanism of fluorescence decrease is most likely due to exclusion of EtBr, not quenching through electron transfer or other non-radiative transmission to the peptide. Therefore,

exclusion of EtBr is representative of peptide-plasmid DNA complex formation[26]. A saturating concentration of EtBr was used to bind to DNA.

The overall trend in fluorescence quenching between the two peptides was similar, with some differences in ability to exclude EtBr as presented in Figure 2.1. The entire fluorescence titration is presented as the fluorescence intensity as a function of increasing N/P ratio. For both peptides most of the fluorescence decrease through EtBr exclusion is occurring between 0 and 3 N/P ratio, 3 being the point of charge neutralization. H3K(+G)4b reaches essentially its final extent of EtBr exclusion by N/P = 3, and is slightly more effective at excluding a greater amount of EtBr at lower peptide concentrations than H3K4b. H3K4b requires more peptide to excluded the same amount of EtBr, but the fluorescence is decreased more at saturating H3K4b than at saturating H3K(+G)4b. Fluorescence never completely reaches the level it starts out with before binding to DNA. This is most likely due to residual EtBr still bound to the DNA even after peptide binding. In the 100 mM NaCl experiments, the average percent of EtBr excluded was calculated to be 97% for H3K4b and 90% for H3K(+G)4b. In the 50 mM NaCl experiments, the average percent of EtBr excluded was calculated to be 98% for H3K4b and 92% for H3K(+G)4b.

To emphasis H3K(+G)4b's greater exclusion of EtBr at lower N/P ratios, the same data was plotted in terms of the final extent of excluded fluorescence at saturating N/P ratios. The last four points of the fluorescence curve were averaged (denoted as $(F)_{Avg}$) to give the averaged signal for plasmid DNA at which, EtBr

exclusions had reached its full extent. The following equation was then used to obtain the percent of EtBr excluded from the DNA at each N/P ratio:

$$\frac{(F)_I - (F)_{Exp}}{(F)_I - \left((F)_{Saturation} - (F)_{avg}\right)} * 100 = \% \text{ EtBr Excluded}$$

Where $(F)_I$ is the initial fluorescence and $(F)_{Exp}$ is the experimental fluorescence at each N/P ratio. The percent excluded EtBr graphs are presented to the right of the original data in Figure 2.1. This presentation of the data clearly demonstrates H3K(+G)4b's enhanced ability to exclude EtBr between N/P = 0.5 and 2. H3K(+G)4b is suggested to allow more of the positive charges on the peptide to come in close contact with the DNA due to the added flexibility afforded by the glycines, which causes more EtBr displacement at lower peptide concentrations. The final extent of EtBr displacement occurs at a higher level of fluorescence for H3K(+G)4b because of protected sites for EtBr dye that are maintained during binding. In comparison, at lower N/P ratios of H3K4b, plasmid DNA is not coated as effectively, leaving exposed regions of plasmid still bound by EtBr. Once enough H3K4b is added, these gaps are bound and more EtBr can be displaced, leading to the greater overall decrease in fluorescence observed. This is the first piece of data to lead to the idea for differences in coating for the two peptides.

In addition to assessing the two peptides binding to plasmid, the EtBr exclusion was done at two NaCl concentrations, 100 mM and 50 mM, to observe possible electrostatic effects on peptide binding. Both NaCl concentrations demonstrate a similar extent of EtBr exclusion from both H3K4b and H3K(+G)4b upon binding plasmid.

2.2.2 Sequential vs. Single-Step addition of H3K4b and H3K(+G)4b to Plasmid

In vivo nanoplexes are prepared by adding the full amounts of plasmid and peptide together needed to form a transfection complex, not by titration. Thus, it was important to observe any differences between sequential titration compared to single-step addition of the peptide in the EtBr fluorescence exclusion analysis (Figure 2.2). An EtBr displacement experiment was conducted in the same manner as presented in section 2.2.1. An experiment was then conducted at the same time, adding in a 1.8 and

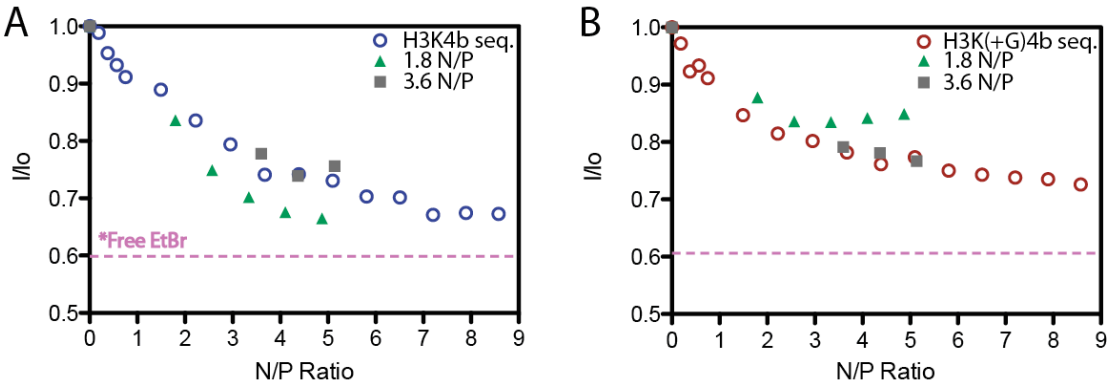


Figure 2.2 ***Sequential compared to single-step addition of H3K4b and H3K(+G)4b peptides to plasmid DNA*****. Fluorescence is plotted as the fluorescence intensity of the sample (I) divided by the starting fluorescence of EtBr bound to DNA (Io) before peptide was added. The plotted blue and red open circles are the results for the sequential addition of H3K4b and H3K(+G)4b to plasmid DNA. The pink dashed pink line is the EtBr fluorescence alone (without DNA). The green triangles are data for a single-step addition of peptide at 1.8 N/P. The single step addition was followed by a titration of increasing peptide to see if we reach the same level of quenching as in the sequential addition. The grey squares present the same single-step experiment starting off with an addition of a 3.6 N/P ratio. A. H3K4b added to plasmid DNA sequentially shows different exclusion for EtBr compared to single-step at N/P ratios of 1.8. B. H3K(+G)4b added sequentially to EtBr bound plasmid shows a similar extent of exclusion compared to single-step addition. The 1.8 N/P ratio shows a deviation from the sequential addition at N/P between 3 and 5.**

3.6 N/P ratio of peptide to DNA instead of titrating in small aliquots of peptide. After the initial single-step addition of peptide was completed, subsequent aliquots of peptide were added as in the titration experiments to observe additional effects on fluorescence. If the peptide bound to DNA consistently excludes the same amount of

EtBr at a given concentration, then the order of peptide addition should not have an impact on the levels of fluorescence observed. We end up observing subtle differences in the levels of fluorescence reached with single-step addition followed by titration compared to titration alone with H3K4b and H3K(+G)4b. With H3K4b at the 1.8 N/P ratio, exclusion of EtBr occurred to a greater extent than in the sequential addition, indicating some variability during complex formation with plasmid. Whole addition of H3K4b presented a similar decrease in fluorescence compared to the sequential addition at a 3.6 N/P ratio. On the other hand, H3K(+G)4b shows EtBr exclusion values for the single-step addition of peptide comparable to sequential addition. The only real variation observed is that after more peptide is added to the 1.8 N/P single-step addition, fluorescence levels off at lower N/P. The larger decrease in fluorescence observed in the EtBr exclusion at N/P = 1.8 of the single-step addition of H3K4b is most likely due to the peptides inability to completely coat the DNA, leaving an open network of DNA for further binding of EtBr. For H3K(+G)4b single-step addition at 1.8 N/P, there is better coverage of the DNA, forming a shell that does not allow any further exclusion of EtBr.

2.2.3 AO bound Plasmid Quenching Via H3K4b and H3K(+G)4b Binding

A similar experiment to the EtBr exclusion assay was conducted with observing exclusion of a different fluorescent dye, Acridine Orange bound to plasmid DNA for two major reasons. The first was to corroborate our findings for the plasmid EtBr exclusions experiments. The second was to develop an assay in which we can observe exclusion of a fluorescence dye bound to siRNA, which we were not able to do with EtBr. The AO fluorescence experiments with plasmid DNA are useful to

assess because of the data's agreement with the results observed for EtBr exclusion during peptide binding. An AO concentration was chosen based on the maximum fluorescence intensity we could obtain while bound to plasmid DNA. Presented in Figure 2.3, H3K4b and H3K(+G)4b quenching of fluorescent AO bound to plasmid DNA at 100 mM and 50 mM NaCl. AO proved to be difficult dye to work

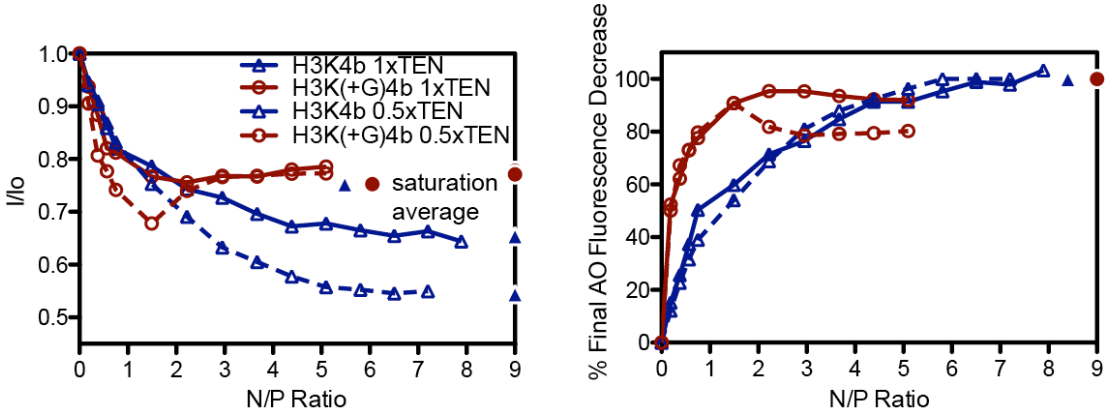


Figure 2.3 ***Fluorescence Decrease of AO bound to DNA through H3K4b and H3K(+G)4b Binding.*** **(Left) H3K4b (blue) and H3K(+G)4b's binding to plasmid DNA was assessed by quenching of bound AO fluorescence. Results proved to be similar to that of the EtBr exclusion assay. H3K(+G)4b demonstrates a greater extent of fluorescence decrease at a lower N/P than H3K4b. H3K4b quenches more AO fluorescence at saturating N/P ratio. These results correlate well with the EtBr results; however, by themselves are not a good measure of peptide binding. The same data was plotted again in the same manner as the EtBr displacement curve in Figure 2.1. The graph to the right present the same data, normalized to the final extent of AO's decrease in fluorescence. H3K(+G)4b demonstrates a significantly greater decrease in AO fluorescence at lower N/P ratios than H3K4b.**

with for two major reasons (which will also be discussed more in section 2.2.4 and 2.3.6). One is that both peptides were found to bind AO and cause a decrease in fluorescence in the absence of nucleic acid. This makes it difficult to assess whether a decrease in fluorescence is coming from exclusion of the AO from the DNA or quenching of the dye from the peptide. The decrease in fluorescence signal observed for the DNA-AO exclusion went below the free AO fluorescence signal. The other is that AO alone fluorescence compared to AO with nucleic acid was not very different. As with EtBr we expected to see a large increase in fluorescence of the dye with

association to nucleic acid, however we did not observe this. H3K4b shows less exclusion and subsequent decrease in fluorescence at lower N/P ratios as peptide is titrated in. At saturating N/P ratios of H3K4b, exclusion of AO is greater than that of H3K(+G)4b. For H3K(+G)4b, full exclusion of AO is reached at lower N/P ratios and plateaus in fluorescence earlier than H3K4b. One result that was not observed in the EtBr assay that is presented here is a significant reemergence of fluorescence upon binding of H3K(+G)4b at higher N/P ratios. We know at lower concentrations, H3K(+G)4b is binding to the DNA which causes fluorescence quenching through exclusion of AO. We believe, as the concentration of H3K(+G)4b increases, more peptide binds to the free AO, therefore an interaction of the peptide with the free AO must be causing the reemergence of fluorescence observed.

2.2.4 AO bound siRNA Quenching Via H3K4b and H3K(+G)4b Binding

Attempts to use EtBr to probe peptide binding to siRNA were unsuccessful and yielded very little increase of fluorescence signal EtBr bound siRNA compared to free dye. It became evident after attempts to observe AO exclusion from plasmid and siRNA, this would also not be an ideal system to assess siRNA binding to peptide. One major problem stems from the fact that AO binding to siRNA provided little increase in the fluorescence signal of the bound dye over that of EtBr. The other is that AO fluorescence is not only decreased in the presence of peptide through exclusion, but the peptide itself also causes quenching of the AO fluorescence in the

absence of nucleic acid (Figure 2.4 A). Figure 2.4 B presents the results from the

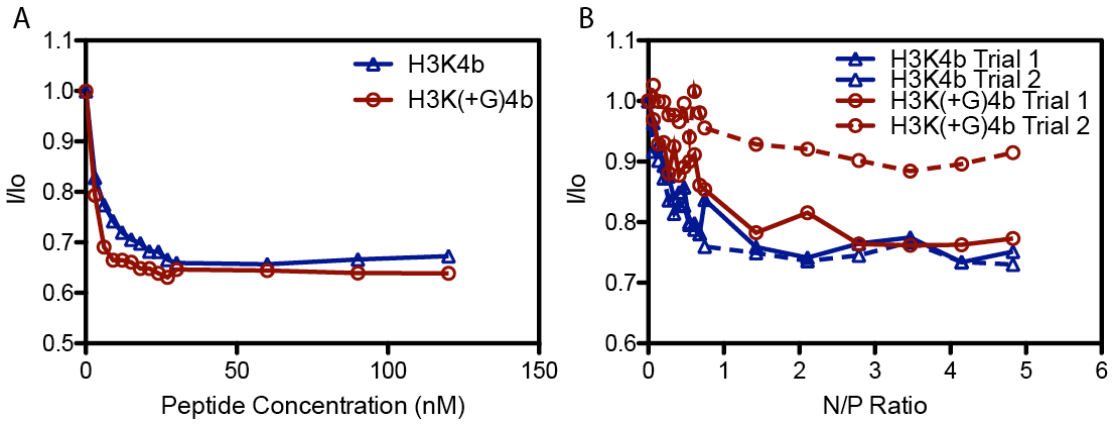


Figure 2.4 ***Fluorescence Decrease of AO in the Absence and Presence of siRNA*****. A. There is a Decrease in AO fluorescence through binding of H3K4b and H3K(+G)4b. B. A Fluorescence Decrease is observed for AO bound to siRNA through H3K4b and H3K(+G)4b Binding. In panel A, Both H3K4b (blue) and H3K(+G)4b show a strong ability to quench peptide in the absence of nucleic acid. The concentrations of peptide used here are in the same range as used for siRNA experiments. This suggests a strong interaction between both of these peptides and the AO dye. This makes analysis of peptide binding to nucleic acid with AO difficult. In panel B, Decreasing in AO fluorescence appears to be greater for H3K4b binding to siRNA than that of H3K(+G)4b. However, there is a large error for the two trials of H3K(+G)4b fluorescence decrease conducted. Excluding the second trial for H3K(+G)4b, results for both peptides binding to siRNA are similar. The fact that both peptides bind and possibly quench AO themselves makes this assay a problematic method for assessing peptide binding to siRNA.**

exclusion of AO from siRNA from H3K4b and H3K(+G)4b binding. A similar degree of AO exclusion and fluorescence decrease is observed for the two trials of H3K4b and one trial of H3K(+G)4b binding siRNA. This assessment is made complicated by the fact one trial for the H3K(+G)4b peptide demonstrates much less of a decrease in fluorescence than the other. The resulting fluorescence decrease during addition of H3K4b and H3K(+G)4b to AO-siRNA prove to be inconclusive in distinguishing differences between peptide binding. There is a definitive interaction of H3K4b and H3K(+G)4b with siRNA but extracting real differences between their binding cannot be accomplished from this experiment.

2.2.5 Peptide Protection of Plasmid from DNase I

The EtBr exclusion results were the first indication that H3K(+G)4b coats plasmid DNA more completley than H3K4b. This motivated a DNase I assay to

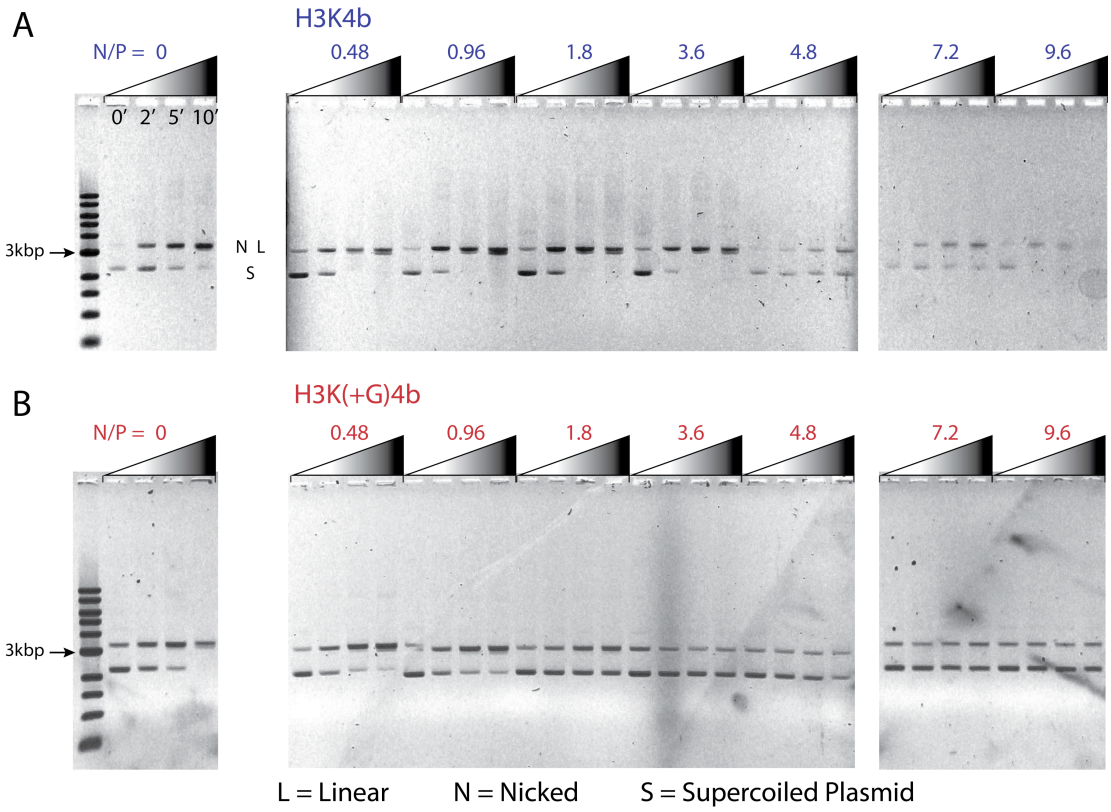


Figure 2.5 ***DNase I degradation of Plasmid in complex with H3K4b and H3K(+G)4b*****. EtBr stained 0.5% Agarose Gels of H3K4b (A) and H3K(+G)4b (B) peptide bound plasmid DNA complex with varying peptide concentrations. Complexes were exposed to DNase I for 0, 2, 5 and 10 minutes as illustrated by the triangle gradient. Lanes are labeled by their corresponding N/P ratio, with increasing concentration of peptide going from lanes left to right, with the furthest left containing DNA alone. Disappearance of the supercoiled band (S) as it gets converted into nicked and linear DNA (N L) indicates the Peptide-DNA complex susceptibility to degradation. A. H3K4b complex formation with DNA demonstrating lessened protection of the DNA. Protection of the DNA with H3K4b appears to peak around a 4.8 N/P ratio of peptide to DNA. At ratios below 4.8 there is poor protection of supercoiled plasmid from DNase I after 2 minutes of incubation. Above a 4.8 N/P ratio there also appears to be some loss in yield or protection of the DNA afforded by higher concentrations of H3K4b. B. H3K(+G)4b complex formation with DNA demonstrates much better protection of DNA. After a 1.8 N/P ratio is reached, there is almost complete retention of the supercoiled DNA, even after 10 minutes of DNase I exposure.**

assess if H3K(+G)4b protected plasmid DNA better than H3K4b. Conducting this experiment at the same time and at the same peptide and DNA concentrations as the

EtBr displacement assay gives us a direct comparison between the extent of protection and the amount of EtBr displaced.

Figure 2.5 presents H3K4b's and H3K(+G)4b's ability to coat and protect plasmid DNA. Plasmid DNA was incubated in the presence of 20 nM-400 nM peptide as in Table 2.1 of the Materials and Methods. These concentrations yielded N/P ratios from 1/6 of the plasmid covered up to 3 times plasmid coverage. The ratios tested allowed us to determine at what ratio of peptide the plasmid showed optimal protection. At much higher ratios of H3K4b, greater than a 3 N/P ratio, (1:1 charge ratio) the possibility exists for formation of positively overcharged particle due to aggregation, which may explain the loss in recovery of the DNA at high N/P for H3K4b. The peptide-DNA complex was allowed to form for 45 minutes and was then incubated with DNase I for 0, 2, 5 and 10 minutes to observe the accessibility of the DNA. The bands in the agarose gel show plasmid DNA after phenol extraction to remove all peptide. DNA alone (not in complex with peptide) is the 0 N/P, which shows that the plasmid is initially found largely as supercoiled (lowest band) with some nicked DNA (upper band). The first lane of each N/P ratio set with actual peptide in it (0.48-9.6), is of DNA that has been incubated with peptide, then extracted without ever coming in contact with DNase I. This was an important control lane because it shows that DNA can be recovered from a complex with either peptide. It also verifies that the peptide itself does not cause degradation of the plasmid, and the only conversion of supercoiled to nicked DNA is coming from DNase I. Plasmid DNA recovered from a peptide complex without DNase I treatment shows roughly the same nicked to supercoiled ratio in the DNA alone

control lanes. DNA at the 0 N/P ratio after 10 minutes of DNase I incubation demonstrates significant conversion of

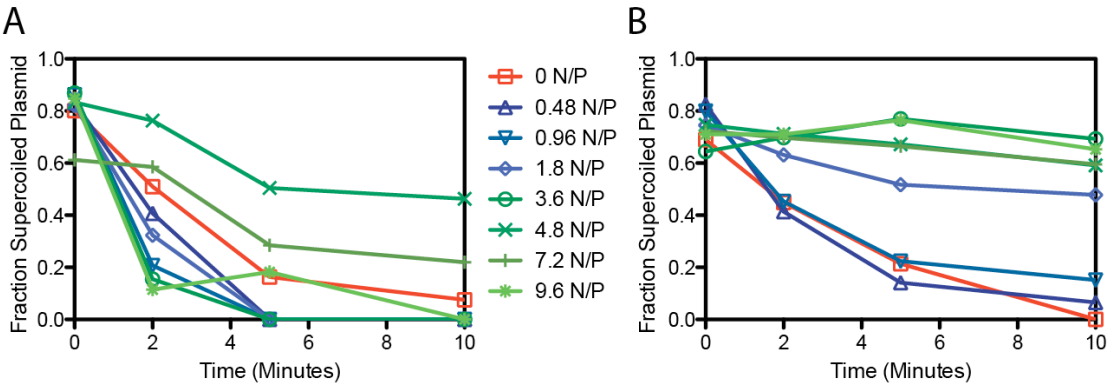


Figure 2.6 ***Fraction of Supercoiled DNA Left After Incubation with DNase I.*** **The fraction of supercoiled DNA that remains at each N/P ratio for H3K4b (A) and H3K(+G)4b (B) is plotted for 0, 2, 5 and 10 minutes of DNase I incubation. The no peptide control is presented in red. N/P ratios below a 1:1 charge ratio are plotted in a blue color and ratios above a 1:1 charge ratio are plotted in green. A. H3K4b's protection of supercoiled plasmid DNA peaks at a 4.8 N/P ratio. Ratios above and below this show a significant decrease in the percentage of supercoiled DNA due to increased accessibility of the plasmid DNA cause by this peptide. B. H3K(+G)4b's protection of supercoiled DNA starts to show enhancement over that of naked plasmid at a 1.8 N/P ratio. Above the 1.8 ratio, there appears to be complete protection of supercoiled plasmid by H3K(+G)4b.**

supercoiled plasmid into nicked and linear. The objective was to use enough DNase I degrade supercoiled into nicked DNA, but not enough to completely destroy the plasmid. A single DNase I cleavage event on the plasmid DNA causes a conversion of supercoiled to nicked DNA, resulting in a very sensitive assay for the completeness of peptide protection of DNA.

Both peptides do provide protection of the DNA, however, at varying degrees with different N/P ratios. Figure 2.6 shows the fraction of supercoiled plasmid DNA remaining due to peptide protection by H3K4b and H3K(+G)4b, obtained by Image J analysis from Figure 2.5. The fraction of supercoiled, nicked and linear population of plasmid DNA were determined by measuring the integrated volume intensity of each band. An example of the peak for the area that is representative of a gel band is

presented in Figure 2.7. As seen in Figure 2.5, the 10 minute lane for

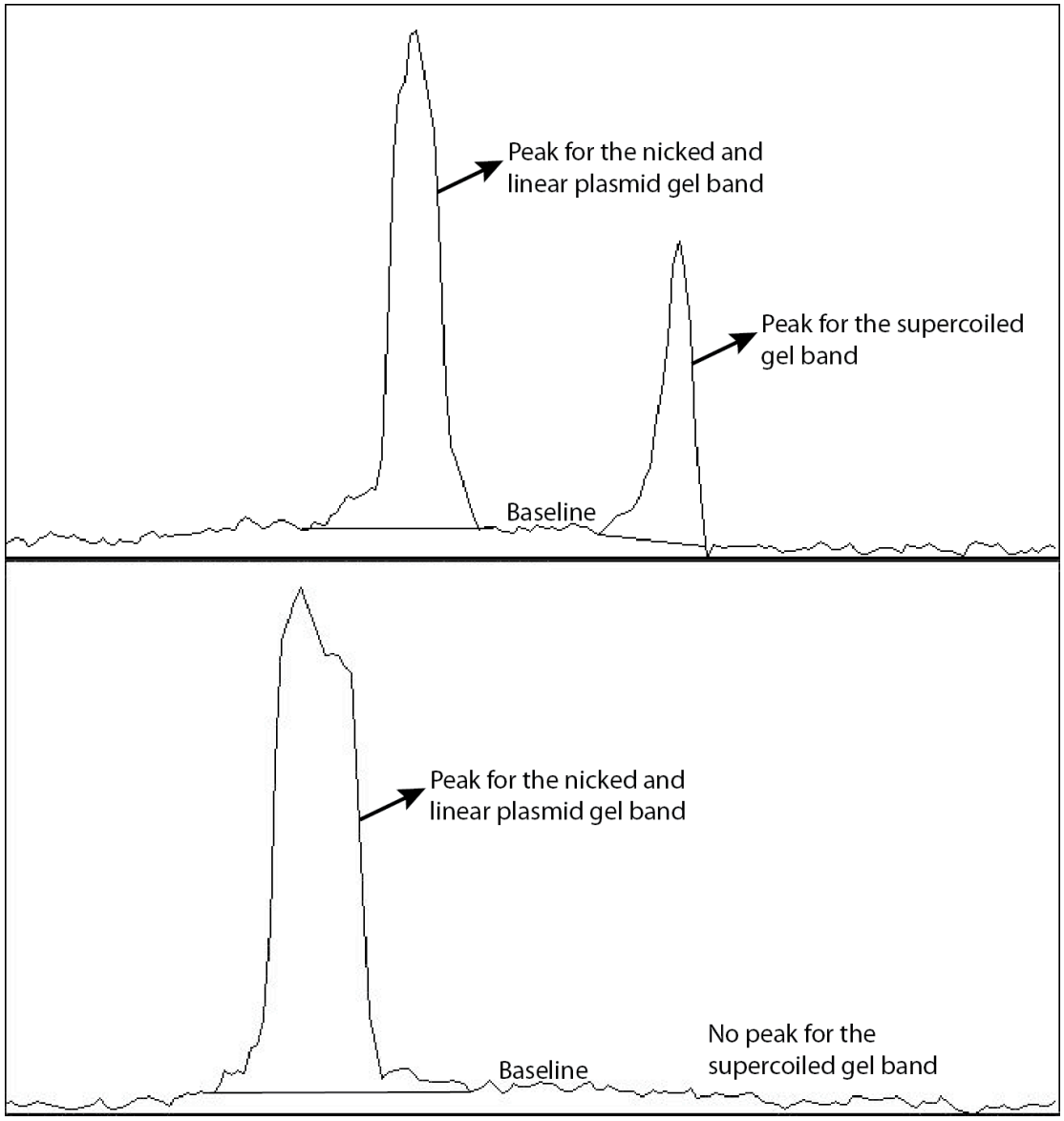


Figure 2.7 ***Example of the Area of a Peak for the Image J analysis of the DNase I Degradation Gel.*** **The area of the peak presented in the top panel is for the nicked, linear and supercoiled gel band for a 1.8 N/P ratio at 2 minutes of DNase I incubation and for the bottom panel, the nicked and linear gel band of the 0.96 N/P ratio of H3K4b at 10 minutes of DNase I incubation. There is a baseline for the background in the gel that is not included in the peak area measurement for each band. A line is drawn under every peak to eliminate the background. In the top panel there is a peak for the nicked, linear and supercoiled DNA. In the bottom panel a second peak would normally be observed for the supercoiled DNA band, however for this sample, all plasmid DNA was converted to nicked and linear. Therefore we observe no peak for the supercoiled DNA.**

0.96 N/P ratio for H3K4b shows only a population for nicked and linear DNA and no band for supercoiled. Figure 2.7 shows that for the 0.96 N/P lane for H3K4b there is

only a peak for the nicked and linear DNA and not the supercoiled, therefore lanes like this one had 0% supercoiled DNA. Charge neutralization of the peptide-DNA complex occurs around a 3 N/P ratio. As shown in by Figure 2.6, ratios of H3K4b up to 3.6 all allow a significant percent conversion of the supercoiled population into nicked and linear plasmid. Conversion of the plasmid DNA into nicked DNA is apparent even after 2 minutes of incubation with DNase I at the lowest ratios of peptide, which is an indication of very poor protection at low concentrations of H3K4. At lower H3K4b concentrations, protection of the plasmid DNA is very similar to that of naked DNA. Optimal coating of DNA appears to occur around a 4.8 ratio of peptide to plasmid. Above a 3.6 N/P ratio, there is a clear decrease in the amount of DNA present in the gel, indicating incomplete recovery of the plasmid from H3K4b. It is difficult to say with certainty because of the incomplete recovery of plasmid, but protection of H3K4b appears to peak at an optimal N/P ratio. Below this ratio, there is insufficient coating within the complex to protect the DNA from DNase I. Above this optimal ratio, loss in DNA recovery could be attributed to excessive aggregation of the peptide with plasmid. At higher concentrations, the peptide may interact with itself in a manner that causes formation of very tight aggregates to occur which makes recovery of the DNA more difficult. The aggregates formed by H3K4b could very well be attributed to the peptides less flexible structure. Overall, coating of DNA with H3K4b more incomplete at increasing concentrations than that of H3K(+G)4b.

H3K(+G)4b demonstrates a better ability to coat with increasing HK:DNA ratios. The results presented for H3K(+G)4b are in sharp contrast to those observed

for H3K4b. At low HK:DNA ratios (0.48 and 0.96), there is a limited amount of protection due to coating by H3K(+G)4b. This is expected since full charge neutralization has not been attained at these ratios. These are the only ratios at which we see accessibility of DNase I to plasmid DNA protected by H3K(+G)4b. However, at a 1.8 N/P ratio (a little over half charge neutralization of plasmid) there is complete protection of the DNA over the full 10 minutes of DNase I exposure. Ratios above N/P of 1.8 all show a similar level of protection for plasmid. Between a ratio of 3.6 and 4.8, there is slightly better protection of the supercoiled plasmid from conversion to nicked and linear DNA, demonstrating superior protection of plasmid DNA compared to H3K4b. H3K(+G)4b completely protects the DNA at a much lower HK:DNA ratio. Once protected, there is very little to no degradation observed for the plasmid.

2.3 Discussion:

2.3.1 Comparison of H3K4b and H3K(+G)4b Protection to other Delivery Agents

A host of gene and siRNA delivery agents exist that function to protect and deliver nucleic acids to their cellular targets. Assessment of a gene delivery peptide's ability to coat a nucleic acid is of primary importance. Without the packaging and protection afforded by the peptide, delivery of siRNA or plasmid for RNA interference or gene expression would be extremely inefficient due to rapid degradation[11; 38]. Assays, either utilizing RNases or DNases to degrade RNA or DNA have proven very effective at probing the protection of nucleic acids formed with protective agents[39; 40]. One assay used nanoparticles to protect plasmid DNA from

cleavage. Plasmid was combined with nanoparticles and a nuclease was added to probe for protection[39]. DNA was isolated from the nanoparticles and then run on a gel[39]. The results demonstrated complete recovery of the supercoiled, linear and nicked plasmid in the same ratios as seen with the undigested DNA, which was almost identical to the results presented for H3K(+G)4bs complete protection of plasmid[39]. Similar experiments have also been conducted for protection of siRNA with PEI. PEI-siRNA complex demonstrates complete protection of the siRNA even at up to 2 hours of exposure to RNase A[40]. Protection assays for siRNA present simple results that report on preservation or degradation of nucleic acid but are not sensitive to nucleic acid topology[40; 41]. These assays form a good foundation for assessment of protection.

2.3.2 Comparison of siRNA and Plasmid DNA Protection Assays

The size difference of plasmid DNA and siRNA present important considerations to the assessment of H3K4b and H3K(+G)4b-nucleic acid complex formation. As a moderate sized nucleic acid, plasmid DNA has distinctive topological attributes in which it can assume supercoiled, nicked and linear forms. It will also take many (~300) peptides to completely coat and package a single 3 kb plasmid molecule. Cleavage of DNA into different topological states make this assay more sensitive than its siRNA counterpart. Nicking of DNA indicates only damage to one strand of DNA has occurred while linearization of the DNA indicates both strands have been cut. H3K4b and H3K(+G)4b demonstrates a different ability to coat DNA based on exposure to DNase I. In order for protection of the DNA from DNase I cleavage to occur, each of these peptides must be excluding the accessibility

of DNase I through coverage of the plasmid. DNase I needs at least 10 base pairs of surface area to effectively and completely cut DNA[42]. This presents the result that H3K(+G)4b leaves sections of DNA less than 10 bp in size exposed, while H3K4b is leaving segments 10 bp or larger exposed. This would suggest that delivery agents coating does not have to be absolute in coverage of the DNA, just enough to exclude access of degradative elements. Observing the transition of supercoiled DNA to nicked and linear provides more information about the assessment of the peptides flexibility and coating ability. This is in contrast to siRNA, in which a very small 21 bp nucleic acid will have very little unique topology to probe in binding of an individual siRNA molecule. A single 21 bp siRNA needs far fewer peptides (about 2) to coat and package. The assessment becomes whether siRNA is either packaged and protected or not properly packaged and exposed to degradation.

2.3.3 EtBr Exclusion via H3K4b and H3K(+G)4b in Light of DNase I Degradation Results

The observations that H3K4b excludes more dye at saturating concentrations than H3K(+G)4b, yet H3K(+G)4b protects the plasmid DNA better than the H3K4b provide an interesting contrast. One might expect instead that the peptide that coats nucleic acid the best will also exclude more dye. The experiments were conducted at several N/P ratios chosen to cover the range that produced the full range of binding based on the EtBr exclusion experiments. The H3K(+G)4b peptide has demonstrated a trend of reaching the final extent of dye exclusion yielding a greater decrease in fluorescence at low N/P ratios in the EtBr and AO fluorescence exclusion experiments. The key N/P ratio to note here is the 3 N/P ratio, which corresponds to

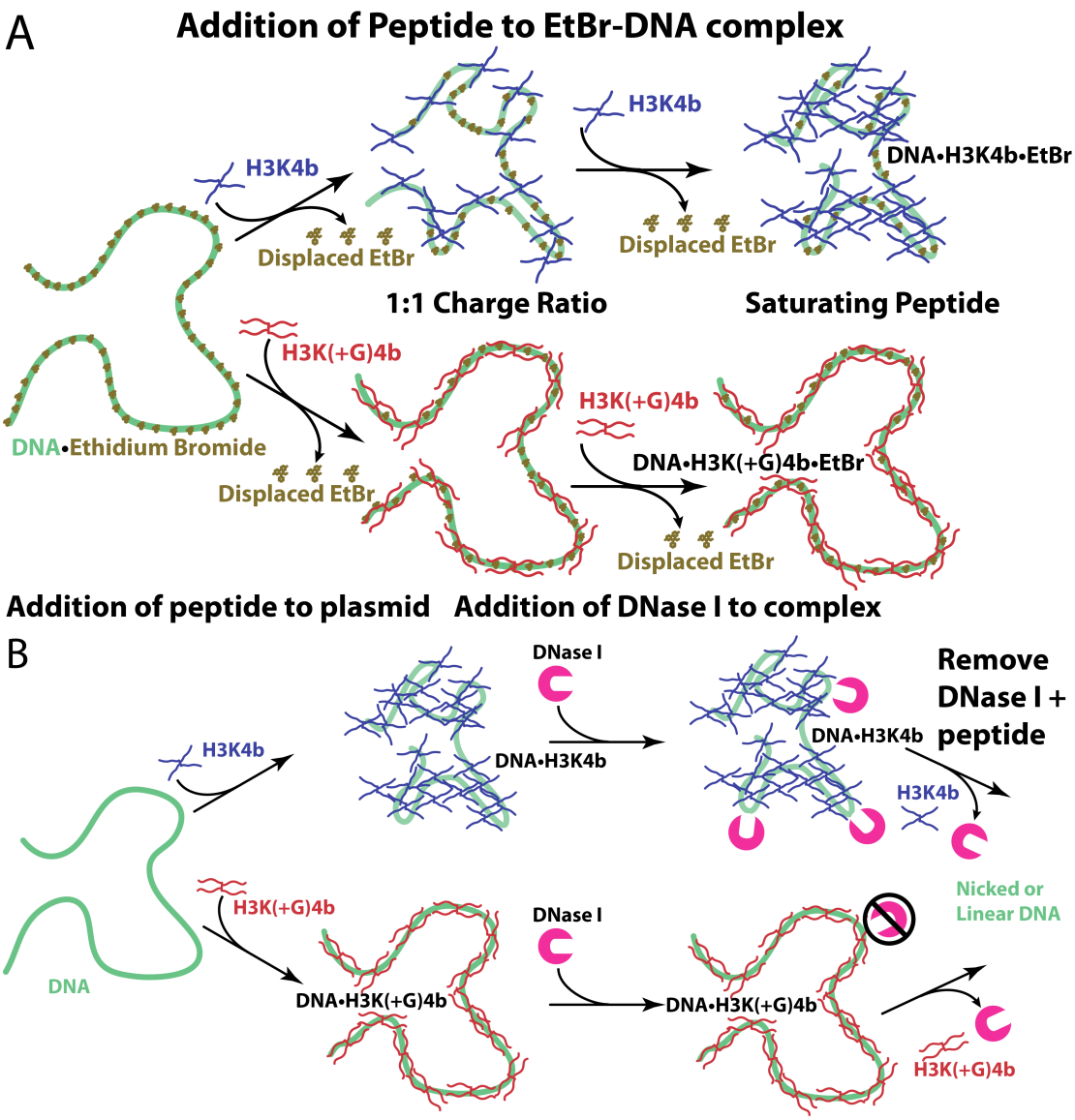


Figure 2.8 ***Proposed mechanisms of H3K4b and H3K(+G)4b EtBr plasmid exclusion and plasmid protection from DNase I Degradation*****. A. H3K4b and H3K(+G)4b displace EtBr from plasmid DNA. H3K4b does not coat the plasmid DNA as effectively leading to less initial exclusion of EtBr than for H3K4b. Upon addition of saturating concentrations of H3k4b, gaps where EtBr was still bound are filled in by peptide, excluding more EtBr than H3K(+G)4b at saturating concentrations. Due to H3K(+G)4bs decreased steric bulk, during binding it traps packets of EtBr still bound to DNA. Upon addition of saturating concentrations of H3K(+G)4b there is less exclusion of EtBr because of these EtBr pockets trapped by peptide. This leads to less overall exclusion of EtBr for H3K(+G)4b. B. DNase I plasmid degradation of H3K4b and H3K(+G)4b protected DNA complexes. H3K4b coats plasmid more unevenly, leaving exposed regions of plasmid for DNase I accessibility and degradation. H3K(+G)4b coats plasmid more evenly and provides better protection, limiting the accessibility of DNase I to the plasmid. The overall result is more degradation of H3K4b protected plasmid compared to the superior protection of plasmid afforded by binding H3K(+G)4b.**

a 1:1 charge to charge ratio of peptide:plasmid DNA. In both the EtBr exclusion and DNase I degradation experiments, this denotes an important ratio where we see the significant differences in binding of the two peptides. At this ratio, EtBr exclusion from H3K(+G)4b reaches its full extent and H3K(+G)4b fully protects plasmid DNA from DNase I. H3K4b is believed to coat the plasmid less evenly, resulting in gaps in the DNA between coated regions at lower peptide concentrations, which will still contain plasmid bound EtBr. Condensation will exclude more EtBr at higher N/P ratios, resulting in an overall greater quenching of fluorescence for H3K4b. H3K(+G)4b results demonstrate better protection of the DNA and more efficient transfection, however, how does this result in less exclusion of EtBr? The only plausible explanation is the presence of some EtBr left bound even after H3K(+G)4b binding to plasmid. It has been previously demonstrated that a nucleic acid-dye-binding agent complex may aggregate before all EtBr is excluded from a complex[43]. Two possible factors contributing to H3K(+G)4b yielding higher end point fluorescence are protected pockets of EtBr trapped in the DNA during peptide binding and small gaps remain after complete assembly, which allow EtBr association with DNA. It has been demonstrated that H3K(+G)4b coats DNA much better at lower N:P ratios, indicating that rapid coating of the plasmid by this peptide can lead to some trapped EtBr molecules in the DNA. In addition coating for this peptide is superior, but may not be complete.

Figure 2.8 illustrates a proposed mechanism for EtBr exclusion of plasmid during the peptide titration of EtBr bound plasmid and DNase I degradation of peptide protected plasmid based on the available data. The enhanced exclusion of

EtBr by H3K(+G)4b at N:P ratios below 3.6 are most likely due to the flexibility afforded by the glycines in the branched regions of this peptide. This allows for better coating of plasmid DNA around a 3.6 ratio and in turn, more EtBr exclusion. At higher N/P ratios, H3K4b is most likely excluding more EtBr because the larger initial gaps in coating are filled in with more H3K4b and exclude more EtBr at the end.

2.3.4 Delivery Agent Flexibility and its Relation to Nucleic acid Coating

The effectiveness of a delivery agent will be strongly linked to the molecule's structure and the complex formed with the nucleic acid[11; 44]. In turn, what determines the type of interaction formed in a complex between plasmid and a peptide is the charge and flexibility of the peptide[35]. Both H3K4b and H3K(+G)4b have the same positive charge at pH 7.5, but very slight structural differences due to the added glycines in the H3K(+G)4b. Based on this subtle structural difference, Figure 2.5 demonstrates the very different abilities of H3K4b and H3K(+G)4b to coat plasmid DNA. H3K(+G)4b should have more backbone rotation and less steric hindrance in each branch of the peptide due to the added glycines. The thought is, with added flexibility would come better coating and protection of the plasmid, which is demonstrated by DNase I protection. The idea that flexibility of a delivery polymer has an impact on coating and transfection ability is not a new concept. The sensitivity of a DNase I based assay to probe coating of λ-DNA has been observed. One study has tested the effect of taking a series of short positively charged lysine peptides (which strongly associate with DNAs negatively charged backbone) and added a scaffolding polymer in which these short peptides are attached[45]. This effectively

creates a branched polymer delivery agent similar to H3K4b and H3K(+G)4b. The resulting assay, which compares the short lysine peptide with and without the scaffolding structure, clearly demonstrates superior protection of λ-DNA with the scaffolding[45]. This strongly indicates that branching is essential for effective coating and protection of the nucleic acid. However, how does flexibility fit into this assessment?

It has been demonstrated that PAMAM dendrimers have very different transfection of plasmid based on how degraded the polymer is[35]. Creation of a polymer that has fewer branches is accomplished by heat degradation[35]. The reduction in the density of branching caused by degradation yields a greater flexibility of the PAMAM molecule[35]. The flexibility of the degraded PAMAM is what's believed to contribute to the enhanced transfection efficiency of plasmid by the polymer[35]. Another important aspect of this study presents non-degraded PAMAM molecules as being more prone to aggregation due to the very dense branching of the PAMAM[35]. This is most likely due to entanglement and interactions between the dense branch regions that cause aggregation. In contrast to the presumably more flexible H3K(+G)4b, H3K4b also presented a poor ability to coat and protect plasmid DNA from DNase I. It is believed that H3K4b cannot coat DNA as well because not all of the lysines are interacting with it. This is most likely due to the slightly greater rigidity of the branches of H3K4b, which form localized aggregates of peptide and DNA clumps. These differences in structural interactions observed for H3K4b compared to H3K(+G)4b is in good agreement with those observed for difference between degraded compared to intact PAMAM polymers[35].

2.3.5 Considerations of the Formation Mechanism of H3K4b and H3K(+G) Nucleic Acid Particles Guided by the Order of Addition

As demonstrated in Figure 2.2 above, there are some important differences to be noted between titrating peptide to nucleic acid compared to mixing all at once. This comparison is key because in-vivo transfection of these complexes is accomplished by mixing the complex all at once, not added in a sequential manner to reach the final N/P ratio. For the fluorescence experiments it was key to do a titration to make an observation about how the complex formed starting from low N/P ratios going to saturating concentrations. For the H3K(+G)4b peptide results for sequential compared to single-step addition look similar. The one deviation from the sequential EtBr displacement is after adding more peptide starting from a 1.8 N/P ratio there is a plateau in the amount of EtBr displaced. However, H3K4b presents more significant differences between sequential addition and combination of the corresponding N/P ratio. When starting at a 1.8 N/P ratio and adding in more peptide from that point, the fluorescence reached a lower level than the sequential addition indicating a continuing exclusion of EtBr. The fluorescence starting at the 3.6 ratio shows a similar level of fluorescence. Continued addition of peptide from the 3.6 ratio demonstrated that full extent of exclusion might have been reached earlier. These results suggest that H3K4b's less effective coating may be causing variability in EtBr excluded because there is no set pattern for binding to plasmid DNA. Since the order of addition may matter, the important consideration from these experiments is the equilibrium of the formed peptide-nucleic acid complex. Demonstrating that order of addition is key for the formation of these complexes future experiments will have to

be conducted on relating transfection efficiency to the order of delivery agent addition.

2.3.6 Complexities and Complications of AO results

Overall the AO data corroborated the results we observed for EtBr exclusion in the plasmid experiments. H3K4b excludes more AO at saturating concentrations, while H3K(+G)4b reaches its final extent of AO exclusion at lower N/P ratios. However, there are some important considerations to keep in mind when assessing this data. AO fluorescence has proven to be very sensitive to a multitude of factors, which include solvent salt concentration, the amount of AO dye used, the ratio of dye to nucleic acid and small molecule effects[46; 47; 48]. The AO dye has been shown to form a dimeric complex, which can occur at lower concentrations of AO in the presence of salt, affecting fluorescence of the dye at 530 nm[48]. Nucleic acid concentration can affect AO dimer formation and fluorescence. At higher nucleic acid to dye ratio, intercalation of the dye can prevent the dimer from forming, yielding an optimal fluorescence of AO[47]. In experiments conducted with AO bound to plasmid and siRNA, a large increase in fluorescence was expected over that of the AO alone. The sensitivity of AO dimer formation in the presence of nucleic acid could explain why what we actually observed was only a modest increase in the bound AO fluorescence signal. Further optimization of the ratio of nucleic acid to AO proved ineffective at yielding a significant increase in AO-nucleic acid fluorescence over AO alone. This made analysis of the AO results more difficult because during addition of peptide to the AO-nucleic acid, the drop in fluorescence

signal went below the AO alone fluorescence. This indicated an interaction of the peptide, not only with nucleic acid, but also the AO.

Quenching of AO was also a concern because of the possible peptide interaction with AO. As seen in Figure 2.4 in section 2.2.4, both peptides interact with AO in the absence of nucleic acid and cause fluorescence quenching. The H3K(+G)4b peptide shows greater quenching of the AO at lower N/P ratio than H3K4b. Molecules like caffeine and Bovine Serum Albumin (BSA) have also been shown to complex with AO and in the case of caffeine have an effect on AO fluorescence[46; 48]. Above saturating concentrations of peptide for the nucleic acid, there may be an abundance of interaction of the peptide with the AO. H3K(+G)4b may be providing a favorable interaction with AO which promotes an enhancement of fluorescence. Since neither of our peptides proved a conjugated ring system, this possibility might be unlikely. Another potential explanation of enhanced AO fluorescence may come from alterations in the plasmid DNA structure during peptide binding[49; 50]. Delivery agents, such as PEI have been proven to cause structural changes in DNA topology upon binding[50]. These structural changes in DNA upon peptide binding may cause an enhancement of AO fluorescence[49]. Since we only see this phenomena in the H3K(+G)4b binding and not the H3K4b, this could be a powerful tool in assessment of H3K(+G)4bs possible effects on DNA topology.

The above factors were a big concern in assessing this data because not only was exclusion of the dye being observed, but a possible binding and quenching interaction between the peptide and dye. This made the AO experiments more complex to assess than the EtBr exclusion because a drop in fluorescence could now

also be attributed to H3K4b and H3K(+G)4b quenching AO fluorescence. The resulting binding equilibrium between AO, DNA and peptide would have made judging true binding of peptide to nucleic acid in this system very difficult. For any future experiments assessing the binding of a packaging peptide to nucleic acid using exclusion of a dye, careful consideration will have to be given to the type of dye used and possible interactions of the packaging agent with the dye.

Chapter 3: Circular Dichroism and Thermodynamic Assessment of Poly-Histidine Lysine-Nucleic Acid Complex Formation

General Introduction:

The structures of H3K4b-DNA and H3K(+G)4b-DNA complexes were assessed by Circular Dichroism (CD) and Isothermal Titration Calorimetry (ITC) to test whether identifiable biophysical characteristics of the complexes correlate with transfection yields of these two peptides. Many studies have observed CD spectral changes in DNA or siRNA induced by nucleic acid packaging molecules and have reported very little signal for the packaging molecule itself [31; 51; 52; 53]. Observing secondary structure formation in H3K4b or H3K(+G)4b upon complexation with nucleic acid was of great interest to see if there is a recognizable structure for the complex. The thermodynamics of the complexes formed with H3K4b and H3K(+G)4b can elucidate whether the type of bonds being formed are electrostatic, hydrogen bonding or van der waals during condensation of the peptide with nucleic acid. The type of bonds that are driving formation of a delivery complex can give insight into how these complexes are formed. The structural results yielded a surprising PII helical structure for the peptide alone. ITC results demonstrated that complex formation is a two-step process with an initially enthalpically driven phase followed by an entropically driven phase.

3.1 Materials and Methods:

3.1.1 CD Analysis of H3K4b and H3K(+G)4b with and without Plasmid DNA or siRNA

A Jasco-810 Spectropolarimeter was used to conduct all CD experiments. Scans were taken from 190-230 nm at a scan speed of 50 nm/min. A 300 uL Cary

quartz cuvette with a pathlength of 0.2 cm was used. The volume used was 300 μL for DNA and 270 μL for siRNA experiments. A 20 mM phosphate buffer pH 7.5 with 2 mM NaCl was used in all experiments: this was the maximum amount of chloride ion that could be used without unacceptably high absorbance at 190 nm. For pH variation experiments 20 mM phosphate buffers at pH 7.5 and pH 5.78 were prepared. Temperature was set to 4 ^{0}C, 25 ^{0}C and 60 ^{0}C using a Peltier element attached to the instrument. For DNA experiments, 2 μM H3K4b or H3K(+G)4b was initially placed in the cuvette. Successive 1 μL aliquots of from a 16 nM DNA stock were titrated into the cuvette yielding 0.05 nM per addition of DNA. After a total of 0.37 nM of DNA had been added to the peptide, aliquots of 1 μL from an 81 nM peptide stock were titrated into the cuvette, yielding 0.27 nM per addition of DNA. For H3K(+G)4b titrations, two 0.75 nM aliquots of DNA were added at the end of the titration. For the siRNA experiments, 5 μM H3K4b or H3K(+G)4b were initially added to the cuvette. Aliquots of 1 μL from a 12.1 μM siRNA stock were titrated into the cuvette yielding 0.045 μM additions of siRNA. After a total of 0.43 μM of siRNA had been added to the peptide, aliquots of 0.5 μL from a 121 μM siRNA stock were titrated into the cuvette yielding 0.22 μM additions of siRNA. After a total of 1.21 μM siRNA was added, 1 μL aliquots from a 121 μM stock were titrated into the cuvette yielding 0.43 μM additions of siRNA. After each aliquot was added, a CD spectrum was taken and changes in the CD signal were observed. All spectra were buffer corrected by subtracting the 20 mM phosphate buffer with 2 mM NaCl spectrum alone from all spectra.

3.1.2 ITC of H3K4b and H3K(+G)4b Plasmid Binding

Szu-Ting (Justin) Chou generously conducted all ITC experiments and graciously allowed us the use of this data. H3K4b, H3K(+G)4b (25 μM), and DNA plasmid (4.8 kb, 3.75 nM) were dissolved in phosphate buffer (10 mM, pH 7.3). The thermodynamic profile of HK peptides DNA binding was obtained with the use of a VP-ITC (Microcal Inc.) at 37 °C. The sample cell was filled with DNA plasmid (1450 μl), the syringe contained peptide solution (250 μl). Injections (5 μl, a total of 56 injections) were carried out at 10 s duration and 360 s intervals. The dilution heat was determined with the titration of peptides into buffer without DNA. Molar enthalpy and final figures were generated in Origin 5.0. The binding study for each HK polymer was performed in duplicate.

3.2 Results:

3.2.1 CD Signal Characteristics of Peptide during Titration with Nucleic Acid

CD analysis was conducted for both H3K4b and H3K(+G)4b peptides in the presence of siRNA and DNA, as seen in Figure 3.1. Previous experiments conducted by Dr. Mixson's laboratory demonstrated that H3K4b could form alpha helicies in the presence of trifluoroethanol (TFE), which is a known inducer of alpha helix formation. Therefore, our expectation was that some alpha helix might form upon nucleation of the peptide on nucleic acid however, the data show no definitive organization into alpha helices or any other structure with increasing addition of siRNA or DNA. H3K4b and H3K(+G)4b exhibit what most closely resembles a poly-proline II (PII) left handed helix conformation, which is characterized by a

strong negative signal around 197 nm

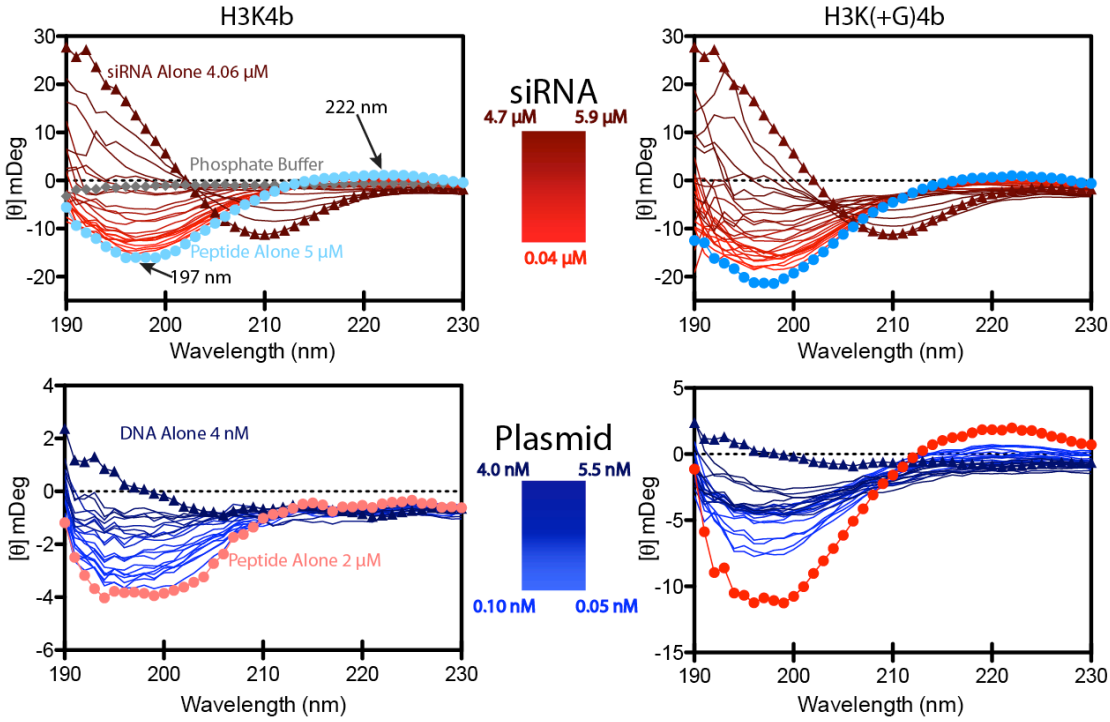


Figure 3.1 ***CD spectra of H3K4b or H3K(+G)4b titrated with siRNA or DNA*****. The buffer used was 20 mM phosphate with 2 mM NaCl at pH 7.5. The top left CD spectrum is a titration of 5 µM H3K4b (powder blue circles) with siRNA ranging from 0.04 µM-4.7 µM (Red Gradient), which corresponds to a N/P ratio ranging from 192-1.8 N/P. The siRNA alone signal at 4.06 µM is presented with dark red triangles. The buffer signal is presented as grey diamonds and is subtracted from all other spectra. (Top right) Titration of 5 µM H3K(+G)4b (blue circles) with siRNA ranging from 0.04 µM-5.9 µM (Red Gradient). (Bottom left) Titration of H3K4b with plasmid DNA ranging from 0.1 nM-4.0 nM, which corresponds to an N/P ratio ranging from 446-6 N/P, (Blue Gradient) into 2 µM H3K4b peptide (salmon circles). The DNA alone signal for 4 nM is presented with dark blue triangles. (Bottom right) Titration of DNA ranging from 0.05 nM-5.5 nM into 2 µM H3K(+G)4b peptide (Red circles).**

and a slight positive maximum around 222 nm. The PII CD spectrum can be easily mistaken for a random coil signal because of the similarity in signal of a negative maximum at 197 nm. However, the CD for a PII helix has a characteristic positive maximum at 222 nm, which decreases upon denaturation and gives a signal greater in magnitude than that of a random coil[54; 55; 56]. There is a change in the signal with increasing amounts of siRNA or DNA. We do not observe any clear structural formation upon complexation of peptide with nucleic acid in any of the CD spectra

presented. There is a clear reduction of the peptide PII structure upon binding of nucleic acid during complex formation. The DNA CD does not contribute a significant signal by itself, but we still see an effect on the peptide upon binding DNA. At the end of the titration, in the case of siRNA, CD signal of the nucleic acid at high concentration is strong enough that it replaces the peptide signal. The peptide alone shows a weaker CD signal in the DNA experiments because a lower concentration was used than in the siRNA experiments (to compensate for the significant differences in binding stoichiometry between siRNA and DNA). There are some differences in the intensity of the CD signal between the two peptides, which is especially true for the 2 μM signal of H3K4b compared to that of H3K(+G)4b. This could be due to difference in structural organization of each peptide. Possible experimental error in the concentrations of each peptide could also be contributing to the differences observed, however checks of the peptide concentration with a UV-visible spectrophotometer proved that the concentrations in the cuvette are what they were believed to be.

Since the only two major CD signals observed are that of the peptide and nucleic acid, we might expect that the a simple addition of the two signals are being observed during the titration. However a Singular Vale Decomposition (SVD) and a 3-state analysis of the data in matlab showed that this is not the case and that there is an independent signal for the peptide-nucleic acid complex.

The CD spectra acquired with titration of siRNA to H3K4b and H3K(+G)4b can be used to assess the different spectral components of peptide-siRNA mixtures. The primary interest in decomposing the CD spectra into different components was to

see if the siRNA-peptide complex formed has an independent CD signal compared to siRNA and peptide alone. Initially, an unbiased singular value decomposition (SVD) was conducted to see if there are multiple spectral components to the CD signals. The SVD was able to pull out at least four independent signals from the siRNA titration of H3K4b (not shown). Two of the spectral contributions closely resemble

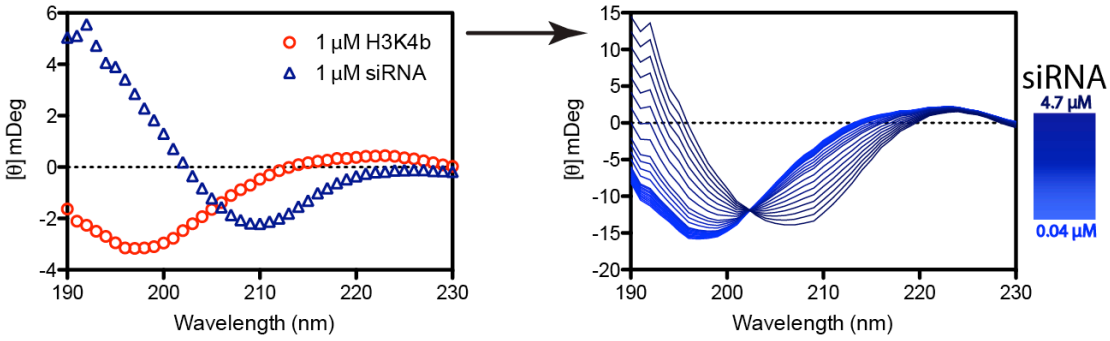


Figure 3.2 ***Added Independent siRNA and H3K4b Signals.*** **The CD spectrum on the left shows the independent signals for siRNA and peptide at 1 µM concentration. When spectra for siRNA are added at their corresponding experimental concentration to 5 µM peptide, a theoretical set of spectra is obtained for the mixture of components (spectrum on right). The additive spectrum does not resemble the experiment spectrum presented in the top left panel of Figure 3.1 suggesting there is at least one missing component.**

the signal for siRNA by itself and the peptide by itself. The other two signals obtained from the SVD were thought to be contributions from a unique signal for the siRNA-H3K4b complex formed. As conformation of the existence of this independent signal, Figure 3.2 presents the spectrum we would expect to obtain if the CD signals for the siRNA and peptide were simply additive. This additive reconstruction of the siRNA titration does not resemble the actual results, meaning there are contributions from at least one other spectrum. A Matlab program was written to extract the unknown spectrum of the siRNA-peptide complex. A general schematic of how the Matlab program was executed is given in Figure 3.3. In the Matlab program a matrix of the known siRNA and peptide spectra and the unknown

theoretical complex spectrum was created. The spectral matrix of references was

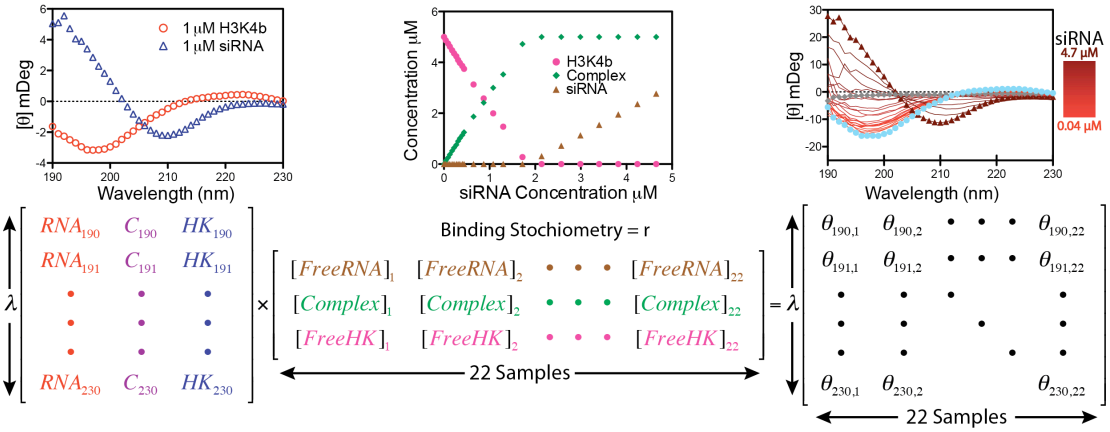


Figure 3.3 ***Components Used to Create the Theoretical Reconstruction of the Experiment.*** **The spectrum on the top left shows the CD for siRNA and peptide (blue) at 1 μM concentration. (Bottom left) A matrix of reference spectra for the peptide, the siRNA and unknown peptide-siRNA (purple) complex was created. (Top Middle) A graph of the concentration of free siRNA (orange), free peptide (Pink) and complex (green) is presented as siRNA is titrated. (Bottom Middle) A matrix of these concentrations was created assuming stoichiometric binding between peptide and siRNA. The reference matrix was multiplied by the concentration matrix, while minimizing the values for the complex spectrum and the binding stoichiometry that optimized the fit of the theoretical reconstructed data to the experiment (Top Right). The minimization yielded a matrix of ellipticities for the theoretical reconstruction (Bottom Right) of the experimental data.**

then multiplied by a matrix containing the concentrations of each component at each point in the titration, assuming an unknown binding stoichiometry, that was also refined. Multiplying these two matrices together generates a theoretical

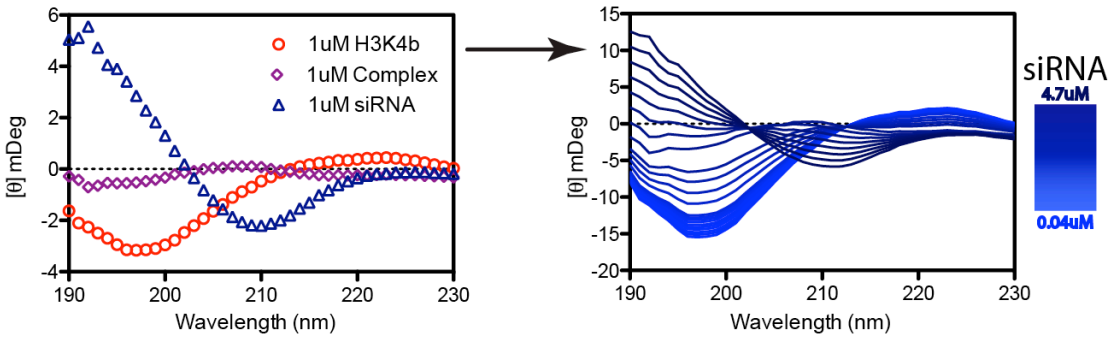


Figure 3.4 ***Theoretical Reconstruction of Experimental Spectra.*** **The spectrum on the left shows the siRNA and peptide spectra along with the spectrum for the siRNA-peptide complex, derived from the minimization described in the text at 1 μM concentration. Multiplying this spectrum by the concentration matrix provided a theoretical reconstruction for the experimental data that's very similar to the experimentally obtained spectra (From Figure 3.1).**

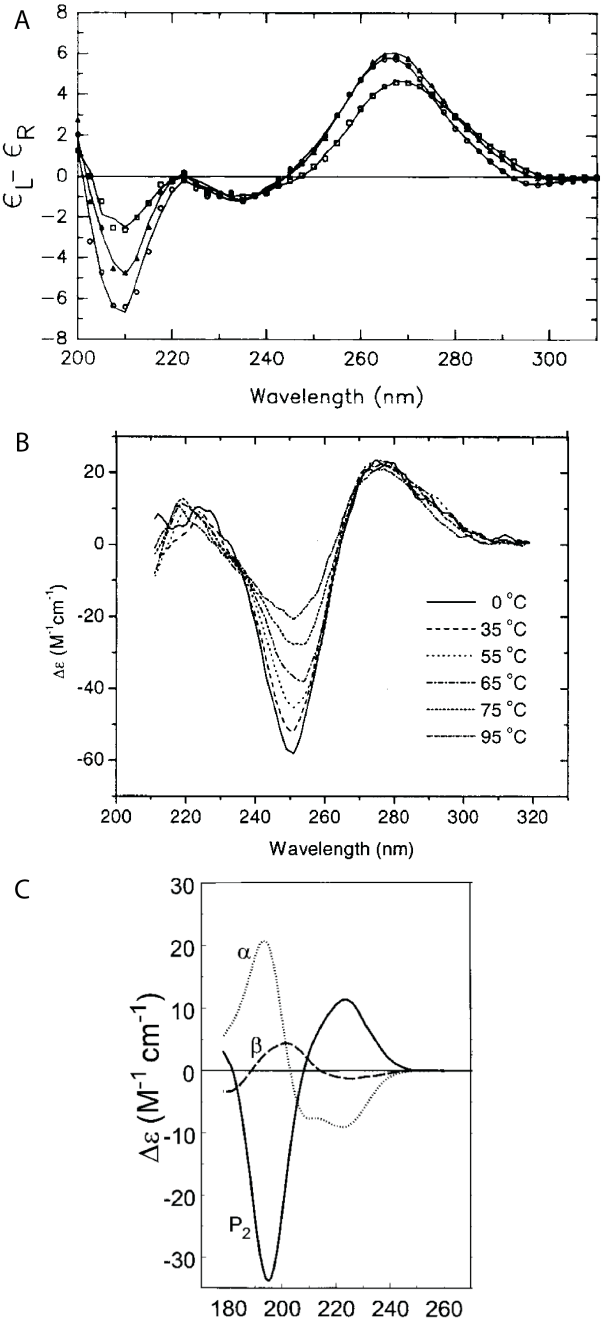


Figure 3.5 ***Reference of Known CD Spectra.*** **CD spectra from three separate papers are presented in A[57], B[58], and C[59]. The spectrum in A shows heat denaturation of an E. coli 5S RNA at 20 ^{0}C, 50 ^{0}C and 65 ^{0}C in circles, triangles and squares, respectively, which is representative of an RNA spectrum. B shows heat denaturation of an $(AT)_3$ duplex from 0 ^{0}C-95 ^{0}C, which is representative of a DNA spectrum. The spectrum in C is a structural deconvolution of alpha helix (α), beta sheet (β) and PII (P_2) helix spectra from reference spectra. The peptide-siRNA complex spectrum obtained does not match any of these reference spectra.**

reconstruction of the experimental data. In order to obtain the H3K4b-siRNA complex spectrum and binding stoichiometry, the difference between the experimental data and the reconstructed theoretical spectrum was minimized using the built in fmincon routine. The sum of squares of the elements of the error matrix was minimized by altering the spectrum for the H3K4b-siRNA complex and binding stoichiometry. Best-fit stoichiometry of binding was 2.9 peptides per siRNA. At charge neutralization, assuming a positive 24 charge for the peptide, we would anticipate 1.75 peptides, per 21 base pair siRNA. We are observing more peptides bound per siRNA than charge neutralization would predict, possibly because of multiple siRNAs bound to a single

peptide. The theoretical spectrum constructed from the siRNA, peptide and complex spectra are presented in Figure 3.4. A series of reference spectra presented from other papers in Figure 3.5 shows CD spectra for thermal denaturation of DNA and RNA and reference spectra for alpha helices, beta sheets and a PII helix[57; 58; 59]. The spectra for the thermal denaturation of DNA and RNA are the same as ideal reference spectra of B-form DNA and siRNA[82; 108]. Our theoretical H3K4b-siRNA complex spectrum has a weak single intensity and does not appear to resemble any of these references. What we are most likely observing is either the disruption of the PII helix structure or possibly some denaturation of the siRNA upon complex formation.

Taking into consideration the siRNA-peptide complex CD spectrum, the re-constructed theoretical spectrum closely resembles that of the experiment. The error between the experimental and theoretical curves is presented in Figure 3.6. The error between the two sets of spectra appears to be relatively random, confirming our theoretical reconstruction. There appears to be a group of slightly positive error curves around 200 nm for the siRNA lower concentrations (light blue lines). For the higher concentrations (dark blue lines) there is a bias signal that is positive at 212 nm and negative at 197 nm. The small bias in the error spectra is believe to come from a missing fourth spectrum for complex that is contributing to the experimental spectra. This fourth spectrum was predicted by the SVD and is believed to be an independent spectrum for further binding of siRNA to an already saturated siRNA-peptide complex. However, an attempt to modify our model to fit for a fourth spectrum did not converge on an answer. Due to errors in concentration and the quality of the data, we believe this spectrum cannot be pulled out from our analysis.

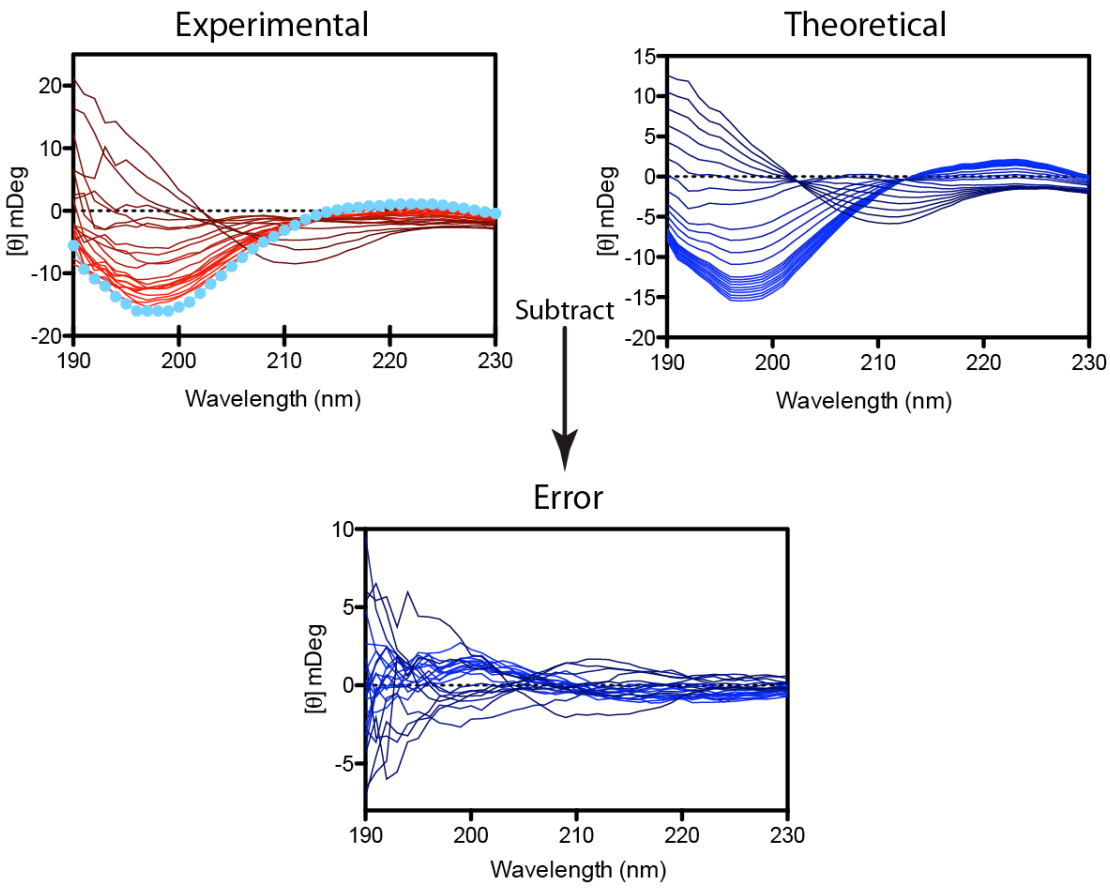


Figure 3.6 ***Error in Experimental versus Theoretical CD Spectra*****. The experimental spectrum (top left) were subtracted from the constructed theoretical titration spectrum (top right). The resulting error spectrum (bottom blue) presents relatively small and random error between the two spectra. There is a slight bias in the error, which may come from a fourth component in the experimental data that was not modeled.**

Both peptides binding to siRNA show very similar characteristics in the peptides' CD signal change. The decrease in peptide signal observed upon addition of siRNA was relatively reproducible and well behaved. Little can be inferred about any difference between structural changes of H3K4b and H3K(+G)4b's with binding of siRNA. This fact alone does not preclude the idea that the two peptides have a different transfection ability based on there structural difference in binding siRNA, simply that we do not see these differences in CD analysis. Local changes in the structural formation of the both complexes observed may be very similar, however,

the differences in transfection observed could be due to large-scale structural differences.

Changes in H3K4b and H3K(+G)4b signal are more inconsistent during DNA binding. The first experiments conducted with addition of DNA to H3K4b, were similar to that of the siRNA. An interesting result is observed during H3K(+G)4b's binding with DNA because of the significant change in CD signal observed with the addition of only a small amount of DNA. This was a very surprising result considering that only a very small fraction of the peptide should be interacting with

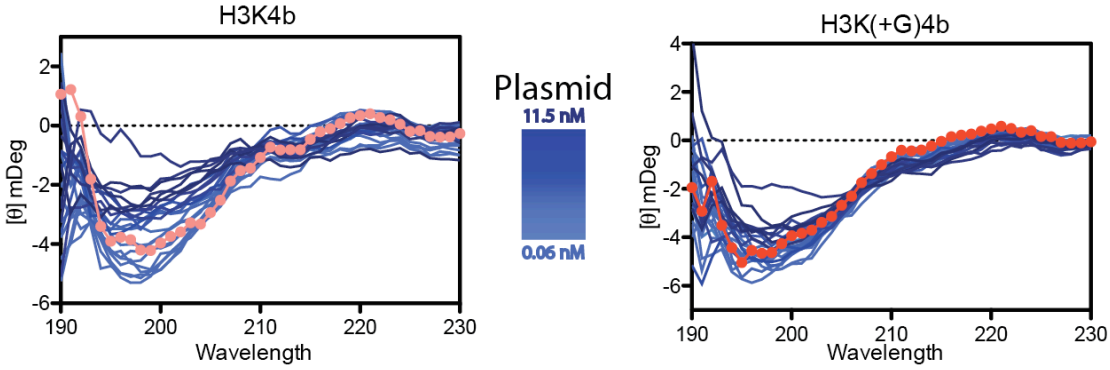


Figure 3.7 ***H3K4b and H3K(+G)4b-DNA Complex Formation Repeated.*** **Titration of H3K4b and H3K(+G)4b with plasmid DNA was repeated to see if the large jump observed in the first DNA titrations could be repeated. Repeated results resembled what is observed for titrations of siRNA to each peptide. There is a slight increase in the peptide CD signal at low concentrations of plasmid DNA with an incremental decrease in peptide CD signal with increasing concentrations. H3K4b and H3K(+G)4b both show similar changes in CD signal with addition of DNA in this experiment.**

the DNA at the 0.05 nM concentration of DNA (about 0.5% of the total peptides binding capacity at 2 μM). When these experiments were repeated (Figure 3.7), the large decrease in CD signal we observed for H3K(+G)4b binding to DNA was not reproducible. A slight increase in PII helix CD signal is observed at low DNA concentrations with a steady decrease in CD signal at higher concentrations. The difference in reproducibility of results between the siRNA and DNA CD experiments suggesting that there might be complicated kinetics for peptide-DNA complex

formation. There is a possibility that plasmid complex formation with each peptide could involve some cooperativity at lower peptide concentrations.

3.2.2 Isothermal Titration Calorimetry (ITC) Results

A large thanks goes to my collaborator, Szu-Ting (Justin) Chou, who generously conducted all ITC (Isothermal Titration Calorimetry) experiments and graciously allowed us the use of the preliminary data presented here with his permission. His contributions to the analysis of this data proved invaluable in our understanding of the mechanics of the peptides' interaction with plasmid DNA and siRNA.

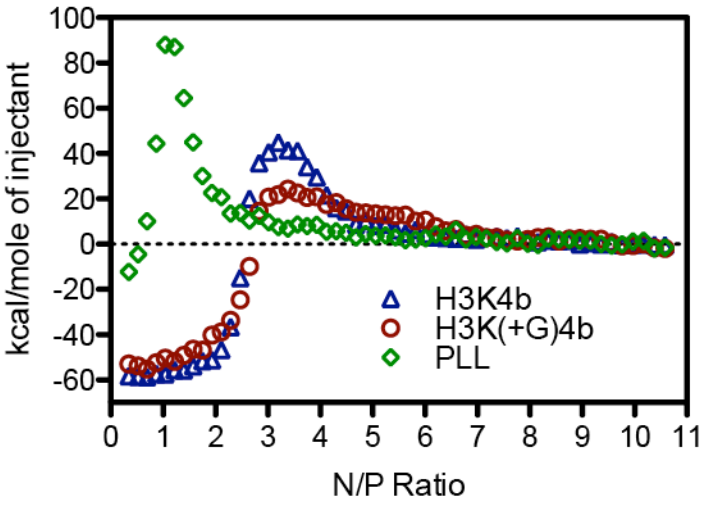


Figure 3.8 ***ITC of H3K4b H3K(+G)4b-DNA and PLL Complex Formation*****. The enthalpy change of complex formation upon addition of H3K4b, H3K(+G)4b or PLL to DNA was measured. PLL (green) complex formation with plasmid DNA demonstrates a very large enthalpy of formation compared to H3K4b or H3K(+G)4b. Both H3K4b (blue) and H3K(+G)4b peptides interactions prove to be exothermic at low N/P ratios, which are characterized by stronger hydrogen bond formation between the peptide and DNA. Upon addition of more peptide, the formation of complex with DNA changes over to being endothermic indicating a different thermodynamic driving force. In this region of the reaction, electrostatic interactions and release of water and cations are driving the reaction. H3K4b has a larger endothermic enthalpy of formation than H3K(+G)4b.**

ITC was used to assess the thermodynamic contributions of PLL, H3K4b and H3K(+G)4b with plasmid DNA. Difference in the binding interaction between H3K4b and H3K(+G)4b to plasmid DNA can be elucidating by observing the complexes enthalpy of formation (Figure 3.8). Each peptide was added to a fixed concentration of plasmid DNA to yield N/P ratios

ranging from 0 to 10. The heat of formation for the peptide-plasmid complex was measured in kcal per mole of injected peptide. The linear PLL peptide was used as a comparison to binding of each of the branched peptides. PLL binding to plasmid demonstrates a large positive enthalpy, with very little exothermic contribution to complex formation. Upon addition of low concentrations (0.5 N/P), binding of both H3K4b and H3K(+G)4b is exothermic. Continued addition, above a 2 N/P ratio of peptide to plasmid switches over to an endothermic reaction. At 1:1 charge neutralization, which is an N/P ratio of 3, the enthalpy of formation reaches a maximum. At the 3 N/P ratio we observe a significant difference between peptides, where H3K(+G)4b has a lower heat of formation, at 26 kcal/mol compared to that of H3K4b at 48 kcal/mol. Past a 3 N/P ratio the formation of the peptide-plasmid complex has completed and drops to 0 kcal/mol by about N/P equal to 6.

3.3 Discussion:

3.3.1 Random Coil and PII Helix Conformation in CD

There is a large body of evidence that suggests what was previously thought to be "random-coil" structure is in fact not-so-random[54]. Studies have been conducted on short poly-L-glutamic acid (PGA) and poly-L-lysine (PLL) peptides exploring this idea[56; 62]. It is known that a polyproline peptide (PPL) assumes a left-handed helix termed a poly-proline II (PII) conformation[55; 60]. The structure of a PII helix as seen in the poly-proline crystal structure and a stretch of PII helix found in a Lignin Peroxidase are shown in Figure 3.9[61]. The CD of the PII conformation tends to resemble that of random coil with a negative signal peak around 200 nm[56],

however there are two prominent differences observed from that of a random coil (Figure 3.10). One is the fact that the negative peak is very strong for the PII

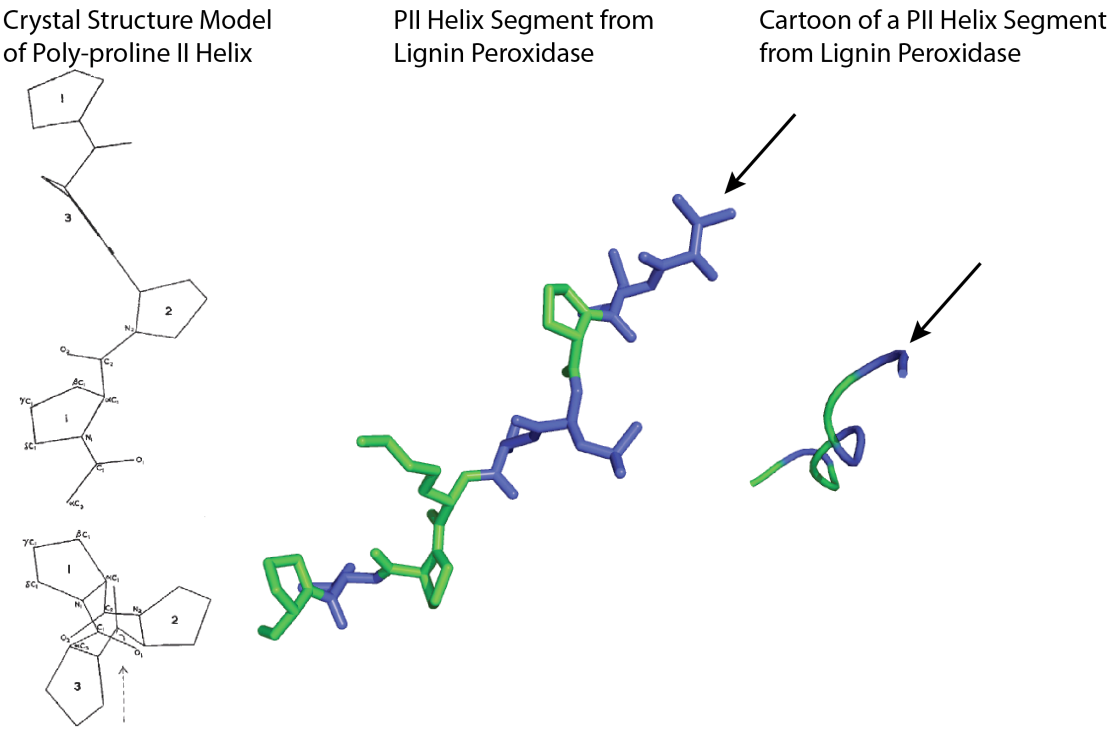


Figure 3.9 ***Views of the PII Helix Structure.*** **Three views of the left handed PII helix structure are shown above. (From left to right) A model of the PII helix determined from the crystal structure which shows three residues per helical turn[60]. (Middle) A stretch of nine residues that forms a PII helical segment in Lignin Peroxidase are shown[61]. The three proline and one lysine residue that strongly correlate to PII helix formation are highlighted in green. The arrow is pointing down the central axis of the PII structure to show the counterclockwise corkscrew of the helix. A cartoon of the same PII helix in Lignin Peroxidase is presented.**

conformation. The other is a slightly positive peak in the 220 nm region. The helical chain structure of the PII conformation has been confirmed by x-ray crystallographic analysis, confirming that the CD is that of a definitively structured peptide and not that of a randomly ordered, denatured structure[60]. Collagen assumes a very similar helical structure to that of PPL[54; 63] and the CD for collagen also shows a characteristic strong negative signal in the 200 nm range and a positive near 220 nm. When collagen is heat denatured, there is a significant drop in the CD signal for the helix conformation. The signal for the heat denatured collagen is that of what a true

disordered structure looks like: it has a much weaker CD signal overall and has only one negative peak around 200 nm[54; 56]. Polyglutamic acid (PGA) and PLL both show a very similar signal to that of PPL[56; 62]. This is a significant observation because PGA and PLL structures have previously been put into the category of random coil, which may have been the predicted result,

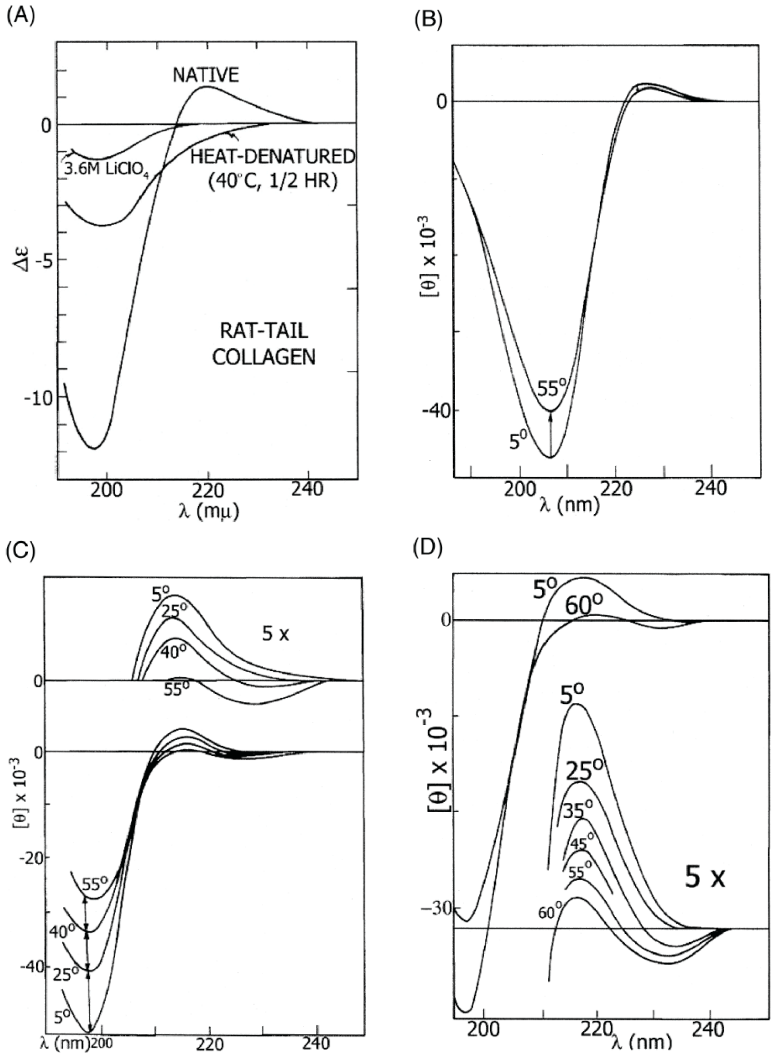


Figure 3.10 ***CD Reference Spectra for a PII Helix Conformation.*** **Reference CD spectra of four different PII helix are used from a review article by Shi et al[54]. Spectra of PII helix structures in their native and denatured state are presented from four different peptides; A, rat tail collagen; B, poly-L-proline II; C, polyglutamic acid; and D, poly-L-lysine. All spectra show a decrease in the amplitude of PII helix spectra upon thermal denaturation.**

because, any significant structural ordering would not have been predicted for a small peptide. Upon heat denaturation at 80 ^{o}C, there is a significant disruption of PGA and PLL's PII conformation to form a disordered structure as seen in collagen[56; 64]. With the evidence presented, we suggest that many short peptides that were thought to have assumed a random coil structure, are in fact a PII left handed helical conformation.

3.3.2 Evidence for a PII Helical Conformation of H3K4b and H3K(+G)4b

The CD data for H3K4b and H3K(+G)4b show the strong negative (197 nm) and a weak positive signal (222 nm) characteristic of the PII helix structure of PLL and PPL. The presence of a definitive PII helical structure would account for why we have such a strong signal for peptide alone, instead of the relatively weak signal to be expected if this peptide were in a truly disordered state[56]. Experimentation at low pH could reveal more about the structure of these peptides because it has been observed that PII helical peptides in the presence of low pH can result in a formation of α-helix structure in these peptides[54]. When the pH was decreased from 7.5 to 5.78, an increase in our peptide signal was observed (Figure 3.11) but we did not observe any α-helix formation. This result indicates the PII conformation of H3K4b and H3K(+G)4b is enhanced when a solution becomes more acidic. A conclusive result contributing to this belief deals with the

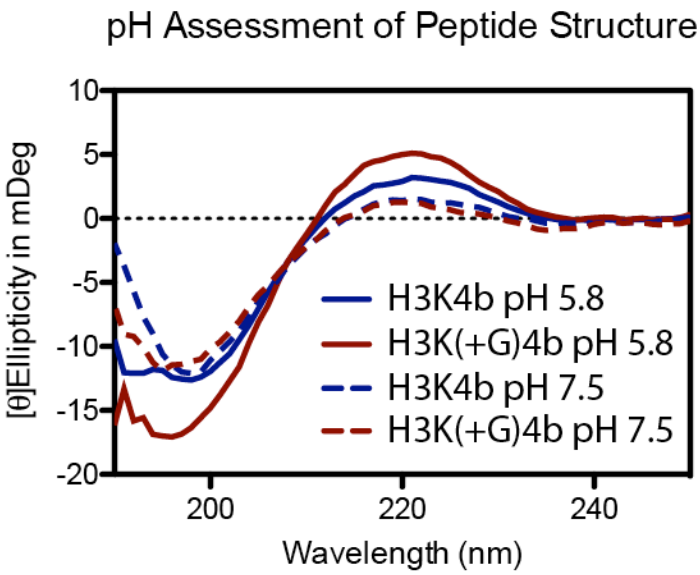


Figure 3.11 ***pH Dependence of PII Helix Structure.*** **A CD spectrum of H3K4b (blue) and H3K(+G)4b (red) in 20 mM phosphate buffer at pH 7.5 (dashed lines) and pH 5.8 (solid lines), 2 mM NaCl is shown. The characteristic PII structure with extremes at 197 nm and 222 nm is seen for the peptide as previously demonstrated. At pH 5.8 there is to be an enhancement of the signal at 197 nm and 222 nm. This could be an indication of stronger PII helical formation when the pH is decreased.**

PLL heat denaturation result presented by Tiffany and Krimm where PII helix CD signal decreases upon heat denaturation of the peptide and increases at lower

temperature. We conducted an experiment where we took each peptide and bound a very small amount of DNA, which still presented a very strong characteristic PII CD signal for the peptide (Figure 3.12). The DNA was originally added to see if heat denaturation had any effect on the peptide-DNA complex, however what we ended up observing were changes in the peptide CD. CD for the peptide was then collected at

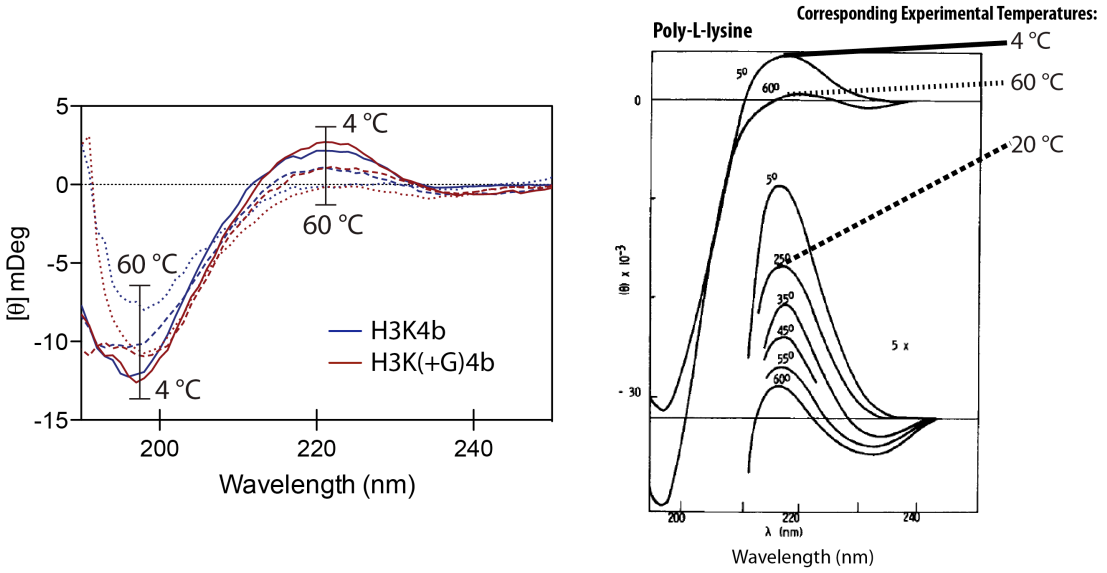


Figure 3.12 ***Temperature Dependence of PII Helix Content of H3K4b and H3K(+G)4b.*** **CD spectra of H3K4b, H3K(+G)4b in the presence of 2.6nM DNA (A) and poly-L-lysine (B) during heat denaturation of the peptides from 4°C-60°C. Both peptides change in structure are compared to a references spectrum of poly-L-lysines change in structure as a function of temperature. The Poly-L-lysines spectrum was used from a paper by Tiffany and Krimm[56]. A. A small amount of DNA was added to both peptides in order to assess whether or not there was a change in the peptide signal while in complex with DNA. Both peptides are only at 8% of their total saturation and still have a very clear PII structure signal from the peptide alone. The two peptides demonstrate an enhancement of PII helix structure when lowered to a temperature of 4°C. Upon heating to 60°C, there is a significant decrease in the PII CD signal seen for the peptides. H3K4b appears to be more strongly affected by thermal denaturation than H3K(+G)4b, highlighting differences in their thermal stability. B. Poly-L-lysine heat denaturation. The change in poly-L-lysine CD signal as a function of temperature shows the exact same trend as for H3K4b and H3K(+G)4b, indicating both of these peptides melt.**

20 °C, then the temperature was dropped to 4 °C and increased to 60 °C and changes in the CD signal were observed, similar to Tiffany and Krimm. The result shown in Figure 3.10 is almost identical to that of Krimm et. all showing a strong increase in PII conformation at 4 °C and a sharp drop at 60 °C. The change in signal is not as

apparent at 197 nm until the temperature is increased to 60 ^{o}C. However, the changes at 222 nm are very clear and distinct, which is a defining characteristic of the PII conformation since random coil has no positive peak in the 220 nm region[54; 65]. All of the evidence provides a strong case for H3K4b and H3K(+G)4b assuming a single or possibly multiple chain (in the case of collagen) PII conformation.

3.3.3 Structural Characteristics of H3K4b and H3K(+G)4b that Allow PII Formation

The major driving force that yields a PII conformation is believed to stem from the peptide backbone's interaction with the solvent, in this case, water[55]. It has been shown that peptides favor a PII conformation near neutral pH and low salt concentrations, which are similar to the conditions we have used to assess H3K4b and H3K(+G)4b[55]. The increased accessibility of solvent to the peptide backbone of small peptides (such as glycine) and peptides with a flexible R group, enables this conformation to occur[55]. Both peptides we have studied, H3K4b and H3K(+G)4b, are small branched peptides that contain histidine but also are comprised of about 32% lysine as well. PLL readily forms a PII helix which has been observed many times, leading to the idea that lysine could be an important amino acid in forming this structure[54; 55]. This may suggest an interaction of the lysines in each of these branched peptides with the solvent. In the case of H3K(+G)4b, the added glycines could be contributing more to solvent interactions. In addition, side chain to backbone interactions could also play a role in forming this helix[55]. A PII helix has Ramachandran angles (ϕ,ψ = -75^{o},145^{o}) consistent with having three amino acid residues per turn[55]. As illustrated previously in Figure 1.4, both of the poly-histidine

lysine branched peptides contain a lysine every four residues on each branch, which may indicate an interaction of the lysine with the backbone every four amino acids.

In light of this comparative analysis, what is most likely happening with H3K4b and H3K(+G)4b peptides upon binding of nucleic acid is as follows: The peptide alone assumes a PII helix conformation. Upon addition of increasing concentrations of siRNA or DNA, there is disruption of this helix structure upon association with nucleic acid. Eventually, enough addition of nucleic acid completely disrupts this structure, and all peptide eventually is bound to siRNA or DNA. DNA's structure changes slightly upon addition of peptide and there is an independent CD signal for the peptide-nucleic acid bound complex. The glycines in the H3K(+G)4b peptide may be contributing to slightly more PII structure formation because of the glycines greater rotational ability, allowing for greater solvent accessibility in this peptide[55].

3.3.4 Thermodynamic Contributions to H3K4b and H3K(+G)4b-Plasmid Complex Formation

Previous results with EtBr exclusion and DNase I degradation assays indicate H3K4b clumps up more on plasmid DNA forming regions of positive overcharge and naked regions of uncoated plasmid. H3K(+G)4b coats plasmid better, forming more evenly distributed, neutrally charged complex. DLS also demonstrates that a complex formed with H3K(+G)4b is larger, perhaps due to better charge neutralization. The ITC data gathered gives some support to the idea of better coating for the H3K(+G)4b peptide compared to the aggregate structure of H3K4b. There are two clear phases during complex formation of both peptides with plasmid DNA as

demonstrated by the ITC. The first phase is presumably initial binding of the peptide to excess DNA and the second is condensation accompanied by aggregation. A two-phase condensation of plasmid by electrostatic interaction with multivalent ions has been demonstrated previously[66]. At low N/P ratios of peptide, an exothermic reaction occurs where heat is given off during complex formation. Strongly exothermic reactions during peptide association with nucleic acid typically correlate to hydrogen bond formation[67; 68]. Another possibility is a hydrophobic interaction between the peptide and DNA. The hydrogen bonding interaction for these peptides are most likely coming from the histidines of the peptide, which has been demonstrated by NMR studies of the peptide-nucleic acid complex (Mixson lab, data unpublished). ITC experiments with PLL binding to DNA do not demonstrate the strong exothermic reaction observed for the branched peptide and primarily show an entropically driven reaction, which occurs from electrostatic binding. The electrostatic interactions from the lysines are also contributing to the exothermic reaction observed for the branched peptide but apparently the total enthalpy for the reaction is negative.

The condensation phase of complex formation is accompanied by a large positive enthalpy change starting around an N/P ratio of 3. The endothermicity of this reaction means that it must be entropically driven[66]. This result came as a surprise because the association of the peptide with the DNA might have been thought to be only an exothermic reaction. A couple of factors that drive this unique complex formation were made clear by this result. The major contribution to unfavorable enthalpy is probably the breaking of hydration bonds between the DNA and water[68; 69]. Another, more speculative influence, is the unfavorable bending and

proximity of the DNA during condensation by the peptide[70]. The entropic driving force for this phase of the reaction is due to counterion release from the DNA upon electrostatic association of the ligand, which plays an important role in peptide and protein binding to nucleic acids[66]. It is also important to mention the release of water from interaction with macromolecules can also play an important role in contributing to entropy[66]. However, the release of counterions or water would be indistinguishable by ITC[66]. This type of condensation through ion release has been experimentally shown by Matulis[66]. In their experiment, multivalent cations used to associate with and condense DNA demonstrated a strongly endothermic, entropically driven reaction. Initial electrostatic interaction of multivalent ions with DNA releases counterions. Upon addition of enough multivalent ions to approach complete charge neutralization of the DNA there is further release of counterions from the DNA upon condensation of the structure. In order for complete condensation of DNA to occur about 90% of the charge on the backbone must be neutralized[70]. Interestingly, we observe a steep rise in endothermicity indicative of condensation between a 3 and 3.5 N/P ratio. Unlike this experiment, other ITC assays where drug binding to a protein site was assessed demonstrated an exothermic reaction that is characterized by hydrogen bond formation[67]. However, the heat of formation for drug binding then plateaus once binding is complete, which we do not observe for binding of H3K4b and H3K(+G)4b to plasmid[67]. The interaction of these peptides with plasmid resembles that of a zipper; it takes enough favorable interactions to get started, but once a critical amount of interactions occur (probably close to charge neutralization), it condenses together. Adding a little bit of peptide (0-0.5 N/P) is exothermic and

starts condensation of the plasmid complex. Upon reaching 90% charge neutralization, there is complete condensation of the complex. Once the complex condenses and is close to neutral charge at a 3 N/P ratio, subsequent aggregation can then occur. The thermodynamic mechanism proposed for H3K4b and H3K(+G)4b complex formation is a prime example of contributing energetic factors that appear relatively simple, but where the underlying chemical contributions are a combination of different factors.

There is one clear difference in the thermodynamics of complex formation for these two peptides. During the condensation phase, H3K4b has a about two times the enthalpy of formation compared to H3K(+G)4b. H3K4b does not coat plasmid DNA well as indicated by the DNase I assay. This leaves regions of uncoated, charged plasmid and overcharged areas of aggregated peptide, which creates a greater opportunity for further electrostatic aggregation demonstrated by H3K4b's higher enthalpy of formation in the ITC. H3K(+G)4b coats the plasmid DNA more evenly, not leaving regions of aggregated peptide or uncoated DNA. This means once the plasmid is coated, there will be little to no electrostatic interactions between the peptide and plasmid because it will be charge neutralized. Charge neutral plasmid complexes formed with H3K(+G)4b will undergo aggregation, however there is no further electrostatic interaction, yielding a less positive enthalpy of formation compared to H3K4b. These data corroborate the idea that H3K4b forms more aggregate structure with more electrostatic interactions between uncoated regions of plasmid. H3K(+G)4b coats the plasmid DNA more effectively forming less electrostatic interactions to accomplish complete coating.

Chapter 4: Morphology of Poly-Histidine Lysine-Peptide Nucleic Acid Complexes

General Introduction:

Understanding morphological characteristics of the peptide-nucleic acid complex may be revealing as to compactness of packaging between individual molecules. The objective was to assess differences in the size and morphology of the complexes formed with H3K4b compared to H3K(+G)4b. Comparison of complex formation between the two peptides using Dynamic Light Scattering (DLS) and Atomic Force Microscopy (AFM) revealed larger sizes for H3K(+G)4b-plasmid and H3K4b-siRNA. DLS analysis demonstrated a change in the size of complex formed with H3K(+G)4b-plasmid and H3K4b-siRNA while only small changes in size were observed for H3K4b-plasmid and H3K(+G)4b-siRNA complexes. AFM analysis presented a greater height to diameter ratio for H3K(+G)4b-plasmid and H3K4b-siRNA, demonstrating a more rigid complex.

4.1 Materials and Methods:

4.1.1 DLS Measurement of H3K4b and H3K(+G)4b-DNA and siRNA Complexes

Experiments were conducted on a Zetasizer Nano ZS90 under the supervision of Dr. Robert Bonenberger who runs the Modern Engineering Materials Instructional Laboratory in the Kim Engineering building. The Zetasizer Nano ZS90 is fitted with a 633 nm laser and scans are taken at 90^{o} to the sample chamber. A siliconized 700 µL Helma fluorescence cuvette was used for these experiments. A 250 µL volume of 1x TEN buffer (10 mM Tris, 1 mM EDTA, 100 mM NaCl pH 7.5) was used in both

the siRNA and DNA experiments containing. This volume was chosen to meet the sample height criteria of the DLS (between 10-15 mm). For the siRNA experiments 1430 nM H3K4b or H3K(+G)4b and 286 nM siRNA (which is 102,960 nM Nitrogens and 12,012 nM Phosphates) were added to the cuvette to form complex. The DNA experiments used 750 nM H3K4b or H3K(+G)4b in combination with 5 nM DNA (Which is 54,000 nM Nitrogens and 30,000 nM Phosphates). The concentrations corresponded to a 8.6 N/P ratio for the siRNA experiments and a 1.8 N/P ratio for plasmid experiments. The N/P ratios for the siRNA experiments were well above 1:1 charge neutralization of the siRNA. N/P ratios for the DNA experiments were over half of 1:1 charge neutralization of the plasmid DNA. The nucleic acid concentrations were chosen as the appropriate concentrations needed to see the complex in the DLS. To see a particle size of 100 nm in diameter, a concentration of 0.01 g/L is required. All components were added to the cuvette and pipetted up and down and finger tapped to mix. For assessment of the effect of salt concentration on complex formation, the 1x TEN, a 0.5x TEN (10 mM Tris, 1 mM EDTA, 50 mM NaCl pH 7.5) and autoclaved filtered water were compared. The samples were allowed to incubate for 30 minutes. Sample readings were taken over the course of two hours due to the time required to take readings of each sample. Over the 2 hour time period, measurements of the effect of salt concentrations on particle size were conducted. DLS measurements of peptide-nucleic acid complex freshly prepared in water were acquired right after sample mixing and were compared to samples prepared in water and incubated for >2 hours. A time course of the experiment was conducted from 0-22 min. Zero minutes corresponded to taking a reading of the

sample immediately after peptide and nucleic acid were mixed. Subsequent readings were taken 10 and 22 min after mixing. The low volume glass cuvette option was used in the instrumentation software. In the software the dispersant selected was water. The material being analyzed was chosen as protein. Each size distribution is a scan of 20 readings taken for 15 seconds each combined by the instrument software. The DLS diameters were reported by the instrument as a log scale and then plotted on a log scale in prism. This made the sizes reported evenly spaced as a histogram distribution when plotted in prism.

4.1.2 AFM Binding to DNA and siRNA

Jason Hustedt generously conducted all AFM visualization experiments and graciously allowed us the use of this preliminary data.

Equal volumes of H3K4b or H3K(+G)4b and siRNA were combined to form complexes for visualization. The same was done for H3K4b or H3K(+G)4b and DNA complex formation. The final peptide and DNA concentration were 120 μM and 65.8 μM, respectively. Peptide and siRNA concentrations were 58.7 μM and 20 μM, respectively. After combining peptide with either siRNA or DNA, the complex was allowed to form and incubate for 30 min. After incubation, the peptide-nucleic acid complexes formed were applied to a freshly cleaved mica surface. The samples were imaged in air with no treatment of the surface. Imaging was conducted with PointProbe Plus Force Modulation Probes using an Agilent 5500 AFM with a 5600LS stage.

4.2 Results:

4.2.1 DLS of Formed Complex

DLS measures the hydrodynamic diameter of a particle, which is a measure of how a particle diffuses in a fluid. The diffusion of a particle is measured by illuminating the sample with a laser and observing the Brownian motion of the particles. Larger particles will diffuse slower than smaller particles. The duration of diffusion in the laser is the basis for the time correlation function, which then used to measures the size of the particle. The DLS analysis was conducted to observe any differences in size of complexes formed with H3K4b and H3K(+G)4b. In addition, large-scale aggregation and precipitation was a concern with the complexes we were forming and the DLS is sensitive to large particles. Both H3K4b and H3K(+G)4b peptides were combined with what was originally thought to be an equal charge ratio of siRNA and DNA to yield a charge to charge ratio of 1:1 and sizes of the complex were measured. Experiments with DNA yielded an N/P ratio of 1.8, which is below charge neutralization of the plasmid. For siRNA experiments, an N/P of 8.6 was used which is well above saturation and charge neutralization of the siRNA. Each peptide shows a different size from each other with binding to DNA or siRNA over a 22 minute time period (Figure 4.1, Table 4.1). Figure 4.1 shows the DLS intensity distribution for the size of particles formed with H3K4b or H3K(+G)4b in complex with DNA or siRNA and the values for the peak intensities of each size distribution are presented in Table 4.1. The Polydispersity (PDI) is a measure of how uniform the distribution of particle sizes are. The lower the PDI, the more uniform the size of a

particle is. The PDI for peptide-DNA complexes were lower on average, around 0.26

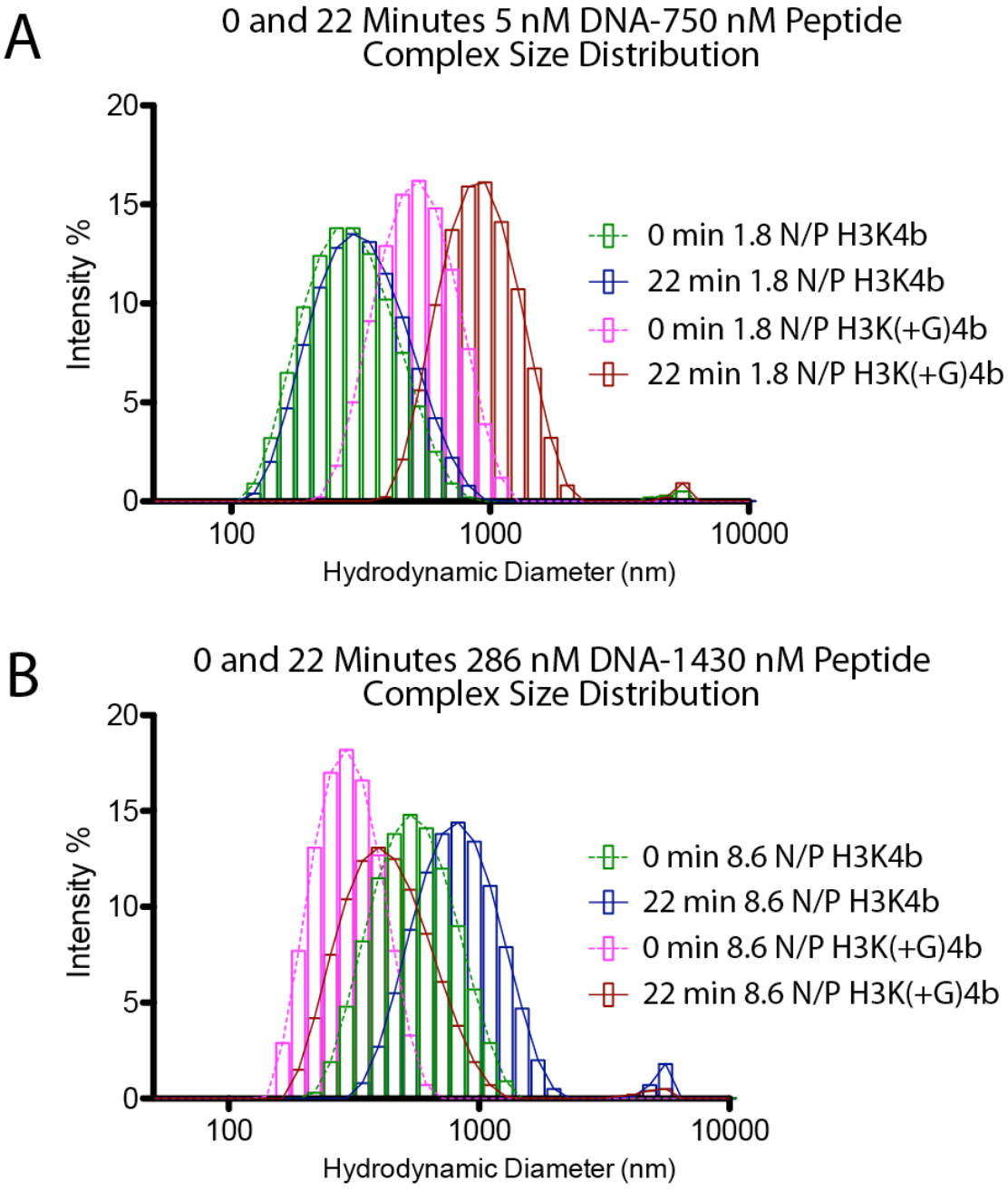


Figure 4.1 ***DLS Analysis of Peptide-Nucleic Acid Complexes Size over Time.*** **H3K4b and H3K(+G)4b-DNA and siRNA complexes formation in 1x TEN for 0 and 22 minutes. The percent intensity on the y-axis is a measure of the intensity of laser light scattering for a given size particle. A. Complex formation with 5 nm DNA and 750 nM H3K4b or H3K(+G)4b. Change in size of the complex is seen by a shift of the maximum peak of the distribution, primarily for H3K(+G)4b. B. Complex formation with 286 nm siRNA and 1430 nM H3K4b and H3K(+G)4b. For siRNA, change in the size of the complex is seen by a shift in the maxima size distribution peak for both peptides. A small peak at 3.6 nm for 22 min H3K(+G)-siRNA complex formation has been omitted.**

	0 min. Particle Size at Peak Intensity	22 min. Particle Size at Peak Intensity
H3K4b-DNA	255-295 nm	295 nm
H3K(+G)4b-DNA	531 nm	955 nm
H3K4b-siRNA	531 nm	825 nm
H3K(+G)4b-siRNA	295 nm	396 nm

Table 4.1 ***Sizes of Peptide-siRNA and Peptide-DNA Complexes.*** **The size associated with the maximum percentage of intensity for each complex is reported in this table. The size of complexes are presented for siRNA and DNA complexes at 0 min and 22 min. Change in size of the H3K(+G)4b-DNA complex is seen by an increase in the maximum peak of the size distribution. For siRNA, change in size of the complex is seen by a shift of the maxima size distribution peak for both peptides.**

for both H3K(+G)4b-DNA and H3K4b-DNA compared to siRNA complexes. DLS measurements of Plasmid DNA, siRNA, H3K4b and H3K(+G)4b alone were taken to ensure that the sizes observed for the complexes were not the same as the individual components (Figure 4.2, Table 4.2). H3K4b-DNA particles stayed smaller in size over time with a peak diameter of 295 nm. In contrast, the H3K(+G)4b-DNA particles appear to be

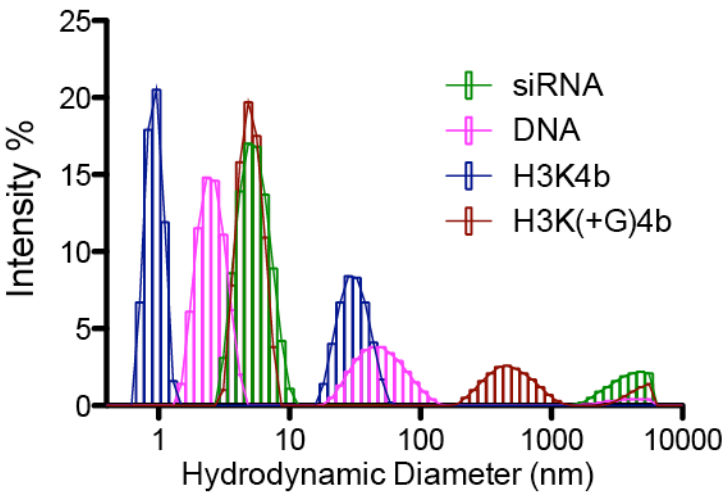


Figure 4.2 ***DLS Analysis of siRNA, DNA and Peptide Sizes.*** **Size distributions for siRNA, DNA, H3K4b and H3K(+G)4b-DNA in 1x TEN are presented. Both peptides proved to be small in size with some aggregates up to a couple hundred nm in diameter. siRNA is also small in size, measuring 5.6 nm in diameter. Plasmid DNA has a peak at 2.3 nm and 51 nm. The peak at 51 nm is most likely for the supercoiled DNA.**

much larger, increasing from 531 to 955 nm in diameter over 22 minutes. This suggests that there are differences in the overall binding and particle aggregation

occurring for the two peptides with DNA. The H3K(+G)4b-DNA complexes appear to aggregate over time, but the H3K4b-DNA complexes do not show aggregation.

	Particle Size at Peak Intensity in 1x TEN
Plasmid DNA	51 nm, 2.3 nm
siRNA	5.6 nm
H3K4b	0.96 nm
H3K(+G)4b	4.8 nm

Table 4.2 ***Control for the Size Measurements of Each Component.*** **The size of plasmid DNA, siRNA, H3K4b and H3K(+G)4b are presented above. All molecules by themselves are significantly smaller in size than any of the complexes formed between peptide and nucleic acid.**

Peptide-siRNA complex formation showed a much higher PDI of 0.3 to 0.5 for H3K4b and H3K(+G)4b, respectively, indicating that there is a very large distribution in particle size for peptide-siRNA particles. With H3K4b-siRNA complexes, there is a change in size from 531 nm to 825 nm in diameter over 22 minutes. For H3K(+G)4b-siRNA complex, there is an increase in the size of complex ranging from 295 to 396 nm. In the case of peptide binding to siRNA, the H3K4b-siRNA complex has an increase in particle size due to aggregation of the complex while the H3K(+G)4b-siRNA size stays relatively constant over time. The size results seem to be flipped between the two peptides depending on whether or not it binds siRNA or DNA. The above-mentioned diameters for the peptide-nucleic acid complexes were obtained in 100 mM NaCl. It is clear from these observations that time and possibly salt concentration had an unanticipated impact on peptide-nucleic acid particle size.

4.2.2 Salt and Time Dependence on Peptide-Nucleic Acid Particle Size

It has been shown that salt concentration can have a large influence on the size of the peptide-nucleic acid complex, most likely due to charge neutralization between particles formed[71]. To probe the effect of NaCl on complex aggregation, the same preparations of peptide and DNA or siRNA were made in a buffer containing 100

mM NaCl as used previously and also in water with no NaCl. It was thought that decreasing the salt concentration would have an effect on the aggregation of the

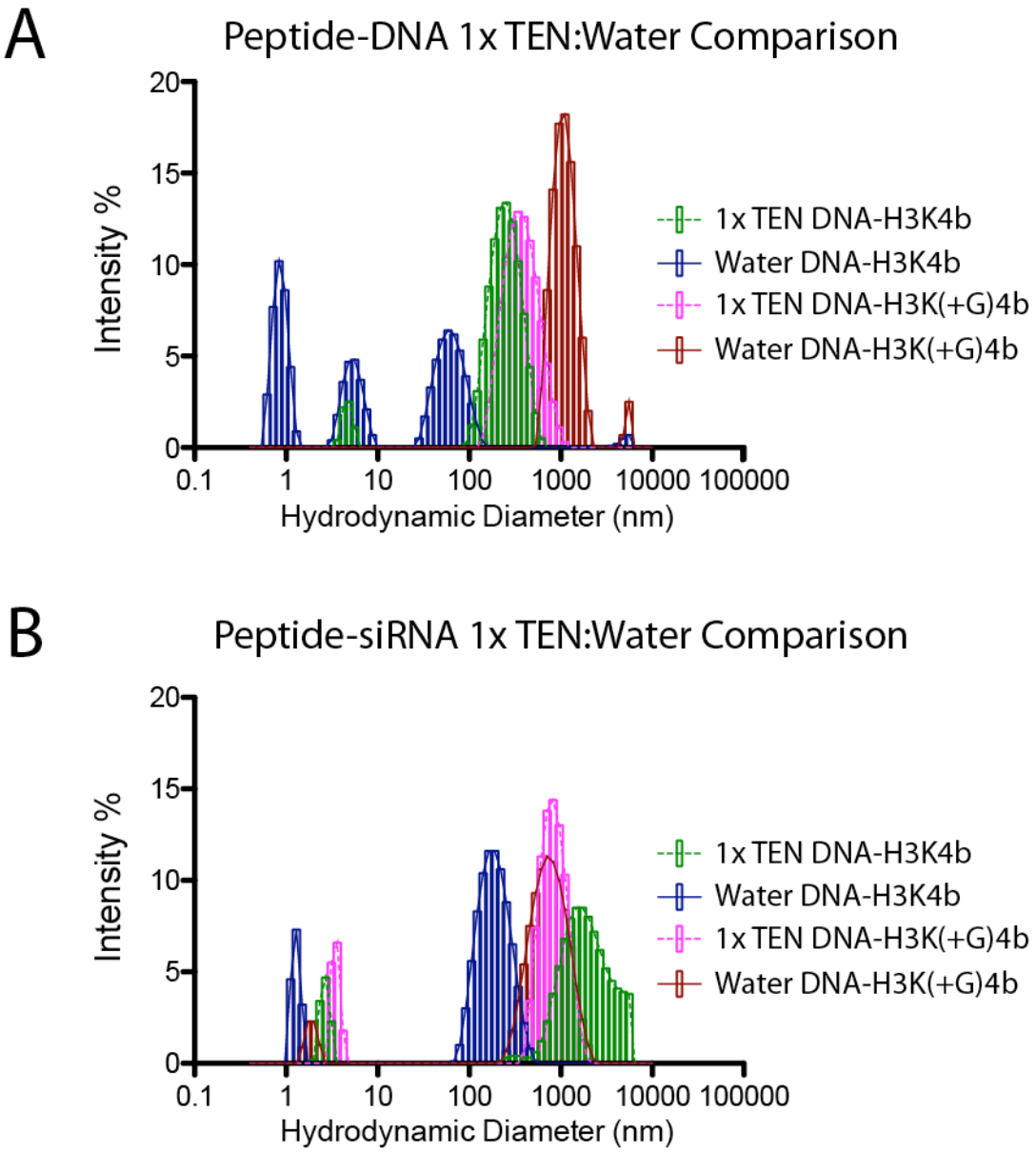


Figure 4.3 ***Comparison of Particle Size in 1x TEN to Water.*** **DLS analysis of H3K4b and H3K(+G)4b-DNA and siRNA complex formation in 1x TEN or water. A. DLS analysis of peptide-DNA complexes. B. DLS analysis of peptide-siRNA complexes. H3K4b complexes show a greater change in size in the presence of water compared to 1x TEN. H3K(+G)4b complexes show more resistance to change in particle size between the two solvent conditions.**

complexes. If there is no salt to shield charge and for counterion release, we should see a smaller particle size. Figure 4.3 shows the results for the particle size

distribution found in water compared to 100 mM NaCl. For H3K4b-siRNA and H3K4b-DNA complex formation, particles were significantly smaller in the absence of 100 mM NaCl (size results are presented in Table 4.3). H3K4b-DNA dropped

	Particle Size at Peak Intensity in 1x TEN	Particle Size at Peak Intensity in H_2O
H3K4b-DNA	255 nm	59, 5.6 and 0.8 nm
H3K(+G)4b-DNA	342 nm	1.1 μm
H3K4b-siRNA	1.5 μm	164-190 nm
H3K(+G)4b-siRNA	825 nm	712 nm

Table 4.3 ***Comparison of Particle Size in Buffer Compared to Water*****. DLS analysis of H3K4b and H3K(+G)4b-DNA or siRNA complex formation in 1x TEN, 100 mM NaCl or H_2O. The size associated with the peak intensity for each complex is reported in this table. A decrease in complex size is very clear for H3K4b with removal of NaCl. H3K(+G)4b-DNA complex size are smaller than seen previously in Table 4.1 and show an increase in size in the absence of NaCl H3K(+G)4b-siRNA demonstrates less change in particle size.**

from 255 nM in size to 59 nm and smaller size particles and H3K4b-siRNA dropped from 1.5 μm to between 164-190 nm. For H3K4b-DNA complexes, the 5.6 and 0.8 nm particles could be small aggregates of peptides and individual peptides based on their small diameter. Sizes observed for H3K4b-DNA and H3K4b-siRNA complexes in NaCl were very similar to what was observed previously in Figure and Table 4.1. In contrast, The H3K(+G)4b peptide showed a greater resistance to change in particle size in the presence of water, which is a surprising result. H3K(+G)4b-DNA complex demonstrated a relatively significant increase in size in water, however the complex size observed in 100 mM NaCl was smaller than observed previously (Table 4.1). H3K(+G)4b-siRNA complex showed a much larger size for complex formation that observed before which stayed relatively the same size in water. The size results for the H3K(+G)4b-DNA and H3K(+G)4b-siRNA complex formed were inconsistent

with what was observed in our previous DLS experiment, demonstrating a much larger size of particle formation than was expected.

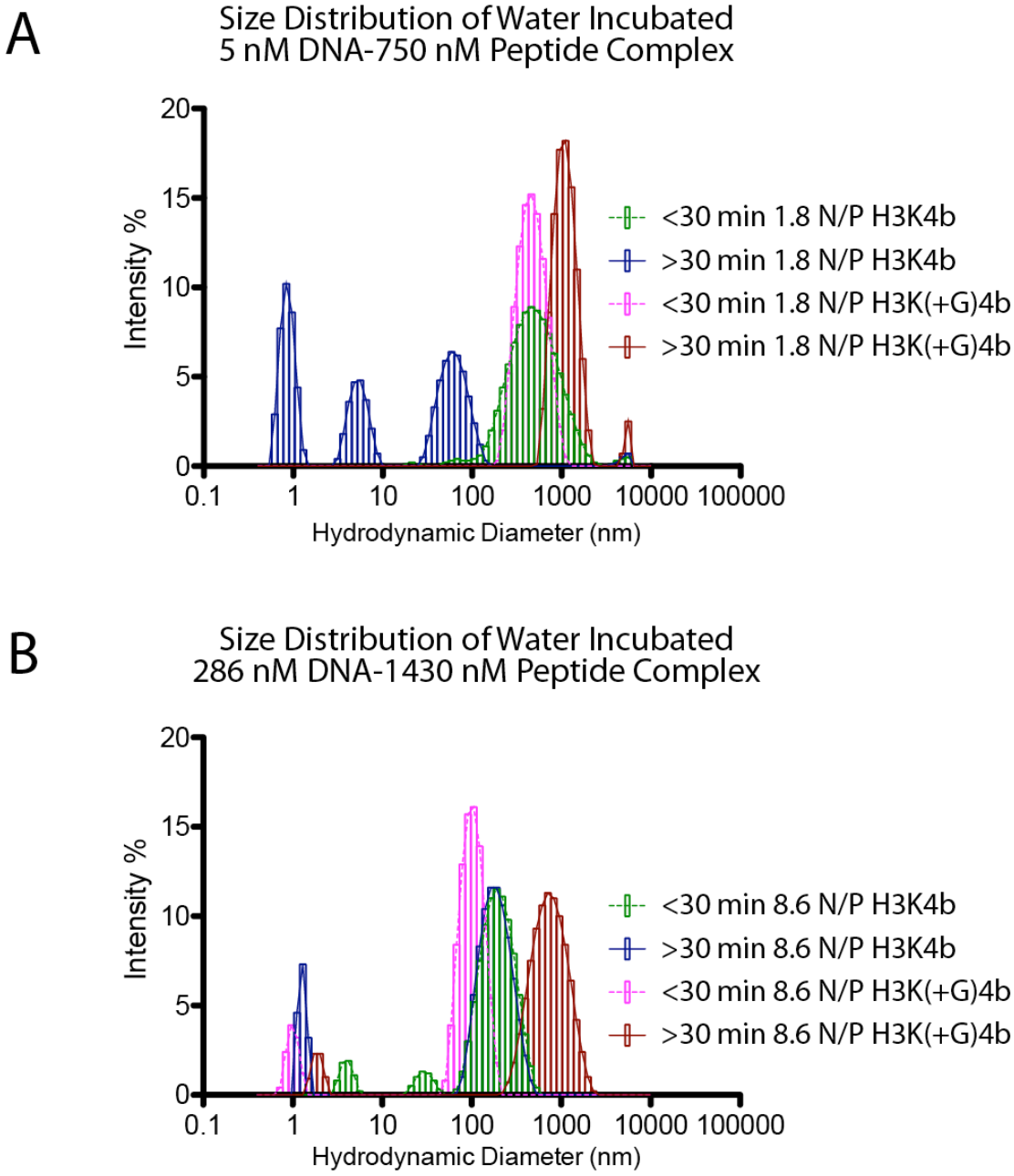


Figure 4.4 ***Time Dependent Formation of Peptide-Nucleic Acid Complexes in Water.*** **DLS analysis of H3K4b and H3K(+G)4b-DNA and siRNA complex formation in water with more or less than 30 min incubation. A. DLS analysis of peptide-DNA complexes. B. DLS analysis of peptide-siRNA complexes. Complexes formed with H3K(+G)4b showed a decease in size when freshly prepared compared to a long incubation period in water. H3K4b complexes demonstrated less change in size with the varying the length of incubation in water.**

The only logical explanation that might describe the differences between observations in Figure 4.1 and 4.3 is the time dependence of forming peptide-nucleic acid complexes. For the DLS experiments conducted in Table 4.3 the first sample that was measured was 30 minutes after preparation. Subsequent measurements took place over a 2 hour time period. This means that the siRNA samples, which were the last to have a reading taken, were allowed to form for 1-2 hours. Previously, the complexes were only allowed to form for up to 22 minutes (Table 4.1). This observation was only a bi-product of the actual experiment that was intended, but, had to be explored in more detail. Therefore, a series of freshly incubated peptide-DNA and siRNA samples were created in water and readings were taken immediately after mixing (Figure 4.4, Table 4.4). Figure 4.4 shows that for H3K(+G)4b-nucleic

	Particle Size at Peak Intensity <30 min Incubation in H_2O	Particle Size at Peak Intesnity >30 min Incubation in H_2O
H3K4b-DNA	459 nm	59, 5.6 and 0.8 nm
H3K(+G)4b-DNA	459 nm	1.1 μm
H3K4b-siRNA	190 nm	164-190 nm
H3K(+G)4b-siRNA	105 nm	712 nm

Table 4.4 ***Sizes of Peptide-Nucleic Acid Complexes with Different Durations of Incubation in Water*****. DLS analysis of H3K4b or H3K(+G)4b-DNA or siRNA complex formation in H_20 with more (scans greater than 30 minutes were take between 1-2hours after combining components) or less (scans less than 30 minutes were taken 4 minutes after combining components) than 30 minutes of incubation. Complexes formed with H3K(+G)4b, which showed an unexpectedly large size in the presence of water in previous results (Table 4.3), decreased in size when freshly prepared. H3K4b-DNA complexes demonstrate an increase in the size of complex formed.**

acid complexes there is a decrease in the size of the complex when freshly prepared. H3K4B-DNA is the only complex where there was an increase in size when freshly prepared. This could just be an increase in particle aggregation for this particular preparation. In all cases there is some change in particle size based on the time of

preparation or the salt content, suggesting time and salt concentration are important factors to consider when looking at the size of these complexes.

4.2.3 AFM Images of H3K4b and H3K(+G)4b-Nucleic Acid Complexes

A huge thanks goes to my collaborator and lab member, Jason Hustedt, for conducting an analysis of all AFM images acquired of H3K4b and H3K(+G)4b with plasmid DNA and siRNA. The AFM analysis here is preliminary data presented with permission from Jason Hustedt. A comprehensive understanding of H3K4b and H3K(+G)4bs interaction would not have been possible without Jason's thoughtful insight and analysis of the AFM data.

AFM images were obtained for H3K4b or H3K(+G)4b with either plasmid DNA or siRNA. Ratios of peptide-nucleic acid prepared in water were the same as in transfection experiments conducted in the Mixson lab. After complexes were allowed to form, they were pipetted onto a mica surface and imaged by the AFM. Topographical and phase images are presented for each complex formed with H3K4b or H3K(+G)4b and plasmid DNA in Figure 4.2. Topography images are acquired by observing differences in oscillation of the cantilever as it scans across the surface. The oscillation amplitude of the cantilever will change as it scans across areas with raised surfaces. Phase imaging looks at the phase shift in the oscillation of the cantilever as it interacts with the material on the surface. The hydrophilic tip interacts with variations in the surface and peptide-nucleic acid complexes we are observing. These variations will cause a shift in the phase of the cantilever oscillation, which can be interpreted by the instrument to develop a more detailed picture of the surface.

Both peptides formed complexes that yielded sphere-like particles on the mica

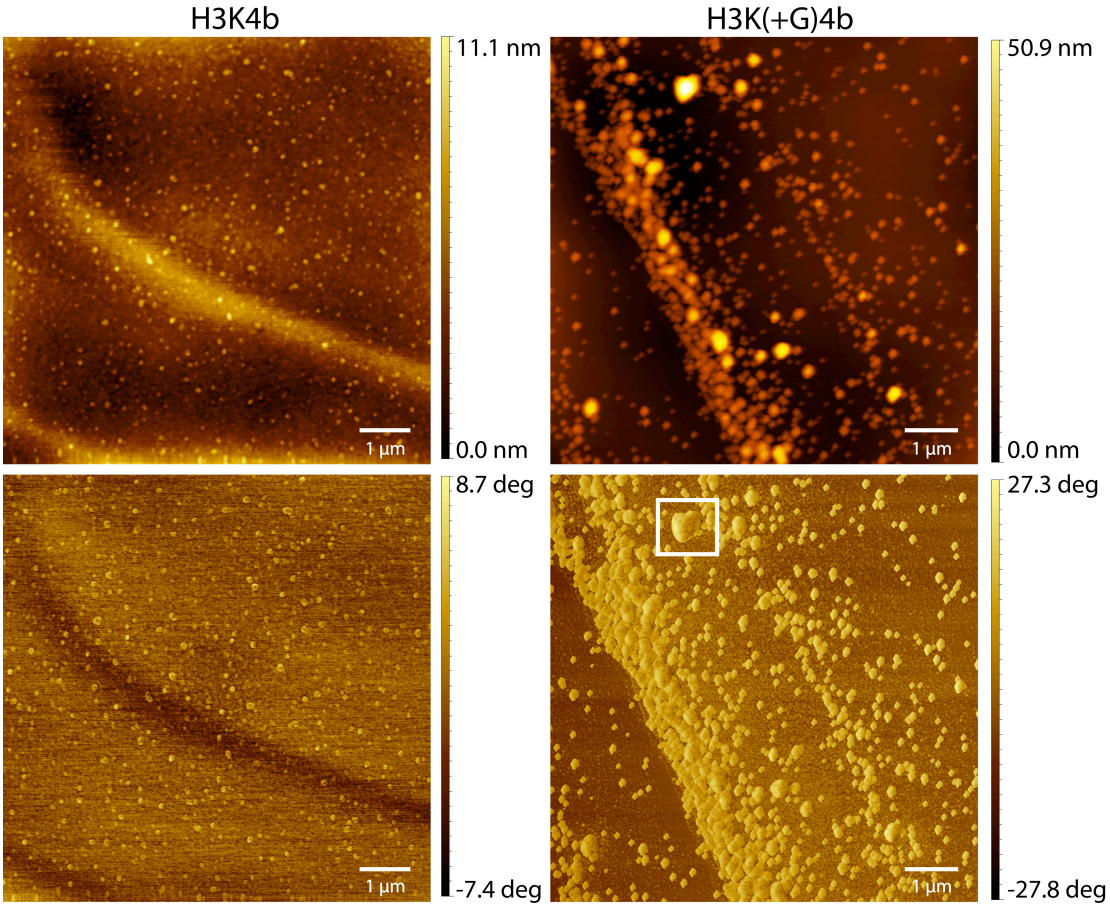


Figure 4.5 ***AFM Images of H3K4b and H3K(+G)4b-DNA Complexes.*** **Topography (Top two panels) and phase (bottom two panels) images are presented for complexes formed with H3K4b and H3K(+G)4b with plasmid DNA. Topography images are presented in the top panels and phase images in the bottom panels. H3K4b forms smaller aggregate structures on the surface. Larger complexes of H3K(+G)4b appear to be an aggregation of smaller, blob-like structures, which can be seen in the large aggregate highlighted by the white box.**

surface. H3K4b-plasmid complexes formed are slightly smaller in size than H3K(+G)4b having a maximum height of about 11.1 nm giving a smaller height to diameter ratio. Formed H3K(+G)4b-plasmid DNA complexes demonstrated a larger structure with a maximum height of 50.9 nm yielding a larger height to diameter ratio. Very large clumps (~1 µm) are comprised of many smaller (~100 nm), globular

structures that are aggregated together. Demarcations of individual complexes that

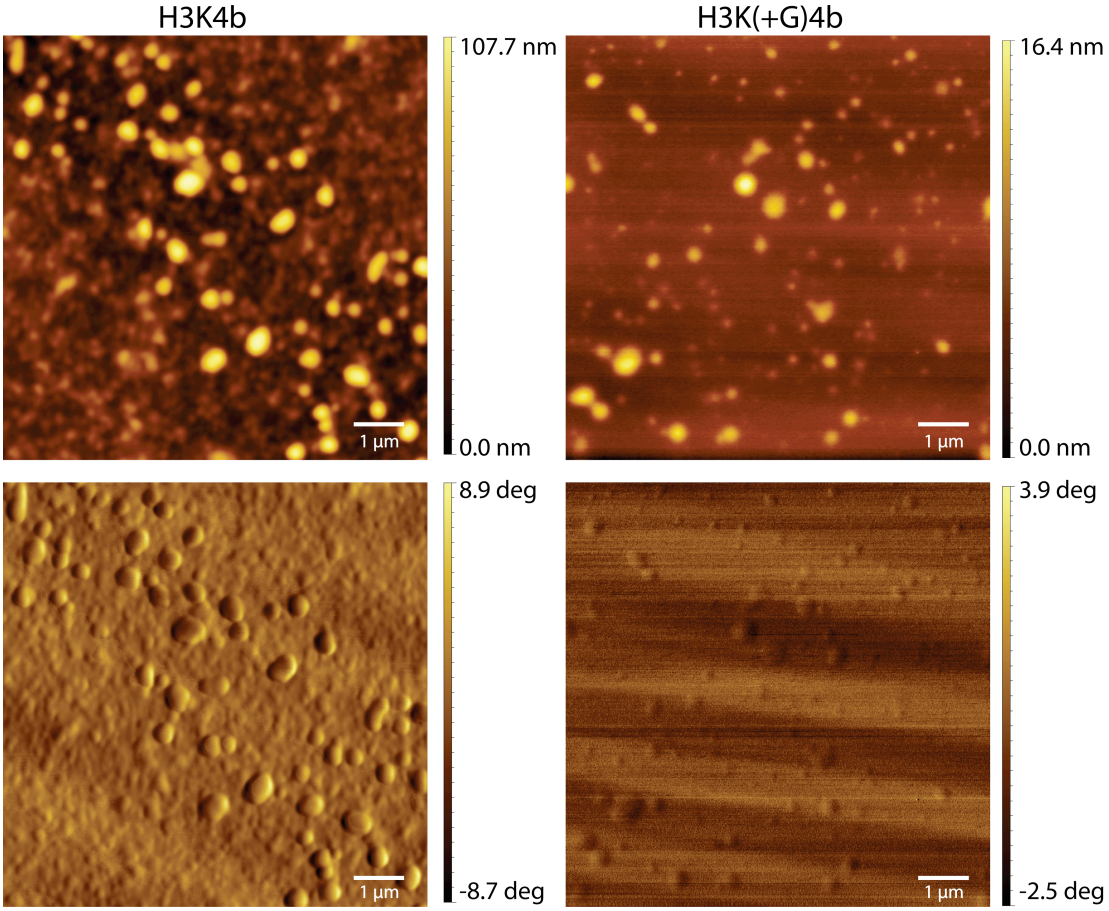


Figure 4.6 ***AFM Images of H3K4b and H3K(+G)4b-siRNA Complexes.*** **Topography (top two panels) and phase (bottom two panels) images are presented for complexes formed with H3K4b and H3K(+G)4b with siRNA. Topography images are presented in the top panels and phase images in the bottom panels. Complexes formed with H3K4b appear to be larger. H3K4b complexes demonstrate a more evenly distributed material. H3K(+G)4b particles are seen to be much smaller in size and are sparsely distributed on the surface.**

are aggregated together are clearly visible in the phase image of H3K(+G)4b-DNA complexes.

Complexes formed with H3K4b and siRNA are much more uniform in size and morphology, compared to H3K(+G)4b-siRNA complexes, based on their appearance (Figure 4.3). Particles formed with H3K4b hold their spherical nature very well with a maximum height of 107.7 nm, giving a larger height to diameter ratio than H3K(+G)4b. H3K(+G)4b-siRNA particles appear to flatten out to a much

greater extent on the surface with a maximum height of only 16.4 nm, yielding a much smaller height to diameter ratio.

4.3 Discussion:

4.3.1 Important Considerations in Peptide-Nucleic Acid Particle Formation

Some limited conclusions can be drawn for the observations made in the DLS analysis between H3K4b and H3K(+G)4b particle formation with nucleic acids. H3K(+G)4b-DNA particle formation was the most consistent in size out of any complex formed, giving the lowest PDI range, between 0.15-0.26. In general, peptide-siRNA complexes had a higher range of PDIs, from 0.16-0.48, indicating a very broad distribution of particle sizes. The difference in PDIs observed could be indicative of the overall stability of particle formed with DNA is higher compared to that of siRNA. Because of the DNA's larger size, it nucleates many more peptides per molecule to form an individual complex. Therefore, larger aggregates of particles are a combination of smaller, more stable nucleated particles. In the case of the siRNA, N/P ratios are relatively high (N/P 8.6) meaning there is a significant excess of peptide. The excess peptide could be having an impact on the consistency of particle size. In addition, the change in particle size over time could make a strong case for charge neutralization of individually packed nucleic acids. For example, H3K(+G)4b-DNA complexes shows a significant change in size over time compared to that of H3K4b-DNA. If H3K(+G)4b is more completely coating the DNA and neutralizing more charge from the phosphate backbone of the DNA, this would allow for more aggregation of individual complexes into a larger polyplex. The results for

DNA complex formation are also conducted below a 1:1 charge neutralization ratio (N/P 1.8) which may indicate that H3K(+G)4b is coating DNA more completely at lower N/P ratios. However, The presence of both NaCl and duration of incubation have a large impact on particle formation with both peptides.

DLS Data gathered previously from the Mixson lab demonstrated the size of these complexes is smaller overall in water than in a buffering system containing salt (Mixson: unpublished results). Mixson has shown that complexes are smaller (on the order of 250 nm) when formed in the presence of water, which correlates to our findings. When complexes are formed in Opti-MEM solution, which is a reduced serum transfection media, they are shown to have a much larger complex diameter, ranging on the order of micrometers. This leads to the conclusion that peptide-nucleic acid particles formed in buffering solutions containing salt and other additives could have the possibility to aggregate into large polyplexes. Extended incubation of H3K4b and H3K(+G)4b with nucleic acid can also lead to the formation of large aggregates as seen in Figure 4.3. Therefore if an ideal size of particle is desired for transfection, time and salt concentration will have to be carefully monitored.

Some studies have indicated that ideal size for optimal uptake of nanoparticles is on the order of 100 nm in diameter[72]. Particles greater than 500 nm in size, aggregate around the cell surface and show limited cellular entry[72]. However, it has been reported in some cases that particles greater than 500 nm in diameter have had better success in delivering transfection than smaller particles[72; 73]. In general, uptake will be a product of the route of entry of the particle. Intuitively, smaller particles might be the ideal choice for optimal transfection because they will aggregate less

and be more readily taken up by the cell. For particles around 100 nm in diameter, caveolin and clathrin-mediated endocytosis are the cellular entry route[73]. For particles much larger in size, greater than 1 μm, the macropinocytosis path can allow entry of these molecules[73]. It will be important to the future analysis of gene delivery particles to take into consideration all potential pathways of entry[73].

4.3.2 Calculations of Nucleic acid and Peptide Contained in a Complex

In addition to the route of entry, the size of the particle can determine the amount of genetic material contained inside. With the data gathered about the hydrodynamic diameter of the complex, a series of simple calculations can be demonstrated to yield an approximate value for how many nucleic acids and peptides are contained in a given size complex. For supercoiled DNA, our DLS results measured a 3 kb plasmid to have a diameter of 51 nm. This is a reasonable approximation considering other supercoiled plasmids of a similar size have measured around 100 nm in diameter[74]. First, we can calculate the volume of our 3 kb plasmid given the diameter and using the calculation for the volume of a sphere, assuming the supercoiled plasmid is roughly spherical in size:

$$Volume = \frac{\pi (d)^3}{6}$$

$$d = 51 \text{ nm}$$

$$V = \frac{\pi (51 \text{ nm})^3}{6} = 69{,}456 \text{ nm}^3$$

To compare to the value obtained above, the volume DNA occupies if it were a 3 kb cylindrical rod will be calculated. B-form DNA is about 0.34 nm per base pair and pRset is 3 kb long.

3,000 bp * 0.34 nm / bp=1,020 nm

A 3 kb plasmid is 1,020 nm in height and B-form DNA has a diameter of 2 nm. Using the formula for the volume of a cylinder, we can calculate the volume occupied by our DNA as a rod.

$$V = \frac{\pi(d)^2 h}{4}$$

$$V = \frac{\pi(2\text{ nm})^2(1{,}020\text{ nm})}{4} = 3{,}204\text{ nm}^3$$

If we take the volume above and plug it into the equation for the volume of a sphere, we can get the diameter of that sphere.

$$d = \sqrt[3]{\frac{3{,}204\text{ nm}^3 * 6}{\pi}} = 18\text{ nm}$$

The diameter for the sphere calculated for the volume of the rod is smaller (18 nm) compared to that of the experimental value for diameter. In addition, the volume occupied by the rod is about 10 fold smaller than that of the sphere estimated from the experimental diameter of a plasmid, presenting a range of sizes for the volume occupied by the DNA.

To calculate the volume occupied by the 21mer siRNA used in the DLS experiments, the estimated dimensions of a 21mer A-form RNA were used to calculate the volume of a cylindrical rod. A-form RNA is about 0.24 nm per base pair and the siRNA is 21 bp long.

$$21 \text{ bp} * 0.24 \text{ nm / bp} = 5.04 \text{ nm}$$

A 21 bp siRNA is 5.04 nm in height and A-form RNA is 2.3 nm in diameter. Again, using the equation for the volume of a cylinder we can calculate the volume.

$$V = \frac{\pi(d)^2 h}{4}$$

$$V = \frac{\pi(2.3 \text{ nm})^2 5.04 \text{ nm}}{4} = 21 \text{ nm}^3 \text{ siRNA}$$

The volume of a 3 kb B-form DNA rod is 3,204 nm^3 and of a 21 bp siRNA rod is 21 nm^3.

The peptide molecule itself is small in size, however, because of the sheer quantity bound, they will likely take up space inside the particle. An assumption will be made that the volume of the peptide will occupy about half of the volume of the complex, if there is complete coating of the nucleic acid. A similar assumption will be made for the siRNA-peptide particles as well. Smaller size H3K4b and H3K(+G)4b-nucleic acid bound particles are around 100 nm in diameter. The volume of a particle of such size is as follows:

$$V = \frac{\pi(100 \text{ nm})^3}{6} = 523{,}599 \text{ nm}^3$$

Only half of this volume will be occupied by the nucleic acid.

$$\frac{523{,}599 \text{ nm}^3}{2} = 261{,}800 \text{ nm}^3$$

Now we know the volume of a 100 nm particle and what volume a 3 kb packaged plasmid DNA occupies according to the experimental data, so we can figure out roughly how many plasmids are inside a particle 100 nm in diameter:

$$\frac{261{,}800\ \text{nm}^3}{69{,}456\ \text{nm}^3 / 3\ \text{kb plasmid}} = 4\ \text{plasmids / 100 nm}$$

We can also calculate the number of plasmids that will be in a 100 nm particle if we assume the volume calculated for our B-form DNA rod is correct.

$$\frac{261{,}800\ \text{nm}^3}{3{,}204\ \text{nm}^3\ / 3\ \text{kb plasmid}} = 82\ \text{plasmids / 100 nm}$$

For siRNA we can do a similar calculation based on the volume an A-form siRNA rod occupies:

$$\frac{261{,}800\ \text{nm}^3}{21\ \text{nm}^3\ / 21\text{mer siRNA}} = 12{,}467\ \text{siRNAs / 100 nm}$$

The two values for the amount of DNA contained in a 100 nm particle, are different based on either the volume occupied by the supercoiled DNA or a DNA as a cylindrical rod. A particle 100 nm in diameter containing 4 plasmids seems reasonable but could be an underestimation. The value of 82 plasmids per 100 nm particle, for the volume of the DNA as a rod, is also a reasonable estimation but could be on the high end. Using the DNA rod volume estimation for the number of nucleic acids per a particle, the number of peptides in a 100 nm particle can be estimated.

If we assume a binding stoichiometry of 250 peptides bound per plasmid for complete charge neutralization (6,000 phosphates for a 3,000 bp DNA), we can also calculate the number of peptides per complex (since the volume of the peptides is half the volume of the sphere).

82 plasmids * 250 peptides / plasmid=20,500 peptides

We assume 1.75 peptides bound per siRNA for charge neutralization (42 phosphates for a 21 bp siRNA):

12,467 siRNA * 1.75 peptides / siRNA=21,817 peptides

With the above calculations we now can estimate that there are about 82 plasmids and 20,500 peptides per 100 nm particle. For siRNA-peptide complexes there are 12,467 siRNAs and 21,817 peptides per 100 nm particle. Typically the sizes seen for the peptide-plasmid or siRNA particles are anywhere between 200-800 nm in diameter, which means there will be more plasmids, siRNAs and peptides (scaling with the cube of the diameter) inside any given complex. There is a large room for error in these estimations, by a factor of ten as shown by difference in the calculated amounts of plasmid based on the volume occupied by a supercoiled DNA as a sphere compared to the volume occupied by a rod.

Biological systems have a much more efficient means of packaging nucleic acids, which H3K4b and H3K(+G)4b may not be able to emulate. A virus particle's capsid, which contains the genetic material can range from 20-120 nm in diameter[75]. In the case of the Herpes-simplex virus, a 152 kb double-stranded linear DNA is stored in a 125 nm in diameter viral capsid[76; 77]. Thus, viruses will be representative of the most efficient use of space for packaging of genetic material that nature can afford. The Herpes capsid is about the same diameter as our 100 nm particle. We can calculate how many 3 kb plasmid DNA molecules can be contained in the capsid based on the fact it holds up to 152 kb of DNA.

$$\frac{152 \text{ kb}}{3 \text{ kb}} = 51 \text{ plasmids}$$

A Herpes capsid 125 nm in diameter could hold up to 51 plasmids 3 kb in size. For comparison to a different virus, a T7 bacteriophage capsid is 60 nm in diameter and holds about 40 kb of DNA[78; 79].

$$\frac{40\text{ kb}}{3\text{ kb}} = 13\text{ plasmids}$$

On average, a virus capsid about 100 nm in diameter would be able to hold between 20-50 3 kb plasmids. Our estimations predicted of about 82 plasmids per 100 nm particle, which is still high compared to the amount of genetic material a virus can hold. It is unlikely that packaging of the H3K4b or H3K(+G)4b with DNA or siRNA will be as effective as a virus, so the amount of nucleic acid held in a complex formed with these peptides will be significantly less. Taking into consideration the calculations made for the virus, the calculated range of 4-82 plasmids held per 100 nm particle and the possibility for further compaction of the plasmid upon complexation with the peptide, the amount of nucleic acid held will most likely be on the order of 10s of plasmids per particle. Overall, the calculated estimates amount of siRNA's and DNA's in a given particle makes available information about the amount of genetic material contained for delivery to the cell.

4.3.3 AFM Morphology and DLS Assessment of H3K4b and H3K(+G)4b-Nucleic Acid Complexes

AFM images and DLS of H3K4b or H3K(+G)4b complex formation with siRNA or DNA provide insight on the morphology of nucleic acid delivery complexes. Differences in H3K4b and H3K(+G)4b morphology can provide a connection between the biological function of these peptides and how they appear. The AFM images of H3K4b-plasmid and H3K(+G)4b-plasmid complexes show differences in their height to volume ratio. In the phase imaging, the large structures formed for H3K(+G)4b appear to be an aggregation of smaller blob-like particles

packed together. Previous results demonstrate H3K(+G)4b better coats plasmid DNA providing better protection and charge neutralization at lower concentrations. The greater extent of charge neutralization for H3K(+G)4b peptide may result in the larger particles observed, consisting of smaller ones aggregated together. The other possibility is that H3K(+G)4b does not form overcharge regions of aggregated peptide, therefore we can observe individual demarcations inside large complexes. The DLS data also demonstrates that H3K(+G)4b forms an aggregating structure that gets larger over a 22 minute time period. H3K(+G)4b's complex formation has a larger height to diameter ratio that could be an indication a more stable structural formation. H3K(+G)4b's better transfection of plasmid DNA could be due to greater structural stability. DNA complexes formed with H3K4b also demonstrate aggregation of individual particles, but to a lesser extent. DLS results confirm that H3K4b complex form smaller structures. Overall the morphologies of H3K4b and H3K(+G)4b present some subtle differences that are in agreement with differences in packaging observed in other experiments.

A large difference is observed between the morphologies of H3K4b and H3K(+G)4b complexes with siRNA. H3K4b has a much larger height to diameter ratio than H3K(+G)4b, suggesting a more ridged structure. H3K4b transfects siRNA better than H3K(+G)4b, perhaps due to a more rigid structure that protects siRNA better. DLS results confirm that H3K4b-siRNA complexes are larger in size.

Chapter 5: Conclusions

5.1 Model for H3K4b and H3K(+G)4b Binding to Plasmid DNA

The data on H3K4b and H3K(+G)4b binding interactions with nucleic acid lead to a model for the complex formation. Each piece of data gathered, as it stands alone, presents a supportive piece of evidence for how both peptides complex with nucleic acid. However, when pieced together, these data provide a very strong body of evidence for the mode of action of each of these peptides. This model reflects the key differences observed for H3K4b compared to H3K(+G)4b in transfection efficiency, the DNase I degradation, fluorescence dye exclusion experiments, DLS, AFM, and ITC. Both the DNase I and EtBr exclusion experimental models have been used as the basis for the overall mechanistic model for peptide-nucleic acid complex formation. This overall model is presented in Figure 5.1 for H3K4b and H3K(+G)4b binding to plasmid. The key observations of each experiment were put into context with one another to explain how this model was formulated.

For H3K4b, the EtBr experiments revealed that more peptide was required to reach the same extent of EtBr exclusion as H3K(+G)4b. However we also observed that H3K4b reached a greater extent of EtBr exclusion than H3K(+G)4b at saturating concentrations of peptide. How can the differences observed here be reconciled? What is believed to occur is that H3K4b, not having the added glycines in each branched region, has more steric hindrance and is not as flexible as H3K(+G)4b. As seen in Figure 5.1, upon initial addition and binding of H3K4b below a 3 N/P (1:1 charge:charge) ratio, there is incomplete coating of plasmid DNA due to peptide inflexibility. Part of the incomplete coating observed could be due to the fact H3K4b cannot bring all of its lysines into contact with the DNA, which might lead to regions

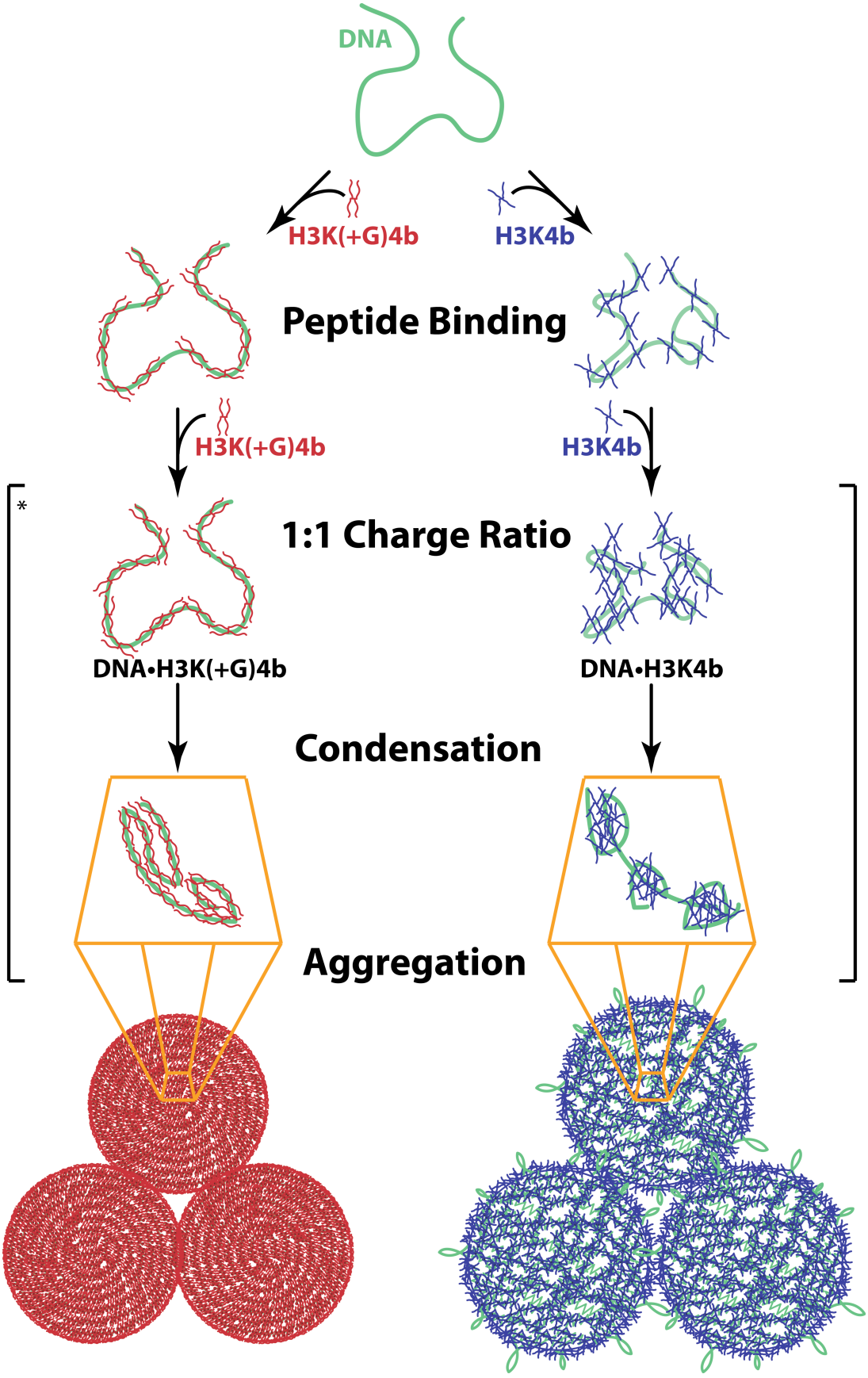

DNA
H3K(+G)4b
H3K4b
Peptide Binding
H3K(+G)4b
H3K4b
*
1:1 Charge Ratio
DNA•H3K(+G)4b
DNA•H3K4b
Condensation
Aggregation

Figure 5.1 ***Coating Versus Clumping Mechanism for H3K4b and H3K(+G)4b Complex Formation with Plasmid DNA*****. Only a section of plasmid DNA binding for both peptides is shown initially in order to illustrate binding on an individual peptide scale. Starting at the top, H3K4b or H3K(+G)4b peptide is added to plasmid DNA. Peptide binding at lower concentration occurs through some hydrogen bonding and electrostatic interaction with the DNA. Small amounts of peptide binding will not yield change in DNA compaction because very little charge neutralization has occurred. As more peptide is added and a 1:1 charge neutralization ratio of peptide to DNA is reached, DNA will start to compact and the complex will begin to aggregate. The starred bracket denotes one single step that is comprised of two simultaneous intermediates. Binding of peptide and compaction of plasmid occur simultaneously once charge neutralization is approached. H3K4b (blue) incompletely coats the plasmid DNA, leaving gaps of naked DNA during complex formation. Condensation of plasmid DNA with H3K4b occurs with pockets of peptide and plasmid DNA clumping together. This can lead to overcharged regions of clumped peptide. Upon aggregation, H3K4b forms a less densely packed aggregate that has regions of clumped peptide and exposed DNA on the surface. H3K(+G)4b coats plasmid DNA more evenly due to the increased flexibility of the added glycines, which leads to less exposed regions of plasmid DNA. During condensation, H3K(+G)4b folds the DNA leading to better charge neutralization and compaction of DNA, forming a densely packed aggregate.**

of excess charge in the complex. Upon addition of more peptide and at saturating concentrations of H3K4b, the peptide fills in some of the naked regions of plasmid DNA, demonstrated by its greater exclusion of EtBr. However, even at the 3.6 N/P ratio, H3K4b never reaches complete coating of the plasmid. The evidence of H3K4bs lesser ability to coat plasmid than H3K(+G)4b was unclear from the EtBr exclusion assay alone because H3K4b ended up excluding more EtBr at saturating concentrations. The DNase I degradation experiment, conducted at a range of N/P ratios corresponding to those of the EtBr experiment, proved that there are exposed regions of plasmid which H3K4b does not protect. The reason these exposed regions of plasmid exist could be due to overly positively charged regions of H3K4b-plasmid complex makings these regions of plasmid inaccessibility to additional positively charged peptide. Zeta potential measurements would tend to corroborate this because other HK peptide complexes have proved to be positively charged in the past, indicating an excess of positive charge[44]. The binding of a 1:1 charge to charge ratio of peptide to DNA is a combination of two simultaneous intermediates, H3K4b binding and condensation of the peptide-plasmid complex. The ITC data

demonstrated initial enthalpically driven binding of the peptide to the DNA, followed by an entropically driven reaction, which is due to electrostatic interaction driven by release of counterions and water into solution. Locally condensed regions of overcharged peptide and plasmid DNA are presumed to resemble "beads on a string", in which there are protected "beads" of peptide-plasmid complex, with exposed "string" regions of naked plasmid in between. Evidence for the "beads on a string" have come from single molecule pulling experiments conducted by Dr. Joonil Seog's lab at the University of Maryland, Department of Material Science and Engineering (unpublished data). These experiments have observed how much force it takes to pull apart a complex formed with a single piece of DNA that has peptide bound to it.

The final part of the electrostatic interaction between H3K4b and DNA binding is aggregation of the overcharged regions of particles formed. DLS and AFM results corroborate condensation of peptide-plasmid complexes. The ITC data for H3K4b complex formation demonstrates a higher enthalpy of formation, possibly between overcharged regions of the complex. DLS results for H3K4b-plasmid complexes revealed a smaller size than those made with H3K(+G)4b. The smaller size observed might be due to a less stable structural formation for the aggregate. The final phase of H3K4b binding includes the aggregation of multiple peptide-DNA complexes together. From DLS and AFM results we know the 300 nM particle size for a complex is composed of many individual complexes aggregated together. As indicated by AFM results, complexes formed with H3K4b tend to be smaller and have a lower height to diameter ratio. Clumped "beaded" regions of peptide on the plasmid may also contribute to less dense packaging of the H3K4b-DNA complex.

The results for H3K4b-DNA complex formation suggest that it is a less stable complex, which has a greater susceptibility to nuclease degradation making it less effective for plasmid transfection.

H3K(+G)4b is believe to be have less steric hindrance than H3K4b due to the added glycines interspersed in between the beginning and middle of each of the four branched regions. The EtBr exclusion assay suggests that H3K(+G)4b in complex with plasmid reaches its full extent of EtBr exclusion at a lower N/P ratio close to 1.8 compared to H3K4b. Full extent of EtBr exclusion at lower N/P ratios indicates H3K(+G)4b's effectiveness at coating and excluding EtBr at lower concentrations suggesting that more lysines can interact with the DNA. Upon reaching a 1:1 charge ratio the peptide has even distributed on the plasmid DNA completely covering it. At saturating concentrations more peptide binds and excludes slightly more EtBr, however, the full extent of EtBr exclusion has mostly (90%) been reached by this point. DNase I results strongly corroborated this theory, demonstrating complete protection of plasmid DNA by H3K(+G)4b at a 3.6/1 N/P ratio. As H3K(+G)4b binds there is also a condensation event in which compaction of the peptide-plasmid complex occurs. Because of the even coating of H3K(+G)4b, there is tighter compaction of the condensed complex formed, with little to no naked regions of plasmid DNA. DLS demonstrates better charge neutralization for an H3K(+G)4b complex because it forms a larger complex than H3K4b, indicating less charge repulsion between complexes due to complete coverage. These charge neutralized complexes come together to form a larger aggregated particle of DNA and peptide. AFM results confirm this because of the larger height to diameter ratio observed,

which demonstrates H3K(+G)4b's larger size. Demarcations inside large complexes can also be observed from the AFM results, presenting that aggregates of peptide and plasmid are comprised of many complexes. The experimental evidence demonstrate H3K(+G)4b contributes to formation of a well-coated and compact plasmid-peptide complex, which is very stable during serum transport. The enhanced stability of H3K(+G)4b-plasmid complex formation makes it a good transfection agent of plasmid.

Overall, the data gathered for siRNA binding did not lead to a conclusive model of peptide binding that describe the difference in siRNA transfection efficiencies presenting H3K4b as more effective than H3K(+G)4b. From the proposed model of H3K4b and H3K(+G)4b binding to plasmid DNA, we can speculate as to what aspects of peptide-siRNA complex formation cause the observed difference in transfection efficiencies. We submit that those properties that make H3K4b an ineffective agent of delivery for plasmid, makes it effective for shorter nucleic acids like siRNAs. It is believed that H3K4b forms more clumped and aggregate regions during complex formation with plasmid. These clumped "bead regions" of peptide are highly positively charged and make it difficult for additional peptides to fill in unprotected regions of the plasmid. However, in the case of siRNA this clumping may be much more effective for delivery. The siRNA being only 21 bp in length is significantly shorter than the plasmid and has a very rigid structure[80]. Plasmid DNA, being much larger, flexible and having a complex topology, would have to rely on effective coating and compaction for delivery. In the case of the siRNA, H3K4b would still be forming clumped regions of peptide and siRNA,

however, no folding of the nucleic acid is needed for protection. The final complex formed could be similar to that of a dendrimeric one where there is a cage that forms around the siRNA and protects it[35]. This way the small, stiff siRNA molecules are imbedded inside an aggregate of H3K4b that is formed by the charge neutralization occurring between the nucleic acid and peptide. We know from the DLS results that complexes formed with siRNA and H3K4b are larger than those of H3K(+G)4b indicating better charge neutralization and aggregation. The AFM matches the DLS results, which present H3K4b-siRNA particles with a larger height to diameter ratio.

The observations made about H3K(+G)4b binding to plasmid DNA describe why it is completely ineffective for siRNA delivery. H3K(+G)4b may have more flexibility allowing for better coating of plasmid, which may not be a desired attribute for small siRNA complex formation and delivery. The enhanced flexibility may not provide the structural support needed to form a stable complex. We know from AFM and DLS results, H3K(+G)4b-siRNA complexes are less uniform and have a smaller height to diameter ratio. In addition, H3K(+G)4b may actually be coating the siRNA better than H3K4b. The problem with effective coating may be that binding is too tight and the siRNA may not be getting released from the complex to its target. These factors contribute to H3K(+G)4b ineffectiveness at delivery of siRNA.

5.2 Implications of Improved Transport afforded by H3K4b and H3K(+G)4b in Serum

The biological relevance of characterizing of H3K4b and H3K(+G)4b is of paramount importance. A connection has to be made between characterization of gene delivery complexes and their biological impact in order for progress to be made

towards creation of better delivery agents. Out of all of the factors for effective deliver, only two major aspects in formation of the complex are of immediate importance to delivery. One of those is transport of the nucleic acid during delivery. Three factors are important when considering travel of the complex to its target: condensation, protection and localized delivery of the nucleic acid. Condensation of nucleic acids is the simplest, yet most important factor in transport and stability of the nucleic acid complex. Without charge neutralization and condensation of nucleic acids, they will never be able to enter the cell to effect therapy. Naked nucleic acids have a high negative charge density and large size, which will make it difficult to penetrate cellular membranes[40; 81]. In the case of larger nucleic acids like plasmid DNA, condensation is crucial for delivery. Polycationic packing agents have been proven to be effective in protection and delivery of small (about 20 base pair) siRNAs[40]. However a, 3 kb plasmid is much larger and will have a more complicated condensation interaction with a polycationic packaging peptide. Compaction of plasmid DNA as a key aspect to delivery is clearly related to the DLS results obtained in this study. H3K(+G)4b peptides ability to condense DNA into a neutrally charged particle was demonstrated by the presence of a larger complex compared to H3K4b. In addition, increase in the size of the plasmid-H3K(+G)4b complex was demonstrated over a 20 minute time period, indicating aggregation of charge neutralized complex. The H3K(+G)4b peptides reduced steric hindrance may be responsible for its better condensation of plasmid DNA. DLS results demonstrated larger sizes for H3K4b-siRNA complexes formed, which may indicate more

condensation and charge neutralization. Once condensed effectively, these complexes can undergo endocytosis to be taken in by the cell[40].

Exposure to a vast variety of biological media during the transport process of a packaged complex presents major hurdles in preventing degradation of siRNA or plasmid. This major difficulty must be overcome in order to keep the nucleic acid intact during transport. Nucleic acids by themselves, especially siRNA, prove to be very unstable during transport through biological serum due to the presence of nucleases[39; 40; 82]. Merkel conducted a series of assays that used PEI as a delivery agent for siRNA in which the lungs were targeted[82]. His experiments address critical aspects for protection of a nucleic acid during transport of a complex formed with a polycationic delivery agent. Survivability of the PEI-siRNA complex proved to be enhanced in two biological media, mucin and lung surfactant, which are representative of conditions found at mucus membranes and the environment in the lungs[82]. Exposure of nucleic acids to biological systems during therapeutic delivery, especially mucus membranes, will be an important barrier to consider during delivery. Upon incubation of PEI-siRNA complexes in mucin and lung surfactant, enhanced stability and survivability was observed for these complexes[82]. Other experiments have also shown direct protection of siRNA and plasmid DNA from degradation by nucleases. PEI in complex with siRNA has proven to protect against degradation by RNase A[40]. Assays have also been conducted looking at DNase I degradation of plasmid DNA in complex with nanoparticles and block-copolymer condensing agents[39; 45]. DNase I and RNase A experiments conducted by other labs are very similar to the ones we conducted on H3K4b and H3K(+G)4bs ability to

protect plasmid DNA[39; 40; 45]. Again H3K(+G)4b increased flexibility lends itself to better coating of plasmid, better protecting it. H3K4b was not completely effective at protection of plasmid and provided protection from DNase I only at an N/P of 4.8. The evidence provided here proves the utility of polycationic agents in increasing stability of siRNA in biological fluids.

Linked to protection is also diffusion of nucleic acid delivery complexes once they are introduced into a biological media. In order for delivery to be effective, the complex must localize nucleic acid release to a desired target region. Naked siRNA shows rapid dispersal and excretion upon introduction to serum[82]. Nucleic acids like siRNA are particularly prone to rapid accumulation in the liver and kidneys leading to excretion. Polycationic polymers like PEI have been shown to improve survivability and localization of delivered siRNA over that of naked siRNA[82]. Another experiment by Merkel observed localized targeting of the PEI-siRNA complex to the lungs[82]. Twenty-four hours after exposure of the PEI-siRNA complex, a significant amount of PEI-siRNA complex was still found in the lungs where it was initially introduced[82]. In this case, PEI proved to be effective at keeping the siRNA release localized to where it was targeted. Much of the survivability and localized targeting also has to deal with controlled release of the nucleic acids of interest. H3K4b and H3K(+G)4b both demonstrate effective binding to nucleic acid as demonstrated by the EtBr exclusion assays. This effective binding may provide localized delivery of nucleic acid in a biological environment. As demonstrated in DLS, complexes that are aggregated will be a combination of hundreds of nucleic acids and thousands of peptides. This will keep them packaged tightly together during transport. Another

advantage of H3K4b and H3K(+G)4b in localized delivery is that they can be modified during their synthesis. These modifications could include molecules and functional groups that are targeted to specific cells or tissues.

The biological significance of condensation, protection and localized delivery of a nucleic acid are critical factors when considering polycationic packaging vectors development.

5.3 Effect of a Peptide Complex Formed on Endosomal Release

Effective transport is only one major factor for effective delivery of a therapeutic nucleic acid. The other factor that influences nucleic acid delivery is once safely transported to its destination, it must be released to its target. The major route of entry for a nucleic acid-catoinic polymer complex is typically some form of endocytosis. There are a few major types of endocytosis, but all involve a similar principle[83]. An invagination of the cells plasma membrane occurs at the site of the particle taken in. This invagination then encompasses the particles and brings it inside the cell. The endosome containing the nucleic acid-polymer complex is then pinched off from the plasma membrane to be internalized into the cell and has many fates as illustrated in Figure 1.3. The one important fate of the endosome to mention here is fusion with a lysosome[84]. The lysosome contains many catalytic and degradative elements, which have a large impact on the contents of the endosome[84]. Delivery complexes can be extremely sensitive to the pH of the endosomal and lysosomal vesicles interior compartment. A typical endosome has a size of around a couple hundred nm in diameter. The average lysosome is about 500 nm in diameter[84]. The very small size of these vesicles means that addition of any kind of

ions will have a large impact on the internal environment. The vesicles sensitivity to pH change is the major proposed mechanisms for release of nucleic acid-polymer complexes into the cytosol.

At physiological pH the average charge of the H3K4b and H3K(+G)4b peptides is +24, which includes only the lysines because they are the only species protonated. The histidines are not protonated at physiological pH because it is a full pH unit above their pKa. Protons are pumped into the endosome from the cytosol by an ATPase enzyme[11]. The endosomal pH drops to about 5 when pinched off from the cell membrane. In this endosomal environment, it is estimated that up to 40% of amines can be protonated, in the case of PEI[36]. If all of the histidines were protonated on the H3K4b and H3K(+G)4b peptide, this would increase the charge to +72. This has a major effect on the charge of the very small internal space of the endosome. To counteract this increase in charge, the endosome takes in chloride ions and also water to balance the increase ion concentration[36]. This sudden influx of positive ions, negative ions and water causes swelling and destabilization of the endosome[36]. This in turn can lead to rupture of the endosome and release of its contents into the cytosol[36]. This mechanism of rupturing the endosome through a sudden influx of ions has been termed the proton sponge hypothesis, which has been mentioned briefly in the introduction (Figure 1.4). This theory has backing by theoretical models and experimental evidence[36; 85]. The primary role of the histidines in the H3K4b and H3K(+G)4b is release of the nucleic acid-peptide complex through this mechanism.

When considering the proton sponge mechanism it is important to note that the sharp changes in pH experienced by the complex during release from the

endosome. Release of the complex from the endosome is important to delivery of the nucleic acid to the cell, however, once inside the cytosol the nucleic acid has to also be released from the peptide. There has been a case where a delivery agent bound to the nucleic acid too tightly, and does not allow its release for effective therapy[82]. Tightness of complex formation is a sensitive balance between adequately protecting the nucleic acid during transport and allowing a weak enough binding interaction to allow dissociation of the agent from the nucleic acid. Once the nucleic acid-peptide complex dissociates, plasmid DNA can be taken up by the nucleus to effect gene delivery or in the case of siRNA for RNAi, released into the cytosol. The large change in pH experienced by the complex during endosomal rupture can be speculated to play an important role in release of the nucleic acid from the peptide. Inside the endosome, at pH 5, the histidines on the peptide will be highly positively charged and may facilitate in tighter binding of the peptide to the nucleic acid. However, once ruptured, the complex will experience a large increase of the pH to about 7.5, causing a significant loss in the positive charge of the peptide. Similar to endosomal rupture, this drastic deprotonation of the peptide due to change in the pH could cause a destabilization in the interaction between the peptide and nucleic acid. The destabilized complex then falls apart, releasing a large majority of the nucleic acid bound up in it.

One more factor that could be aiding in release of H3K4b and H3K(+G)4b complexes from the endosome has to do with the PII helical structure observed for this peptide. From the CD analysis of both H3K4b and H3K(+G)4b, in the absence of nucleic acid, both form a PII helix conformation. There are many antimicrobial

peptides in nature, which form an alpha helix that has been shown to disrupt and rupture the plasma membrane[86]. These antimicrobial peptides work to either disrupt the interaction between lipid molecules or create pores (or holes) in the cells membrane, which cause it to rupture[86]. Since H3K4B and H3K(+G)4b can form a PII structure, the helix they form could be causing a similar effect on the membrane of the endosome. PII helicies have demonstrated the ability to interact with cellular membranes[87; 88]. Excess peptide released from the complex or on its surface could be interacting with the plasma membrane, disrupting it and causing it to rupture and release the contents.

When developing a nucleic acid delivery system, release of the complex from the endosome and the delivery agent is important to successful therapy. Future designs of delivery agents will want to take advantage of the successful release of nucleic acids afforded by H3K4b and H3K(+G)4b.

5.4 Concluding Remarks on H3K4b and H3K(+G)4b Contributions to Effective Transfection

The transport and release of nucleic acids are the two crucial components needed for effective delivery. All characterization of H3K4b and H3K(+G)4b have been completed with those parameters in mind. For H3K(+G)4b, flexibility is key to protection and delivery of larger nucleic acids such as plasmid. H3K(+G)4bs ability to complex and better aggregate plasmid will be an important consideration for gene delivery. However, this same benefit for packaging plasmid DNA, makes it inefficient for delivery of siRNA. In the case of H3K4b, a stiffer, dendrimer-like complex is more efficient for packaging of siRNA. These same attributes of H3K4b

provide ineffective in delivery and protection of plasmid. A delicate balance between adequate protection and release of nucleic acids must occur for effective delivery. Uncovering these characteristics of H3K4b and H3K(+G)4b-nucleic acid complex formation are important in creation of future transfection agents. The model developed from our work is a starting point for future models detailing the chemical interactions and processes necessary for successful nucleic acid delivery. H3K4b's and H3K(+G)4b's current effectiveness as delivery agents will have broad implications for the way effective protection and release of nucleic acids are viewed.

Appendix 1: SafteyCoat Procedure and Matlab Code

1.1: SafteyCoat Procedure For Silicoizing Cuvettes:

- Reproducibility of initial fluorescence quenching experiments with an alexa 555 labeled linear HK peptide proved to be very difficult. Pipetting up and down once with a non-siliconized pipette in a cuvette containing a sample of peptide with DNA, demonstrated vast differences in the fluorescence signal from scan to scan (without any further addition of peptide). It was determined that peptide sticking to the surface and being dislodged upon mixing was the cause for the lack of reproducibility. Siliconization corrected the sticking problem.
- If using glassware that has been previously used or exposed to biomolecules or have used sigma coat previously on a surface, clean off the surface. Removing dirt or particulates can be accomplished by doing an acid wash (concentrated Nitric Acid), which sits for about 1-2 hours or so and followed by a rinse with copious amounts of water.
- Caution: Follow the safety instructions for sigma coat! Use in a well ventilated area and use gloves! Apply the sigma coat to the piece of glassware or just dip the glassware in some of the solution and let it sit for 15 minutes. Typically, with a cuvette just fill it with the solution and let it sit for the 15 minutes.
- Pour the solution out of the cuvette or remove the piece of glassware from the solution and let air-dry. This may take from 5 min to 30 min depending on

ventilation and the surface area being dried, ect. If the sigma coat does not completely dry, simply remove any excess with a Kimwipe.

- This treatment should leave a hazy coating on surface of the piece of glassware coated which can be removed by polishing with a Kimwipe. To polish the inside of the cuvette, use a small gauge needle, about 23, and wrap it with a Kimwipe. Caution: Be careful that the needle does not poke through the Kimwipe and scratch the glass!
- Finally wash the freshly coated surface using ethanol and then water repeatedly to ensure any residue is removed. The formation of many bubbles may occur when using a freshly coated cuvette for the first time. Repeated use and rinsing with water should eventually reduce the amount of bubbles.
- The siliconization of cuvettes that are heavily used lasts about 3 months. If a cuvette has not been used for 4 months, the Siliconization procedure should be repeated. When the siliconization starts to wear out, fluid stays stuck to the cuvette during drying with an air can. This process should be repeated at those time intervals to keep results reproducible.

1.2: Matlab Analysis and Program For CD Spectral Decomposition:

- Import HK_Minimization.mat into matlab. This contains the matrices for the buffer-subtracted experimental data and reference spectra for the peptide and siRNA. The only thing that needs to be done next is to run error_min_search_stoich.m. This program calls on functions that create all reference spectra and runs the minimization to find the spectra of the peptide-

nucleic acid complex and the binding stoichiometry. The two functions that need to be present in your workspace to correctly operate the program are spec_mat_stoich.m and err_vs_exp_stoich.m. Here is the code for the programs needed to run the minimization:

error_min_search_stoich.m:

```
% Script for obtaining optimal reference spectrum and stoichiometry
%Example
%run = @(z)((15-z)^2);
%[x,fval] = fminsearch(run,[1])

% Assumes x5000uM_HK_buffed, HK_titr_buffed, and siRNA_Avrgdsig are already
in the workspace

ref_spect_complex = zeros(size(x5000uM_HK_buffed));
HK_reference_mat = [(x5000uM_HK_buffed)/(5)  ref_spect_complex
(siRNA_Avrgdsig)/(0.45)]; %HK originally divided by (5).  siRNA originally divided
by (0.45) to get 1uM of peptide and siRNA.


%numconcs is the total number of concentrations of siRNA relative to a
%single peptide concentrations (m x n, wavelength x concentations)
siRNA_HKtitr_Conc_values = [0045' , 0089' , 0130' , 0180' , 0220' , 0260' , 0310' ,
0350' , 0390' , 0430' , 0640' , 0830' , 1030' , 1210' , 1620' , 2030' , 2440' , 2850' , 3260'
, 3670' , 4080' , 4490']';
siRNA_conc_vec = siRNA_HKtitr_Conc_values;
pep_per_rna = 1;
pept_init = 5000;

numconcs = size(siRNA_conc_vec,1);

stoich_and_spec = [pep_per_rna ; ref_spect_complex];

err_norm = @(stoich_and_spec) err_vs_exp_stoich(stoich_and_spec, HK_titr_buffed,
HK_reference_mat,pept_init,siRNA_conc_vec);
fminopts = optimset('MaxFunEvals',1000000,'MaxIter',1000000,'TolFun',1e-
10,'FunValCheck','on','TolX',1e-6);
%[stoich_and_spec , fval, exitflag,output] =
fminsearch(err_norm,stoich_and_spec,fminopts)
[stoich_and_spec , fval, exitflag, output] =
```

```
fminunc(err_norm,stoich_and_spec,fminopts)

pep_per_rna = stoich_and_spec(1)
ref_spect_complex = stoich_and_spec(2:end);
new_HK_reference_mat = HK_reference_mat;
new_HK_reference_mat(:,2)=ref_spect_complex;
species_mat = spec_mat_stoich(pept_init, pep_per_rna, siRNA_conc_vec);

figure
plot(ref_spect_complex)
%plot(HK_titr_buffed)
figure
plot(new_HK_reference_mat*species_mat)
figure
plot(HK_titr_buffed - new_HK_reference_mat*species_mat)
```

spec_mat_stoich.m:

```
function [ species_matrix ] = spec_mat_stoich( pept_init, pep_per_rna,
siRNA_conc_vec  )
%UNTITLED2 Summary of this function goes here
%   Detailed explanation goes here
%
%numconcs is the total number of concentrations of siRNA relative to a
%single peptide concentrations (m x n, wavelength x concentations)
%

%siRNA_HKtitr_Conc_values = [0045' , 0089' , 0130' , 0180' , 0220' , 0260' , 0310' ,
0350' , 0390' , 0430' , 0640' , 0830' , 1030' , 1210' , 1620' , 2030' , 2440' , 2850' , 3260'
, 3670' , 4080' , 4490']'
%siRNA_conc_vec = siRNA_HKtitr_Conc_values;
%pep_per_rna = 5;
%pept_init = 5000;

%numconcs = size(siRNA_conc_vec,1)
%
% pep_cons requires two approximated inputs: the initial peptide concentration
spectrum and the stoicheometry of peptide to rna
% defined above in the function as pep_per_rna (i.e. 4:1 = 4)
%
pep_concs = max(pept_init - pep_per_rna*siRNA_conc_vec,0);
comp_concs = min(pep_per_rna*siRNA_conc_vec,pept_init);
siRNA_concs = max(0, siRNA_conc_vec - pept_init/pep_per_rna);
species_matrix = [pep_concs' ; comp_concs' ; siRNA_concs']/(1000);
end
```

err_vs_exp_stoich.m:

```
%We first have to define a function that encompasses all of our matrix
%knowns and variables.  We refer back to this function below in err_norm.
function err_norm = err_vs_exp_stoich(stoich_and_spec, HK_titr_buffed,
HK_reference_mat,pept_init,siRNA_conc_vec)

%First we have to define the variable vector we are going to minimize.  We
%accompish this by defining a new vector, new_HK_reference_mat, that will be our
%new reference spectra that will include the new mimimalize values for the
%theoretical combined siRNA and HK ref specs.  We then defined the row that
%our new values will replace and called it ref_spect_complex, which we definined in
%the function above.  This set up is what needs to be done inorder to
%minimize one vector among a know set of data.
new_HK_reference_mat = HK_reference_mat;

pep_per_rna = stoich_and_spec(1);
ref_spect_complex = stoich_and_spec(2:end);
new_HK_reference_mat(:,2)=ref_spect_complex;
species_mat = spec_mat_stoich(pept_init, pep_per_rna, siRNA_conc_vec);

%Below is the definition for the error of our combined matricies.
err_mat = (HK_titr_buffed-(new_HK_reference_mat*species_mat));

%err_norm = norm(err_mat*(transpose(err_mat)),'fro');
err_norm = sumsqr(err_mat);
```

Appendix 2: Topologically Constrained Minicircles and Linear DNA constructs as Tools for Assaying the TATA Box Binding Protein-DNA Interaction

Overview:

The TATA-Box binding protein (TBP) is a unique binding protein that interacts with the minor groove of DNA. Much has been discovered about this protein because of its key role in transcriptional initiation. Because this proteins binding interaction relies upon a unique structural topology of DNA, uncovering its mechanism of binding can lead to a better understanding of protein-DNA interactions. This is a brief introduction on the TBP research conducted and a synopsis of the results gathered for assessment of TBP's binding mechanism using a series of different DNA constructs.

Introduction:

TBP is roughly 180 amino acids long with a size of about 38kDa[89; 90]. The protein is comprised of two very similar, semi-symmetric domains as seen in Figure 1[89; 90]. It's characterized by a series of beta-sheets which form the hydrophobic interface between the protein and the DNA binding region[89]. As a part of this interface there are a pair of phenylalanine residues that partially intercalate between the first two and last two bases of the TATA repeat[89]. The series of beta-sheets is held in placed by a backbone of four alpha helices[89]. The protein is a rigid C-shaped clamp that locks onto the DNA. This is ideal for its functionality in binding and

bending DNA because it can introduce structural deformation in the TATA box essential for transcription initiation. A key factor in TBPs binding is the fact that it binds in the minor of DNA and not the major groove[89; 91]. This is a curious and

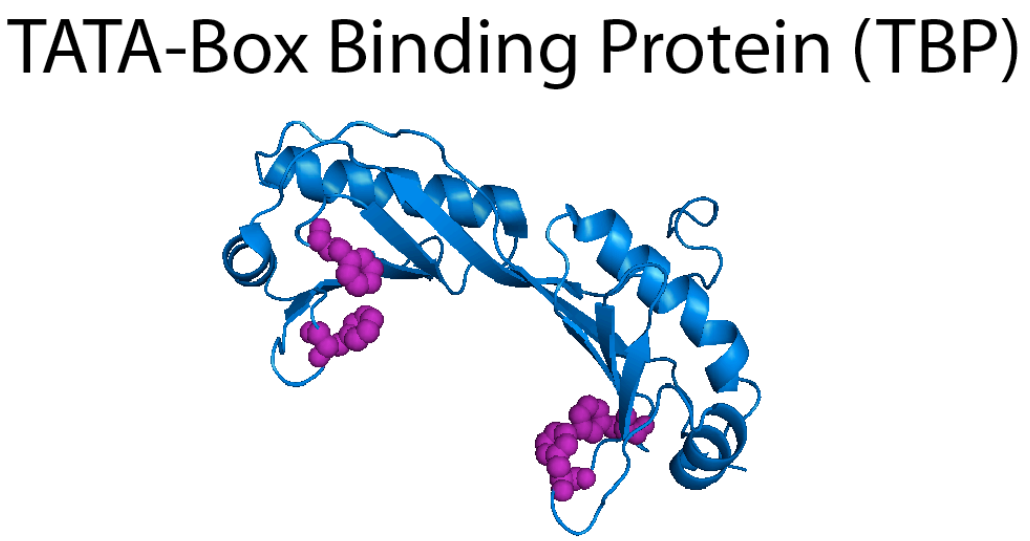


Figure 1. A pymol cartoon illustration of TBP showing its secondary structural motifs of alpha helices and anti-parallel beta sheets. The four purple space-filling residues are the phenylalanines, which partially intercalate between the first and last two base pairs of the TATA-box.

somewhat puzzling aspect of TBPs binding due to the fact that most of the well-characterized DNA binding proteins, such as the lac repressor, bind in the major groove because of the more extensive hydrogen bonding network found there[92; 93]. This is a strong indication that more than just site recognition is guiding TBP binding[94]. This leads us to inquire as to what are the key characteristics of the DNA, which contribute to TBP binding. In combination with the key structural characteristics of TBP, a multitude of unique structural characteristics of the TATA-Box DNA sequence play a major role in TBP binding and transcription initiation. TBP binds to an eight base pair region of DNA characterized by a TATA repeat, followed by WAWR sequence[95]. This TATA repeat is important due to the fact it is

easier to deform a TA sequence because of the steric interaction of the methyl group on the thymine base pairs[96]. Upon binding to the minor groove region of the TATA-Box sequence, TBP induces severe structural deformations of the DNA, which start with an 80° bend of the DNA towards the major groove as seen in Figure 2[89; 91].

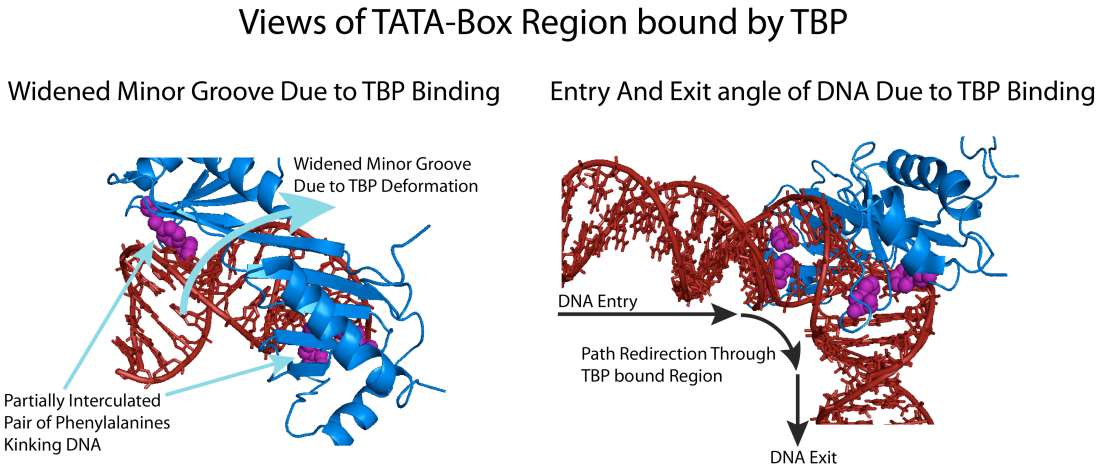


Figure 2. Two different views of TBP bound TATA region illustrate bending and heavy deformation of DNA upon binding. The 80° bending of the DNA can clearly be seen here. The widening of the minor groove is shown which stems from TBP's hydrophobic beta sheet regions interaction. The phenylalanine's, shown in purple, kink the DNA and cause a redirection of the path of DNA through the TBP bound portion. This has a significant contribution to the topology of the DNA, causing it to be slightly writhed upon TBP binding.

Much of this bending comes from the previously mentioned pair of phenylalanine residues partially intercalating between the first two and last two base pairs of the TATA-box region. They also kink the DNA at the ends of the TATA-box , causing the entry and exit points of the of the DNA through the binding site to be very different. This 80° bend is accompanied by a significant untwisting of the DNA by 105° [89; 91]. This leaves the minor groove of the DNA opened up as TBP settles into it[89]. This opening and deformation of the DNA is key to the initiation of the transcription bubble during TBP's heavy deformation of the TATA region, which aids in a stable formation of the PIC during transcription[97].

The overall specificity of TBP recognition of the TATA sequence does not come from typical major groove site recognition or hydrogen bonding recognition, but from the inherent structural aspects afforded by the TATA-box region[91; 98]. The TATA DNA regions great inherent flexibility leads to a more easily deformable DNA, which is key in TBP's recognition and binding to this sequence[91]. This is a strong indication, as pointed out by some very recent papers, that TBP doesn't necessarily have to bind to the exact TATAWAWR region of DNA but a sequence containing TATA-like elements[95]. In general, there is a push towards observing the importance of different structural features of a DNA sequences and the role it plays in initiator recognition[99]. Regardless, the necessary structural features of the TATA DNA play the major role in TBP binding and recognition and *not* the exact sequence. This makes clear that the path to understanding the TBP TATA region binding interaction will be in assessing aspects of TBP binding to different DNA conformational motifs.

A multitude of studies have focused on altering the TATA-box region and observing its subsequent effect on TBP binding and TATA DNA bending. Starr initiated work illustrating that content of the 8 base pair TATA region was critical in bending angle of this region[98]. During this work it was also revealed that the TATA region is inherently bent in the opposite direction of the TBP induced bend[98]. Parvin was among the first to use a circular DNA construct to add to these observations and show a TATA-box sequence pre-bent towards the major groove has preferential TBP binding[100]. These were some of the first indications that the TATA-box regions enhanced flexibility and deformability is critical in TBP binding. This work set up

research conducted later by Davis and Kahn that was comprehensive in proving this. The two major observation made during their research was that the TATA-box region was extremely flexible and that TBP binding induced unexpected topological changes. To further probe the unique dependence on DNA topology for TBP binding, Jung and Kahn used a DNA minicircle construct. Using phenylalanine deleted TBP mutants, a series of experiments were conducted to observe the effect of the phenylalanines on binding to linear and minicircle DNA. Their findings proved that the phenylalanine mutant TBP binds preferentially not only to minicircle as opposed to linear DNA, but to a slightly negatively supercoiled minicircle DNA[101]. The speculated reason for this lies in the phenylalanines partial intercalation in the TATA-box sequence. During the TBP binding process, the phenylalanines play a major role in bending the DNA and kinking it at the sites of intercalation[89]. The result is a very different DNA strand entry and exit angle through the TATA-box region yielding a writhed region of DNA. The only caveat with this writhed segment of DNA is the fact that the entry and exit strands will be attached to a chromosome and

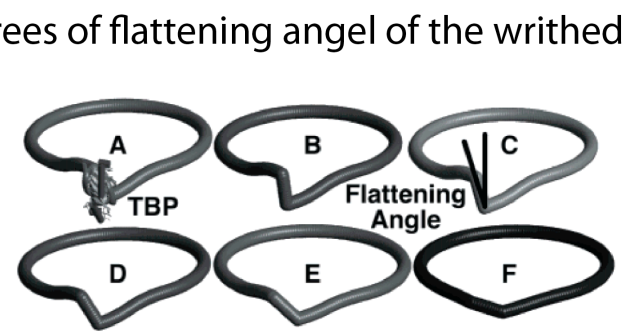


Figure 3. Flattening of the writhed TATA region of DNA, caused by the partial intercalation of a pair of phenylalanines in the start and end of the sequence. A-F, A being most writhed with no phenylalanine retraction and F being completely retracted leading to a flattened DNA, show a series of subsequent flattening angles accomplished through step-wise removal of the phenylalanines from between the bases[102].

locked into place in-vivo, creating an unfavorable writhed DNA topology in a very short piece of DNA. One plausible description proposed by Kahn in 2000 for how TBP deals with this unfavorable DNA topology is through retraction of these phenylalanines from between the DNA in the TATA region[102] (Figure 3). Upon subsequently different degrees of retraction the TBP could allow for flattening of the kinked region of DNA allowing the phenylalanines to accommodate for changes in the DNA topology[102]. These major discoveries of the TBP binding process gives our current proposed view on

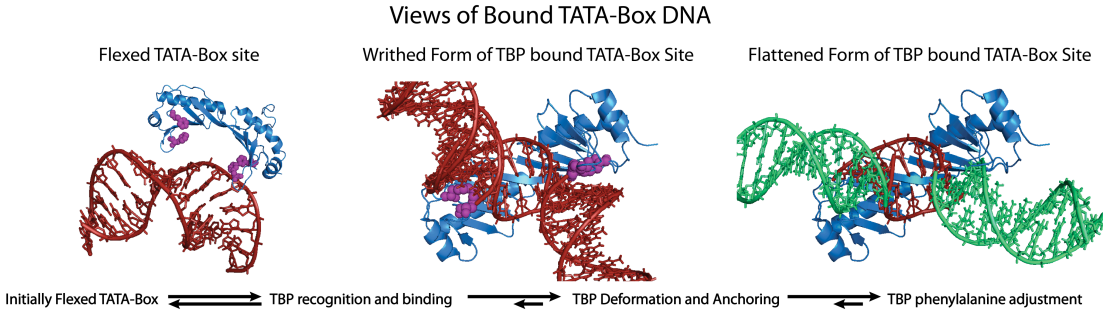


Figure 4. A Proposed mechanism for TBP binding to the TATA-Box region of DNA.

the TBP binding mechanism. The proposed mechanism is outlined as follows (Figure 4):

Initial DNA state:

- TBP recognizes a transiently flexible state of the TATA-Box region. This transiently flexed state is enhanced by its bent topology around a histone core[91; 100].
- TBP recognizes this state and binds to the slightly open DNA. As mentioned this open DNA comes from the bent conformation. The TATA repeat is ideal for deformation[103].

- First intermediate of DNA deformation upon TBP binding:

- The TBP loosely binds to the TATA region upon recognition of the transiently flexed DNA state. No deformation of the DNA has occurred yet[91].
- Phenylalanines of TBP partially intercalate into the first two and last two base pairs of the TATA sequence, kinking and making the two DNA ends 90 degrees perpendicular to one another[103].

- Second intermediate of DNA deformation upon TBP binding:
  - TBP's significant deformation and bending of the DNA is the major driving force for its tight binding. There is a lot of energy put into the system in order to bend the DNA, therefore it's not favorable for the reaction to go backward[91; 104].
  - Upon binding, the hydrophobic backbone of the TBP recognizes the minor groove of the DNA. This backbone interaction is seen (at every AA and nucleotide interaction) over the 8 base pairs of the binding site, unwinding the DNA in the minor groove. This in-turn widens (to the extent of flattening) the minor groove. Widening the DNA and breaking base pair interactions requires a lot of energy[103; 104].
  - Completion of binding and bending of the DNA takes place as TBP settles into place. Final exclusion of water from between the hydrophobic interface may take place[103; 104].
  - Because there are A-tracts in our construct, it takes longer for the TBP to bind to our minicircle. The bend induced by the A-tracts locks the TATA-box region into place, therefore, the bend induced by TBP is

going to take energy. Once TBP binds it is difficult to break the interaction with minicircle DNA because of the energy required.

- This backbone interaction also bends the DNA 80 degrees towards the Major groove[104].

- Final TBP Bound DNA State:
  - TBP uses the phenyalanines to control DNA topology while bound. This relieves strain that may be trapped in the DNA upon binding of the TBP and TFIID complex[102].

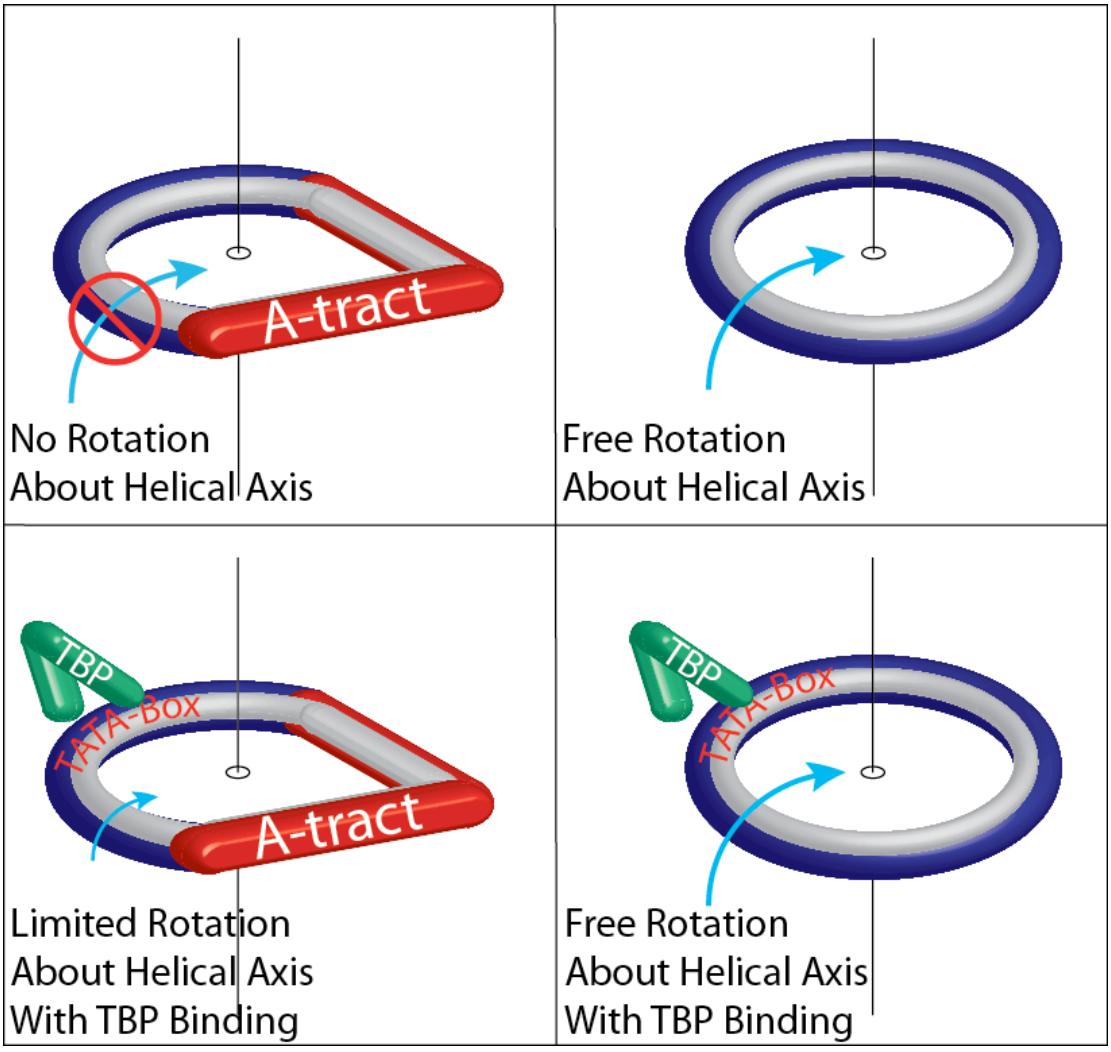
All of the work conducted currently on TBP has given a relatively comprehensive analysis of it's binding mechanism. This leaves us with a few, very focused questioned left to answer about TBP's binding. Many experiments to date have dealt with altering the TATA-box sequence and observing the effects on TBP binding, utilization of some sort of cyclization assay or using short linear fragments of DNA to assess binding[91; 105; 106]. Very little work has been accomplished utilizing a circular DNA construct to observe TBP's binding since Parvin[100]. Jung (2005) explored this interaction to a limited extent, however, mainly focusing on how TBP phenylalanine deletion mutants bind to a TATA-box containing minicircle. We set out to answer three major questions about TBP binding. First, since there was a difference in binding of TBP between linear and minicircle DNA, what aspects of bent DNA topology promote TBP binding? Secondly, if we provide TBP with a very constrained (below one persistence length of DNA) minicircle DNA construct, how will this affect binding? Lastly, how exactly does phenylalanine retraction fit into the overall binding mechanism?

Inspired by work from Kahn, Parvin, Davis and Jung, we set out to further probe the interaction of TBP using different DNA constructs. By exploiting TBP's sensitivity of binding to different DNA conformations we can probe this interaction to a greater extent. We employed the use of a tightly constrained 106bp DNA minicircle with and without A-tracts, a 106bp linear DNA construct and a short 25bp DNA construct containing the TATA-box region. TBP's unique binding mechanism can be assessed by observing its binding to these different DNA molecules.

Results: Study of TBP's bindings interaction with a 106 base pair A-tract minicircle, no A-Tract containing minicircle, 106 base pair linear DNA and 25 base pair linear DNA:

The Rationale for using a 106 base pair DNA A-tract minicircle and no A-tract minicircle to assess TBP binding:

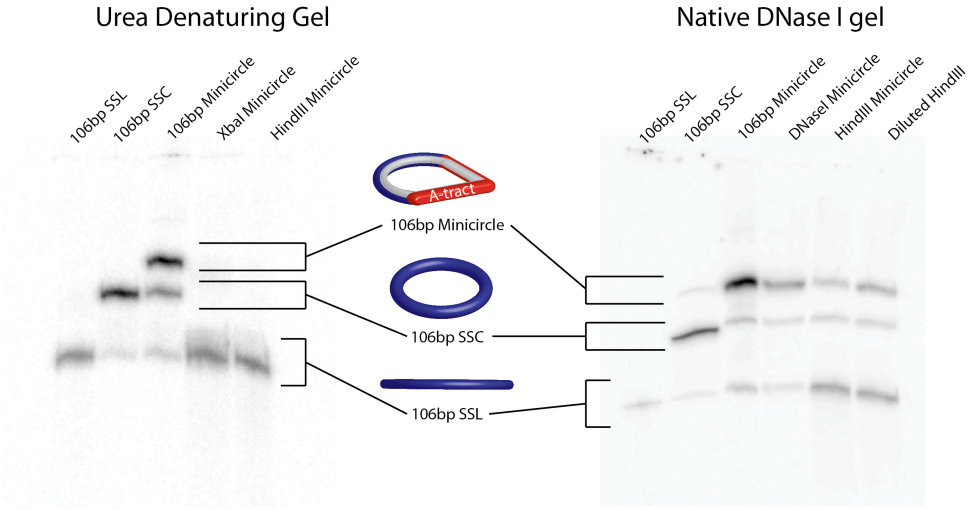
The 106 base pair DNA A-tract minicircle (106 A-tM) was chosen to assess the TBP binding mechanism because of its key topological characteristics. The 106 A-tM is below one persistence length of DNA, which makes it highly strained and the smallest minicircle used to probe the TBP binding interaction. In using such a small minicircle, we are emulating a topologically strained environment reminiscent of DNA bound up in chromatin[107]. The A-tracts in this minicircle lock the TATA-box region of DNA into place half a helical turn out of phase with the A-tract. Once TBP



binds, it will be more difficult for it to open up and bend the DNA. However, once TBP deforms the DNA and fully binds, it will be very difficult to remove. The presentation of the TATA-Box region in the 106 A-tM is ideal for TBP binding.

The idea behind the no A-tract minicircle 106 NA-tM is that there will be free rotation about the DNA helical axis. During TBP's binding process, it will be easier to open up and bend the DNA because the ends of the DNA will not be locked like in the 106 A-tM. The overall result would be better binding of the TBP but not as long lasting. Since there was less energy put into opening and reorienting the 106 NA-tM, then there will be less of a driving force for TBP to stay attached.

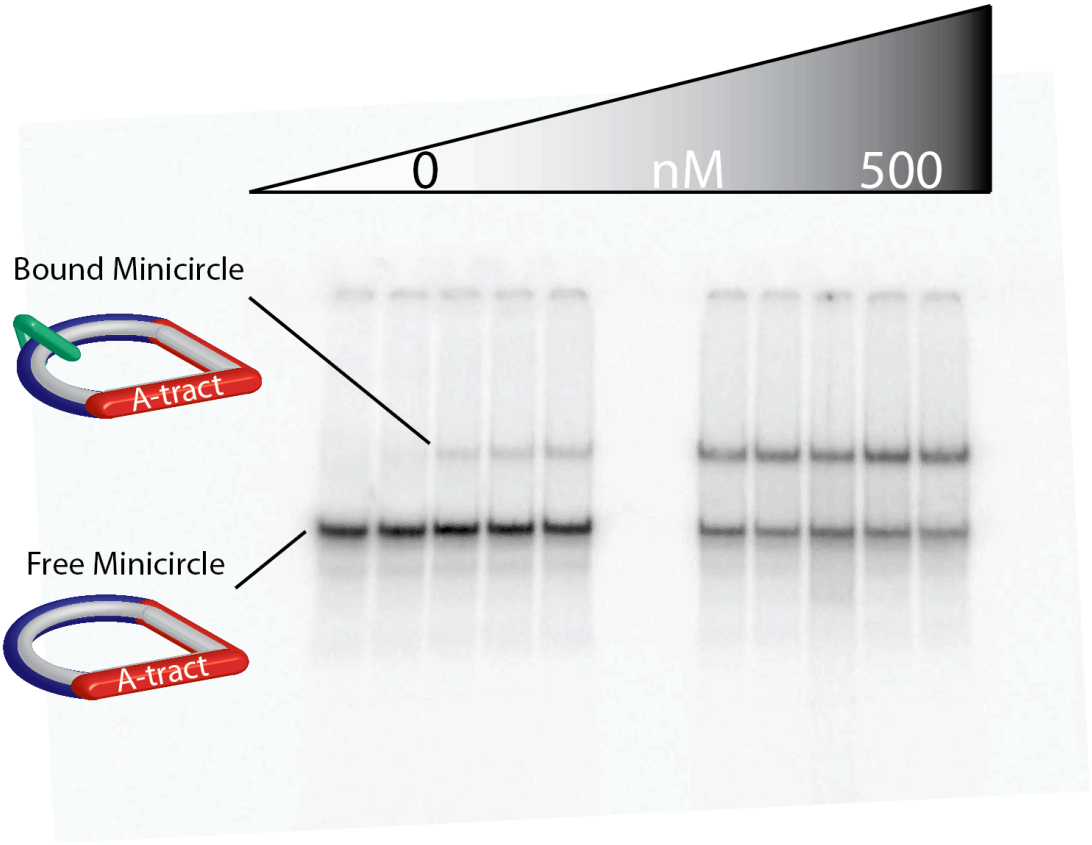
Creation of the106 A-tM:



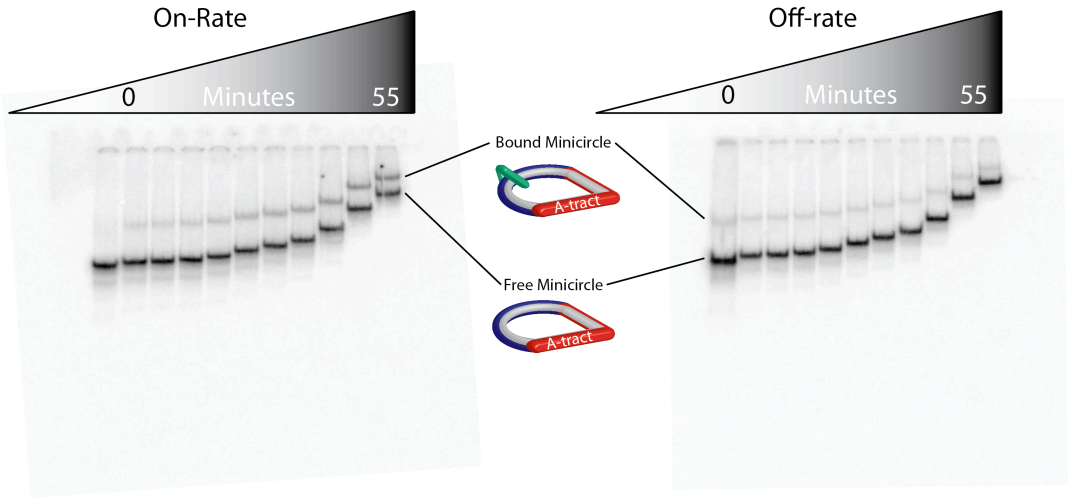
A series of restriction digestion reactions were conducted on the 106 A-tM to confirm its creation. Hind III, Xba I and DNase I digests were performed on the construct. After digestion, a urea denaturing gel and native gel were run. A trend in DNA band movement, in which they move progressively slower, is seen during each ligation reaction in the creation process. The progression in band retardation goes from 106 base pair single stranded linear fragment (106 SSL) to the 106 base pair single stranded product (106 SSC) and finally a shift to the 106 A-tM. Our 106 SSL runs fastest on the gel and is used as the initial reference point for our shift progression. Hind III and Xba I digestions of the 106 A-tM both show a return of the 106 SSL band on a denaturing gel. The control lanes for the 106 SSC and the 106 A-tM both show some residual 106 SSL and 106 SSC bands, respectively. This is most likely due to either incomplete ligation of the 106 SSC or nicking of the 106 A-tM. A Hind III digestion at two different concentrations and a DNase I reaction were

conducted and run on a native gel this time. This time we see a distribution of the three species in each of the digestion reactions. The goal we successfully accomplished was to ensure that a small amount of cleavage enzyme could be used on a native gel to reconstitute all of the products used to create the A-tM.

106 A-tM Electrophoretic Mobility Shift Assay (EMSA):

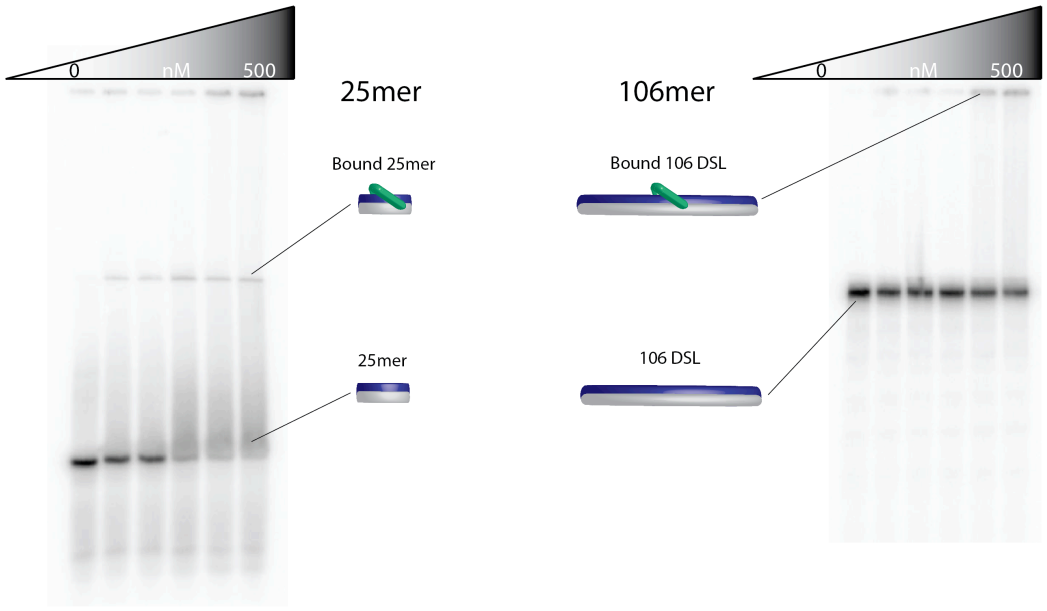


TBP bound 106 A-tM dissociation constant (Kd) EMSA: TBP binds relatively tightly to 106 A-tM. TBP binding saturates between 20-50nM TBP, which suggests a Kd around 30nM.



On and Off-Rates for TBP bound to the 106 A-tM. For the On-rate, a time course of TBP bound to 106 At-M was conducted from 0-55 minutes to assess the kinetics of this binding interaction. The On-rate for TBP shows a rapid binding that saturates within the first 2-5 minutes. This suggests that the initial part of the binding mechanism of TBP is fast . During the Off-rate, TBP was allowed to incubate with the 106 A-tM then cold competitor DNA was added over a course of 0-55 minutes to observe the disappearance of the bound population. The Off-rate for TBP binding illustrates almost no decrease in the bound population once formed. This analysis proves that once TBP is bound to our construct, it's difficult to become unbound.

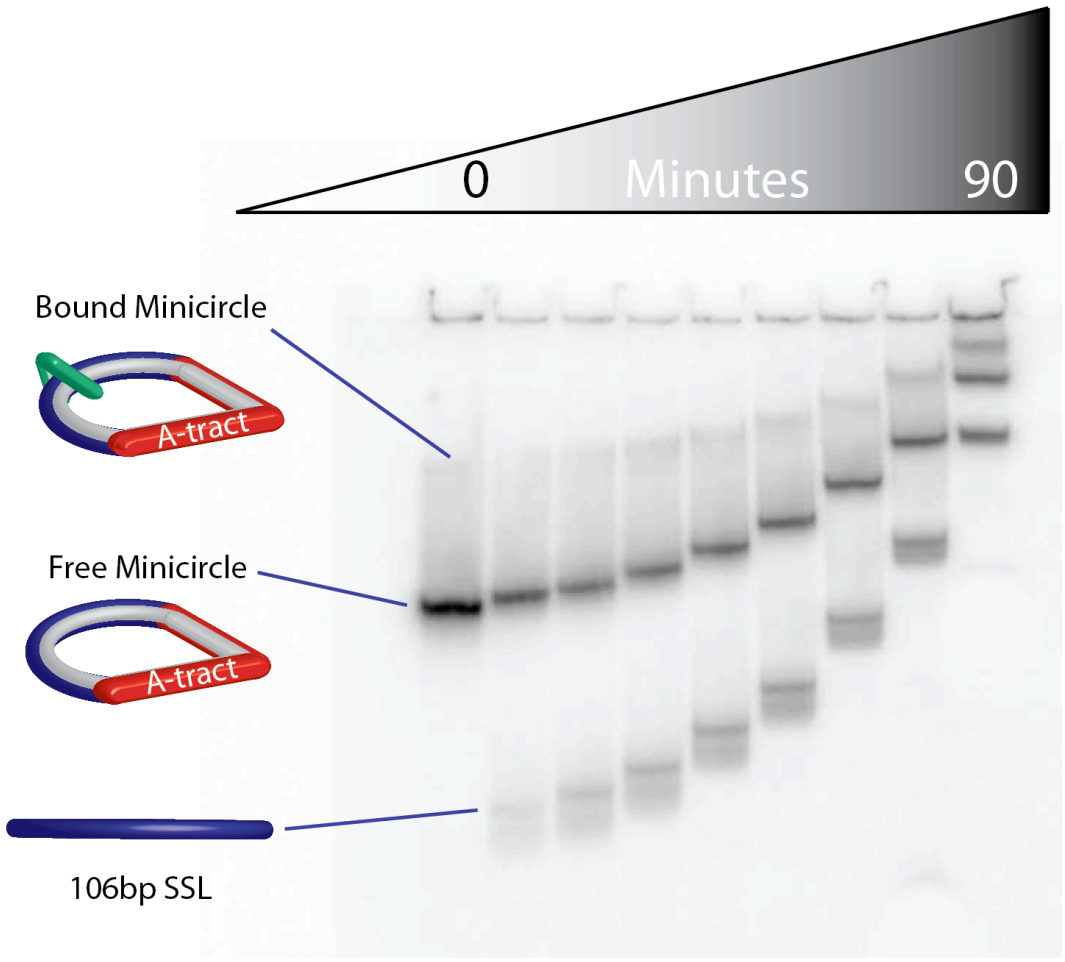
<u>106 base pair (106 DSL) and 25 base pair Linear double-stranded linear (106 DSL) EMSA:</u>



Many attempts to obtain a TBP bound 106 DSL population proved to be unsuccessful. Most of what was observed is a population of DNA molecules, at very high TBP concentration that is trapped in the gel well. This could be due to molecules that get hooked together once bound by TBP due to the A-tract region in the DNA. The reason that TBP may have difficulty binding to 106 DSL could be that the ends of the DNA are not attached to anything in the 106 DSL, which is not an accurate representation of the *in-vivo* environment. TBP would normally be bound to a bent piece of DNA constrained by chromatin[97]. TBP may have a more difficult time opening up the DNA and binding if the arms of the DNA are free in solution.

TBP was added to a short 25mer containing the TATA-box binding region to see if the protein would bind to a short piece of linear DNA. At high concentrations of TBP (150-500nM) a bound population was seen to emerge. The evidence provided for TBP's preferential binding to a short piece of linear DNA compared to a longer one indicates that the length of the DNA free ends may have an impact on TBP binding.

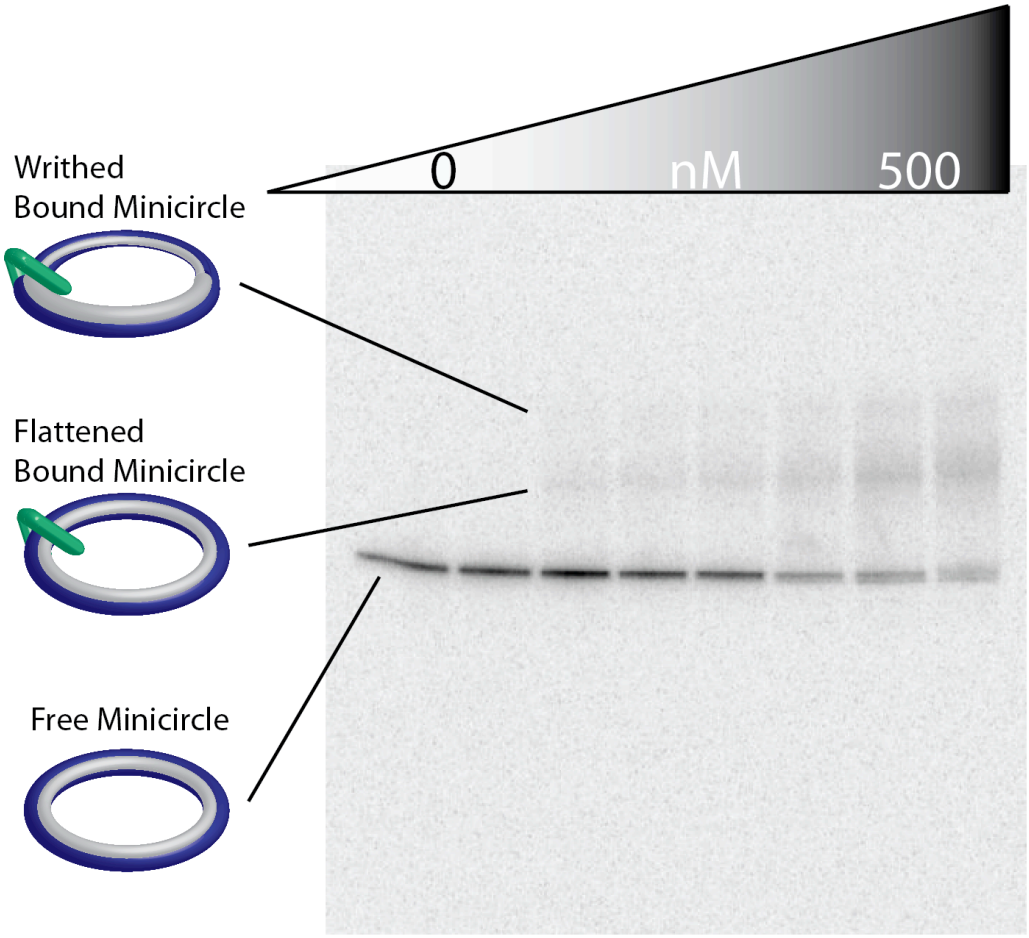
TBP bound 106 A-tM cleavage assays:



Cleavage of a 106 A-tM was conducted with TBP bound. Isolating a TBP bound 106 DSL DNA population had proven to be relatively unsuccessful. In an attempt to capture a TBP bound linear population, TBP was allowed to incubate with the 106 A-tM and then Hind III was added to the reaction for 0-90 minutes. If TPB was still bound upon cleavage, a population of bound and free 106 DSL would be seen. During the cleavage assay, a free species of 106 DSL was seen, but there was no

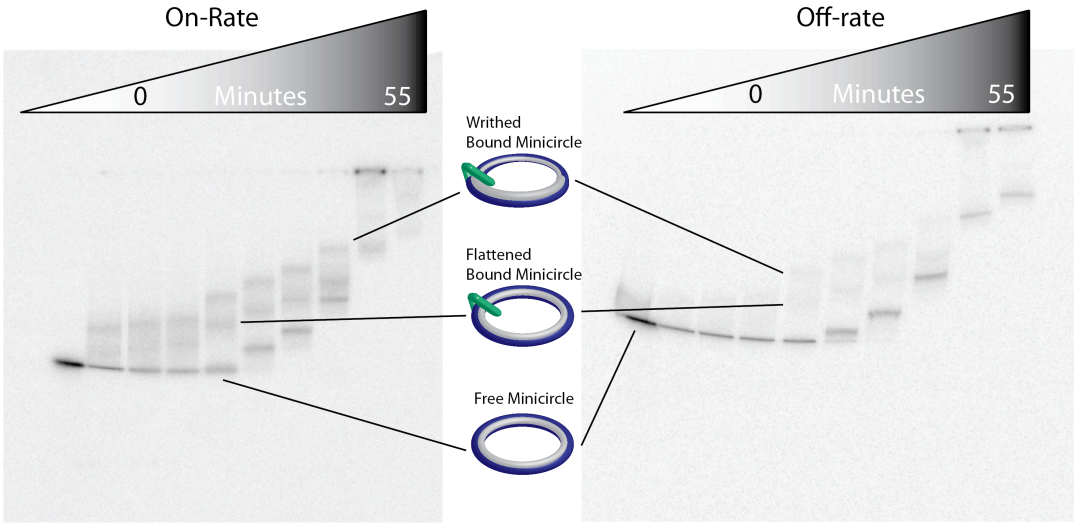
evidence for a bound population. The lack of appearance of a TBP bound linear 106 DSL DNA is another strong indication for the inability of a longer piece of linear DNA to provide a favorable conformation for TBP binding.

106 base pair No A-tract Minicircle (NA-tM):



An EMSA was conducted to determine a dissociation constant (Kd) for TBP bound 106 NA-tM EMSA. TBP binding appears to be weaker to the 106 NA-tM than the 106 A-tM. There is an appearance of a dual bound population, which could be

due to the intercalated phenylalanines that are accommodating for differences in DNA topology. The two TBP bound DNA populations being observed could be of a writhed and flattened minicircle conformation, which was not observed before during binding to the 106 A-tM.



An On and Off-Rate assessment was conducted for TBP bounding to the 106 NA-tM. Evidence for two bound populations of TBP is seen on both of these gel shifts. The On-rate gel shows a rapidly appearing bound population of 106 NA-tM. Binding seems to occur even faster than in the 106 A-tM on rate experiment indicating faster initial binding of the complex. Again for the Off-rate experiment, TBP was allowed to incubate with the 106 A-tM. Subsequently, cold competitor DNA was added over a course of 0-55 minutes to observe the disappearance of the bound population. The Off-rate gel shows another dual bound population that begins to dissipate between 20-30 minutes of incubation with the competitor DNA. This is a strong indication of rapid but weaker binding of TBP to 106 NA-tM.

Conclusions:

- A topologically constrained minicircle was successfully created with a TATA box region to which TBP bound to.
- TBP binds preferentially to the topologically constrained minicircle compared to linear DNA. The differences between A-tract and linear DNA demonstrate the importance of topology for TBP binding to the TATA box.
- Once TBP is bound to the TATA box region it is hard to remove through competition. When the energy is put into deform the DNA and bend it, TBP is firmly attached.
- TBP's binding to the No A-tract minicircle supports the idea of an initial fast intermediate step in binding. There is also a subsequent slower intermediate in opening the DNA, which is important for anchoring the protein to the DNA.
- These constructs present the basis for a topological analysis for a protein binding mechanism.

Bibliography

1. Lewontin, R. C. (1970). The Units of Selection. *Annual Review of Ecology and Systematics* **1**, 1-18.

2. Travers, A. A., Muskhelishvili, G. & Thompson, J. M. (2012). DNA information: from digital code to analogue structure. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **370**, 2960-86.

3. Hood, L. & Galas, D. (2003). The digital code of DNA. *Nature* **421**, 444-8.

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

5. Cox, J. P. (2001). Long-term data storage in DNA. *Trends in biotechnology* **19**, 247-50.

6. Watson, J. D. & Crick, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-8.

7. Crick, F. (1970). Central dogma of molecular biology. *Nature* **227**, 561-3.

8. Sheridan, C. (2011). Gene therapy finds its niche. *Nature biotechnology* **29**, 121-8.

9. Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J. L., Bousso, P., Deist, F. L. & Fischer, A. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669-72.

10. Hacein-Bey-Abina, S., Garrigue, A., Wang, G. P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B. H., Wintergerst, U., Velez, M. C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F. D., Fischer, A. & Cavazzana-Calvo, M. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *The Journal of clinical investigation* **118**, 3132-42.

11. Lavigne, M. D. & Gorecki, D. C. (2006). Emerging vectors and targeting methods for nonviral gene therapy. *Expert opinion on emerging drugs* **11**, 541-57.

12. Friedmann, T. & Roblin, R. (1972). Gene therapy for human genetic disease? *Science* **175**, 949-55.

13. Hahne, S. J., Charlett, A., Purcell, B., Samuelsson, S., Camaroni, I., Ehrhard, I., Heuberger, S., Santamaria, M. & Stuart, J. M. (2006). Effectiveness of antibiotics given before admission in reducing mortality from meningococcal disease: systematic review. *BMJ* **332**, 1299-303.

14. Garrod, L. P. (1960). The relative antibacterial activity of four penicillins. *British medical journal* **2**, 1695-6.

15. Neu, H. C. (1992). The crisis in antibiotic resistance. *Science* **257**, 1064-73.

16. McBride, J. L., Boudreau, R. L., Harper, S. Q., Staber, P. D., Monteys, A. M., Martins, I., Gilmore, B. L., Burstein, H., Peluso, R. W., Polisky, B., Carter, B. J. & Davidson, B. L. (2008). Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 5868-73.

17. Shilling, V., Jenkins, V., Morris, R., Deutsch, G. & Bloomfield, D. (2005). The effects of adjuvant chemotherapy on cognition in women with breast cancer--preliminary results of an observational longitudinal study. *Breast* **14**, 142-50.

18. Bonadio, J., Smiley, E., Patil, P. & Goldstein, S. (1999). Localized, direct plasmid gene delivery in vivo: prolonged therapy results in reproducible tissue regeneration. *Nature medicine* **5**, 753-9.

19. Kay, M. A., Glorioso, J. C. & Naldini, L. (2001). Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nature medicine* **7**, 33-40.

20. Mendez-Ardoy, A., Urbiola, K., Aranda, C., Ortiz-Mellet, C., Garcia-Fernandez, J. M. & de Ilarduya, C. T. (2011). Polycationic amphiphilic cyclodextrin-based nanoparticles for therapeutic gene delivery. *Nanomedicine* **6**, 1697-707.

21. Xiang, S., Fruehauf, J. & Li, C. J. (2006). Short hairpin RNA-expressing bacteria elicit RNA interference in mammals. *Nature biotechnology* **24**, 697-702.

22. Burnett, J. C., Rossi, J. J. & Tiemann, K. (2011). Current progress of siRNA/shRNA therapeutics in clinical trials. *Biotechnology journal* **6**, 1130-46.

23. Tiemann, K. & Rossi, J. J. (2009). RNAi-based therapeutics-current status, challenges and prospects. *EMBO molecular medicine* **1**, 142-51.

24. Padmanabhan, J., Levy, M., Dickson, D. W. & Potter, H. (2006). Alpha1-antichymotrypsin, an inflammatory protein overexpressed in Alzheimer's disease brain, induces tau phosphorylation in neurons. *Brain : a journal of neurology* **129**, 3020-34.

25. Burgess, A., Huang, Y., Querbes, W., Sah, D. W. & Hynynen, K. (2012). Focused ultrasound for targeted delivery of siRNA and efficient knockdown of Htt expression. *Journal of controlled release : official journal of the Controlled Release Society* **163**, 125-9.

26. Grayson, A. C., Doody, A. M. & Putnam, D. (2006). Biophysical and structural characterization of polyethylenimine-mediated siRNA delivery in vitro. *Pharmaceutical research* **23**, 1868-76.

27. Li, S. & Huang, L. (2000). Nonviral gene therapy: promises and challenges. *Gene therapy* **7**, 31-4.

28. Bitounis, D., Fanciullino, R., Iliadis, A. & Ciccolini, J. (2012). Optimizing Druggability through Liposomal Formulations: New Approaches to an Old Concept. *ISRN pharmaceutics* **2012**, 738432.

29. Giljohann, D. A., Seferos, D. S., Prigodich, A. E., Patel, P. C. & Mirkin, C. A. (2009). Gene regulation with polyvalent siRNA-nanoparticle conjugates. *Journal of the American Chemical Society* **131**, 2072-3.

30. Fasbender, A., Marshall, J., Moninger, T. O., Grunst, T., Cheng, S. & Welsh, M. J. (1997). Effect of co-lipids in enhancing cationic lipid-mediated gene transfer in vitro and in vivo. *Gene therapy* **4**, 716-25.

31. Choosakoonkriang, S., Lobo, B. A., Koe, G. S., Koe, J. G. & Middaugh, C. R. (2003). Biophysical characterization of PEI/DNA complexes. *Journal of pharmaceutical sciences* **92**, 1710-22.

32. Tang, M. X. & Szoka, F. C. (1997). The influence of polymer structure on the interactions of cationic polymers with DNA and morphology of the resulting complexes. *Gene therapy* **4**, 823-32.

33. Leng, Q. & Mixson, A. J. (2005). Modified branched peptides with a histidine-rich tail enhance in vitro gene transfection. *Nucleic acids research* **33**, e40.

34. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. & Danielsen, M. (1987). Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7413-7.

35. Tang, M. X., Redemann, C. T. & Szoka, F. C., Jr. (1996). In vitro gene delivery by degraded polyamidoamine dendrimers. *Bioconjugate chemistry* **7**, 703-14.

36. Yang, S. & May, S. (2008). Release of cationic polymer-DNA complexes from the endosome: A theoretical investigation of the proton sponge hypothesis. *Journal of chemical physics* **129**, 185105.

37. Boussif, O., Lezoualc'h, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B. & Behr, J. P. (1995). A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 7297-301.

38. Akhtar, S. & Benter, I. F. (2007). Nonviral delivery of synthetic siRNAs in vivo. *Journal of clinical investigation* **117**, 3623-32.

39. He, X. X., Wang, K., Tan, W., Liu, B., Lin, X., He, C., Li, D., Huang, S. & Li, J. (2003). Bioconjugated nanoparticles for DNA protection from cleavage. *Journal of the American Chemical Society* **125**, 7168-9.

40. Werth, S., Urban-Klein, B., Dai, L., Hobel, S., Grzelinski, M., Bakowsky, U., Czubayko, F. & Aigner, A. (2006). A low molecular weight fraction of polyethylenimine (PEI) displays increased transfection efficiency of DNA and siRNA in fresh or lyophilized complexes. *Journal of controlled release* **112**, 257-70.

41. Urban-Klein, B., Werth, S., Abuharbeid, S., Czubayko, F. & Aigner, A. (2005). RNAi-mediated gene-targeting through systemic application of polyethylenimine (PEI)-complexed siRNA in vivo. *Gene therapy* **12**, 461-6.

42. Suck, D. & Oefner, C. (1986). Structure of DNase I at 2.0 A resolution suggests a mechanism for binding to and cutting DNA. *Nature* **321**, 620-5.

43. Wei Chen, N. J. T., Donald A. Tomalia. (2000). Using Ethidium Bromide To Probe the Interactions between DNA and Dendrimers. *Langmuir* **16**, 15-19.

44. Leng, Q., Kahn, J., Zhu, J., Scaria, P. & Mixson, J. (2007). Needle-like morphology of H2K4b polyplexes associated with increases in transfection in vitro. *Cancer therapy* **5B**, 193-202.

45. Haley, J., Kabiru, P. & Geng, Y. (2010). Effect of clustered peptide binding on DNA condensation. *Molecular bioSystems* **6**, 249-55.

46. Feng, X. Z., Lin, Z., Yang, L. J., Wang, C. & Bai, C. L. (1998). Investigation of the interaction between acridine orange and bovine serum albumin. *Talanta* **47**, 1223-9.

47. Kubota, Y. & Steiner, R. F. (1977). Fluorescence decay and quantum yield characteristics of acridine orange and proflavine bound to DNA. *Biophysical chemistry* **6**, 279-89.

48. Lyles, M. B. & Cameron, I. L. (2002). Interactions of the DNA intercalator acridine orange, with itself, with caffeine, and with double stranded DNA. *Biophysical chemistry* **96**, 53-76.

49. Wiethoff, C. M., Gill, M. L., Koe, G. S., Koe, J. G. & Middaugh, C. R. (2003). A fluorescence study of the structure and accessibility of plasmid DNA condensed with cationic gene delivery vehicles. *Journal of pharmaceutical sciences* **92**, 1272-85.

50. Lobo, B. A., Rogers, S. A., Choosakoonkriang, S., Smith, J. G., Koe, G. & Middaugh, C. R. (2002). Differential scanning calorimetric studies of the thermal stability of plasmid DNA complexed with cationic lipids and polymers. *Journal of pharmaceutical sciences* **91**, 454-66.

51. Tiyaboonchai, W., Woiszwillo, J. & Middaugh, C. R. (2003). Formulation and characterization of DNA-polyethylenimine-dextran sulfate nanoparticles. *European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences* **19**, 191-202.

52. Wang, D. A., Narang, A. S., Kotb, M., Gaber, A. O., Miller, D. D., Kim, S. W. & Mahato, R. I. (2002). Novel branched poly(ethylenimine)-cholesterol water-soluble lipopolymers for gene delivery. *Biomacromolecules* **3**, 1197-207.

53. Midoux, P., Kichler, A., Boutin, V., Maurizot, J. C. & Monsigny, M. (1998). Membrane permeabilization and efficient gene transfer by a peptide containing several histidines. *Bioconjugate chemistry* **9**, 260-7.

54. Shi, Z., Woody, R. W. & Kallenbach, N. R. (2002). Is polyproline II a major backbone conformation in unfolded proteins? *Advances in protein chemistry* **62**, 163-240.

55. Rucker, A. L. & Creamer, T. P. (2002). Polyproline II helical structure in protein unfolded states: lysine peptides revisited. *Protein science* **11**, 980-5.

56. Tiffany, M. L. & Krimm, S. (1972). Effect of temperature on the circular dichroism spectra of polypeptides in the extended state. *Biopolymers* **11**, 2309-16.

57. Johnson, K. H. & Gray, D. M. (1991). An estimate of the nearest neighbor base-pair content of 5S RNA using CD and absorption spectroscopy. *Biopolymers* **31**, 385-95.

58. Mukerji, I. & Williams, A. P. (2002). UV resonance Raman and circular dichroism studies of a DNA duplex containing an A(3)T(3) tract: evidence for a premelting transition and three-centered H-bonds. *Biochemistry* **41**, 69-77.

59. Sreerama, N. & Woody, R. W. (2003). Structural composition of βI- and βII-proteins. *Protein science* **12**, 384-8.

60. Cowan, P.M., McGavin, S. (1955). Structure of Poly-L-Proline. *Nature*, 501-503.

61. Poulos, T. L., Edwards, S. L., Wariishi, H. & Gold, M. H. (1993). Crystallographic refinement of lignin peroxidase at 2 A. *Journal of biological chemistry* **268**, 4429-40.

62. Tiffany, M. L. & Krimm, S. (1968). New chain conformations of poly(glutamic acid) and polylysine. *Biopolymers* **6**, 1379-82.

63. Fraser, R. D., MacRae, T. P. & Suzuki, E. (1979). Chain conformation in the collagen molecule. *Journal of molecular biology* **129**, 463-81.

64. Drake, A. F., Siligardi, G. & Gibbons, W. A. (1988). Reassessment of the electronic circular dichroism criteria for random coil conformations of poly(L-lysine) and the implications for protein folding and denaturation studies. *Biophysical chemistry* **31**, 143-6.

65. Daniel H. A. Corrêa, C. H. I. R. (2009). The use of circular dichroism spectroscopy to study protein folding, form and function. *African Journal of Biochemistry Research* **3**, 164-173.

66. Matulis, D., Rouzina, I. & Bloomfield, V. A. (2000). Thermodynamics of DNA binding and condensation: isothermal titration calorimetry and electrostatic mechanism. *Journal of molecular biology* **296**, 1053-63.

67. Huang, J. X., Cooper, M. A., Baker, M. A., Azad, M. A., Nation, R. L., Li, J. & Velkov, T. (2012). Drug-binding energetics of human alpha-1-acid glycoprotein assessed by isothermal titration calorimetry and molecular docking simulations. *Journal of molecular recognition : JMR* **25**, 642-56.

68. Choi, S. W., Kano, A. & Maruyama, A. (2008). Activation of DNA strand exchange by cationic comb-type copolymers: effect of cationic moieties of the copolymers. *Nucleic acids research* **36**, 342-51.

69. Patel, M. M. & Anchordoquy, T. J. (2005). Contribution of hydrophobicity to thermodynamics of ligand-DNA binding and DNA collapse. *Biophysical journal* **88**, 2089-103.

70. Bloomfield, V. A. (1997). DNA condensation by multivalent cations. *Biopolymers* **44**, 269-82.

71. Nimesh, S., Thibault, M. M., Lavertu, M. & Buschmann, M. D. (2010). Enhanced gene delivery mediated by low molecular weight chitosan/DNA complexes: effect of pH and serum. *Molecular biotechnology* **46**, 182-96.

72. Ogris, M., Steinlein, P., Carotta, S., Brunner, S. & Wagner, E. (2001). DNA/polyethylenimine transfection particles: influence of ligands, polymer size, and PEGylation on internalization and gene expression. *AAPS pharmSci* **3**, E21.

73. Hufnagel, H., Hakim, P., Lima, A. & Hollfelder, F. (2009). Fluid phase endocytosis contributes to transfection of DNA by PEI-25. *Molecular therapy* **17**, 1411-7.

74. Cardenas, M., Schillen, K., Nylander, T., Jasson, J., Lindman, B. (2004). DNA Compaction by cationic surfactant in solution and at polystyrene particle solution interfaces: a dynamic light scattering study. *Phys. Chem. Chem. Phys.* **6**, 1603-1607.

75. Hu, Y., Zandi, R., Anavitarte, A., Knobler, C. M. & Gelbart, W. M. (2008). Packaging of a polymer by a viral capsid: the interplay between polymer length and capsid size. *Biophysical journal* **94**, 1428-36.

76. Brown, J. C. & Newcomb, W. W. (2011). Herpesvirus capsid assembly: insights from structural analysis. *Current opinion in virology* **1**, 142-9.

77. de Silva, S. & Bowers, W. J. (2009). Herpes Virus Amplicon Vectors. *Viruses* **1**, 594-629.

78. Dunn, J. J. & Studier, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *Journal of molecular biology* **166**, 477-535.

79. Khan, S. A., Griess, G. A. & Serwer, P. (1992). Assembly-associated structural changes of bacteriophage T7 capsids. Detection by use of a protein-specific probe. *Biophysical journal* **63**, 1286-92.

80. Hagerman, P. J. (1997). Flexibility of RNA. *Annual review of biophysics and biomolecular structure* **26**, 139-56.

81. Wang, J., Lu, Z., Wientjes, M. G. & Au, J. L. (2010). Delivery of siRNA therapeutics: barriers and carriers. *The AAPS journal* **12**, 492-503.

82. Merkel, O. M., Beyerle, A., Librizzi, D., Pfestroff, A., Behr, T. M., Sproat, B., Barth, P. J. & Kissel, T. (2009). Nonviral siRNA delivery to the lung: investigation of PEG-PEI polyplexes and their in vivo performance. *Molecular pharmaceutics* **6**, 1246-60.

83. Conner, S. D. & Schmid, S. L. (2003). Regulated portals of entry into the cell. *Nature* **422**, 37-44.

84. Luzio, J. P., Rous, B. A., Bright, N. A., Pryor, P. R., Mullock, B. M. & Piper, R. C. (2000). Lysosome-endosome fusion and lysosome biogenesis. *Journal of cell science* **113 (Pt 9)**, 1515-24.

85. Sonawane, N. D., Szoka, F. C., Jr. & Verkman, A. S. (2003). Chloride accumulation and swelling in endosomes enhances DNA transfer by polyamine-DNA polyplexes. *The Journal of biological chemistry* **278**, 44826-31.

86. Shai, Y. (1999). Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochimica et biophysica acta* **1462**, 55-70.

87. Fernandez-Carneado, J., Kogan, M. J., Castel, S. & Giralt, E. (2004). Potential peptide carriers: amphipathic proline-rich peptides derived from the N-terminal domain of gamma-zein. *Angewandte Chemie* **43**, 1811-4.

88. Top, D., Read, J. A., Dawe, S. J., Syvitski, R. T. & Duncan, R. (2012). Cell-cell membrane fusion induced by p15 fusion-associated small transmembrane (FAST) protein requires a novel fusion peptide motif containing a myristoylated polyproline type II helix. *The Journal of biological chemistry* **287**, 3403-14.

89. Kim, J. L. & Burley, S. K. (1994). 1.9 A resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nat Struct Biol* **1**, 638-53.

90. Patikoglou, G. A., Kim, J. L., Sun, L., Yang, S. H., Kodadek, T. & Burley, S. K. (1999). TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev* **13**, 3217-30.

91. Davis, N. A., Majee, S. S. & Kahn, J. D. (1999). TATA box DNA deformation with and without the TATA box-binding protein. *J Mol Biol* **291**, 249-65.

92. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G. & Lu, P. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**, 1247-54.

93. Seeman, N. C., Rosenberg, J. M. & Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 804-8.

94. Zahran, M., Daidone, I., Smith, J. C. & Imhof, P. (2010). Mechanism of DNA recognition by the restriction enzyme EcoRV. *J Mol Biol* **401**, 415-32.

95. Rhee, H. S. & Pugh, B. F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**, 295-301.

96. Pardo, L., Campillo, M., Bosch, D., Pastor, N. & Weinstein, H. (2000). Binding mechanisms of TATA box-binding proteins: DNA kinking is stabilized by specific hydrogen bonds. *Biophys J* **78**, 1988-96.

97. Hoffmann, A., Oelgeschlager, T. & Roeder, R. G. (1997). Considerations of transcriptional control mechanisms: do TFIID-core promoter complexes recapitulate nucleosome-like functions? *Proc Natl Acad Sci U S A* **94**, 8928-35.

98. Starr, D. B., Hoopes, B. C. & Hawley, D. K. (1995). DNA bending is an important component of site-specific recognition by the TATA binding protein. *J Mol Biol* **250**, 434-46.

99. Gan, Y., Guan, J. & Zhou, S. (2012). A comparison study on feature selection of DNA structural properties for promoter prediction. *BMC bioinformatics* **13**, 4.

100. Parvin, J. D., McCormick, R. J., Sharp, P. A. & Fisher, D. E. (1995). Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature* **373**, 724-7.

101. Byun, J. S. (2005). Thesis Dissertation. Characterization of TATA Box Binding Protein Interaction With Minicircle DNA, University of Maryland, College Park.

102. Kahn, J. D. (2000). Topological effects of the TATA box binding protein on minicircle DNA and a possible thermodynamic linkage to chromatin remodeling. *Biochemistry* **39**, 3520-4.

103. Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J. & Dickerson, R. E. (1996). How proteins recognize the TATA box. *J Mol Biol* **261**, 239-54.

104. Lebrun, A., Shakked, Z. & Lavery, R. (1997). Local DNA stretching mimics the distortion caused by the TATA box-binding protein. *Proc Natl Acad Sci U S A* **94**, 2993-8.

105. Hoopes, B. C., LeBlanc, J. F. & Hawley, D. K. (1992). Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway. *J Biol Chem* **267**, 11539-47.

106. Parkhurst, K. M., Richards, R. M., Brenowitz, M. & Parkhurst, L. J. (1999). Intermediate species possessing bent DNA are present along the pathway to formation of a final TBP-TATA complex. *J Mol Biol* **289**, 1327-41.

107. Morse, R. H. & Cantor, C. R. (1985). Nucleosome core particles suppress the thermal untwisting of core DNA and adjacent linker DNA. *Proc Natl Acad Sci U S A* **82**, 4653-7.

108. Jangir, D. K., Charak, S., Mehrotra, R., Kundu, S. (2011). FTIR and circular dichroism spectroscopic study of interaction of 5-fluorouracil with DNA. *Journal of photochemistry and photobiology. B, Biology* **105**, 143-8.