5-1-2011

Calorimetric studies of histone H1 interactions with calf thymus DNA

Sarah Elizabeth Jones

CALORIMETRIC STUDIES OF HISTONE H1 INTERACTIONS WITH CALF THYMUS DNA

By

Sarah Elizabeth Jones

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Chemistry
in Chemistry
in the Department of Chemistry

Mississippi State, Mississippi

August 2011

CALORIMETRIC STUDIES OF HISTONE H1 INTERACTIONS WITH CALF THYMUS DNA

By

Sarah Elizabeth Jones

Approved:

Edwin A. Lewis Professor of Biophysical Chemistry and Head of the Department of Chemistry (Major Professor)	Joseph Emerson Assistant Professor of Inorganic and Bioinorganic Chemistry Department of Chemistry (Committee Member)
Susan Wellman Professor of Pharmacology and Toxicology Department of Pharmacology and Toxicology (Committee Member)	Stephen Foster Associate Professor of Physical Chemistry Department of Chemistry (Graduate Coordinator)

Gary L. Myers
Dean of the College of Arts & Sciences
and Professor of English

Name: Sarah Elizabeth Jones

Date of Degree: August 6, 2011

Institution: Mississippi State University

Major Field: Chemistry

Major Professor: Edwin A. Lewis

Title of Study: CALORIMETRIC STUDIES OF HISTONE H1 INTERACTIONS WITH CALF THYMUS DNA

Pages in Study: 43

Candidate for Degree of Master of Chemistry

In this study we have used isothermal titration calorimetry, ITC, and circular dichroism spectropolarimetry, CD, to directly measure the thermodynamics and the structural changes for binding histones, $H1_1$ and $H1_4$, to DNA. The ITC data have been used to estimate the binding constant, ($K \approx 10^8$) and the enthalpy change ($\Delta H \approx +5$ ($H1_1$ at 25ºC), $\Delta H \approx +20$ kcal/mol ($H1_4$ at 15 ºC) for formation of the H1/DNA complex. CD data indicate that both H1 and DNA are partially unfolded in the H1/DNA complex. Protein and DNA unfolding must contribute to the large unfavorable endothermic enthalpy change for complex formation. The ITC data indicate that the $H1_1$ binding site is comprised 30 DNA base pairs while $H1_4$ interacts with approximately 36 DNA base pairs. At saturation, our data are consistent with 100% of the H1 binding sites being occupied in the H1/DNA complex.

DEDICATION

I would like to dedicate this thesis to my father, Robert Jones, aunt, Theda McCrory, and fiancé, Ken Hendrix, for all their love and support throughout my academic career.

ACKNOWLEDGEMENTS

I would like to thank Dr. Joseph Emerson and his lab group for the generous use of their centrifuges and autoclave. I would like thank Dr. Wellman for the *E. coli* strains, extraction procedure, and protein. I would also like to thank Dr. Lewis for his guidance throughout my graduate career at Mississippi State University.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Histone proteins are DNA binding proteins that are primarily responsible for compacting DNA into chromatin fibers. Chromatin packs large amounts of genetic DNA into the nucleus of a cell [1, 2] . There are five different classes of histone proteins: H1 (and H5), H2A, H2B, H3, and H4. Two of each of the core histones (H2A, H2B, H3, and H4) come together to form an eight histone core known as the nucleosome core, a sub-unit of chromatin [3, 4]. The core histones are comprised of three alpha helices separated by two loops [5]. One hundred and forty six base pairs then wrap around the core like a spool of thread in a 1.65° left handed helix. A region called linker DNA separates each nucleosome from other succeeding or preceding nucleosomes. This linker DNA is stabilized by histone H1. H1 binds to each nucleosome and its adjacent linker DNA to assist in adjoining nucleosomes together, for further super coiling. In other words, the linker proteins (histone H1 and recently discovered histone H5) connect the ends of the DNA at the entry and exit sites [3-5]. The nucleosome then forms the fundamental repeating units of chromatin [4]. A simplified nucleosome structure is illustrated in Figure 1.1. Figure 1.1 illustrates the DNA wrapped around the core histone 1.6 times in a left- handed helical turn. The core octomer comprised of two of each of the core histones is shown in red, while the blue is the dsDNA wrapped around these proteins. The crystal structure of the nucleosome core was obtained from the protein data bank [6]. A NMR structure of H1 was docked with the nucleosome core; this structure was also obtained

from the protein data bank [7]. The DNA tails (yellow) and the N and C terminal regions of H1 (black) were created using Discovery Studios. The image shows the H1 protein (pink) interacting with the minor groove of the DNA at the entry and exit sites of the nucleosome core.

H1 interactions with DNA have attracted significant interest due to the involvement of H1 in chromatin organization and the fact that H1 is more highly modified in cancer cells [8]. Histone H1 has also been shown to regulate gene expression by preventing the binding of transcription factor proteins to DNA [8, 9]. However, questions remain about the actual mechanisms of these histone H1 proteins. In this study we focus on Histone H1/DNA interactions and the effects of sequence in H1 protein/DNA binding interactions.

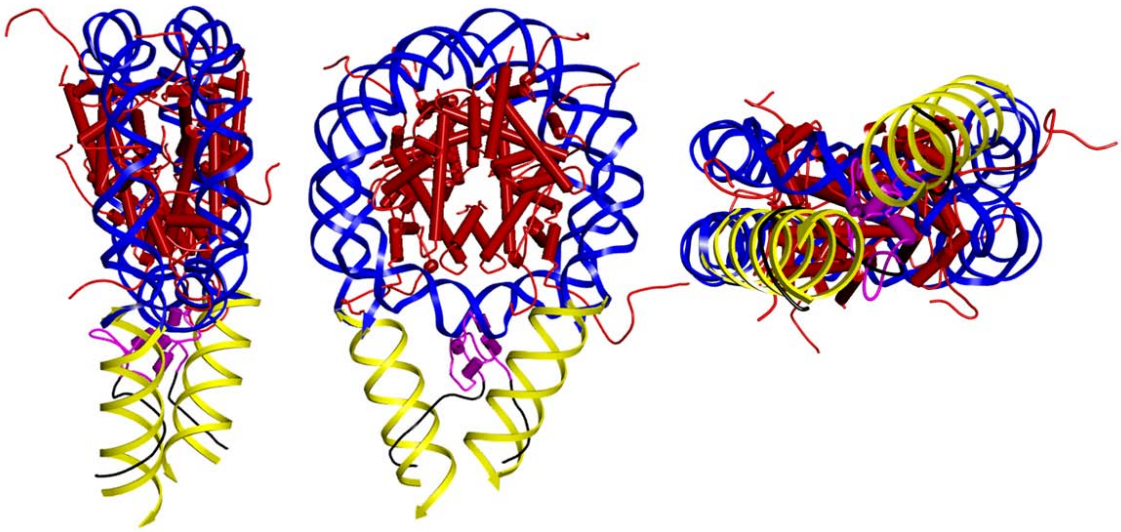


Figure 1.1 Schematic of Nucleosome structure. The nucleosome depicted includes the octomer of the core histones (H2A, H2B, H3, and H4), the H1 histone linker, and dsDNA.

The structural relationship between the H1 and the nucleosome is not clearly understood. H1 is comprised of three regions: a central globular region (containing 80-85

residues), a basic, hydrophilic C-terminus (containing 90-100 residues), and a basic, hydrophilic N-Terminus (containing 20-40 residues) [10-12]. The N-terminus domain is known to be enriched in alanine, lysine, and proline residues. The central globular region is the most highly conserved region in all of the variants [12]. The structure of the central globular domain is similar to that of typical globular proteins, with the exception that histone H1 variants have very few aromatic amino acids [13]. The central globular region is the only region of H1 that is folded in solution without the presence of a physiological binding partner (i.e. dsDNA) [11, 14, 15]. The C-terminal domain, like the N-terminal region is again enriched in lysine, alanine, and proline residues. Lysine makes up almost 50% of the amino acids in the C-terminal domain, while alanine and proline make up another 30% of the amino acids present in this domain [13]. Although the N and C-terminal regions of the protein are thought to be unstructured in solution in the absence of a binding partner, it has been hypothesized that the two terminal regions fold into a defined tertiary structure in the H1/DNA complex [8].

A McGhee-von Hippel binding model has been used to predict the equilibrium constant and the DNA binding site size for binding H1 to ds DNA. In these studies $H1_0$ and $H1_1$ were estimated to have binding constants in the range of 10^8 [10, 11]. Again using a McGhee-von Hippel binding model, it has been suggested that $H1_4$ binds DNA approximately 100 times more tightly than either $H1_0$ or $H1_1$ [13].

All of the linker Histone variants are positive charged at neutral pH ($H1_1$ carries a charge of +54 and $H1_4$ carries a charge of +58 at pH=7) [10]. All H1 variants bind to DNA through a large number of electrostatic interactions between the positive charged terminal regions and the DNA which is a polyanion at neutral pH. We have performed experiments using two different H1 variants ($H1_4$ and $H1_1$), also known as H1e and H1c.

All of the H1 variants have very similar structures. This point is illustrated in Figure 1.2 which provides three different NMR structures for the core domain of the linker histone H1. These structures were obtained from the PDB bank; three out of fourteen structures were chosen randomly to show similarities in structure between the possible variants [7]. This figure illustrates that the core domain of the protein is comprised of three alpha helices connected by beta loops. Presumably, the structures of $H1_1$ and $H1_4$ would also be very similar and subtle changes in amino acid composition, charge and size should have little effect to the gross solution structure. Nevertheless, $H1_1$ and $H1_4$ have been reported to have very different affinities and binding site sizes from the McGhee- von Hippel modeling studies [10].

Figure 1.2 Core domain of Histone H1 [7]

The differences in the two H1 variants used in this study, $H1_1$ and $H1_4$ are described below. $H1_1$ has a molecular weight of 21.205 kDa, an isoelectric point 11.0, and a positive charge of 54.9 at pH 7. The isoelectric points and molecular weights of the proteins were obtained from Dr. Wellman [13, 16]. It is composed of 212 amino acids: 62 strongly basic (Lys and Arg), 7 strongly acidic (Asp and Glu), 83 hydrophobic (Ala,

Ile, Leu, Phe, and Val), and 29 polar amino acids (Asn, Gln, Ser, and Thr). $H1_4$ has a molecular weight of 21.917kDa, an isoelectric point of 11.1, and a charge of 58.9 at pH 7.0. $H1_4$ is comprised of 219 amino acids: 66 strongly basic (Lys and Arg), 7 strongly acidic (Asp and Glu), 76 hydrophobic (Ala, Ile, Leu, and Val), and 36 Polar amino acids (Asn, Gln, Ser, and Thr) [17].

The sequence homology of $H1_1$ in comparison to $H1_4$ is shown in Figure 1.3. $H1_4$ is 6 amino acids longer than $H1_1$, and overall the two sequences are 87% homologous. The higher positive charge for the $H1_4$ histone variant is due solely to the 4 additional lysines in the $H1_4$ sequence.

```
 10        20        30        40        50        60
H1_1   ASETAPAAPAAAPPAEKAPVKKKAAKKAGGTPRKASGPPVSELITKAVAASKERSGVSLA
       ::::::::::  ::::.:::::: :.:::. ::::::::::::::::::::::::::::
H1_4   ASETAPAAPAAPAPAEKTPVKKKARKSAGAAKRKASGPPVSELITKAVAASKERSGVSLA
                10        20        30        40        50        60

                70        80        90       100       110       120
H1_1   ALKKALAAAGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFKLNKKAASGEAKPKV
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::.
H1_4   ALKKALAAAGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFKLNKKAASGEAKPKA
                70        80        90       100       110       120

               130       140       150       160       170       180
H1_1   KKAGGTKPKKPVGAAKKPKKAAGGATPKKSAKKTPKKAKKPAAATVTKKVAKSPKKAKVA
       ::::..: :::.:::::::::.:.:::::::::::::::::::. .:: ::::::::.:
H1_4   KKAGAAKAKKPAGAAKKPKKATGAATPKKSAKKTPKKAKKPAAAAGAKK-AKSPKKAKAA
               130       140       150       160       170

               190         200       210
H1_1   KPKKAAKSAAKA--VKPKAAKPKVVKPKKAAPKK
       ::::: :: :::  ::::::::::.::: : :::
H1_4   KPKKAPKSPAKAKAVKPKAAKPKTAKPKAAKPKK
     180       190       200       210
```

Figure 1.3 Sequence homology for $H1_1$ and $H1_4$ histone proteins.

The one letter amino acid abbreviations denoted in red indicate homologous amino acids residues within the two sequences while residues shown in green denote

semi conserved substitutions, and blue residues denote non homologous residues. The sequence for $H1_1$ is denoted on the top while the sequence for $H1_4$ is denoted on the bottom. When comparing the $H1_1$ and $H1_4$ sequences there are 10 non-conservative substitutions observed at residues 12, 13, 25, 32, 128, 166, 186, 189, 209, 211, and 15 semi conserved substitutions are observed at residues 18, 27, 30, 31, 120, 125, 26, 132, 142, 144, 165, 167, 179, 204, and 205. Amino acid 170 is deleted from the $H1_1$ sequence, and there is an addition of lysine and alanine at residues 193 and 194 in the $H1_4$.

The thermodynamic signatures for the binding of $H1_1$ and $H1_4$ to their physiological binding partner, CTDNA, were previously hypothesized to be exothermic based upon the calculations resulting from Scatchard plots based on the results from Mamoon *et als's* McGhee-von Hippel modeling studies. We report here that the binding $H1_1$ and $H1_4$ to ds DNA is actually an endothermic process that involves structural changes upon formation of the H1protein/dsDNA complex. The binding site size for histone $H1_1$ has been reported to be 15 DNA base pairs while the binding site size for histone $H1_4$ has been reported to be 18 DNA base pairs [10, 11]. Understanding the molecular nature of H1 interactions with DNA will ultimately be important in developing an understanding of the role that H1 plays in condensing DNA in chromatin. More importantly, subtle differences in H1 sequence resulting from post-translational modification reactions may be important in gene regulation.

To date eleven H1 variants have been identified [13]. $H1_1$ and $H1_4$ are two of the most studied H1 variants, and these variants exhibit both the highest positive charges and the largest binding affinities [8, 10, 11, 18]. The determination of the best model for the histone complex can lead to more specific studies in the future. This study attempts to answer several questions regarding the interaction of two different H1 variants with its

physiological binding partner ds DNA. The three main objectives were: 1) to determine the thermodynamic signatures for the binding of $H1_1$ and $H1_4$ to calf thymus DNA (CTDNA), 2) to evaluate whether the binding of these two linker proteins to dsDNA was accompanied by a change in either the protein or DNA structure, and 3) to attempt to attribute any changes in binding affinity, binding site size, or binding driving forces to the changes in the $H1_1$ and $H1_4$ sequences.

CHAPTER II

MATERIALS AND METHODS

Histone H1 expression and purification

The two histone H1 variants, $H1_1$ and $H1_4$, were expressed using a bacterial strain of *E. coli* (Rosetta (DE3) pLysS) infected with a peT-11d expression vector which contained the sequence for either the mouse $H1_1$ protein or the mouse $H1_4$ protein. The construction of the expression strains, induction of protein expression, and extraction and identification of the proteins has been described [18]. These two expression strains were a gift from Dr. Susan Wellman at The University of Mississippi Medical Center. Cultures of the two infected *E. coli* strains containing the expression vectors were maintained in our laboratory, and fresh agar plates were streaked for each strain each month throughout the project period. To express the H1 variants, cells were grown up in TY media at 37 ºC for 12 hours or until they reached mid exponential phase ($Abs_{600} \geq 0.7$).

An Auto-inducing medium was used to produce proteins. The auto-inducing medium was made by supplementing TY medium with a specific mix of sugars (ZYM 5052) to regulate induction of expression, and selected micronutrients to support robust growth of the cells. The TY medium was made in a single batch that contained 10g bactotryptone, 5 g yeast extract, 5 g NaCl, and 800 mL of water. The 800 mL solution was sterilized and divided among four flasks. ZYM 5052 stock solution (25% glycerol, 2.5% glucose, 10% lactose, and 62.5% nanopure water) was made by adding weighed amounts of the glycerol, glucose, and lactose to a 500 ml volumetric flask and adding the

nanopure water (contains no particles larger than one nanometer) to the mark. The ZYM solution was sterilized prior to use. The following stock solutions were prepared, sterilized by filtration, and kept at 5ºC prior to use: 1 M $MgSO_4$, 1 M $FeCl_3$, 100 mg/ml ampicillin, and 25xM solution (17.8g NA_2HPO_4, 17 g KH_2PO_4, 13.4 g NH_4Cl, 8.1 g Na_2SO_4, 200 mL of nanopure water). Sterile pipette tips were used to add 0.8 ml of $MgSO_4$ solution, 8 ml of 25xM solution, 4 ml of ZYM 5052, 0.4 ml of the ampicillin solution, and 0.4 ml of the $FeCl_3$ solution to each flask of sterile TY broth. Each preparation of protein was begun by inoculating one colony of the H1 expression strain from an agar plate into 5 mL of the sterile TY broth. The culture was incubated at 37°C in a shaker for four hours or until the optical density (A_{600}) reached an absorbance value of at least 0.5. The culture was then split and added to the four baffled flasks containing the auto-inducing medium. The flasks were then placed into the incubator where they were shaken for 12 hours at 37°C. After 12 hours, the flasks were removed from the incubator and the contents were poured into two centrifuge bottles. The cells were pelleted in the centrifuge for 30 min at 1800 rpm (1000 x g). The supernatant was removed, and the pellets were allowed to drain for about 30 min before the addition of the lysis buffer.

Lysis

Each of the two *E. coli* pellets (containing the expressed H1 protein) were next resuspended in approximately 40 ml of lysis buffer (50mM Tris pH 8, 50mM EDTA, 50mM NaCl, 0.5 % Triton X-100). The resuspended *E. coli* cells were refrigerated and shaken hourly by hand throughout the remainder of the day (approximately 8 hours) to ensure that the pelleted cells were completely suspended in the lysis buffer before being

frozen. At this point the pellet is stable and can be kept frozen for longer periods of time, if needed. The frozen cells were thawed at room temperature, transferred to a smaller centrifuge bottle (50 ml), and 0.04g of lysozyme was then added to each centrifuge bottle. After lysozyme addition, the *E. coli* cells were incubated on ice for one hour to cleave the cells. The cellular debris was spun at 22,000 x g for one hour (12,000 rpm). The supernatant was discarded and the expressed H1 protein (now found in the lysed cell pellet) was extracted with approximately 40 mL of 5% perchloric acid. The acid extraction of the pellet was done for 60 minutes at 0ºC. After the incubation, the protein solution was spun at 12,000 rpm for 25-30 minutes. The supernatant was collected which contained the histone H1. The pellets were discarded, and the supernatant was refrigerated overnight. The solution was neutralized with 13 mL of 6 M NH_4OH was added to the supernatant. The protein was then concentrated using amicon ultracentrifugal filter devices for volumes with a 5 kDa molecular weight cutoff.

HPLC

High-Performance liquid chromatography is used to identify and purify macromolecules. HPLC has two different phases: a stationary phase (column) and a mobile phase (solvent) [19-22]. The most common detector used is a UV-VIS detector, but they can also be coupled with other detectors such as a mass spectrometer [22]. HPLC can be used for both analytical and preparatory purposes. Each component in the mixture will elute as a separate peak, which allows the user to estimate the purity of the sample collected [20, 22]. The total area of the peak is directly related to the amount of macromolecule present. Reverse-phase HPLC was used for the purification of the histone $H1_1$ and $H1_4$ variants. The full length histone $H1_1$ and $H1_4$ variants were purified by

reverse-phase HPLC using a Dionex P680 HPLC with a Dionex UVD170U detector. A 5 mL sample loop was used therefore only one HPLC was performed per extraction.

The semi-prep reverse-phase column used was a Grace Davison Discovery Science C_{18} hydrophobic interaction column (10mm x 250mm, 5μM). Buffer A was 0.1% trifluoroacetic acid and buffer B was 95% acetonitrile/0.1% trifluoroacetic acid. The gradient was 0% buffer B for 5min, 0-40% buffer B in 20 min, and 40-70% buffer B in 30 min. The flow rate was 3.25 ml/min. Two UV-VIS absorption channels, 220 nm and 254 nm, were used for detection. The 220 nm channel was used for detection of the protein, while the 254 nm channel was used for detection of contaminants such as DNase and RNase. A 220 nm absorption channel was used for detection of the histone proteins due to the lack of aromatic amino acids (i.e. there are no tryptophan or cysteine amino acids present within the $H1_1$ and $H1_4$ protein sequences) [13]. The chromatogram for Histone $H1_1$ is shown in Figure 2.1. The spectrum collected at 220 nm (black) exhibits a large peak at ~ 51 minutes which is consistent with previously reported data for pure Histone $H1_1$ [13, 18]. In comparison to the spectrum collected at 220 nm (black), the small peak observed at 254 nm (red) indicates that there was little or no DNA or RNA present in the purified sample [18]. An absorption channel of 254 nm was used for the detection of RNA and DNA because the aromatic bases absorb in this region of the UV spectrum.

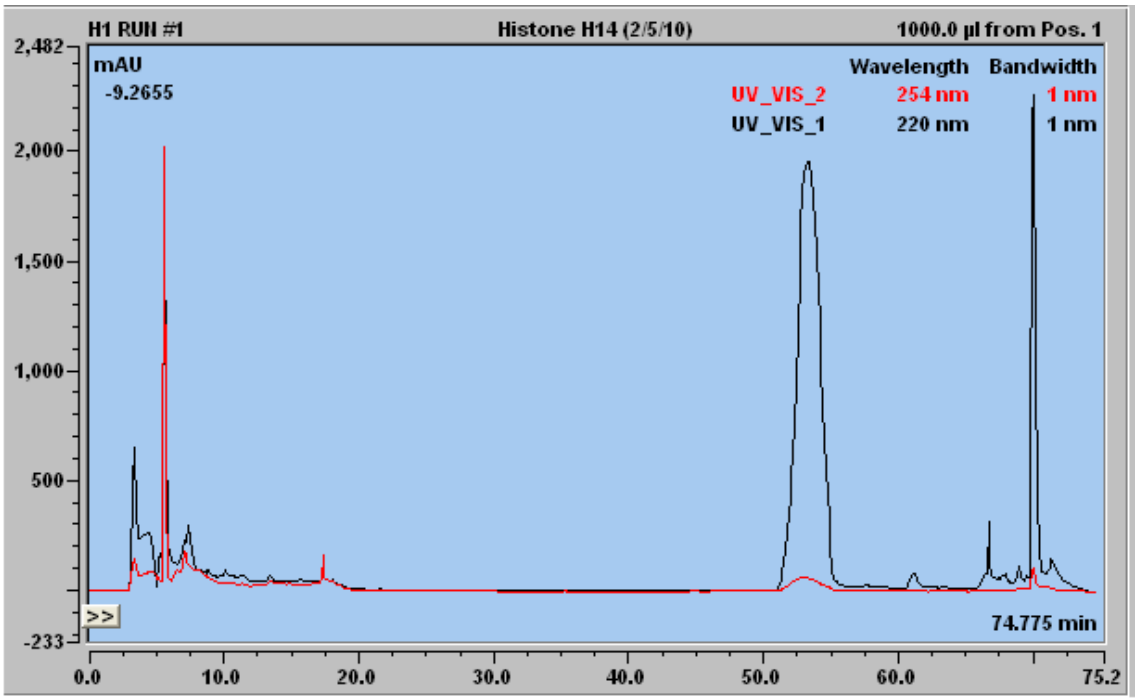


Figure 2.1 Characteristic chromatogram of $H1_1$. The black spectrum is the 220 nm absorption channel, while the red is the 254 nm absorption channel. The $H1_1$ protein eluted around 51 minutes which is consistent with previous data obtained by Wellman [7]. The low absorbance at 254 nm at 51 minutes is indicative of little or no DNase or RNase present in the purified protein sample.

The chromatogram for Histone $H1_4$ is shown in Figure 2.2. The spectrum collected at 220 nm (black) exhibits a large broad peak eluting at ~45 minutes which is consistent with previously reported data for the pure Histone $H1_4$ [13]. The zero absorbance at 254 nm (red line) indicates that there was little or no DNA or RNA present in the purified sample of $H1_4$ [13].

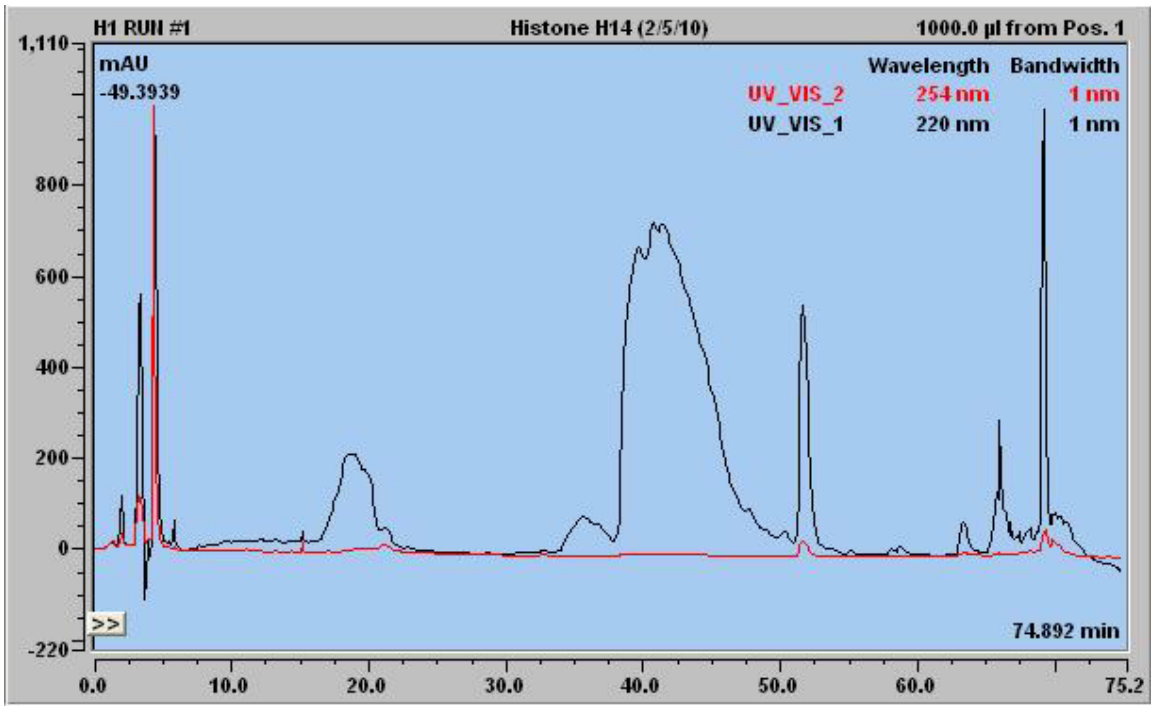


Figure 2.2 Characteristic Chromatogram of $H1_4$. The black line is the 220 nm absorption channel, while the red is the 254 nm absorption channel. The $H1_4$ protein eluted around 45 minutes which is consistent with previous data obtained by Wellman [13]. The zero absorbance at 254 nm at 45 minutes is indicative of little or no DNase or RNase present in the purified protein sample.

Although the sequences are similar, the HPLC chromatograms and respective elution times for the $H1_1$ and $H1_4$ histone protein variants are notably different. The $H1_1$ variant contains 29 polar amino acids whereas the $H1_4$ variant contains 36 polar amino acids. The relative number of polar amino acids greatly impacts the elution time and the interaction with the HPLC stationary phase. $H1_4$ is a more polar protein than $H1_1$, therefore $H1_4$ elutes faster (45 min) in comparison to $H1_1$ (51 min) because of their relative interactions with the column.

SDS-PAGE was used to determine the purity of the fractions. The pure fractions were lyophilized using a Savant SPD 111V speed vac system for 4 hrs at 35°C. After lyophilization, SDS PAGE was performed to determine the purity of the samples.

SDS-PAGE

A common method for separating proteins by electrophoresis uses a polyacrylamide gel as a support medium and sodium dodecyl sulfate (SDS) to denature the proteins. This method is called sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). SDS is an anionic detergent, meaning that when dissolved, its molecules have a net negative charge. A polypeptide chain binds amounts of SDS in proportion to its relative molecular mass [23]. The negative charges on SDS destroy most of the complex structure of proteins, and the charges are strongly attracted toward an anode (positively-charged electrode) in an electric field. As a result, polyacrylamide gels prevent larger molecules from migrating as fast as smaller molecules [24]. Protein separation by SDS-PAGE can be used to estimate relative molecular mass, determine the relative abundance of proteins in a sample, and determine the distribution of proteins among fractions [23, 24]. The purity of the protein was determined before preparation of the ITC and DSC samples.

Gel Electrophoresis was performed using Bio-RAD ready gels (12% TRIS-HCl 10 wells, 30μL). These gels are precast by Bio-Rad for Polyacrylamide electrophoresis; 25 μL of the sample (the protein sample was mixed 1:1 with sample dye purchased from Nusep (TruSep SDS sample Buffer) was loaded into the wells. After loading, the gel was allowed to equilibrate for 10 minutes at 30 V. After 10 minutes the gel was conducted at 120 V for 2 hours. The gel was denatured in 100mL of fixative (50 mL trichloroacetic acid, 50 mL acetic acid, and 150 mL ethanol) for 20 minutes, stained in Coomassie blue solution (Bio-RAD) for 20 minutes, and destained with a destaining solution comprised of methanol, acetic acid, and water until the background was of desired clarity. Figure 2.3 shows a representative gel comparing four different $H1_1$ protein extractions (unpurified

$HClO_4$ extract, Lanes 1-4) and the HPLC purified $H1_1$ protein from the same four extractions (Lanes 5 and 8). Lanes 1-4 that include the unpurified $H1_1$ protein extract clearly contain multiple dark and diffuse bands indicating these extracts consist of a number of disperse protein species in solution. Lanes 5-7 representing the purified $H1_1$ protein exhibit a single intense (clear) band with similar band migration distances. These bands are consistent with the separation and migration for a single purified $H1_1$ protein species [18]. Molecular weight markers were used in previous experiments to determine the size and migration distances of these pure proteins versus their unpurified extract. The protein extractions represented in lanes 5-7 were collectively pooled and then used for ITC, DSC, and CD analysis. Lane 8 includes an $H1_1$ extraction sample that was purified by HPLC. However, this sample exhibited a number of additional bands. This sample was purified an additional time using HPLC and was subsequently not added to the other $H1_1$ protein extracts.

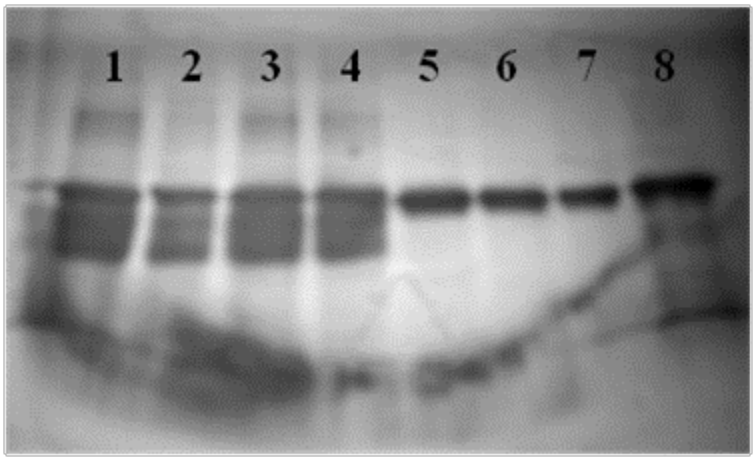


Figure 2.3 Gel Electrophoresis for purified $H1_1$. The lanes 1-4 are $HClO_4$ extract which matches the purified bands on the right in lanes 5-7. Lanes 5-7 are indicative of pure $H1_1$ while lane 8 indicated some impurities in the sample batch. Protein samples 5-7 were pooled together and used for later experiments.

Figure 2.4 shows a representative gel comparing purified $H1_4$ with varying volumes of sample buffer (Lanes 1 and 3) and the $HClO_4$ protein extract prior to HPLC purification (Lane 2). The band in Lane 1 results from a sample mixture that includes 10 µL of purified $H1_4$ protein with 10 µL of the sample buffer, whereas the band in Lane 3 results from a sample mixture that includes 30 µL of purified $H1_4$ protein with 10 µL sample buffer; 25 µL of each of the samples was loaded into the wells. Only a single intense (clear) band is observed for the purified $H1_4$ protein at both concentrations (Lanes 1 and 3) which is consistent with the separation and migration for a single protein species [23]. Lane 2 of the gel clearly contains multiple dark and diffuse bands for the unpurified $H1_4$ protein extract sample. This indicates a non-heterogeneous sample or a sample containing a number of disperse protein species in solution [24].

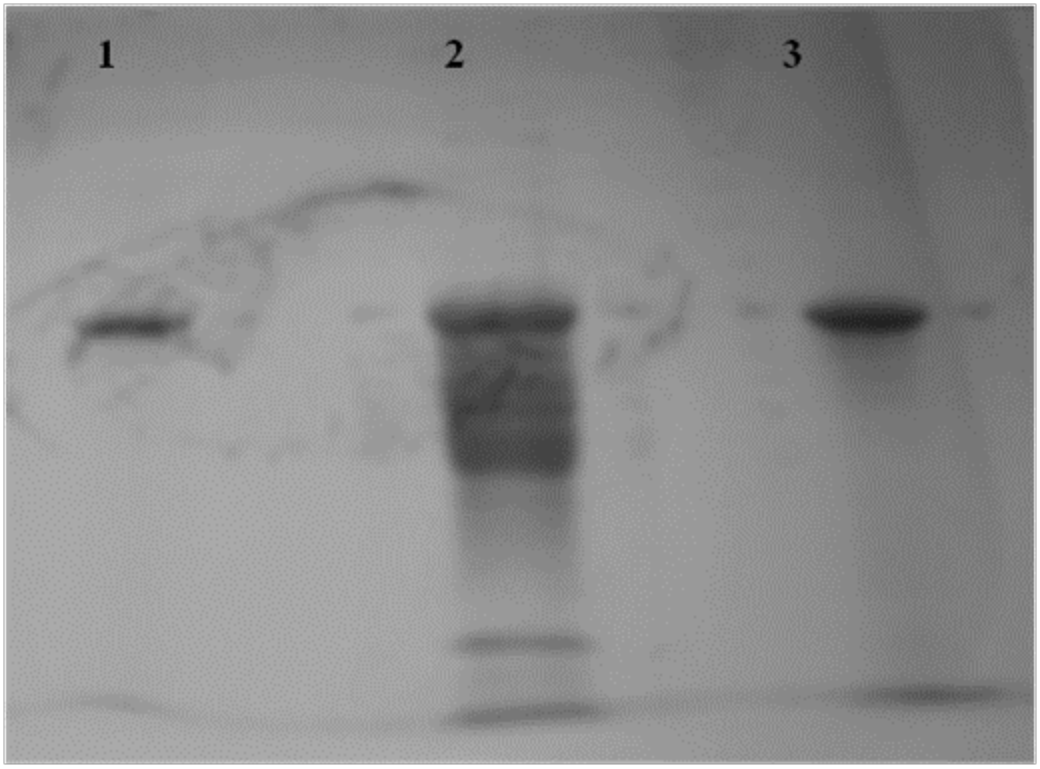


Figure 2.4 Gel Electrophoresis for purified $H1_4$. Lane 1 is $H1_4$ mixed 1:1 with sample buffer (Nusep SDS Sample Buffer), lane 2 is the $HClO_4$ extract for this batch of protein and lane 3 is purified $H1_4$ mixed 3:1 sample/sample dye. Lanes 1 and 3 are indicative of purified $H1_4$ protein after the purification of the $HClO_4$ extract by High Performance Liquid Chromatography.

Calf Thymus DNA (CTDNA)

The CTDNA was purchased from Sigma Aldrich. The CTDNA stock was prepared by weighing out approximately 12 mg of CTDNA and solvated with 1 mL of 130mM K^+ BPES (0.01M KH_2PO_4, 0.001 M EDTA, and 0.02-0.6 M KCl) at pH 7.0. The 1 mL sample was exhaustively dialyzed in the buffer using a 1,000 molecular weight cutoff membrane. Two buffer changes were performed (1 L every 24 hours at 4°C). The concentrations of the stock solutions were determined using an Olis 8452A diode array spectrophotometer at 260 nm. The extinction coefficient determined for CTDNA was $\varepsilon_{260}=1.31 \times 10^4\ M^{-1}cm^{-1}$ [25, 26].

The protein stocks were prepared by weighing out approximately 10 grams of protein and solvated with 1mL of 130mM K^+BPES. It was also exhaustively dialyzed using a 1,000 molecular weight cutoff membrane with two buffer changes (1L every 24 hours at 4°C). The stock concentrations of the protein were determined using UV-Vis. The extinction coefficient used was $\varepsilon_{220} = 152864\ M^{-1}cm^{-1}$ [13].

Circular Dichroism

Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right and left-handed circularly polarized light [27]. Only compounds that are asymmetric, that is, compounds in which the mirror image of the compound cannot be superimposed on itself, will show an absorption in the CD [28-30]. A schematic of the incident beam, left and right polarizing components, and an example of resulting CD spectra is shown below in Figure 2.5.

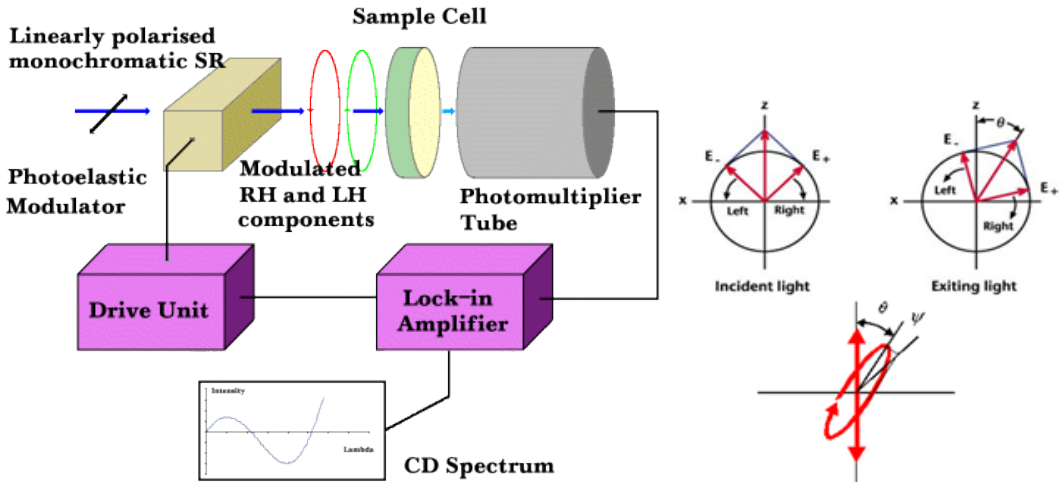


Figure 2.5 Representative diagram of the incident beam, the left and right polarizing components, and an example of the resulting CD spectrum [31].

Circular dichroism spectroscopy can be used for: determining whether a protein is folded and subsequently characterizing its secondary structure, comparing structures for different mutants of the same protein, studying the conformational stability of a protein subjected to pH changes or thermal stress, determining conformation changes with protein-protein or protein-ligand interactions.

Secondary structure can be determined by CD spectroscopy in the "far-UV" spectral region (190-250 nm) [27]. Alpha-helix, beta-sheet, and random coil structures each give rise to a characteristic CD spectrum [27]. Like all spectroscopic techniques, the CD signal reflects an average of the entire molecular population. Although CD can determine the percentage of a particular secondary structure present in a protein; it cannot determine which specific residues are involved in a particular secondary structure [27]. An Olis DSM 20 spectropolarimeter was used to determine the structure of the histone variants ($H1_1$ and $H1_4$). CD experiments were used to determine gross structural changes in the DNA (240 and 280nm regions) and the histone variants upon formation of the histone/DNA complex. The protein concentration was made up at a UV-Vis absorbance of 0.8 which is nominally 5 μM. The concentration of the CTDNA is nominally 5μM based on a 30 and 36 base pair DNA binding site ($H1_1$ and $H1_4$ respectively). The complex was prepared by mixing 1 mL of the histone $H1_1$ or $H1_4$ protein sample with 1 mL of the CTDNA sample, creating a 1:1 molar ratio between the protein and the DNA. The concentrations of both the protein and the CTDNA were verified using UV-Vis. The spectra were collected over a wavelength range of 195 to 305 nm.

Differential Scanning Calorimetry

Differential Scanning Calorimetry (DSC) directly measures the stability and unfolding of a protein, lipid, or nucleotide. In DSC, the biomolecule is heated at a constant rate, which provides a detectable heat change associated with thermal denaturation [32-35]. In a single DSC experiment, you can determine: Transition midpoint - T_m and Enthalpy (ΔH) and heat capacity change (ΔCp) associated with unfolding.

Differential Scanning Calorimetry was used to determine the stability of the histone/DNA complex versus the stability of the dsDNA alone. The DSC experiment can also verified the determination of the binding affinities obtained by Isothermal Titration Calorimetry. A depiction of the instrument is shown in Figure 2.6 [32].

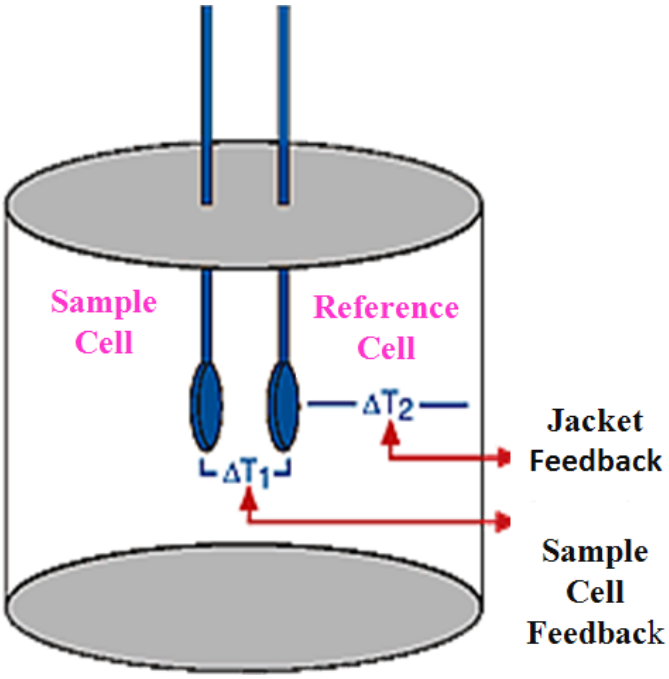


Figure 2.6 Representative diagram of a typical DSC including the sample and reference cells [32].

The differential scanning calorimetry (DSC) experiments were performed using a Microcal VP-DSC. Both the protein and DNA concentrations were nominally 100μM. The scan range was 10 to 95°C with a scan rate of 90°C/hr. The CTDNA denaturation is irreversible; therefore only one scan is performed for each experiment. The DSC data was fit for melting temperature (T_m), the calorimetric enthalpy (ΔH_{cal}), and the van't Hoff enthalpy, (ΔH) for each transition using Origin 7.1 software (Microcal).

Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) is a thermodynamic technique for monitoring chemical reactions and has become the method of choice for characterizing biomolecular interactions [34, 36, 37]. In ITC, a syringe containing a ligand solution is titrated into a cell containing a solution of the macromolecule at constant temperature. When the ligand is injected into the cell, a chemical reaction occurs, and heat is released or absorbed. As the macromolecule in the cell becomes saturated with the ligand, the heat signal diminishes until only the heat of dilution is observed. Isothermal Titration Calorimetry is the only biophysical technique that can measure the binding constant (K), change in enthalpy (ΔH), change in entropy (ΔS), and molar ratio (n) in one experimental procedure. We used a Microcal VP-ITC (Figure 2.7) to determine the binding affinity of the histone $H1_1$ and $H1_4$ variants for CTDNA [38, 39].

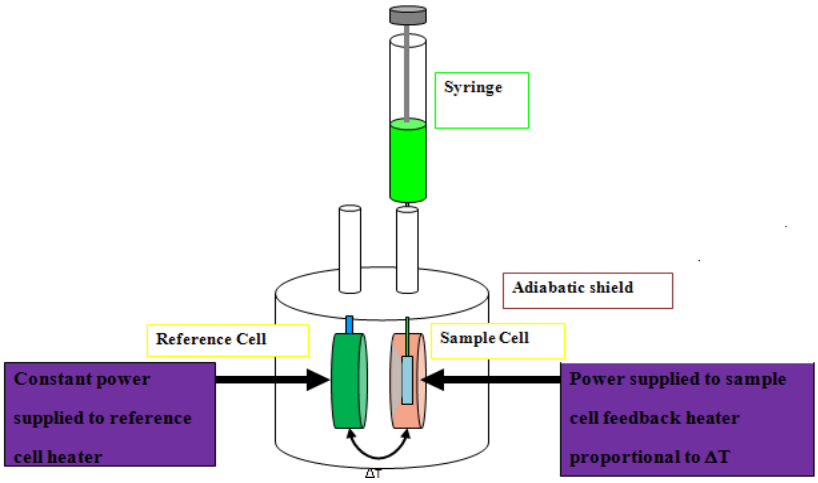


Figure 2.7 Schematic of a typical power compensation Isothermal Titration Calorimeter. The main features are labeled in the figure above (sample cell, reference cell, adiabatic shield, sample cell, reference cell, and syringe) [39].

All of the CTDNA titrations were performed by overfilling the ITC with approximately 1.5 mL of nominally 8 µM CTDNA sample (based on a 30 base pair binding site ($H1_1$) or a 36 base pair binding site ($H1_4$). Approximately 150 µl of the histone variants (nominally 180 µM) were titrated into the DNA in increments of 5 µL. Experiments were performed at 3 different temperatures (15 °C, 25 °C, and 35 °C), and were performed in triplicate. The ionic strength (130 mM K^+) and pH (7.0) remained the same throughout all the experiments. The data was fit using a "two independent sites" model in Origin 7.0 (Microcal). The integration range was adjusted manually to integrate the peaks themselves and not the spacing between each injection. This prevents any noise in the baseline having a large effect on the data.

CHAPTER III

RESULTS

We have previously reported on the use of biophysical techniques including CD, DSC, and ITC for the characterization of other non-histone proteins and their binding interactions with various DNA ligands however this is, to the best of our knowledge, the first report using microcalorimetric and spectroscopic techniques to probe the binding interactions between two histone protein variants $H1_1$ and $H1_4$ and their interaction with dsDNA [33, 35, 40-42]. The results of the biophysical characterization of the histone $H1_1$ and $H1_4$ proteins and their interactions with CTDNA present a consistent picture regarding the structure and the stability of the proteins and DNA alone as well as their structure upon the formation of a stable complex are described in this chapter.

The $H1_1$ and $H1_4$ protein structures and the CTDNA structures are grossly altered upon the formation of the CTDNA/histone complex. Figure 3.1 shows the representative CD spectra for the $H1_1$ histone protein alone and CTDNA alone as well as the spectrum for the 1:1 complex of CTDNA/$H1_1$ protein. The CTDNA spectrum shows a positive molar ellipticity at 280 nm (maximum) and a negative molar ellipticity at 240 nm that is consistent with CD spectra previously reported for CTDNA [43, 44]. The $H1_1$ protein spectrum exhibits a negative molar ellipticity at 208 and 220 nm which is indicates a large amount of α-helix is present in the structure of the histone protein [27]. Although the spectra for β turns are relatively weaker than those of α helix, a characteristic negative band is observed at 210-225 nm which is consistent with the three β turns in the $H1_1$

protein. The CD spectrum obtained for the $H1_1$ protein indicates approximately 70% alpha helices and 30% β turns which is in agreement with the previously reported NMR structures for the H1 variants [8, 17, 45]. In comparison to the spectrum for the CTDNA and $H1_1$ protein alone, the CD spectra for the CTDNA/$H1_1$ complex exhibits a notable attenuation of the 280 nm and the 220 nm ellipticities. A significant loss in secondary structure of both the $H1_1$ histone and the CTDNA are observed upon the formation of the 1:1 complex.

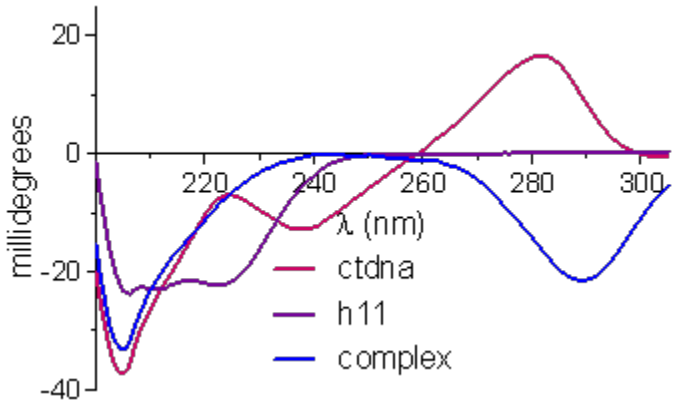


Figure 3.1 Circular Dichroism spectra of Histone $H1_1$ (purple), CTDNA (pink), and 1:1 complex (blue) in 130 mM K^+ BPES at pH 7.0.

Figure 3.2 shows the representative CD spectra for the $H1_4$ histone protein alone and CTDNA alone as well as the spectrum for the 1:1 complex of CTDNA/$H1_4$ protein. The $H1_4$ protein spectrum exhibits a negative molar ellipticity at 208 and 220 nm which is indicative of alpha-helical content and a negative band at 210-225 nm which is indicative of a β turn. Based on the obtained CD spectra, the $H1_4$ protein is very similar to the $H1_1$ consisting of approximately 70% alpha helices and 30% β turns [8, 17, 45, 46]. In comparison to the spectra for the CTDNA and $H1_4$ protein alone, the CD spectra for the CTDNA/$H1_4$ complex exhibits an attenuation of both the 280 nm maximum and the 220

nm minimum molar ellipticities. A significant loss in secondary structure of both the $H1_4$ protein and the CTDNA are observed upon the formation of the 1:1 complex.

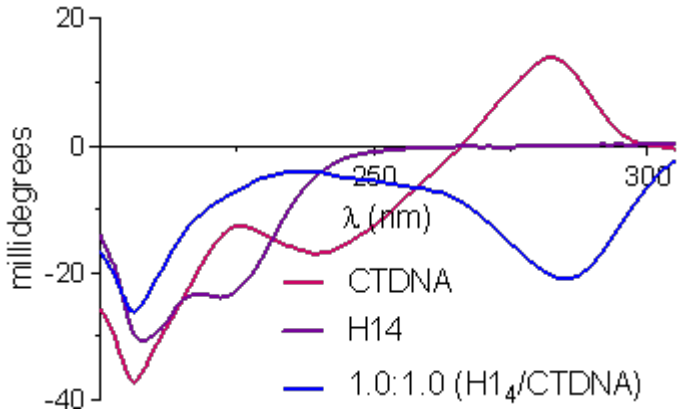


Figure 3.2 Circular Dichroism spectra of Histone $H1_4$ (purple), CTDNA (pink), and 1:1 complex (blue) in 130 mM K^+ BPES at pH 7.0

The results of the DSC and CD experiments present a consistent picture regarding the structure and stability of the H1 variants, the CTDNA, and the histone/DNA complex. The DSC melting profile for the thermal denaturation of the CTDNA alone as well as the DNA complexed with $H1_4$ is shown in Figure 3.3. The CTDNA thermogram is comprised of a predominantly symmetric peak that has been fit for a single "two state" transition with a T_m value of 84.0 °C as shown in Figure 3.3(A).

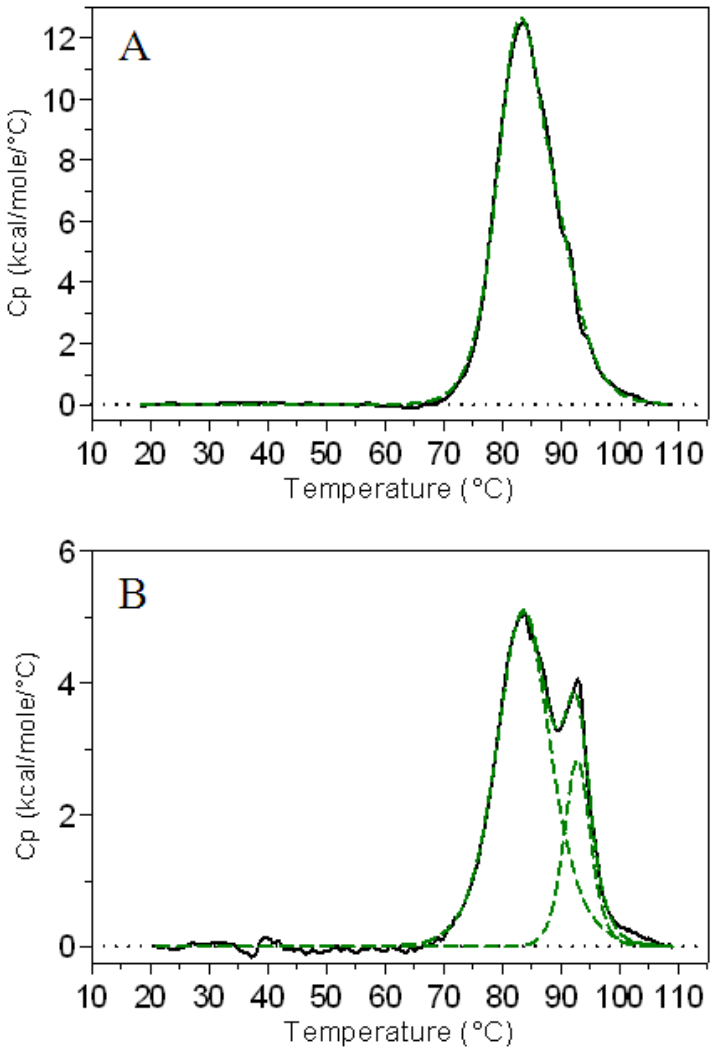


Figure 3.3 A) DSC thermogram for the thermal denaturation of CTDNA. The raw excess heat capacity (*black line*) has been deconvoluted into a single "two-state" process (*dashed green line*). B) DSC thermogram for the thermal denaturation of Histone $H1_4$ with 1:5 mols of CTDNA in 130mM K^+ BPES. The raw excess heat capacity has been deconvoluted into two independent two-state processes. The raw excess heat capacity (*black line*) has been deconvoluted into two independent overlapping "two-state" processes (*dashed green line*). The lower melting profile is attributed to "free" CTDNA while the higher melting transition is for the melting of the $H1_4$/CTDNA stabilized complex

Figure 3.3 (B) shows the melting transitions for the CTDNA in 130 mM K^+ BPES at pH 7.0 after the addition of 1:5 moles of histone $H1_4$ protein/CTDNA. The thermogram for the $H1_4$/CTDNA complex is comprised of two asymmetric peaks that

have been best fit for two independent overlapping "two state" processes with *T*m values of 83.7 °C and 92.8 °C. The lower melting profile (83.7 °C) is attributed to the unbound CTDNA in solution and the higher melting profile (92.8 °C) is attributed to the CTDNA stabilized by $H1_4$ (CTDNA/$H1_4$ complex). In comparison to the melting curve for the CTDNA alone in 130 mM K^+ BPES at pH 7.0 the following changes are observed: 1) there is a notable reduction in the relative peak area (or enthalpy change) for the CTDNA alone and 2) a new thermal transition appears with a T_m of 92.8 °C presumably due to the melting of the CTDNA/$H1_4$ complex. The T_m values for the CTDNA alone as well as for the complex containing 1:5 moles of histone $H1_4$ protein/CTDNA are listed in Table 3.1.

Table 3.1 Tm values for the CTDNA and CTDNA/$H1_4$ complex at 130 mM K^+ BPES pH 7.0.

	T_m1	T_m2
CTDNA	84.0	NA
CTDNA/ $H1_4$ Complex (5:1)	83.7	92.8

Figures 3.4 and 3.5 show data from ITC titrations where histone $H1_1$ was titrated into CTDNA at 15 °C and 25 °C, respectively. Experiments performed at 35 °C did not produce large enough heats upon binding to be properly fit within experimental error. Figure 3.4 shows the integrated heat data and the best fit line for a single titration at 15 °C; however, a minimum of three replicate titrations were completed at both 15 °C and 25 °C. The nonlinear regression fit for the experiment performed at 15 °C (Fig. 3.4 solid red line) is from a two-independent sites model and the average best-fit parameters are listed in Table 3.2.

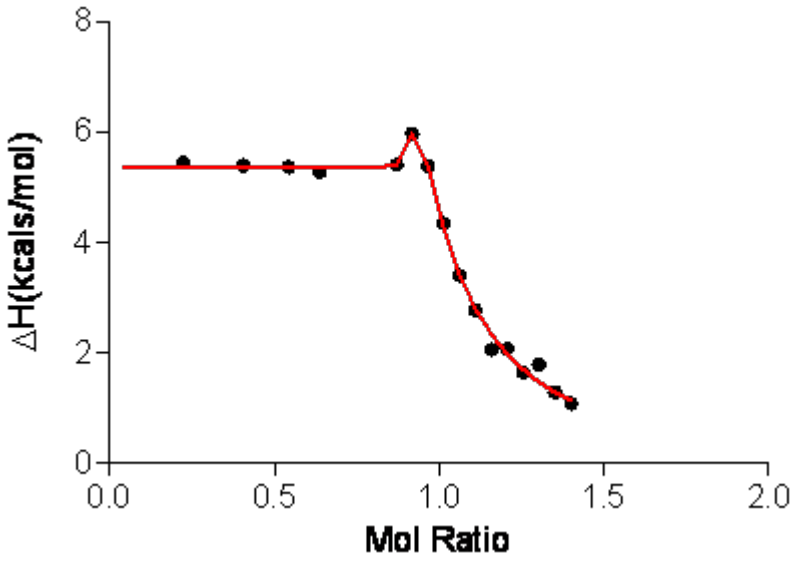


Figure 3.4 ITC data for the addition of $H1_1$ to CTDNA at 15 °C in 130 mM K^+ BPES (pH 7.0) including the integrated heat data (●) and the best-fit non-linear regression line (–). The best-fit nonlinear regression line is for the thermodynamic parameters obtained using a two-independent sites model with two values for K_i, two values for ΔH_i, and two values for n_i.

The nonlinear regression fit for the experiment performed at 25 °C (Fig. 3.5 solid red line) is from a two-independent sites model, and the average best-fit parameters are listed in Table 3.2. The best fit parameters highlighted in yellow in Table 3.2 are not accurate because they were calculated based on a single point in the titration. The large standard deviations indicate a large amount of experimental error in these calculations. The maximum point seen in every $H1_1$ and $H1_4$ titration experiment immediately before the endpoint of the titration essentially determines the best-fit parameters for the second binding mode. Due to the fact this single point is seen in every experiment it cannot be attributed to experimental/instrumental error. The peak observed directly before the end point of the titration is attributed to the structural rearrangement occurring not only in the DNA but the protein as well to accommodate binding of the $H1_1$ protein at all available binding sites.

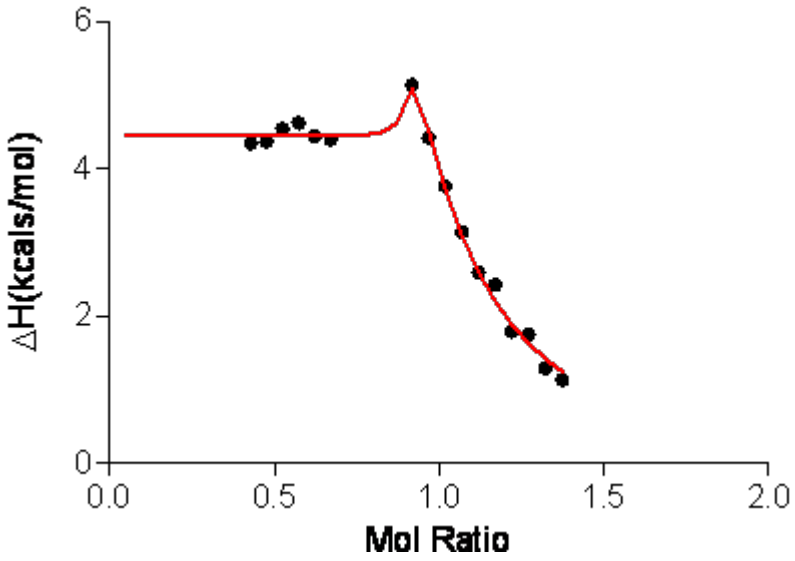


Figure 3.5 ITC data for the addition of $H1_1$ to CTDNA at 25 °C in 130 mM K^+ BPES (pH 7.0) including the integrated heat data (●) and the best-fit non-linear regression line (–). The best-fit nonlinear regression line is for the thermodynamic parameters obtained using a two-independent sites model with two values for K_i, two values for ΔH_i, and two values for n_i.

The values for K_1 are noted in red in Table 3.2. for both temperatures (15 °C and 25 °C) involve standard deviations that are as large or larger than the values determined for K_1. These calculated binding constants are not good approximations for $H1_1$/DNA complex. The raw data were very poor and the integrated heat values were small which required removal of a number of bad data points because of noisy baseline values.

Table 3.2 ITC-derived thermodynamic parameters for $H1_1$ histone protein binding to CTDNA in 130 mM K^+ BPES pH 7.0.

Temperature	n_1	K_1 (M^{-1}) $\times 10^{-8}$	ΔG_1 (kcal/mol)	ΔH_1 (kcal/mol)	$-T\Delta S_1$ (kcal/mol)	n_2	K_2 (M^{-1}) $\times 10^{-5}$	ΔG_2 (kcal/mol)	ΔH_2 (kcal/mol)	$-T\Delta S_2$ (kcal/mol)
15°C	**0.9**	6.69±8.92	-11.6	5.4± 0.1	-17	0.1	2.7 ± .46	-7.15	456± 276	-464
25°C	**0.9**	3.4±3.03	-11.6	4.4 ± 0.0	-16	0.1	1.79± 0.97	-7.2	12.8 ± 35.7	-20

Figures 3.6 through 3.8 show data from ITC titrations where CTDNA was titrated with histone $H1_4$ at 15 °C, 25 °C, and 35 °C. Figure 3.6 shows the integrated heat data

and the best fit line for a single titration at 15 °C; however, a minimum of three replicate titrations were completed at each temperature. The nonlinear regression fit (Fig. 3.6 solid red line) is from a two-independent sites model and the average best-fit parameters calculated are listed in Table 3.3.

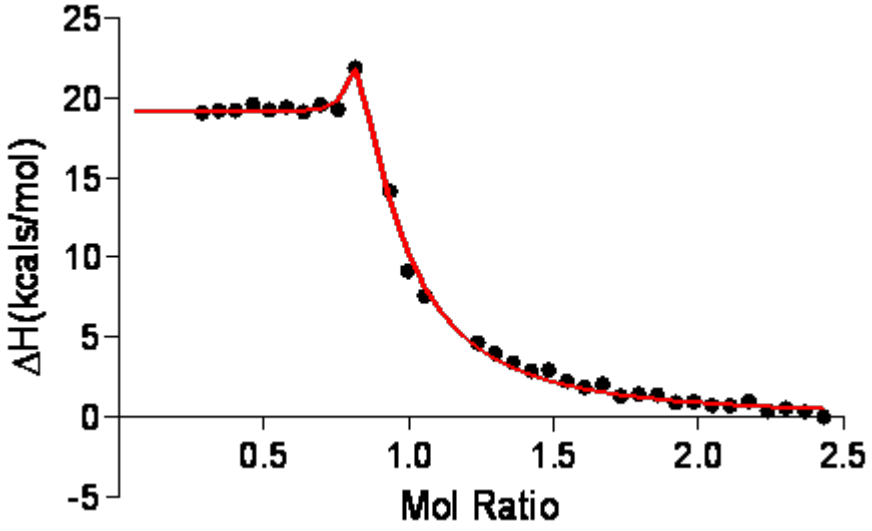
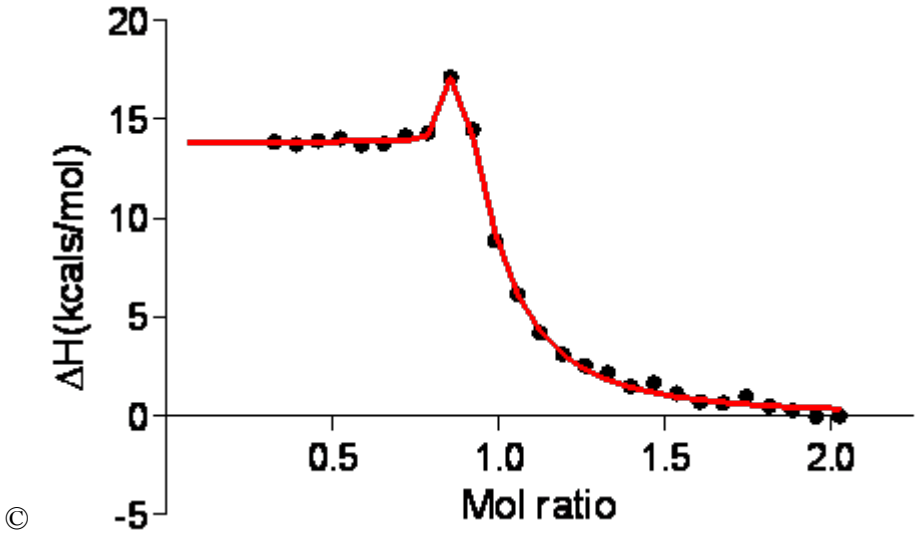


Figure 3.6 ITC data for the addition of $H1_4$ to CTDNA at 15 °C in 130 mM K^+ BPES (pH 7.0) including the integrated heat data (●) and the best-fit non-linear regression line (−). The best-fit nonlinear regression line is for the thermodynamic parameters obtained using a two-independent sites model with two values for K_i, two values for ΔH_i, and two values for n_i.

The ITC data collected at all three temperatures (15 °C, 25 °C, and 35 °C) demonstrate a binding stoichiometry of 1:1 for the $H1_4$/CTDNA complex. However, the microcalorimetric signal and integrated heat data are more complex and are consistent with at least two independent binding processes occurring during complex formation before saturation. Performing titrations over a range of temperatures allowed for us to estimate the ΔCp for binding and further probe the interaction between the histone protein $H1_4$ and CTDNA which is highly endothermic. Figures 3.7 and 3.8 show the integrated heat data and the best fit lines for a single titration at 25 °C and 35 °C, again a

minimum of three replicate titrations were completed at these temperatures with the averaged best-fit parameters listed in Table 3.3.



©

Figure 3.7 ITC data for the addition of $H1_4$ to CTDNA at 25 °C in 130 mM K^+ BPES (pH 7.0) including the integrated heat data (●) and the best-fit non-linear regression line (–). The best-fit nonlinear regression line is for the thermodynamic parameters obtained using a two-independent sites model with two values for K_i, two values for ΔH_i, and two values for n_i.

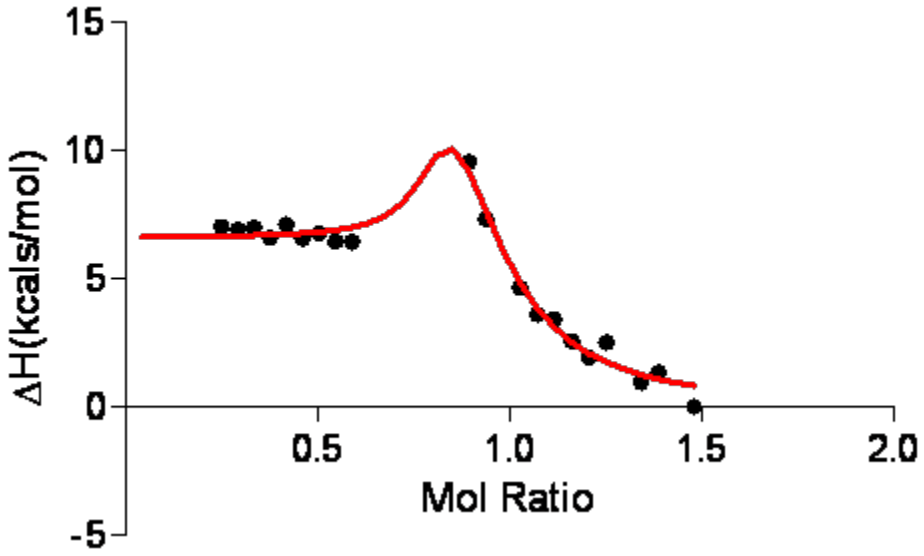


Figure 3.8 ITC data for the addition of $H1_4$ to CTDNA at 35 °C in 130 mM K^+ BPES (pH 7.0) including the integrated heat data (●) and the best-fit non-linear regression line (–). The best-fit nonlinear regression line is for the thermodynamic parameters obtained using a two-independent sites model with two values for K_i, two values for ΔH_i, and two values for n_i.

Table 3.3 ITC-derived thermodynamic parameters for $H1_4$ histone protein binding to CTDNA in 130 mM K^+ BPES pH 7.0.

Temperature	n_1	K_1 (M^{-1}) $\times 10^{-8}$	ΔG_1 (kcal/mol)	ΔH_1 (kcal/mol)	$-T\Delta S_1$ (kcal/mol)	n_2	K_2 (M^{-1}) $\times 10^{-5}$	ΔG_2 (kcal/mol)	ΔH_2 (kcal/mol)	$-T\Delta S_2$ (kcal/mol)
15°C	**0.8**	5.8 ± 1.0	-11.5	20.5 ± 1.8	-32.0	0.1	9.4 ± 7.9	-7.9	68.6± 40.8	-48.7
25°C	**0.8**	2.2 ± 1.2	-11.4	15.6 ± 1.9	-27.0	0.1	.3.6 ± 2.8	-8.9	51.4 ± 18.7	-60.3
35°C	**0.9**	1.3 ± 0.01	-11.4	5.4 ± 1.7	-16.8	0.1	1.48 ± 1.0	-8.7	28.1 ± 11.7	-36.8

Previous experiments indicated binding site sizes of 30 DNA bp for $H1_1$ and 36 DNA bp for $H1_4$. Using DNA concentrations based on 30 DNA bp or 36 DNA bp, the ITC experiments produced a binding stoichiometry of 1:1 for the H1 protein/DNA complex at saturation. The temperature studies allowed us to estimate the ΔCp for the binding of $H1_4$ to CTDNA and probe the role of water in the complex formation. The heat-capacity changes were calculated by plotting the ΔH_i values versus temperature and fitting the data with a simple linear-regression line for each binding mode. The heat-

capacity change calculated for ΔCp_1 was -755 cal/mol K. As mentioned previously, the calculated ΔCp_2 for the second binding mode was not accurate therefore we did not report this value here. The ΔCp_1 value is large and indicates that there is a large net change in structure upon binding.

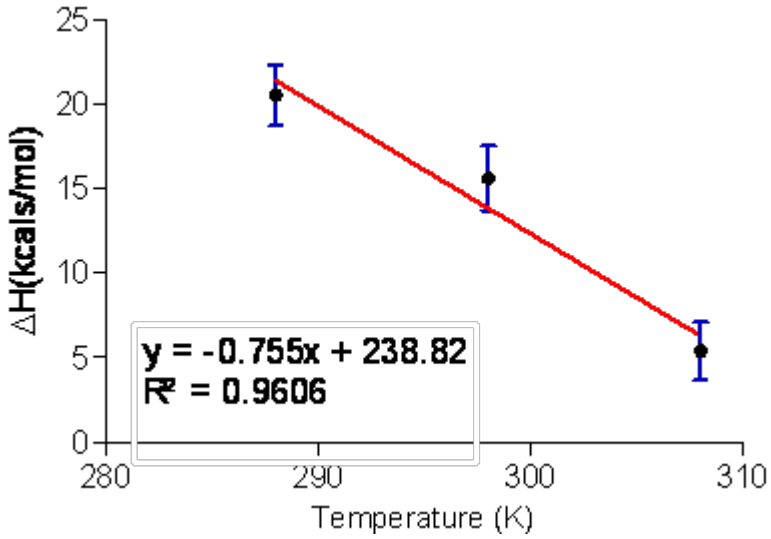


Figure 3.9 Plot of ΔH_1 vs. Temperature (*K*)

CHAPTER IV

DISCUSSION

As stated in the introduction to this thesis, this project attempts to answer several questions regarding the interaction of two different histone H1 variants with its physiological binding partner, dsDNA. The three main objectives were: 1) to determine the thermodynamic signatures for the binding of $H1_1$ and $H1_4$ to CTDNA, 2) to evaluate whether the binding of these two linker proteins to dsDNA was accompanied by a change in either the protein or DNA structure, and 3) to attempt to attribute any changes in binding affinity, binding site size, or binding forces to the changes in the $H1_1$ and $H1_4$ sequences.

Results from a McGhee-von Hippel binding model described previously by Mammoon *et al* were used to establish a theoretical equilibrium constant and the DNA binding site size for binding of histone $H1_1$ and $H1_4$ variants to dsDNA [8, 47, 48]. From these previous reports we were also able to compare some of the sequence and structural differences between the $H1_1$ and $H1_4$ histone proteins and attempt to describe how these minor structural differences affect their binding to dsDNA. The $H1_1$ sequence includes 56 lysines whereas $H1_4$ includes 60 lysines [11, 12, 18]. The increased number of lysines in $H1_4$ in comparison to $H1_1$ is responsible for the higher positive charge for the $H1_4$ histone variant [8, 47, 48].

De *et al* described that the $H1_1$ globular domain is comprised of three α helices connected by beta loops, and the N- and C-terminus are extended without the presence of

the physiological binding partner (i.e. dsDNA) [8, 46, 47]. Our $H1_1$ and $H1_4$ histone variant CD spectra exhibit a broad negative peak (minimum) at 220 nm, which is in agreement with previously reported data for $H1_1$ [11, 18]. Our spectroscopic data is consistent with Mammoon *et al's* reports that the structures of these histone variants are similar [8, 47, 48].

When protein was titrated into the CTDNA, gross structural changes were observed for both the protein and the CTDNA. The peak attributed to the CTDNA seen at 280 nm and the peak attributed to the histone protein variants seen at 220 nm were both attenuated upon the formation of the H1/CTDNA complex [27, 46]. The spectral changes observed for both the DNA and the histone $H1_1$ and $H1_4$ variants are attributed to structural changes in both the CTDNA and the histone $H1_1$ and $H1_4$ variants. From the attenuation of the spectra in the 220 and 280 nm regions we hypothesize the major structural rearrangements occur from 1) the unfolding of the α-helical domains in the central globular domains of the $H1_1$ and $H1_4$ histone proteins, and 2) the restructuring of the CTDNA as it is bound to the protein [27, 43, 44, 46].

The structural stabilization of the $H1_4$ protein/DNA complex was further characterized using DSC. A single (*two state*) transition with a temperature of 84° C was observed for the denaturation of the CTDNA in the absence of the protein. After the addition of 1:5 moles of $H1_4$ protein per mole of DNA, a second melting profile appeared with a T_m value of approximately 93° C. This higher melting transition is assumed to be the result of the complex formed between the CTDNA and $H1_4$ protein. The peak area of the higher melting transition is approximately 1/5 of the area for CTDNA alone which is consistent with 20% of the binding sites being occupied upon formation of the complex.

The 9° C shift in the melting temperature of the DNA is consistent with a binding constant of 10^8 which is consistent with the ITC measured affinities for the $H1_4$ protein.

Previous experiments indicated binding site sizes of 30 DNA bp and 36 DNA bp for $H1_1$ and $H1_4$ respectively. The concentrations of ds DNA were calculated assuming a 30 base pair ($H1_1$) and 36 base pair ($H1_4$) binding site. With this ds DNA binding site size, the calculated saturation stoichiometry for the binding of CTDNA to both the $H1_1$ and $H1_4$ proteins was 1:1 mole of H1 protein/ds DNA. These results were consistent with the CD data that exhibited no further changes in the spectra after 1:1 saturation of H1 protein/ds DNA.

ITC results demonstrated that there were two independent binding processes that occurred upon the addition of CTDNA to the histone proteins. The first endothermic binding process is attributed to the binding of CTDNA to the histone proteins; whereas, the second process is attributed to the subsequent rearrangement of the protein molecules already bound to the DNA near saturation [34, 36]. The large entropic term is presumed to be a result of the structural changes occurring in both the DNA and the protein. It can also be inferred that the release of water molecules and ions from the minor groove of the DNA is a contributing factor to the unfavorable enthalpy and favorable entropy as well. The change in structure for both the H1 protein and the ds DNA is supported by our spectroscopic data and previous reports regarding the activity of the N and C-Terminus upon binding to the physiological binding partner, ds DNA [8, 13, 16, 47].

The binding constant calculated for the first binding process was approximately $K_1 \approx 10^8$ and the binding constant for the second binding process was approximately $K_2 \approx 10^6$. When the protein first binds to the DNA, the protein binds with high affinity and randomly spaced along the DNA. The binding of the protein itself to the ds DNA is not

specific (i.e. not targeting a specific DNA sequence or region). When the protein concentration approaches DNA saturation, the protein is essentially "rearranged" and binds to the ds DNA in a manner which accommodates all of the ds DNA sites. This suggests that all sites are filled with no overlap of the protein from one site to another. This is consistent with the nearest neighbor exclusion theory, which states that every binding site will not be occupied assuming a 15 ($H1_1$) and 18 ($H1_4$) base pair binding site because the protein is too large, therefore there is one protein per every 30 and 36 base pair although not all base pairs in the binding site will be occupied by the protein.

The accuracy of the best-fit parameters highlighted in *yellow* within Tables 3.3 ($H1_1$) and 3.4 ($H1_4$) for the second binding process (the structural rearrangement) cannot be trusted because they have been calculated based on a single point in the titration. The maximum point seen in every $H1_1$ and $H1_4$ titration experiment immediately before the endpoint of the titration essentially determines the best-fit parameters for the second binding mode [36]. Due to the fact that this single point is seen in every experiment it cannot be attributed to experimental/instrumental error. The peak observed directly before the end point of the titration is attributed to the structural rearrangement occurring not only in the DNA, but the protein as well to accommodate binding of the $H1_1$ protein at all available binding sites.

CHAPTER V

CONCLUSION

This is the first study of its kind; many studies have been performed on Histone H1 variants, however previously it has been difficult to collect good calorimetric data for these proteins. The large amount of protein necessary for calorimetric studies has previously been a limiting factor for data collection. Improvements in the extraction procedure and instrumentation have made these studies possible. It was expected that the binding of histone H1 to ds DNA would be exothermic. However, the isothermal titration data indicated an endothermic binding process for the reaction. The endothermic values are attributed to the structural changes associated with both the DNA and the protein. Since it is hypothesized that the protein binds to the minor grove of the DNA, it is possible that the expulsion of water and ions plays a role in the large positive enthalpy and large entropy values associated with the binding process.

We have established that $H1_1$ and $H1_4$, although similar in their sequence and structure, exhibit different, their thermodynamic binding properties. $H1_4$ binds more tightly to double stranded DNA; therefore $H1_4$ would make a better model of study for future experiments focused on drug design and therapeutic methods for diseases such as cancer. We have demonstrated that the binding site for $H1_1$ is 30 base pairs whereas the binding site of $H1_4$ is 36 base pairs. The strength of this study lies in that the binding of the $H1_1$ and $H1_4$ histone proteins to CTDNA induces structural changes in both the DNA and the proteins. This structural information can be used in the future to design and study

the interactions between linker histones, core histones, and DNA. A better understanding of the functional properties of H1 and its interactions with DNA could provide new opportunities for cancer therapies and drug design. If the binding site size is known, drug design may be possible to prevent the histone modifications seen in some cancer cells. A drug specific to the binding site size of these proteins may be useful in the regulation of transcription by $H1_1$ and $H1_4$ histone proteins.

REFERENCES

1. Andrews, A.J. and K. Luger, *Nucleosome Structure(s) and Stability: Variations on a Theme.* Annu Rev Biophys.

2. Woodcock, C.L. and R.P. Ghosh, *Chromatin higher-order structure and dynamics.* Cold Spring Harb Perspect Biol. 2(5): p. a000596.

3. Zlatanova, J., et al., *The nucleosome family: dynamic and growing.* Structure, 2009. 17(2): p. 160-71.

4. Woodcock, C.L., *Chromatin architecture.* Curr Opin Struct Biol, 2006. 16(2): p. 213-20.

5. McBryant, S.J., X. Lu, and J.C. Hansen, *Multifunctionality of the linker histones: an emerging role for protein-protein interactions.* Cell Res. 20(5): p. 519-28.

6. Duggan, M.M. and J.O. Thomas, *Two DNA-binding sites on the globular domain of histone H5 are required for binding to both bulk and 5 S reconstituted nucleosomes.* J Mol Biol, 2000. 304(1): p. 21-33.

7. Ali, T., et al., *Two homologous domains of similar structure but different stability in the yeast linker histone, Hho1p.* J Mol Biol, 2004. 338(1): p. 139-48.

8. De, S., et al., *Histone H1 variants differentially inhibit DNA replication through an affinity for chromatin mediated by their carboxyl-terminal domains.* Gene, 2002. 292(1-2): p. 173-81.

9. Banks, G.C., et al., *Hormone-mediated dephosphorylation of specific histone H1 isoforms.* J Biol Chem, 2001. 276(39): p. 36467-73.

10. Mamoon, N.M., Y. Song, and S.E. Wellman, *Binding of histone H1 to DNA is described by an allosteric model.* Biopolymers, 2005. 77(1): p. 9-17.

11. Mamoon, N.M., Y. Song, and S.E. Wellman, *Histone h1(0) and its carboxyl-terminal domain bind in the major groove of DNA.* Biochemistry, 2002. 41(29): p. 9222-8.

12. Wellman, S.E., *Carboxyl-terminal peptides from histone H1 variants: DNA binding characteristics and solution conformation.* Biopolymers, 1996. 39(4): p. 491-501.

13. Wellman, S.E., *Differences in the Interaction of histone variants, H11, H10, and H14, with DNA.* not submitted, 2011

14. Wellman, S.E., D.B. Sittman, and J.B. Chaires, *Preferential binding of H1e histone to GC-rich DNA.* Biochemistry, 1994. 33(1): p. 384-8.

15. Wellman, S.E., Y. Song, and N.M. Mamoon, *Sequence preference of mouse H1(0) and H1t.* Biochemistry, 1999. 38(40): p. 13112-8.

16. Wellman, S.E., *Differenes in the Interaction of histone variants H11, H10, and H14 with DNA.* not submitted, 2011.

17. Cerf, C., et al., *Homo- and heteronuclear two-dimensional NMR studies of the globular domain of histone H1: full assignment, tertiary structure, and comparison with the globular domain of histone H5.* Biochemistry, 1994. 33(37): p. 11079-86.

18. Wellman, S.E., et al., *Purification of mouse H1 histones expressed in Escherichia coli.* Biotechnol Appl Biochem, 1997. 26 (Pt 2): p. 117-23.

19. Lee, E.D., et al., *High-performance liquid chromatographic chiral stationary phase separation with filament-on thermospray mass spectrometric identification of the enantiomer contaminant (S)-(+)-methamphetamine.* Anal Chem, 1986. 58(7): p. 1349-52.

20. Xiang, Y., Y. Liu, and M.L. Lee, *Ultrahigh pressure liquid chromatography using elevated temperature.* J Chromatogr A, 2006. 1104(1-2): p. 198-202.

21. Horvath, C.G., B.A. Preiss, and S.R. Lipsky, *Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers.* Anal Chem, 1967. 39(12): p. 1422-8.

22. Bendriss, E.K., N. Markoglou, and I.W. Wainer, *High-performance liquid chromatography assay for simultaneous determination of dextromethorphan and its main metabolites in urine and in microsomal preparations.* J Chromatogr B Biomed Sci Appl, 2001. 754(1): p. 209-15.

23. Shapiro, A.L., E. Vinuela, and J.V. Maizel, Jr., *Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels.* Biochem Biophys Res Commun, 1967. 28(5): p. 815-20.

24. Schagger, H. and G. von Jagow, *Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa.* Anal Biochem, 1987. 166(2): p. 368-79.

25. Man, B.Y., et al., *A selective G-quadruplex-based luminescent switch-on probe for the detection of nanomolar silver(I) ions in aqueous solution.* Chem Commun (Camb), 2010. 46(45): p. 8534-6.

26. Kang, J.S. and J.R. Lakowicz, *Fluorescence Resonance Energy Transfer in Calf Thymus DNA from a Long-Lifetime Metal-Ligand Complex to Nile Blue.* Journal of Biochemistry and Molecular Biology, 2001. 34(6): p. 551-558.

27. Martin, S.R. and M.J. Schilstra, *Circular dichroism and its application to the study of biomolecules.* Methods Cell Biol, 2008. 84: p. 263-93.

28. Kanehara, H., et al., *Spectroscopic evidence for the formation of four-stranded solution structure of oligodeoxycytidine phosphorothioate.* Biochemistry, 1997. 36(7): p. 1790-7.

29. Li, W., et al., *Structural competition involving G-quadruplex DNA and its complement.* Biochemistry, 2003. 42(40): p. 11736-44.

30. Kaushik, M., et al., *Possibility of an antiparallel (tetramer) quadruplex exhibited by the double repeat of the human telomere.* Biochemistry, 2007. 46(24): p. 7119-31.

31. Paramasivan, S., I. Rujan, and P.H. Bolton, *Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding.* Methods, 2007. 43(4): p. 324-31.

32. Spink, C.H., *Differential scanning calorimetry.* Methods Cell Biol, 2008. 84: p. 115-41.

33. Freyer, M.W., et al., *Biophysical studies of the c-MYC NHE III1 promoter: model quadruplex interactions with a cationic porphyrin.* Biophys J, 2007. 92(6): p. 2007-15.

34. Freyer, M.W., et al., *Binding of netropsin and 4,6-diamidino-2-phenylindole to an A2T2 DNA hairpin: a comparison of biophysical techniques.* Anal Biochem, 2006. 355(2): p. 259-66.

35. Nagesh, N., et al., *Studies on the site and mode of TMPyP4 interactions with Bcl-2 promoter sequence G-Quadruplexes.* Biophys J, 2010. 98(11): p. 2628-33.

36. Freyer, M.W. and E.A. Lewis, *Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions.* Methods Cell Biol, 2008. 84: p. 79-113.

37. Wiseman, T., et al., *Rapid measurement of binding constants and heats of binding using a new titration calorimeter.* Anal Biochem, 1989. 179(1): p. 131-7.

38. Weber, P.C. and F.R. Salemme, *Applications of calorimetric methods to drug discovery and the study of protein interactions.* Curr Opin Struct Biol, 2003. 13(1): p. 115-21.

39. Lewis, E.A., *Calorimetry*, in *Encylcopedia of Life Sciences*. 2009, John Wiley & Sons, Ltd. p. 1-6.

40. Dettler, J.M., et al., *Biophysical Characterization of an Ensemble of Intramolecular i-Motifs Formed by the Human c-MYC NHE III1 P1 Promoter Mutant Sequence.* Biophysical Journal, 2010. 99: p. 1-7.

41. Dettler, J.M., et al., *DSC Deconvolution of the Structural Complexity of c-MYC P1 Promoter G-Quadruplexes.* Biophys J, 2011. 100(6): p. 1517-25.

42. Cashman, D.J., et al., *Molecular modeling and biophysical analysis of the c-MYC NHE-III1 silencer element.* J Mol Model, 2008. 14(2): p. 93-101.

43. Gray, D.M., A.R. Morgan, and R.L. Ratliff, *A comparison of the circular dichroism spectra of synthetic DNA sequences of the homopurine . homopyrimidine and mixed purine- pyrimidine types.* Nucleic Acids Res, 1978. 5(10): p. 3679-95.

44. Hirschman, S.Z. and G. Felsenfeld, *Determination of DNA composition and concentration by spectral analysis.* J Mol Biol, 1966. 16(2): p. 347-58.

45. Latha, P., et al., *Characterization of histone (H1B) oxalate binding protein in experimental urolithiasis and bioinformatics approach to study its oxalate interaction.* Biochem Biophys Res Commun, 2006. 345(1): p. 345-54.

46. Higashibata, H., et al., *Surface histidine residue of archaeal histone affects DNA compaction and thermostability.* FEMS Microbiol Lett, 2003. 224(1): p. 17-22.

47. Lu, Z.H., et al., *Histone H1 modulates DNA replication through multiple pathways in Xenopus egg extract.* J Cell Sci, 1997. 110 (Pt 21): p. 2745-58.

48. Clark, D.J. and J.O. Thomas, *Differences in the binding of H1 variants to DNA. Cooperativity and linker-length related distribution.* Eur J Biochem, 1988. 178(1): p. 225-33.