1-1-2014

Thermodynamic Characterization of Linker Histone Binding Interactions with ds-DNA

Venkata Ramana Machha

Thermodynamic characterization of linker histone binding interactions with ds-DNA

By

Venkata Ramana Machha

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Chemistry
in theDepartment of Chemistry

Mississippi State, Mississippi

May 2014

Thermodynamic characterization of linker histone binding interactions with ds-DNA

By

Venkata Ramana Machha

Approved:

_____

Edwin A. Lewis
(Major Professor)

_____

Joseph P. Emerson
(Committee Member)

_____

Nicholas Fitzkee
(Committee Member)

_____

Steven R. Gwaltney
(Committee Member)

_____

Gerald Rowland
(Committee Member)

_____

R. Gregory Dunaway
Dean
College of Arts & Sciences

Name: Venkata Ramana Machha

Date of Degree: May 17, 2014

Institution: Mississippi State University

Major Field: Chemistry

Major Professor: Edwin A. Lewis

Title of Study: Thermodynamic characterization of linker histone binding interactions with ds-DNA

Pages in Study: 96

Candidate for Degree of Doctor of Philosophy

Linker histones (H1) are the basic proteins in higher eukaryotes that are responsible for the final condensation of chromatin. H1 also plays an important role in regulating gene expression. H1 has been described as a transcriptional repressor as it limits the access of transcriptional factors to DNA. Linker histone binds to DNA that enters or exits the nucleosome. Several crystal structures have been published for the nucleosome (histone core/DNA complex), and the interactions of the core histone proteins with DNA are well understood. In contrast the location of the linker histone and its interactions with ds-DNA are poorly understood.

In this study we have used isothermal titration calorimetry (ITC), differential scanning calorimetry (DSC), and CD spectropolarimetry to determine the thermodynamic signatures and structural changes that accompany H1 binding to ds-DNA. The thermodynamic parameters for the binding of intact linker histones (H1.1, H1.4, and $H1^0$) to highly polymerized calf-thymus DNA and to short double stranded DNA oligomers have been determined. We have also determined the thermodynamics for binding of $H1^0$ C-terminal tail ($H1^0$-C) and globular domain ($H1^0$-G) to calf-thymus DNA. The real

surprise in the energetics is that the enthalpy change for formation of the H1/DNA complex is very unfavorable and that H1/DNA complex formation is driven by very large positive changes in entropy. The binding site sizes for H1.1, H1.4, and $H1^0$ were determined to be 36bp, 32bp, and 36bp respectively. CD results indicate that CT-DNA is restructured upon complexation with either the full length H1 protein ($H1^0$) or its C-terminal domain ($H1^0$-C). In contrast, the structure of $H1^0$ is largely unchanged in the DNA complex. Temperature dependence of enthalpy change, osmotic stress and ionic strength dependence of K_a were tested using ITC. These results indicate that the entropy driven H1/DNA complexes are a result primarily from the expulsion of bound water molecules from the binding interface.

This study provides new insights into the binding of linker Histone H1 to DNA. A better understanding of the functional properties of H1 and its interactions with DNA could provide new insights in understanding the role H1 in DNA condensation and transcriptional regulation.

Key words: Histone, H1.1, H1.4, $H1^0$, $H1^0$-C, $H1^0$-G, CT-DNA, chromatin, chromosome, nucleosome, ITC, CD, DSC

DEDICATION

I dedicate this PhD Dissertation to Lord Balaji for his blessings, my parents Babu Machha and Lakshmi Devi Machha for their unconditional love and support. I also dedicate this to my brother Sunil kumar Machha for his constant encouragement and also to my Loving wife Vinuthna Machha for her faithful support even during tough times in my Ph.D. pursuit.

ACKNOWLEDGEMENTS

I want to thank my advisor Dr. Edwin Lewis for his help and guidance. He knew where to look for answers for the problems and lead me to right perspective, and I cannot thank him enough for sharing ideas and giving precious and kind advice at the right time.

I very much appreciate my committee members Dr. Joseph Emerson, Dr. Nicholas Fitzkee, Dr. Gerald Rowland and Dr. Steven Gwaltney for reviewing my dissertation and giving their feedback and support. I would like to thank Dr. Emerson and his lab members for allowing me to use their instruments whenever I needed. I would like to thank Dr. Susan Wellman from University of Mississippi Medical Center for collaborating with me and providing advice many times. I thank Ms. Laura Lewis. Her optimism and enthusiasm are contagious. I thank her for being so supportive and caring person. I thank past and present Lewis group members that I have had the pleasure to work with or along, and especially I would like to thank Vu Le for working with me and helping in different aspects of my research.

I would like to thank my dear friend David Wilson for his encouragement and especially his humor made me laugh and keep things light. I would like to acknowledge the staff, faculty, and students of the Chemistry Department at MSU.

Last but not least, I would like to thank my family for all their love and encouragement. I thank my parents for instilling good principles and my dad who taught me about hard work and persistence. I thank my brother Sunil Kumar Machha, my sister

in law Manasa Neelam, brother in law Ajay Karthik Neelam, and my in-laws for their moral support. I also thank my wife Vinuthna Machha for her support in all aspects and sharing many happy and stressful moments throughout this journey. Thank you for believing in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CD	Circular dichroism
CT-DNA	Calf-thymus DNA
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
EDTA	Ethylenediaminetetraacetic acid
$\Delta G°$	Gibbs free energy
$\Delta H°$	Enthalpy change
ITC	Isothermal titration calorimetry
K_a	Equilibrium association constant
$\Delta S°$	Entropy change
TEG	Triethylene glycol
MD	Molecular dynamics

CHAPTER I

INTRODUCTION

1.1 General Introduction

In eukaryotes the genomic DNA is constrained into chromatin in the nucleus of a cell. Chromatin is a dynamic complex structure in which DNA is associated with both histone proteins and numerous non-histone proteins[1]. The basic repeating structural and functional unit of chromatin is known as nucleosome. In each nucleosome the bulk genomic DNA is bound to a core histone octamer which is made up of two each of the core histone proteins, H2A, H2B, H3, and H4[2]. The core histones are comprised of three alpha helices separated by two loops. One hundred and forty-six base pairs of genomic DNA is wrapped around the core histone octamer like a spool of thread in a 1.65° left handed helix, resulting in the first level of DNA condensation. Approximately 10-90 base pairs long DNA separates each nucleosome from other succeeding or preceding nucleosomes is called linker DNA. A polymer of nucleosomes assembled on a single DNA molecule is termed as nucleosomal arrays[3]. Nucleosomal arrays undergo a second level of DNA condensation by binding to the fifth histone protein called a linker histone, forming a chromatin fiber. Self-association of chromatin fibers make up the next level of DNA condensation[4,5].

Linker histones are a family of lysine rich proteins that bind to or near the point at which DNA enters and exits the nucleosomal core in a roughly 1:1 stoichiometry of

linker histone protein to nucleosome and organize an additional 20 base pairs of linker DNA, forming a chromatosome[6]. It has been believed that the primary function of linker histone is to influence the higher order structural transition of condensed structures of chromatin fibers. Although the structure of the nucleosome core particle has been determined at high resolution by X-ray crystallography[6-8], the exact location, function, and dynamics of linker histone binding to nucleosome core are not clearly understood.

A model for nucleosome structure is illustrated in Figure 1.1. This model is generated by using Accelrys Discovery Studio v3.1 (San Diego, CA). The core octamer comprised of two of each of the core histones is shown in red, while the nucleosomal DNA wrapped around these proteins is shown in blue. The crystal structure of the nucleosome core was obtained from the Protein Data Bank (PDB: 1EQZ)[9-11]. An ensemble of NMR structures for H1 linker histone globular domain[12121212121212] was obtained from the Protein Data Bank (PDB accession code: 1GHC), and one representing structure was retained (shown in pink)[12]. The N- and C- terminal regions of H1 (shown in black) were added to the H1 globular domain using the Protein Building Module. Additionally, twenty five base pairs (bp) were added (shown in yellow) using the B-DNA template in Nucleic Acid Building Module. These additional DNA base pairs were fused to each end of the 146 bp core nucleosomal DNA, making the total length of the DNA to be 196 bp. The image shows the complete linker H1 histone protein interacting with the DNA at the entry and exit sites of the nucleosome core.

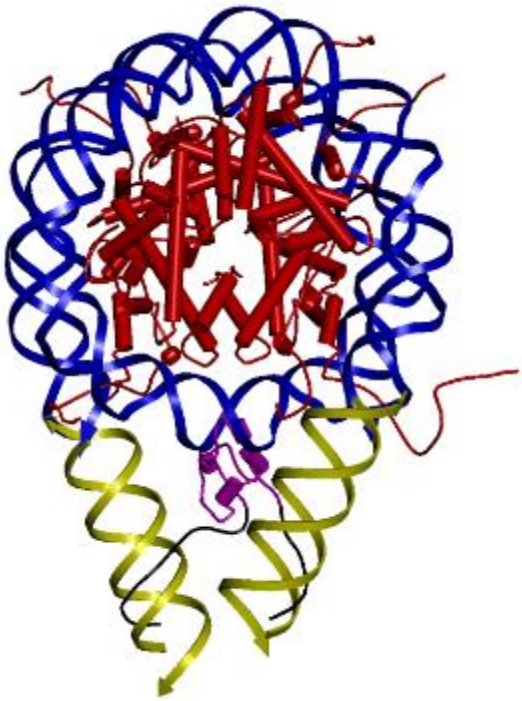

Figure 1.1 Schematic representation of nucleosome (PDB: 1EQZ).

1.2 Linker Histone Family in Mammals

Unlike core histones, linker histones are less evolutionarily conserved[13]. Eukaryotes have different subtypes of linker histones in different tissues. To date eleven different nonallelic variants or subtypes have been identified in mammals[1]. Interestingly, distributions of genes which express these eleven linker histone subtypes are conserved between human and mouse. Initially they were named according to their chromatographic and electrophoretic properties (e.g., H1a, H1b, H1c), but eventually cloned and sequenced each of these sub types and introduced alternative nomenclature (e.g., H1.0 (or $H1^0$), H1.1-H1.5) by Walter *et al*[14]. Among the eleven subtypes, seven are expressed in somatic cells (H1.1-H1.5, H1.0, and H1.X), and four histones are expressed in germ cells: three in sperm (H1t, H1T2, HILS1) and one in oocytes (H1oo)[1].

In both humans and mice, five of the somatic subtypes (H1.1-H1.5) are classified as cell cycle-dependent or replication dependent. This means expression of these subtype genes is strongly connected to a particular phase of the cell cycle. The other somatic subtype $H1^0$ is classified as replication-independent or replacement subtype. $H1^0$ subtype gene is expressed without tight connection to a particular phase of cell cycle by that replace main type histone types depending on particular functional condition. In mammals, a high level of $H1^0$ is found in terminally differentiated cells[15]. H1.X is the least characterized somatic subtype which is localized in the nucleolus and is involved in mitotic progression[16]. Several experimental evidences have been supported that subtypes differ in timing of expression, turnover rate, binding affinity, and DNA condensing capacity. A few researchers have suggested that the functional heterogeneity amongst the subtypes may arises due to the differences in amino acid sequences and sizes of the flanking tails[17]. However, little is known for specific actions of subtypes at a molecular level. Thus, it is important to understand the mechanistic explanation for the functional diversity of these subtypes. In this study, somatic variants such as H1.1, H1.4, and $H1^0$ have been used to determine the thermodynamic parameters of their binding with ds-DNA.

1.3 Linker Histone Domain properties

Linker histones are small basic proteins (~ 21 kDa) that have tripartite domain structure consisting of a short unstructured N-terminal domain approximately 35 amino acid residues in length, a central globular winged helix domain approximately 65 residues in length, and a longer unstructured C-terminal domain approximately 110 amino acid residues in length[18,19]. Amino acid sequence identity and the charge distribution between

the domains of mouse histone H1 variants are listed in Table 1.1 and 1.2. The unstructured N- and C- tails comprise approximately half of the protein molecule. In mammals the sequence of the H1 globular domain is highly conserved while the N- and C-terminal domains are more varied with respect to sequence and amino acid composition[20]. This variability in the terminal domains is thought to be responsible for the functional divergence among the subtypes[19]. However, the actual mechanism and the energetics responsible for each domain binding to the DNA are poorly understood. In this research, the C-terminal tail ($H1^0$-C) and the globular domain ($H1^0$-G) of somatic variant $H1^0$ have been used to determine the DNA binding contributions of each domain to the thermodynamic signatures of the intact protein.

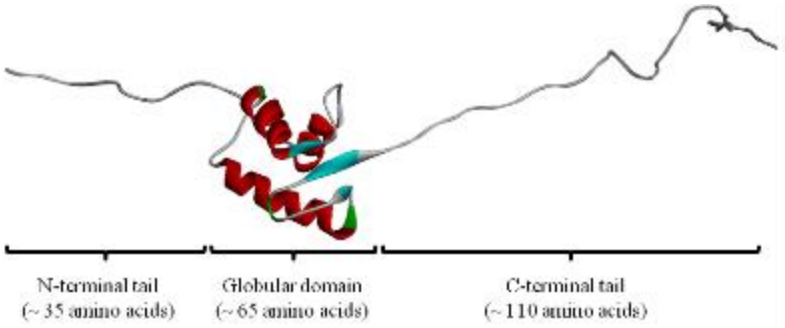

Figure 1.2 Linker histone model

The linker histone (H1) model shown in Figure 1.2 is generated using Accelrys Discovery Studio v3.1 (San Diego, CA). The structure for the globular protein was adapted from the NMR structure obtained from the Protein Data Bank (PDB: 1GHC). In this model the N-and C-terminal tails were added to the H1 globular domain using the Protein Building Module with extended conformation due to high sequence propensity

for random coil[1,21]. Figure 1.2 shows the H1 structure minimized in a solvent shell modeled with periodic boundary condition. The linker histone N-terminal tail is approximately 20-35 amino acid residues in length. Two distinct sub regions are noticed upon examination of the amino acid composition of the N-tail. The first half (distal subregion) is lacking any basic amino acid residues and is enriched in alanine, proline, and highly hydrophobic residues. Unlike the distal subregion, the region proximal to the globular domain is highly basic which contains one arginine and five lysine residues[17]. Because of the basic nature and its close proximity to the globular domain suggest that the second half of N-terminal tail (proximal subregion) may contribute to the binding stability of linker histone in chromatin condensation. Though the N-terminal tail lacks secondary structure in aqueous solution, the N-terminal tails of $H1^0$ and H1.4 acquired a high degree of helical structure in the presence of long DNA and the solvent triflouroethanol (TFE)[20]. N-terminal tail has been proposed to assist in locating and anchoring the globular domain to the nucleosome. Recent FRAP (fluorescence recovery after photobleaching) studies also suggested that the N-terminal tail plays a pivotal role in the differential chromatin binding affinity of $H1^0$ and H1.1[22].

The C-terminal domain has a unique amino acid composition dominated by Lys (~40%), Ala (~17%), and Pro (~12%). The positively-charged residues in the C-domain take part in the neutralization of the linker-DNA phosphate backbone during chromatin condensation. It has been reported that not all of the Lys/Arg residues in the H1 C-terminal tail are equally involved in the stabilization of condensed chromatin and that the change in DNA structure is not simply the result of charge neutralization[4,22]. The 24 residues located closest to the globular domain (97-121) and the 24 residues located

approximately in the middle of the C-terminus (145-169) play the largest role in the condensation and organization of linker-DNA. Although the C-terminal sequence varies with subtype, it is the amino acid composition (*i.e.* 40% lysine content) that is primarily responsible for the interaction of the C-terminal domain and linker-DNA[23]. Although the C-terminal sequence of the various H1 subtypes is variable, all the subtypes have approximately the same charge and a similar characteristic intrinsically disordered structure[23]. It has been reported that the intrinsically disordered C-terminal domain develops some α-helix and/or β-sheet structure upon binding to DNA[4]. It has been proposed that the interactions between the C-domain and linker-DNA influence the orientation of the globular-domain in the H1/DNA complex[24]. In effect the protein and/or DNA conformational changes brought about by the binding of the C-terminal domain to linker-DNA is thought to play a role in determining the interfacial contacts between the H1 globular domain and the nucleosome in chromatin[24]. Truncation studies indicated that the C-terminal tail is required for H1 to recognize protein surfaces[19]. However, the specific sites or residues within the C-terminal tail essential for the chromatin binding and its structural arrangement within the condensed chromatin remain unknown.

Table 1.1 Comparison between amino acid composition, length of each domain, charge on each domain of H1.1, H1.4, H1^{0}[25]

H1 subtype	**N-terminal domain**			**Globular domain**			**C-terminal domain**		
H1.1	S_1-K_{33}	33	+7	T_{34}-A_{115}	82	+7	K_{116}-K_{211}	96	+7
H1.4	S_1-K_{33}	33	+7	T_{34}-A_{115}	82	+7	K_{116}-K_{211}	103	+7
H1^{0}	T_1-S_{21}	21	+7	T_{22}-F_{106}	85	+7	K_{107}-K_{193}	87	+7

Table 1.2 Sequence identity (%) of H1.1, H1.4, H1^{025}

H1 subtype	H1.1 N-tail	H1.1 G-domain	H1.1 C-tail	H1.4 N-tail	H1.4 G-domain	H1.4 C-tail
H1.1				68	98	75
H1^{0}	46	39	39	36	40	35

All the somatic variants of the linker histone proteins share a well-conserved globular domain. H1 and H5 globular domains have been well resolved by single-crystal X-ray crystallography and NMR, respectively[10,11]. The globular domain folds into a classical winged-helix motif and is sufficient for specific binding to the core nucleosome *in vitro*. Although the location of the globular domain is a bit of controversy, it is well accepted that the globular domain binds at least two strands of DNA near the nucleosome dyad where DNA exits and enters the core histone and hence stabilizes DNA wrapping[26,27]. Recent mapping studies revealed that these two binding sites of globular domain are primarily enriched in positively charged amino acids[27,28]. Early experimental and molecular modeling studies suggested that globular domain binds to the minor groove of the DNA at the nucleosome dyad and orients linker DNA to form stem structure. However, it is important to know that the homology of globular domain to the winged-helix family of transcription factors suggested that this domain binds to the major groove of DNA[27,28] . It is also worth mentioning that the much-widened minor groove at the nucleosome dyad axis of symmetry may sufficiently mimics the major groove in this view.

1.4 Linker Histone and Cancer

In addition to the structural role, linker histone proteins have acquired a much more interesting role in transcriptional regulation. Like core histones linker histone are also targets of several posttranslational modifications. The important posttranslational modifications that have been found to date include phosphorylation, acetylation, methylation, and ubiquitination[29]. Through these posttranslational modifications linker histone can both activate and repress transcription. Linker histone C- and N-terminal tail serine and threonine amino acid residues are phosphorylated in response to several cellular stimuli[1]. Reversible phosphorylation of linker histones continuously increases during the cell cycle, becomes maximal during late G2 and mitosis, and decreases distinctly at the end of mitosis in telophase. Recent microarray analysis and gene knockout experiments suggested that the methylated lysine 26 (K26) of H1.4 suppress the transcription[30]. Besides H1 phosphorylation effects on the cell cycle-dependent changes of chromatin structure, H1 also has a functional role in regulating gene expression by preventing the binding of transcription factor proteins to DNA. Different characteristic expression patterns of H1 subtypes were noticed recently in ovarian adenocarcinomas[31]. These findings suggest that H1 variants are important epigenetic factors in modulating chromatin function and gene expression and that they serve as potential epigenetic biomarkers for ovarian cancer[30,31]. Considering the above mentioned studies and evidences, understanding the functional properties of H1 and its interactions with DNA could provide new opportunities for cancer diagnosis, therapies, and drug design.

1.5 Research Specific Objectives

The determination of the best model for the linker histone DNA complex can lead to more specific studies in understanding the nucleosome condensation and gene transcription. This study attempts to answer several questions regarding the interaction of three different mouse H1somatic cell variants (H1.1, H1.4, and $H1^0$) with their physiological binding partner ds-DNA. We also attempt to determine the role of each domain of $H1^0$ in interaction with ds-DNA. The main objectives of this research were: 1) to determine and compare the thermodynamic signatures and structural changes that accompany H1.1, H1.4, and $H1^0$ binding to highly polymerized calf-thymus DNA, 2) to evaluate whether the binding of the complete H1 protein ($H1^0$) or its domain peptides ($H1^0$-C and $H1^0$-G) to ds-DNA is accompanied by a change in structure of the protein, protein fragment, or DNA, 3) to probe the accuracy of binding site size of H1.4 using short ds-DNA oligomers (12 bp, 24 bp, 36 bp, 48 bp, and 72 bp) as its binding partners, 4) to probe the role of water and counterions accompany $H1^0$ binding to DNA using methods like temperature, ionic strength, and osmotic stress dependence of K_a and $\Delta H°$.

CHAPTER II

MATERIALS AND METHODS

This entire dissertation is focused on the determining the energetics and understanding the binding mechanism for the interactions of linker histones (also C-terminal tail and G-domain of the linker histone) with ds-DNA. This chapter is a collective description of the different biophysical tools (*i.e.,* ITC, DSC, and CD) commonly used throughout each of these studies. Detailed descriptions of the instruments, theories, and the experimental methodologies are described in this chapter to avoid repeating of each of these techniques throughout each chapter.

2.1 Isothermal Titration Calorimetry

ITC is used to measure the thermodynamic parameters for a variety of biological interactions, including protein-protein interactions, protein-DNA/RNA interactions, as well as enzymatic reactions. All these biological interactions involve either absorption or evolution of heat. This heat is directly proportional to the total amount of binding that takes place within the calorimetric cell. A typical ITC experiment involves the injection of a known concentration of small volumes of a ligand into a cell filled with a dilute macromolecule solution of known concentration. A typical ITC instrument is displayed in Figure 2.1A depicting the reference cell, sample cell, and syringe. The sample cell and reference cells are made up of the inert alloy Hastelloy and are enclosed in an adiabatic

jacket. Both sample and reference cells are equipped with a power source supplied to the reference and sample cells to maintain a zero temperature difference between the two cells throughout the reaction. As a ligand is titrated into the macromolecule sample either heat is generated or absorbed, creating a transient temperature difference between the sample and reference cells. For our particular ITC instrument (VP-ITC, GE Healthcare) the temperature difference is calibrated to power units (µcal/sec). This power is supplied back to the sample cell so that the temperature between the two cells remains the same. For an exothermic binding process, less power is being supplied to the sample cell in order to keep the sample cell temperature the same as the reference cell. For an endothermic binding process, more power is being supplied to the sample cell in order to keep the sample cell temperature the same as the reference cell. The power difference between the sample and reference cell is then used to determine the stoichiometry (n), the association-binding constant (K_a), and the enthalpy change of binding ($\Delta H°$).

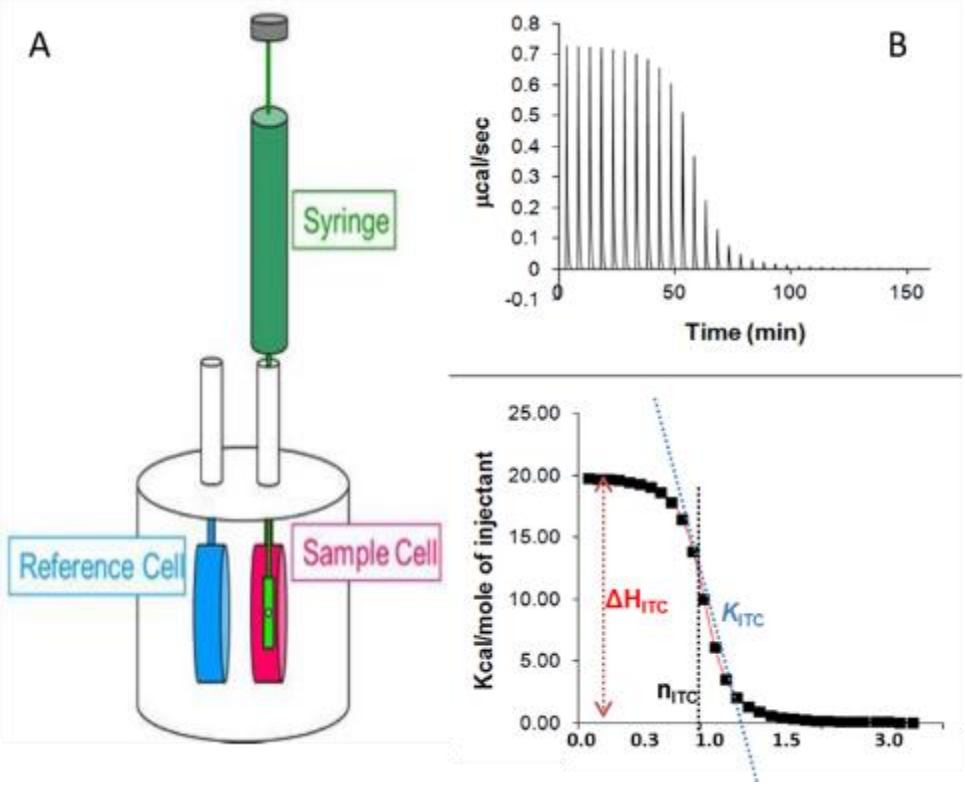

Figure 2.1 Representative diagram of an ITC instrument with reference cell, sample cell, and syringe (A) [32].

Panel B represents typical raw ITC and integrated ΔH° plot for an endothermic process.

2.1.1 ITC determination of n, K_a, and ΔH°

An example of raw data collection from an ITC experiment is shown in Figure 2.1B. A simple protein/DNA interaction can be best explained as one where a protein binds to DNA, an example of the integrated heats of interaction for a 1:1 binding process shown in Figure 2.1B lower panel. In Figure 2.1B, the binding constant is described as the ratio of the concentration of the bound complex, [DP], to that of the free protein, [P], and the concentration of the free DNA, [D] as described below

$$[D] + [P] \rightleftharpoons [DP] \qquad (2.1)$$

$$K_a = \frac{[DP]}{[D][P]} \quad (2.2)$$

The association equilibrium constant, K_a, can be also related to the free energy ($\Delta G°$) using equation 2.3

$$\Delta G° = -RT \ln K_a \quad (2.3)$$

The Gibbs free energy change, ΔG°, can be expressed in terms of the enthalpy of change, ΔH°, entropy change, ΔS°, and the absolute temperature in Kelvin as described by the equation 2.4.

$$\Delta G° = \Delta H° - T\Delta S° \quad (2.4)$$

The power required in maintaining the parity between the sample and reference cell for an exothermic binding reaction is measured for each injection. The initial injection of protein results in the binding of most of the protein molecules to the DNA. This initial injection therefore requires the greatest power compensation and thus generates the greatest amount of heat. On subsequent injections of protein, there is less DNA available for binding and, therefore, less heat of binding is generated. After approximately 20 injections of protein, all the sites on the DNA are bound with protein molecules, and no further heat of binding is observed. The remaining heat generated at this point is as a result of the heat of diluting the protein solution into the DNA solution. Typically, the protein concentration in the injector is at least 10 times more concentrated than in the sample cell. By integrating these deflections with respect to time and correcting for the heat of dilution, the heat of binding per injection (kcal/mol/injection) is calculated and plotted against the molar ratio of the protein to DNA (Figure 2.1B).

The association equilibrium constant (K_a) is related to the shape of the curve, and the binding capacity (n) is determined from the ratio of the protein to DNA at the equivalence point of the curve[33]. The enthalpy change ($\Delta H°$) for the reaction is approximately the intercept at a zero molar ratio. In order to extract thermodynamic parameters from this plot (Figure 2.1B), the data must be fit to an appropriate binding model.

2.1.2 Analysis of ITC raw data

The data generated in Figure 2.1B and the thermodynamic data (n, K_a, and $\Delta H°$) derived from these data are actually apparent values because the measured heat of binding per protein injection can arise from any linked equilibria. This implies that it is the total heat absorbed or released in the sample cell upon each addition of the protein. The heat absorbed or released may also arise from sources other than the binding of the protein to the DNA (equation 2.1). The additional sources of heat from protein-DNA interaction are summarized as follows[33]:

$$\Delta H° \approx \Delta H_{bind} + \Delta H_{dilution} + \Delta H_{matrix} + \Delta H_{ion} \quad (2.5)$$

where ΔH_{bind} represents the heat of binding of the protein to the DNA, $\Delta H_{dilution}$ arises from the dilution of the protein into the matrix of the DNA, ΔH_{matrix} denotes the heat arises from the mixing the protein and DNA matrixes, and ΔH_{ion} represents the heat arise from a change in pH in the sample cell upon addition of protein. Some of these enthalpies (ΔH_{matrix} and ΔH_{ion}) can be eliminated or diminished by setting up the experiment under correct conditions such that they can be neglected. For example, both ΔH_{matrix} and ΔH_{ion} will be approximately zero, when the protein and the DNA solutions

are prepared in the same buffering solution. $\Delta H_{dilution}$ has to be measured in a separate experiment or approximated using the last few heat of binding per injection data points (Figure 2.1B lower panel) and then subtracted from $\Delta H°$ to determine ΔH_{bind}. After the binding data are corrected for contributions arising from nonspecific enthalpies, a model is chosen to fit the data such that the parameters n, K_a, and $\Delta H°$ can be determined.

2.1.3 Single and Multiple Independent Binding Site Model for Fitting Data

Once the raw ITC has been corrected for nonspecific heat contribution, the data are fitted to an appropriate single site binding model. If we define the quantity r as the moles of protein bound per mole of total DNA and use the association constant definition $[PD] = K_a\,[P]\,[D]$ from equation 2.2, then

$$\mathrm{r} = \frac{moles\ protein\ bound}{total\ moles\ DNA} = \frac{[PD]}{[D]+[PD]} = \frac{K_a\,[P][D]}{[D] + K_a\,[P][D]} = \frac{K_a\,[P]}{1 + K_a\,[P]} \qquad (2.6)$$

Hence, the quantity r is the fraction of sites occupied by the protein (P) with an association constant K_a. Rearranging equation 2.6 leads to equation 2.7.

$$K_\mathrm{a} = \frac{r}{(1-r)\,[P]} \qquad (2.7)$$

The total concentration of protein $[P]_T$ is known and can be expressed by equation 2.8, in which $[P]_T$ denotes the total concentration of protein and n represents the capacity number or number of sites.

$$[P]_\mathrm{T} = [P] + \mathrm{n}r[D]_\mathrm{T} \qquad (2.8)$$

It is apparent from equation 2.8 that $nr[D]_T = [PD]$. Combining equations 2.7 and 2.8 gives quadratic equation 2.9:

$$r^2\text{-}r\left[\frac{[P]_T}{n[D]_T}+\frac{1}{nK_a[D]_T}+1\right]+\frac{[P]_T}{n[D]_T}=0 \quad (2.9)$$

Resolving equation. 2.9 for *r* leads to equation 2.10:

$$r=\frac{1}{2}\left[\left(\frac{[P]_T}{n[D]_T}+\frac{1}{nK_a[D]_T}+1\right)-\sqrt{\left(\frac{[P]_T}{n[D]_T}+\frac{1}{nK_a[D]_T}\right)^2-\frac{4[P]_T}{n[D]_T}}\right] \quad (2.10)$$

The total heat content (Q) contained in the sample cell at volume (V) can be defined as

$$Q=nr[D]_T\Delta H°V \quad (2.11)$$

where ΔH° is the heat of binding of the protein to the DNA, and $nr[P]_T=[PD]$. Substituting equation 2.10 into equation 2.11 gives the following equation:

$$Q=\frac{n[D]_T\Delta H^0V}{2}\left[\left(\frac{[P]_T}{n[D]_T}+\frac{1}{nK_a[D]_T}+1\right)-\sqrt{\left(\frac{[P]_T}{n[D]_T}+\frac{1}{nK_a[D]_T}\right)^2-\frac{4[P]_T}{n[D]_T}}\right] \quad (2.12)$$

Therefore, the total heat content (Q) for the reaction is a function of n, K_a, and ΔH° since the $[P]_T$, $[D]_T$, and V are known for each experiment. In Figure 2.1B, the measured heat content represents the change in heat content with the injection of a known volume of protein solution into the sample cell containing a known volume of DNA solution (V). Therefore, the volume of the sample cell changes by a known amount (ΔV_i) after completing the injection of protein. The change in the heat content is defined by equation 2.13

$$\Delta Q(i)=Q(i)+\frac{\Delta V_i}{V}\left[\frac{Q(i)+Q(i-1)}{2}\right]-Q(i-1) \quad (2.13)$$

Analysis of the titration curve shown in the bottom right corner in Figure 2.1B to equation 2.13 using CHASM data analysis software developed in our laboratory yields values for n, K_a, and $\Delta H°$.

For two independent sites, equation 2.6 can be rewritten as a linear combination of these sites:

$$r_1 + r_2 = \frac{K_{a1}[P]}{1 + K_{a1}[P]} + \frac{K_{a2}[P]}{1 + K_{a2}[P]} \quad (2.14)$$

The total concentration of protein $[P]_T$ is rewritten as

$$[P]_T = [P] + (n_1 r_1 + n_2 r_2)[D]_T \quad (2.15)$$

Equation 2.15 can be inserted into equation 2.14 to yield a cubic equation[33]. The above described procedure is followed exactly, except in this case, the cubic equation has to be solved numerically.

2.2 Differential Scanning Calorimetry

DSC is a powerful technique that measures the excess heat capacity (C_p) of a macromolecule solution or macromolecular complexes as a function of temperature. A typical DSC instrumental schematic is shown in Figure 2.2. DSC consists of two independent cells (sample and reference cell) that are isolated, shielded, and their temperatures controlled by a small power source. The sample cell contains the macromolecule and the reference cell contains a reference solution such as water or buffer. Temperature sensors attached to the cells provide feedback regarding the ΔT_m between the sample and reference cells as well as a sensor that monitors the ΔT_m between the cells and the shield[34,35]. The differential heat uptake as a function of temperature

results the difference in the heat capacities of the macromolecule and buffer in the sample and reference cell, respectively. Most biomolecules melt cooperatively, meaning that the equilibrium equation can be described as a two state process where the molecule goes from folded to unfolded state[35]. The melting temperature, T_m, represents the temperature where the fraction of unfolded biomolecule equals the folded biomolecule[35-37]. When protein is added to the DNA solution, the T_m may increase due to stabilization of the higher order protein/DNA complex[36]. The ΔT_m can then be related to the binding affinity for the interaction of the protein with the DNA macromolecule. The calorimetric enthalpy, $\Delta H°_{cal}$, is the calculated as the area under the curve for each melting transition. The energy for the unfolding of the DNA or protein-DNA complex structure is calculated as an integration of the raw excess heat capacity shown in the equation 2.16.

$$\Delta H°_{cal} = \int_{T_0}^{T} C_p\,(ex)dt \qquad (2.16)$$

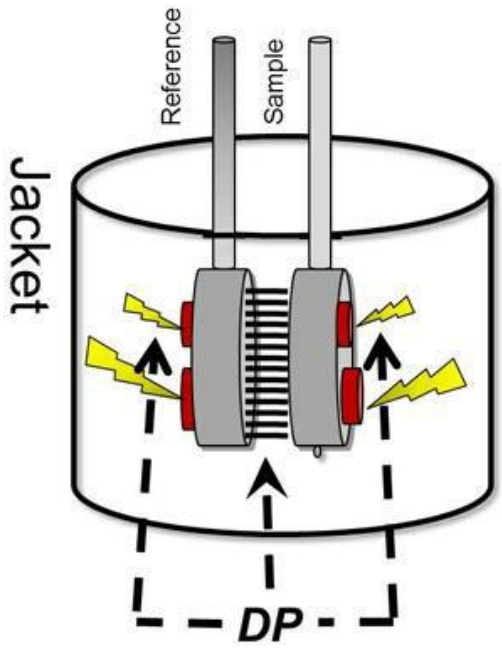

Figure 2.2 Schematic representation of a typical DSC including reference cell and sample cell[37].

2.3 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[38]. CD is highly sensitive to minor structural or conformational differences in the biopolymers because of their asymmetric nature. Biomolecules, which are in asymmetric nature, will only show an absorption in the CD[39-41]. A schematic of the incident beam, left and right polarizing components, and an example of resulting CD spectra are shown below in Figure 2.3. Circular dichroism spectroscopy can be used for: determining whether a protein is folded and subsequently characterizing its secondary structure, studying the conformational stability of a protein subjected to pH changes or thermal

stress, determining conformational changes with protein-protein or protein-DNA interactions.

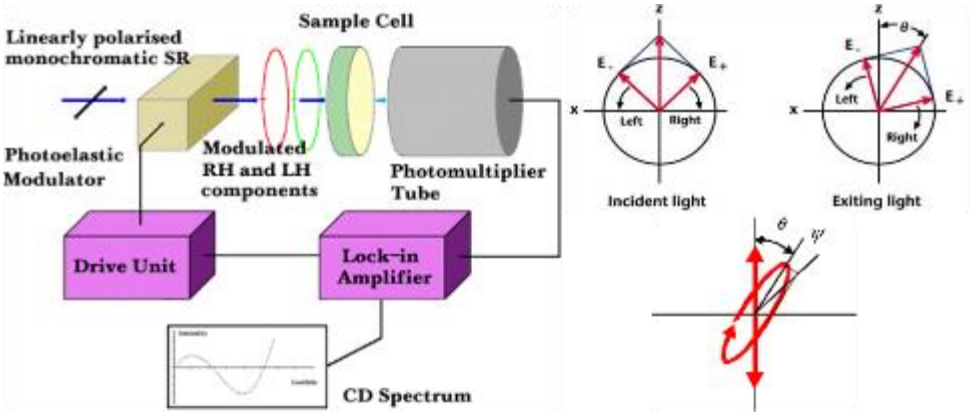

Figure 2.3 Illustratative diagram of the CD spectropolarimeter including incident beam, the left and right polarizing components, and an example of the resulting CD spectra[42].

Secondary structure can be determined by CD spectroscopy in the "far-UV" spectral region (190-250 nm)[38]. Alpha-helix, beta-sheet, and random coil structures each give rise to a characteristic CD spectra[38]. 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[38]. An Olis DSM 20 spectropolarimeter was used to determine the structure of the histone variants (H1.1, H1.4, and $H1^0$). CD experiments were used to determine gross structural changes in the DNA (240 and 280 nm regions) and the histone variants upon formation of the histone/DNA complex.

CHAPTER III

CALORIMETRIC STUDIES OF THE INTERACTIONS OF LINKER HISTONE $H1^0$ AND ITS CARBOXYL ($H1^0$-C) AND GLOBULAR ($H1^0$-G) DOMAINS WITH CALF-THYMUS DNA

3.1 Abstract

Histone H1 is a chromatin protein found in most eukaryotes. ITC and CD have been used to study the binding of $H1^0$, its C-terminal, $H1^0$-C, and its globular, $H1^0$-G, domains to a highly polymerized DNA. ITC results indicate that $H1^0$ and $H1^0$-C bind tightly to DNA ($K_a \approx 1\times10^7$), with an unfavorable $\Delta H°$ ($\Delta H° \approx +22$ kcal/mol) and a favorable $\Delta S°$ ($-T\Delta S° \approx -30$ kcal/mol). Binding $H1^0$-G to DNA at 25º C is calorimetrically silent. A multiple independent sites model fits the ITC data, with the anomaly in the data near saturation attributed to rearrangement of bound H1, maximizing the number of binding sites. CD experiments indicate that $H1^0$/DNA and $H1^0$-C/DNA complexes form with little change in protein structure but with some DNA restructuring. Salt dependent ITC experiments indicate that the electrostatic contribution to binding $H1^0$ or $H1^0$- C is small, ranging from 6% to 17% of the total $\Delta G°$.

3.2 Introduction

Linker histones (H1) are the basic proteins in higher eukaryotes that are responsible for the final condensation of chromatin[43]. H1 also plays an important role in

regulating gene expression. H1 has been described as a transcription repressor as it limits the access of transcriptional factors to DNA[44-49]. Linker histone binds to DNA that enters or exits the nucleosome[3]. Several crystal structures have been published for the nucleosome (histone core/DNA complex), and the interactions of the core histone proteins with DNA are well understood[6-8,13]. In contrast the location of the linker histone and its interactions with DNA are more poorly understood.

H1 exists as multiple isoforms, with eleven different subtypes identified in mammals to date. Among the eleven subtypes, seven are expressed in somatic cells (H1.1-H1.5, $H1^0$, and H1.X), and four histones are expressed in germ cells: three in sperm (H1t, H1T2, HILS1) and one in oocytes (H1oo)[1]. All of the H1 histone subtypes have a similar structure composed of three separate domains, a short disordered N-terminal domain approximately 35 amino acid residues in length, a central globular winged helix domain approximately 65 residues in length, and a longer disordered C-terminal domain approximately 100 amino acids residues in length[18]. In mammals the sequence of the H1 globular domain is highly conserved, while the N- and C-terminal domains are more varied with respect to sequence[50] and amino acid composition[1]. This variability in the terminal domains is thought to be responsible for the functional differences among the subtypes[51].

The C-terminal domain has a unique amino acid composition dominated by Lys (~40%), Ala (~17%), and Pro (~12%). The positively charged residues in the C-domain take part in the neutralization of the linker-DNA phosphate backbone during chromatin condensation[19,23]. It has been reported that not all of the Lys/Arg residues in the H1 C-terminal domain are equally involved in the stabilization of condensed chromatin and that

the change in DNA structure is not simply the result of charge neutralization[5]. The 24 residues located closest to the globular domain (97-121) and the 24 residues located approximately in the middle of the C-terminus (145-169) play the largest role in the condensation and organization of linker-DNA[5]. Although the C-terminal sequence varies with subtype, it is the amino acid composition (*i.e.,* 40% lysine content) that is primarily responsible for the interaction of the C-terminal domain and linker-DNA[5]. Although the C-terminal sequence of the various H1 subtypes is variable, all the subtypes have approximately the same charge and a similar characteristic intrinsically disordered structure[52]. It has been reported that the intrinsically disordered C-terminal domain develops some α-helix and/or β-sheet structure upon binding to DNA[4]. It has been proposed that the interactions between the C-domain and linker-DNA influence the orientation of the globular-domain in the H1/DNA complex[24]. In effect the protein and/or DNA conformational changes brought about by the binding of the C-terminal domain to linker-DNA is thought to play a role in determining the interfacial contacts between the H1 G-domain and the nucleosome in chromatin[22,24,53].

In the present study we have used isothermal titration calorimetry (ITC) to determine the thermodynamics for binding of $H1^0$, $H1^0$-C, and $H1^0$- G to highly polymerized calf-thymus DNA. We have also used circular dichroism (CD) to evaluate whether the binding of the complete H1 protein ($H1^0$) or its domain peptides ($H1^0$-C and $H1^0$-G) to ds-DNA is accompanied by a change in structure of the protein, protein fragment, or DNA. In our ITC studies, we found that the intact protein ($H1^0$) and its C-terminal domain ($H1^0$-C) bind to CT-DNA with approximately the same affinity ($K_a \approx 1\times10^7$). We also observed a surprisingly large endothermic enthalpy change for the

formation of these H1/DNA complexes ($\Delta H° \approx +22$ kcal/(mol $H1^0$ or $H1^0$-C)). There was no ITC signal for the addition of $H1^0$-G to CT-DNA indicating that either $H1^0$-G binding did not occur or that the $H1^0$-G/DNA complex was formed with a very small change in enthalpy ($\Delta H° \approx 0$ kcal/(mol $H1^0$-G)). The $H1^0$/DNA and $H1^0$-C/DNA complexes are driven by a very favorable entropy change ($-T\Delta S° \approx -30$ kcal/mol), and the binding site sizes for $H1^0$ and $H1^0$-C were determined to be 36bp and 28bp, respectively. Our CD results indicate that CT-DNA is restructured upon complex formation with either the full length H1 protein ($H1^0$) or its C-terminal domain ($H1^0$-C). In contrast, the structure of $H1^0$ is largely unchanged in the DNA complex. Even though the ITC does not detect a binding interaction between the globular domain, $H1^0$-G, and CT-DNA, the CD experiments show a significant change in the $H1^0$-G spectrum in the DNA complex, a sure indication of a strong interaction between $H1^0$-G and DNA. The CD results are consistent with significant loss of α-helical structure in the globular domain as it binds to DNA. These results are discussed in more detail in the Results and Discussion sections of the chapter.

3.3 Materials and methods

3.3.1 Proteins and DNA samples

The $H1^0$ intact protein and its C-terminal and globular domains were expressed using a bacterial strain of *E.coli* (*Rosetta2 (De3) pLysS*) transformed with a pET-11d (Novagen) expression vector as described[21]. The methods for expression and purification have been described elsewhere[21,54]. The pure protein fractions were concentrated using a Savant SPD 111V speed vac system for 4 hrs at 35°C to remove the HPLC solvent (5 % acetonitrile/95% water). Typically the sample buffer was BPES which is 30 mM

K_2HPO_4/ KH_2PO_4 (pH = 7.0), 1 mM EDTA, and 100 mM KCl. For the salt dependent studies the amount of added supporting electrolyte, KCl, was 0.0 mM, 30 mM, 70 mM, 130 mM and 290 mM, for the 0.03, 0.06, 0.10, 0.16 and 0.32 M [K^+] solutions respectively. Calf thymus DNA type I was purchased from Sigma (St. Louis, USA) and dissolved in 1mL of the sample buffer. Both protein and DNA stock solutions were exhaustively dialyzed against the sample buffer (24h) at 4˚C. DNA concentrations in base pairs (bp) were determined using measured absorbance at 260 nm and a molar extinction coefficient of ε_{260}=1.31 x 10^4 bp $M^{-1}cm^{-1}$ [55]. The concentrations of $H1^0$, $H1^0$-C and $H1^0$-G were calculated using extinction coefficients 27.8, 31.1, and 28.6 mL mg^{-1} cm^{-1}, respectively at 205 nm[21].

The approximate molecular weights for the $H1^0$ and its domain constructs were estimated from their sequences using the ExPASy ProtParam tool[56] (http://web.expasy.org/protparam): Mw ($H1^0$) ≈ 20.8 kDa, Mw ($H1^0$-C) ≈ 9.55 kDa, Mw ($H1^0$-G) ≈ 9.28 kDa. The approximate average molecular weight of the CT-DNA was 8.42×10^3 kDa (Sigma, St. Louis, USA).

3.3.2 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a Microcal VP-ITC (Northampton, MA, USA). Titrations were done at five salt concentrations (0.03, 0.06, 0.1, 0.16, and 0.32 M [K^+]) and at 25º C. All titrations were performed by overfilling the ITC cell with approximately 1.5 mL of a dilute CT-DNA solution (nominally 480 μM in base pairs). Approximately 250 μL of a dilute solution of $H1^0$, $H1^0$-C, or $H1^0$-G (nominally 150 μM) was titrated into the calorimeter cell. The

injection volume in these titrations was nominally 10 μL, and a typical titration involved the addition of 25 injections of titrant at 600 second intervals. The 1.5 mL added to the VP ITC cell overfills the cell so that there are no air bubbles in the chamber. As a volume of titrant ($H1^0$, $H1^0$-C, or $H1^0$-G) is added, an equivalent volume of solution is displaced from the cell, and titrate concentration is corrected for material loss at each point in the titration. All of our ITC experiments were performed in triplicate at 25 °C. The integrated heat/injection data were fit for to an appropriate thermodynamic model using CHASM data analysis software developed in our laboratory[57]. The non-linear regression fitting process yields best fit parameters for K (or ΔG°), ΔH°, ΔS°, and n.

3.3.3 Circular Dichroism

CD experiments were performed using an Olis DSM 20 spectropolarimeter (Bogart, GA, USA). CT-DNA and protein solutions were prepared with a nominal absorbance of 0.5 AU in a BPES buffer (10 mM KCl, 30 mM Phosphate, 0.1 mM EDTA). The nominal concentrations of both protein and CT-DNA were 1.5 μM and 96 μM in DNA bp, respectively. The H1/DNA complex solutions were prepared to have an approximate 0.5:1 molar ratio between the H1 protein and the DNA binding sites using an approximate site size of 32 bp. This ratio was chosen to avoid the complications that appear near the endpoint in the ITC titrations. We used dilute solutions of both protein and DNA, and used excess moles of DNA to prepare complex samples for the CD experiments to minimize complex aggregation. CD spectra were collected over a wavelength range of from 200 to 300 nm (with measurements every 0.5 nm) in a 1 cm path length cuvette at room temperature. The spectra represent the average of three scans

which were processed using PRISM software (graph-Pad Prism Software, San Diego, CA).

3.4 Results

3.4.1 Isothermal Titration Calorimetry

Figure 3.1 shows typical ITC data for titrations of the $H1^0$ and $H1^0$-C proteins in 0.100 M [K^+] solutions. In these experiments the complete histone $H1^0$ and its C-terminal domain, $H1^0$-C, were titrated into highly polymerized CT-DNA at 25˚C. The ITC thermograms were fit using non-linear regression techniques to a multiple independent site model (one site model) and the average best-fit parameters are listed in Table 3.1. We also attempted to fit these titration data to a nearest neighbor exclusion model [58]; however, the multiple sites do not appear to be interacting and do not exhibit either positive or negative cooperativity (Figure 3.3). This was determined from the independence of enthalpy change on degree of saturation, ($\Delta h \approx 0.0$) where Δh is enthalpy associated with the interaction between nearest neighbor bound proteins. ITC data indicate that although $H1^0$ or $H1^0$-C have a high binding affinity ($K \approx 10^7\ M^{-1}$) for CT-DNA, the enthalpy change is very unfavorable ($\Delta H° \approx +22$ kcal/(mol $H1^0$ or $H1^0$-C)), and complex formation is driven by a large favorable entropy change ($-T\Delta S° \approx -30$ kcal/mol). The maximum in the endothermic heat observed just prior to the end point in both ITC titrations (Figure 3.1A, 3.1B) is attributed to overcoming a steric interaction in which a bound protein is partially occupying two adjacent sites, and this protein must be relocated in order to fully populate all of the potential protein binding sites. The additional endothermic heat observed near saturation represents the energy cost of relocating the one or more already bound proteins that are unevenly distributed along the

linear lattice of the DNA. In effect, as $H1^0$ binds non-specifically (electrostatically) to a long DNA molecule, the placement of the proteins is random and can result in multiple partial binding sites that are vacant. The binding-site size, or the number of base pairs occupied per bound $H1^0$, $H1^0$-C, or $H1^0$-G protein, was calculated from the ratio of added protein at the titration endpoint to the total number of DNA base pairs in the calorimeter. Previous studies have reported a binding-site size of 10 DNA base pairs for both $H1^0$ and $H1^0$-C[21]. Our ITC results are more consistent with a larger binding sites size, 36 bp for $H1^0$ and 28 bp for $H1^0$-C. The interaction between $H1^0$-G and CT-DNA was also studied by ITC but was observed to be calorimetrically silent at 25 ºC (Figure 3.2).

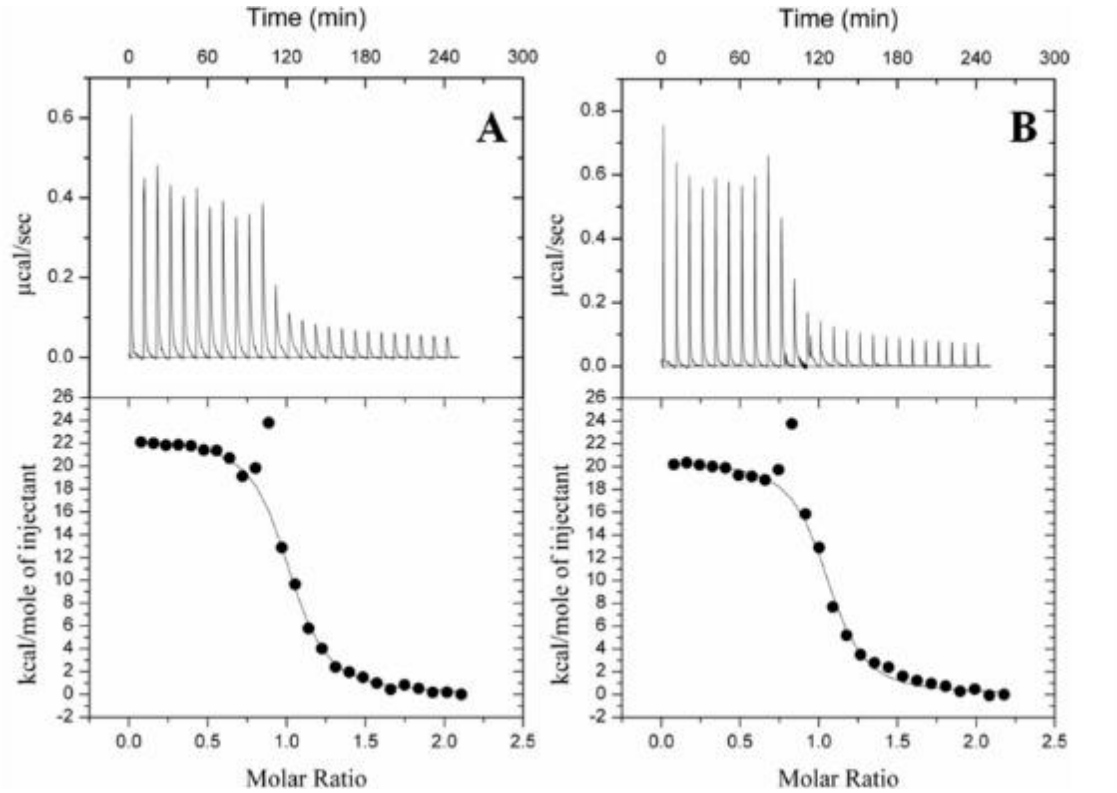

Figure 3.1 ITC thermograms of $H1^0$ and $H1^0$-C binding to CT-DNA at 25 ºC, pH 7.0

Panel **A** shows a typical ITC titration for the addition of $H1^0$ to highly polymerized CT-DNA. The upper half of Panel **A** shows the baseline-corrected raw ITC signal for 25 injections of a dilute $H1^0$ protein solution (10 µL of 150 µM $H1^0$) into the ITC cell filled with a dilute solution of CT-DNA (360 µM bp, 10 µM in $H1^0$ binding sites). The lower half of Panel **A** shows the apparent ΔH° for each injection (-●-) along with the best-fit non-linear regression line (—) for a simple one site binding model. Panel **B** shows a typical ITC titration for the addition of $H1^0$-C to highly polymerized CT-DNA. The upper half of Panel **B** shows the baseline-corrected raw ITC signal for 25 injections of a dilute $H1^0$-C protein solution (10 µL of 150 µM $H1^0$-C) into the ITC cell filled with a dilute solution of CT-DNA (280 µM bp, 10 µM in $H1^0$-C binding sites). The lower half of Panel **B** shows the ΔH° for each injection (-●-) along with the best-fit non-linear regression line (—) for a simple one site binding model.

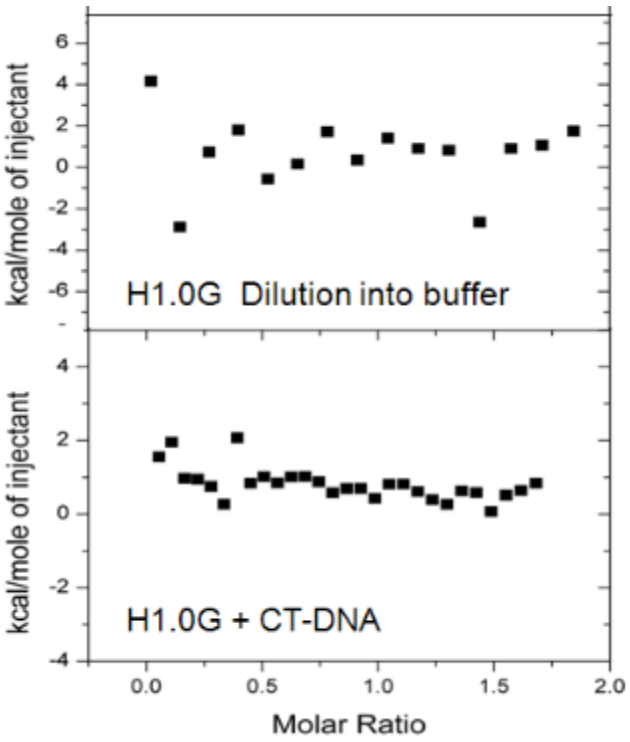

Figure 3.2 ITC thermograms of $H1^0$-G binding to CT-DNA at 25 ºC, pH 7.0.

The upper panel represents the titration of $H1^0$-G into 100 mM [K^+] BPES buffer and the lower panel represents the titration of $H1^0$-G into CT-DNA.

Table 3.1 ITC derived thermodynamic parameters for binding the complete $H1^0$ protein and its carboxyl ($H1^0$-C) and globular domains ($H1^0$-G) to highly polymerized calf thymus DNA.

	K_a (M^{-1}) ×10^{-6}	ΔG° (kcal/mol)	ΔH° (kcal/mol)	-TΔS° (kcal/mol)	Binding site Size (bp)
$H1^0$	7.2±0.2	-9.3	21.8±0.2	-31.1	36
$H1^0$-C	7.5±0.1	-9.4	20.6±0.2	-30.0	28
$H1^0$-G	*ND*	*ND*	*ND*	*ND*	ND

All ITC experiments were performed in triplicate in 100 mM [K^+] BPES buffer at pH 7.0 and 25 °C. The integrated heat/injection data were fit for to a one site thermodynamic model using CHASM data analysis software developed in our laboratory [57]. Errors listed are the standard deviations for the best fit parameters *K* and ΔH° determined in triplicate experiments. Effective binding site size in base pairs was calculated from the titration endpoint, the DNA concentration in base pairs and the assumption that saturation stoichiometry is 1:1 (H1: DNA sites).

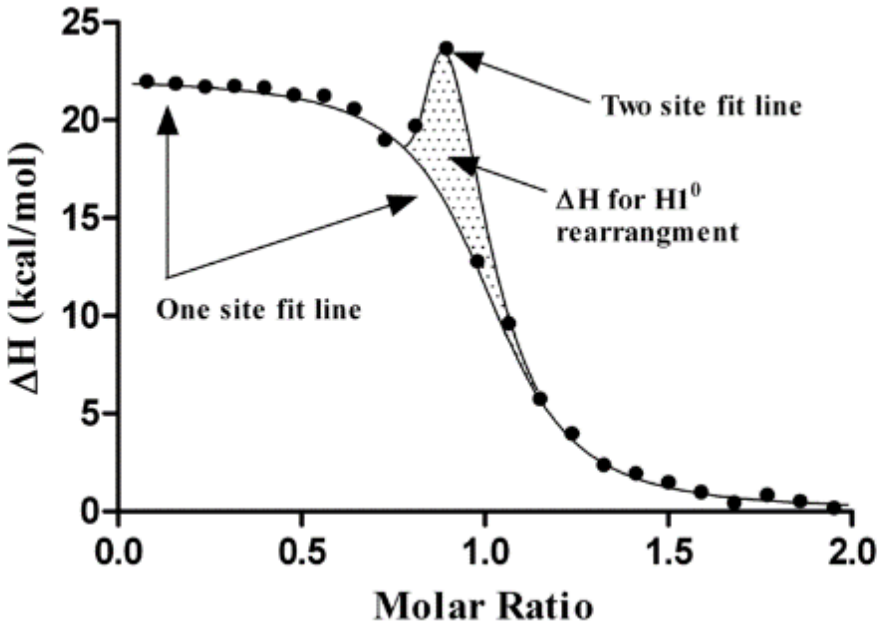

Figure 3.3 ITC data for the addition of $H1^0$-C to CT-DNA.

These are the same data that were shown in Figure 3.1B, fit to two different thermodynamic models, a one site model and a two site model that incorporates the rearrangement reaction for data approaching the endpoint. The shaded area represents the heat for the rearrangement of bound protein.

To further probe the nature of the interactions between $H1^0$ or $H1^0$-C with CT-DNA, we performed a salt dependent study in which the experiments shown in Figure 3.1 were repeated in solutions having total $[K^+]$ of 0.03, 0.06, 0.100, 0.160 and 0.320 M. In Figure 3.4, we have plotted the values for $\Delta G°$, $\Delta H°$, and $-T\Delta S°$ for binding $H1^0$ to CT-DNA as a function the potassium ion concentration. In looking at this plot, it is obvious that increasing the salt concentration results in a continuous but small weakening of the protein/DNA affinity. This plot presents a classic case of enthalpy-entropy compensation wherein larger changes in $\Delta H°$ are compensated for by the opposite change in the $-T\Delta S°$ term. We also show in Figure 3.5, a plot of log K_a, vs. the log of the potassium ion concentration. The salt dependence of the K_a for binding $H1^0$ and $H1^0$-C to CT-DNA has

been analyzed by the polyelectrolyte theory of Record *et al*[59]. Both $H1^0$ and $H1^0$-C exhibit a similar predictable decrease in affinity with increasing ionic strength. The data in Figure 3.5 were used to estimate values for $-Z\psi$, *i.e.,* the slopes of the two lines shown in Figure 3.5, for the salt dependence of K_a for binding $H1^0$ and $H1^0$-C to CT-DNA. The polyelectrolyte effect on the free energy for binding these proteins to CT-DNA was calculated using equations (3.1), (3.2) and (3.3)

$$\log K_{obs} = \log K_t - Z\psi \log[K^+] \quad (3.1)$$

$$\Delta G_{pe} = - Z\psi RT \ln[K^+] \quad (3.2)$$

$$\Delta G_{calc} = \Delta Gt + \Delta G_{pe} \quad (3.3)$$

where ψ is the fraction of potassium counterion associated per DNA phosphate and Z is the apparent charge of the protein. In equation 3.1, K_{obs} is the ITC determined value for $K_{a,}$ and K_t is the non-electrolyte contribution to the overall affinity calculated from the y intercept of the data plotted in Figure 3.5. Table 3.2 lists the values for ΔG_{pe}, the electrostatic contribution to the free energy change, ΔG_t, the non-ionic contribution to the overall binding free energy change, and ΔG_{calc}, the sum of ΔG_{pe} and $\Delta G_{t.,}$ along with the ITC determined value for ΔG for binding either $H1^0$ or $H1^0$-C to CT-DNA, $\Delta G_{obs.}$ These calculations were done at as many as five salt concentrations. The polyelectrolyte effect reduces the affinity at higher salt concentrations, a clear indication that an attractive ligand charge/DNA charge interaction is at least partially responsible for the protein's affinity for DNA. However, the polyelectrolyte contribution to the affinity is very small, ranging from only 17.8% for binding $H1^0$ in 0.03 M $[K^+]$ solution to as low as 6 % for binding $H1^0$-C in 0.320 M $[K^+]$.

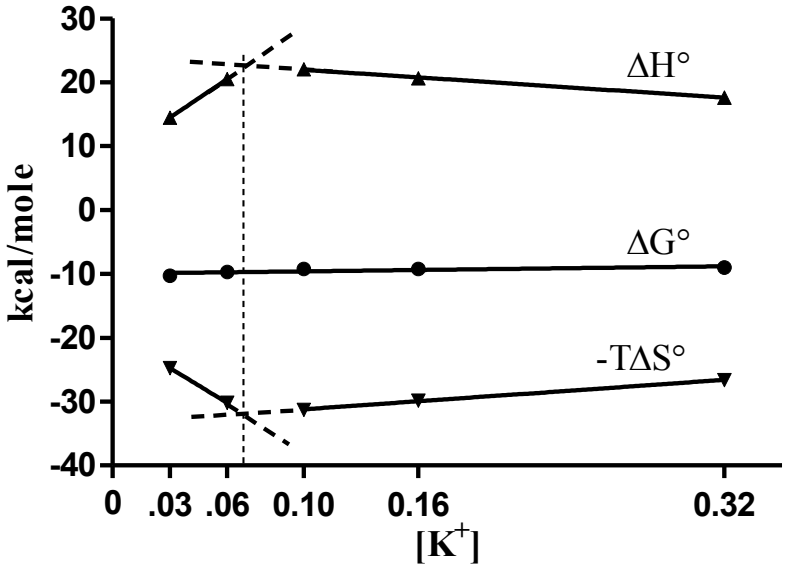

Figure 3.4 A plot of the thermodynamic parameters, ΔG°, ΔH°, and –TΔS° for the binding of $H1^0$ to CT-DNA as a function of salt concentration.

Maximum enthalpy-entropy compensation at 0.06 M [K^+] is denoted by dashed line.

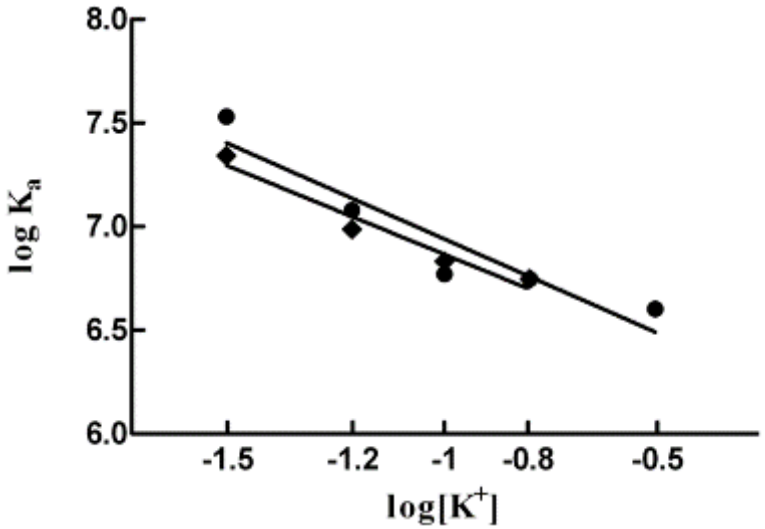

Figure 3.5 A plot of log K_a vs. log [K^+] for the binding of $H1^0$ and $H1^0$-C to CT-DNA.

The data for $H1^0$ are shown as -●-, and the data for $H1^0$-C are shown as -♦-.

Table 3.2 Parsing the free energy change: Calculation of the electrostatic and non-electrostatic contributions to the ΔG° for binding $H1^0$ and $H1^0$-C to CT-DNA.

	$[K^+]$	ΔG_{pe}[a] (kcal/mol)	ΔG_t[b] (kcal/mol)	ΔG_{calc}[c] (kcal/mol)	ΔG_{obs}[d] (kcal/mol)	Ionic contribution[e]
$H1^0$	0.030	-1.8	-8.2	-10.0	-10.3	17.8 %
	0.060	-1.5	8.2	-9.7	-9.7	15.5%
	0.100	-1.2	-8.2	-9.4	-9.2	12.8%
	0.16	-1.0	-8.2	-9.2	-9.2	10.8%
	0.320	-0.6	-8.2	-8.8	-9.0	6.8 %
$H1^0$-C	0.030	-1.7	-8.2	-9.9	-9.9	17.5 %
	0.060	-1.4	-8.2	-9.6	-9.5	14.5 %
	0.100	-1.1	-8.2	-9.3	-9.3	11.8 %
	0.160	-0.9	-8.2	-9.1	-9.2	9.8 %

[a] The polyelectrolyte contribution to ΔG, ΔG_{pe} ,was calculated from equation (3.2)
[b] The non-ionic contribution to ΔG, ΔG_t was calculated from equation (3.3)
[c] The calculated value for ΔG is the sum of ΔG_t and ΔG_{pe} (equation (3.3))
[d] The value for ΔG_{obs} is from the ITC results.
[e] The percent ionic contribution was calculated from $\Delta G_{pe}/\Delta G_{calc}$

3.4.2 Circular Dichroism

CD experiments were used to detect any structural changes in the DNA or in the complete H1 protein ($H1^0$) or its domain peptides ($H1^0$-C and $H1^0$-G) upon complex formation. Figure 3.6 shows representative CD spectra for the complete $H1^0$ protein, CT-DNA, and a 0.5:1 protein/CT-DNA complex. (Figures 3.7A and 3.8A show the same spectra for $H1^0$-C and $H1^0$-G and their complexes with CT-DNA.) Figure 3.5B again shows the complex spectrum for $H1^0$/CT-DNA, as well as two calculated spectra: 1) the summation of the CD spectra for free $H1^0$ and free CT-DNA, and 2) the difference spectrum obtained by subtracting the summation spectrum from the actual spectrum for the complex. From the data shown in Figure 3.6A it is clear that both $H1^0$ and CT-DNA

exhibit significant structure. The CT-DNA spectrum shows a positive molar ellipticity at 280 nm and a negative molar ellipticity in the vicinity of 245 nm. The DNA spectrum is consistent with CD spectra previously reported for CT-DNA[60,61]. The CD spectrum obtained here for $H1^0$ very closely resembles the CD spectrum reported for H1 by Barbero *et al*[62]. Our $H1^0$ spectrum also closely resembles the CD spectrum reported for H5 by Carter and van Holde and it is consistent with the structural contributions of the globular domain, $H1^0$-G[63]. We estimate that approximately 31 % of the $H1^0$ residues are in α-helices, and 13 % of the $H1^0$ residues are in β-turns, while $H1^0$-C is completely disordered, and $H1^0$-G has approximately 70 % of its residues in α-helices and 30 % of its residues in β-turns in agreement with Cerf et al[64]. This estimation was based on the assumption that N- and C- tails have no secondary structure in solution. The structure found here for $H1^0$ exceeds the estimates of 8-9% α-helices and 3-4% β-turns reported previously [33]. Nevertheless, it is clear from the spectrum for the $H1^0$/CT-DNA complex in Figure 3.6A and the calculated spectra shown in Figure 3.6B that the DNA structure is dramatically changed in the $H1^0$/DNA complex, while the $H1^0$ structure remains relatively unchanged. Specifically the DNA peak at approximately 280 nm is completely lost. The DNA negative ellipticity at approximately 245 nm is canceled out by the protein positive ellipticity in the same wavelength range. In contrast, the characteristic CD spectrum for the protein is almost unchanged, especially over the 200 to 230 nm range in the complex. Clearly the $H1^0$ α-helix and β-turn structure persists in the $H1^0$/CT-DNA complex.

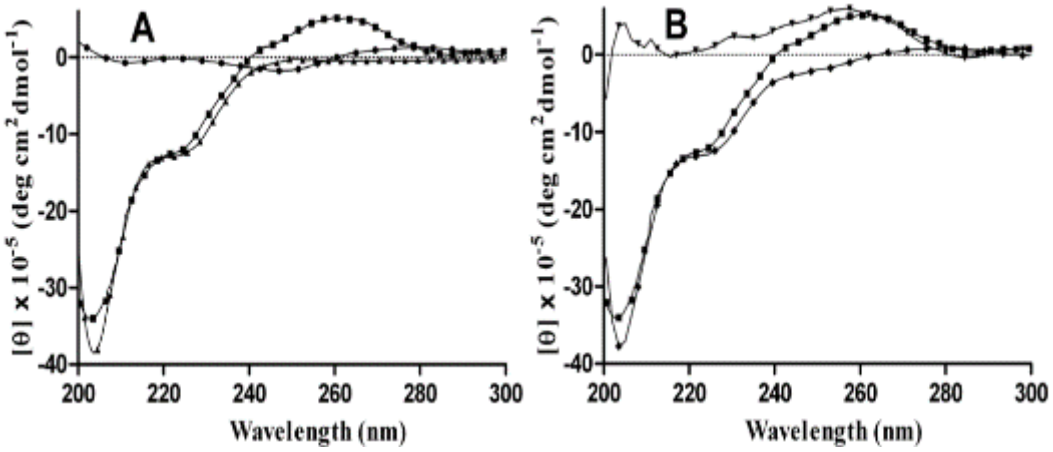

Figure 3.6 Panel **A** shows the CD spectra for the complete protein $H1^0$ (-▲-), CT-DNA (-●-), and the partially saturated $H1^0$/CT-DNA complex (-■-).

Panel **B** shows the CD spectrum for the complex (-■-) along with the calculated summation (-♦-) and difference (-▼-) spectra.

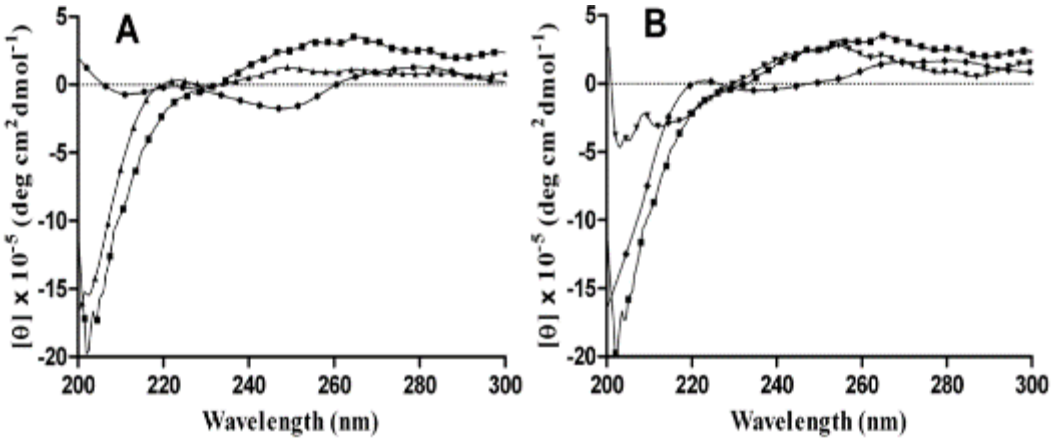

Figure 3.7 Panel **A** shows the CD spectra for the complete protein $H1^0$-C (-▲-), CT-DNA (-●-), and the partially saturated $H1^0$-C/CT-DNA complex (-■-).

Panel **B** shows the CD spectrum for the complex (-■-) along with the calculated summation (-♦-) and difference (-▼-) spectra.

Figure 3.7 Panel A shows representative CD spectra for the C-terminal domain $H1^0$-C, CT-DNA, and a 0.5:1 $H1^0$-C/DNA complex. Figure 3.7 Panel B again shows the complex spectrum for $H1^0$-C/CT-DNA, as well as two calculated spectra: 1) the summation of the CD spectra for free $H1^0$-C and free CT-DNA, and 2) the difference

spectrum obtained by subtracting the summation spectrum from the actual complex spectrum. Again, from the data shown in Figures 3.7 Panel A and 3.7 Panel B, it is clear that the C-terminal domain, $H1^0$-C, is very disordered with a featureless CD spectrum from 220 to 300 nm. It is also clear that $H1^0$-C dramatically changes the structure of the DNA in the $H1^0$-C/CT-DNA complex. Specifically, the DNA peak at approximately 280 nm is unchanged in intensity; however, the DNA trough at 245 nm is completely lost with the complex ellipticity at 245 nm being much greater than the sum of the $H1^0$-C and DNA features at this wavelength. Although in this case, it is difficult to determine whether the increase in structure of the complex, as evident from the increase in ellipticity at wavelengths above 245 nm, is due to a change in the structure of the bound protein, the complexed DNA, or both.

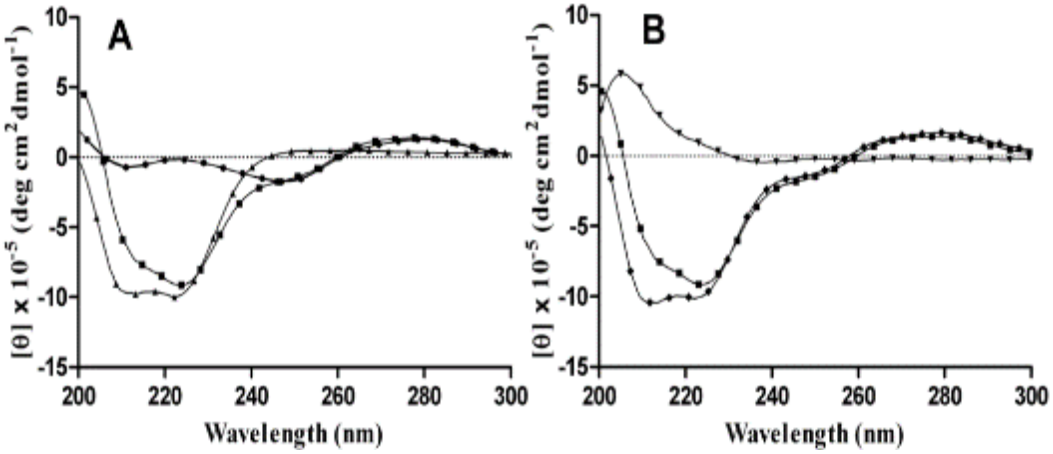

Figure 3.8 Panel **A** shows the CD spectra for the complete protein $H1^0$-G (-▲-), CT-DNA (-●-), and the partially saturated $H1^0$-G/CT-DNA complex (-■-).

Panel **B** shows the CD spectrum for the complex (-■-) along with the calculated summation (-♦-) and difference (-▼-) spectra.

Figure 3.8 Panel A shows representative CD spectra for the globular domain, $H1^0$-G, CT-DNA, and a 0.5:1 $H1^0$-G/CT DNA complex. Figure 3.8 Panel B again shows the complex spectrum for $H1^0$-G/CT-DNA, as well as two calculated spectra: 1) the summation of the CD spectra for $H1^0$-G and CT-DNA, and 2) the difference spectrum obtained by subtracting the summation spectrum from the actual complex spectrum. From the data shown in Figures 3.8 Panel A and 3.8 Panel B it is clear that the G-terminal domain, $H1^0$-G, is highly structured with approximately 70% α-helices and 30% β-turns, which is in agreement with the previously reported CD spectra[64]. The spectrum for the $H1^0$-G/CT-DNA complex (shown in Figure 3.8 Panel A) exhibits significantly reduced ellipticity in the region between 210 and 222 nm, indicating a conformational change in the globular domain upon binding to DNA.

3.5 Discussion

Linker histones are known to affect chromatin dynamics at the DNA replication and transcription sites in eukaryotes[65]. While it is known that H1 binds to DNA at the point where it either enters or exits the nucleosome[3], the structure of the linked nucleosome complex and the energetics responsible for H1 binding are unknown. In this study the thermodynamic parameters for the binding of intact linker histone $H1^0$ and its globular ($H1^0$-G) and C-terminal ($H1^0$-C) domains to DNA have been determined. The real surprise in the energetics is that the enthalpy change for formation of the $H1^0$/DNA or the $H1^0$-C/DNA complex is very unfavorable ($\Delta H° \approx +22$ kcal/(mol H1)) and that $H1^0$/DNA complex formation is driven by a very large change in entropy ($-T\Delta S° \approx -30$ kcal/(mol H1)). We have also determined that the interactions between $H1^0$ (or its globular and C-terminal domains) and DNA result in subtle changes in either the protein

or DNA structure in the H1/DNA complex. All of these results are new and are discussed in detail below.

Figure 3.4 presents a classic example of isothermal enthalpy-entropy compensation driven by changes in ionic strength. At the lowest salt concentrations (0.03 and 0.06 M), the absolute values of the enthalpy and entropy change terms, $|\Delta H°|$ and $|-T\Delta S°|$, for binding $H1^0$ to CT-DNA increase with increasing salt concentration. Upon further increases in the salt concentration, the trend is reversed and the absolute values of the enthalpy and entropy change terms, $|\Delta H°|$ and $|-T\Delta S°|$, for binding $H1^0$ to CT-DNA decrease with increasing salt concentration. Overall, the absolute value of the free energy change, $|\Delta G°|$ decreases monotonically and linearly with increasing salt concentration. There are several references in the literature that indicate that buffers containing 0.07 M $[K^+]$ or $[Na^+]$ are ideal for the study of protein DNA interactions[66,67]. Coincidentally, that is approximately the salt concentration where we see a break (0.068 M $[K^+]$) in our plots of $\Delta H°$ and $-T\Delta S°$ vs. $[K^+]$ as shown in Figure 3.4.

ITC studies demonstrated that the binding of complete protein ($H1^0$) and its C-terminal domain ($H1^0$-C) to CT-DNA was accompanied by a large unfavorable enthalpy change which is compensated by an even larger favorable entropy change. The ITC results for the addition of $H1^0$-G into CT-DNA exhibit no significant heat of interaction (Figure 3.2). This suggests that either the $H1^0$-G/DNA complex does not form (under the conditions of these experiments) or more likely that the complex forms with a very small or near zero change in enthalpy ($\Delta H° \approx 0$ kcal/(mol $H1^0$-G)) at 25° C. The large positive $\Delta S°$ and positive $\Delta H°$ terms for the interaction of $H1^0$ or $H1^0$-C are consistent with changes in the structure of the protein or DNA and/or the release of ordered water

molecules or ions from the DNA backbone and grooves or from the protein upon complex formation with either $H1^0$ or $H1^0$-C. Our CD results appear to rule out any significant macromolecular structural change contributions to the large positive $\Delta S°$ term. The slopes, $-Z\psi$, from the salt dependence data shown in Figure 3.5, *i.e.*, a plot of log K_a vs log $[K^+]$, indicate that the number of ions released upon H1/DNA complex formation is very small (< 1) [30]. The $-Z\psi$ values determined for binding $H1^0$ or $H1^0$-C to linear high molecular weight DNA are -0.89 and -0.82 respectively. In combination, these results point to the loss of bound water as the primary source of the large positive $\Delta S°$ term. Using equation 3.1 and the data in Figure 3.5, we have parsed the free energy change into electrostatic, ΔG_{pe}, and non-electrostatic, ΔG_t, contributions. These data including the calculated and observed values for ΔG are listed in Table 3.2. The agreement between the ΔG_{obs} and ΔG_{calc} is excellent, and the electrostatic contribution to the overall free energy change varies from a maximum of 17 % at 0.03 M $[K^+]$ to a low of 6 % at 0.320 M $[K^+]$. The largest electrostatic contributions to the overall free energy are observed as expected at the lowest salt concentrations, and the smallest electrostatic contributions to the overall free energy change are observed at the highest salt concentrations.

We have used a multiple equivalent sites model to fit the ITC titration data. We used polyelectrolyte theory and the ITC equations developed by Campoy to rule out neighboring site interactions[58]. In effect, the linear DNA is simply represented here as a continuous lattice of non-interacting binding sites. We have used the ITC endpoints and the concentrations of the titrant (ligand or protein) and titrate (DNA bp) to calculate the binding site sizes for $H1^0$ and $H1^0$-C. Whether a binding site is fully covered by bound

protein remains a question. What is known is that there is no interaction between adjacent sites, and the value of the interaction enthalpy parameter, Δh, is zero[58]. One anomaly that is observed in the thermograms for the titration of CT-DNA with either $H1^0$ or $H1^0$-C is a more endothermic peak in the thermogram that is seen just before the DNA is saturated with protein. This effect is shown in Figure 3.3, which includes a fit for the multiple equivalent site model (one site model), a fit for a two non-equivalent sites model, and an area above the equivalent site model line that represents the enthalpy change for rearrangement of bound H1 near saturation. The two non-equivalent sites shows very high uncertainty in the second site parameters as it is based mostly on a single data point, and the rearrangement reaction cannot be fit as a function of added ligand. In actuality, the overall binding process is discontinuous in that the rearrangement reaction only occurs after a critical point is reached in the titration. The rearrangement of bound H1, perhaps a sliding of H1 along the DNA, comes at a cost in ΔH° of approximately 7.5 kcal/mol. Binding of the complete $H1^0$ protein or its C-terminal peptide to ds-DNA is non-specific (i.e. the protein does not target a specific DNA sequence or region)[68]. As the protein (or peptide) concentration approaches DNA saturation, randomly bound protein must be "rearranged" along the DNA in order to minimize the number of partially obscured H1 binding sites and thereby maximize the number of available protein binding sites. This implies that all sites are filled with no overlap of the protein or peptide between adjacent binding sites. This observation is consistent with a recently published small angle X-ray diffraction study which revealed that H1 binding to DNA is a two-step process[69]. The first step, the non-specific electrostatic interaction between the poly-cationic H1 and poly-anionic DNA leads to the formation of primary assemblies having

long columnar hexagonal structure. The second step, the successive rearrangement of H1 molecules in the formed assemblies results in shorter columnar hexagonal structures[69].

CD results indicate that the two peaks attributed to CT-DNA, *i.e.,* 245 and 280 nm, were both attenuated upon formation of the $H1^0$/CT-DNA (or $H1^0$-C/CT-DNA) complex. This is consistent with a conformational change in DNA that is induced by $H1^0$ or $H1^0$-C binding. This result is consistent with several previous studies which have suggested that binding the H1 C-terminal domain to DNA causes bending of chromatosomal DNA to facilitate a stem-like structure[70,71]. In the case of the globular domain, the CD results indicate that the binding of $H1^0$-G to CT-DNA causes a significant decrease in negative ellipticity between 210 and 222 nm, not only suggesting that there is a strong interaction between $H1^0$-G and CT-DNA, but that there is some unfolding of the globular domain as it binds to DNA. These CD results also indicate that binding of $H1^0$-G to CT-DNA has no detectable effect on the DNA conformation in the complex.

In conclusion, this study clearly shows that high affinity complexes are formed between $H1^0$ and $H1^0$-C with CT-DNA. The formation of these protein (or peptide) DNA complexes is driven primarily by large favorable entropy changes. It would appear that these large positive entropy changes must result primarily from the expulsion of bound water molecules from the binding interface.

CHAPTER IV

EXPLORING THE ENERGETICS OF HISTONE H1.1 AND H1.4 DUPLEX DNA INTERACTIONS

4.1 Abstract

H1.1 and H1.4 bind tightly to both short DNA oligomers and to CT-DNA ($K_a \approx 1\times10^7$). Binding is accompanied by an unfavorable enthalpy change ($\Delta H° \approx +22$ kcal/mol) and a favorable entropy change ($-T\Delta S° \approx -30$ kcal/mol). The T_m for the H1.4/CT-DNA complex is increased by 9°C over the T_m for the free DNA. H1.4 titrations of the DNA oligomers yield stoichiometries (H1/DNA) of 0.64, 0.96, 1.29, and 2.04 for 24, 36, 48, and 72-bp DNA oligomers. The stoichiometries are consistent with a binding site size of 37±1 bp. CT-DNA titration data are consistent with binding site sizes of 32 bp for H1.1 and 36 bp for H1.4. The heat capacity changes, $\Delta C°_p$, for formation of the H1.1 and H1.4/CT-DNA complexes are -160 cal mol^{-1} K^{-1} and -192 cal mol^{-1} K^{-1} respectively. The large negative $\Delta C°_p$ values indicate the loss of water from the protein DNA interface in the complex.

4.2 Introduction

H1 interactions with DNA have attracted significant interest due to the involvement of H1 in chromatin compaction and the fact that H1 is highly modified in cancer cells[72,73]. Histone H1 has also been described as a transcription repressor as it

limits the access of transcriptional factor proteins to DNA[46,47,73]. To date, eleven histone H1 subtypes have been identified in mammals. In both humans and mice, five of the somatic subtypes (H1.1-H1.5) are classified as cell cycle-dependent or replication dependent. The expression of these genes is linked to a particular phase of the cell's life cycle. The other somatic subtype ,$H1^0$, is classified as replication-independent or replacement subtype and it is expressed throughout the cell's life cycle in order to maintain a replacement pool of H1 linker histone [1]. H1.1 (mouse H1c) and H1.4 (mouse H1e) carry the highest positive charges and exhibit the strongest DNA binding affinities[1,17]. H1.4 is seven amino acids longer (219 vs. 212) than H1.1, with the two sequences exhibiting 89% identity in the N-terminal and globular domains [1]. The larger net positive charge for H1.4 (+59) compared to H1.1 (+55) is due to four additional lysines in the H1.4 sequence. In comparison, $H1^0$ (mouse H1.0) has a net positive charge at neutral pH of +53 and exhibits 61% sequence identity in comparison to the N-terminal and globular domains of H1.1. The C-terminal tails of H1.1, H1.4 and $H1^0$ (wherein most of the charged residues reside) are apparently scrambled and exhibit no sequence homology from a BLAST sequence comparison, http://blast.ncbi.nlm.nih.gov/Blast.cgi[74,75]. Although these three subtypes show similar net charges, high sequence identity in the N-terminal and globular domains, and only subtle changes in overall amino acid composition, it has been reported that the three subtypes exhibit different DNA binding affinities and occupy different DNA binding site sizes [1,17,76]. If there are real differences in the DNA binding affinities or binding site sizes for the H1.1, H1.4 and $H1^0$ proteins, the differences must be attributed to the

different sequences and or charge and charge distribution in the C-terminal domains of these proteins.

It is generally accepted that H1, or linker histone, binds to DNA as it enters and/or exits the nucleosome[3,6,77,78]. Two different models for H1 binding place the H1 protein across the nucleosome with H1 interacting with two or three patches of the nucleosomal DNA on the same side of the nucleosome[27] or alternatively locate the H1 so that it binds to a continuous and more linear linker DNA region[79]. The generally accepted biological function for bound H1 is that it prevents nucleosomal DNA from either unraveling from or sliding off the nucleosome histone core complex and that it participates in limited chromatin compaction by tethering adjacent nucleosomes together[6]. Although the structure of the nucleosome (histone core/DNA complex) and the interactions of the core histone proteins with DNA are relatively well established[6-8,13,80], several questions remain regarding the structural and functional roles of H1 histone in chromatin. Even such basic information as the H1 DNA binding affinity, the H1 DNA binding site size, and the enthalpy and entropy changes for the formation of the H1/DNA complex are largely absent from the current literature. In the previous chapter, we reported on the thermodynamics for binding $H1^0$, $H1^0$-C (the C-terminal domain), and $H1^0$-G (the globular domain) to highly polymerized calf thymus DNA [76]. Perhaps the most surprising result of the previous study was the fact that the $\Delta H°$ values for formation of the $H1^0$/DNA and $H1^0$-C/DNA complexes were highly endothermic ($\Delta H° \approx +22$ kcal/mol). In contrast, a recent publication by Caterino *et. al.,* reported that the enthalpy change for binding the H1 carboxyl terminal domain to linker DNA is – 12 kcal/mol.

The Caterino enthalpy data were obtained indirectly from a van't Hoff analysis of K_a vs. T experiments with the K_a values determined in gel (GMSA) experiments[81].

In the present study we have used isothermal titration calorimetry (ITC), differential scanning calorimetry (DSC), and CD spectropolarimetry to determine the thermodynamic signatures and structural changes that accompany H1.1 and H1.4 binding to short DNA oligomers and/or H1.1 and H1.4 binding to highly polymerized calf-thymus DNA. The results from these studies on H1.1 and H1.4 were compared to published results of similar studies done on $H1^{076}$. In our ITC studies, we found that the H1.1 and H1.4 bind to CT-DNA with approximately the same affinity observed previously for $H1^0$, ($K_a \approx 1\times10^7$). We also observed large endothermic enthalpy changes for the formation of the H1.1/DNA and H1.4/DNA complexes ($\Delta H \approx +22$ kcal/mol H1.1 or H1.4) which were again similar to the $\Delta H°$ value observed for the formation of the $H1^0$/CT-DNA complex[76]. The binding site sizes for H1.1 and H1.4 were determined to be 32bp and 36bp, respectively. The change in the binding site size between H1.1 and H1.4 seems anomalously large since the two proteins are different by only 3% in molecular weight. Interestingly, the binding site size for the smallest H1 subtype, $H1^0$, was also determined to be 36 DNA bp[76]. Titrations of short double stranded DNA oligomers (having from12 to 72 bp) yield H1.4 binding stoichiometries (H1.4/DNA bp) that are in excellent agreement with the H1.4 binding site size determined in H1.4 CT-DNA titrations, DSC determined T_m values indicate that the melting temperature of the CT-DNA in the H1.4/CT-DNA complex is increased by 9 °C, a result that is in agreement with the high H1.4 DNA binding affinity. CD experiments indicate that DNA is restructured upon complex formation with either the H1.1 or H1.4, again a result that is

similar to the changes in CD spectrum that accompany formation of the $H1^0$/CT-DNA complex. In contrast, the H1 protein structure (H1.1, H1.4, or $H1^0$) is largely unchanged upon formation of the H1/DNA complexes. ΔCp values determined for the formation of the H1.1 and H1.4 CT-DNA complexes are large and negative (e.g. -160 or -192 cal mol^{-1} K^{-1}). Large negative ΔCp values had also been observed previously for the formation of the $H1^0$/CT-DNA complex. These large negative ΔCp values are indicative of the loss of water from either the DNA or the protein as the H1/DNA complexes are formed. All of these results are placed in the context of the current histone literature and are discussed more fully in the later sections of this work.

4.3 Materials and Methods

4.3.1 Proteins and DNA samples

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 (Novagen) expression vector which contained the DNA sequence for either the mouse H1.1 protein or the mouse H1.4 protein. The methods for expression and purification have been described elsewhere[54,82]. The pure protein fractions were concentrated using a Savant SPD 111V speed vac system for 4 hrs at 35°C to remove the HPLC solvent (5 % acetonitrile/95% water). Calf thymus DNA type I was purchased from Sigma (St. Louis, USA) and used without further purification. Short double stranded DNA's designed for the H1.4 ITC binding studies were purchased from Midland Certified reagent Company (Midland TX, USA). (The sequences for the short duplex DNA's are given in Table 3.1) The protein and DNA stock solutions were prepared by dissolution of the concentrated and/or dried purified protein or DNA samples in the standard BPES buffer (30 mM K_2HPO_4/ KH_2PO_4

(pH = 7.0), 1 mM EDTA, and 100 mM KCl). The protein and DNA stock solutions were exhaustively dialyzed against the sample buffer (24h) at 4˚C, using a 1000 Mw cutoff dialysis membrane (Thermo Scientific, USA). Calf thymus DNA concentrations in base pairs (bp) were determined using measured absorbance at 260 nm and a molar extinction coefficient of $\varepsilon_{260}=1.31\times10^4$ bp $M^{-1}cm^{-1}$ [55]. The approximate average molecular weight of the CT-DNA was 8.42×10^3 kDa (Sigma, St. Louis, USA). Short DNA oligomer concentrations were determined using absorbance values at 260nm and molar extinction coefficients (ε_{260}) of 191202.4, 380585.4, 571000.0, 763380.0, and 1142143.3 $M^{-1}cm^{-1}$ for 12-mer, 24-mer, 36-mer, 48-mer, and 72-mer respectively. H1.1 and H1.4 concentrations were determined by UV absorbance measurements at 205 nm and using extinction coefficients of 6.37×10^5 $M^{-1}cm^{-1}$, and 6.11×10^5 $M^{-1}cm^{-1}$ for H1.1 and H1.4 respectively[83]. The approximate molecular weights for the H1.1 and H1.4 were estimated from their sequences using the ExPASy ProtParam tool (http://web.expasy.org/protparam)[56] : Mw (H1.1) ≈ 21.3 kDa, Mw (H1.4) ≈ 21.9 kDa.

Table 4.1 Short synthetic duplex DNA's designed for H1.4 ITC binding studies

12-mer (T_m=36.0 ºC)
5′-ATCAAGCTACGC-3′ 3′-TAGTTCGATGCG-5′
24-mer (T_m=57.4 ºC)
5′-ATCAAGCTACGCCTGAAGAGTCTG-3′ 3′-TAGTTCGATGCGGACTTCTCAGAC-5′
36-mer (T_m=67.9 ºC)
5′-ATCAAGCTACGCCTGAAGAGTCTGGTGAGCAAGGGT-3′ 3′-TAGTTCGATGCGGACTTCTCAGACCACTCGTTCCCA-5′
48-mer (T_m=72.2 ºC)
5′-ATCAAGCTACGCCTGAAGAGTCTGGTGAGCAAGGGTACTCTGGTGTAG-3′ 3′-TAGTTCGATGCGGACTTCTCAGACCACTCGTTCCCATGAGACCACATC-5′
72-mer (T_m=76.6 ºC)
5′- ATCAAGCTACGCCTGAAGAGTCTGGTGAGCAAGGGTACTCTGGTGTAGACCAAGTGCACTGGTCCTTCAATC-3′ 3′- TAGTTCGATGCGGACTTCTCAGACCACTCGTTCCCATGAGACCACATCTGGTTCACGTGACCAGGAAGTTAG-5′

4.3.2 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a Microcal VP-ITC (Northampton, MA, USA). All titrations were performed by overfilling the ITC cell with approximately 1.5 mL of dilute DNA solution (nominally 540 μM in bp). During a typical ITC titration, approximately 250 μL of the H1.1 or H1.4 protein solution was added to the DNA solution in the micro-calorimeter cell in 25 to 50 increments delivered at 600 second intervals. All of our ITC experiments were performed in triplicate and at three different temperatures, 15˚C, 25˚C, and 35˚C. The integrated heat/injection data were fit for to an appropriate thermodynamic model using CHASM data analysis software developed in our laboratory [57]. The non-linear regression fitting process yields best fit parameters for K (or ΔG), ΔH, ΔS, and n for each independent interaction.

4.3.3 Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) experiments were performed using a Microcal VP-DSC (Northampton, MA, USA). The melting temperatures, T_m, were determined for CT-DNA and H1.4/CT-DNA in separate DSC experiments. The nominal DNA concentration for these experiments was 2.16 mM (in bp), and the complex was prepared by the addition H1.4 protein to the 2.16 mM DNA solution to yield a final H1.4 concentration of 0.012 mM. (Assuming an approximate DNA binding site size of 36 bp/H1.4, this solution had a nominal composition of 0.2 equivalents of H1.4 protein per equivalent of DNA sites.) The temperature scan range in these experiments was 10 to 110°C with a scan rate of 90°C/hr. The CT-DNA denaturation over this temperature range was irreversible; therefore only one heating scan was performed in each

experiment. The irreversibility in these experiments was due to aggregation and precipitation of the high molecular weight DNA. Because the aggregation and precipitation is slow and displace in time from the helix to coil transition, the determination of the thermodynamic parameters for DNA denaturation is valid. The DSC data from the first scan were therefore fit for $\Delta H°_{cal}$, $\Delta H°_{VH}$, and the melting temperature (T_m) using the "two-state" model in Origin 7.1 software (Microcal, Northampton, MA).

4.3.4 Circular Dichroism

CD experiments were performed using an Olis DSM 20 spectropolarimeter (Bogart, GA). CT-DNA and protein solutions were prepared with a nominal absorbance of 0.5 AU in a nominal 50 mM KBPES buffer (pH=7.0, 10 mM KCl, 30 mM Phosphate, 0.1 mM EDTA). In these experiments, the nominal concentration of the protein was 1.5 μM, and the concentration of the CT-DNA was 3.0 μM in H1 binding sites. The 0.5:1 mole ratio for protein to DNA binding sites was chosen to avoid the complications that appear near the endpoint in the ITC titrations. We used diluted solutions of both protein and DNA, and used excess moles of DNA to prepare complex samples for the CD experiments to minimize complex aggregation. CD spectra were collected over a wavelength range of from 200 to 300 nm (with measurements every 0.5 nm) in a 1 cm path length cuvette at room temperature. The spectra represent the average of three scans which were processed using PRISM software (graph-Pad Prism Software, San Diego, CA).

4.4 Results

4.4.1 Isothermal Titration Calorimetry

Figure 4.1 shows typical ITC data for titrations of the H1.1 and H1.4 proteins into highly polymerized CT-DNA at 25˚C. The ITC thermograms were fit using non-linear regression techniques to a multiple independent site model (one site model) and the average best-fit parameters are listed in Table 4.1. We also attempted to fit these titration data to a nearest neighbor exclusion model[58]; however, the multiple sites do not appear to be interacting and do not exhibit either positive or negative cooperativity. This was determined from the independence of ΔH on degree of saturation, ($\Delta h \approx 0.0$). ITC data indicate that although H1.1 or H1.4 have a high binding affinity ($K \approx 10^7\ M^{-1}$) for CT-DNA, the enthalpy change is very unfavorable ($\Delta H \approx +22$ kcal/(mol H1.1 or H1.4)), and complex formation is driven by a large favorable entropy change ($-T\Delta S \approx -30$ kcal/mol). The anomalously large endothermic heat values that are observed as one or two points that are significantly above the best fit line and just prior to the end point in both ITC titrations (Figure 4.1 A and B) are attributed to overcoming a steric interaction in which a bound protein is partially occupying two adjacent sites, and this protein must be relocated in order to fully populate all of the potential protein binding sites on the DNA. The additional endothermic heat observed near saturation represents the energy cost of relocating the one or more already bound proteins that are unevenly distributed along the linear lattice of the DNA. In effect as H1.1 (or) H1.4 binds non-specifically (electrostatically) to a long DNA molecule, the placement of the proteins is random and can result in multiple partial binding sites that are vacant. The binding-site size, or the number of base pairs occupied per bound H1.1 or H1.4 protein, was calculated from the

ratio of added protein at the titration endpoint to the total number of DNA base pairs in the calorimeter. The saturation stoichiometry indicated that one mole of H1.1 binds to 32 bp, and one mole of H1.4 binds to 36 bp.

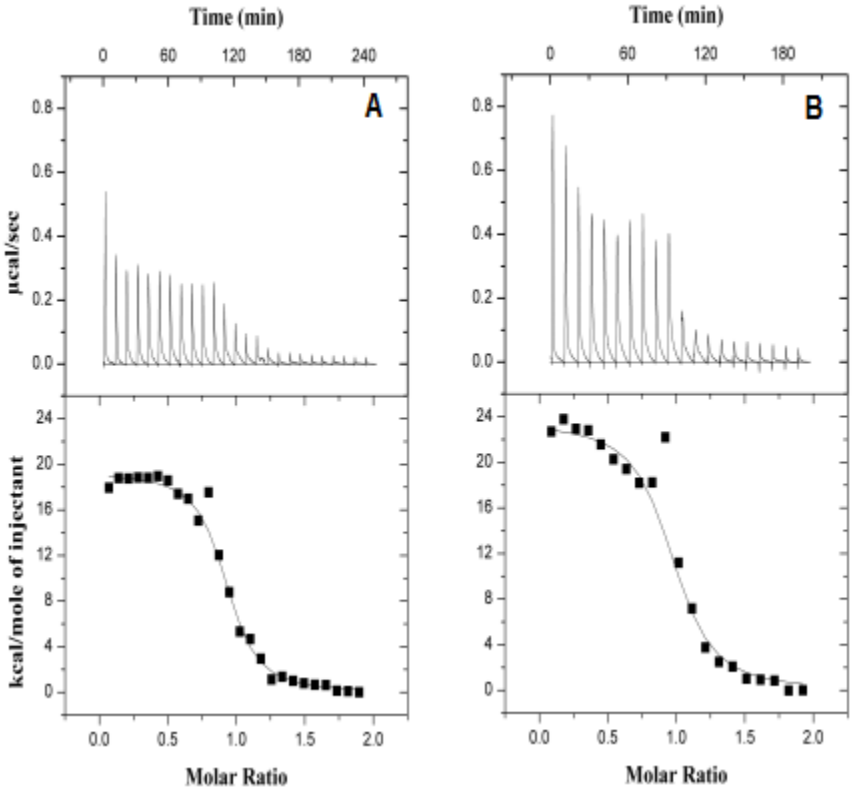

Figure 4.1 ITC thermograms of H1.1 and H1.4 binding to CT-DNA at 25 ºC, pH 7.0

Panel **A** shows a typical ITC titration for the addition of H1.1 to highly polymerized CT-DNA. The upper half of Panel **A** shows the baseline-corrected raw ITC signal for 25 injections of a dilute H1.1 protein solution (10 µL of 135 µM H1.1) into the ITC cell filled with a dilute solution of CT-DNA (416 µM bp, 13 µM in H1.1 binding sites). The lower half of Panel **A** shows the apparent ΔH° for each injection (-■-) along with the best-fit non-linear regression line (—) for a simple one site binding model. Panel **B** shows a typical ITC titration for the addition of H1.4 to highly polymerized CT-DNA. The upper half of Panel **B** shows the baseline-corrected raw ITC signal for 20 injections of a dilute H1.4 protein solution (14 µL of 140 µM $H1^0$-C) into the ITC cell filled with a dilute solution of CT-DNA (540 µM bp, 15 µM in H1.4 binding sites). The lower half of Panel **B** shows the ΔH° for each injection (-■-) along with the best-fit non-linear regression line (—) for a simple one site binding model.

Table 4.2 ITC derived thermodynamic parameters for H1.1 and H1.4 histone protein binding to CT-DNA in 100 mM [K^+] BPES pH 7.0 at 25°C.

H1 subtype	K_a (M^{-1}) $\times 10^{-6}$	ΔG° (kcal/mol)	ΔH° (kcal/mol)	-TΔS° (kcal/mol)	Binding site Size (bp)
H1.1	6.30±0.14	-9.27	18.22±0.83	-27.49	32
H1.4	3.20±0.28	-8.87	22.52±0.50	-31.37	36

All ITC experiments were performed in triplicate in 100 mM [K^+] BPES buffer at pH 7.0 and 25˚C. The integrated heat/injection data were fit for to a one site thermodynamic model using CHASM data analysis software developed in our laboratory. Errors listed are the standard deviations for the best fit parameters *K* and ΔH° determined in triplicate experiments. Effective binding site size in base pairs was calculated from the titration endpoint, the DNA concentration in base pairs, and the assumption that saturation stoichiometry is 1:1 (H1: DNA sites).

To further probe the accuracy of binding site size in more detail, we titrated H1.4 into short random sequence ds-DNA solutions (12 bp, 24 bp, 36 bp, 48 bp, and 72 bp). In Figure 4.2, we have plotted the apparent ΔH values for each injection along with the best-fit non-linear regression line against to the molar ratio (H1/DNA oligomer) for each titration. Table 4.2 lists the stoichiometries and the enthalpy values for the titrations of H1.4 into short ds-DNA sequences. These results are discussed in more detail in the discussion section of this chapter. We also performed a temperature dependent study in which the experiments shown in Figure 4.1 were repeated at 15°C and 35°C. In Figure 4.3, we have plotted the values for ΔH binding H1.1 and H1.4 to CT-DNA as a function of temperature. Both H1.1 and H1.4 showed a linear decrease in enthalpy change with increasing temperature. The slope of the linear least squares fits to all data for both H1.1 and H1.4 yielded large negative ΔC_p values of -160 and -192 cal mol^{-1} K^{-1} respectively.

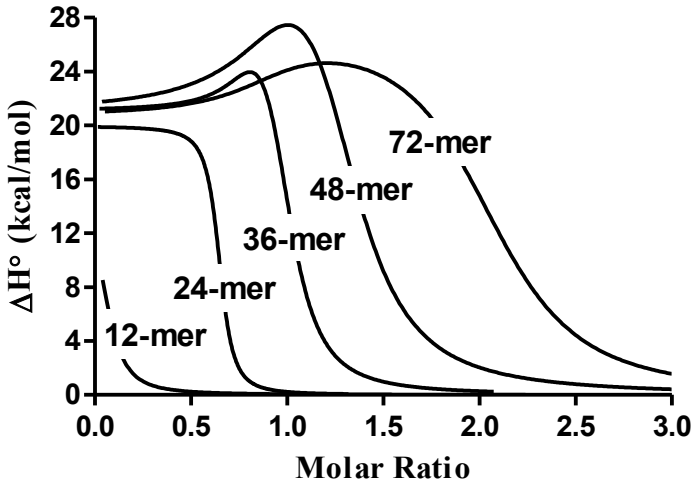

Figure 4.2 ITC data for the addition of H1.4 into a dilute solution containing a short DNA duplex oligomer.

The data for the 12-bp and 24-bp oligomers were fit to a one site thermodynamic model whereas the 36-bp , 48-bp , and 72-bp oligomer data were fit to a two site thermodynamic model. The best fit parameters (n_1, $\Delta H°_1$, n_2, and $\Delta H°_2$) are listed in Table 3.3.

Table 4.3 ITC derived thermodynamic parameters for H1.4 histone protein binding to short ds-DNA oligomers in 100 mM [K^+] BPES pH 7.0 at 25 °C.

Oligo	n_{total}	Binding site Size (bp)	n_1	$\Delta H°_1$ (kcal/mol)	n_2	$\Delta H°_2$ (kcal/mol)
24-mer	0.64	37.5	0.64±0.00	20.0±0.2	-	-
36-mer	0.96	37.5	0.84±0.06	21.4±0.3	0.12±0.08	55±28
48-mer	1.29	37.2	0.94±0.10	21.1±0.7	0.35±0.10	50±15
72-mer	2.04	35.3	0.87±0.08	20.7±0.6	1.18±0.09	27.8±1.2

All ITC experiments were performed in triplicate in 100 mM [K^+] BPES buffer at pH 7.0 and 25˚C. The integrated heat/injection data were fit for to a fractional site thermodynamic model in which the first fractional reaction is for the simple binding of the protein to the DNA and the second fractional interaction includes both binding and the energetic for rearrangement of the bound protein to maximize the available binding site area. The total n (n_1+n_2) is determined from the titration endpoint. The non-linear regression fits were done using CHASM data analysis software developed in our laboratory. Errors listed are the standard deviations for the best fit parameters n_1, n_2, $\Delta H°_1$, and $\Delta H°_2$ determined in triplicate experiments. Effective binding site sizes in base pairs were calculated from the titration endpoint, the DNA concentration in base pairs and the assumption that saturation stoichiometry is 1:1 (H1: DNA sites).

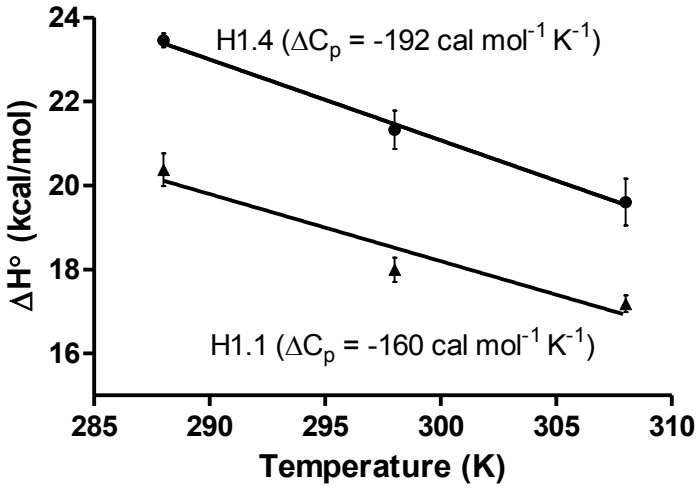

Figure 4.3 A plot of the ITC derived ΔH° values for the formation of the H1.1 and H1.4/CT-DNA complexes vs. temperature.

Data are shown for three temperatures: 15, 25, and 35 ºC. The slopes of the two lines yield estimates for the $\Delta C°_p$ values that accompany the formation of the H1.1 and H1.4 CT-DNA complexes.

4.4.2 Circular Dichroism

CD experiments were used to detect gross structural changes in the DNA and the histone variants upon formation of the histone/DNA complex. Figure 4.4A shows representative CD spectra for the H1.1 histone protein, CT-DNA and the spectrum for the 0.5:1 complex of H1.1/CT-DNA. Figure 4.4B shows representative CD spectra for the H1.4 histone protein, CT-DNA and the spectrum for the 0.5:1 complex of H1.4/CT-DNA. The CT-DNA spectrum shows a positive molar ellipticity at 280 nm and a negative molar ellipticity at 245 nm that is consistent with CD spectra previously reported for CT-DNA [60,61]. The H1 protein spectrum exhibits a negative molar ellipticity at 208 and 222 nm indicative of a significant amount of α-helix and β-turn present in the structure of the histone protein [84]. The CD spectra obtained for the H1.1 and H1.4 closely resemble the CD spectrum reported for H1 by Barbero *et al* [62]. Both H1.1 and

H1.4 also exhibit CD spectra that are similar to the spectrum reported for H5, which includes structural contributions from the globular domain [63]. The complex spectra for H1.1/CT-DNA and H1.4/CT-DNA (Figure 4.4 A and B) exhibit some changes to the DNA structure while the structures of the H1.1 and H1.4 proteins appear to be largely unchanged. Specifically the DNA peak at approximately 280 nm is completely lost. The DNA negative ellipticity at approximately 245 nm is canceled out by the protein positive ellipticity in the same wavelength range. In contrast, the characteristic CD spectrum for the protein (H1.1 and H1.4) is almost unchanged especially over the 200 to 230 nm range in the complex. Clearly this indicates that the H1 α-helix and β-turn structure persists in the H1/CT-DNA complex.

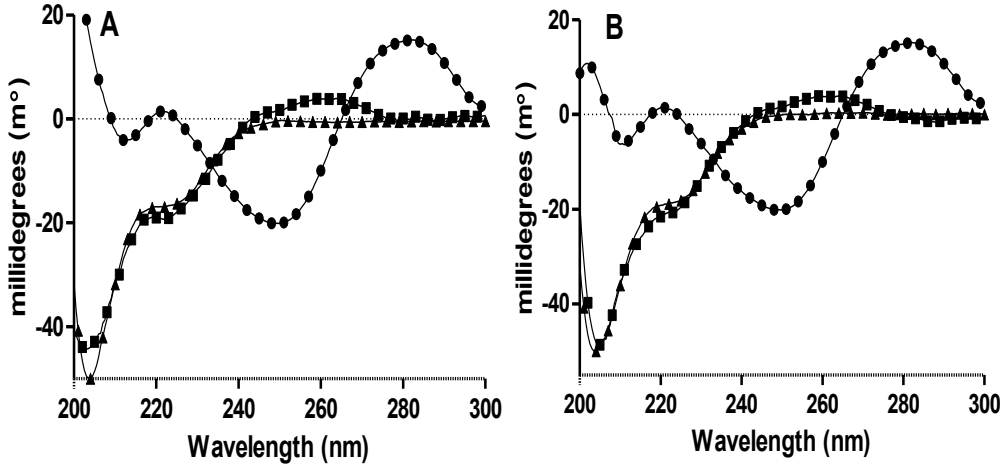

Figure 4.4 Panel **A** shows the CD spectra for the H1.1 protein (-▲-), CT-DNA (-●-), and the 0.5:1 H1.1/CT-DNA complex (-■-).

Panel **B** shows the CD spectrum for the H1.4 protein (-▲-) along with the spectra for CT-DNA (-●-), and the 0.5:1 H1.4/CT-DNA complex (-■-).

4.4.3 Differential Scanning calorimetry

Differential Scanning Calorimetry was used to determine the stability of the histone/DNA complex versus the stability of the uncomplexed CT-DNA. The DSC melting profiles for the thermal denaturation of the CT-DNA as well as the H1.4/CT-DNA complex are shown in Figure 4.5. The CT-DNA thermogram is comprised of a single symmetric peak that has been fit for a "two state" transition (see Figure 4.5A). The CT-DNA melting temperature determined here (84.0°C), is in reasonable agreement with previously reported T_m values for CT-DNA [85,86]. The DSC data shown in Figure 4.5B show melting transitions for CT-DNA and the H1.4/CT-DNA complex after the addition of approximately 0.2 equivalents of protein per equivalent of protein binding sites. This thermogram has been fit for two independent overlapping "two-state" processes with T_m values of 83.7°C and 92.8°C. The lower melting peak is attributed to denaturation of the free CT-DNA in solution while the higher melting peak is attributed to denaturation of the protein stabilized DNA in the H1.4/CT-DNA complex. The lower temperature peak corresponds to the melting of approximately 80% of the total DNA while the higher temperature peak corresponds to melting approximately 20% of the total DNA. These values are consistent with the approximate composition of the sample.

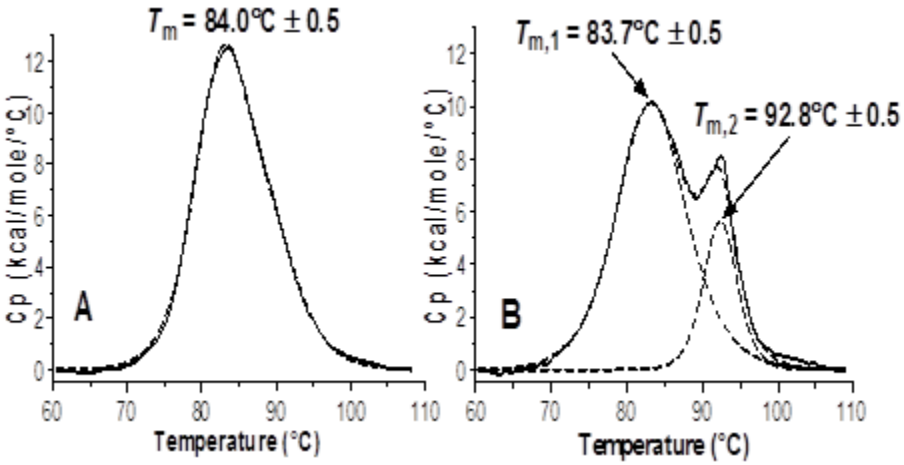

Figure 4.5 DSC thermograms for CT-DNA and CT-DNA/H1.4 complex.

Panel A shows DSC thermogram for the thermal denaturation of CTDNA. The raw excess heat capacity (solid line) has been deconvoluted into a single "non two-state" process (dashed line). Panel B shows DSC thermogram for the thermal denaturation of Histone H1.4 and its complex with CT-DNA in 100mM KCl BPES. The excess heat capacity for the H1.4/CT-DNA complex has been deconvoluted into two independent overlapping "non two-state" processes (dashed line). The lower melting profile is attributed to "free" CT-DNA while the higher melting transition is for the melting of the H1.4/CT-DNA stabilized complex.

4.5 Discussion

The main purpose of this study was to determine the thermodynamic parameters for the binding of H1.1 and H1.4 variants to variable length ds-DNA such as short sequence DNA oligomers and to highly polymerized calf-thymus DNA, and to compare the thermodynamics for these two H1 variants to H1[0] [76]. ITC studies demonstrated that the binding of H1.1 and H1.4 to CT-DNA was accompanied by a large unfavorable enthalpy change which is compensated by an even larger favorable entropy change. The large positive entropy change ($-T\Delta S \approx -30$ kcal/(mol H1)) and positive enthalpy change ($\Delta H \approx +22$ kcal/(mol H1)) terms are consistent with the release of ordered water molecules from the backbone and grooves of the DNA upon H1 complex formation. We have used a multiple equivalent sites (one site) model to fit the ITC titration data. One

anomaly that is observed in the thermograms for the titration of CT-DNA with $H1^0$, H1.1, and H1.4 is one or more larger endothermic peak in the thermogram that are seen just before the DNA is saturated with protein. We attribute the anomalously large endothermic point(s) observed immediately before the end point in the $H1^0$, H1.1, and H1.4 DNA titrations to the rearrangement of non-specifically bound H1 along the length of the DNA. Binding of the H1 protein to ds-DNA is non-specific (*i.e.,* the protein does not target a specific DNA sequence or region) [18,19,68]. As the protein concentration approaches DNA saturation, randomly bound protein may be "rearranged" along the DNA in order to minimize the number of partially obscured H1 binding sites and thereby maximize the number of available protein binding sites. The endpoints in the ITC thermograms were used along with the DNA concentration in base pairs to estimate the number of base pairs covered by the binding of 1 mole of H1 protein. This analysis yielded estimates of 32 bp and 36 bp for the H1.1 and H1.4 binding site sizes respectively. Since some H1 binding sites (or partial binding sites) may be unoccupied at the ITC endpoint (saturation point), these values represent an upper limit on the number of DNA base pairs that may correspond to the actual binding site size.

To further probe the issue of binding site size for the interactions of H1 proteins with duplex DNA, H1.4 was titrated into solutions containing short ds-DNA oligomers having 12, 24, 36, 48 or 72 bp. These data, which are shown in Figure 4.2, have been fit to a thermodynamic model in which the saturated complex is the result of at least two competing processes. Fitting the ITC data to a fractional sites model yields two values for each of the thermodynamic parameters (K_a, $\Delta H°$, and n), with the first set of parameters applicable to the interaction of the protein and DNA at low site occupancy and the second

set of parameters applicable for the interaction of the protein and DNA at high site occupancy or near saturation. The thermodynamic parameters listed in Table 4.3 assume a one to one correspondence between the number of bound protein molecules and the number of sites provided on one mole of the ds-DNA oligomer. In effect the shorter DNA oligomers (12 and 24-mers) were only long enough to bind a short part of a single protein molecule, while the longer DNA oligomers (48 and 72-mers) were long enough to bind more than one complete protein molecule. The end points for the ITC titration curves shown in Figure 4.2 can thus be used to directly estimate the binding site size in DNA base pairs for binding one mole of H1.4. The CT-DNA titration data yielded an approximate binding site size of 36 bp for H1.4, while the short DNA titrations yielded an average value of 36.9 bp. Obviously, the agreement in binding site size as determined in titration experiments done on poly-disperse, highly polymerized CT-DNA and the experiments employing 5 different mono-disperse low molecular weight DNAs is very good.

The anomaly in the enthalpy curve only begins to appear when more than one mol of protein (e.g. 1.2 to 2.0 mols of protein) is bound to a single DNA oligomer. In effect, it is only when the DNA is fully covered with protein (or when available sites are fully populated) with bound protein that the excess endotherm for protein rearrangement along the DNA is observed. In the high molecular weight CT-DNA the anomaly in apparent $\Delta H°$ is limited to one or two points in the thermogram, whereas in the short oligomer titrations the effect is more pronounced since the binding of a second protein (even if it is only partially overlapping available base pairs) occurs almost throughout the titration.

In our previous study of $H1^0$ binding to CT-DNA we had determined that the release of bound counterions (e.g. Na^+, or K^+) upon $H1^0$ complex formation was limited, ($\delta n_{K^+} \approx 1$)[76]. Since the net charges, protein sizes, and amino acid composition of all three proteins are very similar, we have made the assumption that the counterions released upon binding H1.1 or H1.4 to CT-DNA would be similar to what we had previously seen for $H1^0$. The slopes from the temperature dependent ITC data shown in Figure 4.3 for binding H1.1 and H1.4 to CT-DNA, *i.e.*, plots of $\Delta H°$ vs. temperature, are indicative of large negative heat capacity changes, $\Delta C°_p$, for the formation of the H1.1 and H1.4 CT-DNA complexes. We had also observed a large negative heat capacity change for formation of the $H1^0$/CT-DNA complex. Our CD results appear to rule out any significant macromolecular structural change contributions to either the large positive $\Delta S°$ values or to the large negative $\Delta C°_p$ values. Since macromolecular structural changes and the release of bound counterions have been ruled out, these large negative $\Delta C°_p$ values must be due to the release of a large number of water molecules upon formation of the $H1^0$, H1.1, and H1.4 complexes with CT-DNA.

CD results indicate that the two peaks attributed to CT-DNA, i.e. 245 and 280 nm, were both attenuated upon formation of the H1.1/CT-DNA (or H1.4/CT-DNA) complex. This is consistent with a conformational change in DNA that is induced by $H1^0$, H1.1, or H1.4 binding. This result is consistent with several previous studies which have suggested that binding the H1 C-terminal domain to DNA causes bending of chromatosomal DNA to facilitate a stem-like structure[70,71]. The structural stabilization of the H1.4 protein/DNA complex was further characterized using DSC. A single ("two-state") transition occurring at 84°C was observed for the denaturation of the CT-DNA in

the absence of the protein. After the addition of 0.2 equivalents of H1.4 protein per equivalent of DNA protein binding site, a second melting profile appeared with a T_m value of approximately 93°C. This higher melting transition is for the melting of the H1.4 protein stabilized DNA in the complex. The peak area of the higher melting transition is approximately 1/5 of the area for CT-DNA alone, which is consistent with 20% of the binding sites being occupied in the 20% saturated H1.4/CT-DNA complex. The 9°C shift in the melting temperature of the protein bound DNA is consistent with a binding constant of 10^7, which is in excellent agreement with K_a value determined for H1.4 binding in the ITC experiments.

There are two distinctly different models in the current literature which describe the function of H1 linker in compacting and stabilizing the nucleosome[27,79] In the dyad axis model, the linker H1 globular domain is bound to two widely separated DNA domains or patches on one side of the nucleosome. The bound H1 globular domain is approximately centered on the nucleosome's dyad axis[27]. In this model, the two DNA H1 binding sequences are separated along the length of the nucleosomal DNA by from a minimum of approximately 82 bp to a maximum of approximately 168 bp . In the off dyad axis model, the linker H1 globular domain binds to only one patch of DNA located at either the point where DNA is beginning to wrap on the nucleosome or the point where DNA is leaving the nucleosome[79]. In this model only one DNA domain is interacting with globular domain of bound H1, and the H1 binding site consists of a continuous stretch of DNA. In both models, the H1 N-terminal tail binds to a short continuous fragment of nucleosomal DNA, while H1 C-terminal tail binds to a longer continuous fragment of linker DNA bridging between two adjacent nucleosomes. Although both the

nucleosomal DNA and linker DNA are probably bent, these stretches of continuous DNA that interact with the H1 tails should be reasonably well modeled by the DNA substrates used in this study. Although not bent, DNA oligomers as short as 36 bp would correspond to one third of one DNA turn around the nucleosome. While actual nucleosomes may have provided a better protein binding substrate (or template) for exploring H1 binding to bent or curved DNA, the binding of $H1^0$, H1.1, and H1.4 to linear DNA (either short double stranded DNA's or CT-DNA) seems to provide a consistent picture with respect to binding affinity, the enthalpy and entropy changes for H1 binding, and the number of DNA bp covered by bound H1 in formation of these complexes. Our most puzzling result is that the enthalpy change results determined by Caterino et al are completely opposite in sign, *i.e.,* exothermic rather than endothermic as determined here [81]. We can only suggest that the gel mobility shift assays done as a function of temperature by Caterino were flawed in some way. Perhaps changes in gel porosity with temperature lead to their conclusion that the binding of Histone H1 to nucleosomal DNA is exothermic. The large differences in the reported $\Delta H°$ values for formation of the H1 DNA complexes, ranging from -12 kcal/mol to +22 kcal/mol, would serve to demonstrate the need for studies like the one reported here[81].

In conclusion, this study clearly shows that H1.1 and H1.4 behave very similarly to $H1^0$ in forming high affinity complexes with highly polymerized calf-thymus DNA. In addition, the binding of H1.4 to short DNA oligomers confirms both the limiting binding site size and the thermodynamics for the H1 DNA interactions. The formation of the $H1^0$, H1.1, and H1.4 DNA complexes is driven primarily by large favorable entropy changes.

It would appear that these large positive entropy changes result primarily from the expulsion of bound water molecules from the protein/DNA binding interface.

CHAPTER V

TEMPERATURE AND OSMOTIC STRESS DEPENDENCE OF THE THERMODYNAMICS FOR BINDING HISTONE $H1^0$, ITS CARBOXYL DOMAIN ($H1^0$-C) OR GLOBULAR DOMAIN ($H1^0$-G) TO DS-DNA

5.1 Abstract

Linker histones (H1) are the basic proteins in higher eukaryotes that are responsible for the final condensation of chromatin. In contrast to the nucleosome core histone proteins, the role of H1 in compacting DNA is not clearly understood. In this study we used isothermal titration Calorimetry (ITC) to directly measure the binding constant, enthalpy change, and stoichiometry or binding site size at different temperatures (*i.e.*, 15, 20, 25, 30, and 35 °C) for binding of $H1^0$ or its C-terminal ($H1^0$-C) and globular ($H1^0$-G) domains to highly polymerized calf-thymus DNA.

The temperature studies allowed us to estimate the $\Delta C°_p$ values for the binding of $H1^0$ and its carboxyl ($H1^0$-C) and globular domains ($H1^0$-G) to CT-DNA. The enthalpy changes for binding $H1^0$, $H1^0$-C, or $H1^0$-G to CT-DNA became increasingly exothermic as the temperature was increased from 15 °C to 35 °C. The interaction between $H1^0$-G with CT-DNA was observed to be calorimetrically silent at 25˚C. At higher temperatures, $\Delta H°$ for the $H1^0$-G/CT-DNA interaction becomes exothermic (e.g. $\Delta H° \approx$ -8 kcal/(mol $H1^0$-G) at 40 °C). Osmotic stress experiments revealed that the binding

constant for $H1^0$ is only slightly dependent on the osmolyte concentration. However, by plotting ln *K* vs. (Osm) we were able to estimate that 35 water molecules are released upon formation of the $H1^0$/CT-DNA complex.

5.2 Introduction

In the previous chapter we used isothermal titration calorimetry (ITC) to determine the thermodynamics for binding of $H1^0$, $H1^0$-C, and $H1^0$- G to highly polymerized calf-thymus DNA at 25 ºC in solutions havening a nominal salt concentration approximately 0.1 M. In our ITC studies, we found that the intact protein ($H1^0$) and its C-terminal domain ($H1^0$-C) bind to CT-DNA with approximately the same high affinity ($K_a \approx 1\times10^7$). We also observed large unfavorable enthalpy changes for the formation of these H1/DNA complexes ($\Delta H° \approx +22$ kcal/(mol $H1^0$ or $H1^0$-C)). There was no significant ITC signal for the addition of $H1^0$-G to CT-DNA at 25 ºC indicating that the $H1^0$-G/DNA complex was either not-formed or formed with a very small change in enthalpy at this temperature ($\Delta H° \approx 0$ kcal/(mol $H1^0$-G)). On the other hand, CD measurements indicated significant binding between $H1^0$-G and CT-DNA. The free energy change for formation of the $H1^0$/DNA and $H1^0$-C/DNA complexes at 25 ºC is driven by a very favorable entropy change ($-T\Delta S° \approx -30$ kcal/mol), and the binding site sizes for $H1^0$ and $H1^0$-C were determined to be 36bp and 28bp, respectively[76]. Using the polyelectrolyte theory of Record *et al.* the electrostatic contribution to the free energy change for binding $H1^0$ or $H1^0$-C to CT-DNA, ΔG_{elec}, was estimated to range from 6% to 17% of the total $\Delta G°$. In addition, the release of bound counterions (e.g. K^+) upon formation of the $H1^0$ and $H1^0$-C/CT-DNA complexes was estimated to be only one potassium ion[76]. We speculated large favorable entropy term for the $H1^0$ and $H1^0$-C

DNA complexes was due largely to the expulsion of bound water molecules from the protein-DNA interaction interface.

In the present study, we performed ITC titration experiments over the temperature range of 15 ºC to 40 ºC. In contrast to ITC experiments done previously at 25 °C where ΔH° was found to be approximately zero for formation of the $H1^0$-G/CT-DNA complex, we noted that ΔH° for formation of the $H1^0$-G/CT-DNA complex at higher temperatures was exothermic with ΔH° ≈ -8 kcal/ (mol $H1^0$-G) at 40 ºC. Analysis of $H1^0$-G/CT-DNA ITC data (at 40 ºC), using our fractional sites binding model, suggests that the binding mechanism for the interaction of the $H1^0$-G with CT-DNA may involve the formation of two different complexes (Complex A and B).

The ITC experiments done at different temperatures allowed us to determine ΔCp values for the formation of the $H1^0$ and $H1^0$-C/CT-DNA complexes. The ΔCp values determined here were found to be large and negative ($\Delta C°_p \approx -430$ cal mol^{-1} K^{-1}). This result is consistent with the loss of structure in the protein or DNA and/or the loss of bound water molecules as these complexes are formed. In this study, we also performed ITC experiments with TEG added as a co-solute or osmolyte. These experiments done at osmolalities of from 0.2 to 1.2 molal allowed us to probe the role of water and water release in the formation of the $H1^0$/CT-DNA complex. The result of the osmotic stress experiments is that the overall change in hydration, (ΔN_w), for formation of the $H1^0$/CT-DNA complex is –35 ± 8 water molecules. In effect approximately 35 water molecules are released upon complex formation.

5.3 Materials and Methods

5.3.1 Proteins and DNA samples

The $H1^0$ intact protein and its C-terminal and Globular domains were expressed using a bacterial strain of *E.coli* (*Rosetta2 (De3) pLysS*) transformed with a pET-11d (Novagen) expression vector as described [21]. The methods for expression and purification have been described elsewhere[21,54]. The pure HPLC protein fractions were concentrated using a Savant SPD 111V speed vac system for 4 hrs at 35°C to remove the HPLC solvents. Typically the sample buffer was BPES which is 30 mM K_2HPO_4/ KH_2PO_4 (pH = 7.0), 1 mM EDTA, and 100 mM KCl. For the osmotic stress dependent studies the amount of added osmolyte, TEG, was 0.2 m, 0.4 m, 0.6 m, 0.8 m, 1.0 m and 1.2 m. Calf thymus DNA type I was purchased from Sigma (St. Louis, USA) and dissolved in 1mL of the sample buffer. Both protein and DNA stock solutions were exhaustively dialyzed against the sample buffer (24h) at 4˚C. DNA concentrations in base pairs (bp) were determined using measured absorbance at 260 nm and a molar extinction coefficient of ε_{260}=1.31 x 10^4 bp $M^{-1}cm^{-1}$ [55]. The concentrations of $H1^0$, $H1^0$-C and $H1^0$-G were calculated using extinction coefficients 27.8, 31.1, and 28.6 mL mg^{-1} cm^{-1}, respectively at 205 nm[21].

The approximate molecular weights for the H1 and H1 domain constructs were estimated from their sequences using the ExPASy ProtParam tool[56] (http://web.expasy.org/protparam): Mw ($H1^0$) ≈ 20.8 kDa, Mw ($H1^0$-C) ≈ 9.55 kDa, Mw ($H1^0$-G) ≈ 9.28 kDa. The approximate average molecular weight of the CT-DNA was 8.42×10^3 kDa (Sigma, St. Louis, USA).

5.3.2 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a Microcal VP-ITC (Northampton, MA, USA). Titrations were done at five osmolyte concentrations (0.2, 0.4, 0.6, 0.8, 1.0, and 1.2 m TEG) and at 25º C. All titrations were performed by overfilling the ITC cell with approximately 1.5 mL of a dilute CT-DNA solution (nominally 480 μM in bp). Approximately 250 μL of a dilute solution of $H1^0$ (nominally 150 μM) was titrated into the calorimeter cell. The injection volume in these titrations was nominally 10 μL and a typical titration involved the addition of 25 injections of titrant at 600 second intervals. The 1.5 mL added to the VP-ITC cell overfills the cell so that there are no air bubbles in the chamber. As a volume of titrant ($H1^0$) is added, an equivalent volume of solution is displaced from the cell, and titrate concentration is corrected for material loss at each point in the titration. Titrations were also done at various temperatures (15, 20, 25, 30, 35 °C) for $H1^0$, $H1^0$-C, and $H1^0$-G. All of our ITC experiments were performed in triplicate at 25 °C. The integrated heat/injection data were fit for to an appropriate thermodynamic model using CHASM data analysis software developed in our laboratory[57]. The non-linear regression fitting process yields best fit parameters for K (or ΔG°), ΔH°, ΔS°, and n.

5.3.3 Molecular Modeling Study

Molecular modeling and MD simulations were performed using Accelrys Discovery Studio v.3.1 (San Diego, CA, USA). The crystal structure for the globular domain of linker Histone H1 protein was adapted from the Protein Data Bank (PDB accession code 1HST)[11]. The globular H1 was typed with the CHARMm27 force field

using the Momany-Rone partial charge method[87]. The entire system was solvated using an Explicit Periodic Boundary condition to using an orthorhombic shell extending 10 Å away from the boundary. Counterions were added to a concentration of 0.15 M. The system was subjected to a minimization routine using the Smart Minimizer algorithm and involving as many as 8,000 steps using a RMS gradient of less than 0.1 and a Spherical cutoff electrostatics model.

DNA-protein interactions were modeled according to the proposed binding sites based on a homology model as described by Ramakrishnan *et al.*[11]. A nucleosomal B-DNA fragment was extracted from the X-ray structure of nucleosome core particle (PDB accession code 1AOI) and was used as a substrate for linker histone protein binding. Based on the homology binding model described by Ramakrishnan *et al.*, there are two possible Histone H1 binding sites: a primary and a secondary binding site. The non-covalently DNA bound residues in the globular domain according to the proposed model were manually brought into contact with DNA backbones to facilitate (major) DNA groove binding. The Intermolecular Monitor feature was employed to assist with visualizing the intermolecular contacts between the protein residues and the bases in the major grooves. Specifically, for the primary binding site, residues Lys69, Lys85, and Arg73 are in close proximity with the DNA backbones, while in the hypothetical secondary binding site, residues Lys40, Arg42, Lys52, and Arg94 are brought to close proximity with the DNA backbones. The protein-DNA complex was again subjected to minimization same as described above. Molecular dynamics (MD) simulations was initiated with heating dynamics to raise the temperature of the system from 50 to 300 K, followed by an equilibration period and lastly a production step both executed at 300 K.

The total equilibration time was simulated for 20 ps at 1fs step followed by an additional of 100 ps of production step. NPT production type and the Leapfrog Verlet algorithm were used to integrate the dynamics. A representive structure of the last 5 ps in the production step was captured and Interaction Energy was calculated for this DNA-protein complex.

5.4 Results

The heat capacity changes (ΔC_p) associated with $H1^0$ or $H1^0$-C CT-DNA binding interactions can be determined directly from the temperature dependence of binding enthalpy using $\Delta C°_p=\delta(\Delta H°)/\delta T$[88]. We performed a temperature dependent ITC experiments in which $H1^0$ and $H1^0$-C were titrated into CT-DNA at temperatures ranging from 15 ºC to 35 ºC. The ITC thermograms at various temperatures were fit using nonlinear regression techniques to an independent site model (one site model) and the average best-fit parameters are listed in Table 5.1.

In Figure 5.1 shows a plot for ΔH values for binding $H1^0$ and $H1^0$-C to CT-DNA at the different measurement temperatures. The ΔH values for binding both $H1^0$ and $H1^0$-C to CT-DNA exhibit a linear decrease in the endothermic enthalpy change with increasing temperature. As shown in Figure 5.1, the temperature dependence of the enthalpy change, $\Delta H°$, for binding either $H1^0$ or $H1^0$-C to CT-DNA is similar for both proteins. The slope of the least squares line in Figure 5.1 corresponds to an estimated value for ΔC_p of -430 cal mol^{-1} K^{-1}.

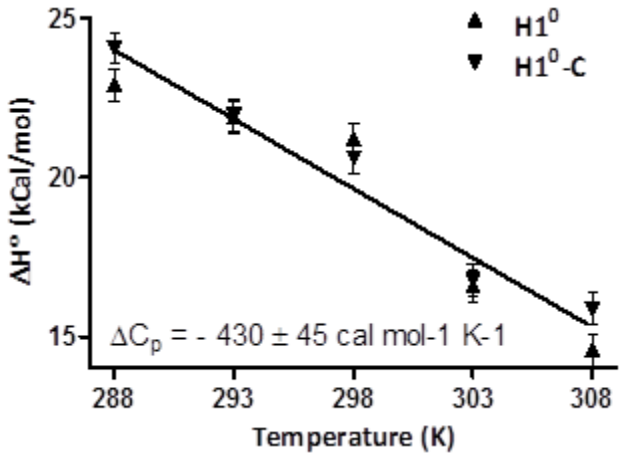

Figure 5.1 A plot of the ITC derived ΔH° values for the formation of the $H1^0$ and $H1^0$-C/CT-DNA complexes vs. temperature.

Data are shown for three temperatures: 15, 20, 25, 30 and 35 ºC. The slopes of the two lines yield estimates for the $\Delta C°_p$ values that accompany the formation of the $H1^0$ and $H1^0$-C CT-DNA complexes.

Table 5.1 ITC derived thermodynamic parameters for $H1^0$ and $H1^0$-C binding to CT-DNA at 15, 20, 25, 30 and 35 ºC in 100 mM [K^+] BPES pH 7.0

	Temp (K)	K_a (M^{-1}) ×10^{-7}	ΔG° (kcal/mol)	ΔH ° (kcal/mol)	-TΔS° (kcal/mol)
$H1^0$	288	1.6 ± 0.2	-9.5	22.9 ± 0.2	-32.4
	293	1.6 ± 0.1	-9.5	21.9 ± 0.3	-31.5
	298	0.7 ± 0.1	-9.4	21.2 ± 0.1	-31.0
	303	0.6 ± 0.1	-9.1	16.6 ± 0.3	-26.4
	308	0.5 ± 0.1	-9.2	14.6 ± 0.5	-23.5
$H1^0$-C	288	1.7 ± 0.1	-9.6	24.1 ± 0.4	-33.6
	293	1.7 ± 0.2	-9.6	21.9 ± 0.2	-31.6
	298	0.6 ± 0.1	-9.1	20.6 ± 0.2	-29.8
	303	1.6 ± 0.2	-9.5	16.8 ± 0.3	-26.3
	308	1.6 ± 0.1	-9.5	15.9 ± 0.3	-25.1

Figure 5.2 shows both the raw ITC signal (upper panel) and the apparent heat data for the titration of the $H1^0$-G into CT-DNA at 40 ºC. The integrated heat data were fit using a "fractional-sites" binding model where the total number of protein binding sites

was set to one (i.e. saturation stoichiometry of 1 mol or protein per 1 mole of binding site). The size of a protein binding site was determined to be 7 DNA base pairs from the ITC endpoint and the concentration of DNA in bp. The thermogram is consistent with the formation of two different $H1^0$-G/DNA complexes. Taken literally, the thermogram is consistent with the formation of 0.34 mol of a higher affinity complex exhibiting a smaller exothermic $\Delta H°$ value (-3 kcal mol^{-1}) and the formation of 0.63 mol of a weaker affinity complex which exhibits a larger exothermic $\Delta H°$ (-8 kcal mol^{-1}) for every mol of protein that is added or bound. Based on our fundamental understanding of chemical equilibria and free energy, this model cannot be correct as the high affinity complex must form to a greater extent than the low affinity complex. The nonlinear regression fit of the heat data to a simple fractional sites model consistent with the formation of two different complexes (Complexes A and B) is shown as the solid line in Figure 5.2. This model yields values for $\Delta H°_A$, K_B, $\Delta G°_B$, ΔH_B, and $-T\Delta S°_B$ that are well determined. However, the values for K_A, $\Delta G°_A$, and $-T\Delta S°_A$ are at best estimates since K_A must be approximately equal to K_B and therefore $\Delta G°_A \approx \Delta G°_B$, and calculated values for $-T\Delta S°_A$ would be fraught with error. Even though Table 5.2 lists the best fit thermodynamic parameters for the formation of both $H1^0$-G/CT-DNA complexes at 25 ºC, 30 ºC, 35 ºC, and 40 ºC, only the thermodynamic parameters for the formation of the weaker $H1^0$-G/CT-DNA complex, Complex B, should be considered reliable. The values shown in red in Table 5.2 were determined from the extrapolation of the best fit thermodynamic parameters obtained at 30 ºC, 35 ºC, and 40 ºC back to 25 °C.

In Figure 5.3 we have plotted the values of ΔH, ΔG, and $-T\Delta S$ for the binding of $H1^0$-G to CT-DNA as a function of temperature. For the reasons discussed above, Figure

5.3 Panel A shows only $\Delta H°_A$ for binding H1 0-G to CT-DNA to form Complex A. The value of $\Delta H°_A$ becomes more exothermic with increasing temperature. The value of $\Delta C°_p$ for formation of Complex A is estimated from the slope of the enthalpy change data in Figure 5.3 to be -260 cal mol^{-1} K^{-1}In Figure 5.3 Panel B, the thermodynamic parameters for binding H1 0-G to CT-DNA, forming Complex B. The change in $\Delta G°_B$ with temperature is attenuated by enthalpy/entropy compensation with $\Delta H°_B$ becoming more favorable with increasing temperature while the $-T\Delta S°_B$ term is becoming less favorable with increasing temperature. The enthalpy/entropy compensation is not perfect and $\Delta G°_B$ becomes somewhat less favorable as the temperature is increased. Again, the formation of Complex B exhibits an increasingly exothermic enthalpy change with increasing temperature and an estimated value for $\Delta C°_p$ of -590 cal mol^{-1} K^{-1}. The large negative heat capacity changes associated with the formation of both Complex A and B are the result of the loss of water from the complexes and bending of the DNA. The DNA bending angles are estimated to be approximately 118º in Complex A and 81º in Complex B. The larger DNA deformation (approximately 100º) in complex B may account for the larger change in heat capacity.

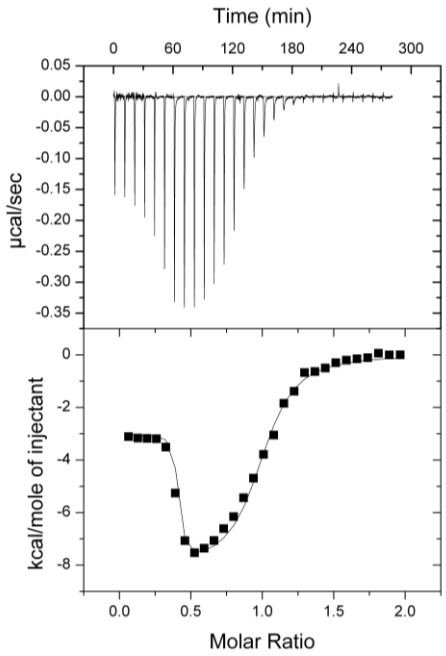

Figure 5.2 A typical ITC titration for the addition of $H1^0$-G to highly polymerized CT-DNA at 40 °C.

The upper half of Panel shows the baseline-corrected raw ITC signal for 25 injections of a dilute $H1^0$-G protein solution (10 µL of 70 µM $H1^0$-G) into the ITC cell filled with a dilute solution of CT-DNA (416 µM bp, 10 µM in $H1^0$-G binding sites). The lower half of shows the apparent ΔH° for each injection (-■-) along with the best-fit non-linear regression line (—) for a simple fractional sites binding model.

Table 5.2 ITC derived thermodynamic parameters for $H1^0$-G to CT-DNA at 30, 35, and 40 ºC in 100 mM [K^+] BPES pH 7.0.

Temp (K)	K_A (M^{-1}) ×10^{-9}	ΔG°$_A$ (kcal/mol)	ΔH°$_A$ (kcal/mol)	-TΔS°$_A$ (kcal/mol)	K_B(M^{-1}) ×10^{-6}	ΔG°$_B$ (kcal/mol)	ΔH°$_B$ (kcal/mol)	-TΔS°$_B$ (kcal/mol)
313	2.9±1.6	-12.9	-2.9±0.3	-9.9	2.4±1.6	-8.7	-8.1±0.5	-0.6
308	1.4±1.0	-12.5	-1.1±0.1	-11.3	4.0±1.5	-8.9	-4.6±0.2	-4.3
303	0.9±0.6	-12.2	-0.3±0.1	-12.6	1.8±0.6	-9.7	-2.2±0.3	-7.5
298	-	-11.8	1.2	-13.0	-	-10.1	0.9	-11.0

Thermodynamic values for the formation of Complex A and B are designated as ΔH°$_A$, K_A, ΔG°$_A$, –TΔS°$_A$ and ΔH°$_B$, K_B, ΔG°$_B$, –TΔS°$_B$. Thermodynamic values shown in red are ITC derived extrapolated values back to 25 ºC.

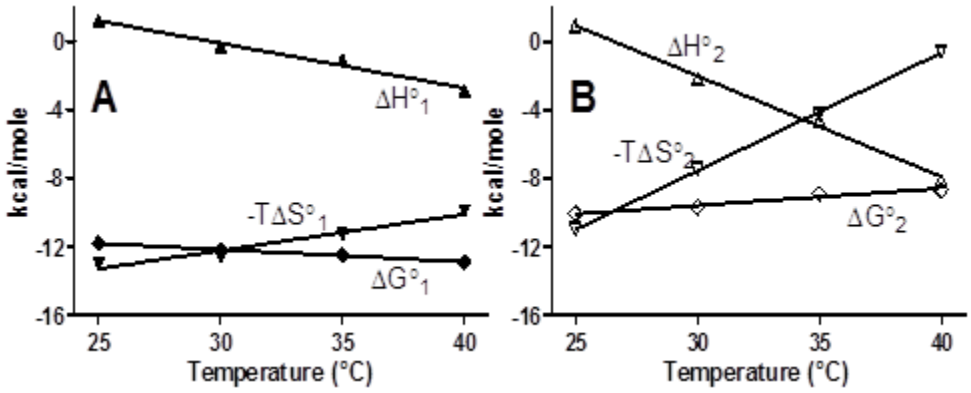

Figure 5.3 A plot of the thermodynamic parameters, ΔG°, ΔH°, and –TΔS° for the formation of Complex A (A) and Complex B (B) for the binding of $H1^0$-G to CT-DNA as a function of temperature.

The effect of increasing concentrations of TEG on equilibrium constant at 0.10 M $[K^+]$ was measured by ITC at 25 ºC. All the thermodynamic parameters for the titrations performed at different concentration of osmolyte ranging from 0.2 to 1.2 m are listed in Table 5.4. In Figure 5.4 we have plotted the values of ΔH°, ΔG°, and -TΔS° as a function of buffer osmolality. This plot demonstrates that the binding free energies are weakly dependent on the TEG concentrations. The binding of $H1^0$ to CT-DNA with the increasing concentration of osmolyte (TEG) is noticed as enthalpically less unfavorable, but this decrease in enthalpy is almost perfectly compensated by favorable increase of the binding entropy. In Figure 5.5 we have also plotted the natural logarithm of the K_a values as a function of TEG concentrations to determine the net hydration change. Osmolyte dependence of the equilibrium constant coupled with hydration changes has been analyzed by[89]

$$\frac{d(\ln K)}{d\,(Osm)} = \frac{-\Delta N_w}{55.6} \qquad (5.1)$$

where K_a is the equilibrium constant, Osm is the osmolality (moles of TEG/kg of solvent) of the buffer, and ΔN_w is change in the number of water molecules for the association of $H1^0$ with CT-DNA. A linear-least-square fit of the data points in figure 5.4 using Eq. 5.1 gives ΔN_w value as -35 ± 8. This means approximately 35 water molecules are released from complexation of $H1^0$ and CT-DNA.

Table 5.3 ITC derived thermodynamic parameters for $H1^0$ binding to CT-DNA in 100 mM [K^+] BPES pH 7.0 buffer solutions containing varying amount of TEG.

Osmolyte (TEG) m	**K_a (M^{-1}) ×10^{-6}**	**ΔG° (kcal/mol)**	**ΔH° (kcal/mol)**	**-TΔS° (kcal/mol)**
0.0	7.2 ± 0.1	-9.4	-21.8± 0.3	7.2
0.2	6.6 ± 0.1	-9.3	-18.4± 0.1	6.6
0.4	9.4 ± 0.1	-9.5	-16.2± 0.2	9.4
0.6	11.0 ± 0.3	-9.6	-16.2± 0.4	11
0.8	5.7 ± 0.2	-9.2	-12.6± 0.3	5.7
1.0	10.0 ± 0.4	-9.6	-11.6± 0.2	10
1.2	17.0 ± 0.2	-9.9	-8.7± 0.1	17

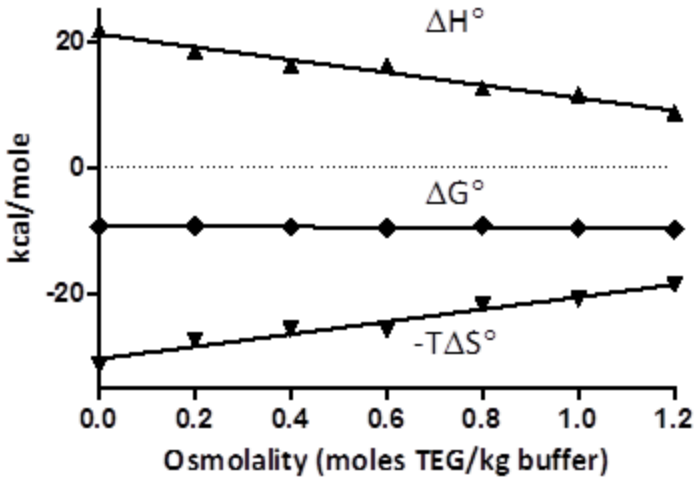

Figure 5.4 A plot of the thermodynamic parameters, ΔG°, ΔH°, and -TΔS° for the binding of $H1^0$ to CT-DNA as a function of osmolyte (TEG) concentration.

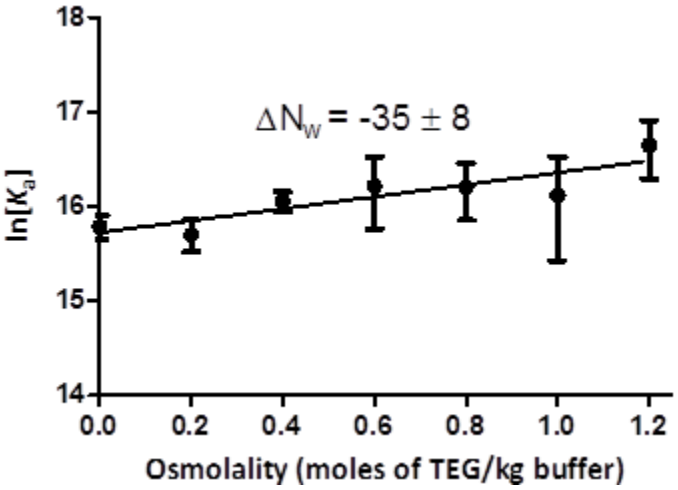

Figure 5.5 A plot of ln[K_a] vs osmolyte concentration (moles of TEG/kg buffer) for the binding of $H1^0$ to CT-DNA.

The data for $H1^0$ are shown as -●-.

5.4.1 Modeling Study

Results from our molecular modelling study are presented in Figures 5.6 and 5.7. A hyothetical model for the higher affinity, lower enthalpy, $H1^0$-G/CT-DNA complex, Complex A, is shown in Figure 5.6, while a hypothetical model for the lower affinity, higher enthalpy, $H1^0$-G/CT-DNA complex, Complex B, is shown in Figure 5.7. In both proposed models for Complex A and Complex B, duplex DNA remained in the bent conformation after extensive minimization. Formation of Complex A results form the slight bending of ds-DNA in order to contact with the three amino acid residues (Lys69, Arg73, and Lys85) of $H1^0$-G. The angle defined by entring and leaving vectors on the DNA is changed from a linear DNA (180°) to a slightly bent DNA of 118°. In constrast, the formation of Complex B results from the significant bending of the DNA in order to contact four amino acid residues (Lys40, Arg42, Lys52, and Arg94) of H10-G. The angle defined by entring and leaving vectors on the DNA is changed from a linear DNA

(180°) to a slightly bent DNA of 81°. As it was suggested in the previous studies by Ramakrishnan *et al.* we also observed binding of $H1^0$ helix III to the major groove of the DNA in Complex A model.

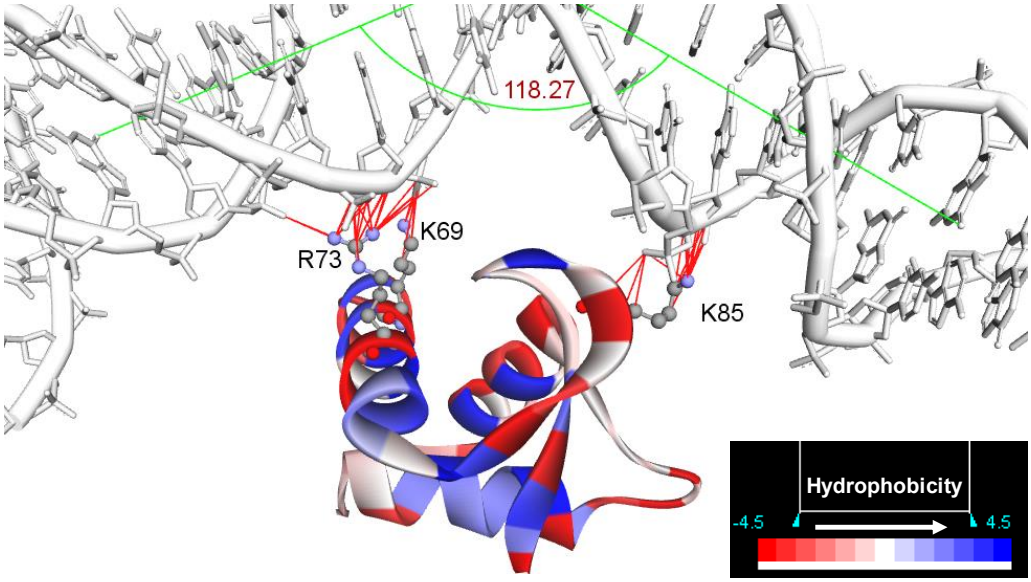

Figure 5.6 A hypothetical model for the formation of H^0-G/CT-DNA Complex A.

The protein is displayed using ribbon representation and colored according to the hydrophobicity of the residues. Intermolecular contacts are displayed in red.

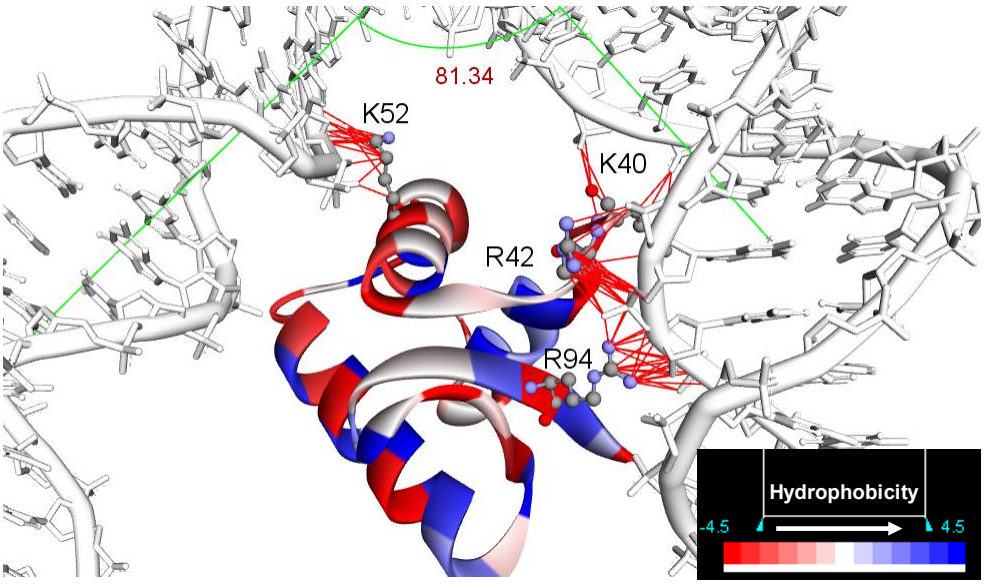

Figure 5.7 A hypothetical model for the formation of H^0-G/CT-DNA Complex B. The protein is displayed using ribbon representation and colored according to the hydrophobicity of the residues. Intermolecular contacts are displayed in red.

5.5 Discussion

Experiments performed previously in our lab using ITC revealed that $H1^0$ and $H1^0$-C bind tightly to CT-DNA ($K_a \approx 1\times10^7$). In both cases the enthalpy change is highly endothermic ($\Delta H° \approx +22$ kcal/mol). Obviously a large positive entropy change is responsible for the tight binding between $H1^0$ (including the C-terminus of H1.0, $H1^0$-C) and CT-DNA. We speculate that the energy for formation of the H1/CT-DNA complexes is largely due to the liberation of water molecules and ions from the highly charged C-terminus of the H1 protein and the DNA backbone in the vicinity of the binding site. In the subsequent study, dependence of K_a on ionic strength using ITC revealed that ligand charge/DNA charge interaction is at least partially responsible for the protein's affinity for DNA, and the counterions released from the $H1^0$/CT-DNA complex formation is very small (< 1). These results allowed us to speculate that the releasing of water molecules

from the binding surfaces of protein and DNA is the major driving force for the formation of the H1/DNA complex. In the current study, we used ITC to investigate the role of accompanying dehydration effects upon non-specific binding of $H1^0$ to CT-DNA.

Based on the linear relationship between the $\Delta C°_p$ and the changes in the solvent exposure of the hydrophobic and hydrophilic groups[90,91], we planned to repeat the ITC binding experiments for $H1^0$ (or H^0-C) with CT-DNA at various temperatures and under varying solvent conditions. The dependence of binding enthalpy on the temperature provides the heat capacity effect on the binding[92].

In Figure 5.1, both $H1^0$ and $H1^0$-C exhibit the strong and similar temperature dependence pattern of $\Delta H°$ which indicates the major contribution to $\Delta C°_p$ of $H1^0$ is coming from the C-terminal tail of the protein. Assuming a temperature-independent heat capacity change, linear square fit to the temperature dependent $\Delta H°$ values yielded large negative heat capacity change values. The observation of large negative heat capacity change values for the binding of H^0 to CT-DNA was unexpected because close to zero heat capacity changes have been observed for several non-sequence-specific DNA binding proteins[93,94]. It is well known that sequence specific DNA binding proteins interactions involving tight and solvent excluded interfaces are often associated with large negative heat capacity changes for the association[95,96]. Since no salt dependence of the relative equilibrium constant for $H1^0$/CT-DNA and $H1^0$-C/CT-DNA was observed, we believed dehydration of interacting surfaces is the major driving force for H1/DNA complexation.

In our earlier studies we have determined the thermodynamic parameters and the binding site sizes for $H1^0$ and $H1^0$-C binding to CT-DNA at 25 °C. We also attempted to

measure the interaction between $H1^0$-G and CT-DNA, but the interaction is calorimetrically silent at 25 °C. Single complex formation has been observed for either binding of $H1^0$ or $H1^0$-C to CT-DNA. However, binding of $H1^0$-G to CT-DNA at higher temperatures (30, 35, and 40 °C) clearly shows the formation of two complexes with characteristic thermodynamic signatures (Complex A and B). ITC experiments carried out at higher temperatures (30, 35, and 40 ºC) result exothermic heat changes for the binding of $H1^0$-G to CT-DNA. The two-fractional-sites model fits for the $H1^0$-G/CT-DNA ITC data at higher temperatures yield thermodynamic signatures which are consistent with the formation of two different complexes. Though the simple fractional sites model fits the ITC data very well, observation of less populated ($n_A = 0.34$) higher affinity complexes (Complex A) led us consider the thermodynamic parameters K_A and $-T\Delta S°_A$ are incorrect. The entropy change ($-T\Delta S°_B$) for the formation of Complex B is becoming more favorable with increasing temperature. The slopes from the temperature dependent study, *i.e.,* plot of $\Delta H°$ vs. temperature, are indicative of large negative heat capacity changes for the formation complex A and B. In the previous studies by Ramakrishna *et.al,* it has been proposed that H1 globular domain has two binding sites for the DNA [28]. Positively charged amino acid cluster in the primary binding site are located in the helix III of the H1 globular domain, which is similar to that of CAP (catabolite activator protein) recognition helix[27,28]. The secondary binding site amino acid cluster of H1 globular domain is located on the helix I and II. Based on these findings we proposed two hypothetical models in which binding of $H1^0$-G to DNA forms two different complexes (Complex A and B). In Complex B DNA is observed to be more bent than the DNA in Complex A. Formation of Complex B associated with

significant bending of DNA is consistent with the larger negative heat capacity value observed for Complex B. Bending of DNA by $H1^0$-G observed from the modeling study contradicts the CD result of H^0-G binding to DNA. Our previous CD results indicated that the binding of $H1^0$-G to CT-DNA causes no detectable structural changes in the CT-DNA. Ligation assays done by Maria *et al.* revealed that binding of H1 globular domain causes unwinding of superhelical DNA with the unwinding angle of 8°[97]. We speculate our CD experiments of $H1^0$-G binding to CT-DNA are not sensitive enough to detect the DNA bending.

The role of dehydration on the binding equilibrium was probed by osmotic stress method. In this method replacing the part of water with neutral osmolyte (TEG), thereby lowering the water activity of the binding buffer[89]. The dependence of equilibrium binding constant on water activity allowed us estimating the net volume of released surface water upon complexation of $H1^0$/CT-DNA. We observed only a minor effect on the binding equilibrium constant in the presence of increasing concentration of TEG. The free energies of $H1^0$-G binding to CT-DNA are almost independent of TEG concentrations reflecting compensation of enthalpic and entropic terms. Osmotic stress studies yield an estimate of the hydration changes (ΔN_w) occurring upon formation of the $H1^0$/CT-DNA complex. The difference in hydration between the free $H1^0$and free CT-DNA and the $H1^0$/CT-DNA complex is –35 ± 8, in effect approximately 35 water molecules are released upon complex formation.

In conclusion, the negative heat capacity changes for the formation of $H1^0$/CT-DNA, $H1^0$-C/CT-DNA, and $H1^0$-G/CT-DNA complexes support the desolvation of the binding surface functional groups on protein and DNA and significant deformation of the

DNA. The loss of water molecules from the binding interface is further supported by the osmotic stress study.

CHAPTER VI

CONCLUSIONS

Many studies have been conducted on histone H1 variants to understand the H1 interactions with DNA, however previously it has been difficult to collect good calorimetric data for these protein DNA interactions. 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. This study clearly shows that H1.1, H1.4, and $H1^0$ form high affinity complexes with variable length ds-DNA such as short DNA oligomers and with highly polymerized calf-thymus DNA. 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 large positive entropy values are attributed to the release of water from the binding interface and structural changes associated with the DNA upon H1/DNA complexation.

Though H1.4 and $H1^0$ exhibit differences in charge and charge distribution of amino acid sequence, similar binding site sizes have been determined for binding of H1.4 and $H1^0$ to ds-DNA. H1.4 Binding site size (36bp) determined from the titration experiments performed on poly-disperse, highly polymerized CT-DNA is in very good agreement with the experiments employing five different mono-disperse low molecular weight DNAs. The thermodynamic parameters observed for binding either the whole

$H1^0$ protein or the C-terminal domain, $H1^0$-C were very similar. Although the thermodynamic parameters observed for binding $H1^0$-G to CT-DNA are quite different than for either $H1^0$ or $H1^0$-C, the $\Delta H°$ for binding $H1^0$ is approximately equal to the sum of the $\Delta H°$ values for binding $H1^0$-C and $H1^0$-G. The binding site size for interaction of the complete protein (36 bp) is also approximately equal to the sum of the bps covered in binding $H1^0$-C (28 bp) and $H1^0$-G (7bp).

Binding of $H1^0$ and/or $H1^0$-C to CT-DNA is associated with significant conformational changes in the DNA. In the case of the globular domain, binding of $H1^0$-G to CT-DNA results little unfolding of globular and no structural changes in the DNA. The binding of the globular domain ($H1^0$-G) to CT-DNA is tight with little unfolding of the globular domain, yet the enthalpy change for this interaction is near zero ($\Delta H° = 0$) at room temperature. But at higher temperatures, exothermic enthalpy change for the binding of $H1^0$-G to CT-DNA was observed. A very small electrostatic contribution to the overall free energy change for the formation of $H1^0$/CT-DNA complex is observed. Formation of H1/DNA complexes associated with large negative ΔCp values is consistent with the loss of structure in the DNA and the loss of approximately 35 bound water molecules.

A better understanding of the functional properties of H1 and its interactions with DNA could provide new insights in understanding the role H1 in DNA condensation and transcriptional regulation. Since linker histone is highly altered in cancer cells, if the binding site size is known, drug design may be possible to prevent the linker histone modifications seen in some cancerous cells. A drug specific to the binding site size of

these proteins may be useful in regulation of transcription by H1.1 and H1.4, $H1^0$ histone proteins.

REFERENCES

1. Happel N., Doenecke D. "Histone H1 and its isoforms: contribution to chromatin structure and function." *Gene*. **2009**; 431(1-2): 1-12.

2. Khochbin S., Wolffe A. P. "Developmentally regulated expression of linker-histone variants in vertebrates." *Eur J Biochem*. **1994**; 225(2): 501-10.

3. Woodcock C. L. "Chromatin architecture." *Curr Opin Struct Biol*. **2006**; 16(2): 213-20.

4. Hansen J. C., Lu X., Ross E. D., Woody R. W. "Intrinsic protein disorder, amino acid composition, and histone terminal domains." *J Biol Chem*. **2006**; 281(4): 1853-6.

5. Lu X., Hansen J. C. "Identification of specific functional subdomains within the linker histone H10 C-terminal domain." *J Biol Chem*. **2004**; 279(10): 8701-7.

6. Ramaswamy A., Bahar I., Ioshikhes I. "Structural dynamics of nucleosome core particle: comparison with nucleosomes containing histone variants." *Proteins*. **2005**; 58(3): 683-96.

7. Suto R. K., Clarkson M. J., Tremethick D. J., Luger K. "Crystal structure of a nucleosome core particle containing the variant histone H2A.Z." *Nat Struct Biol*. **2000**; 7(12): 1121-4.

8. Zheng C., Hayes J. J. "Intra- and inter-nucleosomal protein-DNA interactions of the core histone tail domains in a model system." *J Biol Chem*. **2003**; 278(26): 24217-24.

9. Duggan M. M., Thomas J. O. "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): 21-33.

10. Goytisolo F. A., Gerchman S. E., Yu X., Rees C., Graziano V., Ramakrishnan V., et al. "Identification of two DNA-binding sites on the globular domain of histone H5." *EMBO J*. **1996**; 15(13): 3421-9.

11. Ramakrishnan V., Finch J. T., Graziano V., Lee P. L., Sweet R. M. "Crystal structure of globular domain of histone H5 and its implications for nucleosome binding." *Nature*. **1993**; 362(6417): 219-23.

12. Ali T., Coles P., Stevens T. J., Stott K., Thomas J. O. "Two homologous domains of similar structure but different stability in the yeast linker histone, Hho1p." *J Mol Biol*. **2004**; 338(1): 139-48.

13. Marino-Ramirez L., Kann M. G., Shoemaker B. A., Landsman D. "Histone structure and nucleosome stability." *Expert Rev Proteomics*. **2005**; 2(5): 719-29.

14. Walter L., Klinga-Levan K., Helou K., Albig W., Drabent B., Doenecke D., et al. "Chromosome mapping of rat histone genes H1fv, H1d, H1t, Th2a and Th2b." *Cytogenet Cell Genet*. **1996**; 75(2-3): 136-9.

15. Zlatanova J., Doenecke D. "Histone H1 zero: a major player in cell differentiation?" *FASEB J*. **1994**; 8(15): 1260-8.

16. Happel N., Schulze E., Doenecke D. "Characterisation of human histone H1x." *Biol Chem*. **2005**; 386(6): 541-51.

17. Orrego M., Ponte I., Roque A., Buschati N., Mora X., Suau P. "Differential affinity of mammalian histone H1 somatic subtypes for DNA and chromatin." *BMC Biol*. **2007**; 5: 22.

18. Allan J., Hartman P. G., Crane-Robinson C., Aviles F. X. "The structure of histone H1 and its location in chromatin." *Nature*. **1980**; 288(5792): 675-9.

19. Subirana J. A. "Analysis of the charge distribution in the C-terminal region of histone H1 as related to its interaction with DNA." *Biopolymers*. **1990**; 29(10-11): 1351-7.

20. Vila R., Ponte I., Jimenez M. A., Rico M., Suau P. "An inducible helix-Gly-Gly-helix motif in the N-terminal domain of histone H1e: a CD and NMR study." *Protein Sci*. **2002**; 11(2): 214-20.

21. Mamoon N. M., Song Y., Wellman S. E. "Histone H10 and Its Carboxyl-Terminal Domain Bind in the Major Groove of DNA." *Biochemistry*. **2002**; 41(29): 9222-8.

22. Vyas P., Brown D. T. "N- and C-terminal domains determine differential nucleosomal binding geometry and affinity of linker histone isotypes H1(0) and H1c." *J Biol Chem*. **2012**; 287(15): 11778-87.

23. Ponte I., Vila R., Suau P. "Sequence complexity of histone H1 subtypes." *Mol Biol Evol*. **2003**; 20(3): 371-80.

24. Misteli T., Gunjan A., Hock R., Bustin M., Brown D. T. "Dynamic binding of histone H1 to chromatin in living cells." *Nature*. **2000**; 408(6814): 877-81.

25. Roque A., Orrego M., Ponte I., Suau P. "The preferential binding of histone H1 to DNA scaffold-associated regions is determined by its C-terminal domain." *Nucleic Acids Res*. **2004**; 32(20): 6111-9.

26. Vignali M., Workman J. L. "Location and function of linker histones." *Nat Struct Biol*. **1998**; 5(12): 1025-8.

27. Zhou Y. B., Gerchman S. E., Ramakrishnan V., Travers A., Muyldermans S. "Position and orientation of the globular domain of linker histone H5 on the nucleosome." *Nature*. **1998**; 395(6700): 402-5.

28. Ramakrishnan V. "Histone structure and the organization of the nucleosome." *Annu Rev Biophys Biomol Struct*. **1997**; 26: 83-112.

29. Pham A. D., Sauer F. "Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in Drosophila." *Science*. **2000**; 289(5488): 2357-60.

30. Trojer P., Reinberg D. "Facultative heterochromatin: is there a distinctive molecular signature?" *Mol Cell*. **2007**; 28(1): 1-13.

31. Gurley L. R., Valdez J. G., Buchanan J. S. "Characterization of the mitotic specific phosphorylation site of histone H1. Absence of a consensus sequence for the p34cdc2/cyclin B kinase." *J Biol Chem*. **1995**; 270(46): 27653-60.

32. Sigurskjold B. W. "Exact analysis of competition ligand binding by displacement isothermal titration calorimetry." *Anal Biochem*. **2000**; 277(2): 260-6.

33. Buurma N. J., Haq I. "Advances in the analysis of isothermal titration calorimetry data for ligand-DNA interactions." *Methods*. **2007**; 42(2): 162-72.

34. Kensal E. van Holde J., W. C., P. Shing Ho. *Principles of Physical Biochemistry* Second Edition ed. Carlson G, editor. Upper Saddle River: Pearson Prentice Hall; **2006**. 710 p.

35. Spink C. H. "Differential scanning calorimetry." *Methods Cell Biol*. **2008**; 84: 115-41.

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

37. Clas S. D., Dalton C. R., Hancock B. C. "Differential scanning calorimetry: applications in drug development." *Pharm Sci Technolo Today*. **1999**; 2(8): 311-20.

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

39. Kanehara H., Mizuguchi M., Tajima K., Kanaori K., Makino K. "Spectroscopic evidence for the formation of four-stranded solution structure of oligodeoxycytidine phosphorothioate." *Biochemistry*. **1997**; 36(7): 1790-7.

40. Li W., Miyoshi D., Nakano S., Sugimoto N. "Structural competition involving G-quadruplex DNA and its complement." *Biochemistry*. **2003**; 42(40): 11736-44.

41. Kaushik M., Bansal A., Saxena S., Kukreti S. "Possibility of an antiparallel (tetramer) quadruplex exhibited by the double repeat of the human telomere." *Biochemistry*. **2007**; 46(24): 7119-31.

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

43. Carruthers L. M., Bednar J., Woodcock C. L., Hansen J. C. "Linker histones stabilize the intrinsic salt-dependent folding of nucleosomal arrays: mechanistic ramifications for higher-order chromatin folding." *Biochemistry*. **1998**; 37(42): 14776-87.

44. Shen X., Yu L., Weir J. W., Gorovsky M. A. "Linker histones are not essential and affect chromatin condensation in vivo." *Cell*. **1995**; 82(1): 47-56.

45. Thoma F., Koller T., Klug A. "Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin." *J Cell Biol*. **1979**; 83(2 Pt 1): 403-27.

46. Fan Y., Nikitina T., Zhao J., Fleury T. J., Bhattacharyya R., Bouhassira E. E., et al. "Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation." *Cell*. **2005**; 123(7): 1199-212.

47. Croston G. E., Kerrigan L. A., Lira L. M., Marshak D. R., Kadonaga J. T. "Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription." *Science*. **1991**; 251(4994): 643-9.

48. Zlatanova J., van Holde K. "Linker histones versus HMG1/2: a struggle for dominance?" *Bioessays*. **1998**; 20(7): 584-8.

49. Strahl B. D., Allis C. D. "The language of covalent histone modifications." *Nature*. **2000**; 403(6765): 41-5.

50. Parseghian M. H., Hamkalo B. A. "A compendium of the histone H1 family of somatic subtypes: an elusive cast of characters and their characteristics." *Biochem Cell Biol*. **2001**; 79(3): 289-304.

51. Izzo A., Kamieniarz K., Schneider R. "The histone H1 family: specific members, specific functions?" *Biol Chem*. **2008**; 389(4): 333-43.

52. Lu X., Hamkalo B., Parseghian M. H., Hansen J. C. "Chromatin condensing functions of the linker histone C-terminal domain are mediated by specific amino acid composition and intrinsic protein disorder." *Biochemistry*. **2009**; 48(1): 164-72.

53. George E. M., Izard T., Anderson S. D., Brown D. T. "Nucleosome interaction surface of linker histone H1c is distinct from that of H1(0)." *J Biol Chem*. **2010**; 285(27): 20891-6.

54. Wellman S. E., Song Y., Su D., Mamoon N. M. "Purification of mouse H1 histones expressed in Escherichia coli." *Biotechnol Appl Biochem*. **1997**; 26 (Pt 2): 117-23.

55. Wells R. D., Larson J. E., Grant R. C., Shortle B. E., Cantor C. R. "Physicochemical studies on polydeoxyribonucleotides containing defined repeating nucleotide sequences." *J Mol Biol*. **1970**; 54(3): 465-97.

56. Pace C. N., Vajdos F., Fee L., Grimsley G., Gray T. "How to measure and predict the molar absorption coefficient of a protein." *Protein Sci*. **1995**; 4(11): 2411-23.

57. Le V. H., Buscaglia R., Chaires J. B., Lewis E. A. "Modeling complex equilibria in isothermal titration calorimetry experiments: thermodynamic parameters estimation for a three-binding-site model." *Anal Biochem*. **2013**; 434(2): 233-41.

58. Velazquez-Campoy A. "Ligand binding to one-dimensional lattice-like macromolecules: analysis of the McGhee-von Hippel theory implemented in isothermal titration calorimetry." *Anal Biochem*. **2006**; 348(1): 94-104.

59. Record M. T., Jr., Anderson C. F., Lohman T. M. "Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity." *Q Rev Biophys*. **1978**; 11(2): 103-78.

60. Gray D. M., Morgan A. R., Ratliff R. L. "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): 3679-95.

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

62. Barbero J. L., Franco L., Montero F., Moran F. "Structural studies on histones H1. Circular dichroism and difference spectroscopy of the histones H1 and their trypsin-resistant cores from calf thymus and from the fruit fly Ceratitis capitata." *Biochemistry*. **1980**; 19(17): 4080-7.

63. Carter G. J., van Holde K. "Self-association of linker histone H5 and of its globular domain: evidence for specific self-contacts." *Biochemistry*. **1998**; 37(36): 12477-88.

64. Cerf C., Lippens G., Muyldermans S., Segers A., Ramakrishnan V., Wodak S. J., et al. "Homo- and heteronuclear two-dimensional NMR studies of the globular domain of histone H1: sequential assignment and secondary structure." *Biochemistry*. **1993**; 32(42): 11345-51.

65. Hashimoto H., Takami Y., Sonoda E., Iwasaki T., Iwano H., Tachibana M., et al. "Histone H1 null vertebrate cells exhibit altered nucleosome architecture." *Nucleic Acids Res*. **2010**; 38(11): 3533-45.

66. Read M., Harrison R. J., Romagnoli B., Tanious F. A., Gowan S. H., Reszka A. P., et al. "Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors." *Proc Natl Acad Sci U S A*. **2001**; 98(9): 4844-9.

67. Sparapani S., Haider S. M., Doria F., Gunaratnam M., Neidle S. "Rational design of acridine-based ligands with selectivity for human telomeric quadruplexes." *J Am Chem Soc*. **2010**; 132(35): 12263-72.

68. Allan J., Mitchell T., Harborne N., Bohm L., Crane-Robinson C. "Roles of H1 domains in determining higher order chromatin structure and H1 location." *J Mol Biol*. **1986**; 187(4): 591-601.

69. Dootz R., Toma A. C., Pfohl T. "Structural and dynamic properties of linker histone H1 binding to DNA." *Biomicrofluidics*. **2011**; 5(2): 24104.

70. Bednar J., Horowitz R. A., Grigoryev S. A., Carruthers L. M., Hansen J. C., Koster A. J., et al. "Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin." *Proceedings of the National Academy of Sciences*. **1998**; 95(24): 14173-8.

71. Zlatanova J., Seebart C., Tomschik M. "The linker-protein network: control of nucleosomal DNA accessibility." *Trends Biochem Sci*. **2008**; 33(6): 247-53.

72. De S., Brown D. T., Lu Z. H., Leno G. H., Wellman S. E., Sittman D. B. "Histone H1 variants differentially inhibit DNA replication through an affinity for chromatin mediated by their carboxyl-terminal domains." *Gene*. **2002**; 292(1-2): 173-81.

73. Espino P. S., Drobic B., Dunn K. L., Davie J. R. "Histone modifications as a platform for cancer therapy." *J Cell Biochem*. **2005**; 94(6): 1088-102.

74. Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W., et al. "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic Acids Res*. **1997**; 25(17): 3389-402.

75. Altschul S. F., Wootton J. C., Gertz E. M., Agarwala R., Morgulis A., Schaffer A. A., et al. "Protein database searches using compositionally adjusted substitution matrices." *Febs j*. **2005**; 272(20): 5101-9.

76. Machha V., Waddle J. R., Turner A. L., Wellman S., Le V. H., Lewis E. A. "Calorimetric Studies of the Interactions of Linker Histone H10 and its Carboxyl (H10-C) and Globular (H10-G) Domains with Calf-Thymus DNA." *Biophysical Chemistry*. **2013**; (0).

77. Andrews A. J., Luger K. "Nucleosome structure(s) and stability: variations on a theme." *Annu Rev Biophys*. 40: 99-117.

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

79. Pruss D., Wolffe A. P. "Histone-DNA contacts in a nucleosome core containing a Xenopus 5S rRNA gene." *Biochemistry*. **1993**; 32(27): 6810-4.

80. Luger K., Mader A. W., Richmond R. K., Sargent D. F., Richmond T. J. "Crystal structure of the nucleosome core particle at 2.8 A resolution." *Nature*. **1997**; 389(6648): 251-60.

81. Caterino T. L., Fang H., Hayes J. J. "Nucleosome linker DNA contacts and induces specific folding of the intrinsically disordered H1 carboxyl-terminal domain." *Mol Cell Biol*. **2011**; 31(11): 2341-8.

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

83. Schlitz H. K. "Protein purification: Principles and practice, by R Scopes. Pp 282. Springer-Verlag, Berlin, Heidelberg and New York. 1982. DM 79. ISBN 3–540–90726–2." *Biochemical Education*. **1984**; 12(3): 143-.

84. Johnson W. C., Jr. "Protein secondary structure and circular dichroism: a practical guide." *Proteins*. **1990**; 7(3): 205-14.

85. Modukuru N. K., Snow K. J., Perrin Jr B. S., Bhambhani A., Duff M., Kumar C. V. "Tuning the DNA binding modes of an anthracene derivative with salt." *Journal of Photochemistry and Photobiology A: Chemistry*. **2006**; 177(1): 43-54.

86. Tan W. B., Cheng W., Webber A., Bhambhani A., Duff M. R., Kumar C. V., et al. "Endonuclease-like activity of heme proteins." *J Biol Inorg Chem*. **2005**; 10(7): 790-9.

87. Brooks B. R., Bruccoleri R. E., Olafson B. D., States D. J., Swaminathan S., Karplus M. "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations." *Journal of Computational Chemistry*. **1983**; 4(2): 187-217.

88. Angell C. A., Tucker J. C. "Anomalous heat capacities of supercooled water and heavy water." *Science*. **1973**; 181(4097): 342-4.

89. Parsegian V. A., Rand R. P., Rau D. C. "Macromolecules and water: probing with osmotic stress." *Methods Enzymol*. **1995**; 259: 43-94.

90. Chaires J. B. "Energetics of drug-DNA interactions." *Biopolymers*. **1997**; 44(3): 201-15.

91. Prabhu N. V., Sharp K. A. "Heat capacity in proteins." *Annu Rev Phys Chem*. **2005**; 56: 521-48.

92. Liu C. C., Richard A. J., Datta K., LiCata V. J. "Prevalence of temperature-dependent heat capacity changes in protein-DNA interactions." *Biophys J*. **2008**; 94(8): 3258-65.

93. Ladbury J. E., Wright J. G., Sturtevant J. M., Sigler P. B. "A thermodynamic study of the trp repressor-operator interaction." *J Mol Biol*. **1994**; 238(5): 669-81.

94. Takeda Y., Sarai A., Rivera V. M. "Analysis of the sequence-specific interactions between Cro repressor and operator DNA by systematic base substitution experiments." *Proc Natl Acad Sci U S A*. **1989**; 86(2): 439-43.

95. Bailly C., Chessari G., Carrasco C., Joubert A., Mann J., Wilson W. D., et al. "Sequence-specific minor groove binding by bis-benzimidazoles: water molecules in ligand recognition." *Nucleic Acids Res*. **2003**; 31(5): 1514-24.

96. Jayaram B., Jain T. "The role of water in protein-DNA recognition." *Annu Rev Biophys Biomol Struct*. **2004**; 33: 343-61.

97. Ivanchenko M., Hassan A., van Holde K., Zlatanova J. "H1 binding unwinds DNA. Evidence from topological assays." *J Biol Chem*. **1996**; 271(51): 32580-5.