Investigating HMGB1 binding dynamics on Nucleosomes and DNA using single-molecule TIRF microscopy

by

Calvin K. Voong

B.S., California State University Long Beach, 2018

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosopy

Department of Biochemistry

2024

Committee Members:

James A. Goodrich, Chair

Jennifer F. Kugel, Co–Chair

Joseph Falke

Karolin Luger

Thomas Perkins

Voong, Calvin K. (Ph.D., Biochemistry)

Investigating HMGB1 binding dynamics on Nucleosomes and DNA using single-molecule TIRF microscopy

Thesis directed by Professor James A. Goodrich and Professor Jennifer F. Kugel

HMGB1 is an architectural DNA binding protein that plays an important role in modulating chromatin compaction, facilitating protein factor binding to chromatin, and numerous other nuclear processes. In this thesis, I explored HMGB1 movement and binding dynamics on DNA and nucleosomes to better understand how HMGB1 interacts with these nucleic acids and nucleoprotein complexes.

Within the nucleus, HMGB1 is capable of rapidly navigating through the complexities of the chromatin environment. However, it is unclear what interaction mechanisms facilitates HMGB1s high mobility within the nucleus. In Chapter 2, I used single molecule Total Internal Reflection Fluorescence Microscopy (TIRF-M) to explore the roles that each of the individual HMGB1 domains play in modulating its movement speed on a dense surface of DNA molecules. In addition, I studied how varying the density and length of the immobilized DNA influenced HMGB1 movement characteristics and examined the implications that electrostatic interactions have on driving HMGB1 movement. I discovered that each of HMGB1's domains work together to modulate its movement speed, that reducing the density or length of the immobilized DNA drastically reduced its presence on the DNA surface, and that electrostatic interactions facilitate HMGB1's movement. The culmination of these findings provides a framework for describing the mechanisms that allow for HMGB1 to move between DNA and ultimately contribute to our understanding of how this protein interacts with chromatin.

Nuclear HMGB1 can function as an architectural factor by operating as a DNA chaperone to facilitate the binding of nearby transcription factors to their cognate DNA binding sites. Currently, it is known that HMGB1 is capable of interacting with the nucleosome by binding to the linker

DNA at the entry/exit junctions and that its negatively charged C-terminal tail is capable of interacting with the histone H3 N-terminal tail. However, it is unclear how these interactions facilitate HMGB1's positioning onto the nucleosome and kinetics with which it interacts with nucleosomes. In Chapter 3, I used single molecule TIRF-M to study the on/off binding kinetics of HMGB1 on nucleosome particles. HMGB1 exhibited on/off binding behaviors on nucleosomes; however, the data were not reproducible across multiple replicates. In an attempt to overcome the irreproducibility issue, I simplified the system by examining HMGB1 on/off binding on free DNA, but HMGB1 on/off binding kinetics was still inconsistent between experimental replicates. It is unclear what causes the reproducibility issues, but these could result from heterogeneity in protein preparations over time.

Dedication

Working on this PhD has been quite a challenge, to say the least. However, it has been an incredibly rewarding experience. I am deeply grateful to have had the chance to interact and chat with the brilliant minds of the faculty members and colleagues that have helped me become the scientist I am today. Thank you for taking time out of your day to chat with me about science, food, and life - I truly appreciate it. Thank you to the members of my cohort for all the memories filled with warmth, belly laughters, and tears of joy. Thank you to my lab members for dealing with my chaotic shinegans, as well as chatting about life and random topics (even if the topic changed 25 times within the span of 10 minutes). Although I am sadden that my chapter here has come to an end, I look forward to seeing how this vibrant community continues to grow and change decades beyond my departure.

I would also like to thank my partner, Esther, for being one of my major pillars of support. Thank you being patient with me when I didn't have the bandwidth to make a decision on things (such as food, places to go, etc.) throughout this writing process. Thank you for listening to me when I was frustrated, burnt out, and on the brink of complete exhaustion. Thank you for showing me what it was like to appreciate the smaller things in life. Thank you for having faith in me. And most importantly, thank you for loving me for the person I am. I feel incredibly fortunate to have you in my life, and I look forward to creating many more memories together with you through this 4-dimensional space called life.

Lastly, I would like to thank my family for playing an instrumental role in shaping me into the individual I am today. Thank you A Ba and A Ma for all the sacrifices you made to keep a

roof over our heads and food on the table for us to eat growing up. Thank you for buying my elementary school uniform and my school supplies, even it meant having to spend less money on food for that week. Thank you for the small moments of connection in silence we had on the days you came home exhausted, lingering still the smell of the fabrics you worked with at the sweatshop around the corner. Thank you for being kind to me and loving me. Thank you to my siblings for raising me during times when A Ba and A Ma couldn't. Thank you Guan Mei Zi for teaching me how to reflect on my feelings during turbulent times. Thank you Hien Zi for instilling the sense of tenacity and discipline into me. Thank you Si Ge for showing me what commitment looks like. Thank you Long Jie for supporting and encouraging me to pursue my dreams. I am truly grateful to you all.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Interactions of HMGB proteins with the Genome and the Impact on Disease

Adapted from: Voong,C.K., Goodrich,J.A., and Kugel,J.F.(2021) Interactions of HMGB proteins with the genome and the impact on disease.*Biomolecules*.11:1451

High Mobility Group Box (HMGB) proteins are small architectural DNA binding proteins that regulate multiple genomic processes such as DNA damage repair, nucleosome sliding, telomere homeostasis, and transcription. In doing so they control both normal cellular functions and impact a myriad of disease states, including cancers and autoimmune diseases. HMGB proteins bind to DNA and nucleosomes to modulate the local chromatin environment, which facilitates the binding of regulatory protein factors to the genome and modulates higher order chromosomal organization. Numerous studies over the years have characterized the structure and function of interactions between HMGB proteins and DNA, both biochemically and inside cells, providing valuable mechanistic insight as well as evidence these interactions influence pathological processes. This review highlights recent studies supporting the roles of HMGB1 and HMGB2 in global organization of the genome, as well as roles in transcriptional regulation and telomere maintenance via interactions with G quadruplex structures. Moreover, emerging models for how HMGB proteins function as RNA binding proteins are presented. Nuclear HMGB proteins have broad regulatory potential to impact numerous aspects of cellular metabolism in normal and disease states.

1.1 Introduction

High mobility group (HMG) proteins play essential roles in normal cellular biology by functioning in several intracellular and extracellular capacities. These proteins also contribute to the molecular mechanisms behind the pathology of many disease states, making understanding their regulatory mechanisms particularly important. There are three superfamilies of HMG proteins: HMG AT-Hook (HMGA), HMG nucleosome binding (HMGN), and HMG box (HMGB) proteins. The first members of the HMG family were discovered in the 1970s and were initially characterized by their fast migration in an acidic polyacrylamide gel (hence their name) [1]. Within the nucleus, HMG proteins function as architectural DNA binding proteins that dynamically interact with chromatin, thereby modulating DNA-dependent processes. This includes impacting the accessibility of nucleosomal DNA to facilitate protein factor binding, as well as impacting the higher order organization of chromatin structure, thereby regulating DNA replication, DNA repair, transcription, and chromatin remodeling [2]. Each protein within the HMG super family contains one or more HMG domains that define its mechanism of interaction with DNA and/or the nucleosome core particle [3] (Figure 1.1). In addition, most HMG proteins contain regulatory domains, typically a C-terminal acidic tail, that can control interactions with the genome or other proteins. After briefly introducing the HMGA and HMGN families of proteins, this Chapter focuses on the HMGB proteins and recent work describing their roles in organizing higher order chromatin structure genome -wide and in binding specific DNA structures and cellular RNAs, ultimately contributing to the regulation of genomic processes in specific disease states.

1.1.1 The DNA and Nucleosome Binding Properties of HMGA and HMGN Proteins

The HMGA protein family consists of HMGA1, its splicing variants HMGA1a, HMGA1b, HMGA1c, and HMGA2, which are encoded by the HMGA1 and HMGA2 genes respectively [4]. The HMGA proteins interact with DNA and nucleosomes through their unstructured AT hook

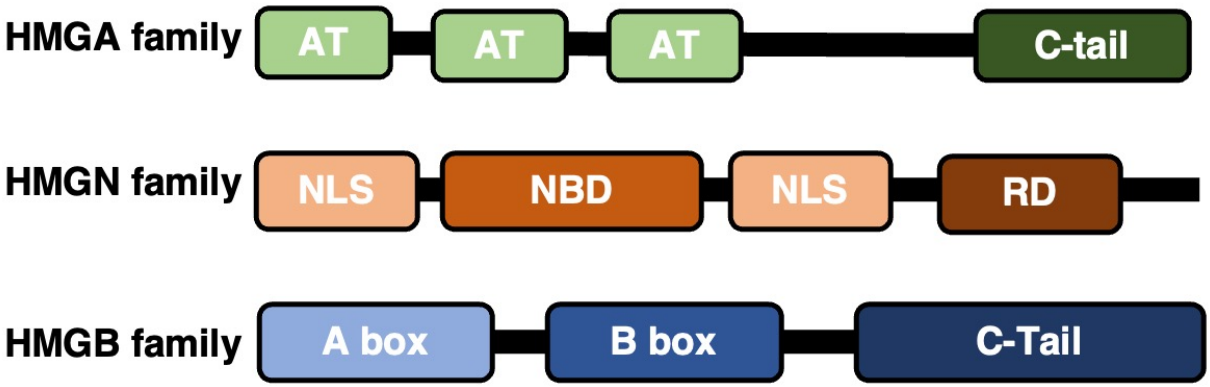

Figure 1.1: Schematic showing the domains in HMG family proteins. Abbreviations: AT, AT hook; C-tail, C-terminal acidic tail; NLS, nuclear localization sequence; NBD, nucleosome binding domain; RD, regulatory domain; A box, HMG box A; and B box, HMG box B.

domains (Figure 1.1), which become structured upon binding to DNA[5, 6]. HMGA proteins preferentially bind to the minor groove of DNA containing AT-rich stretches [7, 8]. They mediate transcriptional activation by promoting the decompaction of chromatin by binding the linker DNA between histones in competition with histone H1[9, 10]. HMGA proteins also coordinate with other members of the HMG family to modulate the compaction and higher order architecture of chromatin, thereby impacting a number of biological processes that occur on the genome[4, 6].

The HMGN protein family consists of five members: HMGN1, HMGN2, HMGN3, HMGN4, and HMGN5. These proteins contain a conserved nucleosomal binding domain, a nuclear localization signal, and a C-terminal regulatory domain (Figure 1.1). The HMGN proteins directly bind nucleosome core particles via a conserved eight amino acid sequence in the nucleosome binding domain that interacts with an acidic patch on the H2A/H2B dimer, in conjunction with DNA interactions that occur near the entry/exit point of the nucleosome core[11]. HMGN proteins do not displace H1, but instead bind simultaneously [12]. In addition, HMG proteins can promote the post translational modifications of the core histones by enhancing the activity of histone modifiers [13]. This helps de-compact chromatin structure and affects the expression of genes and other biological processes that occur on the genome.

1.1.2 The HMGB Proteins Bind Diverse DNA Structures and Interact with Histone Proteins

There are four members in the HMGB family: HMGB1, HMGB2, HMGB3, HMGB4. These proteins are characterized by their domains that bind to DNA in a structure-specific manner [14]. Within the nucleus HMGB proteins are highly abundant and ubiquitously expressed; there are approximately 10^6 molecules of HMGB1, averaging 1 molecule per 10-15 nucleosomes [14, 15]. HMGB proteins function as architectural DNA binding proteins by modulating the local environment of the chromatin to facilitate the binding of regulatory proteins that are important for genomic processes such as transcription and DNA repair. Mammalian HMGB proteins contain two HMG box domains and an acidic C-terminal tail (Figure 1.1), except for HMGB4, which lacks

the acidic C-terminal tail [16]. HMGB proteins can be post-translationally modified to influence their interactions with DNA, protein partners, and cellular localization. For example, HMGB1 can undergo acetylation, phosphorylation, methylation, ribosylation, and oxidation, creating a complex repertoire of potential regulatory mechanisms [17, 18].

The structure of each HMG box consists of three alpha helices folded into an L shaped conformation [19–21]. Each individual HMG box can bind to DNA in a manner that is not sequence specific, but is sensitive to the structure of the DNA [14]. The proteins show an increased affinity for non-B form distorted DNA structures (illustrated in Figure 1.2, such as cisplatinated DNA [22], cruciform DNA [23], single stranded DNA [24], supercoiled DNA [25], hemi catenated DNA [26], and DNA mini circles [27]. Moreover, atypical DNA structures can form in cells and play a role in regulating genomic processes through interactions with architectural and regulatory proteins [28, 29]. When the HMG boxes bind double stranded DNA, they induce the DNA to bend via specific amino acids that intercalate into the double helix and induce deformation [30]. The acidic C-terminal tails on HMGB proteins are unstructured and consist of streteches of glutamine and aspartate amino acid residues. The acidic C-terminal tail of HMGB1 makes extensive intramolecular interactions with its two DNA binding domains, which dampen its ability to bind and bend DNA [31–37], impacting the ability of HMGB1 to regulate transcriptional activity [38, 39].

In the context of chromatin, HMGB proteins can make specific interactions with both the DNA and histone proteins. HMGB proteins have been shown to bind to the linker DNA near the entry/exit junction of the nucleosome [40]. The C-terminal tail of HMGB1 can bind to histone H1, facilitating displacement of the linker histone, allowing HMGB1 to bind the linker DNA to induce nucleosome remodeling and organization of the chromatin [9, 41–43]. Furthermore, studies have shown that HMGB1 interacts with the histone H3 N-terminal tail, which could facilitate its binding to chromatin [38, 44]. It has been proposed that when HMGB proteins bind chromatin they locally modify the structure by bending the DNA and facilitating the binding of regulatory proteins such as transcription factors, chromatin remodelers, and DNA damage repair machinery [14, 45, 46]. For example, data suggest that HMGB1 facilitates binding of p53 [47] and estrogen

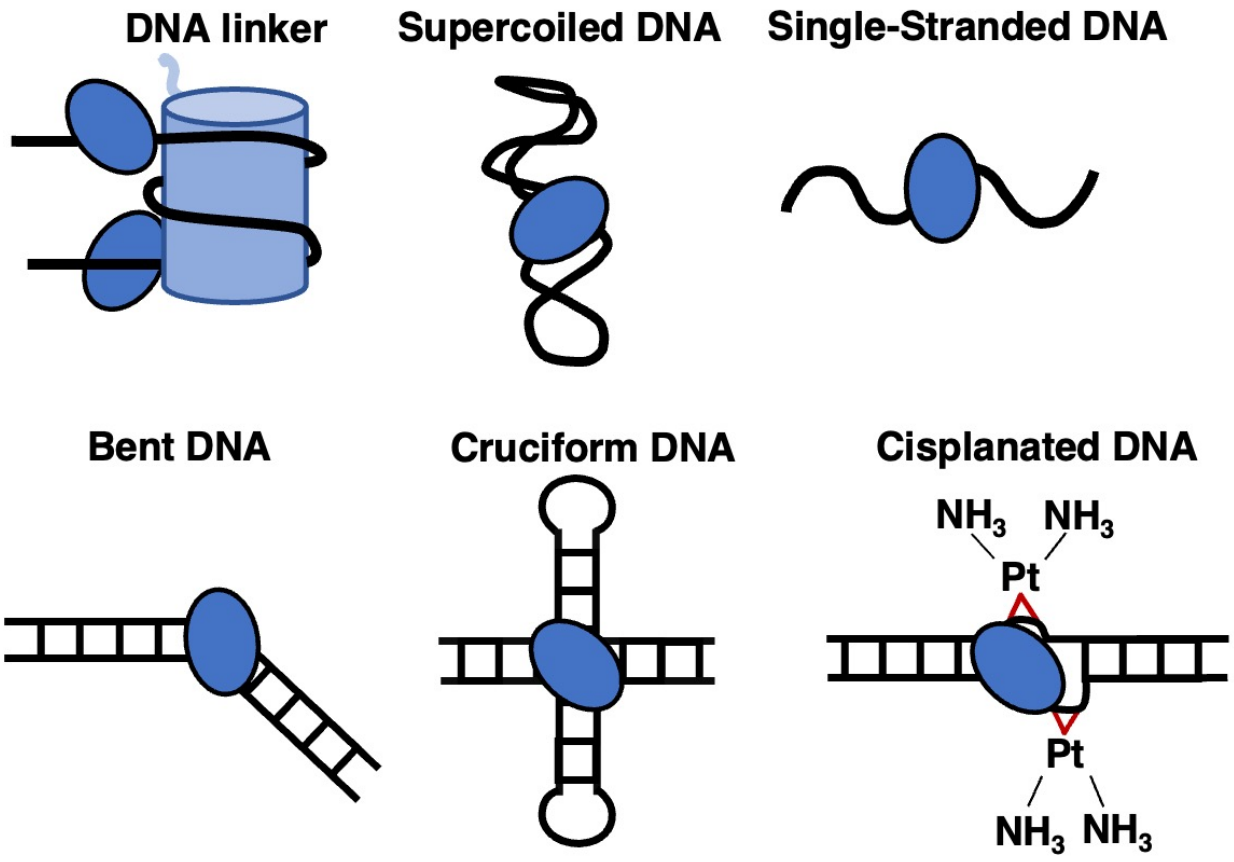

Figure 1.2: HMGB1/2 can bind to a myriad of DNA structures. The light blue cylinder represents a nucleosome, and the dark blue oval represent HMGB1/2

receptor [48], among other transcription factors[49].

As a function of their interaction with chromatin, HMGB proteins control several genomic processes in response to specific biological cues; therefore, they are essential regulators of cellular programs, for example senescence [50–54] as well as disease states, including cancer, autoimmune diseases, and inflammation [55]. Interestingly, HMGB1, in addition to its roles on the genome, has critical roles in both the cytoplasm and extracellularly. Within the cytoplasm, HMGB1 plays a role in the induction of autophagy, in mitochondrial quality control, and in mitochondrial DNA repair [56–58]. The secretion of HMGB1 into the extracellular matrix has been widely studied, where HMGB1 can act as a pro-inflammatory cytokine molecule by binding to cell surface receptors, such as the receptors for advanced glycation end products (RAGE) and toll like receptor 4 (TLR4), which stimulate the pro-inflammatory response [18].

The focus of this review is on the roles of HMGB proteins in the nucleus, emphasizing recent studies that demonstrate how HMGB interactions with the genome control a diversity of cellular programs. Also discussed is recent evidence characterizing HMGB protein interactions with G-quadruplex structures that form in promoters and at telomeres, as well as a growing body of data showing that different types of cellular RNA could be important regulatory partners for HMGB proteins. Although the focus of this review is mammalian HMGB proteins, it is worth noting that Saccharomyces cerevisiae Nhp6, a yeast counterpart to HMGB1/2, has been extensively investigated using structural, biochemical, and biophysical approaches, which provide important insight into how HMG boxes interact with DNA and chromatin [59–62].

1.2 HMGB Proteins Regulate Genome Organization

An important aspect of regulating gene expression is controlling the higher order three-dimensional organization of the genome, such as inter- and intra-chromosomal interactions. Architectural proteins play an important role in this process. For example, topologically associating domains (TADs) are large looped regions of chromatin in which distal areas of chromosomes are brought together with one another to create chromosomal territories within the nucleus that mod-

ulate gene transcription [63]. This three-dimensional architecture can be mapped genome wide using chromosome conformation capture techniques such as Hi-C. Hi-C can be used to identify TAD boundaries, which are often demarcated by the binding of the architectural protein CTCF (CCCTC-binding factor) and the protein complex cohesion [63]. In addition to higher-order genome organization, gene expression is controlled by regulating the compaction of chromatin [64]. Highly compacted heterochromatin silences gene expression and is marked by specific repressive histone modifications such as H3K9 methylation, while euchromatin is considered less compacted and contains more activating histone modifications such as acetylation and H3K4 methylation [65]. Only recently have genome-wide views of HMGB protein occupancy has been revealed, providing new insight into how these proteins control both higher order genome organization and chromatin structure.

1.2.1 Regulation of Chromatin Organization and Structure During Senescence

Cellular senescence is a stable arrest of the proliferative state that can be induced by a variety of internal and external stresses, such as prolonged DNA damage, shortening the telomere ends, and oncogene activation [66]. One signature of cellular senescence is the alteration of chromatin structure. HMGB2 is important during replicative senescence. Recent work has shown that HMGB2 is depleted from the nucleus upon entry into cellular senescence, resulting in reorganization of the genome and changes in transcriptional activity [51]. Specifically, using a combination of chromosome conformation capture (Hi-C assays) and ChIP-seq, it was found that that HMGB2 functions to modulate the global chromatin structure and expression of genes found within topologically associating domains (TADs) by insulating against the clustering of CTCF proteins (Figure 1.3). Upon senescence entry when HMGB2 was exported from the nucleus, loss of HMGB2 resulted in the clustering of CTCF molecules, as seen in fluorescence imaging, and the compaction of chromatin [51]. The reorganization of TADs was further validated by ChIP-seq data that showed HMGB2 bound to looped TAD regions prior to senescence, helping demarcate the boundaries of sub-TAD regions. HMGB2 was bound to genes involved in senescence processes, such as extracellular matrix

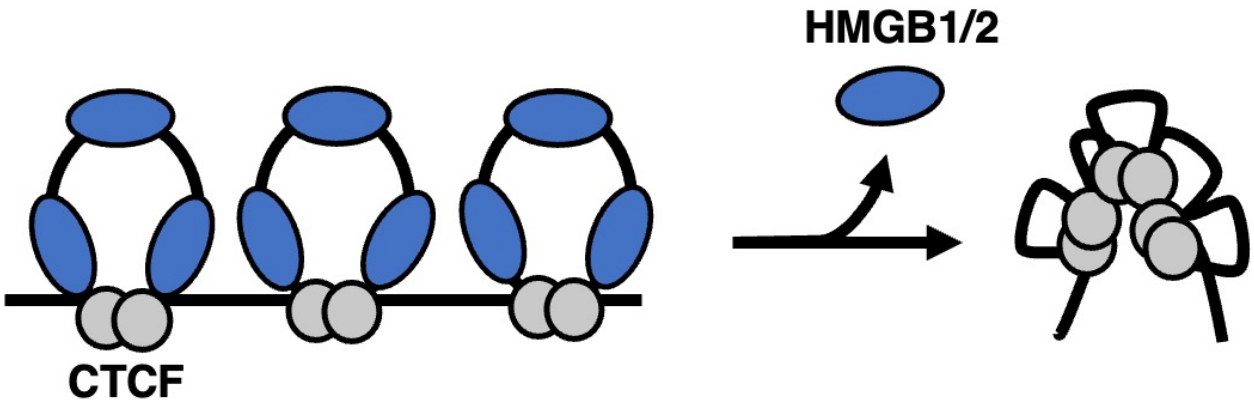

Figure 1.3: HMGB1/2 serves to insulate the interactions between CTCF proteins and demarcate the boundaries of TADS

reorganization and cellular aging [51]. Upon senescence and the loss of genome-bound HMGB2, the change in genome architecture led to changes in heterochromatin formation and ultimately gene expression. This included an increase in HP1α, a heterochromatic binding protein, and changes in core histone markers, such as a decrease in H3K27me3, consistent with a transition from facultative to constitutive heterochromatin upon senescence entry [51].

Recent work shows that HMGB1 also interacts with specific regions of the genome in proliferating cells and is depleted from the nucleus during entry into senescence [50]. Moreover, HMGB1 can be secreted into the extracellular matrix and induce senescence of nearby cells through paracrine signaling, thus playing a dual role in regulating senescence [50]. Within the nucleus, ChIP-seq of HMGB1 showed that the level of chromatin bound HMGB1 decreased upon entry into senescence [50]. Combining these data with Hi-C data revealed that in proliferative cells HMGB1 bound to TAD looped regions and marked a subset of senescence-associated TAD boundaries. Moreover, the HMGB1 clusters were bound to senescence-associated genes [50]. Upon cellular senescence, HMGB1 clusters were disrupted, and the TADs collapsed to form larger domains, similar to what was observed with HMGB2 [51]. This reorganization of the TAD domains as a function of HMGB1 loss during cellular senescence drives the expression of senescence associated genes and the down-regulation of proliferation associated genes.

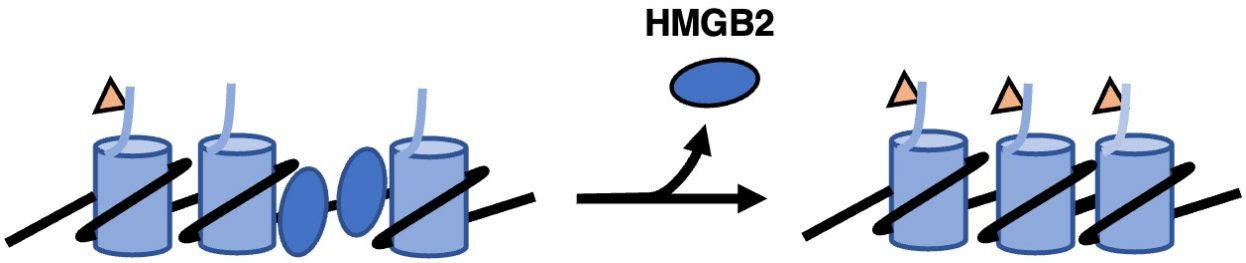

Figure 1.4: Loss of HMGB2 during Oncogene Induced Senescence promotes the spread of heterochromatic marks (triangles) into the SASP gene loci.

During oncogene induced senescence (OIS), which can be caused by mutations in oncogenes such as RAS, BRAF, AKT, E2F1, and cyclin E [66], chromosome condensation occurs to form senescence-associated heterochromatic foci (SAHF), which can be detected as bright punctum in DAPI (4'-6'-Diamidino-2-phenylindole) cell staining [67]. The formation of SAHF has been found to proceed the assembly of HP1 proteins, macroH2A, and H3K9Me onto chromosomes [68], which are all markers of heterochromatin and gene silencing. The binding of HMGB proteins to chromatin during OIS plays an important role in modulating the structure of chromatin during the formation of SAHF. During OIS HMGB2 modulates the expression of senescence-associated secretory phenotype (SASP) protein factors through regulating the chromatin structure of these genes. SASP involves the expression and secretion of cytokines, chemokines, growth factors, metalloproteases, and lipids by senescing cells, which can influence immune responses via paracrine signaling [66]. Using chromatin immunoprecipitation followed by sequencing (ChIP-seq), HMGB2 was found bound to SASP gene loci to promote the expression of SASP factors, such as IL1α, IL8, and IL6 [52]. Knockdown of HMGB2 influenced the global formation of SAHFs by decreasing the percentage of identifiable heterochromatic foci and increasing the spread of heterochromatic marks into the SASP gene loci, thereby silencing the expression of SASP factors and further compacting the genome [52] (Figure 1.4). Thus, HMGB2 promotes SASP gene expression by keeping SASP gene loci from being silenced by heterochromatin.

1.2.2 Regulation of the Cardiac Genome by HMGB2

A relationship between HMGB2 and cardiovascular disease was revealed through mass spectrometry experiments that identified chromatin-associated proteins in the hearts of mice in various stages of heart disease [69]. Changes in the chromatin proteome were consistent with global changes to chromatin structure and accessibility. For example, the reprogramming of post-translational modifications on histones and changes in the levels of architectural proteins that control chromatin structure, including HMGB2, were observed. Knockdown of HMGB2 in cardiac myocytes changed gene expression programs related to cell growth and cardiac function and changed overall levels of histone modifications, which suggested a shift toward euchromatin [69]. Further investigation of the molecular role of HMGB2 in orchestrating chromatin organization and gene expression revealed locations across the genome in cardiac cells that were bound by HMGB2 [70]. A comparison with the genomic occupancy of CTCF showed that HMGB2 and CTCF share gene targets; however, these proteins did not bind genes simultaneously in ChIP re-ChIP experiments and they did not colocalize in super-resolution imaging of myocyte nuclei. This leads to a model whereby there is a reciprocal relationship between these two architectural proteins in regulating chromatin organization of the cardiac genome [70]. This in turn helps define the changes in gene expression that occur during the development of cardiac disease.

1.2.3 Regulation of Genome Organization in a Human Malaria Parasite

A recent study has highlighted a conserved role for HMGB proteins in organizing the genome of a parasite during infection. HMGB1 from the malaria parasite Plasmodium falciparum (pfHMGB1) is important for organizing the higher order conformation of the parasite genome [71]. During infection, pathogenesis involves the transcription of a family of variant malaria genes that express different surface antigens to help escape host immune responses [72]. Expression of family members is mutually exclusive; thus, each malaria parasite only expresses a single variant gene at a given time. This differential expression is controlled by the architectural organization of the

variant genes and the localization of individual genes within centromeric or telomeric chromatin domains [73–75]. Recent experiments showed that pfHMGB1 is critical for maintaining the centromere/telomere genome organization that controls expression of the variant genes that are critical for pathogenesis [71]. pfHMGB1 was found localized to the centromeres of all 14 P. Falciparum chromosomes by ChIP-seq, along with the centromere-specific H3 variant CenH3 [71]. Upon knockout of pfHMGB1, the structural composition of the centromeres became destabilized, decreasing the intermolecular interactions between centromeres. The disrupted nuclear organization was restored by complementation with the pfHMGB1 gene. Moreover, the HMGB1-dependent changes in genome organization correlated with HMGB1-dependent changes in expression of variant genes. pfHMGB1 knockout downregulated expression of variant genes, which could also be rescued by complementation. Therefore, in the context of the malaria genome, pfHMGB1 is a critical regulator of genomic architecture, which in turn impacts the pathogenesis of malaria infection.

1.3 HMGB Proteins Bind G-Quadruplexes with Potential Effects on Cancer

HMGB1 and HMGB2 are capable of binding to non-B form DNA structures with high affinity, such as those found in cisplatinated DNA, cruciform DNA, and bent-DNA (see Figure 1.2). More recently, HMGB1 has been shown to bind G-quadruplex (G4) DNA structures, and in doing so could influence the activity of the enzyme telomerase [76, 77] and transcription of the KRAS oncogene [78] (Figure 1.5). G4 structures form through the self-association of guanines to form stacked tetrads and could serve to regulate cellular processes that occur on the genome [79, 80]. Within the human genome, there are approximately 700,000 G4 structures experimentally identified by G4-seq [81], which are enriched in transcriptionally regulated regions [82]. G4 structures are often found in oncogene promoters, such as the KRAS oncogene [83] and on the ends of telomeres [84] suggesting the G4 structures are important in cancer progression and telomere maintenance.

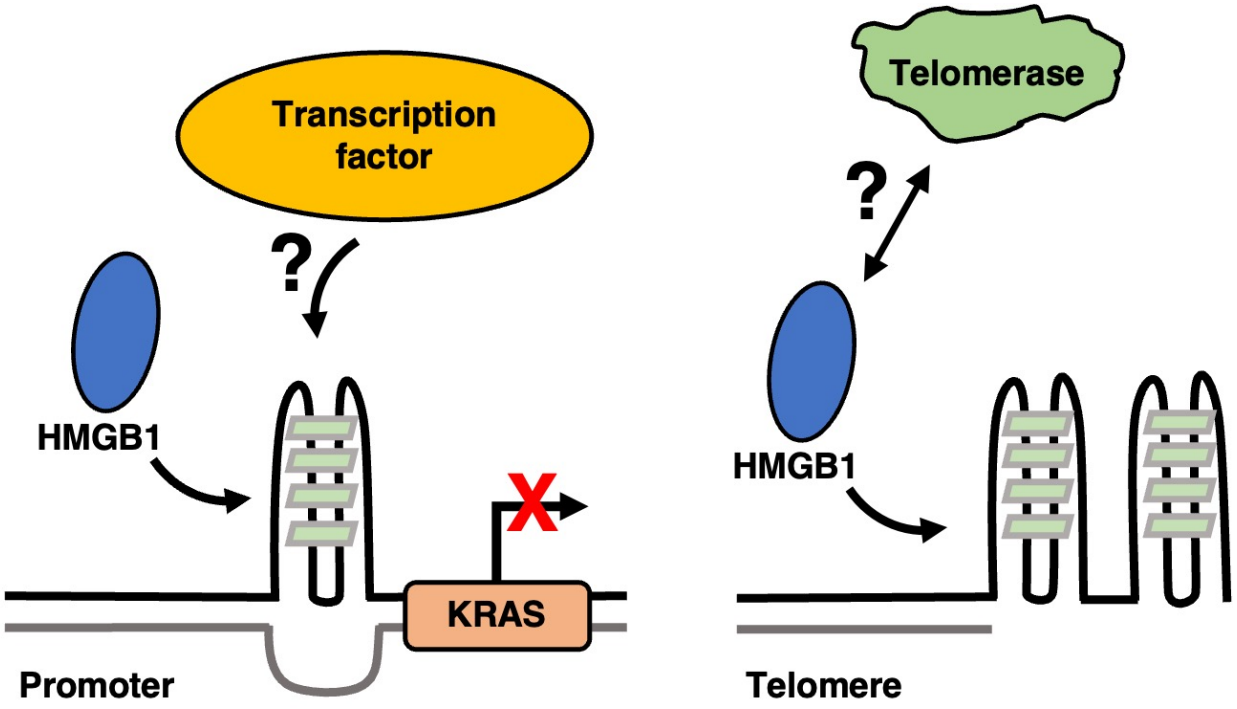

Figure 1.5: HMGB1 binding to the KRAS and Telomeric G4-quadruplex.

1.3.1 The HMGB1 Protein Binds Telomeric G-Quadruplex DNA and Affects the Activity of Telomerase

The dysregulation of telomerase in tumor cells can affect the progression of tumor growth and cellular immortalization due to disruption of the maintenance and extension of the telomeric ends of chromosomes [85]. Proper maintenance of telomeres requires the catalytic activity of the enzyme telomerase, which contains an RNA template that complements the single stranded telomeric overhangs that form after each successive DNA replication. A role for HMGB proteins in telomere homeostasis is beginning to be uncovered. Knockout of HMGB1 in mouse embryonic fibroblasts resulted in decreased telomerase activity, whereas its overexpression increased telomerase activity [76]. Further studies showed that down-regulation of HMGB1 increased the radiosensitivity of human breast cancer cells by dysregulating telomere homeostasis [86]. However, how HMGB1 affects telomerase activity and whether this requires interaction with telomeric DNA were unclear. Recently, HMGB1 was shown to colocalize with TRF1, a telomeric binding protein, and to localize to telomeres via ChIP [78]. Knockdown of HMGB1 in these cells resulted in an increase in DNA damage at telomere ends as measured by the accumulation of γH2AX, a histone variant that is a marker for DNA damage. These data suggest that the binding of HMGB1 at telomeric ends could serve to protect the ends of chromosomes from damage [78, 85]. Telomere ends contain guanine-rich DNA repeat sequences (TTAGGG) that have the propensity to fold into G4 structures [87] (Figure 1.5, right panel). Recent studies investigated the binding of telomeric G4 structures by HMGB1 using structural and biophysical approaches, characterizing in detail the parameters that govern this interaction [78]). Although HMGB1 binds to the G4 telomeric structures and localizes to the telomere ends, it remains unclear whether HMGB1 can facilitate the assembly of the telomerase machinery onto G4 structures.

1.3.2 HMGB1 Binds a G-Quadruplex in the Promoter of the KRAS Oncogene and Regulates Its Transcription

The KRAS gene encodes for a small GTPase transducer that is implicated in the RAS/MAPK signaling pathway. KRAS is upregulated in many cancers, and mutations in the KRAS gene can cause expression of aberrant KRAS proteins that promote the transformation of normal cells into cancerous cells by promoting cellular proliferation, survival, and cancer progression [88]. The promoter of the KRAS gene contains a guanine–rich element that folds into a parallel G4 DNA structure [89]. Stabilization of this G4 structure silences KRAS transcription [89, 90]. The protein hnRNPA1 has been shown to bind and destabilize the KRAS G4 structure, resulting in transcriptional activation [91]. Recently, biophysical studies found that HMGB1 can bind to the KRAS G-quadruplex structure with high affinity and stabilize the structure [92] (Figure 1.5, left panel). Importantly, knockdown of HMGB1 led to an increase in KRAS expression, supporting the model that HMGB1 binding to the KRAS G4 structure in cells results in transcriptional repression of the KRAS gene and decreased expression of the KRAS protein [92]. Given the role of KRAS in cancer, and the propensity of other oncogene promoters to contain G4 structures, this regulatory mechanism involving HMGB1 warrants further investigation.

1.4 A New Regulatory Role for HMGB Proteins as RNA Binding Factors

HMGB proteins are known to be architectural DNA binding proteins; however, a growing body of evidence shows that HMGB proteins can also function as RNA binding proteins. Although the biological roles and diversity of RNA targets remain to be revealed, data suggest that HMGB/RNA complexes have the potential to play important regulatory roles in the cell. Early evidence of HMG proteins binding to RNA was provided by the discovery that the drosophila HMG protein (HMG-D) binds to the minor groove of the transactivation response region (TAR) A-form RNA from HIV-1 and to the rev binding protein element RNA [93]. Other studies found that recombinant rat HMGB1 bound to the long branched *E.Coli* 5S rRNA and the *Azoarcus* ribozyme in

vitro with high affinity, and in the latter case HMGB1 influenced RNA splicing activity [94]. More recently HMG box proteins were found in proteomics screens aimed at comprehensively identifying RNA binding domains in human cells [95, 96]. As studies of interactions between specific RNAs and HMGB proteins grow, a clearer view of this regulatory interaction is likely to be revealed (Figure 1.6).

1.4.1 HMGB1 Coordinates RNA Metabolism During Senescence Entry

sCLIP (simplified Cross-Linking and Immuno-Precipitation), a high-throughput sequencing method used to identify RNA-protein interactions after crosslinking [97], was used to identify RNA binding partners for HMGB1 in proliferating IMR90 cells [50]. This resulted in identification of non-coding RNAs and senescence-related mRNAs, as well as mRNAs that encode regulators of splicing and chromatin reorganization [50]. Knockdown of HMGB1 altered the splicing of many of its target mRNAs, a portion of which also showed changes in splicing during senescence, suggesting a functional role for HMGB1 in mRNA processing during cellular senescence. In addition, identifying the HMGB1 interactome revealed a myriad of RNA binding proteins (RBPs) that are regulated during senescence, and the mRNAs of some of the RBPs were also bound by HMGB1 [50]. It is possible that a regulatory circuit involving HMGB1/mRNA/RBP interactions participates in regulating senescence and mRNA splicing. Interestingly, studies of the HMGB1 interactome in prostrate epithelial cells identified, among other proteins, members of the SR protein family (SRSFs) [98], which are involved in the regulation of RNA splicing, and were also identified as HMGB1 partners in IMR90 cells during senescence [50]. Other studies have shown a relationship between HMGB1 and SRSF3 in controlling the expression and secretion of the IL-1β mRNA [99], a proinflammatory cytokine that plays a role in SASP activation [100, 101]. Further investigations probing how HMGB1 interactions with RNA binding proteins and mRNAs modulate gene expression will likely provide novel insight into new regulatory mechanisms in diverse cell types and diseases.

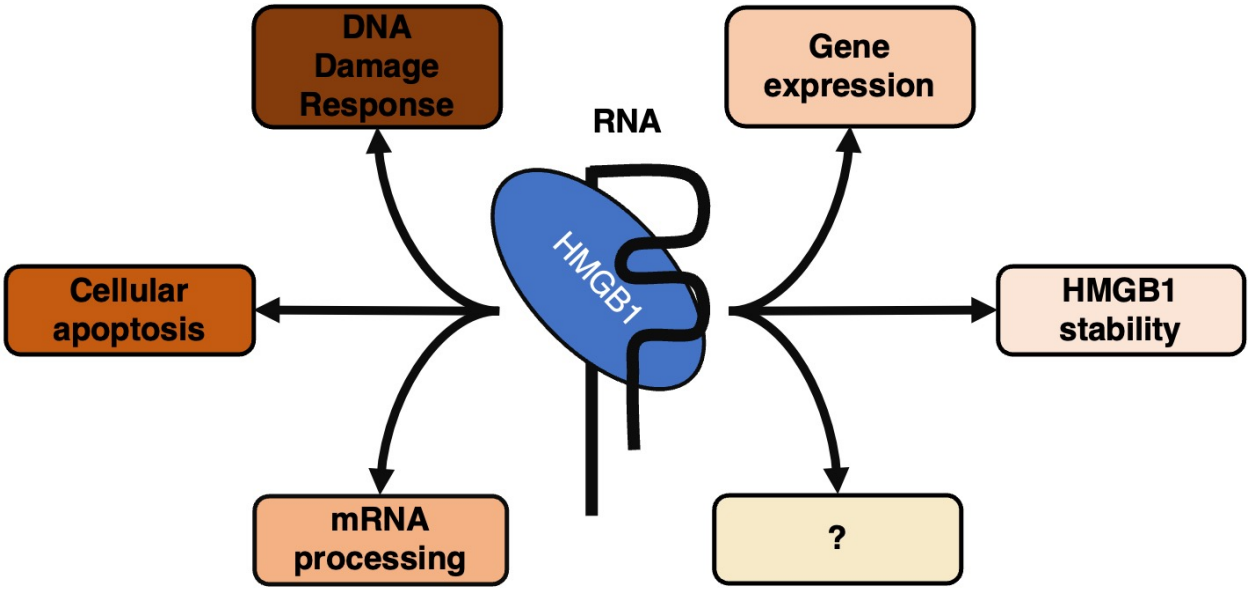

Figure 1.6: HMGB1/RNA interactions can modulate the stability of the HMGB1 protein within the cell as well as a number of cellular processes

1.4.2 HMGB1 Interacts with Long Noncoding RNAs to Control Disease States

Long non-coding RNAs (lncRNAs) have been implicated to have roles in a diversity of cellular processes and diseases [102]. It was recently reported that HMGB1 associates with the brain specific DNA damage related lncRNA1 (BS-DRL1) [103]. Data suggest that binding of HMGB1 to BS-DRL1 in neuronal cells directs HMGB1 to sites of DNA damage on chromatin to help facilitate the repair process. Knockdown of BS-DRL1 resulted in decreased HMGB1 binding to chromatin and an increase in DNA breaks, but decreased DNA damage response signaling [103]. Association of HMGB1 with BS-DRL1 is mediated through the N-terminus of the protein, and disruption of this interaction impacted the integrity of the genome by increasing the accumulation of DNA damage, motor dysfunction, and neurodegeneration. Together these results provide compelling evidence that an interaction with a non-coding RNA controls the ability of HMGB1 to bind chromatin in response to DNA damage in neurons.

In a different biological system, the interaction of HMGB1 with an lncRNA was found to reduce degradation of HMGB1 within the cell and promote tumor cell growth [104]. In multiple myeloma cells, HMGB1 was found to associate with the lncRNA MALAT-1 in a pull-down assay. Interestingly, knockdown of MALAT1 increased the degradation of HMGB1, and increased its ubiquitination, which suggested a potential mechanism for degradation. Treatment with siRNA against MALAT-1 increased apoptosis of the multiple myeloma tumor cells, which was attenuated by expression of HMGB1 [104]. This regulatory network provides new insight into the pathological process of multiple myeloma. It will be interesting to learn of other interactions between HMGB1 and lncRNAs and the mechanisms by which these interactions influence multiple diseases.

1.5 Conclusions and Future Directions

To summarize, HMGB proteins are the most abundant non-histone chromatin binding proteins in the nuclei of mammalian cells. The HMGB proteins are well known for modulating the local chromatin environment, facilitating the binding of other proteins to chromatin, and control-

ling nuclear processes such as transcription, DNA damage repair, and nucleosome sliding. Recent evidence shows that the role of HMGB proteins in the nucleus extends to regulating global chromatin architecture, telomere maintenance, cellular senescence, and RNA biology. For example, the loss of HMGB1/2 from the nucleus results in the reorganization of genome architecture, which has direct implications for gene expression and ultimately the progression of cancers [51, 52], cardiovascular diseases [70, 105], and parasitic immune evasion [71]. The binding of HMGB1 to G4 quadruplex DNA stabilizes the DNA structure; at the KRAS oncogene this inhibited expression and at telomere ends this protected the DNA from damage [78, 92]. Newly identified interactions between HMGB1 and various mRNAs or lncRNAs modulate splicing, gene expression, and the formation of ribonucleoprotein complexes within the cell [50, 103, 104]. The varying roles of nuclear HMGB1/2 proteins within different cell types highlights the complexity of this family of proteins in regulating genomic processes.

Much remains to be learned about the regulatory roles of nuclear HMGB proteins. It will be interesting to unravel the relationship between genome organization and the HMGB-facilitated binding of transcription factors. This could involve interplay with the formation of local DNA structures such as G4 quadruplexes or cruciform DNA. For example, both p53 and HMGB1 exhibit enhanced binding to cruciform DNA, suggesting that the local DNA structure may play a regulatory role in transcription factor binding facilitated by HMGB proteins [28]. With over 700,000 experimentally identified G4-quadruplex structures in the human genome [81], this is an intriguing possibility. Future studies are required to unravel the relationship between the genomic positioning of local DNA structures, HMGB1/2 binding, localization of specific transcription factors, and the presence of looped regions of chromatin. In addition to revealing how HMGB proteins function on the genome, it will be important to identify the breadth of potential RNA targets of HMGB proteins. Only then will we realize how widespread RNA-mediated regulatory mechanisms for HMGB proteins might be. Lastly, characterizing how these regulatory mechanisms differ in healthy versus diseased cells is critical for future understanding of how HMGB proteins contribute to disease eitology and progression.

Studies have suggested that HMGB1 has the potential to be a therapeutic target, as reviewed in detail elsewhere [106]. In addition, it is possible nuclear HMGB1 could serve as a prognostic biomarker, given its elevated levels in cancer cells that reflect the proliferative state of the cell [107]. The majority of studies of nuclear HMGB proteins have been performed in cells and in vitro. In the future, such studies could extend to in vivo models. Indeed, a conditional knockout mouse for HMGB1 has been developed [108] Future work will undoubtedly reveal the potential of HMGB proteins to serve as therapeutic and/or diagnostic molecules.

Chapter 2

Understanding how HMGB1 navigates DNA using single-molecule TIRF microscopy

HMGB1 is a highly mobile architectural DNA binding protein that is implicated in several nuclear processes, including transcription. How HMGB1 navigates the nuclear environment to search for DNA binding sites remains unclear. To better understand this, I employed a single molecule TIRF microscopy system to study the movement of HMGB1 across a dense surface of DNA molecules. The mean speeds of the moving molecules were quantified for full-length HMGB1. To gain insight into mechanisms of movement, mean speeds of movement were also analyzed for HMGB1 deletions, decreased surface DNA densities, shortened DNA lengths, and different ionic conditions. In most of the conditions tested, the distribution of the mean speeds of tracked molecules increased compared to full length HMGB1. The exception was the C-terminal tail deletion mutant, which exhibited a slight decrease in the distribution of the mean speeds. The average number of moving molecules per region greatly decreased upon decreasing the density of immobilized DNA, shortening the DNA, increasing the ionic strength, and using a dense surface of heparin in place of DNA. Collectively, the data support the model that each domain of HMGB1 contributes to setting its speed of its movement across a surface of DNA. Moreover, movement of HMGB1 is mediated by electrostatic interactions, and robust movement requires longer lengths of DNA within close proximity of each other.

2.1 Introduction

The family of High Mobility Group (HMG) proteins was first discovered by Goodwin et al. in 1973 as a group of chromatin associated proteins that exhibited fast migration on an acidic polyacrylamide gel [1]. HMG proteins play an important role in modulating chromatin dynamics and facilitating numerous nuclear processes to influence cellular proliferation and metabolism [3]. Among the three HMG protein families - HMG AT-Hook (HMGA), HMG box (HMGB), and HMGN - the HMGB protein family is the most abundant group within the nucleus [109]. The HMGB protein family currently consists of four protein variants (HMGB1, HMGB2, HMGB3, and HMGB4) [14]. Within the HMGB protein family, HMGB1 is the most abundant protein among the four, where 10^6 molecules are found within nuclei, averaging 1 HMGB1 per 10-15 nucleosomes [14, 15]. HMGB1 also has important functions in the cytoplasm and extracellularly [110], although its role as a nuclear DNA binding protein is the focus of the studies described here.

HMGB1 is a 26kDa non-sequence specific DNA binding protein. It consists of two DNA binding domains, the A-box and B-box, and an unstructured acidic C-terminal tail that contains a stretch of glutamate and aspartate residues [14]. The C-terminal tail is implicated in dampening HMGB1's affinity for DNA by interacting intramolecularly with its DNA binding domains [31–34, 36, 37]. HMGB1 is capable of bending linear DNA when it binds by intercalating specific amino acid residues within each of the HMG boxes in between the bases of the DNA double helix to deform the DNA structure [30]. These intercalating residues consist of a phenylalanine within each of the HMG boxes (A-box: Phe38 and B-box: Phe103) and an isoleucine within the B-box (Ile122) [30, 36]. HMGB1 exhibits a higher affinity for non B-form DNA, such as cisplatinated DNA [22], cruciform DNA [23], single stranded DNA [24], supercoiled DNA [25], hemi-catenated DNA [26], and DNA mini circles [27].

Within the nucleus, HMGB1 is thought to modulate the local chromatin environment to facilitate the accessibility of DNA to various protein factors that control transcriptional activation [38, 39], DNA damage repair [46], and other processes that happen on chromatin. For example,

previous studies have shown that HMGB1 facilitates the binding of p53 [47], estrogen receptor [48], and numerous other transcription factors [49] to their cognate DNA binding sites. It has been proposed that HMGB1 facilitates transcription factors binding to their cognate DNA binding sites using a "hit-and-run" mechanism [48, 111, 112], where HMGB1 would present a bent DNA to a nuclear factor and then dissociate once this factor binds.

The association of HMGB1 with nuclear chromatin is thought to be very dynamic and transient. Fluorescence-based studies showed that the majority of nuclear HMGB1 stays bound to chromatin for an average time of about 5 seconds [113]. Additional studies have shown that HMGB1 is capable of traversing the entire nucleus within 1-2 seconds [111, 114]. HMGB1's dynamic behavior could be attributed to its weak association with interphase chromatin [115]. Some in vitro studies investigating HMGB1's DNA binding and bending mechanisms showed that once bound to DNA, the HMGB1-DNA complex remains stable [36, 116]; however, in the presence of excess DNA, HMGB1 appears to readily dissociate from the HMGB1-DNA complex, possibly due to a direct transfer mechanism that hands HMGB1 off to a nearby DNA molecule [116]. The mechanisms that HMGB1 uses to navigate the nuclear environment in search of DNA binding sites remains unclear.

In the research described in this chapter, I used single molecule Total Internal Reflection Fluorescence (TIRF) microscopy to better understand the properties that facilitate HMGB1 movement on a dense DNA surface. This provided a well-defined, controlled system that mimics the characteristics of a dense DNA environment found within the nucleus. The mean speeds of single molecule tracks of HMGB1 moving across the surface were measured under various protein and surface conditions. I first examined how HMGB1's movement changed with protein truncation mutants. This was followed by investigating how decreasing the DNA surface density and shortening the length of the DNA impacted HMGB1's movement. Lastly, I tested the importance of electrostatic interactions in facilitating HMGB1 movement by varying the salt concentration and densely coating the surface of the microscope chamber with heparin in place of DNA. The findings in this study suggest a model where both DNA binding domains connected in tandem set the speed

with which HMGB1 searches for its DNA binding sites, and that the electrostatic interactions are the major drivers that facilitate HMGB1 movement between DNA molecules.

2.2 Results

2.2.1 A single molecule system for studying HMGB1 movement on DNA

To study HMGB1 movement on DNA, I created a dense surface of immobilized DNA on a slide. A mixture of 10pM biotinylated-Cy3 157 bp DNA and 25 nM biotinylated unlabeled 157 bp DNA was flowed onto a microscopy slide functionalized with biotin-PEG (Figure 2.1). Because the unlabeled DNA is not visible under TIRF conditions, the Cy3-DNA was used to measure the density of the DNA on the surface. This was done by imaging and collecting movies in the Cy3 channel (Figure 2.2) which were then used to quantify the density of the Cy3-DNA molecules. This was then used to calculate the number of unlabeled DNA molecules per pixel based on the relative concentrations of labeled and unlabeled DNA in the sample flowed into the slide chamber. The density of unlabeled DNA on the surface of the microscopy slide was 40±5 molecules per pixel. Considering a pixel size with a diameter of 200nm and assuming a surface of evenly spaced DNA, this translates to about 31.7±2.3 nm of distance between adjacent DNA molecules. The average length of a DNA base pair is 3.4 Å and the estimated length of the biotinylated 157 bp DNA construct is 54 nm. With the possibility of DNA swaying on the surface of the slide, there is potential for the ends of two swaying DNA molecules to overlap with one another, which may be sufficiently close to allow for an HMGB1 bound to one DNA to move to a neighboring DNA molecule without fully dissociating from the first DNA prior to interacting with the second.

Various HMGB1 constructs were studied in this chapter (Figure 2.3). All HMGB1 constructs had all three of its native cysteine residues mutated to serine (C23S, C45S, C106S). The three cysteine residues each reside on the inner face of a helix within the HMG boxes and is positioned away from the DNA binding surfaces. A single cysteine was engineered into the protein sequence to allow for AF647 maleimide labeling. For HMGB1 constructs that contain an A-box domain, the cysteine

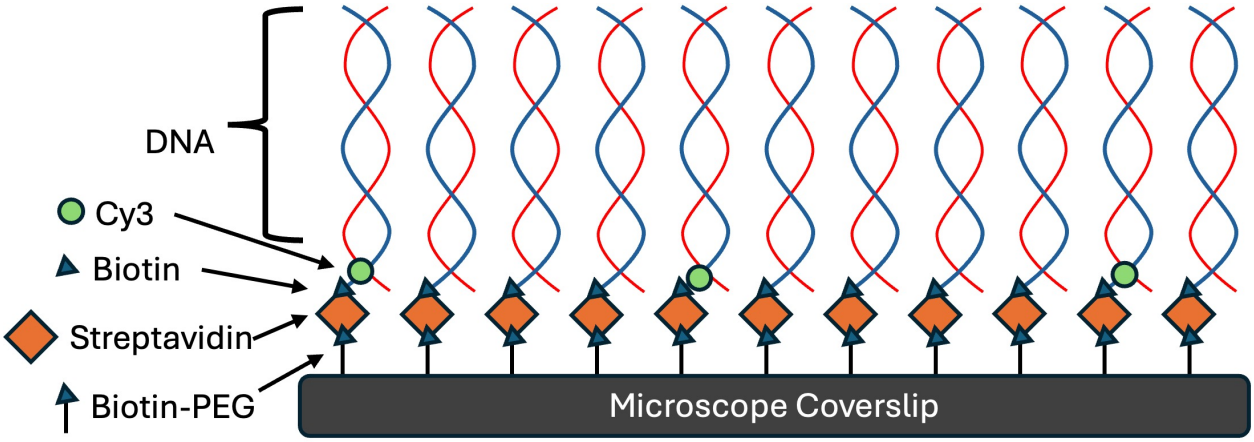

Figure 2.1: A schematic showing a microscopy cover slip (Gray Rectangle), densely coated with a mixture of unlabeled and Cy3-labeled biotinylated DNA

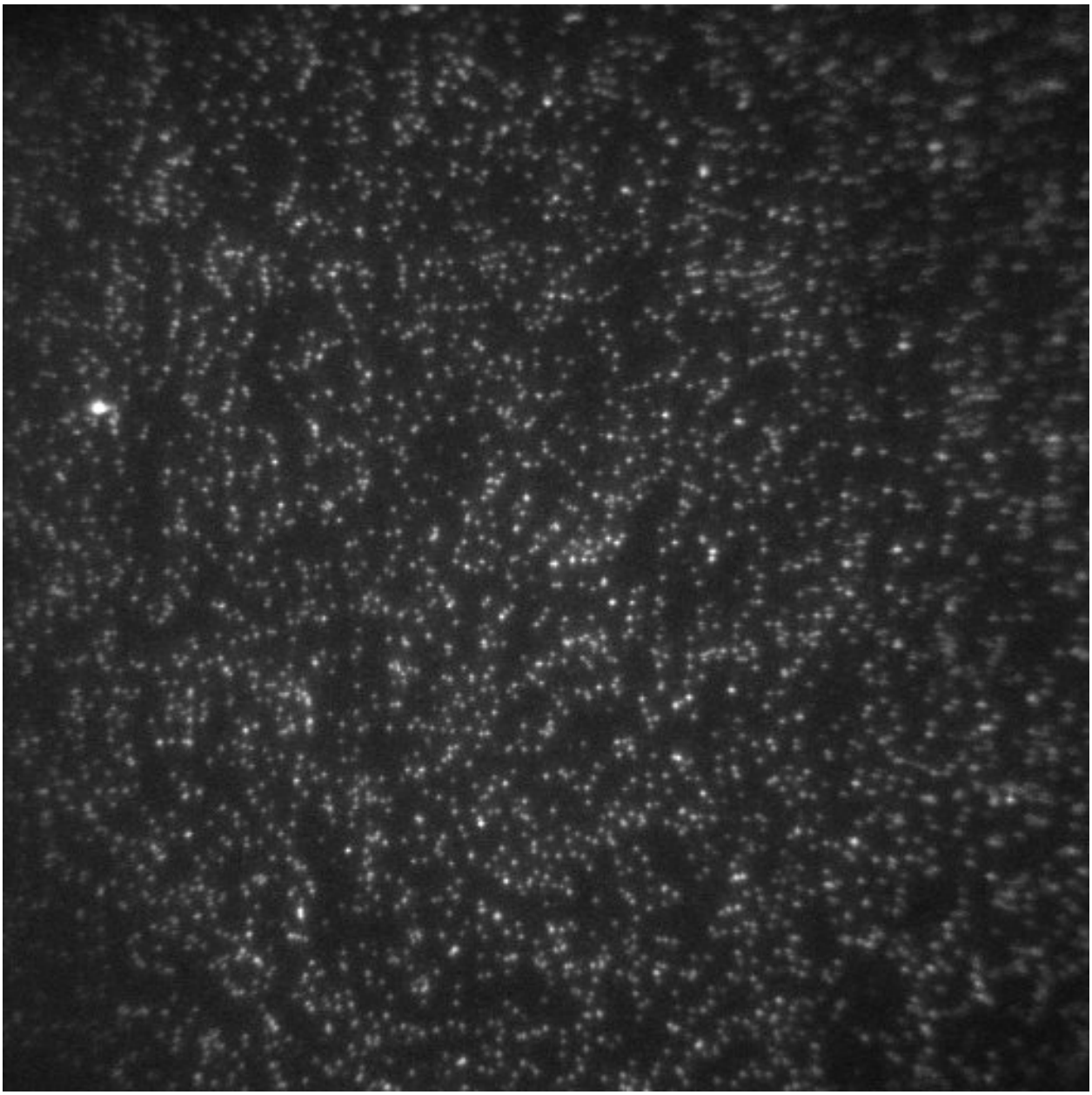

Figure 2.2: A TIRF-M image of immobilized 157 bp Cy3 DNA introduced into the slide chamber at a concentration of 10 pM along with 25 nM unlabeled DNA. The Cy3 DNA was used to measure the surface density of DNA, which was 40±5 molecules per pixel.

residue was engineered away from the DNA binding surfaces by mutating serine-35 to cysteine-35 (S35C). For the HMGB1 B-box truncation mutant, the cysteine was similarly engineered away from the DNA binding surfaces by mutating alanine-137 to cysteine-137 (A137C). The HMGB1 constructs were expressed, purified, and labeled with maleimide-AF647 dye. Afterwards, excess free maleimide-AF647 dye was removed from the protein sample and the protein concentration was determined with denaturing gel electrophoresis using known concentrations of BSA to form a standard curve.

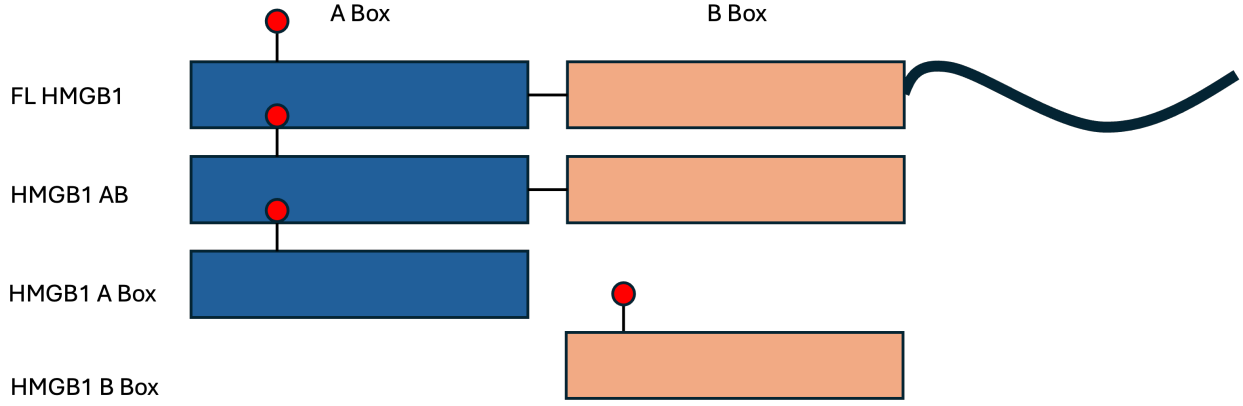

Figure 2.3: A Schematic of the various HMGB1 constructs used to study movement on a dense surface of DNA

2.2.2 Full Length HMGB1 moves on a dense surface of DNA

AF647-HMGB1 at a concentration of 25pM was added into the flow chamber and movies were collected in the AF647 channel. These movies were then analyzed via TrackMate [117], single molecule tracking and analysis software that comes with the ImageJ package. Detection of AF647-HMGB1 was done by applying an intensity thresholding parameter using the Differences of Gaussian detector to filter out background signal. Once the single molecule dyes were identified, the simple Linear Assignment Problem (LAP) tracking algorithm was used to track each molecule over subsequent frames. Afterwards, the tracks were further filtered to remove noise from the data by excluding molecules that persisted on the surface for less than 5 frames (< 300 ms) and traveled

a maximum distance less than 15 pixels from the starting pixel. Figure 2.4 shows a representative track following the movement of a single AF647-HMGB1.

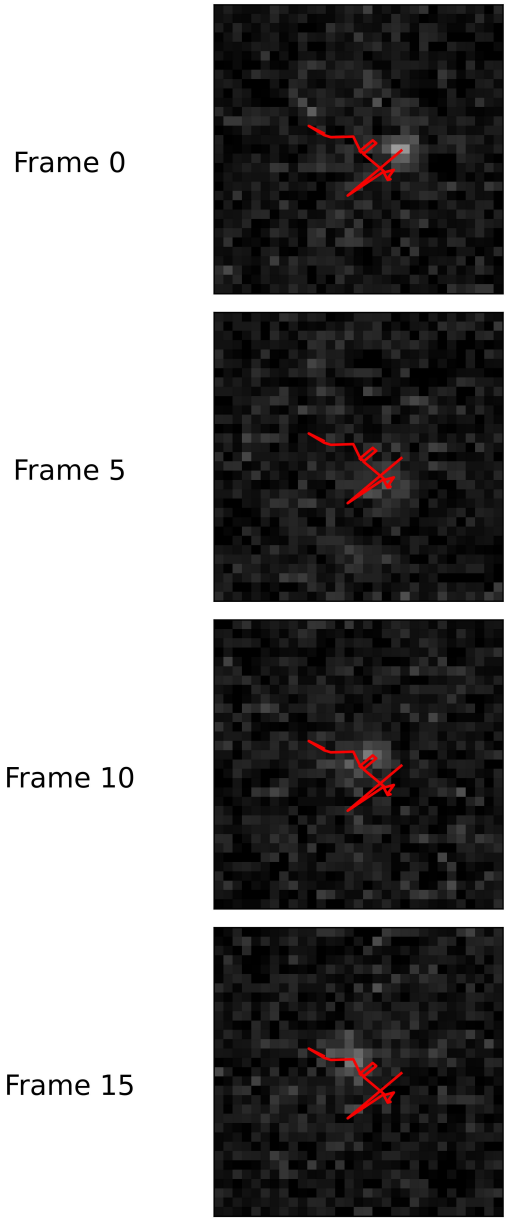

Figure 2.4: A representation of a track (red) overlayed onto a single molecule of FL HMGB1 moving on a dense surface of DNA (the bright puncta close to the center of each frame).

On a dense surface of DNA, AF647-HMGB1 produced an average of 124 tracks per region. By contrast, on a no DNA control slide (n = 3), AF647-HMGB1 produced only 2 tracks per region

(Table 2.1). Therefore, observing HMGB1 movement on the surface required DNA. For each tracked molecule, I focused on the track mean speed (μm/s) as it allows me to quantitatively assess and describe how HMGB1 behaved on a DNA surface. The mean speed distribution for thousands of tracked molecules across five biological replicates of AF647-HMGB1 were summed and visualized in a histogram (Figure 2.5A). As shown in Figure 2.5B, the experimental replicates were reproducible.

To assess the distribution characteristics of the histogram in Figure 2.5A, I calculated the average, median, interquartile range, excess kurtosis, and skewedness, which are shown in Table 2.1. The average (4.8 μm/s) is greater than the median (4.0 μm/s), consistent with the histogram appearing to be skewed to the right. The excess kurtosis and skewedness values are two descriptive statistical measurements that can be used to describe the histogram shape [118]. The excess kurtosis measures how the density of data disperses between the distribution's center and tails, reflecting the density of the more extreme values within the histogram in comparison to a normally distributed data set. Skewedness measures how the observed histogram deviates, or skews, from a normal distribution. For example, a normally distributed population would have excess kurtosis and skewedness values of 0. A negative excess kurtosis can be described by the broadening of the central peak and fewer extremes along the tails compared to a normal distribution. By contrast, a positive excess kurtosis produces a sharper central peak and heavier tail with a greater density of extreme values compared to a normal distribution. Regarding skewedness and how it relates to the shape of the population distribution, a negative skewedness value would result from a left skewed population distribution with a tail that extends to the left of the central peak, whereas a positive skewedness value would result from a right skewed population with a tail that extends to the right of the central peak. The shape of the mean speed distribution for FL HMGB1 on DNA produced a right skewed distribution with a corresponding skewedness of 1.36. AF647-HMGB1 movement produced several tracks with high speeds, reflected by the heavier tail which is described by its excess kurtosis of 2.13.

HMGB1 Construct	Surface Conditions	Total Tracks	Average Tracks per Region	Average of Mean Speed (μm/s)	Median of Mean Speed (μm/s)	Interquartile Range (μm/s)	Excess Kurtosis	Skewedness
FL	Dense 157 bp DNA	3457	133	4.8	4.0	2.5 - 6.2	2.13	1.36
FL	No DNA Control	18	2	-	-	-	-	-
AB	Dense 157 bp DNA	2405	200	4.1	3.4	2.2 - 5.3	3.22	1.54
A-Box	Dense 157 bp DNA	1757	146	8.5	8.2	5.1 - 11.6	-0.63	0.30
B-Box	Dense 157 bp DNA	3214	134	9.9	10.0	6.6 - 13.0	-0.74	0.02
FL	0.5x 157 bp DNA	324	20	12.1	12.5	9.3 - 15.3	-0.32	-0.36
FL	0.25x 157 bp DNA	260	17	13.2	14.1	11.3 - 16.0	0.82	-0.98
FL	Dense 30 bp DNA	667	30	7.8	6.9	4.2 - 11.1	-0.55	0.58
FL	High salt	33	8	-	-	-	-	-
FL	Dense Heparin	646	27	5.9	5.4	3.5 - 7.9	1.07	0.92

Table 2.1: Parameters that describe the speed measurements from different experimental conditions.

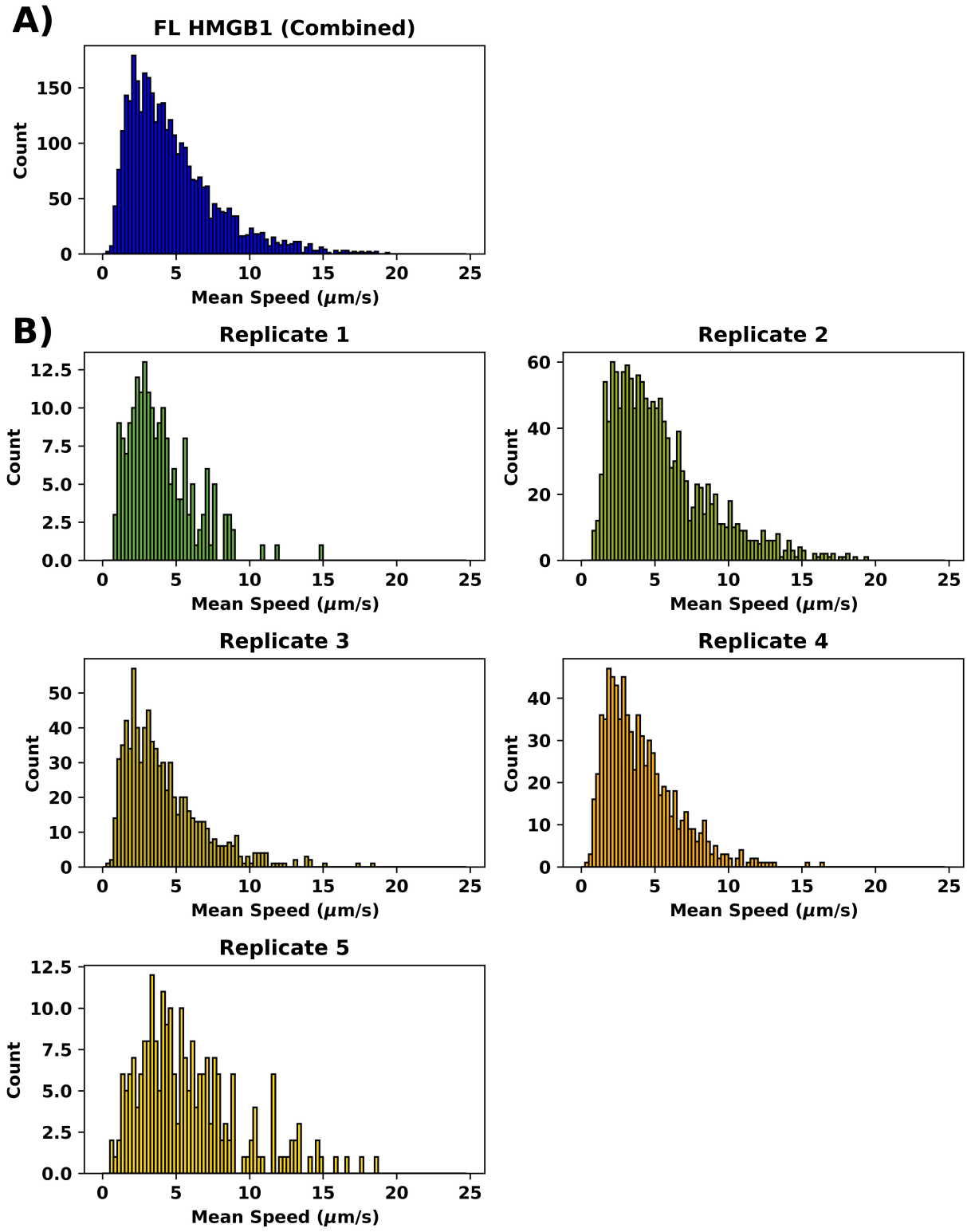

Figure 2.5: A) The mean speed from thousands of tracks over several replicates (n = 5) were combined and plotted to generate a right-skewed histogram. B) The mean speed distribution between five expeirmental replicates on 157 bp DNA were reproducible.

2.2.3 HMGB1 movement on dense DNA is controlled by both of HMGB1's DNA binding domains

I next asked how each of the individual HMGB1 domains contributes to its ability to move on a dense surface of DNA. To better understand how each of the HMGB1 domains participate in movement, I expressed and purified 3 HMGB1 truncation mutants - the HMGB1 C-tail deletion mutant (HMGB1 AB), HMGB1 A-Box, and the HMGB1 B-Box. Each truncation protein was labeled with AF647 using maleimide chemistry. Using the same DNA surface and tracking analysis system described above, I first examined how AF647-HMGB1 AB behaved on DNA. The removal of the C-terminal tail slightly reduced HMGB1's average of the mean speed by 13.6% on a dense surface of DNA (Table 2.1). Plotted in Figure 2.6A is the distribution of mean speeds for thousands of tracked molecules of AF647-HMGB1 AB from three biological replicates. The histogram is right skewed similar to what was observed for FL HMGB1. To compare the shapes of the speed distributions between FL HMGB1 and HMGB1 AB, the counts for both data sets were normalized by dividing each count by the highest count in each respective data set. The outlines of the speed distribution for both datasets were then overlaid on a single plot as seen in Figure 2.6B. The right skew of the HMGB1 AB data was slightly greater, the central peak sharper, and the tail heavier when compared to FL HMGB1. This is reflected by HMGB1 AB's increase in skewedness and excess kurtosis over FL HMGB1 (Table 2.1). HMGB1 AB's tail quickly trails off when compared to FL HMGB1, reflecting fewer molecules with extremely fast mean track speeds; however, there are gaps between binned data points towards the right end of its mean speed distribution (Figure 2.6A). These gaps between the binned data towards the ends of the tail likely increased the excess kurtosis served as large outliers.

I next examined how the individual HMG boxes, the A-box and the B-box, behave on a dense surface of DNA. The individual HMG boxes were both more than 2-fold faster than HMGB1 AB (Table 2.1) It appears that the linkage of both HMGB1's DNA binding domains (i.e. the AB construct) slows its speed when compared to the individual HMG boxes, and that both protein

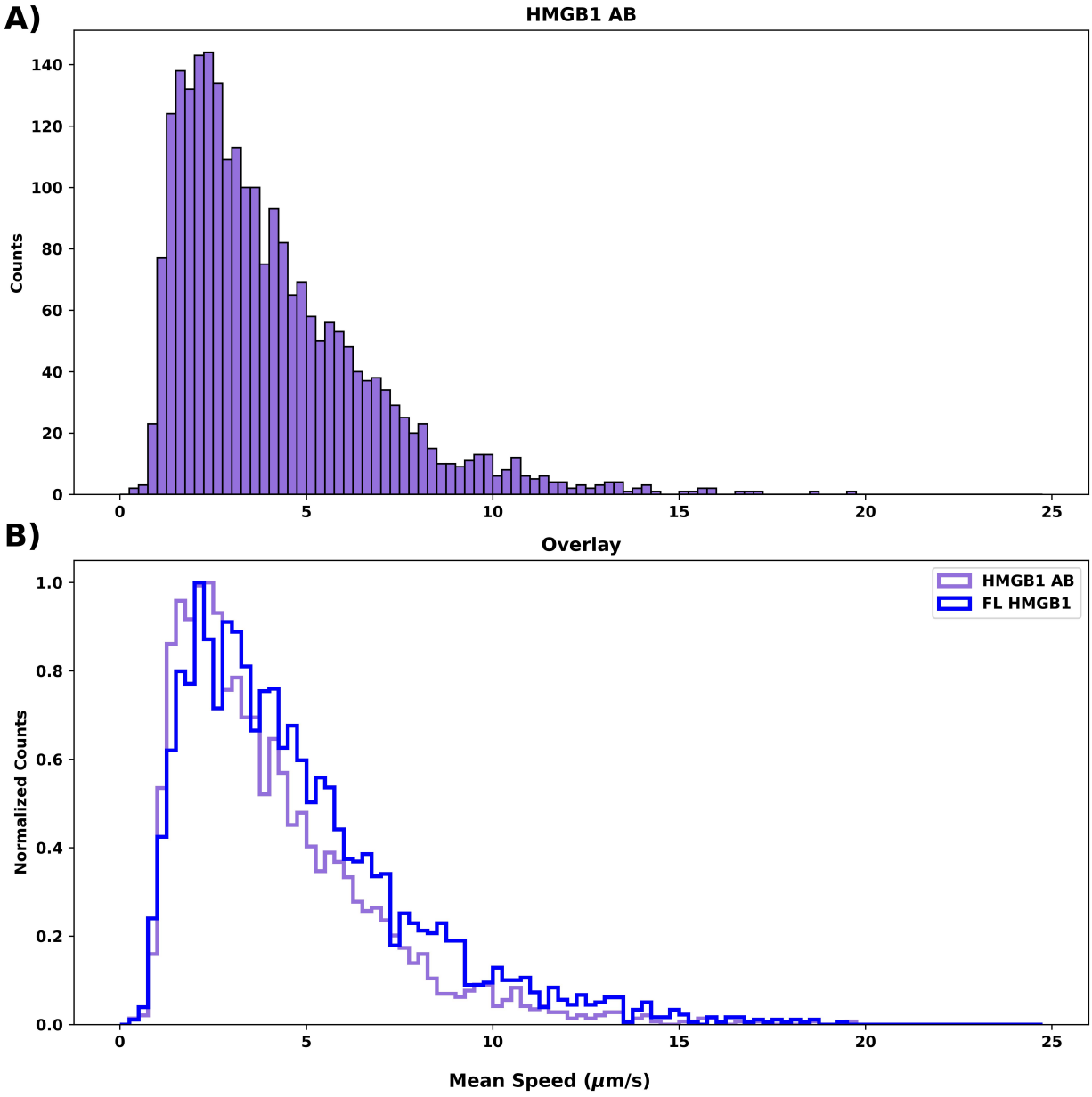

Figure 2.6: (A) The mean speed distribution of thousands of tracks for HMGB1 AB (n = 3) were plotted as a histogram. (B) The removal of the C-terminal tail generates a right skewed histogram with a shape similar to FL HMGB1. The truncation slightly decreased the mean speed relative to FL HMGB1.

domains in tandem may be required to set the speed with which HMGB1 moves on a dense DNA surface. The A-box produced a histogram of mean speeds that closely resembled a broad normal distribution as seen in Figure 2.7A. The excess kurtosis of -0.63 supports the broadening of the central peak relative to a normal distribution and a skewedness of 0.3 shows that the A-box presents a slight right skewed distribution. The B-box mean speeds produced a broad normal distribution (Figure 2.8A) like that of the A-box with an excess kurtosis of -0.74 and even smaller skewedness of 0.02. Figure 2.7B and Figure 2.8B show the A-box and B-box histograms overlaid with the AB data, normalized to the same number of tracks to compare the histogram shapes. Both the A-box and B-Box constructs behaved similarly to each other but differ quite drastically from the HMGB1 AB construct. The excess kurtosis and skewedness for the A-Box and B-box were lower than that of HMGB1 AB, which is consistent with the broadening of the peaks decreasing the excess kurtosis value and the normal distribution like characteristics brings the skewedness values closer to 0.

2.2.4 HMGB1 movement is dependent on DNA density and DNA length

Next, I tested whether the speed of HMGB1 movement is dependent on the density of immobilized DNA. I hypothesized that decreasing the DNA surface density should decrease the number of HMGB1 tracks recorded. I decreased the concentration of DNA added to slide chambers 2-fold, from 25 nM to 12.5 nM, and examined the movement of FL HMGB1 on the less dense DNA surface. This 2-fold decrease in surface density decreased the average number of tracks per region nearly 7-fold (Table 2.1), as hypothesized. Hundreds of tracks were identified across biological replicates. The average and median of the mean track speeds increased as the surface density decreased (Table 2.1 and Figure 2.9). Hence, the mean speed distribution shifted to the right compared to FL HMGB1 on 25 nM of DNA and produced slightly left skewed histogram (skewedness of -0.36). Decreasing the surface density from 25 nM DNA to 12.5 nM generated far fewer scanning molecules and increased the average of the mean track speed of HMGB1.

I assessed whether decreasing the DNA concentration further would continue to change the speed distribution characteristics or ablate the number of observable events. To do this, I decreased

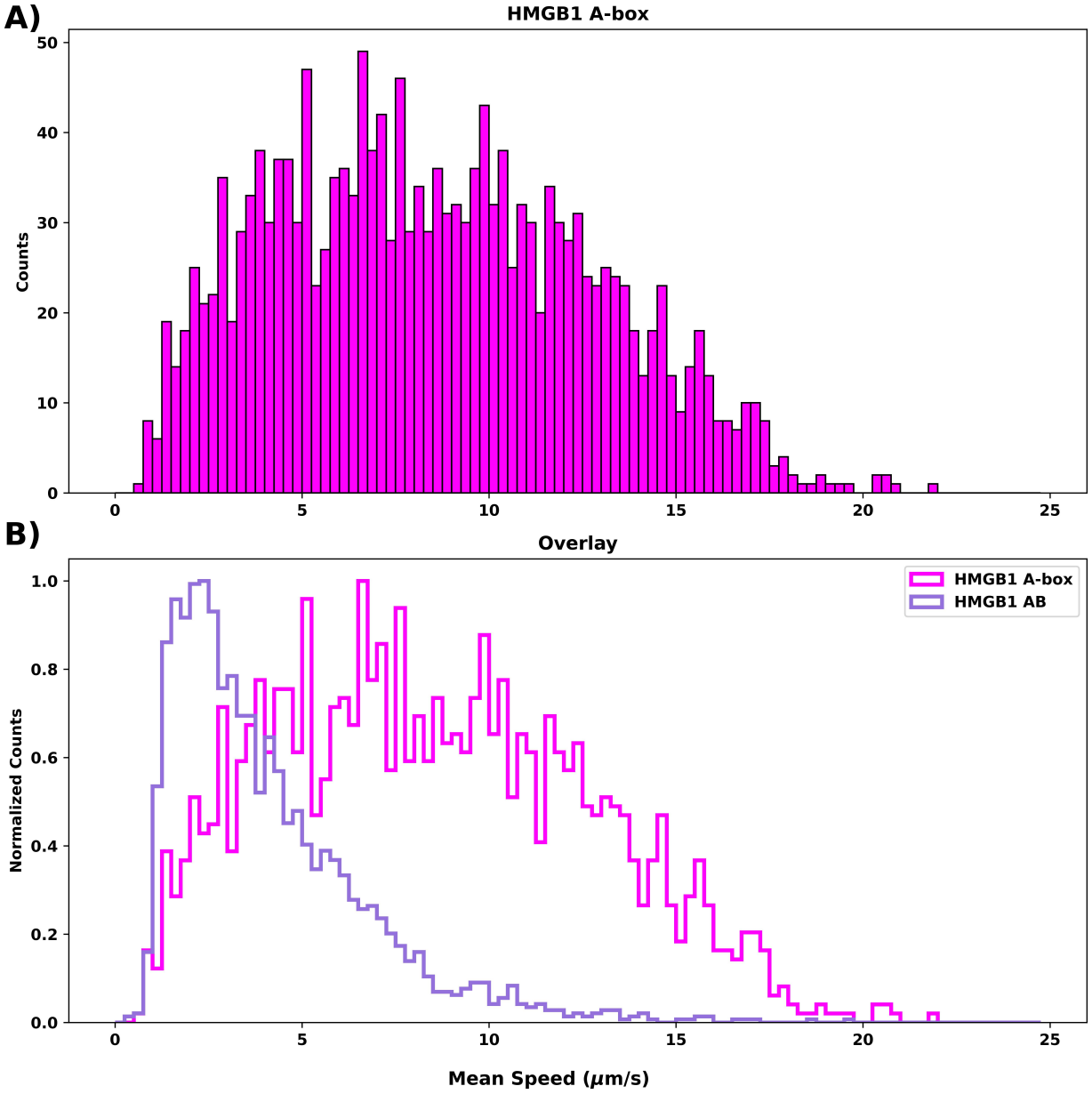

Figure 2.7: A) The distribution of mean speeds for HMGB1 A-box from 3 replicates produce a broad slightly right skewed distribution. B) The distribution characteristics of HMGB1 A-box drastically differ from HMGB1 AB. HMGB1 A-box moves faster than HMGB1 AB.

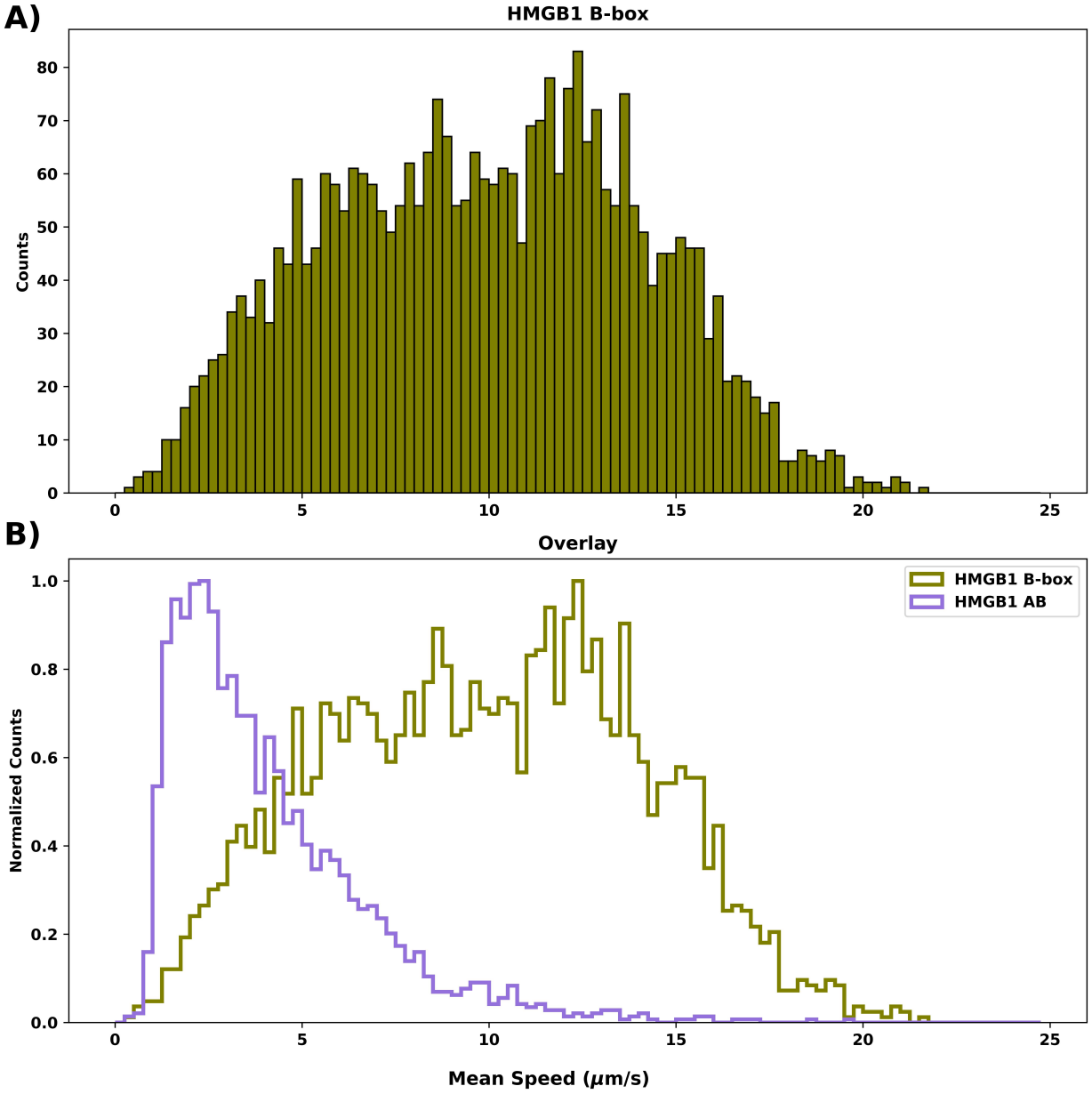

Figure 2.8: A) The mean speeds of HMGB1 B-box from 4 replicates produced a broad normal distribution. B) Similar to the A-box, the shape of the speed distribution drastically differs from HMGB1 AB, shifting the distribution to the right to signify increased movement speeds.

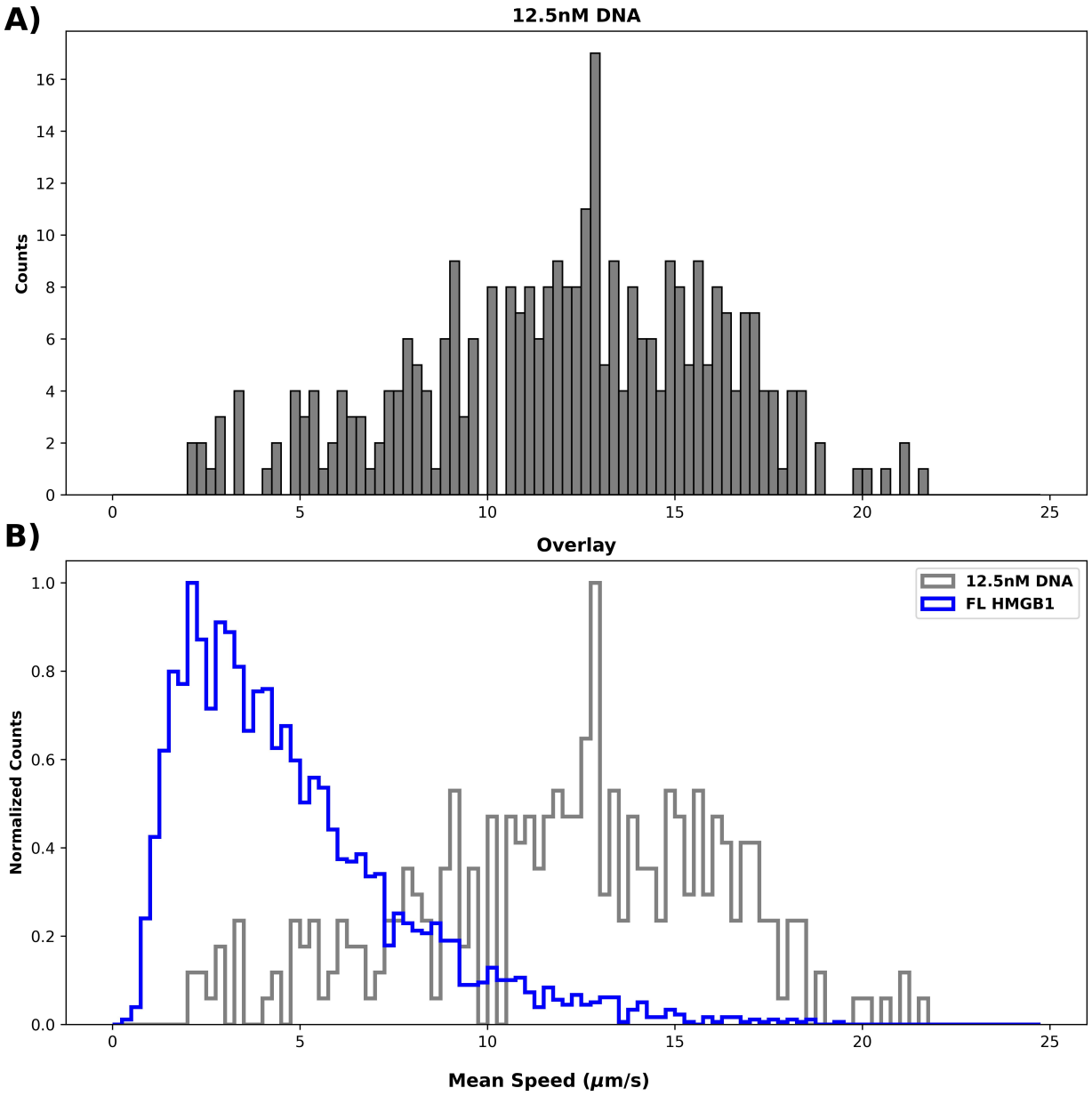

Figure 2.9: A) The mean speed distribution of 2 replicates of FL HMGB1 on 12.5 nM DNA produced a slight left skewed distribution. B) When compared to FL HMGB1, the mean speed distribution shifted to the right upon decreasing the surface density.

the amount of DNA added to the surface by another 2-fold, changing the concentration flowed into slide chambers to 6.25 nM of DNA. When compared to the 12.5 nM DNA surface, movement on the 6.25 nM DNA surface showed only small changes in the number of tracks observed per region, and in the average of the mean track speeds (Table 2.1). On the 6.25 nM DNA surface, the distribution of mean track speeds was further shifted to the right and generated a slightly more left-skewed dataset with a skewedness of -0.98 compared to the 12.5 nM data (Table 2.1 and Figure 2.10). Furthermore, decreasing the DNA surface concentration slightly increased HMGB1's average mean speed from 24.2 μm/s on the 12.5 nM DNA surface to 26.4 μm/s on the 6.25 nM DNA surface. Together the data in Figures 2.9 and 2.10 show that the initial decrease in the density of DNA drastically decreased the number of single molecule tracks per region and increased the HMGB1's speed, with smaller subsequent changes as the density decreased further.

I next asked if shortening the immobilized DNA, while maintaining a high DNA density, would impact HMGB1's movement. I examined FL HMGB1 movement on a dense surface of 30 bp DNA, which is still longer than required for HMGB1 binding, and compared its movement behavior to FL HMGB1 on 157 bp DNA. The shortening of the DNA to 30 bp decreased the average tracks per region by more than 4-fold and increased the average mean speed 1.6-fold (Table 2.1). The distribution of mean track speeds was right skewed as seen in Figure 2.11A. When compared to the longer 157 bp DNA construct, the mean speed histogram for the 30 bp DNA construct has a broadened central peak as described by its decreased excess kurtosis of -0.55 and maintained a moderately right skewed distribution with a skewedness of 0.58 (Figure 2.11B). This reflects that the movement of FL HMGB1 on the 30 bp DNA was generally faster than on the 157 bp DNA. Together these data show that shortening the DNA greatly reduced the frequency of HMGB1 interacting with the DNA surface and increased its speed of movement.

2.2.5 HMGB1 movement on a dense surface of DNA involves ionic interactions

Given that HMGB1 showed movement across all the surfaces tested I was curious whether electrostatic interactions contributed to its movement. I hypothesized that high salt conditions

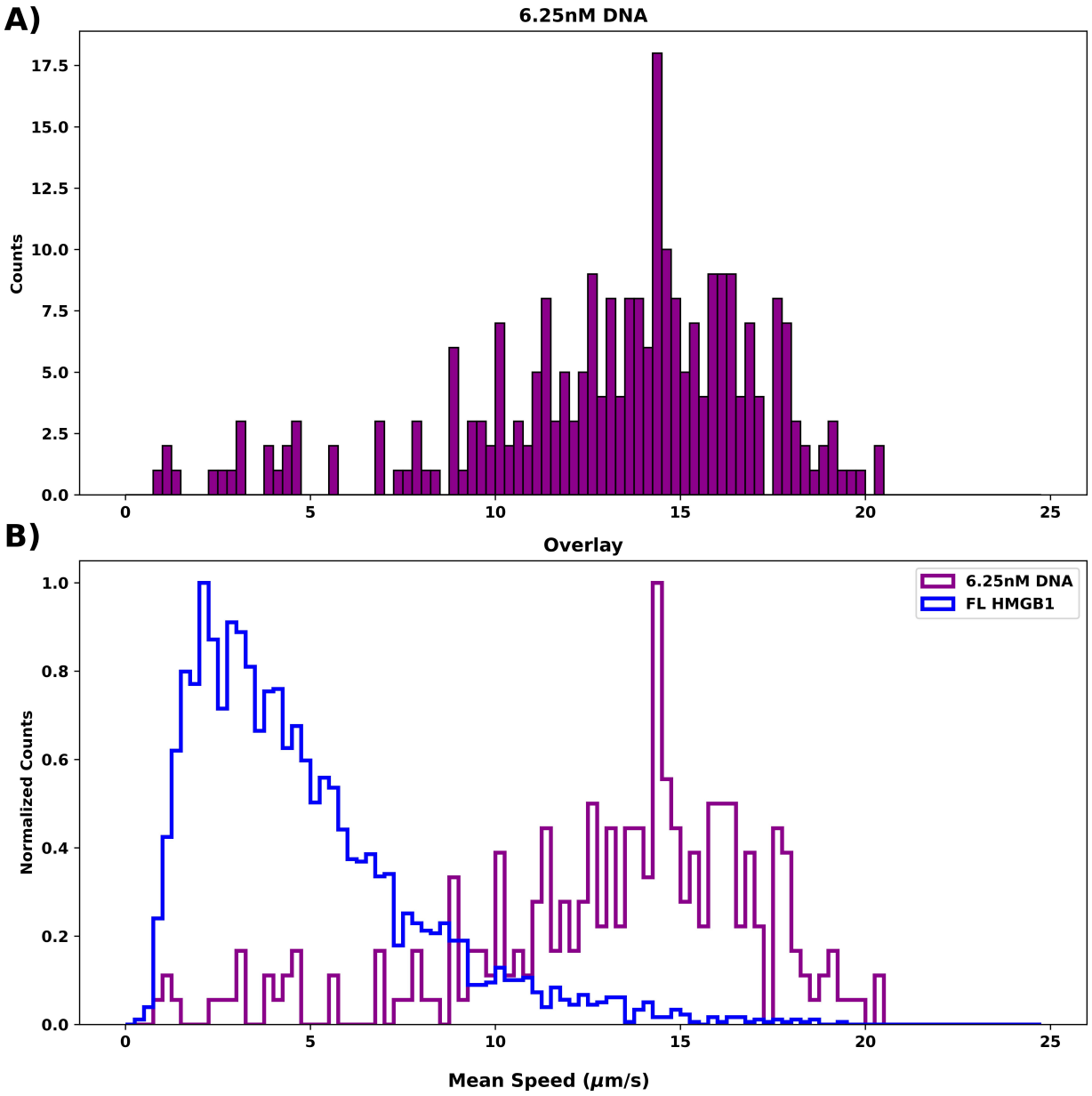

Figure 2.10: A) The mean speed distribution of 2 replicates of FL HMGB1 on 6.25 nM DNA produced a left skewed distribution. B) When compared to FL HMGB1, the mean speed distribution shifted to the right upon decreasing the surface density.

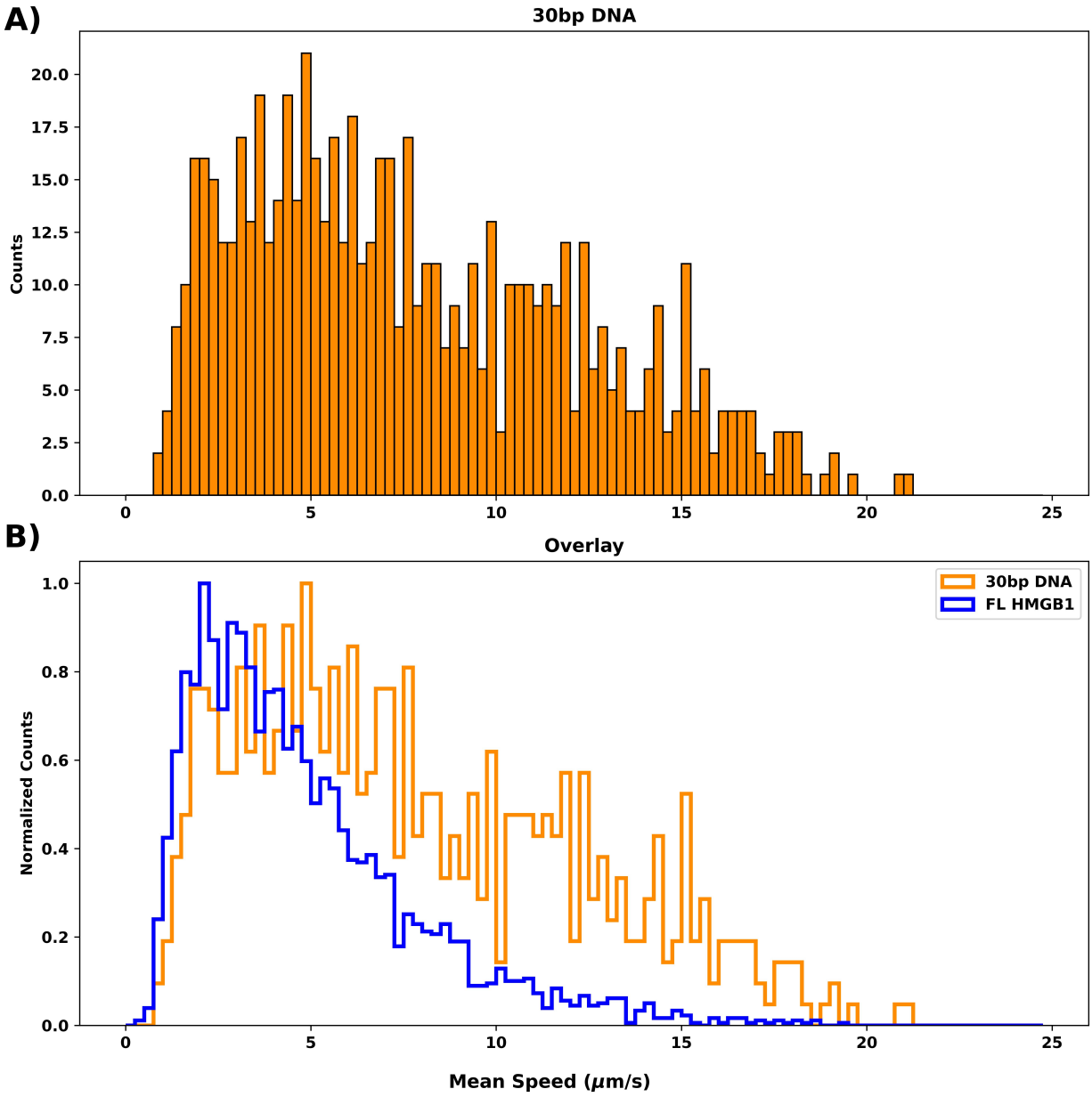

Figure 2.11: A) The combined track mean speed of 3 replicates of FL HMGB1 on a dense surface of 30 bp DNA produced a right skewed distribution. B) The shortening of the DNA broadened the track mean speed distribution for HMGB1.

would reduce HMGB1's movement. To test this, I densely immobilized the long 157 bp DNA and added FL HMGB1 under conditions where the KCl concentration was increased from 50 mM to 150 mM. Interestingly, increasing the salt concentration largely ablated the number of detectable tracks, producing on average only 8 tracks per region. The total number of tracks produced from this high salt condition had insufficient data to produce a histogram.

Observing that HMGB1 movement was ablated under high salt conditions suggested that its scanning behavior depends on charge-charge interactions. To further test this I asked whether HMGB1 would exhibit movement on dense surface of immobilized heparin, which is a negatively charged polysaccharide. Interestingly, FL HMGB1 showed movement behavior on a dense surface of heparin, albeit with far fewer tracks per region than with the dense 157 bp DNA surface (Table 2.1). On heparin, the mean speed distribution of HMGB1 movement showed a slightly broader right-skewed histogram when compared to its movement on the dense 157 bp DNA surface (Figure 2.12). This means that HMGB1 molecules generally moved faster on heparin than on DNA, also reflected by the increase in average and median of the mean track speeds. HMGB1 exhibiting movement behaviors on heparin suggests that electrostatic interactions are sufficient to drive its movement between negatively charged molecules; however, the 5-fold decrease in the average tracks per region suggests that DNA binding by HMGB1 controls its level of interaction with the surface.

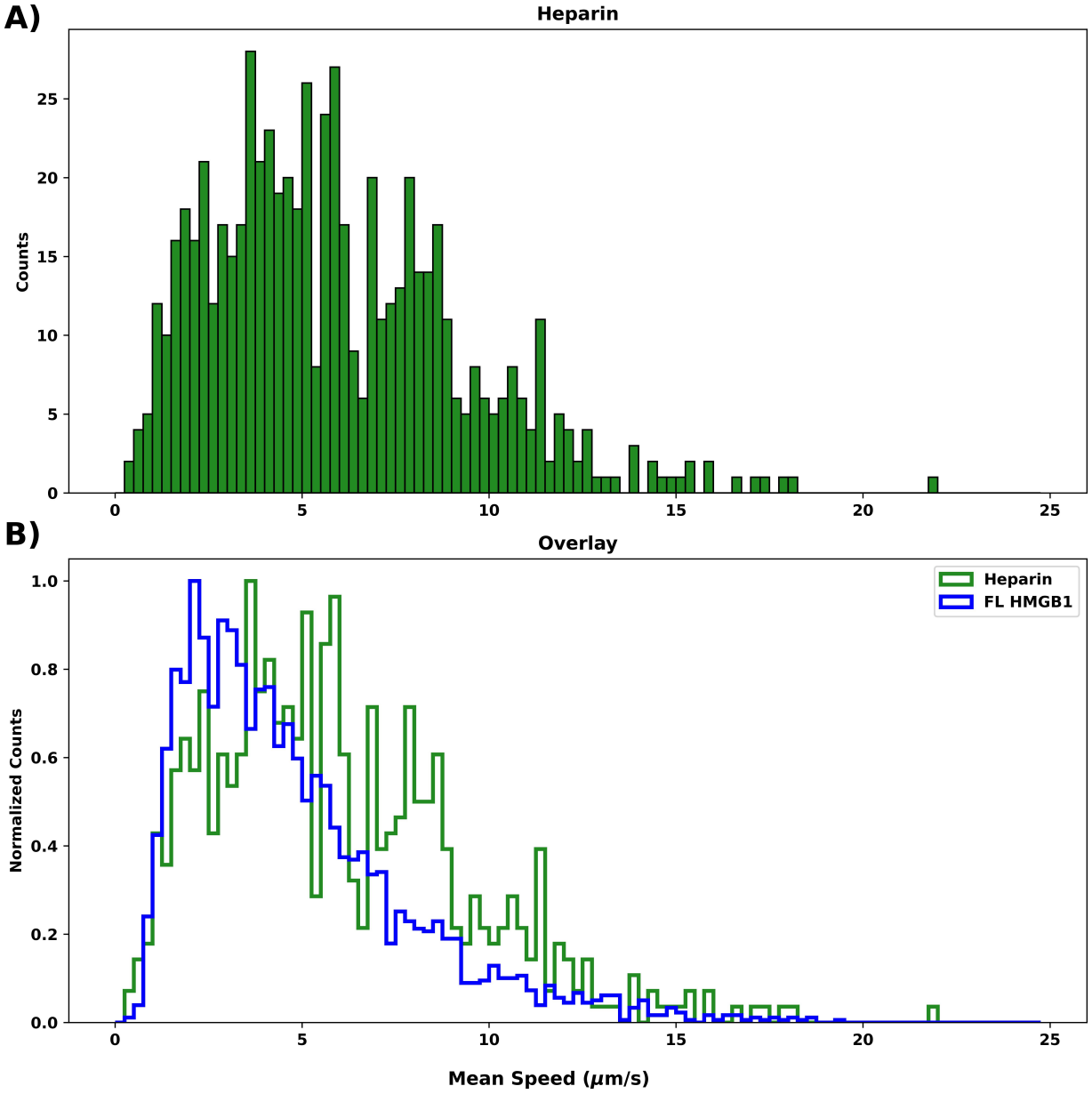

Figure 2.12: A) The track mean speeds from 3 replicates of FL HMGB1 on heparin were plotted, producing a right skewed distribution. B) FL HMGB1 showed increased speed when moving on a surface of heparin compared to the 157 bp DNA surface.

2.3 Discussion

HMGB1 is an architectural DNA binding protein that plays important roles in modulating numerous nuclear processes, such as transcription, DNA damage repair, telomere maintenance, and others [106]. Within the nucleus, HMGB1 is thought to bind and bend DNA to facilitate the binding of nuclear protein factors such as transcription factors to their cognate DNA binding sites. HMGB1 is known to weakly associate with interphase chromatin [115] and previous studies employing Fluorescence Recovery After Photobleaching (FRAP) have shown that HMGB1 is a highly dynamic protein capable of rapidly navigating the chromatin environment to traverse across the nucleus within 1-2 seconds [114]. However, it is unclear the mechanisms used by HMGB1 to navigate the densely packed nuclear environment. In this chapter, I used single molecule TIRF microscopy to investigate how HMGB1 moves on a dense surface of DNA. Specifically, I determined the roles played by individual HMGB1 domains, and how surface characteristics influence HMGB1's movement behavior.

Because HMGB1 consists of three domains – the A-box, the B-box, and the unstructured C-terminal tail (CT-tail) – we wondered how each of these influence HMGB1's movement on a dense surface of DNA. Removal of the CT-tail, yielding the AB construct, slightly reduced the speed of HMGB1's movement and increased the number of tracks per region. HMGB1's CT-tail dampens its DNA binding and bending capabilities by intramolecularly interacting with its DNA binding domains, particularly the B-box [36]. Removal of this domain alleviates the intramolecular interactions that transiently shield one of the DNA binding domains from its substrates. With both DNA binding domains free to bind and bend DNA, the frequency in which both DNA binding domains interact with a DNA substrate is increased, which could be the cause of the increased number of average tracks per region compared to FL HMGB1 (200 versus 133) and the slightly decreased movement speed.

The largest effect on the speed of movement was observed by further deleting each of the DNA binding domains. The A-box and B-box constructs moved substantially faster than either the

FL or AB constructs. A potential explanation for the increased speed observed with the individual HMGB1 DNA binding domains is that each of the single domain constructs can freely dissociate from a DNA substrate and is not constrained by the DNA binding of a second linked DNA binding domain. This could increase the frequency with which the single box proteins jump between DNA molecules, and result in a faster average track speed. Furthermore, movement of the A-box was slightly slower than the B-box.

Another question I asked was how does the surface DNA density influence HMGB1's movement. At an initial working concentration of 25 nM DNA, the surface density was calculated to be 40±5 molecules per pixel which translated to 31.7±2.3 nm of distance between neighboring DNA molecules. Decreasing the amount of DNA flowed onto the surface by half drastically reduced the number of moving molecules from 133 tracks per region to 20 tracks per region. The reduction in the number of moving events per region may be a result of the increased distance between the immobilized DNA, where the likelihood of HMGB1 moving between less closely spaced DNA molecules is decreased. At 12.5 nM DNA, the distance between individual DNA molecules likely increased from approximately 32 nm to 45 nm. The immobilized 157 bp DNA construct used in this experiment is approximately 54 nm in length. It may sway around a pivot point on the surface, thereby bringing portions of the DNA within proximity of a neighboring DNA for short periods of time. Increasing the distance between neighboring DNA molecules decreased the frequency and extent to which two DNA molecules could come into proximity, resulting in the observed decrease in the number of moving HMGB1 events. Further reducing the surface density should also decrease the number of moving events per region; however, upon reducing the concentration of DNA added to the slide chamber from 12.5 nM to 6.25 nM, the number of moving events marginally decreased from 20 to 17 tracks per region. It possible there is a critical DNA density for obtaining dose-dependent changes in movement by HMGB1 that is achieved at an added DNA concentration between 12.5 nM and 25 nM; and a density at which all HMGB1 movement is lost. Additional experiments could be performed to more clearly define the relationship between DNA density and the number of HMGB1 tracks per region. It could also be interesting to determine how increasing

the density beyond what was tested here affects the number of moving HMGB1 molecules. In addition to reducing the number of tracked molecules, decreasing the DNA density also increased the speed of FL HMGB1 movement. This is likely due to HMGB1 having to make larger jumps between DNA molecules. Thus, it appears that when the DNA surface density was decreased, very few HMGB1 molecules move between DNAs, but those that do appear to move, do so at a faster speed.

I assessed how HMGB1 behaves on a dense surface of a shorter 30 bp DNA. Upon shortening the DNA to 30 bp, the number of mobile events drastically decreased when compared to the 157 bp DNA construct. HMGB1 has a DNA binding footprint of approximately 18 bp [109]. When compared to the shorter DNA construct, the 157 bp DNA provides a many more binding sites for a single HMGB1 molecule, which would promote more interactions. Moreover, when considering DNA motion, the longer 157 bp DNA construct would be more effective at facilitating HMGB1 movement on the surface as the frequency of two DNA molecules being within proximity would be greater than that of the shorter 30 bp DNA construct. The shorter 30 bp DNA construct decreased the length of biotinylated-DNA from about 54 nm to about 11 nm. Under the extreme scenario where neighboring DNA molecules lay flat on the slide surface pointing at one another, there would still be a distance of about 10 nm of between the ends of two 30 bp DNA molecules. This would result in HMGB1 having to make larger jumps between proximal DNA molecules which could explain for the broadening of the speed distribution population. As a result, of the molecules of FL HMGB1 that were able to move on the shoter DNA, they generally moved faster, reflected by a relatively broad left skewed distribution of mean speeds.

Lastly, I asked whether HMGB1 movement was charge-charge dependent by increasing the salt concentration from 50 mM KCl to 150 mM KCl. Previous studies have suggested that increased salt concentration decreases HMGB1's affinity for DNA [109] as well decreases the interaction between its CT-tail and its DNA binding domains [33]. Thus, I hypothesized that the average number of tracks per region should decrease as salt increased. Indeed, at the higher salt concentration there was a very small number of mobile events, approaching that of the no DNA control. This shows

that HMGB1 movement on the surface likely requires electrostatic interactions. Because of this observation, I tested whether HMGB1 was able to move on a surface of heparin, a negatively charged polysaccharide. On a dense surface of heparin, HMGB1 produced about 5-fold less tracks per region when compared to a dense surface of 157 bp DNA. Additionally, HMGB1 move slightly faster on the heparin surface despite having a broader mean speed distribution when compared to the 157 bp DNA surface. These observations suggest that DNA binding is important for initially attracting HMGB1 to the surface, and that the speed of HMGB1 movement between substrates is largely facilitated by electrostatic interactions.

Overall, my results suggest that the two HMGB1 DNA binding domains work together to modulate its movement speed on a dense surface of DNA. Changing the surface characteristics, such as decreasing the surface density of DNA drastically reduced the number of tracked events per region and increased HMGB1's movement speed, highlighting the importance of having a dense DNA environment to facilitate HMGB1's movement between DNA substrates. Shortening of the DNA also impacted HMGB1's movement behavior by drastically reducing the number of tracked events per region and broadening its mean speed distribution. This suggests that DNA functions as a platform that facilitates both HMGB1's ability to bind and to move between DNA molecules. Lastly, HMGB1's recruitment to the surface and its movement speed are dependent on both DNA binding and electrostatic interactions, as the number of tracked events per regions drastically decreased on a surface of heparin while eliciting a similar movement speed when compared to a DNA surface. This body of work highlights the dynamic nature of HMGB1 while dissecting the factors the facilitate HMGB1 movement.

2.4 Materials and Methods

2.4.1 DNA constructs

The 157 bp DNA construct was generated by performing a PCR reaction on a pUC19 vector containing the Widom601 sequence using the following primers (Integrated DNA Technologies):

Forward 5' – Biotin – C – Cy3 – TG ATA TCT GAG AAT – 3 and reverse 5' – ATT CTC AGA TAT CAG – 3'. The unlabeled DNA construct was generated using the same primers sequences listed above, but without the Cy3 modification within the forward primer. The PCR products were then purified by using Q Sepharose Fast Flow resin (Cytiva) packed into a gravity follow column. The Q Sepharose resin was equilibrated with 10x resin volume of TE buffer [20 mM Tris (pH 7.9) and 1 mM EDTA (pH 8)]. The DNA was then passed through the column, then washed with 10x resin volume TE buffer containing 600mM KCl. The DNA was eluted from the column with TE buffer containing 800 mM KCl. The DNA fractions were then assessed on a 10% non-denaturing polyacrylamide gel run at 150V in 0.2x TBE. The fractions containing the most 157 bp DNA were pooled, and the salt concentration was reduced 1000-fold by buffer exchanging the sample with MilliQ water using Pierce 10K MWCO protein concentrators (Thermo Fisher Scientific). The DNA concentration was then calculated by measuring the A260 utilizing the Beer-Lambert law. The DNA was then stored a -20^{o}C until ready to use.

The 30 bp DNA construct was generated by annealing two primers (Integrated DNA technologies). The Forward primer [5'- Biotin- CTG ATA TCT GAG AAT CCG GTG CCG AGG CCG – 3'] and reverse primer [5'- CGG CCT CGG CAC CGG ATT CTC AGA TAT CAG – 3'] were mixed in a 1 M forward primer to 4 M reverse primer ratio in annealing buffer [20 mM Tris pH 7.9, 2 mM $MgCl_2$, 50 mM KCl]. Samples were annealed by heating in a heating block at 95^{o}C for 10 minutes. The heating block was then removed from the heating element and the DNA was allowed to slowly cool to room temperature. The annealed DNAs were then run on a 1.5% agarose gel at 150V in 1x TBE and stained with ethidium bromide. The DNA bands were then cut out of the gel, flash frozen in liquid nitrogen, and isolated using Spin-X columns (Fisher Scientific). Afterwards, the DNA was ethanol precipitated, the concentration calculated from its A260, and the purified sample was stored at -20^{o}C until ready to use.

2.4.2 Protein expression, purification, and fluorescence labeling

A previously engineered pET19b vector containing the sequence for an N-terminal 6x His-tagged human FL HMGB1 (C23S, C45S, C106S) was used, as well as constructs for the AB protein, A-box, and B-box that also had the Cys to Ser mutations [36]. A single cysteine was engineered into the constructs for the FL protein, AB, and A-box using PCR to change Serine-35 to Cysteine-35 (S35C) to allow for protein-labeling via maleimide conjugation. The B-box construct had a single cysteine engineered via PCR to change Alanine-137 to Cysteine-137(A137C). The expression vectors from FL HMGB1 and the various protein truncation constructs were transformed into *Escherichia coli* BL-21 DE3 pLysS cells and were grown overnight at 37^{o}C on LB agar plates containing 100 μg/mL ampicillin. Afterwards, for each construct, a single colony from the LB agar plate was picked to inoculate a starter culture containing 5 mL of LB containing 2 g/L D-glucose and 100 μg/mL ampicillin. The starter culture was grown overnight in a shaking incubator set to a speed of 200 rpm and at a temperature of 37^{o}C. The starter culture was then transferred to a 500 mL flask of LB containing 2 g/L D-glucose and 100 μg/mL ampicillin and was grown to an OD_{600} of 0.5 – 0.6 at the same settings mentioned above. Protein expression was then induced by adding 0.5 mM IPTG into the culture for 2 hours. Afterwards, the cells were pelleted by centrifugation at a speed of 2,400 x g at 4^{o}C for 30 minutes. Once the cells were pelleted, the supernatant was decanted and discarded.

The cell pellets were then resuspended in 5 mL buffer Ni-A [20 mM Tris (pH 7.9), 5 mM β-mercaptoethanol, 500 mM NaCl, 10% glycerol, 1x EDTA-free protease inhibitors (Roche), 0.2 mM PMSF, and 10 mM imidazole], transferred to a 15 mL conical tube, and sonicated (5 times, 30s on and 2 min off) on ice. The lysed cells were then centrifuged at 26,000 x g using a F21 rotor for 15 minutes at 4^{o}C to pellet the cell debris. A 1.5 cm diameter by 15 cm length gravity flow column (Biorad) was packed with a 1.3 mL slurry of the HisPur Ni-NTA resin (Thermo scientific) that was equilibrated with buffer Ni-A at 4^{o}C. The supernatant from the centrifugation step was passed through the column and the flow through was collected. The column was washed with 10x

column volume of buffer Ni-A followed by 10x column volume of buffer Ni-A with 50 mM imidazole. Afterwards, the protein was eluted from the column with 10x column volume of buffer Ni-A with 250 mM imidazole and was collected in 700 μL fractions. The elution fractions were analyzed by 12% SDS-PAGE and the fractions containing the most HMGB1 protein were pooled together and dialyzed for 2 hours at 4^{o}C in 20 mM Tris (pH 7.9), 50 mM KCl, 10% glycerol, 0.2 mM PMSF, 5 mM $MgCl_2$, and 1 mM DTT.

After dialysis, the HMGB1 proteins were further purified on a dsDNA cellulose column. 250 mg of calf thymus dsDNA cellulose (Sigma) was equilibrated in buffer ds-A [20 mM Tris (pH 7.9), 50 mM KCl, 10% glycerol, 5 mM $MgCl_2$, 1 mM DTT] then added to a gravity flow column and was washed with 7.5 mL of buffer ds-A. The dialyzed eluate from the Ni column was centrifuged at 16,100 x g for 15 minutes at 4^{o}C to pellet any precipitants that may have formed during dialysis. The supernatant was then passed through the dsDNA column and washed with 3 mL of buffer ds-A. The protein was then eluted off the column with buffer ds-A containing 500 mM KCl, and collected in 300 μL fractions. The eluted fractions were analyzed by 12% SDS-PAGE and the fractions with the most HMGB1 protein were pooled and dialyzed overnight against buffer ds-B [20 mM Tris (pH 7.9), 50 mM KCl, 10% glycerol, 0.2 mM PMSF, and 5 mM $MgCl_2$]. The dialyzed fractions were centrifuged at 16,100 x g for 15 minutes at 4^{o}C to pellet any protein precipitants present. The concentration of the protein was then assessed via 12% SDS-PAGE using known concentrations of BSA to create a standard curve.

The proteins were then reduced by adding a 10-fold molar excess of TCEP (pH 7) and the sample was nutated for 30 minutes at 4^{o}C. A 2-fold molar excess of Alexa Fluor 647 (AF647) C2 maleimide dye dissolved in DMSO (Fisher Scientific) was then added to the reduced protein sample and was nutated for 1 hour at 4^{o}C. The excess dye was then removed using the Pierce dye removal columns (Thermo Scientific). Afterwards, the protein concentration and percent labeling were determined via 12% SDS PAGE using standard curves of BSA and a AF647 labeled primer (Integrated DNA Technologies). The proteins were then aliquoted, snap frozen in liquid nitrogen, and stored at -80^{o}C until ready to use.

2.4.3 Single molecule tracking experiments

The functionalization of glass slides, coverslips, and the preparations of the single molecule imaging reagents (e.g. the oxygen scavenging system and Trolox to minimize photoblinking) are described elsewhere [119]. The slide chambers were first washed twice with MilliQ water and twice with RB buffer [10% glycerol, 25 mM Tris (pH 7.9), 50 mM KCl, 5 mM $MgCl_2$, 0.05 mg/mL BSA, 1 mM DTT, and 0.1% NP-40]. Afterwards, a streptavidin mixture [1 mg/mL streptavidin, 0.8 mg/mL BSA in RB buffer] was flowed into the sample chamber and was incubated for 5 minutes. The slide chamber was then washed twice with RB buffer. Biotinylated DNA (25 nM unlabeled, plus 10 pM Cy3-labeled) was then flowed into the slide chamber and was allowed to incubate for 5 minutes to enable DNA immobilization. For the experiment with heparin, 25 nM of a 27kDa biotinylated heparin (Creative PEGWorks) was used in place of DNA. The slide chamber was then washed twice with RB buffer. Cy3-DNA movies were then collected after flowing in imaging buffer [2 mM Trolox, 0.8% (w/v) D-glucose, 1 mg/mL glucose oxidase, and 11 μg/mL catalase in RB buffer]. The slide chamber was then washed twice with RB buffer. 25 pM AF647 HMGB1, prepared in imaging buffer with 1 mg/mL BSA, was flown into the slide chamber and movies were collected, as described below.

2.4.4 Single molecule tracking data collection and analysis

The single molecule tracking data was collected using a Nikon TE2000-U microscope equipped with a 1.49 numerical aperture immersion total internal reflection fluorescence objective. All movie captures were done using the NIS-Elements (Nikon) software. The Cy3-DNA movies were collected using a 532 nm laser with an exposure rate of 60 ms for 10 frames. The AF647-HMGB1 movies were collected using a 640 nm laser with an exposure rate of 60 ms for 1000 frames. Both laser lines each used an Andor iXon Life EMCCD camera. The movies were summed every 2 frames between every other frame to improve the signal of the single molcules and were then analyzed using TrackMate [117], an ImageJ plugin, to identify and track spots over subsequent frames. The spots

were detected by thresholding for visible spots using the Difference of Gaussian (DoG) detector and the tracks were linked using the simple Linear Assignment Problem (LAP). The data was then filtered to exclude tracks that persisted on the surface for less than 5 frames (< 300 ms) and traveled a maximum distance less than 15 pixels from the starting pixel. Afterwards, the mean speed for each track was recorded. The mean speeds of each track were binned and plotted as a histogram for each experimental condition, combining tracks obtained from biological replicates. The statistical parameters shown in Table 2.1 were calculated from the mean speeds of each track across all replicates.

Chapter 3

Investigating HMGB1 on/off binding dynamics on nucleosomes and DNA

The tight wrapping of DNA around the positively charged surface of the histone octamer serves as a physical barrier to transcription factors that need access to their cognate DNA binding sites, and therefore serves as a point of gene expression regulation. The accessibility of DNA to transcription factors can be modulated by numerous protein that can alter the chromatin environment to impact both its organization as well as the composition of the nucleosome particle, the fundamental subunits of chromatin. Some of these nuclear protein factors include chromatin remodelers, pioneer factors, and architectural factors. Among this list is High Mobility Group Box 1 (HMGB1) - an architectural protein factor that binds DNA without sequence specificity. Ensemble biochemical assays have shown that HMGB1 binds nucleosomal linker DNA near its entry/exit junction and interact with the histone H3 N-terminal tail peptide to help position HMGB1 onto the nucleosome. However, it remains unclear what the underlying mechanisms are that drive HMGB1 to interact with the nucleosome particle. In this chapter, I explored the association/dissociation kinetics of HMGB1 interacting with nucleosomes and free DNA sparsely immobilized on slide surfaces using single molecule Total Internal Reflection Fluorescence (TIRF) microscopy. However, there were challenges in reproducing the kinetic measurements between experimental replicates, making it difficult to draw conclusions about HMGB1 behavior with confidence.

3.1 Introduction

HMGB1 plays an important role in facilitating transcription factor binding to their cognate DNA binding sites to stimulate transcriptional activation [106]. HMGB1 can accomplish this by interacting with chromatin through binding and bending DNA to distort the local chromatin environment [43]. This is believed to create pockets of DNA accessible for transcription factor binding within chromatin. Within the nucleus, HMGB1 can modulate the local chromatin environment to facilitate the accessibility of DNA to various protein factors to control transcriptional activation [38, 39]. For example, previous studies have shown that HMGB1 facilitates the binding of p53 [47], estrogen receptor [48], and numerous other transcription factors [49] to their cognate DNA binding sites. In addition, HMGB1 facilitates DNA damage and repair [46], and other processes that happen on chromatin. Many of these processes are driven by HMGB1s ability to interact with the nucleosome particles.

The packing of DNA into chromatin serves both to protect the DNA from damaging elements within the nucleus, such as nucleases, as well as control gene expression by modulating the accessibility of its DNA to nuclear protein factors. Chromatin compaction and DNA accessibility can be modulated by a number of protein factors, such as chromatin remodelers [120], histone modifiers [121], pioneer factors [122], and architectural factors [123]. At the molecular level chromatin is composed of a series of individual nucleosome particles where the DNA is wrapped around a protein octamer complex consisting of two H2A/H2B dimers and two H3/H4 dimers. A high resolution 2.8 Å crystal structure of the nucleosome core particle showed that the DNA wrapped around the histone octamer with 1.65 super helical turns and was organized as a left-handed DNA helix [124]. This high-resolution structure revealed that the histone octamers make a series of contact with the negatively charged DNA phosphate backbone to bend the DNA and form nucleosome particles. Several DNA sequences have been discovered and/or engineered for nucleosome reconstitution *in vitro*, such as the human α satellite centromeric repeat [125], the 5s RNA coding sequence [126], and the Widom601 positioning sequence [127]. Particularly, the SELEX engineered Widom 601

positioning sequence has been widely used due to its high affinity for histone proteins [127] and improved nucleosome stability when compared other DNA constructs such as the human α satellite sequence [128].

Biochemical studies have shown that HMGB1 interacts with nucleosomes by binding the entry/exit DNA junction [40] and also bind histone H3's N-terminal tail via interaction with HMGB1's C-terminal tail [38, 44]. At the DNA junction, HMGB1 competes with linker histone H1 for binding nucleosomes at the linker DNA [9, 129]. The linker histone H1 plays an important role in the condensation of chromatin [130], whereas HMGB1 functions to destabilize the local chromatin environment to facilitate nuclear processes such as transcription activation [15, 129]. Histone H1 binds nucleosomes at the dyad axis, where it interacts with the linker DNA and stabilize the nucleosome particle using its basic N-terminal and C-terminal tails [131–133]. HMGB1 uses its acidic C-terminal tail to interact with histone H1's basic N-terminal tail [41], which could potentially function as a mechanism, in addition to competing for DNA binding, for how HMGB1 displaces histone H1 from the linker DNA. However, it is unclear from the current body of literature the interplay between HMGB1 interacting with the nucleosome to facilitate the binding of nearby protein factors, and interacting with histone H1 to modulate the local chromatin environment.

In this chapter, I explored HMGB1 on/off binding dynamics on nucleosomes reconstituted with the Widom601 sequence using single molecule Total Internal Reflection Fluorescence Microscopy (TIRF-M). However, I experienced difficulties in being able to reproduce the kinetics of HMGB1's on/off binding behaviors between experimental replicates. Therefore, I simplified the system to use free DNA rather than Nucleosomes. This, however, did not resolve the reproducibility issue as the binding behaviors and kinetics remained inconsistent between experimental replicates. With the goal of further simplifying the system to eliminate issues of reproducibility, I shortened the DNA only system from 157 bp to 32 bp and explored HMGB1 binding kinetics, only to once again have trouble with reproducibility.

3.2 Results

3.2.1 Designing a DNA construct for nucleosome assembly and TIRF-M

To assess how HMGB1 interacts with nucleosmes, I designed a minimal 157bp DNA construct containing the Widom 601 positioning sequence [127] with a 5bp DNA linker on each end (Figure 3.1). The rationale for this design was two-fold: (1) To determine how HMGB1 behaves on mononucleosomes with a short DNA overhang, and (2) to minimize the binding of HMGB1 to the linker near the surface. Point (2) was important because longer DNA linkers could accommodate one or more molecules of HMGB1 between the surface and the nucleosome. To immobilize the DNA construct to the surface of the slide and to visualize the individual DNA molecules, a single 5'-end of the DNA construct was conjugated with a biotin molecule and a Cy3 dye. The positioning of the Cy3 dye proximal to the biotin molecule greatly improved the fluorescence signal to noise, allowing for clear identification of individual DNA molecules via TIRF-M.

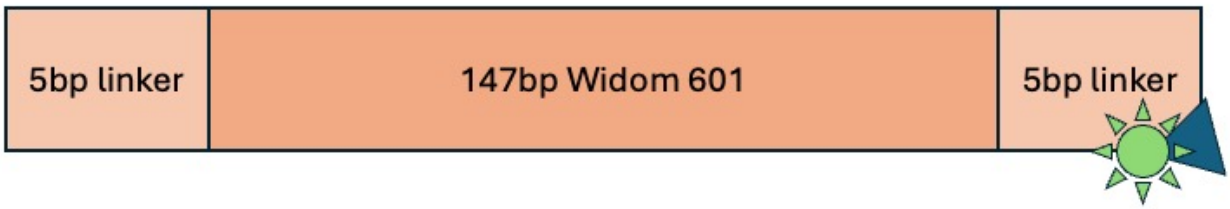

Figure 3.1: A schematic of the DNA construct containing the Widom601 sequence labeled with a Cy3 dye (green sun) and a biotin molecule (blue triangle) on a 5'-end.

Upon purifying the DNA construct, I then assessed the appropriate ratio of DNA to histone octamer ratio that formed nucleosomes while minimizing the free DNA and the formation of hexasomes and free DNA using a stepwise salt dilution protocol [134, 135]. The microscale nucleosome assembly reactions consisted of fixed DNA concentrations and increasing amounts of human histone octamers. To facilitate the assembly of nucleosomes, the samples were diluted in a step-wise manner from a high salt condition (2 M KCl) to a lower salt condition (50 mM KCl). The nucleosome

assembly reactions were assessed by native gel electrophoresis, as shown in Figure 3.2. Here, the lowest concentration of histone octamers (lane 1) produced a nucleosome, hexasome, and a DNA band, indicating that the concentration of histone octamer was not sufficient to drive complete nucleosome formation. The formation of hexasomes was likely due to the lack of integration of the histone H2A/H2B dimer into the nucleosome, which can occur at lower octamer-DNA ratios or higher salt concentrations [136]. As I increased the concentration of histone octamers (Lane 2 and 3), the DNA and hexasome band disappears to leave only the nucleosome band. Continuing to titrate more histone octamer into the reaction decreases nucleosome band intensity (Lane 3). This decrease in nucleosome band intensity is a result of excess of histone octamers sequestering DNA into higher-order complexes instead of assembling into a single nucleosome particle.

3.2.2 A single molecule system to study HMGB1 on/off binding dynamics on nucleosomes

To study HMGB1's on/off binding dynamics on nucleosomes, I used a previously developed single molecule TIRF-M system and analysis pipeline with the central goal of quantifying the kinetics with which HMGB1 interacts with nucleosome particles. To do this, I sparsely immobilized Cy3-nucleosome molecules to the surface of microscopy slides functionalied with Biotin-PEG utilizing a biotin-streptavidin-biotin linkage as represented in Figure 3.3. It was important to keep the surface to 1000 spots or less per imaging area as a higher surface density decreased the ability to resolve individual molecules due to overlapping signals from other fluorescent molecules that may be within proximity. A representative image showing the characteristics of a sparse surface of nucleosomes can be seen in Figure 3.4.

Once the nucleosomes were immobilized, the sample was excited with a 532 nm laser line under TIRF conditions. Emission from the Cy3-nucleosomes was then collected as reference movies over four regions using a piezo stage. AF647-HMGB1 was added into the flow chamber and the sample was excited with a 640 nm laser line under TIRF conditions, and the emission of AF647-HMGB1 was collected over the same four regions marked by the pieozo stage. Aftewards, in-house

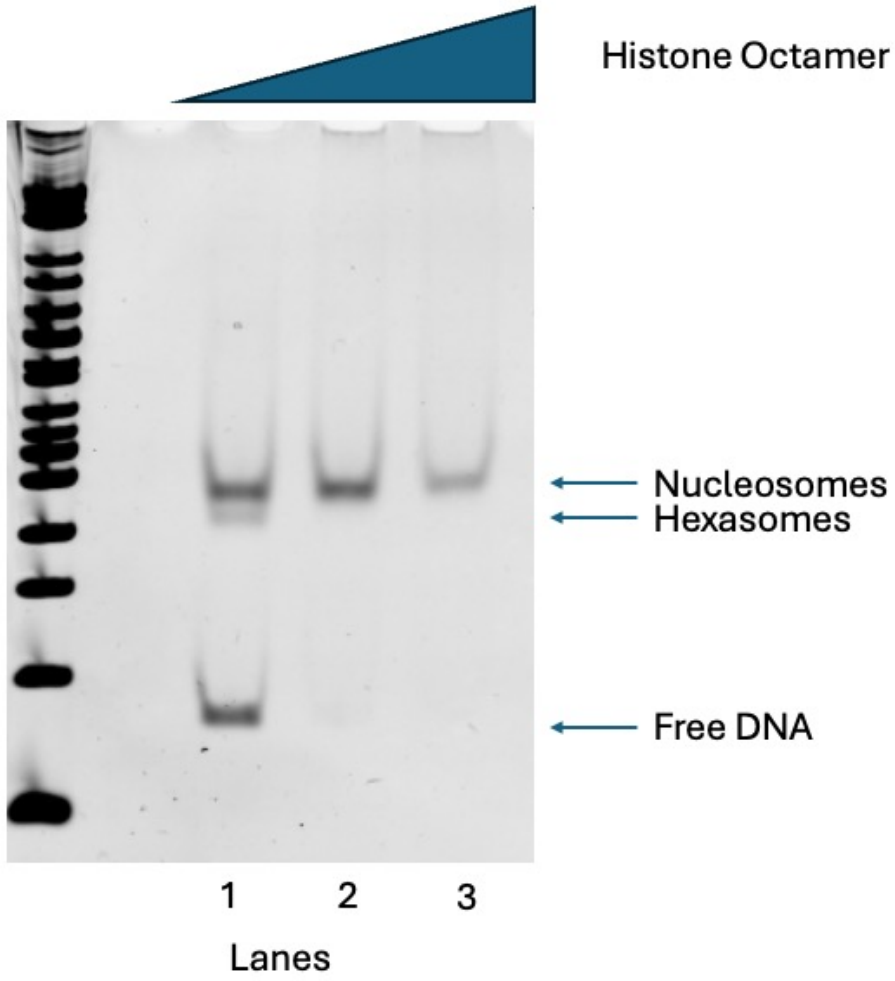

Figure 3.2: A representative nucleosome assembly reaction showing the presence of free DNA, hexasomes, and nucleosomes. The DNA Concentrations were kept constant at 2.7 μM and the histone octamer concentrations were increased by 0.27 Molar increment from 4.66 μM to 5.21 μM for each lanes seen above.

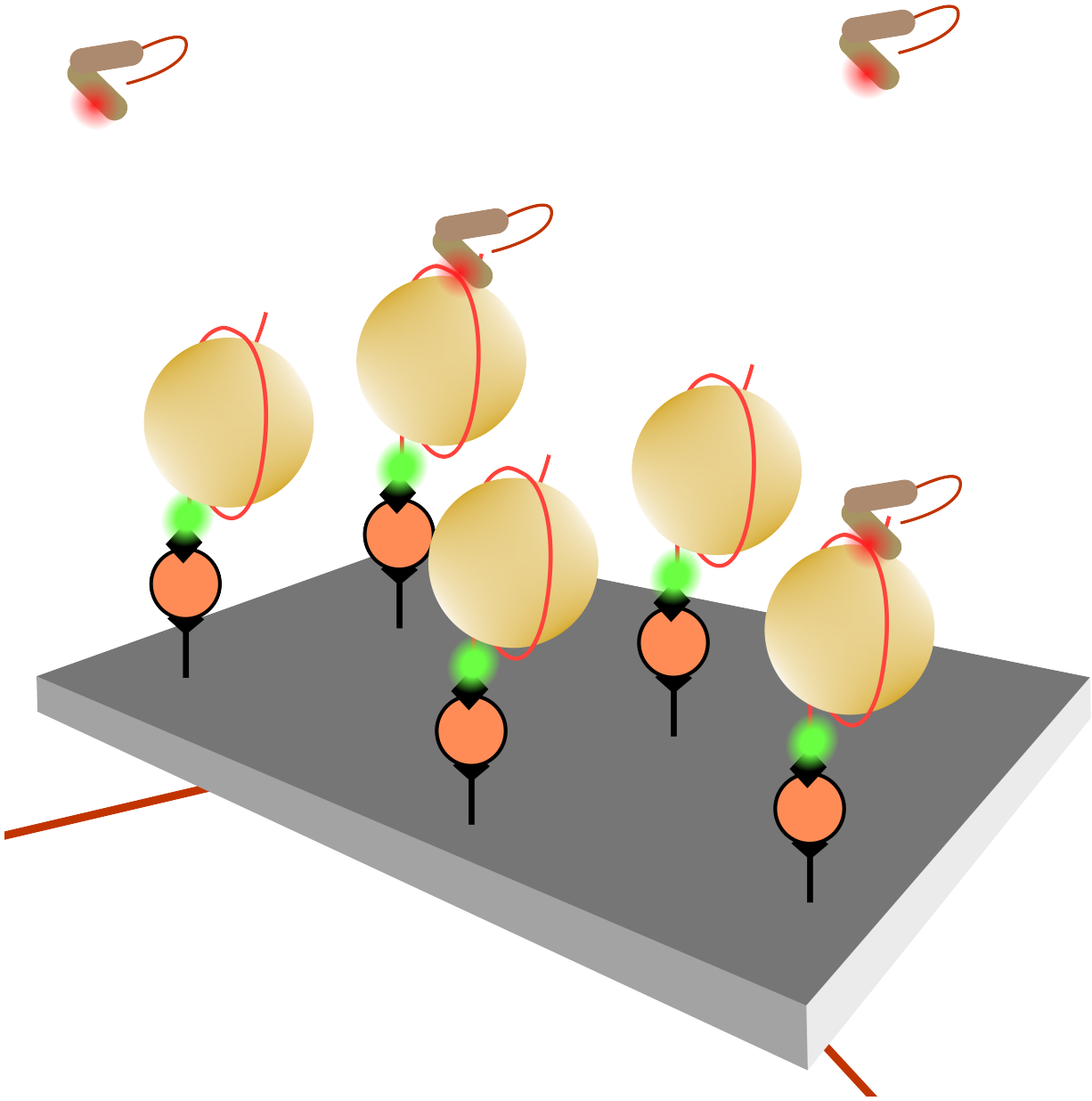

Figure 3.3: A schematic showing HMGB1 (brown with red dot representing AF647) interacting with nucleosomes (yellow spheres with green dot representing Cy3) that are immobilized by a biotin-streptavidin-biotin (orange circle complex) linkage on microscopy slides functionalized with PEG-biotin.

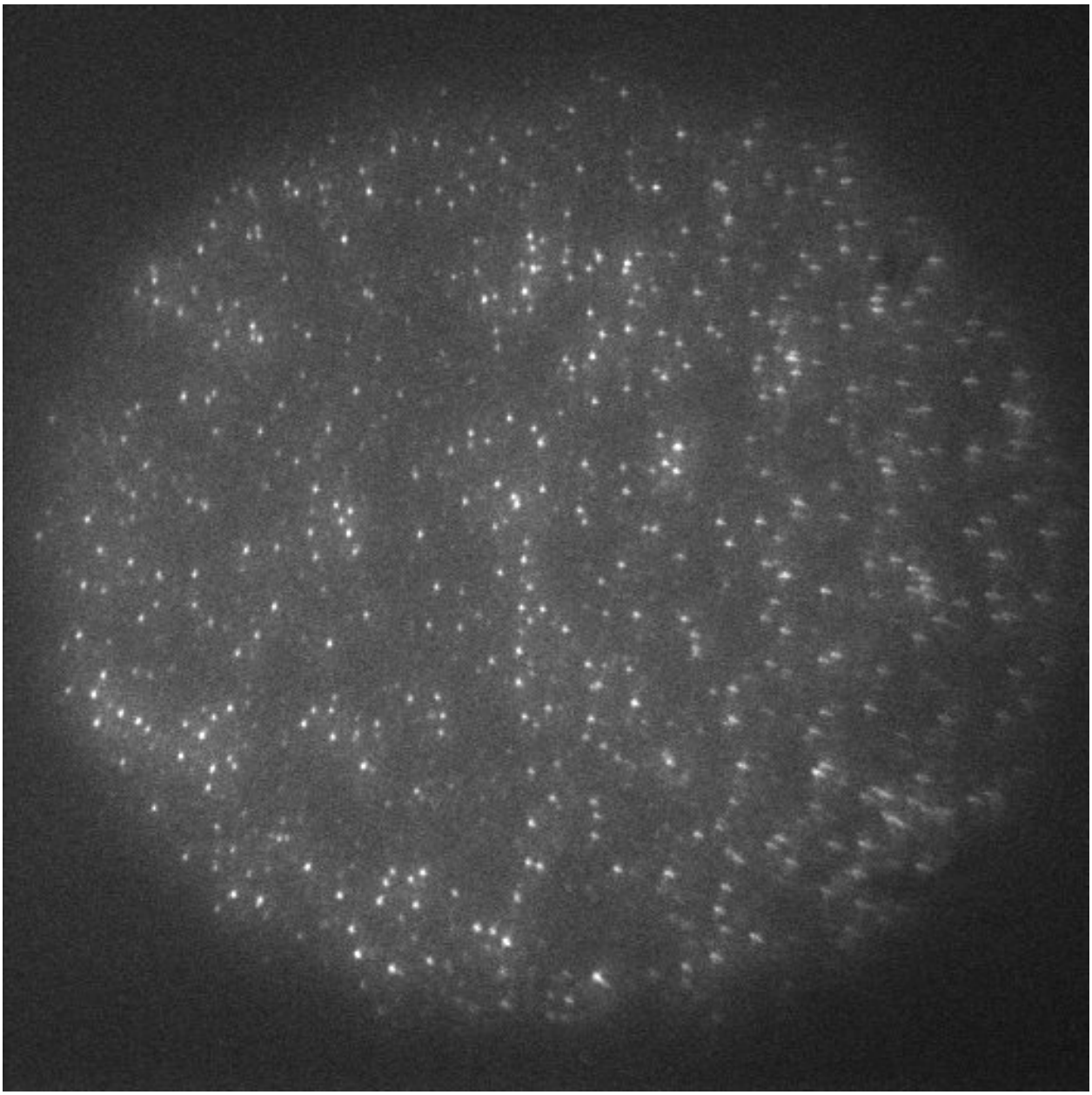

Figure 3.4: A representative frame from a movie of sparsely immobilized Cy3-nucleosomes

single molecule analysis software was then used to identify and record the coordinates of all visible Cy3-nucleosomes in the reference movies. After having identified the locations of the nucleosomes, the coordinates of individual AF647-HMGB1 molecules were identifeid and colocalized with individual Cy3-nucleosome molecules based on matching coordinate pairs. The emission intensity traces for each colocalized pairs was then extracted and plotted as a time series, where the intensity values were plotted on the y-axis and time on the x-axis (Figure 3.5).

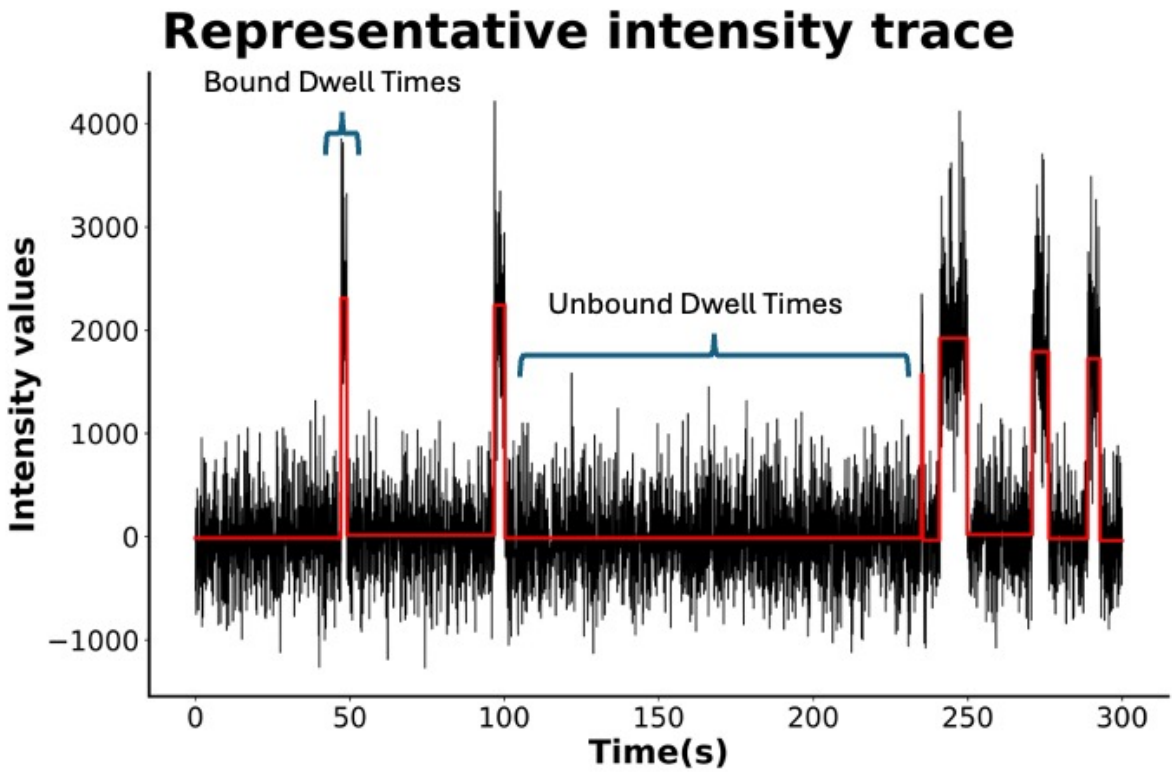

Figure 3.5: The fluorescence intensity of single AF647-HMGB1 molecules binding and releasing from nucleosome was plotted as a function of time (black line). The time series was fitted (red line) to identify the inflection points as binding and unbinding events.

Afterwards, the binding kinetics of HMGB1 were quantified based on the time series data. To calculate the binding kinetics, the time series for each spot pair was fitted (red line as shown in Figure 3.5) to identify the inflection points in the data set. A time point and a change in intensity value were associated with each of these inflection points to allow assignment and duration length

of a binding (positive change) or dissociation (negative change) event. Using this information a bound dwell time, characterized by a binding event followed by a dissociation event, was calculated. Conversely, an unbound dwell time, characterized by a dissociation event followed by a binding event, was calculated. Upon collecting hundreds of bound and unbound dwell times, the rate constants and 95% confidence intervals were determined by plotting the data as a cumulative sums and fitting with an exponential function using non-linear least squares regression analysis (Figure 3.6).

I first attempted to fit the cumulative sums plot with a single exponential function; however, a single exponential function was a poor fit for both the bound and unbound dwell times. A double exponential function was a better fit for both datasets, suggesting that HMGB1 exists as two kinetic population in both the rate of association and dissociation. The two rate constants for dissociation of HMGB1 from nucleosomes, which are indepedent of protein concentration, were k_{off1} of 0.20 s^{-1} and a k_{off2} of 0.71 s^{-1} (Figure 3.6, upper panel). HMGB1's observed rate constants for association on nucleosomes, which would be affected by protein concentration, were $k_{on1(obs)}$ of 0.0431 s^{-1} and a $k_{on2(obs)}$ of 0.59 s^{-1}(Figure 3.6, lower panel). In considering the influence of the protein concentration on the rate of association, the calculated apparent rate constants for association were determined to be 4x10^7 $M^{-1}s^{-1}$ and 5.9x10^8 $M^{-1}s^{-1}$ at 1 nM of HMGB1.

3.2.3 Inconsistencies in on/off binding kinetics between experimental replicates for HMGB1 on a sparse surface of nucleosomes

I collected data from multiple expeirments with HMGB1 and immobilized nucleosomes that were performed on different days, but in the same manner. When plotting data from the individual experimental replicates (n = 6) as a normalized cumulative sums, the bound dwell time data were somewhat reproducible (Figure 3.7, top panel); whereas the unbound dwell time replicates were inconsistent with one another (Figure 3.7, bottom panel). Prior to the 1 second time point, the data closely overlapped with one another (Figure 3.7, top panel, inset). However, at longer time points, some of the replciates deviated from the group. Due to the inconsistences between

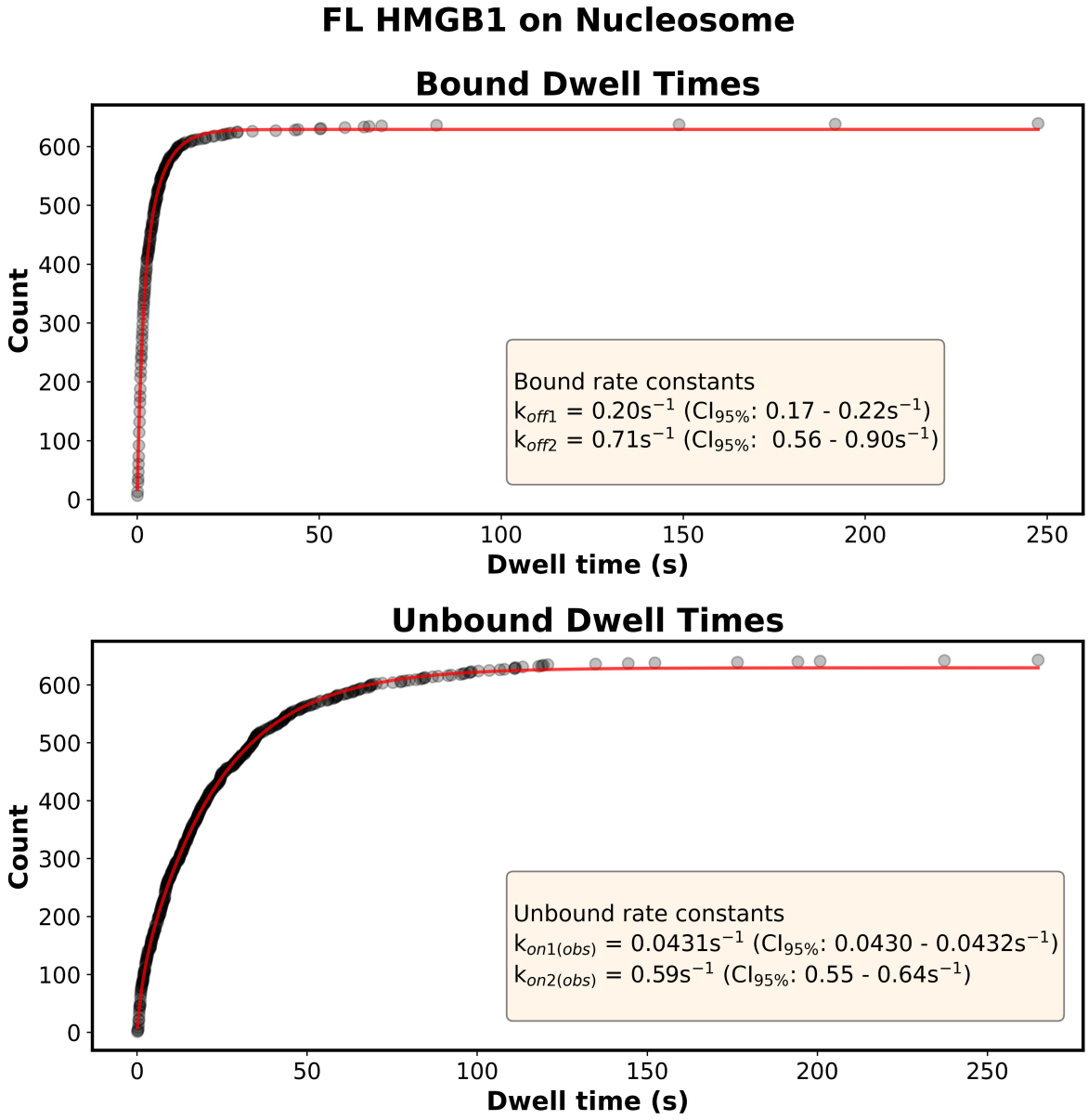

Figure 3.6: The combined bound (top) and unbound (bottom) dwell times from 6 experimental replicates of HMGB1 on a sparse surface of nucleosomes plotted as a cumulative sum (black circles) and fit with a double exponential function(red line). The rate constants and 95% confidence intervals of the curve fit are shown.

replicates for both the bound and unbound dwell times, it was difficult to accurately assess the rate constants of association and dissociation with confidence. When comparing the data across 6 replicates (Table 3.1), the apparent rate constants of association exhibited a broad range ($k_{on1(obs)}$: 0.002 - 0.057 s^{-1}, $k_{on2(obs)}$: 0.023 - 1.778 s^{-1}). This broad distribution persisted even for the rate constants of dissociation (k_{off1}: 0.085 - 0.367 s^{-1}, k_{off2}: 0.367 - 1.493 s^{-1}). While most of the replicates were best fit with a double exponential function, there was a single instance in which the observed rate of associate for replicate 3 was better fit with a single exponential function, further adding to the matrix of inconsistent results. It was unclear why the measured rate constants were inconsistent between experimental replicates, which warrants an investigation into why it was difficult to reproduce measurements of HMGB1 on/off binding dynamics on nucleosomes.

Replicate	$k_{on1}(s^{-1})$	$k_{on2}(s^{-1})$	$k_{off1}(s^{-1})$	$k_{off2}(s^{-1})$
Combined data	0.0431	0.59	0.20	0.71
1	0.057	0.290	0.129	0.594
2	0.033	1.215	0.211	0.506
3	0.023	-	0.230	1.238
4	0.002	1.267	0.367	0.367
5	0.034	0.595	0.192	1.493
6	0.045	1.778	0.085	0.839

Table 3.1: Most of the rate constants across 6 experimental replicates of FL HMGB1 on nucleosomes were best fit with a double exponential function, except for the observed rate of association for replicate 3 where it was better fit with a single exponential function.

3.2.4 HMGB1 on a sparse surface of 157bp DNA also exhibits inconsistencies between experimental replicates

To better understand why it was difficult to reproduce HMGB1 on/off binding on nucleosomes, I simplified the system by using the DNA construct used to assemble nucleosomes (i.e. no histone octamer, only free DNA). Under the same experimental conditions mentioned above, I investigated HMGB1 binding dynamics on a sparse surface of free 157 bp DNA (Figure 3.8). Four replicates of data were collected, combined, and was plotted as a cumulative sum that was

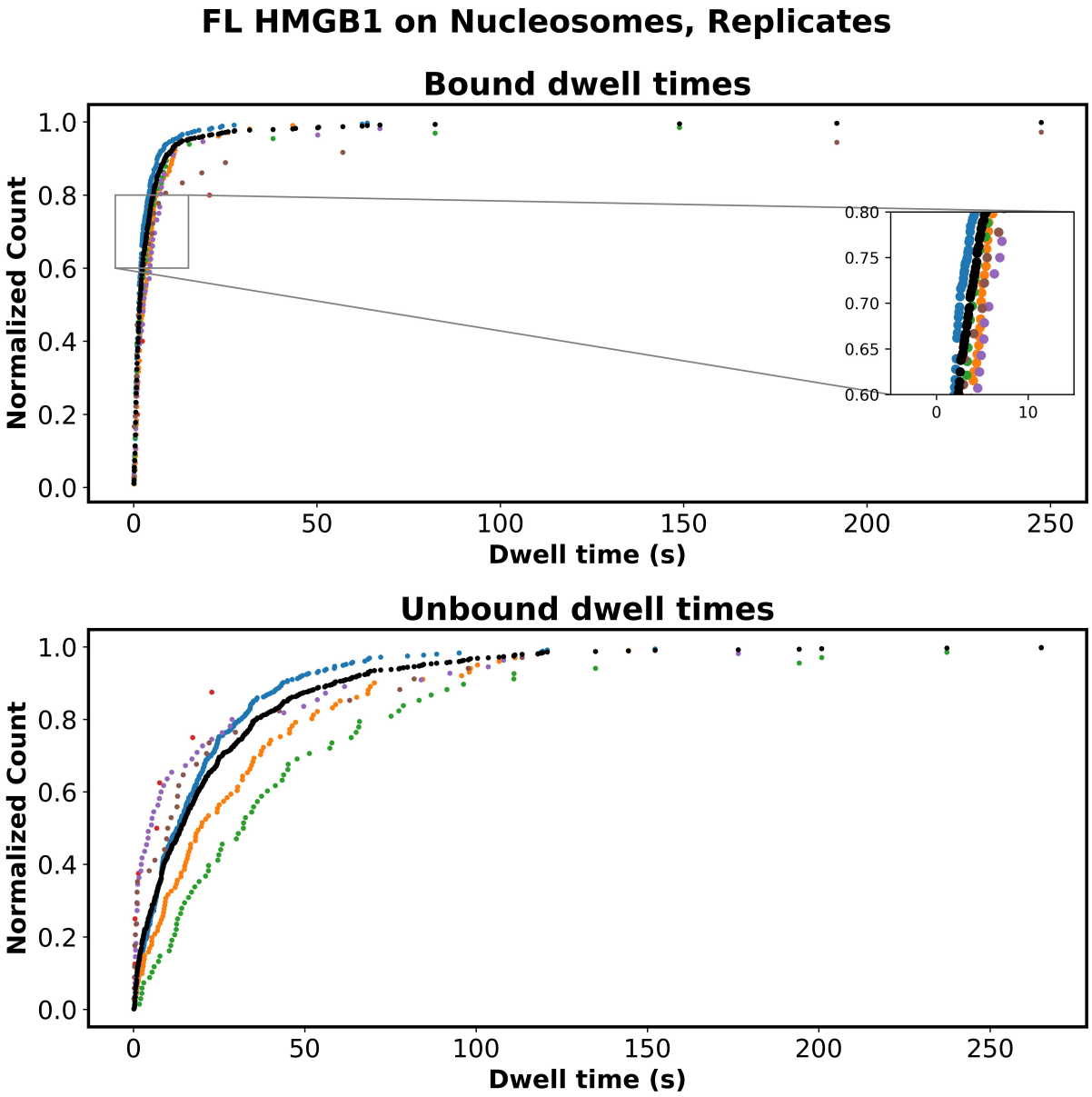

Figure 3.7: The individual experiments from 6 replicates were individually plotted as normalized cumulative sums to show the inconsistencies between experimental replicates. The combined replicates for the bound and unbound dwell times are also shown here (black circles).

best fit with a double exponential function (Figure 3.9). However, despite simplifying the system from nucleosomes to a DNA only surface, the cumulative sums plot lack of reproducibility for both the bound and unbound dwell times (Figure 3.10). Again, like with the nucleosomes, most of the bound and unbound dwell times were better fit with a double exponential function, which again indicated that there were two kinetic populations present (Table 3.2). However, replicate 2's rate of dissociation was better fitted with a single exponential function, suggesting that there was a single kinetic population that contributed to the measured rate of dissociation. Even after simplifying the system from a surface of nucleosomes to a surface of DNA, the data between experimental replicates continued to be irreproducible.

Replicate	$k_{on1}(s^{-1})$	$k_{on2}(s^{-1})$	$k_{off1}(s^{-1}1)$	$k_{off2}(s^{-1})$
Combined data	0.0502	0.72	0.133	0.97
1	0.014	0.074	0.033	0.561
2	0.070	0.945	1.153	-
3	0.043	0.815	0.149	0.911
4	0.051	1.222	0.088	1.205

Table 3.2: The rate constants from mixed exponential fits from the experimental replicates of FL HMGB1 on DNA were inconsistent with one another. Most of the replicates were best fit with a double exponential function, except for replicate 2 where its rate of dissociation was best fit with a single exponential function.

3.2.5 Experiments with HMGB1 binding to a shortened 32bp DNA were not reproducible

Despite simplifying the system from nucleosomes to free DNA, the rate constants were still inconsistent between experimental replicates. To better understand the inconsistencies behind HMGB1 on/off binding dynamics on nucleosomes and also free DNA, I used a short 32 bp DNA construct to mitigate any potential influence the longer 157 bp DNA may have on HMGB1 binding. By shortening the DNA, I decreased the number of potential binding site where HMGB1 could position itself which would theoretically allow for a more consistent HMGB1 positioning on the DNA upon binding and perhaps reduce variability in the rate constants.

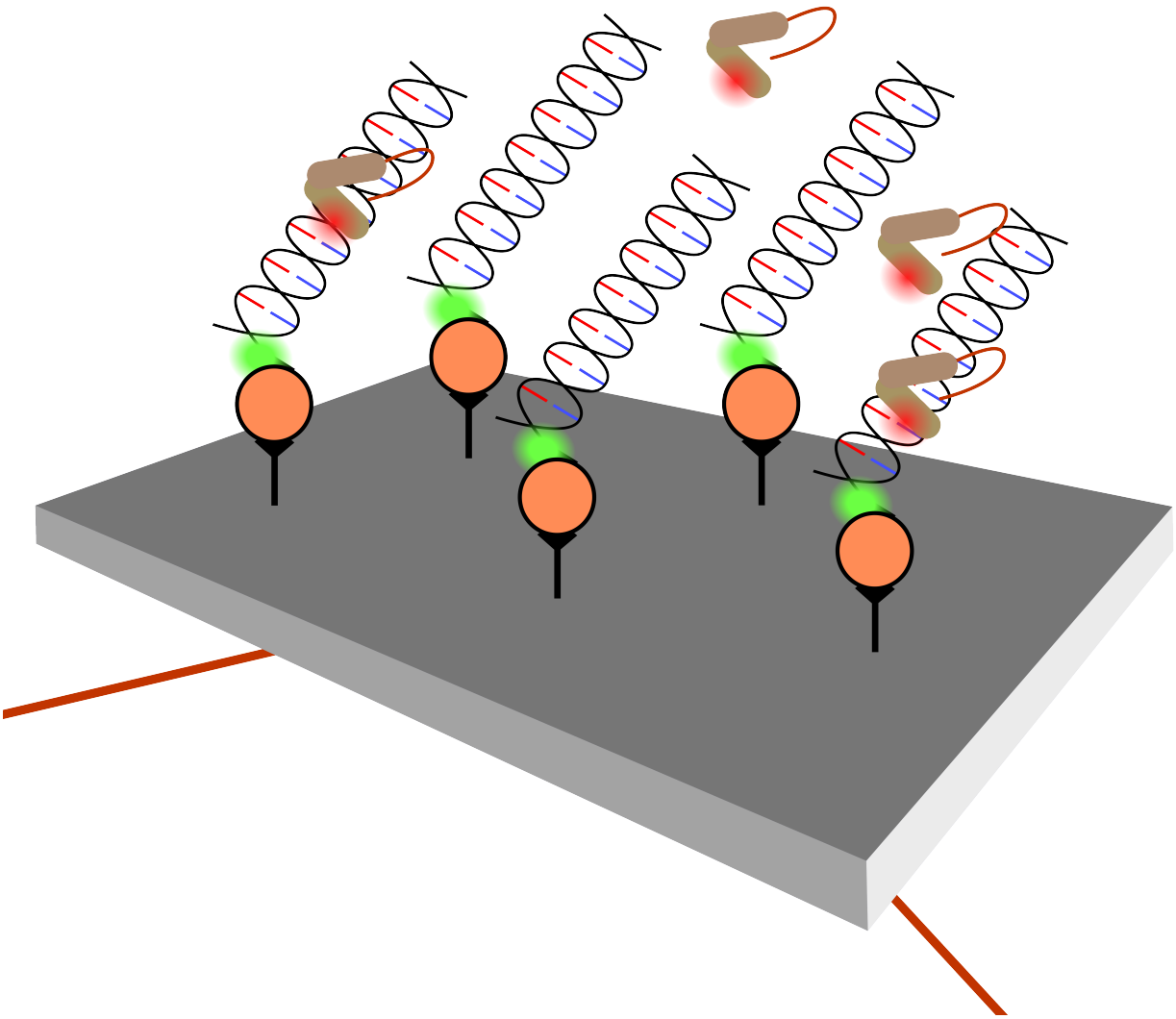

Figure 3.8: A schematic showing HMGB1 interacting with DNA sparsely immobilized to a microscopy slide functionalized with PEG-biotin through a biotin-streptavidin-biotin linkage.

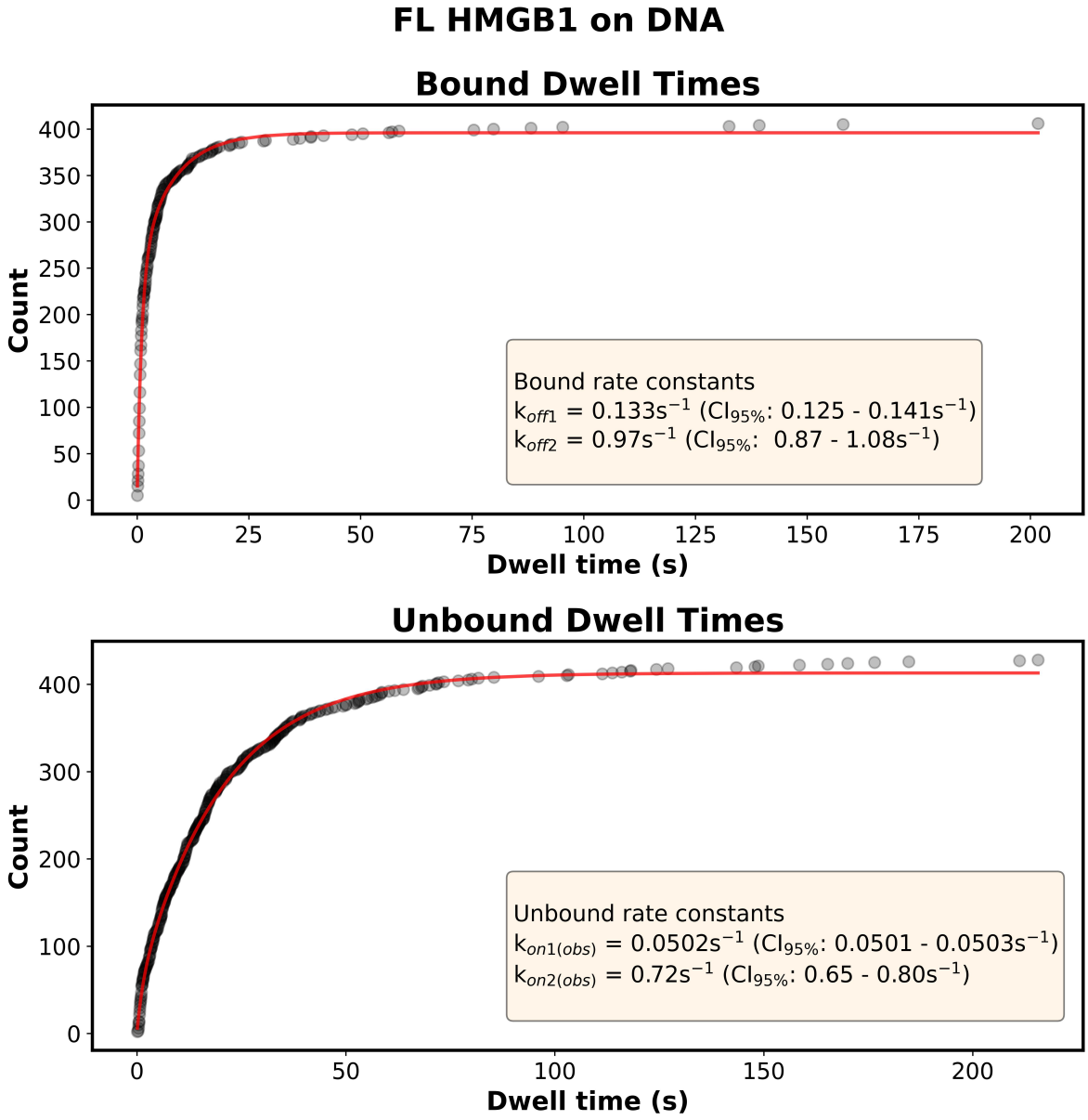

Figure 3.9: The cumulative sums plots of bound and unbound dwell times for 4 experimental replicates of HMGB1 binding to 157bp DNA were combined and plotted here. The data were fit with a double exponential function, suggesting that there are two kinetic populations present. The rate constants and 95% confidence intervals from the curve fits are shown.

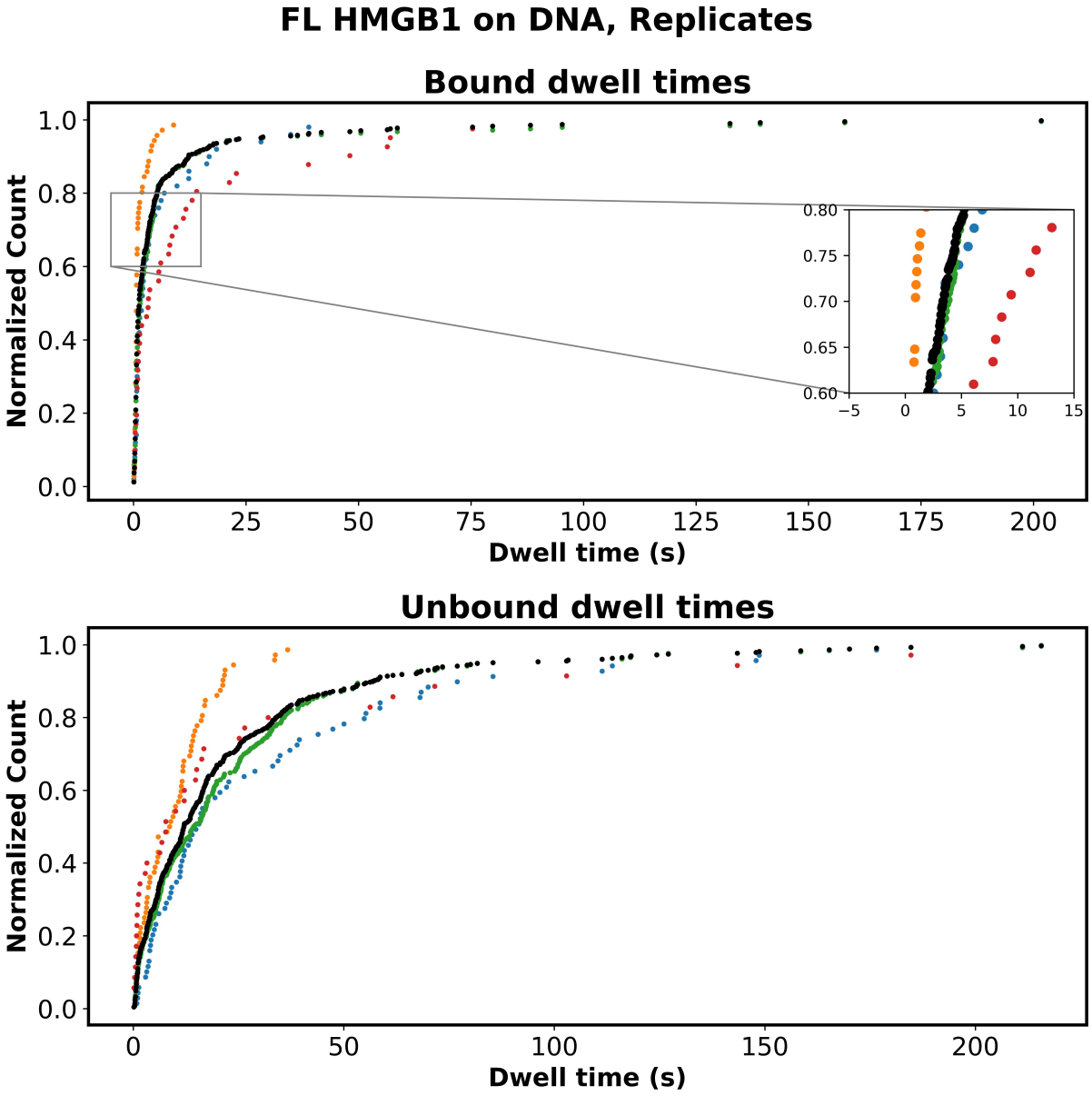

Figure 3.10: Bound and unbound dwell times from the 4 individual biological replicates of HMGB1 binding to a sparse surface of 157bp DNA plotted as normalized cumulative sums show inconsistences between experimental replicates. The combined replicates of the bound and unbound dwell times are also shown here (black circles).

The cumulative sums plot from a single experiment using 32 bp DNA was best fit with a double exponential for both the bound and unbound dwell times, again suggesting that there are two kinetic populations for both association and dissociation (Figure 3.11). However, when the experiment was repeated for a second time, there were no detectable dynamic events when the movies were analyzed. As a result, HMGB1 on/off binding dynamics remained irreproducible even on a short 32 bp DNA, suggesting that there are other underlying issues regarding the system that impacts HMGB1s binding to both nucleosomes and DNA.

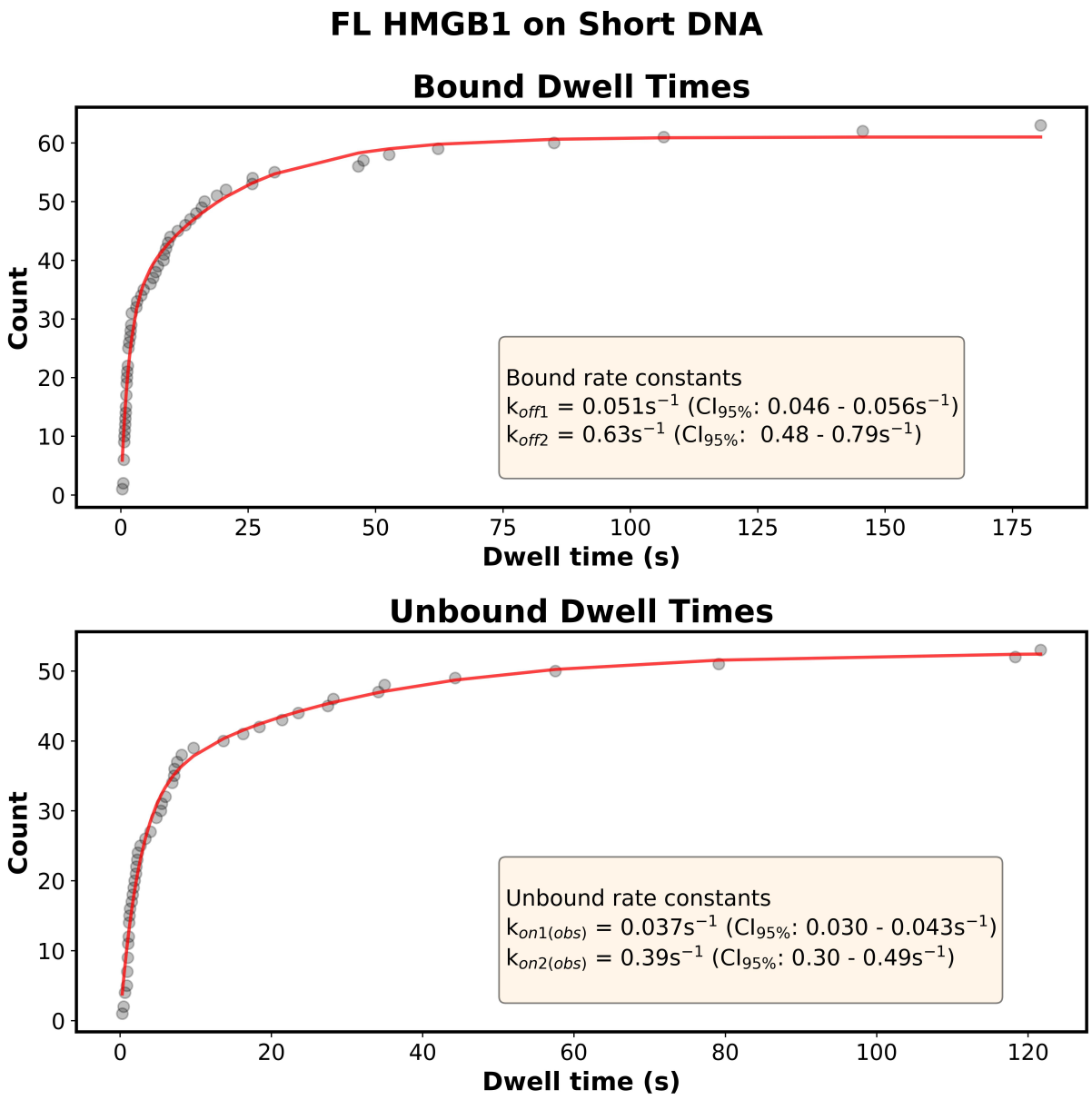

Figure 3.11: Cumulative sums plots fit with a double exponential function showing HMGB1 interacting with a short 32 bp DNA(n = 1).

3.3 Discussion

Within the nucleus, HMGB1 functions as an architectural DNA binding protein that plays an important role in chromatin organization [51] and facilitating the binding of transcription factors to their cognate DNA binding sites [49]. HMGB1 can bind to the linker DNA at the entry/exit junction of the nucleosome and interact with the histone H3 N-terminal tail peptide to help position itself onto the nucleosome [40, 44]. It is unclear how HMGB1 utilizes this combination of interactions to aid in its search and binding to nucleosomes. To better understand how HMGB1 searches for its binding sites on nucleosomes, I used single molecule TIRF microscopy to explore HMGB1 on/off binding dynamics on both nucleosomes and DNA. However, I experienced difficulties in reproducing data between experimental replicates, which persisted even upon simplifying the system from a nucleosome to a short 32 bp DNA. Ultimately, it appeared that this issue of irreproducibility persisted regardless of the nucleic acid substrate, and may be attributed to the characteristics of HMGB1 itself.

I first explored HMGB1 on/off binding dynamics on a 157bp nucleosome reconstituted with the Widom601 sequence. The DNA construct was designed to accommodate a 5bp DNA linker on each end of the nucleosome, where one end served to anchor the nucleosomes to the surface of a microscopy slide functionalized with biotin-PEG. The 5 bp DNA linker was shorter than the minimal DNA length required to accommodate both of the HMGB1s DNA binding domains [109]. This allowed me to explore how HMGB1 would interacts with nucleosomes, as opposed to observing binding exclusively to linker DNA.

Initially, I was able to collect a moderate number of bound and unbound dwell times from an experiment to produce a cumulative sums plot that was fitted with an exponential function to calculate rate constants for association and dissociation. However, I had trouble in reproducing the results from these initial experiments and believed that there may have been complications with the nucleosomes that I was not initially aware of. One potential complication may have to do with the integrity of the nucleosomes when immobilized to the surface of the slide. It is possible that the

nucleosome particle may be destabilized upon immobilization to produce a mixture of hexasomes and free DNA, despite looking homogenous on a native gel prior to dilution and immobilization. The extent of destabilization could have been variable between experiments. Another potential complication driving these observed inconsistencies could be nucleosome breathing [137]. Previous studies have shown that the nucleosomal DNA generally stays wrapped for about 250 ms and would stay in the unwrapped state for about 10-50 ms prior to rewrapping around the nucleosome [138]. The nucleosomal DNA would unwrap in 5 bp steps [139] and the extent of the wrapping might vary in length based on molecular dynamics analysis [140]. This variation in DNA length during nucleosome breathing can either facilitate or block HMGB1 binding, depending on how much DNA is accessible. The combination of nucleosome breathing and the inability to determine which of the hetereogenous population of nucleosomes, hexasomes, and DNA that HMGB1 was interacting with may have contributed to the inconsistences between the observed rate constants for association of experimental replicates. Conversely, the rate constants for dissociation seems less likely to depend on DNA availibility, thus would explain why the cumulative sums plot of bound dwell times were more tightly clustered between experiments. However, it continues to remains unclear how the nucleosomes influenced HMGB1 on/off binding dynamics aside from the inconsistencies observed between experimental replicates.

To potentially alleviate HMGB1 binding inconsistently on nucleosomes, I simplified the system by studying HMGB1 on/off binding dynamics on the same DNA construct used to reconstitute nucleosomes, but in its free (i.e. non-nucleosomal) state. Simplifying the system, removed any heterogeneity that may have been caused by the dynamic nature of the nucleosome, the structural composition of the nucleosome, as well as any potential interactions HMGB1 may have with the histone proteins. However, the inconsistencies between experimental replicates persisted even in the absence of histone proteins, and even more so with the bound dwell times. DNA sticking to the surface of the functionalized microscopy slide was not an issue as its immobilization was dependent on the biotin-streptavidin interactions (data not shown). Given that HMGB1 has a DNA binding footprint of 18 bp when using its two DNA binding domains [109], the 157 bp DNA construct could

theoretically accommodate up to 9 molecules of HMGB1. Althought it was possible that multiple HMGB1 molecules could bind to the DNA construct; however, this was infrequently observed in the intensity traces (data not shown) and is an unlikely explanation for the inconsistencies observed.

I further simplified the system by shortening the DNA from 157 bp construct to a 32 bp construct, reducing the area to which HMGB1 could bind. If the length of the DNA influences HMGB1 on/off binding dynamics, then the reduction in DNA length should localize HMGB1 to a smaller region on the DNA and improve the reproducibility of its binding behaviors. It was unlikely more than one HMGB1 molecule would occupy the short DNA due to steric constrains, and in the first experiment performed nearly all dynamic binding events within an intensity trace reflected the signature of a single HMGB1 molecule binding (data not shown). However, the second replicate produced no identifying dynamic events. This was concerning because the second replicate was done shortly after the first replicate, which used same protein and 32 bp DNA from a single preparation. More experiments of HMGB1 on the 32 bp DNA construct is needed to better understand how HMGB1 is interacting with this DNA construct and whether the binding kinetics is reproducible.

It remains unclear why it has been difficult to reproduce HMGB1 on/off binding dynamics on all three constructs (nucleosomes, long DNA, and short DNA). The inconsistencies may be linked to HMGB1 itself. It's C-terminal tail regulates HMGB1 accessibility to DNA substrates by interacting with its DNA binding domains [31–37] and each DNA binding domains can independently bind and bend DNA [14]. The structure for FL HMGB1 remains unsolved; however, the individual DNA binding domains for rat HMGB1 have been solved [19, 20], signifying that full length HMGB1 is structurally dynamic. It is possible that HMGB1 might adopt various structural conformation that modulates the accessibility of its DNA binding domain for DNA binding. Furthermore, it is possible that when HMGB1 is not stably bound to its DNA substrate, that its C-terminal tail may be able to displace itself. In combination with a sparse DNA surface, it is possible that even though the experiments were done in replicates, I may have not collected enough events per experiment to properly represent HMGB1s binding kinetics on nucleosomes and DNA. The summation of dwells may be representative of HMGB1 binding kinetics. It may be worthwhile to collect more replicates

of HMGB1 on any of the conditions attempted above and assess whether this summation of dwell times from this set of replicates differ from the previous set of replicates.

Currently, we have few insights into the binding dynamics of individual HMG boxes on nucleosomes and DNA. Thus, better understanding how a single HMG box interacts with nucleosomes and DNA could provide insight into the issues I experienced with full length HMGB1. It would be worthwhile to also isolate and explore proteins with sequence specific HMG boxes, like specific proteins within the SOX protein family [141], as well as explore the importance of DNA bending in its interaction with nucleosomes and DNA. Gaining a better understanding into the how an individual HMG box interacts with nucleosomes and DNA will provide insight into how HMG boxes from different proteins are fine tuned for their substrates.

3.4 Materials and Methods

3.4.1 DNA constructs

The preparation of the 157bp DNA construct containing the Widom601 positioning sequence was described in Chapter 2. The preparation of AF647 32 bp DNA used in this chapter was described elsewhere [119].

3.4.2 Nucleosome assembly

The microscale nucleosome assembly reaction was adapted from Hayes and Lee [134, 135] and was modified for the single molecule experiments. In a 10 μL starting reaction with buffer N-RB1 [20 mM Tris (pH 7.5), 1 mM EDTA (pH 8.0), 1 mM DTT, 2 M KCl], a single-fixed concentration within a range of 1-10 μM of the 157 bp Widom601 DNA in 2 M KCl was combined with varying molar ratios of human histone octamers (The Histone Source at Colorado State University). The molar ratio of octamer:DNA was titrated over the range of 0.6:1 to 2.0:1 and was allowed to incubate on ice for 10 minutes. After the 10-minute incubation, 5 μL of buffer N-RB2 [20 mM Tris (pH 7.5), 1 mM EDTA (pH 8.0), 1 mM DTT] was added to the reaction and was incubated on ice for 10

minutes. For every 10-minute timepoint thereafter – 10, 25, 50, 100, and 200 μL of buffer N-RB2 was added to the reaction. After the final addition, the incubation continued for an additional 10 minutes, and then was centrifuged at 16,100 x g to pellet any precipitants. The supernatant containing the nucleosomes was then assessed on a 4% native polyacrylamide gel run at 150 V for 30 minutes in 0.2x TBE on ice. The nucleosome assembly reactions were then stored at 4^{o}C on ice and at a final salt concentration of 50 mM KCl until ready to use or up to 1 month.

3.4.3 Protein expression, purification, and labeling

The protein preparation has been described in Chapter 2 under Protein expression, purification, and labeling. Briefly, HMGB1 was expressed in BL21 DE3 *E.Coli* cells and purified over a Ni-NTA resin followed by second purification over a dsDNA cellulose resin, each packed into a gravity column. HMGB1 was then fluorescently labeled with AF647 via maleimide conjugation, and the excess dye was removed using a fluorescent dye removal kit. The protein concentration and fraction labeled was then quantified via SDS-PAGE. The protein was aliquoted, snap frozen in liquid nitrogen, and stored at -80^{o}C until ready to use.

3.4.4 Slide preparation and single molecule imaging

The slide chamber preparation, DNA immobilization, and imaging buffers and conditions were described in Chapter 2.

3.4.5 Single molecule colocalization and data analysis

To perform two color colocalization, a piezo stage was used to set the coordinates and rotate through the same four regions of a slide to first collect green emission movies of the DNA, then red emission movies of the protein. After the first four regions of DNA were imaged, the slide chamber was washed twice with RB buffer and 1 nM AF647-HMGB1 prepared in imaging buffer was flowed into the slide chamber. HMGB1 movies were then collected over the same four regions set by the piezo stage. For data analysis the movies were then pulled into an in-house single molecule

analysis software written in IDL. The HMGB1 movies were colocalized with the DNA movies. The intensity traces (emission over time) of the AF647-HMGB1 from each colocalized spot pairs were fit to determine state changes (i.e. bound versus unbound). The dwell times for bound and unbound states were extracted and plotted as cumulative sums. The data were fit with exponential equations to determine rate constants.

Chapter 4

Summary and Future Directions

I investigated how HMGB1 navigated various surface substrates using smTIRF microscopy. In Chapter 2, I tracked the movement of fluorescently labeled HMGB1 and different truncation constructs on a dense surface of DNA and heparin to understand the properties that allow for HMGB1 to scan DNA. In Chapter 3, I examined the on/off binding behaviors of HMGB1 on nucleosomes and DNA to better understand HMGB1 binding dynamics. Collectively, these studies were aimed at unraveling the mechanisms that facilitate HMGB1's interactions with DNA as a steppingstone to better understand how it functions as an architectural factor within the nucleus. In this chapter, I summarize the findings of my investigation and discuss future directions aimed at expanding our understanding of HMGB1 as a dynamic DNA binding protein and architectural factor.

4.1 HMGB1 movement on DNA

I first explored the roles of that each of the three HMGB1 domains play in facilitating its movement on a dense surface of DNA. Removal of the C-terminal tail (HMGB1 AB) slightly reduced its movement speed; however, both single domain constructs (A-box and B-box) exhibited a large increase in movement speed when compared to either the full length or HMGB1 AB construct. The change in movement speed in response to the truncation of the protein domains highlights important attributes regarding HMGB1 movement where (1) The C-terminal tail plays a subtle role in modulating speed, likely by interacting with the B-Box domain to dampen its binding to

DNA, and (2) the linking of two DNA binding domains is important for setting the rate of movement of the full length protein, as the single domain constructs moved significantly faster.

Changing the DNA surface characteristics also highlighted some important HMGB1 movement attributes. Decreasing the concentration of DNA flowed into the slide chamber greatly increased HMGB1's movement speed. Increasing the salt concentration in the slide chamber appeared to ablate HMGB1's movement activity. Shortening the DNA construct to 30bp DNA or immobilization of a dense surface of heparin moderately increased HMGB1's movement speed. Most importantly, changing the surface characteristics like the ones mentioned above greatly reduced the number of identifiable HMGB1 movement tracks. Together these findings show that efficient HMGB1 binding across dense DNA is driven by its interaction with the DNA, and that movement between immobilized DNA molecules is facilitated by electrostatic interactions.

These findings highlight some of the mechanisms that facilitate HMGB1's interaction and movement on its substrates, but there remain several parameters to explore. For example, HMGB1 is known to bend DNA by intercalating specific residues between the bases within its major groove [30], potentially stabilizing HMGB1 interactions with DNA. It would be worthwhile to explore how DNA bending influences HMGB1's movement between DNA molecules, which could be done by nullifying its DNA bending capabilities through mutation of the intercalating residues. It is unclear whether HMGB1 is capable of traversing along a single DNA molecule via a 1D diffusion mechanism; it would also be interesting to explore this avenue. If so, this could provide additional insight into why longer DNAs enable more tracked molecules of HMGB1 on the surface. We can also analyze additional parameters from the TrackMate output, such as distance traveled, to gain additional insight into how these various conditions influences HMGB1's movement Additionally, HMGB1 has been shown to facilitate the binding of nuclear protein factors (e.g. transcriptional activators) to their cognate DNA binding sites [111]. It is unclear whether these transcription factors are themselves capable of movement on a dense surface of DNA, and whether the presence of HMGB1 could promote transcription factor movement. Thus, it would be interesting to explore the mechanisms by which HMGB1 facilitates transcription factors binding and potentially moving

on a dense surface of DNA containing the transcription factor's cognate DNA binding site dispersed throughout the surface.

4.2 HMGB1 on/off binding

I explored HMGB1 on/off binding dynamics on nucleosomes and DNA. I first examined how HMGB1 behaved on a surface of sparsely immobilized nucleosomes. The rate constants measured for both binding and release were inconsistent between experimental replicates for the nucleosome condition. I then attempted to resolve these inconsistencies by simplifying the system to a DNA only, but the rate constant data remained inconsistent between replicates. I then attempted to further simplify the system by using a shorter 32bp DNA construct. The first replicate produced a number of dynamic on/off binding events that were curve fit to determine rate constants, but the second replicate produced no identifiable binding events.

The inconsistencies between replicates observed with the nucleosomes may have been due to differences in the quality/heterogeneity of the nucleosomes on different days impacting HMGB1 binding kinetics. However, this cannot be the sole reason as simplifying the system to contain only DNA on the surface continued to produce inconsistent results. Even after further simplifying the system by using a shorter DNA construct, there was a large discrepancy between replicate one and two. It was surprising that one replicate produced enough data points to generate a nice curve fit, while the other produced no binding events. It is possible the latter replicate did not show binding for an unrecognized technical reason. Therefore additional experiments are needed on this short DNA to draw conclusions.

A possible explanation for the inconsistencies may be attributed the surface density. As mentioned in Chapter 2, there appears to be a critical surface density that facilitates HMGB1 movement. The surface density fell far below this threshold on the sparse surface of nucleosomes and DNA, thereby preventing movement. Thus, it is possible that the movement is important to the mechanisms of HMGB1 interacting with DNA/nucleosomes, leading to inconsistent kinetics when movement cannot occur. However, increasing the surface density would complicate the analysis

process, as a combination of HMGB1 movement and on/off binding can occur. This would make it difficult to separate the two events to analyze the results. Thus, it is possible that the experimental conditions tested in Chapter 3 were not optimal for HMGB1 binding dynamics. A potentially parameter to improve the efficacy of data collection and analysis is to optimize the signal-to-noise ratio by experimenting with fluorescent dyes that have a greater quantum yield or photostability than AF647. Increasing the signal-to-noise ratio may improve the detection of the inflection points in the intensity traces that serves as the basis for assigning association and dissociation events. Another potential parameter that can help optimize the system is exploring conditions that reduces HMGB1 sticking to the slide surface. Currently, HMGB1 sticking to the slide surface ($\leq$100 spots per region) detracts from the detection of HMGB1-DNA colocalized spot pairs. Reducing the number of surface-stuck HMGB1 molecules will improve the colocalization percentage and thus improve the efficacy of data collection and analysis.

Furthermore, this system can still be of great use to investigate the on/off binding behaviors of transcription factors interacting with their recognition sites in the presence of HMGB1. For example, we have previously used smTIRF-M co-localization on a sparse surface of DNA to measure the kinetics of human p53 binding/unbinding its response element [119, 142]. It would be interesting to explore how HMGB1 impacts the binding kinetics of p53 to its cognate DNA binding site on free DNA and nucleosomes to unravel the mechanisms that influences these processes. This would involve flowing in a mixture of fluorescently labeled protein factor with an excess of unlabeled HMGB1 and collect movies using the co-localization as described in Chapter 3.

References

1. Goodwin, G. H., Sanders, C. & Johns, E. W. A New Group of Chromatin-Associated Proteins with a High Content of Acidic and Basic Amino Acids. **Eur. J. Biochem. 38,** 14–19 (1973).

2. Reeves, R. Nuclear functions of the HMG proteins. **BBA Gene Regulatory Mechanisms 1799,** 3–14 (2010).

3. Hock, R., Furusawa, T., Ueda, T. & Bustin, M. HMG chromosomal proteins in development and disease. **Trends Cell. Biol. 17,** 72–79 (2007).

4. Vignali, R. & Marracci, S. HMGA Genes and Proteins in Development and Evolution. **Int. J. Mol. Sci. 21,** 654 (2020).

5. Evans, J. N., Zajicek, J., Nissen, M. S., Munske, G., Smith, V. & Reeves, R. 1H and 13C NMR assignments and molecular modelling of a minor groove DNA-binding peptide from the HMG-I protein. **Int. J. Pept. Protein Res. 45,** 554–560 (1995).

6. Ozturk, N., Singh, I., Mehta, A., Braun, T. & Barreto, G. HMGA proteins as modulators of chromatin structure during transcriptional activation. **Front. Cell Dev. Biol. 2** (2014).

7. Reeves, R. & Nissen, M. S. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. **JBC 265,** 8573–8582 (1990).

8. Winter, N., Nimzyk, R., Bösche, C., Meyer, A. & Bullerdiek, J. Chromatin Immunoprecipitation to Analyze DNA Binding Sites of HMGA2. **PLOS ONE 6,** e18837 (2011).

9. Catez, F., Yang, H., Tracey, K. J., Reeves, R., Misteli, T. & Bustin, M. Network of Dynamic Interactions between Histone H1 and High-Mobility-Group Proteins in Chromatin. **Mol. Cell. Biol. 24,** 4321–4328 (2004).

10. Kishi, Y., Fujii, Y., Hirabayashi, Y. & Gotoh, Y. HMGA regulates the global chromatin state and neurogenic potential in neocortical precursor cells. **Nat. Neurosci. 15,** 1127–1133 (2012).

11. Kato, H., van Ingen, H., Zhou, B.-R., Feng, H., Bustin, M., Kay, L. E. & Bai, Y. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. **Proc. Natl. Acad. Sci. U.S.A. 108,** 12283–12288 (2011).

12. Mallik, R., Kundu, A. & Chaudhuri, S. High mobility group proteins: the multifaceted regulators of chromatin dynamics. **Nucleus 61,** 213–226 (2018).

13. Nanduri, R., Furusawa, T. & Bustin, M. Biological Functions of HMGN Chromosomal Proteins. **Int. J. Mol. Sci. 21,** 449 (2020).

14. Stros, M. HMGB proteins: Interactions with DNA and chromatin. **BBA Gene Regulatory Mechanisms 1799,** 101–113 (2010).

15. Travers, A. A. Priming the nucleosome: a role for HMGB proteins? **EMBO Rep. 4,** 131–136 (2003).

16. Catena, R., Escoffier, E., Caron, C., Khochbin, S., Martianov, I. & Davidson, I. HMGB4, a Novel Member of the HMGB Family, Is Preferentially Expressed in the Mouse Testis and Localizes to the Basal Pole of Elongating Spermatids1. **Biol. Reprod. 80,** 358–366 (2009).

17. Richard, S. A., Jiang, Y., Xiang, L. H., Zhou, S., Wang, J., Su, Z. & Xu, H. Post-translational modifications of high mobility group box 1 and cancer. **Am. J. Transl. Res. 9,** 5181–5196 (2017).

18. Kwak, M. S., Kim, H. S., Lee, B., Kim, Y. H., Son, M. & Shin, J.-S. Immunological Significance of HMGB1 Post-Translational Modification and Redox Biology. **Front. Immunol. 11** (2020).

19. Hardman, C. H., Broadhurst, R. W., Raine, A. R. C., Grasser, K. D., Thomas, J. O. & Laue, E. D. Structure of the A-Domain of HMG1 and Its Interaction with DNA as Studied by Heteronuclear Three- and Four-Dimensional NMR Spectroscopy. **Biochemistry 34,** 16596–16607 (1995).

20. Weir, H., Kraulis, P., Hill, C., Raine, A., Laue, E. & Thomas, J. Structure of the HMG box motif in the B-domain of HMG1. **EMBO J. 12,** 1311–1319 (1993).

21. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C. & Norman, D. G. Solution structure of a DNA-binding domain from HMG1. **Nucleic Acids Res. 21,** 3427–3436 (1993).

22. Ohndorf, U.-M., Rould, M. A., He, Q., Pabo, C. O. & Lippard, S. J. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. **Nature 399,** 708–712 (1999).

23. Bianchi, M. E., Beltrame, M. & Paonessa, G. Specific Recognition of Cruciform DNA by Nuclear Protein HMG1. **Science 243,** 1056–1059 (1989).

24. Isackson, P. J., Fishback, J. L., Bidney, D. L. & Reeck, G. R. Preferential affinity of high molecular weight high mobility group non-histone chromatin proteins for single-stranded DNA. **JBC 254,** 5569–5572 (1979).

25. Hamada, H. & Bustin, M. Hierarchy of binding sites for chromosomal proteins HMG 1 and 2 in supercoiled deoxyribonucleic acid. **Biochemistry 24,** 1428–1433 (1985).

26. Gaillard, C. & Strauss, F. High affinity binding of proteins HMG1 and HMG2 to semicatenated DNA loops. **BMC Mol. Biol. 1,** 1 (2000).

27. Webb, M., Payet, D., Lee, K.-B., Travers, A. A. & Thomas, J. O. Structural requirements for cooperative binding of HMG1 to DNA minicircles1. **J. Mol. Biol. 309,** 79–88 (2001).

28. Brázda, V., Laister, R. C., Jagelská, E. B. & Arrowsmith, C. Cruciform structures are a common DNA feature important for regulating biological processes. **BMC Mol. Biol. 12,** 33 (2011).

29. Lin, Y., Wu, K., Jia, F., Chen, L., Wang, Z., Zhang, Y., Luo, Q., Liu, S., Qi, L., Li, N., Dong, P., Gao, F., Zheng, W., Fang, X., Zhao, Y. & Wang, F. Single cell imaging reveals cisplatin regulating interactions between transcription (co)factors and DNA. **Chem. Sci. 12,** 5419–5429 (2021).

30. Thomas, J. O. & Travers, A. A. HMG1 and 2, and related 'architectural' DNA-binding proteins. **TIBS 26,** 167–174 (2001).

31. Knapp, S., Müller, S., Digilio, G., Bonaldi, T., Bianchi, M. E. & Musco, G. The Long Acidic Tail of High Mobility Group Box 1 (HMGB1) Protein Forms an Extended and Flexible Structure That Interacts with Specific Residues within and between the HMG Boxes. **Biochemistry 43,** 11992–11997 (2004).

32. Watson, M., Stott, K. & Thomas, J. O. Mapping Intramolecular Interactions between Domains in HMGB1 using a Tail-truncation Approach. **J. Mol. Biol. 374,** 1286–1297 (2007).

33. Stott, K., Watson, M., Howe, F. S., Grossmann, J. G. & Thomas, J. O. Tail-Mediated Collapse of HMGB1 Is Dynamic and Occurs *via* Differential Binding of the Acidic Tail to the A and B Domains. **J. Mol. Biol. 403,** 706–722 (2010).

34. Lee, K.-B. & Thomas, J. O. The effect of the acidic tail on the DNA-binding properties of the HMG1,2 class of proteins: insights from tail switching and tail removal1. **J. Mol. Biol. 304,** 135–149 (2000).

35. Stros, M., Stokrova, J. & Thomas, J. O. DNA looping by the HMG-box domains of HMG1 and modulation of DNA binding by the acidic C-terminal domain. **Nucleic Acids Res. 22,** 1044–1051 (1994).

36. Blair, R. H., Horn, A. E., Pazhani, Y., Grado, L., Goodrich, J. A. & Kugel, J. F. The HMGB1 C-Terminal Tail Regulates DNA Bending. **J. Mol. Biol. 428,** 4060–4072 (2016).

37. Belgrano, F. S., Silva, I. C. d. A. d., Oliveira, F. M. B. d., Fantappie, M. R. & Mohana-Borges, R. Role of the Acidic Tail of High Mobility Group Protein B1 (HMGB1) in Protein Stability and DNA Bending. **PLOS ONE 8,** e79572 (2013).

38. Ueda, T., Chou, H., Kawase, T., Shirakawa, H. & Yoshida, M. Acidic C-Tail of HMGB1 Is Required for Its Target Binding to Nucleosome Linker DNA and Transcription Stimulation. **Biochemistry 43,** 9901–9908 (2004).

39. Aizawa, S., Nishino, H., Saito, K., Kimura, K., Shirakawa, H. & Yoshida, M. Stimulation of Transcription in Cultured Cells by High Mobility Group Protein 1: Essential Role of the Acidic Carboxyl-Terminal Region. **Biochemistry 33,** 14690–14695 (1994).

40. An, W., van Holde, K. & Zlatanova, J. The Non-histone Chromatin Protein HMG1 Protects Linker DNA on the Side Opposite to That Protected by Linker Histones. **JBC 273,** 26289–26291 (1998).

41. Cato, L., Stott, K., Watson, M. & Thomas, J. O. The Interaction of HMGB1 and Linker Histones Occurs Through their Acidic and Basic Tails. **J. Mol. Biol. 384,** 1262–1272 (2008).

42. Stros, M., Polanska, E., Kucirek, M. & Pospisilova, S. Histone H1 Differentially Inhibits DNA Bending by Reduced and Oxidized HMGB1 Protein. **PLOS ONE 10,** e0138774 (2015).

43. Thomas, J. O. & Stott, K. H1 and HMGB1: modulators of chromatin structure. **Biochem. Soc. Trans. 40,** 341–346 (2012).

44. Watson, M., Stott, K., Fischl, H., Cato, L. & Thomas, J. O. Characterization of the interaction between HMGB1 and H3—a possible means of positioning HMGB1 in chromatin. **Nucleic Acids Res. 42,** 848–859 (2014).

45. Bonaldi, T., Längst, G., Strohner, R., Becker, P. B. & Bianchi, M. E. The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. **EMBO J. 21,** 6865–6873 (2002).

46. Lange, S. S. & Vasquez, K. M. HMGB1: The jack-of-all-trades protein is a master DNA repair mechanic. **Mol. Carcinog. 48,** 571–580 (2009).

47. Jayaraman, L., Moorthy, N. C., Murthy, K. G. K., Manley, J. L., Bustin, M. & Prives, C. High mobility group protein-1 (HMG-1) is a unique activator of p53. **Genes Dev. 12,** 462–472 (1998).

48. Das, D., Peterson, R. C. & Scovell, W. M. High Mobility Group B Proteins Facilitate Strong Estrogen Receptor Binding to Classical and Half-Site Estrogen Response Elements and Relax Binding Selectivity. **J. Mol. Endocrinol. 18,** 2616–2632 (2004).

49. Ueda, T. & Yoshida, M. HMGB proteins and transcriptional regulation. **BBA Gene Regulatory Mechanisms 1799,** 114–118 (2010).

50. Sofiadis, K., Josipovic, N., Nikolic, M., Kargapolova, Y., Übelmesser, N., Varamogianni-Mamatsi, V., Zirkel, A., Papadionysiou, I., Loughran, G., Keane, J., Michel, A., Gusmao, E. G., Becker, C., Altmüller, J., Georgomanolis, T., Mizi, A. & Papantonis, A. HMGB1 coordinates SASP-related chromatin folding and RNA homeostasis on the path to senescence. **Mol. Syst. Biol. 17,** e9760 (2021).

51. Zirkel, A., Nikolic, M., Sofiadis, K., Mallm, J.-P., Brackley, C. A., Gothe, H., Drechsel, O., Becker, C., Altmüller, J., Josipovic, N., Georgomanolis, T., Brant, L., Franzen, J., Koker, M., Gusmao, E. G., Costa, I. G., Ullrich, R. T., Wagner, W., Roukos, V., Nürnberg, P., Marenduzzo, D., Rippe, K. & Papantonis, A. HMGB2 Loss upon Senescence Entry Disrupts Genomic Organization and Induces CTCF Clustering across Cell Types. **Mol. Cell 70,** 730–744.e6 (2018).

52. Aird, K. M., Iwasaki, O., Kossenkov, A. V., Tanizawa, H., Fatkhutdinov, N., Bitler, B. G., Le, L., Alicea, G., Yang, T.-L., Johnson, F. B., Noma, K.-i. & Zhang, R. HMGB2 orchestrates

the chromatin landscape of senescence-associated secretory phenotype gene loci. **JCB 215,** 325–334 (2016).

53. Guerrero, A. & Gil, J. HMGB2 holds the key to the senescence-associated secretory phenotype. **JCB 215,** 297–299 (2016).

54. Davalos, A. R., Kawahara, M., Malhotra, G. K., Schaum, N., Huang, J., Ved, U., Beausejour, C. M., Coppe, J.-P., Rodier, F. & Campisi, J. p53-dependent release of Alarmin HMGB1 is a central mediator of senescent phenotypes. **JCB 201,** 613–629 (2013).

55. Xue, J., Suarez, J. S., Minaai, M., Li, S., Gaudino, G., Pass, H. I., Carbone, M. & Yang, H. HMGB1 as a therapeutic target in disease. **J. Cell. Physiol. 236,** 3406–3419 (2021).

56. Tang, D., Kang, R., Livesey, K. M., Cheh, C.-W., Farkas, A., Loughran, P., Hoppe, G., Bianchi, M. E., Tracey, K. J., Zeh III, H. J. & Lotze, M. T. Endogenous HMGB1 regulates autophagy. **JCB 190,** 881–892 (2010).

57. Tang, D., Kang, R., Livesey, K. M., Kroemer, G., Billiar, T. R., Van Houten, B., Zeh, H. J. & Lotze, M. T. High-Mobility Group Box 1 Is Essential for Mitochondrial Quality Control. **Cell Metab. 13,** 701–711 (2011).

58. Ito, H., Fujita, K., Tagawa, K., Chen, X., Homma, H., Sasabe, T., Shimizu, J., Shimizu, S., Tamura, T., Muramatsu, S.-i. & Okazawa, H. HMGB1 facilitates repair of mitochondrial DNA damage and extends the lifespan of mutant ataxin-1 knock-in mice. **EMBO Mol. Med. 7,** 78–101 (2015).

59. Kamagata, K., Itoh, Y., Tan, C., Mano, E., Wu, Y., Mandali, S., Takada, S. & Johnson, R. C. Testing mechanisms of DNA sliding by architectural DNA-binding proteins: dynamics of single wild-type and mutant protein molecules in vitro and in vivo. **Nucleic Acids Res. 49,** 8642–8664 (2021).

60. Kamagata, K., Ouchi, K., Tan, C., Mano, E., Mandali, S., Wu, Y., Takada, S., Takahashi, S. & Johnson, R. C. The HMGB chromatin protein Nhp6A can bypass obstacles when traveling on DNA. **Nucleic Acids Res. 48,** 10820–10831 (2020).

61. Yen, Y.-M., Wong, B. & Johnson, R. C. Determinants of DNA Binding and Bending by the*Saccharomyces cerevisiae* High Mobility Group Protein NHP6A That Are Important for Its Biological Activities. **JBC 273,** 4424–4435 (1998).

62. Masse, J. E., Wong, B., Yen, Y.-M., Allain, F. H. .-., Johnson, R. C. & Feigon, J. The *S. cerevisiae* Architectural HMGB Protein NHP6A Complexed with DNA: DNA and Protein Conformational Changes upon Binding. **J. Mol. Biol. 323,** 263–284 (2002).

63. Szabo, Q., Bantignies, F. & Cavalli, G. Principles of genome folding into topologically associating domains. **Adv. Sci.** (2019).

64. Felsenfeld, G., Boyes, J., Chung, J., Clark, D. & Studitsky, V. Chromatin structure and gene expression. **Proc. Natl. Acad. Sci. U.S.A. 93,** 9384–9388 (1996).

65. Morrison, O. & Thakur, J. Molecular Complexes at Euchromatin, Heterochromatin and Centromeric Chromatin. **Int. J. Mol. Sci. 22,** 6922 (2021).

66. Liu, X.-l., Ding, J. & Meng, L.-h. Oncogene-induced senescence: a double edged sword in cancer. **Acta Pharmacol. Sin. 39,** 1553–1558 (2018).

67. Narita, M., Nuñez, S., Heard, E., Narita, M., Lin, A. W., Hearn, S. A., Spector, D. L., Hannon, G. J. & Lowe, S. W. Rb-Mediated Heterochromatin Formation and Silencing of E2F Target Genes during Cellular Senescence. **Cell 113,** 703–716 (2003).

68. Zhang, R., Poustovoitov, M. V., Ye, X., Santos, H. A., Chen, W., Daganzo, S. M., Erzberger, J. P., Serebriiskii, I. G., Canutescu, A. A., Dunbrack, R. L., Pehrson, J. R., Berger, J. M., Kaufman, P. D. & Adams, P. D. Formation of MacroH2A-Containing Senescence-Associated Heterochromatin Foci and Senescence Driven by ASF1a and HIRA. **Dev. Cell 8,** 19–30 (2005).

69. Franklin, S., Chen, H., Mitchell-Jordan, S., Ren, S., Wang, Y. & Vondriska, T. M. Quantitative Analysis of the Chromatin Proteome in Disease Reveals Remodeling Principles and Identifies High Mobility Group Protein B2 as a Regulator of Hypertrophic Growth. **MCP 11,** M111.014258 (2012).

70. Monte, E., Rosa-Garrido, M., Karbassi, E., Chen, H., Lopez, R., Rau, C. D., Wang, J., Nelson, S. F., Wu, Y., Stefani, E., Lusis, A. J., Wang, Y., Kurdistani, S. K., Franklin, S. & Vondriska, T. M. Reciprocal Regulation of the Cardiac Epigenome by Chromatin Structural Proteins Hmgb and Ctcf. **JBC 291,** 15428–15446 (2016).

71. Lu, B., Liu, M., Gu, L., Li, Y., Shen, S., Guo, G., Wang, F., He, X., Zhao, Y., Shang, X., Wang, L., Yang, G., Zhu, Q., Cao, J., Jiang, C., Culleton, R., Wei, G. & Zhang, Q. The Architectural Factor HMGB1 Is Involved in Genome Organization in the Human Malaria Parasite Plasmodium falciparum. **MBio** (2021).

72. Deitsch, K. W. & Dzikowski, R. Variant Gene Expression and Antigenic Variation by Malaria Parasites. **Annu. Rev. Microbiol. 71,** 625–641 (Volume 71, 2017 2017).

73. Batugedara, G., Lu, X. M., Bunnik, E. M. & Le Roch, K. G. The Role of Chromatin Structure in Gene Regulation of the Human Malaria Parasite. **Trends Parasitol. 33,** 364–377 (2017).

74. Bunnik, E. M., Cook, K. B., Varoquaux, N., Batugedara, G., Prudhomme, J., Cort, A., Shi, L., Andolina, C., Ross, L. S., Brady, D., Fidock, D. A., Nosten, F., Tewari, R., Sinnis, P., Ay, F., Vert, J.-P., Noble, W. S. & Le Roch, K. G. Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. **Nat. Commun. 9,** 1910 (2018).

75. Bunnik, E. M., Venkat, A., Shao, J., McGovern, K. E., Batugedara, G., Worth, D., Prudhomme, J., Lapp, S. A., Andolina, C., Ross, L. S., Lawres, L., Brady, D., Sinnis, P., Nosten, F., Fidock, D. A., Wilson, E. H., Tewari, R., Galinski, M. R., Ben Mamoun, C., Ay, F. & Le Roch, K. G. Comparative 3D genome organization in apicomplexan parasites. **Proc. Natl. Acad. Sci. U.S.A. 116,** 3183–3192 (2019).

76. Polanska, E., Dobsakova, Z., Dvorackova, M., Fajkus, J. & Stros, M. HMGB1 gene knockout in mouse embryonic fibroblasts results in reduced telomerase activity and telomere dysfunction. **Chromosoma 121,** 419–431 (2012).

77. Pagano, B., Margarucci, L., Zizza, P., Amato, J., Iaccarino, N., Cassiano, C., Salvati, E., Novellino, E., Biroccio, A., Casapullo, A. & Randazzo, A. Identification of novel interactors of human telomeric G-quadruplex DNA. **Chem. Commun. 51,** 2964–2967 (2015).

78. Amato, J., Cerofolini, L., Brancaccio, D., Giuntini, S., Iaccarino, N., Zizza, P., Iachettini, S., Biroccio, A., Novellino, E., Rosato, A., Fragai, M., Luchinat, C., Randazzo, A. & Pagano, B. Insights into telomeric G-quadruplex DNA recognition by HMGB1 protein. **Nucleic Acids Res. 47,** 9950–9966 (2019).

79. Bochman, M. L., Paeschke, K. & Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. **Nat. Rev. Genet. 13,** 770–780 (2012).

80. Spiegel, J., Adhikari, S. & Balasubramanian, S. The Structure and Function of DNA G-Quadruplexes. **TRENDS CHEM. 2,** 123–136 (2020).

81. Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P. & Balasubramanian, S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. **Nat. Biotechnol. 33,** 877–881 (2015).

82. Huppert, J. L. & Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. **Nucleic Acids Res. 35,** 406–413 (2007).

83. Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? **Nat. Rev. Drug. Discov. 10,** 261–275 (2011).

84. Moye, A. L., Porter, K. C., Cohen, S. B., Phan, T., Zyner, K. G., Sasaki, N., Lovrecz, G. O., Beck, J. L. & Bryan, T. M. Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. **Nat. Commun. 6,** 7643 (2015).

85. Dogan, F. & Forsyth, N. R. Telomerase Regulation: A Role for Epigenetics. **Cancers 13,** 1213 (2021).

86. Ke, S., Zhou, F., Yang, H., Wei, Y., Gong, J., Mei, Z., Wu, L., Yu, H. & Zhou, Y. Down-regulation of high mobility group box 1 modulates telomere homeostasis and increases the radiosensitivity of human breast cancer cells. **Int. J. Oncol. 46,** 1051–1058 (2015).

87. Li, J., Correia, J. J., Wang, L., Trent, J. O. & Chaires, J. B. Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. **Nucleic Acids Res. 33,** 4649–4659 (2005).

88. Tsuchida, N., Murugan, A. K. & Grieco, M. Kirsten Ras oncogene: Significance of its discovery in human cancer research. **Oncotarget 7,** 46717–46733 (2016).

89. Cogoi, S. & Xodo, L. E. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. **Nucleic Acids Res. 34,** 2536–2549 (2006).

90. Kaiser, C. E., Van Ert, N. A., Agrawal, P., Chawla, R., Yang, D. & Hurley, L. H. Insight into the Complexity of the i-Motif and G-Quadruplex DNA Structures Formed in the KRAS Promoter and Subsequent Drug-Induced Gene Repression. **J. Am. Chem. Soc. 139,** 8522–8536 (2017).

91. Paramasivam, M., Membrino, A., Cogoi, S., Fukuda, H., Nakagama, H. & Xodo, L. E. Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription. **Nucleic Acids Res. 37,** 2841–2853 (2009).

92. Amato, J., Madanayake, T. W., Iaccarino, N., Novellino, E., Randazzo, A., Hurley, L. H. & Pagano, B. HMGB1 binds to the KRAS promoter G-quadruplex: a new player in oncogene transcriptional regulation? **Chem. Commun. 54,** 9442–9445 (2018).

93. Arimondo, P. B., Gelus, N., Hamy, F., Payet, D., Travers, A. & Bailly, C. The chromosomal protein HMG-D binds to the TAR and RBE RNA of HIV-1. **FEBS Lett. 485,** 47–52 (2000).

94. Bell, A. J., Chauhan, S., Woodson, S. A. & Kallenbach, N. R. Interactions of recombinant HMGB proteins with branched RNA substrates. **Biochem. Biophys. Res. Commun. 377,** 262–267 (2008).

95. Castello, A., Fischer, B., Frese, C. K., Horos, R., Alleaume, A.-M., Foehr, S., Curk, T., Krijgsveld, J. & Hentze, M. W. Comprehensive Identification of RNA-Binding Domains in Human Cells. **Mol. Cell 63,** 696–710 (2016).

96. He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D. C., Wilusz, J. E., Garcia, B. A. & Bonasio, R. High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. **Mol. Cell 64,** 416–430 (2016).

97. Kargapolova, Y., Levin, M., Lackner, K. & Danckwardt, S. sCLIP—an integrated platform to study RNA–protein interactomes in biomedical research: identification of CSTF2tau in alternative processing of small nuclear RNAs. **Nucleic Acids Res. 45,** 6074–6086 (2017).

98. Barreiro-Alonso, A., Lamas-Maceiras, M., García-Díaz, R., Rodríguez-Belmonte, E., Yu, L., Pardo, M., Choudhary, J. S. & Cerdán, M. E. Delineating the HMGB1 and HMGB2 interactome in prostate and ovary epithelial cells and its relationship with cancer. **Oncotarget 9,** 19050–19064 (2018).

99. Moura-Alves, P., Neves-Costa, A., Raquel, H., Pacheco, T. R., D'Almeida, B., Rodrigues, R., Cadima-Couto, I., Chora, Â., Oliveira, M., Gama-Carvalho, M., Hacohen, N. & Moita, L. F. An shRNA-Based Screen of Splicing Regulators Identifies SFRS3 as a Negative Regulator of IL-1β Secretion. **PLOS ONE 6,** e19829 (2011).

100. Lau, L., Porciuncula, A., Yu, A., Iwakura, Y. & David, G. Uncoupling the Senescence-Associated Secretory Phenotype from Cell Cycle Exit via Interleukin-1 Inactivation Unveils Its Protumorigenic Role. **Mol. Cell. Biol. 39,** e00586–18 (2019).

101. Shang, D., Hong, Y., Xie, W., Tu, Z. & Xu, J. Interleukin-1β Drives Cellular Senescence of Rat Astrocytes Induced by Oligomerized Amyloid β Peptide and Oxidative Stress. **Front. Neurol. 11** (2020).

102. Kopp, F. & Mendell, J. T. Functional Classification and Experimental Dissection of Long Noncoding RNAs. **Cell 172,** 393–407 (2018).

103. Lou, M.-M., Tang, X.-Q., Wang, G.-M., He, J., Luo, F., Guan, M.-F., Wang, F., Zou, H., Wang, J.-Y., Zhang, Q., Xu, M.-J., Shi, Q.-L., Shen, L.-B., Ma, G.-M., Wu, Y., Zhang, Y.-Y., Liang, A.-b., Wang, T.-H., Xiong, L.-L., Wang, J., Xu, J. & Wang, W.-Y. Long noncoding RNA BS-DRL1 modulates the DNA damage response and genome stability by interacting with HMGB1 in neurons. **Nat. Commun. 12,** 4075 (2021).

104. Gao, D., Lv, A.-e., Li, H.-P., Han, D.-H. & Zhang, Y.-P. LncRNA MALAT-1 Elevates HMGB1 to Promote Autophagy Resulting in Inhibition of Tumor Cell Apoptosis in Multiple Myeloma. **J. Cell. Biochem. 118,** 3341–3348 (2017).

105. Tang, X., Li, P.-H. & Chen, H.-Z. Cardiomyocyte Senescence and Cellular Communications Within Myocardial Microenvironments. **Front. Endocrinol. 11** (2020).

106. Kang, R., Chen, R., Zhang, Q., Hou, W., Wu, S., Cao, L., Huang, J., Yu, Y., Fan, X.-g., Yan, Z., Sun, X., Wang, H., Wang, Q., Tsung, A., Billiar, T. R., Zeh, H. J., Lotze, M. T. & Tang, D. HMGB1 in health and disease. **Mol. Asp. Med. 40,** 1–116 (2014).

107. Mukherjee, A. & Vasquez, K. M. Targeting Chromosomal Architectural HMGB Proteins Could Be the Next Frontier in Cancer Therapy. **Cancer Res. 80,** 2075–2082 (2020).

108. Yanai, H., Matsuda, A., An, J., Koshiba, R., Nishio, J., Negishi, H., Ikushima, H., Onoe, T., Ohdan, H., Yoshida, N. & Taniguchi, T. Conditional ablation of HMGB1 in mice reveals its protective function against endotoxemia and bacterial infection. **Proc. Natl. Acad. Sci. U.S.A. 110,** 20699–20704 (2013).

109. McCauley, M. J., Zimmerman, J., Maher, L. J. & C.Williams, M. HMGB Binding to DNA: Single and Double Box Motifs. **J. Mol. Biol. 374,** 993–1004 (2007).

110. Yang, H., Antoine, D. J., Andersson, U. & Tracey, K. J. The many faces of HMGB1: molecular structure-functional activity in inflammation, apoptosis, and chemotaxis. **J Leukoc Biol 93,** 865–873 (2013).

111. Agresti, A. & Bianchi, M. E. HMGB proteins and gene expression. **Curr. Opin. Genet. Dev. 13,** 170–178 (2003).

112. Ner, S. S., Travers, A. A. & Churchill, M. E. Harnessing the writhe: a role for DNA chaperones in nucleoprotein-complex formation. **Trends Biochem Sci 19,** 185–187 (1994).

113. Phair, R. D., Scaffidi, P., Elbi, C., Vecerová, J., Dey, A., Ozato, K., Brown, D. T., Hager, G., Bustin, M. & Misteli, T. Global Nature of Dynamic Protein-Chromatin Interactions In Vivo: Three-Dimensional Genome Scanning and Dynamic Interaction Networks of Chromatin Proteins. **Mol Cell Biol 24,** 6393–6402 (2004).

114. Scaffidi, P., Misteli, T. & Bianchi, M. E. Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. **Nature 418,** 191–195 (2002).

115. Falciola, L., Spada, F., Calogero, S., Längst, G., Voit, R., Grummt, I. & Bianchi, M. E. High Mobility Group 1 Protein Is Not Stably Associated with the Chromosomes of Somatic Cells. **JCB 137,** 19–26 (1997).

116. Skoko, D., Wong, B., Johnson, R. C. & Marko, J. F. Micromechanical Analysis of the Binding of DNA-Bending Proteins HMGB1, NHP6A, and HU Reveals Their Ability To Form Highly Stable DNA-Protein Complexes. **Biochemistry 43,** 13867–13874 (2004).

117. Ershov, D., Phan, M.-S., Pylvänäinen, J. W., Rigaud, S. U., Le Blanc, L., Charles-Orszag, A., Conway, J. R. W., Laine, R. F., Roy, N. H., Bonazzi, D., Duménil, G., Jacquemet, G. & Tinevez, J.-Y. TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. **Nat Methods 19,** 829–832 (2022).

118. Groeneveld, R. A. & Meeden, G. Measuring Skewness and Kurtosis. **J.R. Stat. 33,** 391–399 (1984).

119. Ly, E., Kugel, J. F. & Goodrich, J. A. Single molecule studies reveal that p53 tetramers dynamically bind response elements containing one or two half sites. **Sci Rep 10,** 16176 (2020).

120. Magaña-Acosta, M. & Valadez-Graham, V. Chromatin Remodelers in the 3D Nuclear Compartment. **Front. Genet. 11** (2020).

121. Ito, T. Role of Histone Modification in Chromatin Dynamics. **JB 141,** 609–614 (2007).

122. Iwafuchi-Doi, M. The mechanistic basis for chromatin regulation by pioneer transcription factors. **WIREs Systems Biology and Medicine 11,** e1427 (2019).

123. Gómez-Díaz, E. & Corces, V. G. Architectural proteins: regulators of 3D genome organization in cell fate. **Trends Cell. Biol. 24,** 703–711 (2014).

124. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. **Nature 389,** 251–260 (1997).

125. Yang, T. P., Hansen, S. K., Oishi, K. K., Ryder, O. A. & Hamkalo, B. A. Characterization of a cloned repetitive DNA sequence concentrated on the human X chromosome. **Proc. Natl. Acad. Sci. U.S.A. 79,** 6593–6597 (1982).

126. Gottesfeld, J. M. DNA sequence-directed nucleosome reconstitution on 5S RNA genes of Xenopus laevis. **Mol. Cell. Biol. 7,** 1612–1622 (1987).

127. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning1. **J. Mol. Biol. 276,** 19–42 (1998).

128. Chua, E. Y. D., Vasudevan, D., Davey, G. E., Wu, B. & Davey, C. A. The mechanics behind DNA sequence-dependent properties of the nucleosome. **Nucleic Acids Res. 40,** 6338–6352 (2012).

129. Ju, B.-G., Lunyak, V. V., Perissi, V., Garcia-Bassets, I., Rose, D. W., Glass, C. K. & Rosenfeld, M. G. A Topoisomerase IIβ-Mediated dsDNA Break Required for Regulated Transcription. **Science 312,** 1798–1802 (2006).

130. Willcockson, M. A., Healton, S. E., Weiss, C. N., Bartholdy, B. A., Botbol, Y., Mishra, L. N., Sidhwani, D. S., Wilson, T. J., Pinto, H. B., Maron, M. I., Skalina, K. A., Toro, L. N., Zhao, J., Lee, C.-H., Hou, H., Yusufova, N., Meydan, C., Osunsade, A., David, Y., Cesarman, E., Melnick, A. M., Sidoli, S., Garcia, B. A., Edelmann, W., Macian, F. & Skoultchi, A. I. H1 histones control the epigenetic landscape by local chromatin compaction. **Nature 589,** 293–298 (2021).

131. Allan, J., Hartman, P. G., Crane-Robinson, C. & Aviles, F. X. The structure of histone H1 and its location in chromatin. **Nature 288,** 675–679 (1980).

132. Syed, S. H., Goutte-Gattat, D., Becker, N., Meyer, S., Shukla, M. S., Hayes, J. J., Everaers, R., Angelov, D., Bednar, J. & Dimitrov, S. Single-base resolution mapping of H1–nucleosome interactions and 3D organization of the nucleosome. **Proc. Natl. Acad. Sci. U.S.A. 107,** 9620–9625 (2010).

133. Zhou, B.-R., Feng, H., Kato, H., Dai, L., Yang, Y., Zhou, Y. & Bai, Y. Structural insights into the histone H1-nucleosome complex. **Proc. Natl. Acad. Sci. U.S.A. 110,** 19390–19395 (2013).

134. Hayes, J. J. & Lee, K.-M. In Vitro Reconstitution and Analysis of Mononucleosomes Containing Defined DNAs and Proteins. **Methods 12,** 2–9 (1997).

135. Böhm, V., Hieb, A. R., Andrews, A. J., Gansen, A., Rocker, A., Tóth, K., Luger, K. & Langowski, J. Nucleosome accessibility governed by the dimer/tetramer interface. **Nucleic Acids Res. 39,** 3093–3102 (2011).

136. Arimura, Y., Tachiwana, H., Oda, T., Sato, M. & Kurumizaka, H. Structural Analysis of the Hexasome, Lacking One Histone H2A/H2B Dimer from the Conventional Nucleosome. **Biochemistry 51,** 3302–3309 (2012).

137. Polach, K. J. & Widom, J. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. **J Mol Biol 254,** 130–149 (1995).

138. Li, G., Levitus, M., Bustamante, C. & Widom, J. Rapid spontaneous accessibility of nucleosomal DNA. **Nat. Struct. Mol. Biol. 12,** 46–53 (2005).

139. Hall, M. A., Shundrovsky, A., Bai, L., Fulbright, R. M., Lis, J. T. & Wang, M. D. High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. **Nat. Struct. Mol. Biol. 16,** 124–129 (2009).

140. Armeev, G. A., Kniazeva, A. S., Komarova, G. A., Kirpichnikov, M. P. & Shaytan, A. K. Histone dynamics mediate DNA unwrapping and sliding in nucleosomes. **Nat Commun 12,** 2387 (2021).

141. Weiss, M. A. Floppy SOX: Mutual Induced Fit in HMG (High-Mobility Group) Box-DNA Recognition. **Mol. Endocrinol. 15,** 353–362 (2001).

142. Suwita, J. P., Voong, C. K., Ly, E., Goodrich, J. A. & Kugel, J. F. Single molecule studies characterize the kinetic mechanism of tetrameric p53 binding to different native response elements. **PLOS ONE 18,** e0286193 (2023).