Reconstitution of FACT-Subnucleosome Complexes to Capture the Missing Domains of FACT

Alexander W. Black

Luger Laboratory

April 6, 2020

Thesis Advisor:

Dr. Karolin Luger

Department of Biochemistry

Committee Members:

Dr. Jeffrey Cameron

Department of Biochemistry

Dr. Jennifer Martin

Department of Molecular, Cellular and Developmental Biology

Table of Contents

Acknowledgements

I would like to first and foremost extend my deepest appreciation to Dr. Karolin Luger for the opportunity to pursue my research in her lab and for her mentorship and care. Without her, none of this would have been possible. I also extend my gratitude to Dr. Keda Zhou and Dr. Yang Liu for the countless hours they have both spent mentoring and guiding me along this path. I would also like to offer special thanks to Garrett Edwards for operating the AUC and processing the data. Finally, I'd like to thank everyone in the Luger lab for their support and inspiration that has helped to motivate me to pursue my honors research.

Abstract

Facilitates **C**hromatin **T**ranscription (FACT) is a histone chaperone which relieves the barrier that nucleosomes represent to many DNA related processes in a cell. Human FACT exists as a heterodimer consisting of the SPT16 and SSRP1 subunits and binds to nucleosome components. Previous attempts on the FACT-subnucleosome complex through cryo-EM captured most of FACT, however a few crucial domains remain missing. Here, we asymmetrically extend the DNA wrapping the histone proteins on the distal side to form 110 and 128 base pair (bp) subnucleosomes. FACT-(H2A–H2B) was added to these substrates and homogeneity and stability of the complexes were observed through gel shift assays and Sedimentation velocity analytical ultracentrifugation. We found that the 128 bp subnucleosome as substrate forms the most homogenous and stable FACT-subnucleosome complex. While further optimization of the homogeneity and stability of the 128 bp complex is recommended, this investigation demonstrates that the 128 bp subnucleosome is a suitable substrate for FACT. Using this subnucleosome, the missing domains of FACT could be captured through cryo-EM. A complete structure of FACT bound to a subnucleosome complex will provide further insight into the mechanism behind FACT's important biological role of nucleosome assembly and disassembly in transcription.

Introduction

This investigation aims to improve upon a previous study conducted by Lui et al 2019 where the histone chaperone protein FACT (**FA**cilitates **C**hromatin **T**ranscription), bound in a complex with a subnucleosome, was imaged through cryogenic electron microscopy.[1] FACT is a protein that is required for many DNA-related processes; such as transcription, where the presence of nucleosomes remains a barrier for cellular machinery.[2] FACT has been shown to interact with nucleosome components, and studies using FACT have shown that it can facilitate both the disassembly and assembly of nucleosomes.[3] Transcription, and other DNA related processes that require FACT are essential for the cell's survival, which highlights the importance of FACT. Despite its importance, relatively little is known about mechanism of how FACT mediated nucleosome assembly/disassembly occurs. This results from an incomplete understanding of how FACT binds to its subnucleosome substrates.

For the first time, FACT was captured in complex with its subnucleosome substrate through cryogenic electron microscopy by Lui et al 2019. However, large portions of the FACT protein could not be resolved due to limitations in their subnucleosome substrate. This project builds upon their work by extending the DNA used to wrap the histone components. Some parts of the missing FACT domains are known to bind DNA, so they will likely interact with the extended DNA. By tethering these regions of FACT to the complex, those regions would show a decrease in flexibility and an increase in intrinsic stability. Hence, the aim of this project is to reconstitute stable subnucleosome complexes that contain enough DNA to capture the missing regions of the bound FACT protein. Achieving this would further our understanding of how FACT binding occurs and lets us infer the mechanism of FACT mediated nucleosome assembly and disassembly. Knowing how FACT functions is a step towards understanding FACT recruitment, activation, and helps foster a larger view of how cells bypass nucleosome barriers. In addition, FACT is known to be upregulated in cancer cells.[4] Understanding how FACT functions can lead to successful inhibitor design and an idea of how nucleosome assembly and disassembly occurs in cancerous cells.

Background

Across all domains of life, genetic information is stored in the form of nucleic acids. Nucleic acids are biopolymers, made up of individual monomers that can stretch to enormous lengths. DNA is an example of a nucleic acid, and for the vast majority of organisms, DNA is the predominant medium of information storage. This is due in part to its quaternary code, representing the four bases in DNA, as well as its stability. In eukaryotic organisms, DNA is packaged and stored in the nucleus of the cell and is bound to histone proteins to form chromatin. Compaction of DNA into chromatin allows a 40,000-fold reduction in length.[5] This is essential since an entire organism's genome must be stored in the nucleus of a cell. Chromatin contains multiple levels of structure, but the fundamental repeating unit was discovered in the 1970's to be a nucleoprotein complex appropriately denoted as the nucleosome.[6] The histone proteins essential to the chromatin structure have regions of highly-conserved structure.

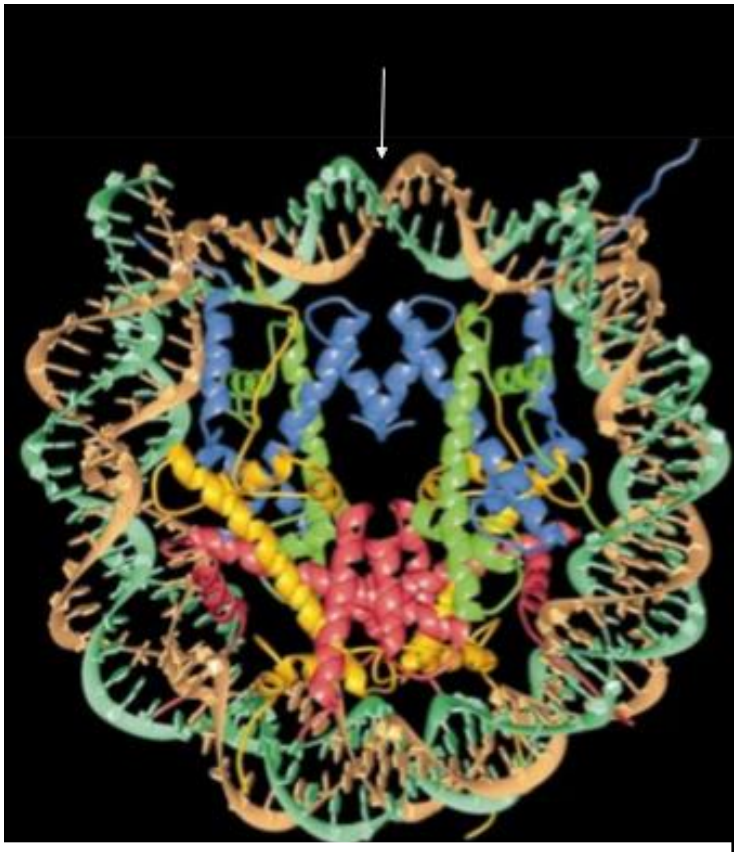

Figure 1. Nucleosome Core
146 bp of DNA wraps around the nucleosome core comprised of the histone octamer. Individual histones are labelled: H2A: yellow, H2B: red, H3: blue, H4: green. Figure taken from Luger et al.,1997. The white arrow marks the dyad.

Nucleosome structure was solved at 2.8 Å through X-ray crystallography in Luger 1997.[7] The crystallized nucleosome wrapped 146 base pairs (bp) of a palindromic human α-satellite DNA sequence 1.65 times with a radius measured at 41.8 Å. Nucleosome complexes consist of two copies of histones H2A and H2B assembled as

two dimers. Two copies of each H3 and H4 proteins are assembled as a tetramer. The two H2A–H2B dimers along with the $(H3–H4)_2$ tetramer form a histone octamer which wraps 147 bp of DNA. Figure 1 depicts the nucleosome schematic with labelled histone proteins. This histone octamer constitutes the "core" of the nucleosome. The nucleosome dyad can be seen at the apex marked by the white arrow.[7]

The core histones share many similarities. They all have a basic, flexible domain on their amino terminal end, as well as a highly structured domain.[8] In addition, the core histones all contain the same folding motif. This folding pattern of helix-loop-helix-loop-helix is used to generate the 'handshake motif' which allows assembly into heterodimers.[9] Heterodimer formation however, only exists between H2A–H2B and H3–H4.[3] In addition, the H3–H4 heterodimer can itself dimerize to form the $(H3–H4)_2$ tetramer. The prevention of homodimer or other heterodimer formations is accomplished through complementary packing. The L1 loop lengths of H3 and H2B are longer than H4 and H2A, which complements the early termination of the $\alpha 1$ helix in H4 and H2A. The histone fold pairs are also used in DNA binding. Two DNA binding regions fall into two categories; $\alpha 1\alpha 1$ or L1L2, named after the histone motifs, used to bind the DNA. The fold pairs each wrap 27-28 bp DNA. Histone core proteins also contain conserved lysine residues which form ionic interactions with the phosphodiester backbone wrapping an additional ~35 bp of DNA.

A fifth histone, often H1, is the linker histone which binds the nucleosome by interacting with the linker DNA outside the core 147 bp.[10] The addition of H1 on the nucleosome is essential for the construction of higher order chromatin structures. H1 is involved with chromatin condensation and protects an extra 20 bp of linker DNA in the nucleosome.[11] While the H1 is important for more complex structures, the core nucleosome octamer is stable in solution without H1.

Higher order structure of chromatin exists when nucleosomes are assembled into a three-dimensional, repeatable unit.[12] The simplest form of chromatin consists of nucleosomes connected by linker DNA in a 'beads on a string' structure. While the negatively charged DNA

naturally repels away other DNA molecules through coulombic forces, solutions with higher ionic strengths can drive condensation *in vitro*. A commonly recognized higher-order structure is the 30 nm fiber which is formed by polynucleosome condensation. The linker histone H1 is required for 30 nm fiber stabilization. This is most likely due to the charge neutralization that the linker histone provides, reducing the coulombic repulsion between DNA.[12]

While higher order chromatin structures may be advantageous for DNA packing, they are barriers to many DNA-related processes such as replication, repair and transcription.[13] For transcription, the multi-enzyme complex that is required, has difficulty in accessing tightly packed chromatin. This can be partially alleviated by positioning transcriptionally active genes on the periphery of the chromosomes, though not all genes follow this pattern.[14] While a degree of chromatin decondensing is generally seen, some genes may, paradoxically be transcriptionally active when parts of the chromatin is more compact.[10, 15] This may be explained by chromatin dynamics permitting the entry of transcription factors and nuclear proteins into condensed chromatin and heterochromatin.[16, 17]

While the exact degree of chromatin compaction required for transcription remains unclear, the levels of compaction may be modulated through various methods. One such method is through histone modification. These post-translational covalent modifications include acetylation, methylation, phosphorylation and many others.[18] These modifications have been correlated with specific effects, such as acetylation being used to promote transcription. The correlation between histone modification, typically at the promoter region, can be generalized across multiple cell types.[19] Histone modification remains an important cellular tool for the regulation of transcription. Another prominent method of chromatin compaction and relaxation is through chromatin remodelers, which are crucial for effective transcription.

ATP-dependent chromatin remodelers are a class of complexes which use ATP to alter nucleosome structure and/or stability.[20] SWI2/SNF2 is one of the major groups, named after the classification of the ATPase subunit of the complexes. This group has evolutionary analogs across eukaryotic organisms from yeast to humans. The yeast version (SNI/SNF) has been shown to tightly bind both DNA and nucleosomes.[21, 22]The DNA binding by SNI/SNF is nonspecific but

also length dependent and seemingly mediated through interactions with the minor groove. Nucleosome binding by SNI/SNF is tighter (albeit only slightly) than with free DNA. Additionally, binding interactions seem to expand beyond the minor groove.[23] Interactions between the core histones and SNI/SNF is a potential explanation. Through these chromatin remodelers, initiation of transcription is made possible.

In-vitro transcription assays through 'beads on a string' chromatin structure using many known transcription factors (general, sequence-specific and coactivators) have shown elongation rates significantly slower than what was physiologically expected.[24] It follows that clearance of chromatin at the promoter and transcription start site are not sufficient for transcription. The transcriptional polymerase PolII, cannot transcribe DNA templates with nucleosomes present, even if they occur downstream of the start site.[25] This is obviously problematic since transcription occurs within chromatin, and the nucleosomes would have to be dismantled and reassembled after transcription. It was found that histones were not degraded during transcription; however, instead they were displaced on the DNA.[26] While this was performed using the prokaryotic polymerase SP6, it was believed that results could be applied to eukaryotic organisms. However, this was not quite true. The in-vitro yeast PolII would transcribe the template, but was unable to displace the nucleosome, instead only resulting in the loss of a single H2A–H2B dimer.[27] The finding provided distinct evidence that PolII uses a different mechanism than the prokaryotic SP6, and that other factors may be involved.

Previous research had identified a protein necessary for transcription through chromatin.[2] The protein was shown to **FA**cilitate **C**hromatin **T**ranscription and was given the name FACT. The known chromatin remodelers could not facilitate transcription through chromatin without the FACT protein. FACT does not belong to histone remodeler class and shows many distinctions from the remodelers, but the most obvious difference is that FACT activity is ATP-independent. While the FACT mechanism was unknown, it was postulated that FACT binds to and destabilizes the core nucleosome, acting as a histone chaperone.[2]

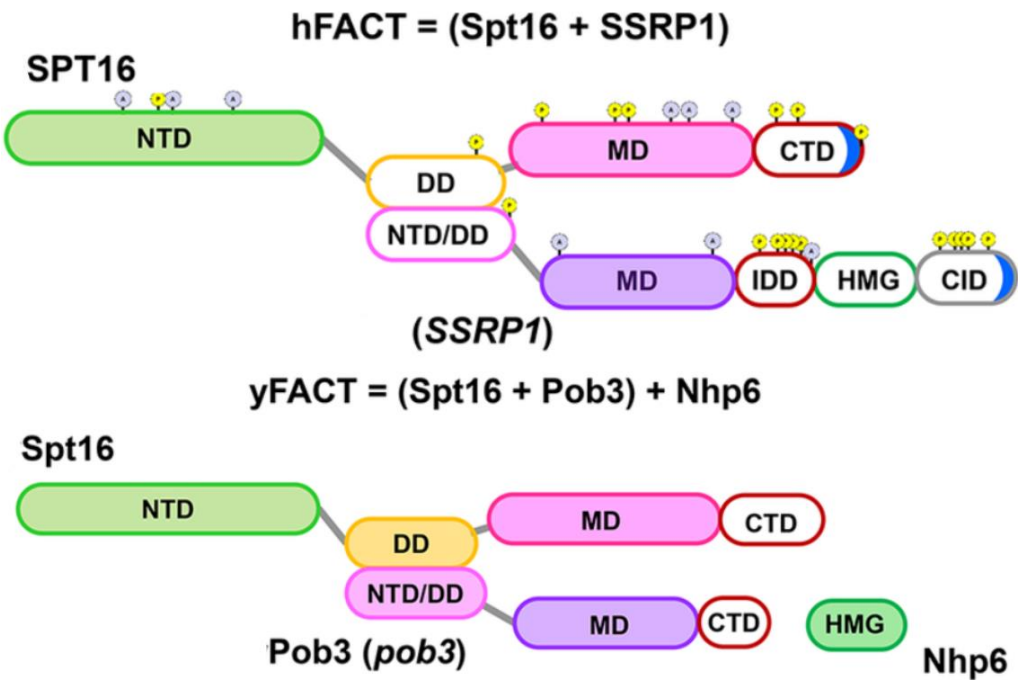

Figure 2. Structural Domains of Human FACT (hFACT) and Yeast FACT (yFACT) hFACT and yFACT domains are depicted containing the SPT16 and SSRP1 polypeptides and yeast Pob3 and Nhp6 analogs. Specific structural domains are labeled; N-terminal domain (NTD), dimerization domain (DD), middle domain (MD), intrinsically disordered domain (IDD), High mobility group (HMG), C-terminal domain (CTD) and C-terminal intrinsically disordered domain (CID). Potential hFACT Phosphorylation sites are marked as yellow circles, acetylation sites are marked as purple circles, and nuclear localization signal is marked by the blue crescents. Figure taken from [34].

During purification of human FACT, two proteins were eluted and detected by SDS-PAGE, a 140 kDa and an 80 kDa protein. FACT appeared to be a heterodimer of the two, when native gel electrophoresis showed a single 230 kDa protein. The two subunits of FACT would be denoted Suppressor of Ty 16 (SPT16) (140 kDa) and structure specific recognition protein 1 (SSRP1) (80 kDa). The role of SPT16 was theorized after mutations that disrupted H2A–H2B dimer and $(H3–H4)_2$ tetramer binding and showed the same result as SPT16 knockouts.[28] This would imply that SPT16 is involved with the H2A–H2B dimer and $(H3–H4)_2$ interface.[29]The SSRP1 subunit was found to contain a high mobility group (HMG) domain. HMG domains have been shown to both bend and bind DNA in a sequence independent manner.[30]

Both the function and structure of FACT is highly conserved in eukaryotic organisms.[31] The SPT16 protein contains four conserved domains illustrated in Fig. 2. The N-terminal domain (NTD) contains homology with aminopeptidases sequence and was found to interact with both the highly structured domains and histone tails of $(H3–H4)_2$ independently.[32] However, no

aminopeptidase activity could be found. The middle domain (MD) has two pleckstrin homology (PH1 & PH2) motifs which are used to bind $(H3–H4)_2$.[33] The C-terminal domain of SPT16 (CTD) contains a nuclear localization signal, and loss of this domain, (and a section of the MD) results in a loss of histone chaperone function in FACT.[34]

SSRP1, while also generally well-conserved, has major differences across eukaryotic organisms, due to the inclusion or removal of certain domains. The intrinsically disordered domain (IDD) and C-terminal intrinsically disordered domain (CID) are only present in mammalian SSRP1. Similar to SPT16, the SSRP1 MD contains two PH motifs that theoretically could bind $(H3–H4)_2$, although pull-down assays have shown extremely weak or no binding at all to histones.[35] While the HMG domain was known to be a DNA binding region, the MD can also bind DNA in a sequence independent manner. The CTD domains of SPT16 and Pob3 both bind H2A–H2B on overlapping sites, allowing the heterodimer to bind two H2A–H2B dimers at once.[36] This H2A–H2B dimer binding is critical for the histone chaperone function of FACT *in vivo*. Therefore, the two subunits of FACT can bind H2A–H2B dimer, $(H3–H4)_2$ and DNA, allowing the heterodimer complex to bind nucleosomes.

While the sequence of FACT along with crystal structures of certain domains (MD) are known, the overall architecture of the FACT-nucleosome complex remains a mystery. Curiously enough, human FACT (hFACT) does not bind to nucleosomes unless the H2A–H2B dimer is destabilized.[37] Further inquiry into the role of H2A–H2B in the functions of FACT, found that an H2A–H2B-FACT complex can bind the $(H3–H4)_2$ tetramer and this binding was promoted by the H2A–H2B interaction with FACT.[38] The FACT-H2A–H2B complex was found to bind to subnucleosome structures consisting of the $(H3–H4)_2$ tetramer wrapped with 79 bp of DNA to form a tetrasome, though FACT alone could not bind the tetramer. DNA was found to compete off FACT from the H2A–H2B dimer, which would logically follow as a method of depositing H2A–H2B into the tetrasome complex to release FACT. This potential mechanism is described

in Fig.3. The deposition of H2A–H2B onto an $(H3–H4)_2$ tetrasome creates the subnucleosome known as hexasome.

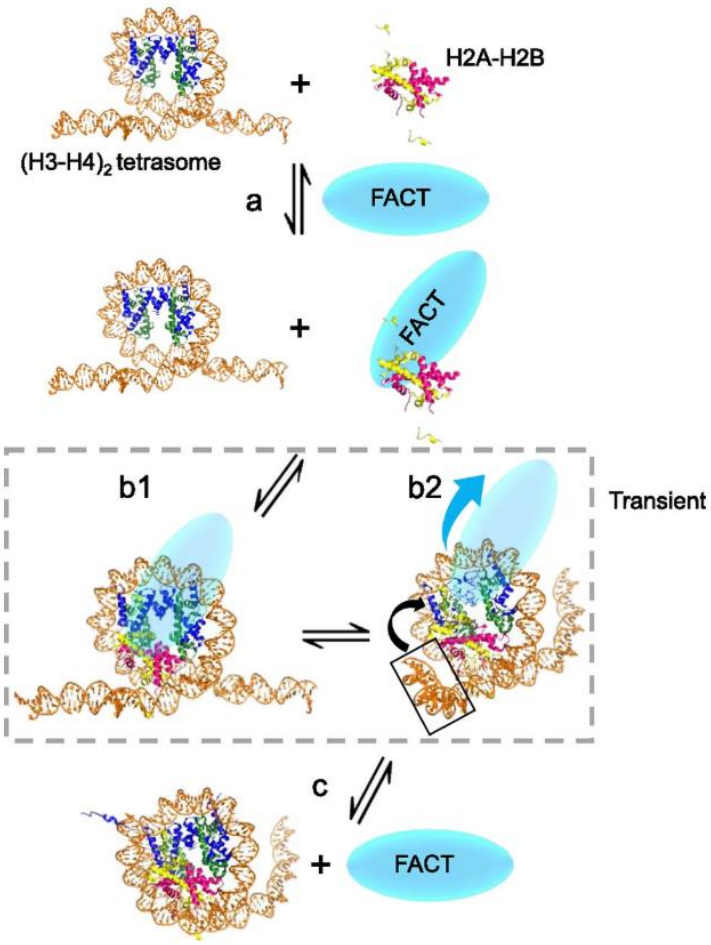

Figure 3. Proposed Mechanism of FACT in Hexasome Assembly
Fact binds free H2AB dimer (Step a.) Then FACT-(H2A–H2B) binds to $(H3–H4)_2$ tetrasome (Step b1.) DNA competes FACT off H2A–H2B (b2.) H2A–H2B is now deposited on the hexasome and FACT is released. Figure taken from [37].

The global structure of the FACT-subnucleosome complex was solved using single particle cryo electron microscopy (Cryo-EM).[1] The complex was described as a 'unicycle' where FACT resembled the seat, the $(H3–H4)_2$ tetrasome was the wheel and the H2A–H2B dimer was a peddle. The depiction of the FACT-subnucleosome is shown in Fig.4. Surprisingly, FACT was shown to have extensive interactions with the DNA wrapping the nucleosome. SPT16 dimerization domain (DD) and both MD contributed to the 19 bp of DNA binding. This was not predicted since phosphorylated FACT does not bind free DNA directly without histones present. The removal of the CTD from both SPT16 and SSRP1 domains allowed FACT to bind free DNA, hence the H2A–H2B binding to FACT relieves the inhibitory effects of the CTDs. The acidic CTD of SPT16 was found to interact with the DNA binding motif of the H2A–H2B dimer. Hydrogen deuterium exchange (HDX) indicates the L2 loop of H2A binds to the CTD of SPT16. The L2 of H2B which binds DNA and H4 interface also binds to the CTD. The inclusion of a second H2A–H2B dimer can be accomplished through the relocation of the PH2 motif on the MD of SSRP1. In the absence of the second H2A–H2B dimer, the PH2 domain forms an interface across the $(H3–H4)_2$ tetramer and interacts with H3 $\alpha 2$. This positioning prevents the inclusion of the second H2A–

H2B dimer. Upon successful addition of a second H2A–H2B dimer, the SSRP1 domain is translocated 10.9 Å and 9.5°, revealing an additional 5 bp of DNA which, along with the H3 αN, bind the L1L2 motifs on H2A–H2B.

The structural data gave insights into

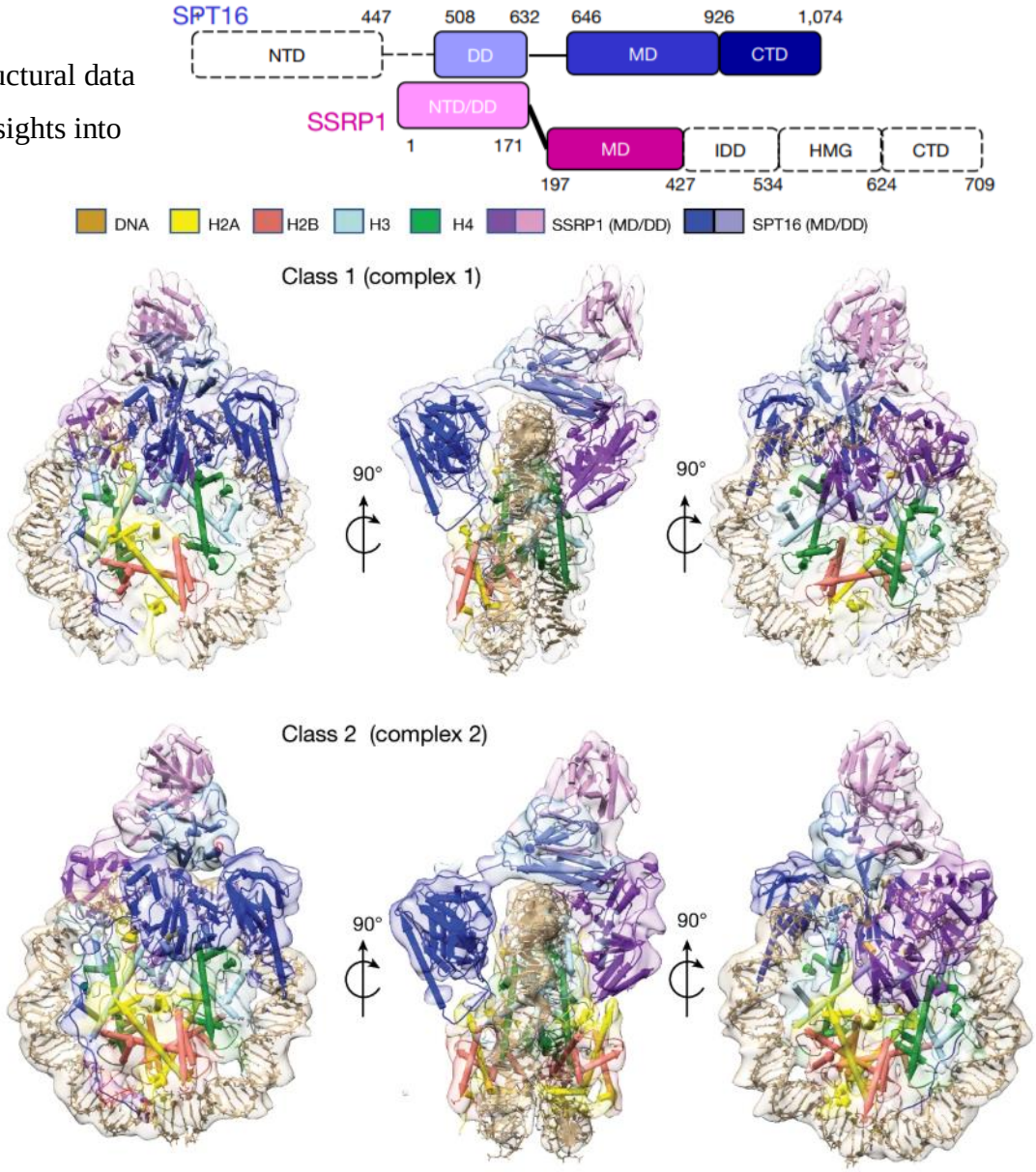

Figure 4. Cryo-EM Structures of 79 bp FACT-Subnucleosome Complex
Cryo-EM illustrates the partial FACT protein forming a complex with 79 bp subnucleosomes. Only the MD and DD domains of the SPT16 and SSRP1, and CTD of SPT16 could be captured. The class 1 complex contains a single H2A–H2B dimer whereas the class 2 complex contains both H2A–H2B dimers. Figure taken from Liu et al 2019.

the mechanism of how FACT facilitates nucleosome assembly and disassembly. However, this insight is limited in a few ways. Firstly, the resolution of the FACT-79 bp subnucleosome complex is limited to 4.9 Å. Increased resolution may uncover additional binding interactions within the complex. More importantly, a large portion of the SSRP1 subunit and the N-terminus of the SPT16 subunit are both missing. One of the missing SSRP1 domains is the HMG box which is known to bind DNA. The 79 bp DNA is likely too short to interact with the HMG domain, as well as any other missing SSRP1 domains. lack of electron density in these domains cloaks the potential interactions between these domains and their function in the FACT mechanism.

Capturing the FACT-subnucleosome complex under cryo-EM was a breakthrough in understanding how FACT bind to the subnucleosome and provided numerous mechanistic insights on how FACT facilitates assembly and disassembly of the nucleosome. Here, I aim to build upon their results by adding additional DNA to the distal side (SSRP1 side) of the subnucleosome to fully wrap the distal H2A–H2B dimer with linker DNA that extends past the subnucleosome dyad. Figure 5 depicts this asymmetric addition of DNA to the subnucleosome. Fig.5A, C both represent the 79 bp subnucleosome which only partially wraps the distal H2A–H2B dimer. Fig.5B, D depicts an extension of DNA to 100 bp which fully wraps the H2A–H2B dimer and extends towards the dyad. This DNA extension should interact with at least the HMG domain of the FACT SSRP1 subunit, which would in turn allow the capture of that domain, and potentially others through cryo-EM. The stability of this larger FACT-subnucleosome structure would likely be more stable, and lead to higher resolution images. The stability of the complex should increase for two reasons. Firstly, the wrapping of the H2A–H2B dimer by the extended DNA decreases the probability of H2A–H2B dissociation. The second reason is that the flexibility of the SSRP1 subunit would decrease if the HMG domain is tethered to the FACT subnucleosome complex.

For this investigation, two different DNA lengths were designed to reconstitute the subnucleosome. These include a 110 bp and 128 bp DNA sequence that are 10 bp and 28 bp longer than DNA shown in Fig.5B, D. Since both DNA lengths fully wrap the $(H3–H4)_2$ tetramer

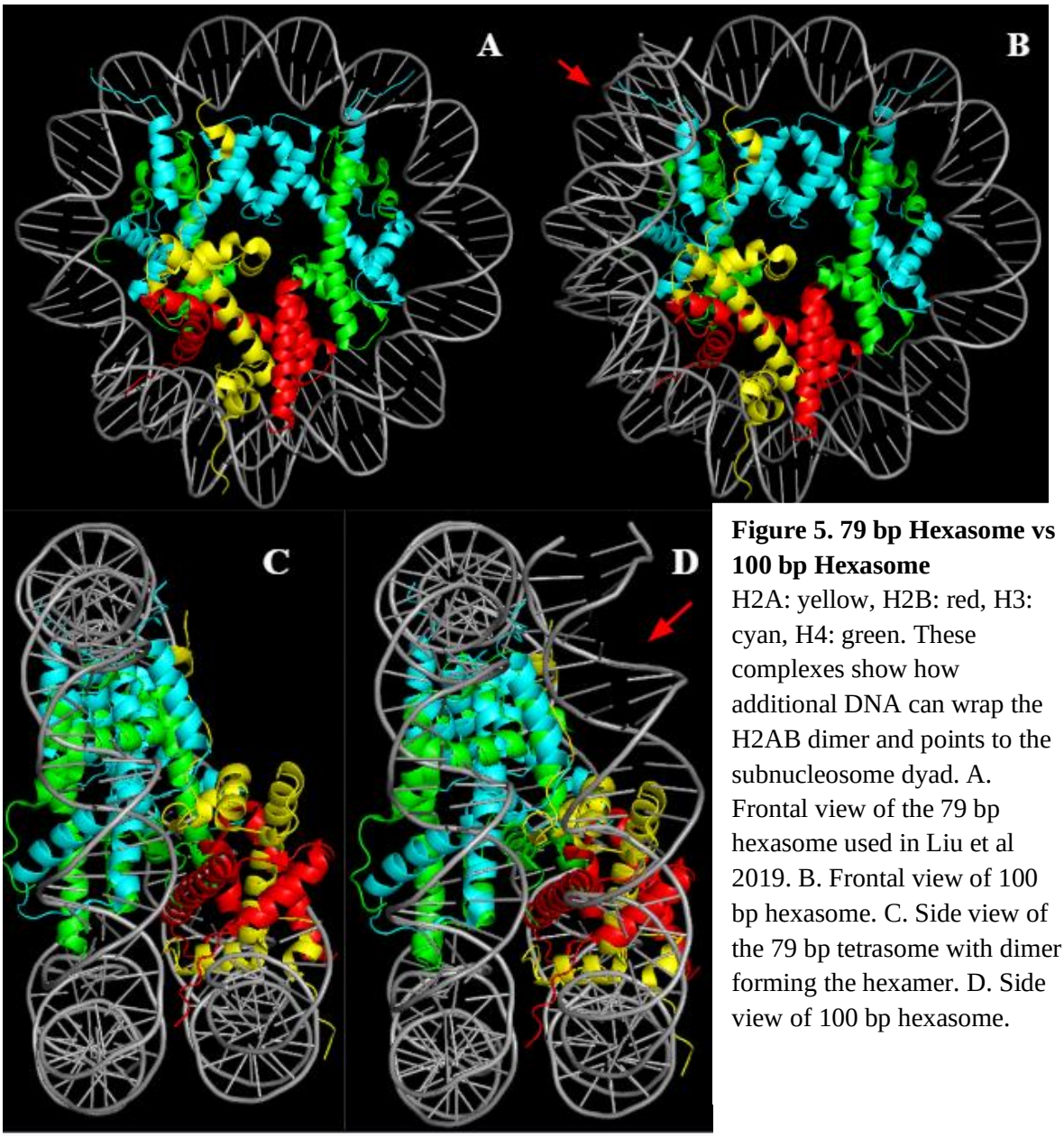

Figure 5. 79 bp Hexasome vs 100 bp Hexasome
H2A: yellow, H2B: red, H3: cyan, H4: green. These complexes show how additional DNA can wrap the H2AB dimer and points to the subnucleosome dyad. A. Frontal view of the 79 bp hexasome used in Liu et al 2019. B. Frontal view of 100 bp hexasome. C. Side view of the 79 bp tetrasome with dimer forming the hexamer. D. Side view of 100 bp hexasome.

and the H2A–H2B dimer, their reconstituted subnucleosomes will henceforth be referred as a hexasome. While the NTD of SPT16 is also missing in the cryo-EM images, this investigation will not attempt to capture it. The NTD of SPT16 is unlikely to be involved in complex formation, given its positioning above the subnucleosome dyad.[35] There is also no documented DNA interactions with the NTD, and elongation of DNA on the proximal and distal side has been shown to compete FACT off the complex.[1]

Methods and Materials

DNA Preparation

DNA used to reconstitute the subnucleosomes was prepared by amplifying regions of the 601 'Widom' nucleosome positioning sequence.[39] This was done by designing primers to amplify the parts of the widom sequence with additional linker DNA of various lengths. DNA was amplified through PCR with PFU enzyme and purified using a prepacked HiTrap Q HP column with buffer QA containing 20 mM Tris-Cl and 1 mM EDTA and buffer QB containing 20 mM Tris-Cl, 1 mM EDTA, and 1 M NaCl.

Hexasome Reconstitution

Sub-Nucleosome reconstitution was done by combining molar ratios of 1:1, 1:1.2, 1:0.9 and 1: 0.8 of DNA to $(H3–H4)_2$ tetramer, and H2A–H2B dimer. The solution was diluted to 1 μM of DNA with a buffer containing 10 mM Tris-Cl, 2M NaCl, 1 mm EDTA, and 1 mm DTT. The Solution was dialyzed for 16+ hours at 4°C into a solution with only 50 mM NaCl. Reconstituted hexasomes were characterized using native PAGE.

FACT Expression and Purification

The purification of human FACT was adapted from published work with minor changes.[40] Six histidine residue tag (6XHis tag) was fused on the N-terminus of SPT16 and FACT was purified first using 5 mL prepacked HisTrap HP column, followed by 5 mL HiTrap Q HP column. FACT was stored in 150 mM NaCl, 20 mM Tris, pH 7.8, 10% glycerol, 0.01% CHAPS, 0.01% octyl glucoside, and 1 mM TCEP. All columns were purchased from GE healthcare. The purified FACT protein was visualized through SDS-PAGE and stained using the BlazinBlue protein stain.

Histone Refolding

Purified human histone proteins H2A, H2B, H3, and H4 were first unfolded in a 2 mL buffer containing 20 mM Tris-Cl, 6 M guanidinium HCl, and 5 mM DTT at pH 7.5 for thirty minutes. Equimolar amounts of histones in the unfolding buffer were transferred to a separate tube. Refolding buffer was added to each tube at the ratio of one mL per one mg of histone protein. The histone solutions were then dialyzed for 18 hours at 4 ℃ into 2 L of a refolding buffer, consisting of 2 M NaCl, 10 mM Tris-Cl, 0.5 M EDTA, and 5 mM 2-mercaptoethanol. The refolding buffer was replaced twice during dialysis. Refolded histones were again purified using the Superdex 200 size exclusion column in 2 M NaCl, 10 mM Tris-Cl, 0.5 M EDTA, and 5 mM 2-mercaptoethanol. The refolded histones were visualized using SDS-PAGE and stained using the BlazinBlue protein stain.

FACT-Hexasome Complex Formation

Purified FACT (1 μM) was combined with equimolar amounts of the purified H2A–H2B dimer and incubated at room temperature for 5 minutes. Then equimolar or 50% equimolar 110 or 128 bp hexasomes were added to pre-mixed FACT-H2A–H2B in a buffer containing 20 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM EDTA and 1 mM DTT. Complex formation was evaluated using 5% native PAGE and stained for at least thirty minutes in SYBR Gold DNA stain.

Sedimentation Velocity Analytical Ultracentrifugation (SV-AUC)

SV-AUC was used to evaluate the homogeneity of the FACT-Hexasome complexes in solution to infer their stability. SV-ACU procedure was adapted from Wang et al 2018.[38] SV-AUC with absorbance optics was performed and spun at 30–35,000 rpm at 20°C in a Beckman XL-A ultracentrifuge, using an An60Ti rotor. Partial specific volumes of the samples were determined using UltraScan 3 version 2.0 (Department of Biochemistry, The University of Texas Health Science Center at San Antonio). Time invariant and radially invariant noise was subtracted by 2-dimensional-spectrum analysis followed by genetic algorithm refinement and Monte Carlo analysis to resolve the sedimentation coefficient s. Sedimentation coefficient were converted to Svedberg's and plotted using van Holde-Weischet plots. (Demeler & van Holde, 2004)[41].

Analyses were performed using Ultrascan 3 version 2.0 and distributions were plotted using GraphPad Prism.

Results

To reconstitute the subnucleosome and the FACT-subnucleosome complexes, recombinantly expressed human histone proteins H2A, H2B, H3, and H4 were refolded and purified. The refolding was necessary to achieve dimerization of the histone proteins. Refolding of the histone proteins was accomplished by first unfolding them in a 6 M guanidinium buffer, before dialyzing in a 2 M NaCl refolding buffer. The refolded histone proteins were purified using a S-200 size exclusion column and visualized on an SDS-PAGE, as shown in Fig. 6. H2A and H2B proteins visualized in Fig.6A ran just above 15 kDa on the SDS-gel. The histone proteins bands align with the documented sizes of human H2A and H2B. These are 14.1 kDa and 15.6 kDa respectively based on known literature.[41]

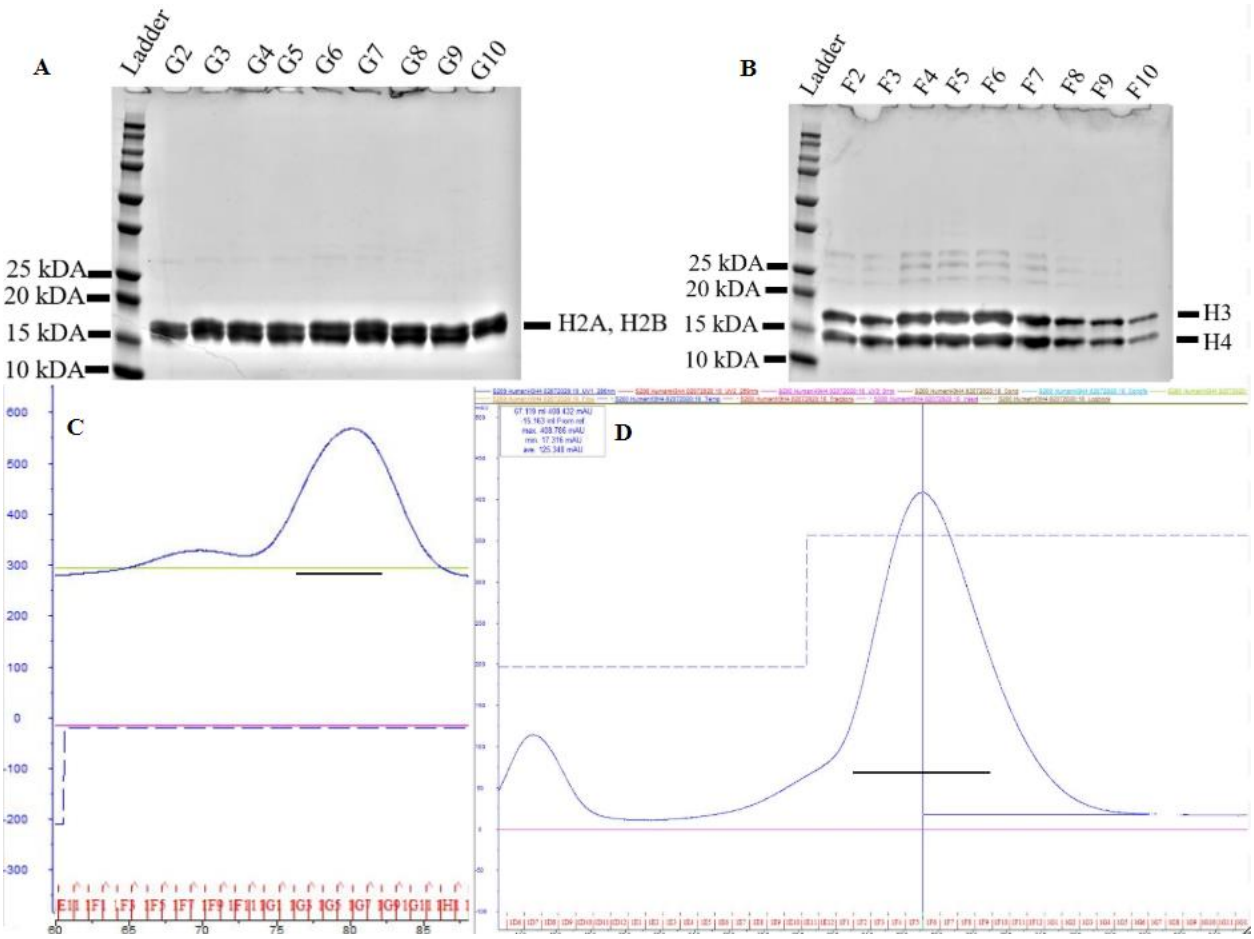

Figure 6. Histone Refolding and Purification

Histone proteins were refolded and purified and visualized on an SDS gel. H2A and H2B, as well as the H3 and H4 fractions are seen on the SDS gel (A), (B) at the appropriate lengths. H2A and H2B, along with H3 and H4 A280 peaks are shown during purification by S-200 size exclusion (C), (D). The black lines represent the fractions shown in the respective gels.

While minor impurities were present above 25 kDa, the minute concentrations relative to the histone proteins renders them negligible. The H3 and H4 proteins visualized in Fig.6B ran above 15 kDa and below 15 kDa respectively. Human H3 and H4 are known to be 15.4 kDa and 11.3 kDa respectively.[42] There were more impurities visualized in the H3, H4 purification, most around 25 kDa. Despite this, their relative concentrations were too small to be a concern. The ratio of H2A to H2B and H3 to H4 from the SDS gel are all visibly equal. Therefore, we assume that these histones dimerize into their dimeric (H2A–H2B) and tetrameric $(H3–H4)_2$ forms in solution. Histone proteins were then stored at -80℃ in 100 μM aliquots until usage.

Human FACT protein was expressed in insect cells and purified by HisTrap HP and HiTrap Q columns. The purified FACT protein was visualized through SDS-PAGE as shown in Figure 7. Both FACT subunits SPT16 and SSRP1 could be seen on the gel with distinct bands. These subunits ran just below 150 kDa and above 75 kDa, matching their reported size of 140 kDa and 80 kDa for SPT16 and SSRP1. Purification of the FACT protein left only trace amounts of other proteins seen mostly below 75 kDa. The FACT protein itself may have also been slightly degraded, contributing to the extra bands. However, the FACT subunit bands are more distinct than the other bands indicating a large majority of FACT subunits present as a heterodimer (1:1).

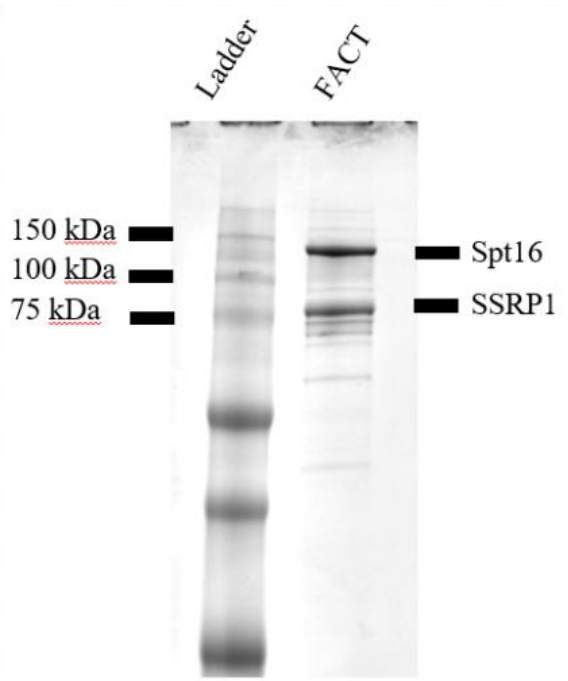

Figure 7. FACT Purification and Visualization Purified human FACT protein is visualized by SDS-PAGE. The two subunits; SPT16 and SSRP1 are both visible at their appropriate molecular weight. Other contaminating proteins can be seen along with slight FACT subunit degradation

Hexasomes were reconstituted for FACT binding assays. Reconstitution was done by dialyzing the subnucleosome components from 2 M NaCl to 50 mM NaCl overnight at 4°C. 110 and 128 bp DNA were used to wrap the histone proteins to create the hexasome. Figure 8 illustrates the shifting of the DNA through the hexasome formation made from 1:1 and 1:1.2 DNA to histones. The gel was stained for 30 minutes in SYBR gold to visualize the DNA. Two hexasome samples were made: a 110 bp hexasome consisting of the $(H3–H4)_2$ tetramer and H2A–H2B dimer wrapped by 110 bp of DNA, and a 128 bp consisting of the $(H3–H4)_2$ tetramer and H2A–H2B dimer wrapped by 128 bp of DNA. A large fraction of the 110 bp hexasome aggregated during reconstitution and could not be used. This results in a weaker 110 bp hexasome band due to less material added. Similar aggregation was not seen in the 128 bp hexasome sample. The 128 bp hexasome was shown to be largely homogenous, though faint bands under 200 bp are visible along with a general smear. The 110 bp hexasome appears to be completely homogenous with a well-defined band at ~175

bp. Though the 110 bp hexasome is around 10-fold less concentrated than the 128 bp hexasome, so comparison is limited. The nucleosome control shows two bands corresponding to the nucleosome and a 147 bp hexasome.

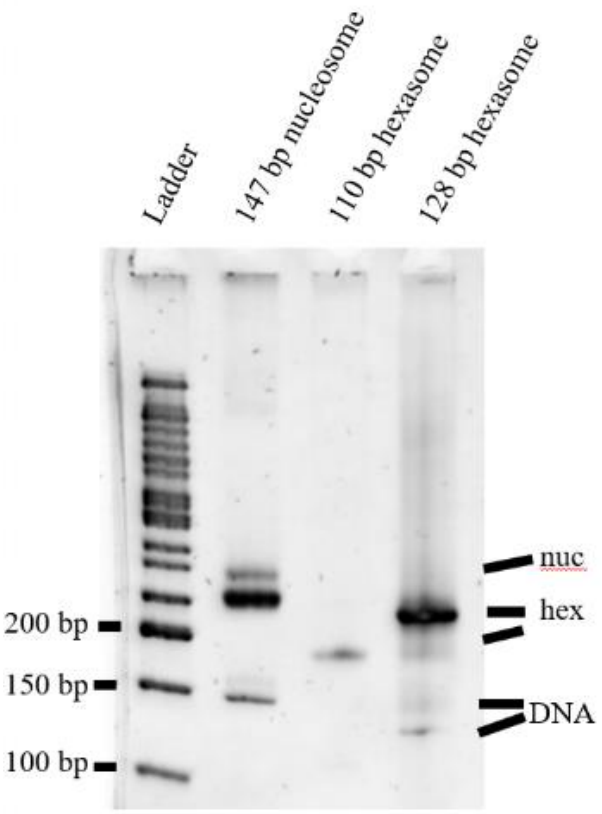

Figure 8. Subnucleosome Reconstitution Visualized by Native Gel.
Hexasomes with 110 and 128 bp DNA were reconstituted with one equivalent of H2A–H2B dimer, and $(H3–H4)_2$ tetramer. Samples were visualized through native PAGE. A previously reconstituted 147 bp nucleosome was run for comparison.

FACT-hexasome complexes were formed by first combining an equimolar amount of FACT with H2A–H2B dimer and incubating at room temperature for 5 minutes. Then, equimolar amounts of 110 and 128 bp hexasome were added to the FACT-(H2A–H2B). The visualization of the complexes through native PAGE is shown in Figure 9. The 110 and 128 bp hexasomes are seen to be shifted with the addition of the FACT-(H2A–H2B), indicative of the FACT-hexasome complex formation. The 128 bp hexasome alone exhibits a strong single band above 200 bp, which depicts the hexasome, and a smaller band in between 150 and 110 bp which represents the free DNA. The 110 bp hexasome alone shows a large band at 200 bp, representing the actual hexasome, along with a much fainter band just above 110 bp showing the free DNA. There is another faint peak above the 110 hexasome which could be the formation of a larger complex through nonspecific binding of DNA and histones. The FACT-hexasome complexes are seen to be largely shifted. While a majority of the 110 bp hexasome seems to be shifted by FACT-(H2A–H2B), the 128 bp hexasome still shows a strong band at the free hexasome mark. The 110 bp FACT-hexasome complex shows some free hexasome remaining, though not as much as the 128 bp. The shifted complexes are also not homogenous. There seems to be two or three separate, yet closely-packed bands of around equal strength. This pattern is seen for both

the 110 and the 128 bp FACT-hexasome complexes. While two bands are expected, as they are the class 1 and class 2 complexes seen in Fig.4, the third band is representative of degradation or nonspecific binding.

Attempts to improve the homogeneity of the FACT-hexasome complexes was done by increasing the ratio of FACT-(H2A–H2B) to 1.5 that of the hexasome. The resulting 1.5:1 FACT-hexasome complexes were visualized using native PAGE and are shown in Figure 10. The 1.5:1 110 bp FACT-hexasome complex exhibited nearly complete hexasome shifting, indicated by the loss of the 110 bp hexasome band. Both the 1:1 FACT-hexasome complexes seem to have become fainter, which was expected after a day of storage at 0℃. The 128 bp 1.5:1 FACT-hexasome complex seems to have shifted the free hexasome more than 1:1. Though this increase is more apparent in the 110 bp complex. The homogeneity seems to have improved for the 128 bp 1.5:1 FACT-hexasome complex, as indicated by the strong single band for the complex. This contrasts well with the previous result of three equally strong bands for the 128 bp 1:1 FACT-hexasome complex. The 110 bp 1.5:1 FACT-hexasome complex did not see the same increase in homogeneity. While there are now two clear bands, there is still a smear, though this does appear to be a small improvement. The lack of any free hexasome in the 110 bp 1.5:1 complex seems to be the main advantage of a higher ratio for the 110 bp FACT-hexasome complex formation.

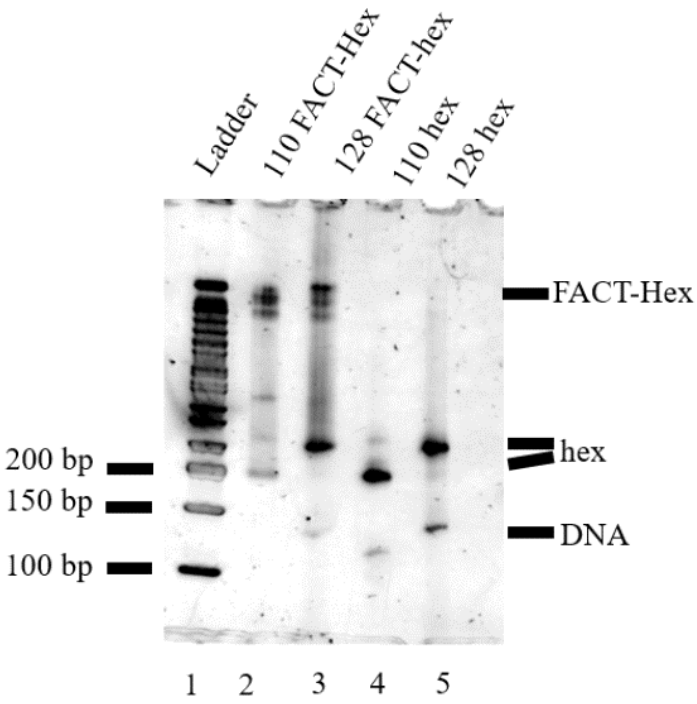

Figure 9. FACT-Hexasome Complex Visualized on Native Gel.
110 and 128 bp hexasomes were combined with FACT-dimer at a 1:1 ratio to generate FACT-hexasome complexes, which were run on a 5% native gel. The gel was stained using SYBR Gold to visualize the DNA and DNA-protein complexes.

Large scale versions of the FACT-hexasome complexes using the 1.5:1 FACT-(H2A–H2B) to hexasome ratio were made for SV-AUC analysis. The complexes were constituted in the same manner as before and visualized through native PAGE prior to spinning, as seen in Figure 11. The large scale of FACT-hexasome complexes look identical to the previously made 1.5:1 complex on the native gel, showing that complex formation is possible at higher concentrations. The samples including the 110 bp hexasome, 110 bp complex, 128 bp hexasome and 128 bp complex were then diluted down to 300 μM and analyzed through AUC. Analysis of the SV-AUC is represented using a van Holde–Weischet plot as shown in Fig.11B. The sedimentation coefficient is represented using the standard Svedberg units S ($_{20,w}$). Both hexasome seemed to sediment at similar values of 9 S. Addition of FACT-(H2A–H2B) (which sediments at 8.3S)[37] increases the sedimentation to around 13 and 14 S for the 110 bp complex and 128 bp complex respectively. Homogeneity of the 128 bp free hexasome was higher than the 110 bp hexasome as indicated by the steeper sedimentation curve. The AUC 'tails' are considerably larger for the 110 bp hexasome and complex. This is indicative of increased free components such as histones, DNA and FACT. The increased tails in the 110 bp sample indicate a decrease in stability. FACT-(H2A–H2B) addition appears to increase the homogeneity of the 110 bp complex, as seen by the decrease in the tail. However, the 110 bp complex still shows lower homogeneity than the 128 bp complex. The SV-AUC results as a whole show that both the hexasome and the complex are moderately homogeneous in solution, with the better homogeneity and stability for the 128 bp version.

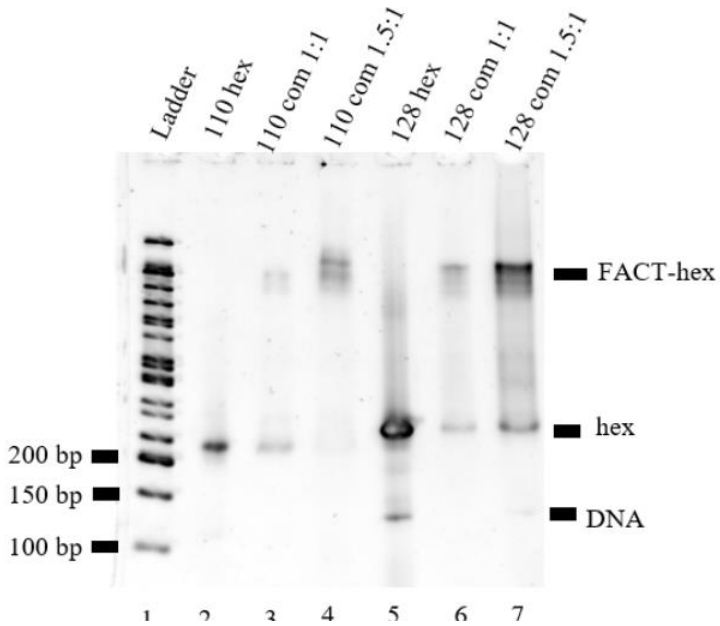

Figure 10. 1.5:1 FACT-Hexasome Complex Formation
FACT-hexasome complexes were reconstituted with a 1.5 FACT-dimer to hexasome ratio and visualized on a 5% native gel with SYBR Gold DNA stain. The 1:1 FACT hexasome complexes made the day prior, and the free hexasomes were shown for comparison.

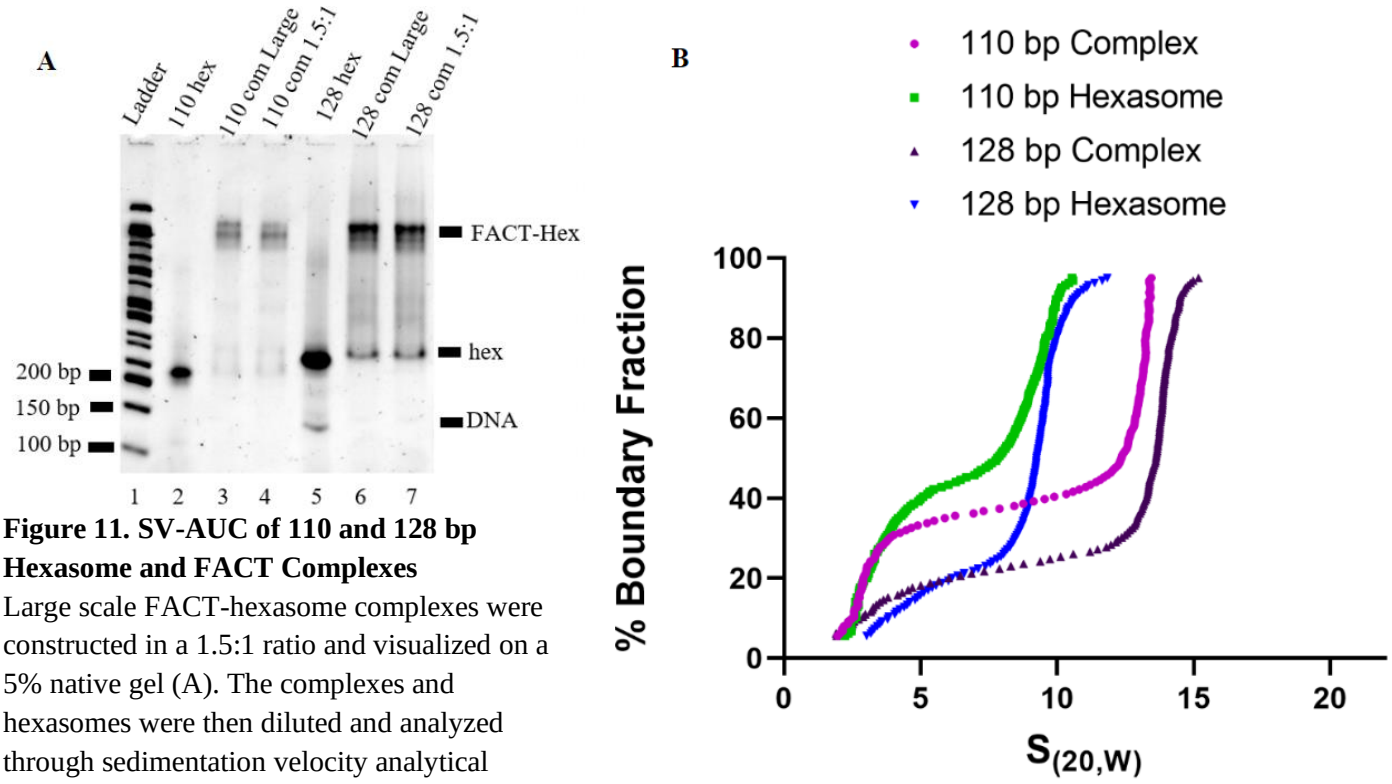

Figure 11. SV-AUC of 110 and 128 bp Hexasome and FACT Complexes

Large scale FACT-hexasome complexes were constructed in a 1.5:1 ratio and visualized on a 5% native gel (A). The complexes and hexasomes were then diluted and analyzed through sedimentation velocity analytical ultracentrifugation (SV-AUC). SV-AUC results were shown using a van Holde-Weischet plot, which plotted the boundary fraction percentage against the sedimentation coefficient S (B).

Discussion

Refolding and purification of human histones resulted in complexes that ran at the expected size and are bound to free DNA. Small contaminants were found during visualization through SDS PAGE after histone purification as seen in Fig.6. These contaminating proteins were vastly less concentrated than the histone products. While the refolded histone proteins should dimerize, it is highly unlikely that the contaminating proteins are dimerized histones. While the contaminating proteins run at around twice the molecular weight of the purified histones, the SDS buffer denatures the histones, preventing dimerization. Since the individual histone proteins seem to be at equimolar quantities, effective dimerization in their native form is highly likely. The formation of subnucleosome complexes with free DNA verifies that the functionality of the histones remains intact.

Reconstitution of both the 110 and 128 bp hexasomes were successful. Comparison of the hexasomes against a 147 bp nucleosome indicates that the hexasome does not contain a second H2A–H2B dimer. The hexasome does not shift as high as the nucleosome on the native gel (Fig.8) also shows no aggregation. This indicates that the vast majority of the histone components were incorporated. If the histone proteins instead formed a partially wrapped nucleosome, then two distinct bands would appear. Formation of a nucleosome would use two equivalents of H2A–H2B dimer, and since H2A–H2B are in a 1:1 ratio with the $(H3–H4)_2$ tetramer, octamer formation would deposit tetrasomes. The nucleosome and tetrasome would then form two distinct bands on the native gel. Since there is only a single distinct hexasome band, nucleosome formation is highly unlikely. Another possibility is the formation of the tetrasome with no H2A–H2B dimer interaction. Formation of the tetrasome would result in

H2A–H2B dimer crashing out completely, or an extra H2A–H2B-DNA band. Neither of these were seen, further supporting the assertion that the histone DNA complex was a hexasome. However, aggregation of the hexasome structure was apparent at high histone concentrations. This aggregation was only seen in 110 bp hexasome reconstituted above 1:1 ratios of histone to DNA. Reconstitution of 110 bp hexasomes with lower than one equivalent of histone proteins worsened yield with no further reduction of aggregation. 128 bp hexasomes exhibited limited to no aggregation in both the 1.2:1 and 1:1 fraction. This disparity is likely due to the improved wrapping of the H2A–H2B dimer due to the additional 18 bp of DNA in the 128 bp hexasome.

FACT-hexasome complex formation was most successful when using 1.5:1 ratio of FACT-(H2A–H2B) to hexasome. 1.5:1 FACT-(H2A–H2B) completely shifted the 110 bp hexasome whereas the 1:1 FACT-(H2A–H2B) had free 110 bp hexasome present. Higher FACT-(H2A–H2B) ratios improved the 128 bp complex homogeneity, albeit only slightly. The improvements are likely seen due to a lower binding affinity of FACT to the hexasomes. An important distinction must be made regarding the double bands present in the FACT complex. These double bands are expected because they represent the class 1 and class 2 depicted in Fig. 4. The presence of the second H2A–H2B dimer is likely influenced by the FACT-(H2A–H2B) ratio, with increased FACT favoring class 2. It does not matter which class predominates, since capturing the hidden SSRP1 domains is unaffected by the presence of a second H2A–H2B dimer on the SPT16 side. Any additional bands however represent nonspecific interaction and their removal is important for constituting a homogenous complex.

The van Holde-Weischet plot of the SV-AUC results indicate the FACT-(H2A–H2B) can bind to both 110 and 128 bp hexasomes. The free hexasomes exhibit modest heterogeneity, with the 110 bp hexasome being the most heterogeneous. This is indicated by the flatter sedimentation curve and larger tail. However, addition of FACT-(H2A–H2B) seems to improve homogeneity. The FACT complexes are largely homogenous which is evidenced by the uniform sedimentation curve. In addition, the tail seems to be smaller for the 110 bp FACT complex relative to the hexasome, which indicates an improvement to stability. The 128 bp complex also sees a small decrease in tail size, though this change is less noticeable. While homogeneity of the complexes is similar for both 110 and 128 bp, the difference in tail length indicates that the 128 bp complex is more homogeneous.

SV-AUC of 79 bp subnucleosome with fluorescently labelled H4 was previously performed in Wang et al 2018.[38] This investigation seeks to improve upon the complexes of the FACT-(H2A–H2B)-79 bp tetrasome shown in Fig.4. For successful cryo-EM images, the FACT complex must have the same or improved stability. The 79 bp FACT complex appears to have similar homogeneity due to its sharp sedimentation curve. However, the key difference is the lack of a tail. This is exaggerated by the lack of detection of free H2A, H2B, H3, DNA and FACT since only H4 is fluorescently labeled. While exaggerated, the lack of a tail indicates a highly homogeneous and stable complex. Therefore, further optimization of complex formation will likely be required for successful FACT-hexasome cryo EM.

Going forward, improving the homogeneity of the FACT-hexasome complex is essential. To investigate the FACT-hexasome structure through cryo-EM requires a stable, homogenous

complex. Improving the complex stability may be achieved by optimizing the FACT-(H2A–H2B) and hexasome ratio. Increasing the FACT-(H2A–H2B) ratio to 1.5:1 showed improvements in Fig.10. While ratios above 2.:1 did not improve stability (data not shown), there is room for optimization. For the DNA lengths, the better substrate candidate for the FACT binding is the 128 bp hexasome. This structure is more homogenous than the 110 bp hexasome both free and when bound to FACT. There may however be room to optimize the length of DNA wrapping the hexasome. While hexasome reconstitution using 143 bp and 99 bp DNA failed to create a homogenous hexasome and, there may be a length between or slightly above/below 128 and 110 bp that is better suited for hexasome or complex formation. Overall, this research demonstrates that formation of stable FACT-hexasome complexes are possible. With further optimization, cryo-EM imaging of the complex could resolve previously hidden SSRP1 domains, and further our collective understanding of how FACT facilitates nucleosome assembly and disassembly.

Conclusion

Reconstitution of 110 and 128 bp hexasomes formed relatively stable substrates that could bind to FACT-(H2A–H2B) to form homogenous complexes. Both the 128 bp hexasome and complex demonstrated an increased homogeneity and stability over the 110 bp version on native gels and in SV-AUC. While FACT complex formation should be further optimized for increased homogeneity, these results describe a strong substrate candidate for FACT complex in cryo-EM. Depositing these complexes onto the cryo-EM grids will now be the main challenge and maximizing the stability of the complexes will be paramount to a successful image. Using the 128 bp complex, the HMG domain of FACT SSRP1 will likely be captured, then a nearly complete model of FACT binding to subnucleosomes will be attained.

References

1. Liu, Y. *et al.* FACT caught in the act of manipulating the nucleosome. *Nature* **577**, 426–431 (2020).
2. Orphanides, G., LeRoy, G., Chang, C.-H., Luse, D. S. & Reinberg, D. FACT, a Factor that Facilitates Transcript Elongation through Nucleosomes. *Cell* **92**, 105–116 (1998).
3. Belotserkovskaya, R. FACT Facilitates Transcription-Dependent Nucleosome Alteration. *Science* **301**, 1090–1093 (2003).
4. Fleyshman, D. *et al.* Level of FACT defines the transcriptional landscape and aggressive phenotype of breast cancer cells. *Oncotarget* **8**, (2017).
5. Bazzi, M. & Badros, A. Multiple myeloma: Implementing signaling pathways and molecular biology in clinical trials. *Cancer Biology & Therapy* **10**, 830–838 (2010).
6. Wolffe, A. *Chromatin: structure and function*. (Academic Press, 1998).
7. 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).
8. Smith, M. M. Histone structure and function. *Current Opinion in Cell Biology* **3**, 429–437 (1991).
9. Alva, V., Ammelburg, M., Söding, J. & Lupas, A. N. On the origin of the histone fold. *BMC Struct Biol* **7**, 17 (2007).
10. Thoma, F. & Koller, Th. Influence of histone H1 on chromatin structure. *Cell* **12**, 101–107 (1977).
11. Allan, J., Hartman, P. G., Crane-Robinson, C. & Aviles, F. X. Nature 288, 675–679 (1980).
12. Woodcock, C. L. & Ghosh, R. P. Chromatin Higher-order Structure and Dynamics. *Cold Spring Harbor Perspectives in Biology* **2**, a000596–a000596 (2010).
13. Teif, V. B. & Rippe, K. Predicting nucleosome positions on the DNA: combining intrinsic sequence preferences and remodeler activities. *Nucleic Acids Research* **37**, 5641–5655 (2009).
14. Dietzel, S. The 3D Positioning of ANT2 and ANT3 Genes within Female X Chromosome Territories Correlates with Gene Activity. *Experimental Cell Research* **252**, 363–375 (1999).

15. Müller, W. G., Walker, D., Hager, G. L. & McNally, J. G. Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. *Journal of Cell Biology* **154**, 33–48 (2001).
16. Burke, L. J. *et al.* CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. *EMBO J* **24**, 3291–3300 (2005).
17. Bancaud, A. *et al.* Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J* **28**, 3785–3798 (2009).
18. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128**, 693–705 (2007).
19. Karlic, R., Chung, H.-R., Lasserre, J., Vlahovicek, K. & Vingron, M. Histone modification levels are predictive for gene expression. *Proceedings of the National Academy of Sciences* **107**, 2926–2931 (2010).
20. Vignali, M., Hassan, A. H., Neely, K. E. & Workman, J. L. ATP-Dependent Chromatin-Remodeling Complexes. *Molecular and Cellular Biology* **20**, 1899–1910 (2000).
21. Lorch, Y., Cairns, B. R., Zhang, M. & Kornberg, R. D. Activated RSC–Nucleosome Complex and Persistently Altered Form of the Nucleosome. *Cell* **94**, 29–34 (1998).
22. Quinn, J., Fyrberg, A. M., Ganster, R. W., Schmidt, M. C. & Peterson, C. L. DNA-binding properties of the yeast SWI/SNF complex. *Nature* **379**, 844–847 (1996).
23. Cote, J., Quinn, J., Workman, J. & Peterson, C. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* **265**, 53–60 (1994).
24. Ucker, D. S. & Yamamoto, K. R. Early events in the stimulation of mammary tumor virus RNA synthesis by glucocorticoids. Novel assays of transcription rates. *J. Biol. Chem.* **259**, 7416–7420 (1984).
25. Knezetic, J. A. & Luse, D. S. The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell* **45**, 95–104 (1986).
26. Felsenfeld, G., Clark, D. & Studitsky, V. Transcription through nucleosomes. *Biophysical Chemistry* **86**, 231–237 (2000).

27. Kireeva, M. L. *et al.* Nucleosome Remodeling Induced by RNA Polymerase II. *Molecular Cell* **9**, 541–552 (2002).
28. Santisteban, M. S., Arents, G., Moudrianakis, E. N. & Smith, M. M. Histone octamer function in vivo: mutations in the dimer-tetramer interfaces disrupt both gene activation and repression. *EMBO J.* **16**, 2493–2506 (1997).
29. Belotserkovskaya, R., Saunders, A., Lis, J. T. & Reinberg, D. Transcription through chromatin: understanding a complex FACT. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1677**, 87–99 (2004).
30. Thomas, J. O. & Travers, A. A. HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends in Biochemical Sciences* **26**, 167–174 (2001).
31. Formosa, T. Changing the DNA Landscape: Putting a SPN on Chromatin. in *Protein Complexes that Modify Chromatin* (ed. Workman, J. L.) vol. 274 171–201 (Springer Berlin Heidelberg, 2003).
32. Stuwe, T. *et al.* The FACT Spt16 "peptidase" domain is a histone H3–H4 binding module. *PNAS* **105**, 8884–8889 (2008).
33. Kemble, D. J. *et al.* Structure of the Spt16 Middle Domain Reveals Functional Features of the Histone Chaperone FACT. *J. Biol. Chem.* **288**, 10188–10194 (2013).
34. Gurova, K., Chang, H.-W., Valieva, M. E., Sandlesh, P. & Studitsky, V. M. Structure and function of the histone chaperone FACT - Resolving FACTual issues. *Biochim Biophys Acta Gene Regul Mech* (2018) doi:10.1016/j.bbagrm.2018.07.008.
35. Zhang, W. *et al.* Crystal Structure of Human SSRP1 Middle Domain Reveals a Role in DNA Binding. *Sci Rep* **5**, 18688 (2016).
36. Kemble, D. J., McCullough, L. L., Whitby, F. G., Formosa, T. & Hill, C. P. FACT Disrupts Nucleosome Structure by Binding H2A–H2B with Conserved Peptide Motifs. *Molecular Cell* **60**, 294–306 (2015).
37. Tsunaka, Y., Fujiwara, Y., Oyama, T., Hirose, S. & Morikawa, K. Integrated molecular mechanism directing nucleosome reorganization by human FACT. *Genes Dev.* **30**, 673–686 (2016).

38. Wang, T. *et al.* The histone chaperone FACT modulates nucleosome structure by tethering its components. *Life Sci. Alliance* **1**, e201800107 (2018).

39. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *Journal of Molecular Biology* **276**, 19–42 (1998).

40. Winkler, D. D., Muthurajan, U. M., Hieb, A. R. & Luger, K. Histone Chaperone FACT Coordinates Nucleosome Interaction through Multiple Synergistic Binding Events. *J. Biol. Chem.* **286**, 41883–41892 (2011).

41. Demeler, B. & van Holde, K. E. Sedimentation velocity analysis of highly heterogeneous systems. *Analytical Biochemistry* **335**, 279–288 (2004).

42. Uniprot: A worldwide hub of protein knowledge. doi:https://doi.org/10.1093/nar/gky1049.