Quantifying nucleosome dynamics and protein binding with PIE-FCCS and spFRET

Martens, C.L.G.

CHAPTER 5

H3K36me3 DECREASES NUCLEOSOME STABILITY AND INCREASES LEDGF AFFINITY IN CHROMATIN

Posttranslational modifications (PTMs) of the histone tails are widely associated with modulation of gene expression. By altering the local charge and structure of the amino acids in histone tails through acetylation, methylation or phosphorylation, these modifications change binding affinity of proteins for nucleosomal DNA or the histone tails themselves. The trimethylation of lysine 36 in the H3 tail, H3K36me3, is associated with DNA repair and transcription as well as HIV viral DNA integration. In the latter case protein LEDGF/p75 is hijacked by a pre-integration complex (PIC) containing the viral DNA and used as a bridge to the host DNA. The interaction mechanism between LEDGF/p75, PIC and DNA is not yet completely elucidated, as is the effect of H3K36me3 on nucleosomal dynamics and stability.
We combined the results of PIE-FCCS and PIE-FCS-FRET measurements to quantify the effect of H3K36me3 on nucleosome dynamics, stability, and the affinity of LEDGF/p75 to bare and nucleosomal DNA. We used wild-type and an AT-hook deficient variant of the protein to show that the dominant factor in affinity is the H3K36me3 modification. We also quantified the decrease in nucleosomal stability caused by H3K36me3. Lastly, we propose a possible interaction mechanism based on the decrease in FRET efficiency and binding affinity measurements.

This chapter is a collaboration with:
T. Brouns, W. Frederickx, S. De Feyter, W. Vanderlinden, KU Leuven.

5.1 Introduction

The nucleosome is the first level of DNA compaction in eukaryotic cells. Compacted DNA is sterically occluded from DNA-binding proteins involved in processes like transcription and repair[203][174]. DNA accessibility is modulated by spontaneous reversible unwrapping (nucleosome breathing) and DNA remodelers[39][50], the latter being proteins and enzymes attracted to specific DNA sequences or chemical modifications of histone tails. The posttranslational modifications (PTMs) of the amino acids comprising histone tails act as markers for a myriad of DNA processes[204] [205] [206] [207] [208]. Through covalent binding a chemical functional group is added or modified, thereby changing the local charge density or conformation of the amino acid, or both. The modifications alter the affinity for proteins to the histone tail and consequently to the nearby compacted DNA[209][210].

Although the role of PTMs in cellular processes is understood on a causal level, the underlying mechanisms are less clear[211][212][213][214]. This hiatus in mechanical insight is largely due to experimental limitations. However, in recent years experiments focusing on the role of PTMs in DNA accessibility have provided some insight. Cross-linking experiments by Mutskov et al. have shown the nucleosome stability is decreased by hyperacetylation of the H4 tails, while binding of protein GAL4 to its binding site in the wrapped DNA remained unaffected[215]. Work from the Widom group has demonstrated that completely removing all histone tails increased accessibility of nucleosomal DNA for restriction enzymes and in a lesser degree for protein GAL4[216]. More recent sm-FRET experiments have shown that acetylating H3K56 (H3K56Ac) increased nucleosome breathing 7-fold but did not decrease nucleosomal stability[217].

Here we aim to combine several single-molecule methods to answer (1) how does a PTM affect nucleosomal stability and dynamics, (2) how does a PTM affect the binding affinity of a protein, and (3) can we quantify from FRET experiments if and how nucleosomal dynamics is affected by this binding. Because we want as little disruption of the nucleosome structure as possible, we have chosen the trimethylation of lysine 36 in histone 3 (H3K36me3) as modification. Like acetylation, methylation of histones occurs on NH+ groups of lysine residues and is mediated by histone methyltransferases. Unlike acetylation, lysines can bind up to three methyl groups, and methylation preserves the positive amine charge of the lysine,

creating only a sterical bulk. Hence methylation of lysine residues should not interfere with the electrostatic histone–DNA interactions[218].
Next to having an active role in DNA repair and transcription events, H3K36me3 has also been found in pericentromeric (constitutive) heterochromatin, which are areas not associated with activity. This hints at H3K36me3 being involved in the structure of heterochromatin[74]. For DNA repair and transcriptional processes H3K36me3 recruits proteins such as p75, better known as LEDGF (Lens Epithelial-Derived Growth Factor)[203][219]. LEDGF/p75 is a chromatin-binding protein that has been tied to integration of HIV-1 cDNA into human chromosomes and reduction of cellular stress-induced apoptosis[220]. The structural basis for LEDGF binding to chromatin through the proteins' PWWP domain has been elucidated and the domain seems to be essential for binding[221]. However, the AT-hook domains of LEDGF have been shown to interact specifically as well as non-specifically with DNA[222]; deletion of the domains resulted in contradicting observations regarding affinity for compacted DNA[223][224]. Despite these observations, a lot is still not known about the exact binding mechanism and effects of LEDGF interactions with chromatin.
In this chapter we detail the effects of LEDGF binding to nucleosomes by combining single-molecule spectroscopy methods with FRET. We observe that LEDGF affinity for chromatin depends mostly on the presence of H3K36me3.

5.2 Materials and methods

5.2.1 Sample preparation and nucleosome reconstitution

A 197 bp long DNA construct containing one Widom 601 sequence and a Cy3B-ATTO647N fluorophore pair was produced with PCR. The fluorophores were positioned 80 base pairs apart, making FRET possible only when the DNA was reconstituted into a nucleosome. DNA containing one Widom 601 sequence and only the Atto647N fluorophore was also produced with PCR. All nucleosomes were reconstituted by salt gradient dialysis from 2 M to 0 mM NaCl overnight. DNA was mixed with human recombinant histones in a titration of molar ratios ranging from 1:1 to 1:3. Only titrations where no unreconstituted DNA substrates were detected through gel electrophoreris were used for FCS experiments. Nucleosomes

containing H3 trimethylated K36 were reconstituted following the same salt gradient dialysis protocol. Measurement buffers contained 10 mM Tris and 15 mM NaCl, unless stated otherwise. Nucleosome concentrations in FCS measurements were between 3 and 7 nM. Samples of 20 to 40 μl were placed in a closed flowcell to minimize evaporation.

5.2.2 LEDGF purification and labeling

LEDGF/p75 wild-type and ΔAThook were purified as described by Bartholomeeusen [225]. Expression of LEDGF/p75 forms was induced in E.Coli cells (grown in LB medium) by 0.5% IPTG (Isopropyl β-D-1-thiogalactopyranoside) followed by incubation at 30C for 4 hours. Cells were collected and lysed in 20 mL lysis buffer (50 mM Tris, 100 mM NaCl, 1 mM DTT, 10 μg/ml proteinase K and 1 U/10 ml DNase, pH 7.4). After lysis cells were centrifuged and the supernatant was purified over a Heparin column and eluated with a 50 mM Tris 2M NaCl, 1 mM DTT buffer (pH 7.4). Flag-tagged LEDGF/p75(C373A) was labeled by incubating 50μM of the protein with a 10-fold excess of ATTO532-maleimide for 3 hours at room temperature. Unbound dye was removed using Amicon ultra-0.5 centrifugal filters (10 kDa MWCO). The **d**egree **o**f **l**abeling was calculated as

$$dol = \frac{A_{532} \cdot \epsilon_{LEDGF}}{(A_{280} - CF_{280} \cdot A_{532}) \cdot \epsilon_{Atto532}}$$

with $\epsilon_{LEDGF} = 16960 M^{-1} cm^{-1}$, $\epsilon_{ATTO532} = 11500 M^{-1} cm^{-1}$ and $CF_{280} = 0.11$ which corrects for absorbance of ATTO532 at 280nm. Labelling efficiencies were 84% for LEDGF$_{WT}$ and 89% for LEDGF$_{\Delta ATh}$.

5.2.3 Single-molecule fluorescence microscopy

PIE-F(C)CS measurements were done for a least 60 minutes, in recordings of 10 minutes. Measurements were performed on a home-built confocal microscope with a water-immersion objective (60x, NA 1.2, Olympus), using an ICHROME MLE-SFG laser module as excitation source. The excitation beam was directed via fiber coupler and a dichroic mirror (z514/640rpc, Chroma) through the objective and focused 50 μm above the glass-sample interface. FCS experiments were performed in pulsed interleaved excitation (PIE) mode by alternating 514 (30 μW) and 632 nm (20 μW) excitation

pulses of 100 ns long with 300 ns intermittent dark periods. Collected fluorescence was spatially filtered with a 50 μm pinhole in the image plane and split by a second dichroic mirror (640dcxr, Chroma). The fluorescent signals were further filtered (hq570/100nm and hq700/75nm, resp.) and focused on the active area of single photon avalanche photodiodes (SPADs, SPCM AQR-14, Perkin Elmer). The photodiodes were read out with a TimeHarp 200 photon counting board (Picoquant), and the arrival times of the collected photons were stored in t3r (time-tagged to time-resolved) files. These files were further processed with home-built Python analysis programs.

5.2.4 Fluorescence Correlation Spectroscopy

Fluorescently labeled molecules diffusing through the confocal focus cause the intensity of the fluorescent signal to fluctuate in time. In Fluorescence Correlation Spectroscopy (FCS) these fluctuations are used to determine the concentration, diffusion constant and when possible dynamical properties of molecules. The fluctuations in intensity are analyzed by correlating photon arrival times over increasing time-lag τ:

$$G(\tau) = \frac{\langle \delta I_1(t) \cdot \delta I_2(t+\tau) \rangle}{\langle I_1(t) \rangle \cdot \langle I_2(t) \rangle} \tag{5.1}$$

To assess the diffusion of a molecule, photon arrival times of one channel are correlated to generate an autocorrelation curve ($I_1 = I_2$). To quantify the fraction of two differently labeled molecules diffusing through the focus at the same time (i.e. as a complex) the signal of one molecule (I_1) is correlated with the signal of another molecule (I_2) to generate a crosscorrelation curve. The correlation function that fits the diffusional part of a autocorrelation curve is formulated in terms of the concentration and diffusion time of the population of molecules labeled with the same fluorophore, taking into account the confocal volume which they diffuse through:

$$G_{diff}(\tau) = N^{-1} \cdot (1 + \tau/\tau_D)^{-1} \cdot (1 + a^{-2} \cdot \tau/\tau_D)^{-1/2} \tag{5.2}$$

where τ_D is the diffusion time and N the average number of molecules in the confocal volume. Parameter a is the ratio between the axial and radial size of the confocal volume. The value of a for the setup used for the measurements presented here was determined through calibration experiments to be 8.

Physical interpretation of the crosscorrelation functions requires additional calculations and will be discussed further on. The Python pycorrelate module developed by Ingargiola et al.[133] was used to calculate all correlation curves. The correlation algorithm used in this module was developed by Laurence et al.[134]. The algorithm is based on rewriting the correlation as a counting operation on photon pairs and can be used with arbitrary bin widths and spacing.
The diffusion time τ_D of a molecule is determined by its size and the viscosity of the solvent η. The hydrodynamic radius of a molecule r_H can be obtained using the Stokes-Einstein equation:

$$r_H = \frac{k_B T}{6\pi\eta D} \tag{5.3}$$

where diffusion constant $D = \frac{w^2}{4\tau_D}$ with k_B the Boltzmann constant, T temperature and w the radius of the confocal spot in the radial (x,y) direction. Equation **5.3** shows the hydrodynamic radius scales proportional with diffusion time, implying that larger molecules move slower through the focus. This property was used to analyze correlation curves constructed from signals of molecules of different sizes. If we assume the molecule to have a spherical shape, the radius scales with the molecular mass as $r_H \propto M^{\frac{1}{3}}$. In practice this means for the diffusion time to increase two-fold, the mass of a molecule needs to increase a factor of 8.
Photophysics of the fluorophore, i.e. transiting to a triplet or dark state, as well as afterpulsing effects of the APD's need to be included in the fit of a correlation curve

$$G_{total}(\tau) = G_{diff}(\tau) \cdot G_{tr}(\tau) \cdot G_{ap}(\tau) \tag{5.4}$$

where the latter two terms are defined as

$$G_{tr}(\tau) = 1 + \left(\frac{F_{tr}}{1 - F_{tr}} \cdot e^{\frac{-\tau}{\tau_{tr}}} \right)$$

and

$$G_{ap}(\tau) = 1 + \left(\frac{F_{ap}}{1 - F_{ap}} \cdot e^{\frac{-\tau}{\tau_{ap}}} \right)$$

with F_{tr}, F_{ap} the fractions of molecules associated with either triplet state (**tr**) or afterpulsing (**ap**), and τ_{tr}, τ_{ap} their characteristic timescales. As

fluorophore photophysics and afterpulsing take place on different timescales sensible boundaries were set for fitting these parameters.

5.2.5 Fluorescence Cross-Correlation Spectroscopy and binding affinity

Contrary to the value of G(0) of an autocorrelation curve, the plateau of a crosscorrelation curve at small τ does not directly correspond to the concentration of molecules in complex. Rather, $G(0)_{cc}$ represents the complex molecules as a percentage of the population of molecules present at a higher percentage[139]. To calculate the real number of complexed molecules from the crosscorrelation curve, N_{CC} is first corrected for background photons from both channels involved in the cross-correlation:

$$N_{CC,corr} = \frac{N_{CC} \cdot (I_{514G} - bg_{514G}) \cdot (I_{632R} - bg_{632R})}{I_{514G} \cdot I_{632R}} \tag{5.5}$$

Here channel I_{514G} corresponds to the signal of ATTO532-labeled proteins and I_{532R} with ATTO647N-labeled DNA or nucleosomes. We need to take into account that confocal spots from different excitation wavelengths do not completely overlap. Not compensating for this incomplete overlap would mean underestimating the number of molecules in complex. The actual number of molecules in a complex is then calculated as

$$N_{514Gx632R} = c_{over}^{-1} \cdot \frac{N_{514G} \cdot N_{632R}}{N_{CC,corr}} \tag{5.6}$$

where c_{over} is used to correct for the incomplete overlap. Calibration experiments after each alignment of the setup showed $c_{over} = 0.9 \pm 0.03$, implying without correction $\sim$10% of the number of molecules in complex would be missed. With the actual number of molecules in complex determined, the dissociation constant K_d is calculated as

$$K_d = \frac{[molecule1] \cdot [molecule2]}{[complex1+2]} \tag{5.7}$$

The dissociation constant is a measure of the binding affinity and is equal to the concentration of molecule 1 at which half of its available binding sites are occupied by molecule 2.

5.2.6 Multi-component fit and nucleosome dynamics

To determine stoichiometry of populations of two interacting molecules, the number of molecules N in equation **5.2** is split in fractions (for instance closed and open, or bound and free) as represented in the diffusional part of $G(\tau)$:

$$G(\tau)_{diff} = N_{total}^{-1} \cdot \left(F_1 \cdot (1 + \tau/\tau_{D1}) \cdot (1 + a^{-2} \cdot \tau/\tau_{D1})\right)^{-1/2} \cdot \left(F_2 \cdot (1 + \tau/\tau_{D2}) \cdot (1 + a^{-2} \cdot \tau/\tau_{D2})\right)^{-1/2} \quad (5.8)$$

with fractions $F_1 + F_2 = 1$.
For nucleosomes the closed and open fractions are known by previously fitting the autocorrelation curve from the FRET channel ($F_{closed} = \frac{N_{closed}}{N_{total}}$). From this initial fit the characteristic diffusion time ($\tau_{D,closed}$) is also known, leaving only the diffusion time of open nucleosomes $\tau_{D,open}$ to be fitted by the multi-population fit algorithm.
To quantify the effect of protein binding to nucleosomal dynamics, equation **5.8** is extended to include a third population following the same fractioning principle. This means two new variables, F_3 and τ_3, need to be fitted. PIE-FCS experiments on double-labeled nucleosomes with unlabeled protein however only result in two usable (auto)correlation curves: 514R514R and 632R632R. The first curve represents closed nucleosomes, the second all nucleosomes, open and closed, as well as free and bound to protein. To fit the 632R autocorrelation curve we assumed that (1) only open nucleosomes bind to proteins, (2) the dissociation constant (K_d) between open and closed nucleosomes was not altered, and (3) the diffusion times τ_D of unbound (free) open and closed nucleosomes remained the same. Also, if FRET efficiency per closed nucleosome was the same as in a measurement on nucleosomes only, it is not necessary to include a fourth population in the fitting algorithm (closed nucleosomes bound to protein). If the FRET efficiency is however lower, then the 514R autocorrelation curve needs to be fitted with two populations, with $\tau_{D-closednucl-free}$ as a fixed parameter. We can also set the lower bound of $\tau_{D-closednucl-bound}$ to $\tau_{D-closednucl-free}$ as we assume a closed nucleosome bound to a protein is at least diffusing as fast as an unbound one, if not slower due to increased size. The 632R curve is then fitted including this additional population.

5.3 Results

5.3.1 Effect of H3K36me3 on the stability of nucleosomes

To measure the effect of the H3K36me3 modification on a mononucleosomes' intrinsic and salt-dependent equilibria, nucleosomes 39-12-H3(K36me3) and 39-12-WT were diluted to nanomolar concentration in 10 mM Tris and NaCl titrated up to 50 mM. The fluorescent signals of the labels were acquired in PIE-FCS measurements to quantify the concentration and diffusion of the nucleosome and the rates of nucleosome opening and closing.
The average FRET value <E> derived from the total fluorescent signals (corrected for background photons) showed for nucleosome 39-12-H3 a significantly lower value compared to 39-12-WT for all salt concentrations. At 1 mM NaCl the difference was almost 50%; 0.25 $\pm$ 0.005 compared to 0.42 $\pm$ 0.01 for 39-12-WT. Fitting with the Hill function resulted in transition concentration $c_{1/2}$ of 15.83 $\pm$ 0.51 mM NaCl and Hill coefficient H = 0.08 $\pm$ 0.006. For 39-12-WT $c_{1/2}$ = 21 $\pm$ 1 mM and H = 0.14 $\pm$ 0.03, indicating stronger interactions between the nucleosomal DNA exit and the histone core. At low salt concentrations 39-12-H3 was found more often in an open state; 52% at 1 mM NaCl (39-12-WT: 44%).
Fitting the inverted equilibrium constant from population fractions resulted in a salt-independent equilibrium K_{eq} of 1.16 $\pm$ 0.393, a transition concentration x_0 of 12.97 $\pm$ 0.2 mM and salt stoichiometry m of 2.04 $\pm$ 0.23 ion pairs. Compared to 39-12-WT (K_{eq} = 0.83 $\pm$ 0.07, c_0 = 22.1 $\pm$ 0.04 and m = 3.6 $\pm$ 0.1) nucleosomes containing H3K36me3 have a 25% lower intrinsic equilibrium constant, an almost 50% lower transition concentration and about 1.5 less electrostatic interactions between DNA and histone core.
Evaluating nucleosomal dynamics over the salt titration range revealed below the transition concentration c_0 39-12-H3 opened on average at a rate of 24 $\pm$ 15 s^{-1} and closed 23 $\pm$ 15 times per second, translating to an equilibrium constant of appr. 1. Above c_0 the opening rate rapidly increased to 3000 s^{-1} with the closed rate decreasing to 10 times per second, on average. For comparison, below their c_0 39-12-WT nucleosomes on average opened 15 and closed 32 times per second, and above c_0 still closed 40 s^{-1} while opening 1200 times per second at 50 mM NaCl.

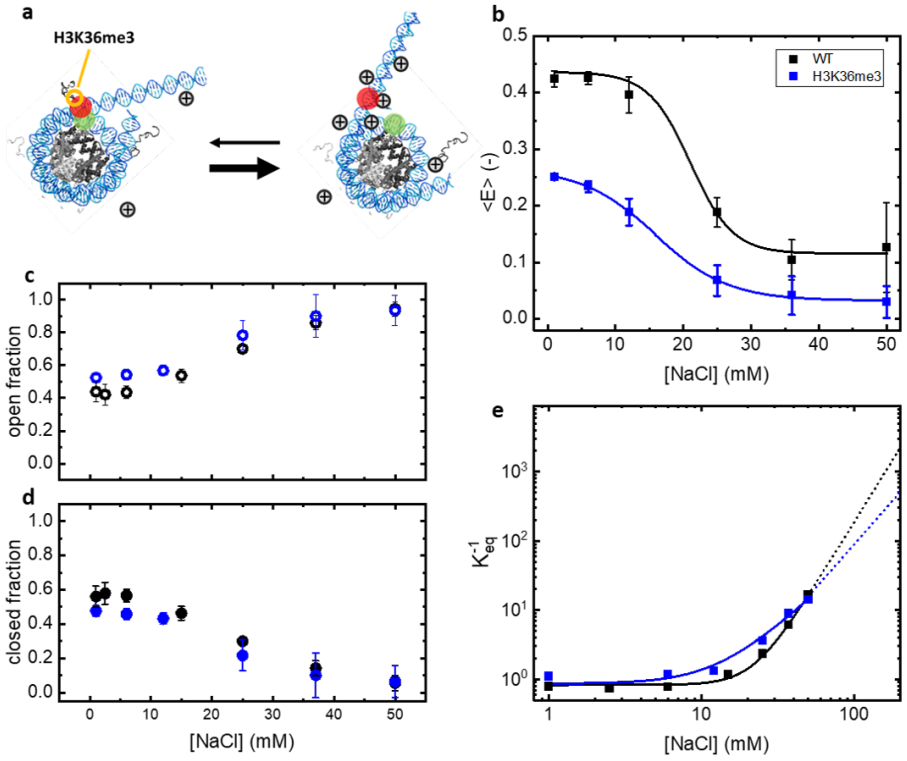

FIGURE 5.1: **Post-translational modification H3K36me3 drives nucleosomes to open conformation. a)** Salt ions interact with the nucleosome, cancelling electrostatic interactions between DNA and histone core, thereby driving the nucleosome to the open conformation. **b)** Comparing average FRET <E> from 39-12-H3 with an unmodified nucleosome shows lower <E> at both low and high salt concentrations. **c) and d)** Open and closed fractions showed nucleosomes containing H3K36me3 are also at low salt concentrations more in the open conformation. **e)** Fitting the inverse of the equilibrium constant from conformational states resulted in a K_X of 0.004 $\pm$ 0.004, a K_0 of 0.9 $\pm$ 0.4, c_0 of 13.0 $\pm$ 0.2 mM and m = 2.0 $\pm$ 0.2. All these values are significantly lower compared to those found for 39-12 nucleosomes containing WT histones.

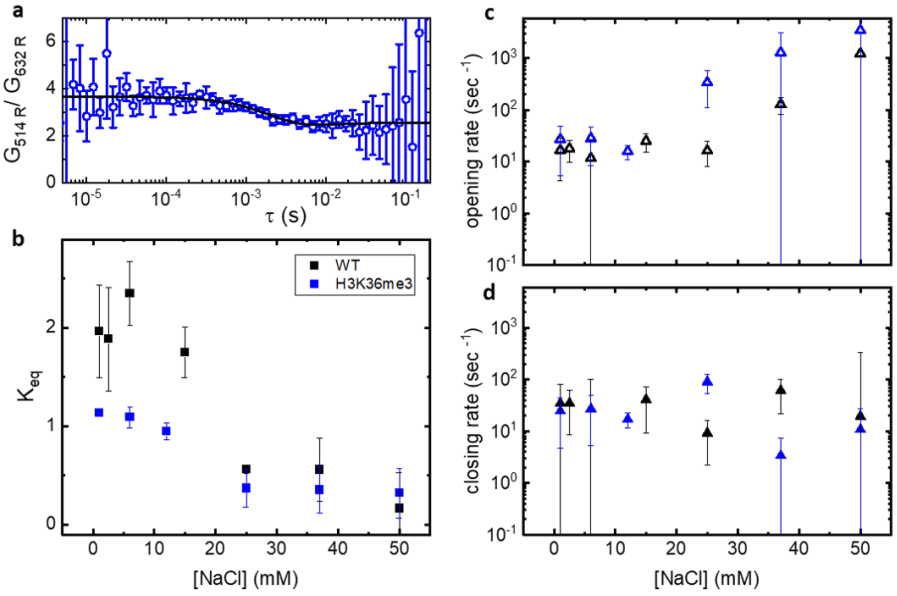

FIGURE 5.2: **Post-translational modification H3K36me3 increases opening rate, but not closing rate of nucleosome. a)** Characteristic plot of dynamics correlation curve of 39-12-H3K36me3. **b)** The equilibrium constant derived from opening (**c)**) and closing rates (**d**) confirmed result from open and closed fractions that nucleosome 39-12-H3 is more often in an open conformation for all salt concentrations. Under c_0 the average opening rate of 39-12-H3 is 24 s^{-1} (15 s^{-1} for 39-12-WT) and average closing rate is 23 s^{-1} (32 s^{-1} for 39-12-WT).

5.3.2 LEDGF binding to DNA and nucleosomes, and resulting ATTO532 quenching

PIE-FCCS measurements were performed on single-labeled LEDGF with ATTO532 and nucleosomes reconstituted from 601 DNA sequence labeled with ATTO647N on one nucleosomal exit (**figure 5.4a**). FRET between ATTO532 and ATTO647N is possible when the labels get close enough (r_F = 5.1nm), which was in principle possible in our construct. However, we did not observe any evidence of this. Nevertheless, a decrease of the fluorescent signal of LEDGF-ATTO532 was clearly visible over time. Depending on concentration the signal stabilized after around 20 minutes. This phenomenon did not seem to be caused by bleaching, as ATTO532

dye was measured at same intensities and showed no decrease. Also non-specific sticking to surface was ruled out as the decrease also occurred for LEDGF-ATTO532 on glasses coated with either PLL-PEG or BSA (shown in **fig. 5.3a**). Fitting showed that both concentration and diffusion time seemed unaffected, meaning protein concentration was not decreasing, but the ATTO532 molecules were being quenched. This process is described in literature as well[226][227][228]. Hence we concluded concentrations and diffusion times derived from ATTO532 signal were reliable to use for affinity calculations.
The loss of fluorescent signal from LEDGF-ATTO532 needed to be taken into consideration when examining the crosscorrelation curve to detect complex formation between protein and nucleosome. As this correlation curve was constructed from correlating photons from their respective fluorescent channels, the complex concentration, as well as its diffusion time, might be underestimated due to signal decrease over longer timescales from the green channel. We observed a small decline of complex concentration in interaction experiments (**fig. 5.2e**). In the case shown however, this is most likely due to decline in red signal /molecules, most likely caused by sample holder drift, a microscope artefact. Since both nucleosome and complex concentrations decrease at similar rates this however did not influence the dissociation constant K_D. Omitting the first 5 minutes, we observed that diffusion times during the next 25 minutes of the measurement were also stable, at 2.2 ms for nucleosomes (39-12-WT), 1.8 ms for LEDGF-WT and 3 ms for the complex. The significantly slower diffusion time of LEDGF in complex in this measurement compared to that of free LEDGF at 1 ms (**fig. 5.2c**) could be used as a separate method to calculate complex formation by fitting the autocorrelation curve from LEDGF with 2 populations. However, 30 minutes into the measurement the diffusion time of LEDGF appeared to decrease to an average of 1.64 ms, and the diffusion time of the complex decreased to a similar time as that of nucleosomes. Similar observations were made during measurements of the other combinations of nucleosomes and LEDGF variants. While diffusion times of nucleosomes and proteins were comparable to those of 39-12-WT and LEDGF-WT as in **fig. 5.3f**, the diffusion times of the complexes were shorter: 1.8$\pm$0.7 ms for 39-12-WT with LEDGF-ΔAth, and 2.0$\pm$0.5 ms for LEDGF-WT and 1.8$\pm$0.4 ms for LEDGF-ΔATh with nucleosomes containing H3K36me3. The errors increased not only from the decreased diffusion time but also because of

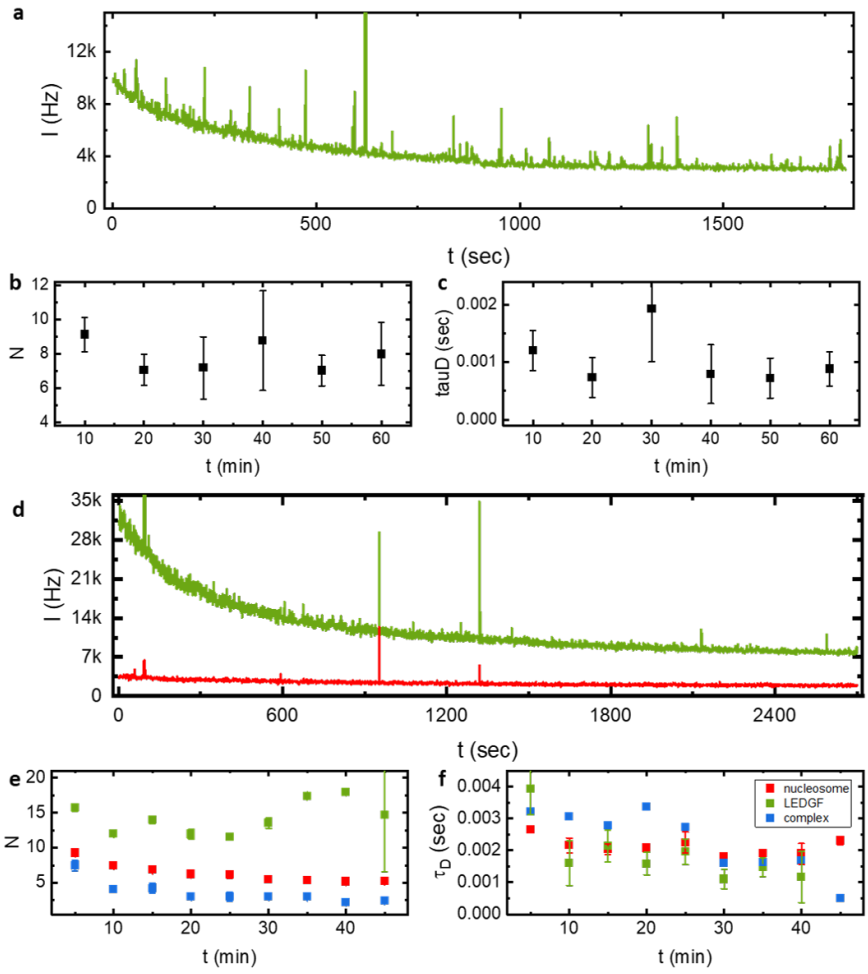

FIGURE 5.3: **ATTO532 is prone to quenching when LEDGF is over-saturated with labelling dye. a)** Decrease of the fluorescent signal from ATTO532 labelled LEDGF shows an irreversible loss of fluorescence on glasses coated with PLL-PEG or BSA. **b)** However, bleaching and sticking can be ruled out, because the concentration of protein remained stable over time, as did the diffusion time (**c)**). **d)** Fluorescent signal of an interaction experiment of single-labeled nucleosomes (red) with LEDGF (green) decreased somewhat in signal from the nucleosome, likely due to sample holder drift. **e)** The number of nucleosomes decreased according to the decrease in fluorescence signal although the LEDGF concentration was remained constant. **f)** The diffusion times of both nucleosomes and LEDGF remained stable over time, the diffusion of the complex however seemed to decrease. (Measurement shown in **d-e-f** is on nucleosome 39-12-WT with LEDGF-WT)

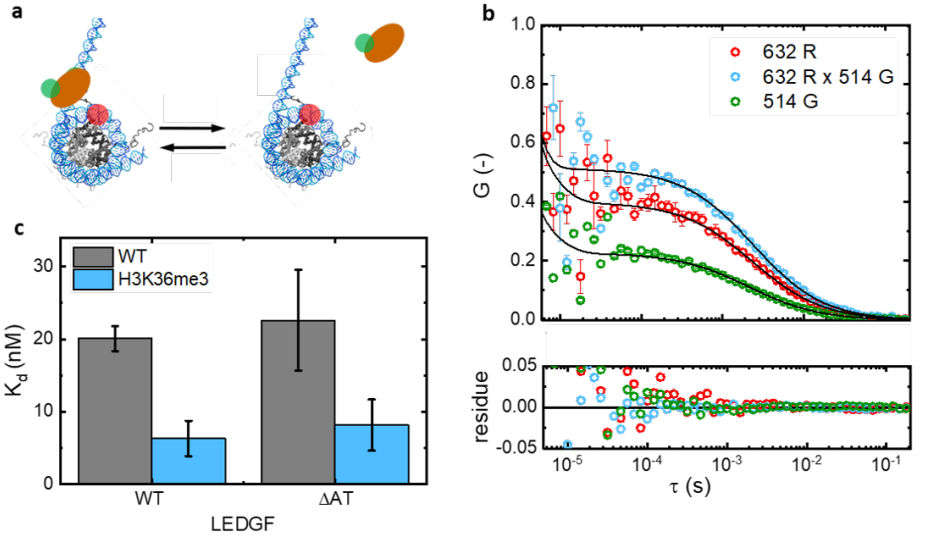

FIGURE 5.4: **H3K36me3 tripled LEDGF affinity for nucleosomes, deletion of AThook domain does not decrease affinity. a)** Representation of LEDGF labeled with ATTO532 binding to nucleosome labeled with ATTO647N. This method cannot distinguish if LEDGF binds to closed or open nucleosomes, or to both conformations. **b)** Representative auto- and cross-correlation curves from measurements, corrected for concentrations of molecules associated with each curves (red: nucleosomes, green: LEDGF, blue: complex) **c)** Dissociation constants derived from concentrations showed both WT-LEDGF and ΔAThook LEDGF had a significantly higher affinity for nucleosomes containing H3K36me3.

protein aggregates. Dissociation constants (**fig. 5.4c**) were calculated from averaging affinity at the start and at the end of the measurement. Note though that at the end fluorescent signals were stable, but complex concentration might be underestimated due to quenching. At the start however, complex concentration might be overestimated due to the rapid decrease of the fluorescent signal. Averaging these two instances may cancel out these uncertainties. The results showed both LEDGF-WT and LEDGF-ΔATh had a higher affinity for nucleosomes containing H3K36me3 (6±2 nM and 8±4 nM, resp.) than for nucleosomes containing recombinant/WT histones (LEDGF-WT: 20±2 nM, LEDGF-ΔATh: 23±7 nM). The error bars in **fig. 5.4c** are the propagated standard deviations after averaging. The size of the errors result from larger differences between affinities calculated at the

start and at the end of the experiments for LEDGF-ΔATh than for LEDGF-WT.

5.3.3 Effect of LEDGF binding on nucleosome dynamics

PIE-FCS was used to measure the changes in nucleosome conformation stoichiometry, FRET efficiency and diffusion time upon adding WT or ΔATh LEDGF to nucleosomes containing or lacking H3K36me3 (39-12-H3 and 39-12-WT). Proteins were added to nucleosomes in a 4 to 1 ratio.
A decrease in FRET signal was immediately noticeable in the first 5-10 minutes (fast binding), then slowed down (**fig. 5.5a**). The rapid decrease at the start of the measurement was however only visible for nucleosome 39-12-H3 with LEDGF-ΔATh and to a lesser degree for nucleosome 39-12-WT with LEDGF-WT. In other measurements it was probably missed because a larger time was used for transferring sample to microscope and starting the measurements, in which time nucleosomes may have been depleted by sticking to the surface of the flowcell. We therefore corrected the FRET intensity by dividing it by the number of closed nucleosomes, determined from fitting the correlation curves. When adding LEDGF, FRET per nucleosome decreased, but only significantly for LEDGF-ΔATh binding to 39-12-WT and 39-12-H3 (**fig. 5.5b**). The fraction of closed nucleosomes decreased over time as well (**fig. 5.5c**), so loss of FRET signal was mostly from losing closed nucleosomes, and only marginally from nucleosomes becoming more open through binding of LEDGF. The equilibrium constants decreased accordingly (**fig. 5.5d**). Interesting to note is that LEDGF binding lowers the equilibrium constant K_{eq} between open and closed nucleosomes faster and more for 39-12-WT nucleosomes than for 39-12-H3 nucleosomes. This could however be caused by differences in protein concentration, since it was not possible to verify directly from measurements.
Fig. 5.6a shows the reaction scheme for combined nucleosome breathing and LEDGF binding. The loss of overall FRET, but minimal loss of FRET per nucleosome implied route k_{14}, i.e. from free closed nucleosome to bound open nucleosome via bound closed nucleosome, was not the preferred pathway. Also, in all measurements the diffusion times of closed nucleosomes (with LEDGF and complexes present) were not significantly slower than those found for free closed nucleosomes, indicating no detectable binding of proteins to closed nucleosomes. Lastly, the decrease in FRET signal as well as the decrease in closed nucleosomes did not seem to reverse, implying pathway k_{32} is likely never taken. Fitting in a complex fraction in addition

to an open and a closed fraction on autocorrelation curve 632R (**fig. 5.6b**) resulted in diffusion times of the complex similar to the times found in single-labeled experiments for binding to 39-12-WT nucleosomes (**fig. 5.6c**); 3.6±0.3 ms (LEDGF-WT) and 2.2±0.4 ms (LEDGF-ΔATh). For nucleosomes containing H3K36me3 the diffusion times of complexes appeared to be faster (1.1±0.8 ms (LEDGF-WT) and 1.7±1.8 ms (LEDGF-ΔATh)), comparable to the times of closed nucleosomes. Faster diffusion times might indicate bound 39-12-H3 nucleosomes were more wrapped through LEDGF binding, but the nucleosomal exit was not close enough for FRET to occur.
Combining the found diffusion times for the complex fractions with the decrease in I_{514R} per nucleosome showed an interesting connection; the decrease in FRET was almost proportional with the decrease in diffusion time. The complex with the slowest diffusion time, 39-12-WT with LEDGF-WT, had the smallest decrease of 8±2%. Complex 39-12-WT with LEDGF-ΔATh decreased 22±9%. Complexes with 39-12-H3 showed FRET decreasing with 14±2% (LEDGF-WT) and 26±7% (LEDGF-ΔATh). This might mean that, depending on the LEDGF variant and the presence of H3K36me3, a small subset of proteins were able to bind to closed nucleosomes (taking pathway k_{14}). The 3-component fit resulted in similar complex fractions for nucleosome 39-12-WT with both LEDGF variants; 0.33±0.02 with open: 0.38±0.02 and closed: 0.29±0.02 (LEDGF-WT) and 0.26±0.02 with open: 0.42±0.02 and closed: 0.32±0.02 (LEDGF-ΔATh). For nucleosome 39-12-H3 the fit resulted in a larger complex fraction of 0.48±0.04 (with open: 0.11±0.04 and closed: 0.41±0.04) for binding with LEDGF-WT and a small fraction of non-bound nucleosomes 0.14±0.05 (with open: 0.46±0.05 and closed: 0.40±0.05) for LEDGF-ΔATh.
Considering that diffusion times yielded large errors and resembled the times of closed nucleosomes, the fitting algorithm may not have been able to distinguish between the contributions of the complex fraction to the autocorrelation curve from those of closed unbound nucleosomes. An obvious solution would have been to fit the 632R autocorrelation curve with 4 populations. This would have resulted in a objectively better fit of the curve, but would unlikely had given better results for the stoichiometry of the population.

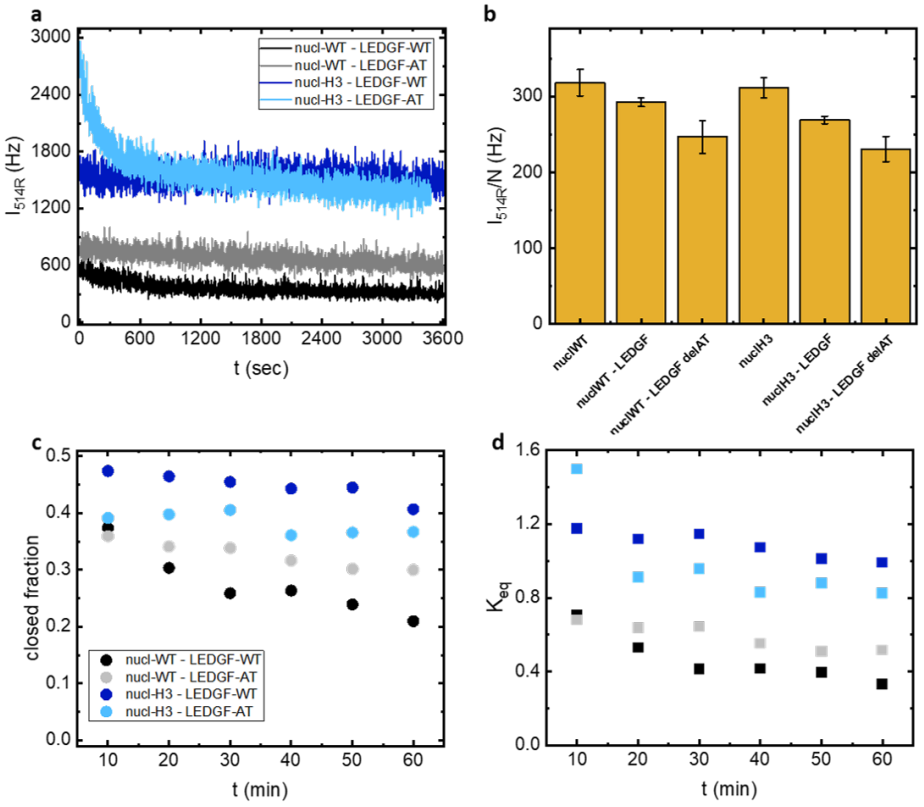

FIGURE 5.5: **Unlabeled LEDGF variants binding to double-labeled nucleosomes became visible from loss of FRET signal. a)** Upon adding unlabeled LEDGF, for all combinations of proteins and nucleosomes a decrease in total FRET signal was visible **b)** Also FRET per nucleosome decreased for each combination between nucleosomes and proteins, however the loss was less since also the numbers of closed nucleosome (**c)**) and consequently the equilibrium constants (**d)**) decreased. Error bars were omitted from figures **c)** and **d)** for clarity. For nucleosome 39-12-WT (318±17 ph/s), the loss of FRET was 8±2% with LEDGF-WT and 22±9% with LEDGF-ΔATh. For nucleosome 39-12-H3 (312±14 ph/s) the loss was 14±2% with LEDGF-WT and 26±7% with LEDGF-ΔATh.

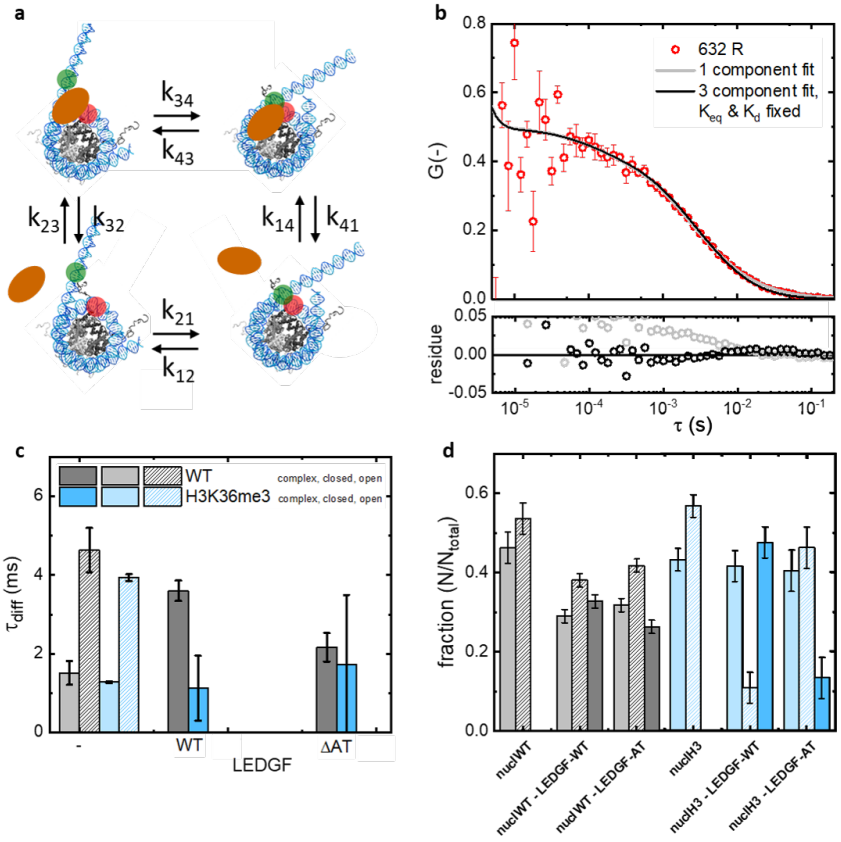

FIGURE 5.6: **Binding of LEDGF variants to nucleosomes containing H3K36me3 decreased diffusion times of the formed complex.a)** 4-state reaction scheme of protein binding to open and closed nucleosomes. **b)** Fitting the autocorrelation with 3 components (open, closed and nucleosomes in complex) improved the fits, especially around the millisecond timescale, characteristic for nucleosome and LEDGF (and their complex) diffusion times. **c)** Complex diffusion times from 3-component fit resulted in times for nucleosome 39-12-WT with either LEDGF variant resembling times found in measurements on single-labeled nucleosomes and proteins (**fig. 5.3f**).**d)** Fractions from fit resulted in significant complex fractions for 39-12-WT with LEDGF-WT (0.33±0.02) and LEDGF-ΔAth (0.26±0.02), and 39-12-H3 with LEDGF-WT (0.48±0.04). For 39-12-H3 with LEDGF-ΔATh only a small fraction of complex was observed (0.14±0.05).

5.4 Discussion and conclusions

Here we studied the effect of H3K36me3 on nucleosomal dynamics and stability, and how it increased affinity of the protein LEDGF to nucleosomes. Nucleosomes containing H3K36 trimethylation seemed more open from average FRET and equilibrium constants. But the FRET signal shows that the trimethylation does not inhibit bending of the nucleosomal exit towards the histone core. It does however lower the number of electrostatic interactions m from 3.6 to 2. This decrease in DNA-histone interactions is supported by the observation of faster dynamics of H3K36me3 nucleosomes in both low and high salt conditions. Increased kinetics through less (strong) interactions between DNA arms and histone core may be how the trimethylation facilitates binding to nucleosomal DNA in processes such as transcription and DNA repair. Morrison et al.[229] have recently suggested that the H3 histone tails interact with compacted DNA as a 'fuzzy' complex, interacting robustly but adopting a dynamic ensemble of DNA-bound states[230]. It is possible that trimethylation of lysine 36, although not altering the local charge of the histone tail, does change the range of binding states with the condensed DNA.

We have also presented evidence that H3K36me3 increases LEDGF affinity to nucleosomes by using PIE-FCCS on single-labeled nucleosomes and single-labeled variants of the LEDGF protein (wild type, WT and AT-hook deficient, ΔATh). PIE-FCCS and our data analysis algorithm could circumvent the effects of ATTO532 quenching and generate reliable results for concentrations and diffusion times of nucleosomes, proteins as well as the complex formed between them. The diffusion times we found for the four different complexes differed between one another with more than one millisecond, indicating different modes of interaction, depending on the LEDGF variant and if H3K36me3 was present or not in the nucleosome. Complexes of AT-hook deficient LEDGF with nucleosomes diffused fastest at around 1.8 ms, almost as fast as closed nucleosomes (1.2-1.5 ms). It might be through interaction of the PWWP domain, and the lack of the 20+ amino acids of the AT-hooks domain, that the LEDGF variant brings the DNA exit closer to the nucleosome core. The slowest complex, composed of LEDGF-WT and nucleosomes lacking H3K36me3, still diffused faster than open nucleosomes (3 ms vs. 4 ms). The complex of LEDGF-WT with H3K36me3 nucleosomes is somewhat faster, around 2 ms.

We found similar diffusion times when investigating the effect of LEDGF

variants on nucleosome dynamics in PIE-FCS measurements on double-labeled nucleosomes and unlabeled proteins. As the LEDGF variants were unlabeled we had to determine the bound fractions by fitting the autocorrelation curve from the signal of all states of nucleosomes. The results for fractions and diffusion times for experiments on WT nucleosomes seemed to be more accurate than those of experiments on H3K36me3 nucleosomes; both fraction sizes and diffusion times had large errors. Since these values are based on the correlation curve, representing the nucleosome in certain conformational states, these large errors might be caused by 'fuzzy' states due to the altered interactions of the H3 tail with nucleosomal DNA.
We also observed that despite losing closed nucleosomes at the start of an experiment, loss of FRET per nucleosome is minimal, about 10-25% depending on the protein and nucleosome variants, implying the pathway from free nucleosome to bound nucleosome is predominantly via free open nucleosomes. The steady decrease in FRET signal in measurements of 60 minutes pointed to the depletion being irreversible. This observation, combined with the diffusion times of the complex that were faster than those of open nucleosomes, suggests that LEDGF binding stabilizes open nucleosomes by remodelling rather than unwrapping them. It would be an interesting follow-up experiment to titrate salt to determine if the electrostatic interactions in LEDGF-bound nucleosomes, and thus nucleosomal stability, have increased through binding.