Heterochromatin protein 1 recruitment and chromatin conformational dynamics at the single-molecule level

THÈSE N^{O} 7417 (2017)

PRÉSENTÉE LE 9 MARS 2017

À LA FACULTÉ DES SCIENCES DE BASE
LABORATOIRE DE CHIMIE BIOPHYSIQUE DES MACROMOLÉCULES
PROGRAMME DOCTORAL EN CHIMIE ET GÉNIE CHIMIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Sinan KILIC

acceptée sur proposition du jury:

Prof. A.-C. Corminboeuf, présidente du jury
Prof. B. Fierz, directeur de thèse
Prof. M. G. Poirier, rapporteur
Prof. C. Hackenberger, rapporteur
Prof. K. Johnsson, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Suisse
2017

Summary

Chromatin is the template on which DNA-associated transactions take place in eukaryotic organisms. Nucleosomes consisting of the four histones H2A, H2B, H3 and H4 each organize ~150bp of DNA and constitute a first layer of chromatin. The three-dimensional organization of chromatin as well as histone post-translational modifications (PTMs) regulate recruitment of chromatin-associated effector proteins (effectors).

Heterochromatin protein 1 (HP1) is an effector associated with silenced genome regions. HP1 recognizes histone H3 trimethylated at lysine 9 (H3 K9me3) and can dimerize. This results in a protein with two binding domains allowing multivalent engagement of target chromatin. HP1 can further promote chromatin condensation and inter-fiber contacts.

The effector p53 binding protein (53BP1) is a key regulator in the DNA damage repair pathway. It is known to target a trio of PTMs; H4 dimethylated at K20 (H4 K20me2), H2A(.X) ubiquitylated at K15 (H2A.X K15ub) and H2A.X phosphorylated at S139 (H2AX S139ph).

Although details about the function of the individual domains of these proteins have been uncovered, little is known about the binding mechanism of the holoproteins and how this affects chromatin conformation.

The aim of this thesis was to develop chromatin engineering and single-molecule fluorescence based methods to; i) Interrogate recruitment kinetics of HP1 to post-translationally modified chromatin. ii) Understand how HP1 binding alters the conformational dynamics of chromatin secondary structure. iii) Prepare histones carrying the PTM signature recognized by 53BP1.

We established an assay to monitor binding kinetics of HP1α to modified chromatin using co-localization single-molecule microscopy (CoSM). H3 K9me3 octamers, labeled HP1α, dimerized HP1α and labeled array DNA formed the basis for this. With this, we found that HP1α multivalency induced by dimerization functions as a platform to enhance HP1α binding to target chromatin up to 9-fold by both accelerating association and prolonging retention. This was further corroborated by FRAP measurements using specific mutants in live mouse fibroblasts.

Chromatin conformational dynamics were investigated by ensemble and single-molecule FRET (smFRET). Multiple combinations of FRET positions allowed us to obtain multi-perspective information on the conformational changes in chromatin upon compaction. This showed distinct steps

in the local folding of chromatin. Histone acetylation of histone H4 prevented the last steps of this folding pathway. HP1-mediated compaction promotes earlier steps of compaction while maintaining conformational dynamics.

Finally, towards similar studies with 53BP1, we devised schemes for synthesis of histones containing the target PTMs. Histones with the individual PTMs were prepared by ligation and desulfurization. H2A.X with both an N-terminal ubiquitin and C-terminal phosphorylation was prepared through a convergent route using recombinant SUMO and a split intein from *Nostoc Punctiforme* as orthogonal recombinant protection groups.

Together the work described in this thesis combines advanced protein and chromatin engineering with CoSM and smFRET. This resulted in mechanistic insight into the spatio-temporal regulation of HP1 recruitment, chromatin conformational dynamics, templates for similar investigations with 53BP1, and tools for further investigation of nucleosome function in the context of chromatin.

Keywords: *HP1*, *histone post-translational modifications*, *multivalent chromatin effectors*, *chromatin conformations*, *expressed protein ligation*, *53BP1, co-localization single-molecule microscopy*, *single-molecule Förster resonance energy transfer*.

Résumé

Les nucléosomes sont constitués de quatre histones, organisent ~150bp d'ADN et composent la première couche d'empaquetage de la chromatine. L'organisation tridimensionnelle de la chromatine ainsi que les modifications des histones régulent le recrutement des protéines effectrices associées à la chromatine.

La protéine HP1 est une protéine effectrice associée aux régions hétérochromatiques du génome. Cette protéine qui a la capacité de formé des homodimères reconnaît l'histone H3 triméthylée à la Lysine 9 (H3 K9me3). La dimérisation permettant un engagement multivalent sur la chromatine cible, cette interaction promeut condensation de la chromatine.

La protéine effectrice 53BP1 est un régulateur clé dans le processus de réparation de l'ADN. Elle reconnaît un trio de modifications incluant H4 diméthylée à la Lysine 20 (H4K20me2), H2A(.X) ubiquitylée à la Lysine 15 (H2AX K15ub) et H2AX phosphorylée à la Sérine 139 (H2AX S139ph). Bien que ces sites de reconnaissance aient été identifiés *in vivo*, on en sait relativement peu à propos de l'effet combiné de ceux-ci sur le recrutement de 53BP1.

Le but de cette thèse était de développer des approches basées sur l'ingénierie de la chromatine et sur la fluorescence à molécule unique pour; i) Interroger la cinétique de recrutement de HP1 sur la chromatine modifiée post-traductionnellement (PTM). ii) Comprendre comment la liaison de HP1 pourrait altérer la dynamique de conformation de la structure secondaire de la chromatine et iii) Développer une route synthétique pour la préparation des histones portant les PTMs reconnu par 53BP1.

De la chromatine synthétique marquées avec des fluorophores, des nucléosomes contenant la PTM H3 K9me3, HP1 ainsi que HP1 dimerisé ont été préparés en utilisant plusieurs méthodes. En utilisant ceux-ci, nous avons trouvé par le biais de la microscopie de colocalisation à molécule unique (CoSM) que la plurivalence de HP1 induite par homodimérisation améliorerait la liaison de HP1 sur la chromatine cible en accélérant l'association et en prolongeant la rétention. Cela a ensuite été confirmé *in vivo* par des mesures de FRAP dans des fibroblastes de souris mutant.

La préparation de la chromatine synthétique avec des marqueurs FRET a mené au développement d'une méthode d'évaluation de la dynamique de conformation de la chromatine au niveau de la molécule unique. Plusieurs combinaisons de positions FRET nous ont permis de démontrer des étapes distinctes de conformation locale lors de la compaction de la chromatine. L'acétylation de l'histone

H4 empêche la dernière étape de cette voie de pliage. La compression induite par HP1 promeut les premières étapes de compression en rapprochant les dyades de nucléosomes.

Finalement nous avons procédé à des expériences similaires avec 53BP1 en utilisant des histones comportant les PTMs cibles de 53BP1. H2AX comportant une ubiquitine à l'extrémité N-terminale et une phosphorylation à l'extrémité C-terminale a été préparée par une route convergente utilisant SUMO recombinante et l'intéine divisée Npu comme groupes protecteurs orthogonaux.

En somme, le travail décrit dans cette thèse combine une ingénierie avancée de la chromatine avec du CoSM et smFRET. Cela a conduit à une connaissance du mécanisme de la régulation spatio-temporelle du recrutement de HP1, de la dynamique conformationnelle de la chromatine, une base pour des investigations similaires avec 53BP1 ainsi que des outils pour étudier les nucléosomes dans le contexte de la chromatine.

Mots-clés: *HP1, modifications des histones post-traductionnellement, protéines effectrices multivalent associées à la chromatine, conformations de la chromatin, ligation de protéines exprimée, 53BP1, colocalisation à molécule unique, transfert d'énergie par résonance de type Förster sur molécule unique*

Acknowledgements

I owe several people my thanks for contributing in various ways to this thesis.

I want to thank Beat for having me in his lab. His optimism and support has been greatly motivating at times when experiments have been challenging. The continuous discussions that we have had have been stimulating and educational and have contributed significantly to shaping my scientific mindset. His tireless development of software and work on setting up the infrastructure for microscopy measurements has been inspirational.

I also want to thank the jury members Prof. Kai Johnsson, Prof. Christian Hackenberger and Prof. Michael Poirier, for taking some of their precious time to review this thesis and to Prof. Clémence Corminboeuf for being the president of the jury.

A number of colleagues have been instrumental in making my PhD both enjoyable and educational.

First, I want to thank Carolin Lechner for being a great friend and colleague, who I have greatly appreciated talking to both about scientific and less scientific topics.

My colleagues in Düsseldorf have also greatly contributed to this work. Thanks to Claus Seidel, Olga Doroshenko and Suren Felekyan for welcoming me to their lab and for introducing me to single-molecule FRET in solution. Their work in acquiring and analyzing data have been a great help.

Thanks a lot to Aurore and Maxime for their help with the French translation of my summary, and to them, and the rest of the lab for creating and maintaining a good spirit in the lab.

I further thank Maeva Tobbler and Nicolas Sambiago for their help with preparative work, and their positive attitudes. The help of Marie Munoz, Jacques Gremaud and Marie Jirousek with administrative and ordering tasks as well as their friendly demeanor is also greatly appreciated.

I greatly acknowledge the financial support, communication training and network I gained through the Boehringer Inhelheim Fonds, thanks to Claudia Walther, Anja Hoffmann and Sandra Schedler.

I am thankful to my parents and brother who have emboldened and supported me regardless of what I have been up to. Having them always back me up made it easier to get through challenging times.

Finally I am endlessly grateful to Inkatuuli Heikkinen. Thank you for your support, advice and encouragement in all aspects of my life. It was particularly a pleasure looking at fluorescent cells with you on weekend evenings. It means more than you imagine to know that you are always there for me.

Abbreviations

53BP1	p53 binding protein 1
5C	carbon copy chromatin conformation capture
acm	acetamidomethyl
ACN	acetonitrile
ASF1	anti-silencing factor 1
ATP	adenosine triphosphate
AUC	analytical ultracentrifugation
Boc	tert-butyloxycarbonyl
bp	basepair
BRCA1	breast-cancer associated gene 1
BRCT	BRCA1 C-terminus
BSA	bovine serum albumin
CAF1	chromatin assembly factor 1
CAT	catalase
CD	chromo domain
CoA	coenzyme A
CoSM	co-localization single-molecule microscopy
COT	cyclooctatetraene
CPEC	circular polymerase extension cloning
CSD	chromo shadow domain
Cy3	cyanine 3
Da	Dalton
DAXX	death domain associated protein
Dbz	diamino-benzoic acid
DCM	dichloromethane
DDR	DNA damage repair
DHS	dNaseI hypersensitive sites
DIPEA	diisopropylethylamine
DMEM	dulbeccos modified eagle medium
DMF	dimethylformamide

DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
Dnase1	deoxyribonuclease 1
dNTP	deoxyribonucleic acid triphosphate
DOT1L	disruptor of telomeric silencing 1-like
DSB	double-strand break
DTT	dithiothreitol
E.Coli	*Escherischia Coli*
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
EPL	expressed protein ligation
ESI-MS	electrospray ionization mass spectrometry
EtOH	ethanol
FACT	facilitates chromatin transcription
FCS	fluorescence correlation spectroscopy
FISH	fluorescence in situ hybridization
Fmoc	9-fluorenylmethoxycarbonyl
FPLC	fast protein liquid chromatography
FRAP	fluorescence recovery after photobleaching
FRET	Förster resonance energy transfer
GlcNAc	N-acetyl glucosamine
GOD	glucose oxidase
GSH	glutathione
GuaHCl	guanidine hydrochloride
H1, H2A, H2B, H3, H4	histone proteins
HAT	histone acetyltransferase
HATU	2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HDAC	histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HOBt	hydroxybenzotriazole
HOMO	higher occupied molecular orbital
HP1	heterochromatin protein 1
HR	homologous recombination
IPTG	isopropyl b-D-1-thiogalactopyranoside
Kac	potassium acetate
KAP1	KRAB-associated protein 1
kDa	kilodalton
KDM	lysine demethylase
KMT	lysine methyltransferase
LUMO	lower unoccupied molecular orbital
MCS	multiple cloning site
MEF	mouse embryonic fibroblast
MeOH	methanol
MESNA	mercaptoethanesulfonate sodium
MFD	multiparameter fluorescence detection
Micro-C	micrococcal nuclease capture
MLA	methyl-lysine analogue
MMTV	mouse mammary tumor virus
MNaseI	micrococcal nuclease I
MPAA	mercaptophenyl acetic acid
MS	mass spectrometry
MST	microscale thermophoresis
MTG	methylthioglycolate
MW	molecular weight
MWCO	molecular weight cutoff
Mxe GyrA	*Mycobacterium xenopy* DNA gyrase
NA	neutravidin
NAD	nicotinamide adenine dinucleotide
Nap1	nucleosome assembly protein 1
NaPi	sodium phosphate

NBA	4-nitrobenzyl alchohol
Nbz	N-acyl-benzimidazolinone
NCL	native chemical ligation
NCP	nucleosome core particle
NHEJ	non-homologous end joining
NHS	N-hydroxysuccinimide
Ni-NTA	nickel nitrilotriacetic acid
NPS	nucleosome positioning sequence
Npu	*Nostoc punctiforme*
NTE	N-terminal extension
NuRD	nucleosome remodelling and deactylation
PAGE	polyacrylamide gel electrophoresis
PALM	photoactivation localization microscopy
PCA	protocatechuic acid
PCD	protocatechuate dioxygenase
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIE	pulsed interleaved excitation
PMSF	phenylmethane sulfonyl fluoride
POI	protein of interest
PSF	point spread function
PTM	post-translational modification
PTS	protein transsplicing
RAD6	radiation sensitivity protein 6
RNA	ribonucleic acid
RP-HPLC	reverse phase high pressure liquid chromatography
RT	room temperature
SDS	sodium dodecyl sulphate
SHL	super-helix location
smFRET	Single-molecule förster resonance energy transfer
SOB	super optimal broth

SOC	super optimal broth with catabolite repression
SPPS	solid phase peptide synthesis
STORM	stochastic optical reconstruction microscopy
SUMO	small ubiquitin modifier
SVA	succinimidyl valerate
TAD	topologically associated domain
TAMRA	tetramethylrhodamine
TBE	tris-boric acid-EDTA
t-BuSH	tert-butylthiol
TCEP	tricarboxyethylphosphine
TCSPC	time-correlated single photon counting
TE	tris-EDTA
TEV	tobacco etch virus
TEAA	triethylamine acetic acid
TF	transcription factor
TFA	trifluoroacetic acid
TFET	trifluoroethanethiol
TFP	Tetrafluorophenyl
Thz	thiazolidine
TIRF	total internal reflection fluorescence
TIS	Triisopropylsilane
TONSL-MMS22L	tonsoku-like-MMS22 like complex
TSQ	triplet state quencher
TTD	tandem tudor domain
UDR	ubiquitin dependent region
UHRF1	ubiquitin like with PHD and ring finger domains 1
Ulp1	ubiquitin like protease 1
WT	wild-type
βME	β-mercaptoethanol

Index

1. Introduction

1.1. Chromatin biology

1.1.1. DNA packaging in eukaryotes

DNA in interphase eukaryotic cells is localized in the cell nucleus as decondensed post-mitotic chromosomes. The chromatin that each chromosome is built from is a nucleoprotein complex between DNA, histone proteins along with linker histones and a range of proteins binding to this nucleoprotein complex. The fundamental repeating unit is the nucleosome core particle (NCP) consisting of 145-150bp of DNA wrapped in 1.65 negative superhelical turns around the exterior of a disk-shaped histone octamer [1-5] (**Figure 1.1.1**). The nucleosome is often bound by the linker histone H1[6].

Figure 1.1.1*. Chromatin organization in the eukaryotic cell nucleus. Chromatin is densely packed in the cell nucleus with the fundamental unit being the nucleosome consisting of DNA wrapped around the core histones H2A, H2B, H3 and H4.*

The histone octamer in nucleosomes is composed of two copies each of the four core histones H2A, H2B, H3 and H4 that are highly conserved in organisms ranging from yeast to plants and mammalians [7,8]. The 102-145 amino acid residue length of the histones are highly enriched in lysines and arginines that form essential contacts with the DNA phosphate backbone, resulting in overall isoelectric points of 10.6-11.4. The histones form two distinct domains within the structure of the nucleosome; The central folded domains and the N-terminal tails protruding from the core as unstructured regions. The folded core is an intricate bundle of α-helices referred to as the handshake fold. These form a central $(H3\text{-}H4)_2$ tetramer flanked on each lateral surface of the nucleosome by H2A-H2B dimers[9,10]. A range of histone variants homologous to the core histones contribute to the diversification of the core of the histone octamer. Such variants exist in particular for H2A (H2A.X, H2A.Z, macroH2A and H2A.BBd) and H3 (H3.1, H3.2, H3.3, H3.1t, CENP-A, H3.X and H3.Y)[11,12]. The N-terminal tails

protrude either directly from the surface of the histone octamer, or through the gyres of the DNA. Post-translational modifications (PTMs) of these tails as well as some of the residues within the core of the histones serve essential functions in regulating DNA dynamics and interaction with downstream functioning proteins[13].

The DNA wrapped around the octamer interacts at 14 points either directly with the histones or through water-mediated contacts each time the minor groove points towards the octamer through phosphate backbone interactions and not with the DNA bases[14,15]. The central position of the 147bp DNA (i.e. ~74bp in) is referred to as the nucleosome dyad and each DNA superhelix (~10bp distance) referred to as superhelix locations (SHL) +1 to +7 downstream of the dyad and SHL -1 to -7 upstream of the dyad[10,16]. The interaction between DNA and histones is principally independent of the DNA sequence. Despite this, there are sequence dependent patterns of nucleosome formation linked to the flexibility and curvature of A/T rich sequences vs G/C rich sequences and the negative superhelical turning of the DNA around the histones[17-19]. A natural sequence found to function as a strong nucleosome positioning sequence (NPS) is the DNA encoding 5S rRNA in sea urchin[20]. Through selection experiments, the 601 DNA sequence was identified[18]. This is one of the strongest known NPSs and has therefore been used in almost all subsequent *in vitro* studies, requiring homogenous positioning. The DNA-based preferences for nucleosome positioning also significantly affect the genomic positioning of nucleosomes[21,22]. The number of interactions and the strength of these vary at the 14 contact points. Near the nucleosome dyad and extending 50-55bp in each direction, the DNA is tightly wrapped around the histone octamer. The exterior 20bp on each side of the nucleosome (The entry/exit region) however is more flexible.

Dynamics of the nucleosome

Histones are assembled onto DNA to form nucleosomes in tandem with the replication fork to prevent exposure of naked DNA[23]. Several histone chaperones facilitate the stepwise assembly of nucleosomes. Two pairs of H3-H4 to form the $(H3\text{-}H4)_2$ tetramer are deposited by shuffling between the associated chaperones chromatin assembly factor 1 (CAF-1), anti-silencing factor 1 (ASF1) and tonsoku-like-MMS22 like complex (TONSL-MMS22L) (replication-coupled assembly) as well as HIRA and DAXX (replication-independent assembly with H3.3)[24-28]. H2A-H2B dimers are incorporated on each side by chaperones nucleosome assembly protein 1 (Nap1) or mobilized during transcription and DNA damage by facilitates chromatin transcription (FACT)[29-32]. Coupled to their

assembly, nucleosomes are repositioned by a range of ATP-dependent chromatin remodelling enzymes[33]. Together these chaperones ensure the proper assembly and distribution of nucleosomes along DNA to form chromatin with a regular repeat length[34]. When fully assembled, nucleosomes are stable and show very little spontaneous dissociation/turnover of the core histones, although histone H1 remains mobile on the timescale of minutes[19,35,36]. They form a barrier to recognition of the underlying DNA. At the entry/exit regions however, experiments suggest that the DNA is undergoing breathing motions providing a starting point for recognition, unwinding and invasion of the nucleosomes[37,38]. The disassembly of nucleosomes and the associated intermediates have been extensively studied. These studies rely in particular on nucleosomes reconstituted in vitro from recombinant histones and DNA containing NPS, followed by disassembly either by dilution or by increasing monovalent salt concentrations[39-42]. With this, the key steps of disassembly have been shown to be the reverse of assembly; H2A-H2B dimers are lost at a salt concentration of about 700mM NaCl, whereas $(H3\text{-}H4)_2$ is lost as a tetramer at around 950mM NaCl. Further, the interface between the H2A-H2B dimers and the H3-H4 tetramer is opened up at salt concentrations of ~600mM, functioning as an intermediate prior to H2A-H2B loss[43,44].

1.1.2. Chromatin structure beyond the nucleosome

Hierarchical layers of chromatin organization

Several layers of organization and epigenetic information beyond the nucleosome core particle allow the packing of the several meters of DNA within the confines of the interphase nucleus and to contain cell-specific developmental programs. Distinct modifications and conformational topologies are linked to each structural level and further explanation of these is provided in the following sections (**Figure 1.1.2**)[45].

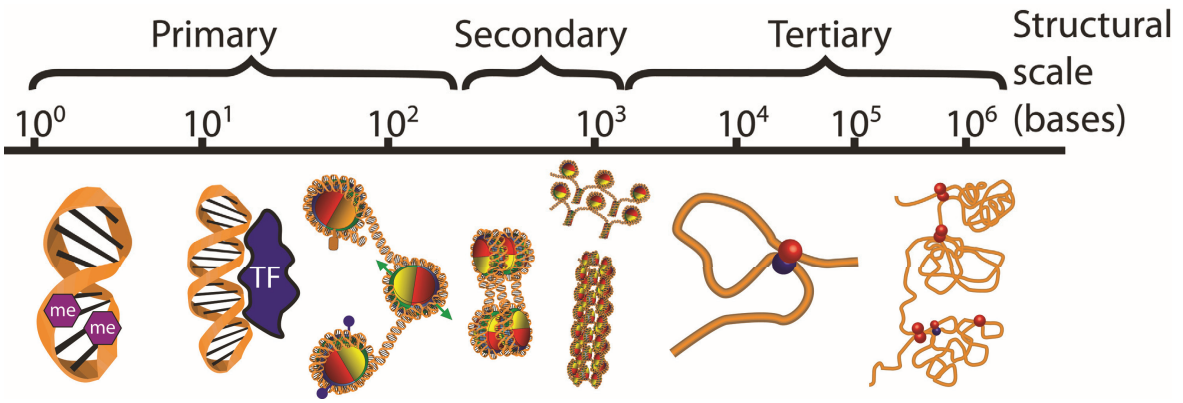

Figure 1.1.2. *Hierarchical layers of chromatin organization. On the scale of bases. At the scale of individual bases (10^0 bases), single cytosines or adenines can be methylated encoding information beyond the DNA sequence. Transcription factors can bind stretches of ~5-30bp DNA (10^1 bases) sequence-specifically. Nucleosomes form on approximately 150bp of DNA (~10^2 bases) and carry PTMs or histone variants as an additional layer of instructions. Four or more nucleosomes (0.5-1·10^3 bases) can pack into distinct local structures facilitated by inter-nucleosomal interactions. Stretches of chromatin (10^4-10^5 bases) are the basis of anchoring points resulting in chromatin extrusions and loops. Multiple loops in distinct territories (10^5-10^6 bases) are referred to as topologically associated domains (TADs). Figure based on Risca and Greenleaf, 2015.*

Despite more than half a decade of intense research, aspects about even the most elementary aspects of chromatin structure and regulation are still being discovered as exemplified by the recent discovery of adenine methylation in mammalian genomes[46]. Several aspects of the primary structural scale and the involvement of nucleosomes in this have been elucidated as outlined above. Beyond the primary level however the principles governing secondary and tertiary level structure formation and dynamics remain poorly understood.

Chromatin secondary structure

Each nucleosome core particle is interconnected by 20-80bp of DNA referred to as linker DNA. Because of this linkage, chromatin isolated from nuclei can appear as extended arrays of beads on a string[47]. Chromatin conformation plays an important role both in packaging, but particularly in controlling access to the underlying DNA. The accessibility of the underlying DNA has often been investigated in a cellular context using deoxyribonuclese I (DNAse I) to identify sites sensitive to enzymatic digestion (DNAse hypersensitive sites, DHS), revealing correlations between accessibility and transcriptional activity[48]. Early work on determining the structure of chromatin involved brief digestion with micrococcal nuclease I (MNaseI), which is a restriction enzyme cleaving only non-nucleosomal DNA. Subsequent isolation of chromatin segments combined with electron microscopy revealed a propensity of chromatin to form a fiber-like structure of ~30nm in diameter. Based on the microscopy images, this led to suggestions that chromatin adopts either a zig-zag arrangement of nucleosomes in a two-start helix or forms the basis of a solenoid-like structure (**Figure 1.1.3**)[49-51].

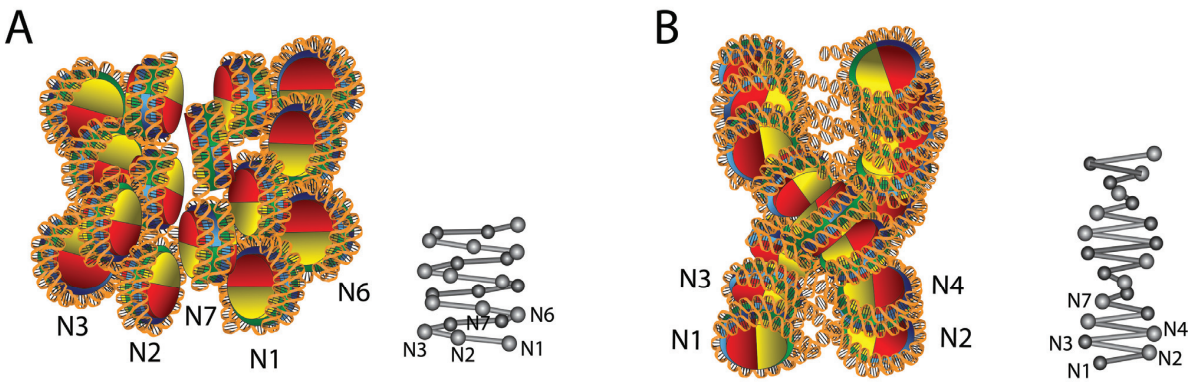

Figure 1.1.3. *Secondary chromatin structures based on 30nm-fiber formation.* ***A.*** *Solenoid-like chromatin structure organizing the nucleosomes in large helical conformations with nucleosomes N+6 stacked between nucleosomes N and N+1. Simplified scheme showed in ball-stick format with nucleosomes being balls and linker DNA being sticks.* ***B.*** *Two-start helix structure showing formation of two intertwined helical columns of nucleosomes with stacking of nucleosomes N and N+2. Ball and stick format shows the DNA zig-zagging back and forth between the two stacks. Based on Robinson, Rhodes et al. 2006.*

Complementary to top-down approaches of imaging and isolation from nuclei, bottom-up methods using reconstitution of chromatin from isolated histones on short 12-mer cloned repeats of NPS (5S RNA) allowed further studies of the reversible condensation of chromatin using analytical ultracentrifugation (AUC)[52,53] . With AUC, specific chromatin states induced by varying salt concentrations correlated with defined sedimentation velocities showing a gradual transition from decompacted to compacted chromatin[53,54].

Combining defined NPS arrays with recombinant histone proteins allowed reconstitution of fully recombinant chromatin samples thereby overcoming challenges with heterogeneity and endowing full control of the components for mechanistic and structural investigations. Employing this along with AUC revealed that the H4 N-terminal tail is highly important for packaging of chromatin and that acetylation at lysine 16 disrupts chromatin folding[55,56]. Mutants of H4 and H2B allowing disulfide formation in a compacted fiber between adjacent nucleosomes followed by digestion of linker DNA resulted in two stacks of nucleosomes, showing the formation of a two-start helix as the secondary structure of DNA(**Figure 1.1.3.A**)[57]. Evidence for such a packing of the chromatin fiber was also obtained with the crystal structure of a tetranucleosome[58]. Further structural evidence pointing towards such a model was obtained with the cryo-em structures of 12-mer and 24-mer arrays (177bp and 187bp repeats) with or without the H1 linker histone. In addition to the tetranucleosome stacking, this structure suggested that higher-order structure may be repeats of tetranucleosomes with slight kinks between each tetranucleosome stack[59]. Similar has recently been observed from force-extension of tetranucleosomes, where such structures extend in a fashion compatible with two-start stacking[60].

Other electron microscopy and force spectroscopy measurements using similar reagents however with longer linker lengths and more nucleosomes (30-60) point towards the nucleosomes arranging into the solenoid-like structure[61,62]. Studies using reversible compaction, crosslinking and decompaction, suggest rather that the contacts formed is of a relatively heterogenous form, with possible interactions between nucleosome N with N+1-7, however with a predominance of contacts with nucleosomes N1-3[63]. Super-resolution imaging and sequencing based methods point towards the possibility for dense local packaging of neighbouring nucleosomes in chromatin in two-start conformations *in vivo*[64-66]. Nevertheless, such packaging is probably limited to short segments of chromatin, as data from investigations on tertiary structure in a cellular context are incompatible with long contiguous segments of chromatin 30nm-fiber conformations[67,68]. Regardless, of the conformation adopted, there is consensus that the arrays can be reversibly compacted and decompacted directly by PTMs and chromatin effectors binding either to PTMs or to other parts of the chromatin topology[69-71].

As suggested from these studies with often contrasting conclusions, the secondary conformations adopted by chromatin remain highly disputed. Although chromatin arrays can be reconstituted and their global structures investigated by e.g. AUC, imaging and structural analysis, not much is known about the local folding pathways from decompacted to compacted chromatin. Further these static methods disallow investigations into the dynamic transitions and intermediates observed in the chromatin secondary structure. Since the conformational fluctuations regulate access to the nucleosomal surface and DNA, it is essential to understand the nature of these movements. Still, only two studies with di-nucleosomes and tri-nucleosomes respectively have investigated the effect of neighbouring nucleosomes on DNA unwrapping and likely folding pathways in these short arrays[71,72]. Two to three nucleosomes however have limited potential to form higher-order structures.

Chromatin tertiary structure

Beyond secondary structure, chromatin folds extensively to result in contacts between DNA segments separated by more than 10kbp in sequence[68]. The investigation of these have been facilitated by variations of chromatin conformation capture methods (3C)[73]. 3C involves cross-linking of chromatin in intact cells, chromatin solubilization and digestion, followed by ligation, reversing crosslinks and quantitative PCR using pairwise primers for each of the two genome locations expected to interact[74]. These steps produce interaction frequency maps revealing long-range interactions as increases in

interaction frequency disproportional to the linear sequence distance in the genome. This procedure has been extended by modification of the digestion, ligation steps and/or primers used for capture and sequencing to allow mapping of one DNA region vs all (4C), many by many (5C) or all vs all (Hi-C and Micro-C)[64,67,75-78]. Such chromatin capture strategies have shown that chromatin is organized into loops of varying sizes ranging from a few kbp to the MBp scale, e.g. between enhancers, promoters and their underlying genes. Particularly Hi-C has been instrumental in revealing several important aspects of higher-order chromatin folding. Maps produced by Hi-C show that chromatin is segmented into distinct topologically associated domains (TADs), that has an increase in interaction frequency with itself relative to other chromatin regions. Formation of TADs occur by extrusion of multiple loops across the genome by scaffold proteins condensin and cohesion[79-81]. Transcriptionally active chromatin regions further has an increased propensity for interaction with other active regions, whereas interaction is reduced between active and silenced regions. Disruption of TADs have recently been directly linked to the development of malformed limbs, underlining the importance of higher-order chromatin organization[82,83]. At a chromosome-level, the location of each chromosome within the cell nucleus as non-random has been revealed by in vivo mapping of genomic locations using DNA-FISH[84]. Each chromosome occupies distinct territories relative to the nuclear center that are specific to cell-type and tissue[79].

1.1.3. Chromatin modifications and their function

The cells in a eukaryotic organism principally contain the same genotype. This however gives rise to multiple different phenotypes as exemplified by the cells found throughout the tissue of a multicellular organism[85]. Also, several traits can be effectuated by exposure to changes in the environment and passed on from generation to generation, either through mitosis or meiosis[86,87]. Classic examples of such environmental changes altering phenotypic traits include twin studies and studies on children in the Netherlands born from mothers pregnant during the dutch famine in 1944-1945[88,89]. These children subsequently had an increased risk of cardiovascular and metabolic diseases owing to the in utero nutritional deprivation[88]. Such abnormalities have also been documented in grandchildren of those mothers. Similarly, monozygotic twins carrying the same genotype exposed to different environments can develop very differently[89]. Collectively, phenotypic differentiation from one genotype and transgenerational non-DNA based inheritance are linked to epigenetics. Epigenetics describe the changes occurring in chromatin beyond the DNA sequence. These include

changes in DNA methylation patterns, histone PTMs, distribution and function of chromatin-associated effector proteins and changes in chromatin structure. The importance of epigenetics for normal functioning of eukaryotic organisms has made it an area of significant research interest.

Histone PTMs

The discovery that acetylation and deacetylation of core histone proteins were directly associated with activated and repressed transcription respectively, sparked a significantly increased interest in understanding the mechanisms of action associated with histone post-translational modifications (PTMs)[90-93]. More than ~550 PTMs and sites in the core histones have been identified and several of them associated with specific functions (Few examples in **Figure 1.1.4**)[94].

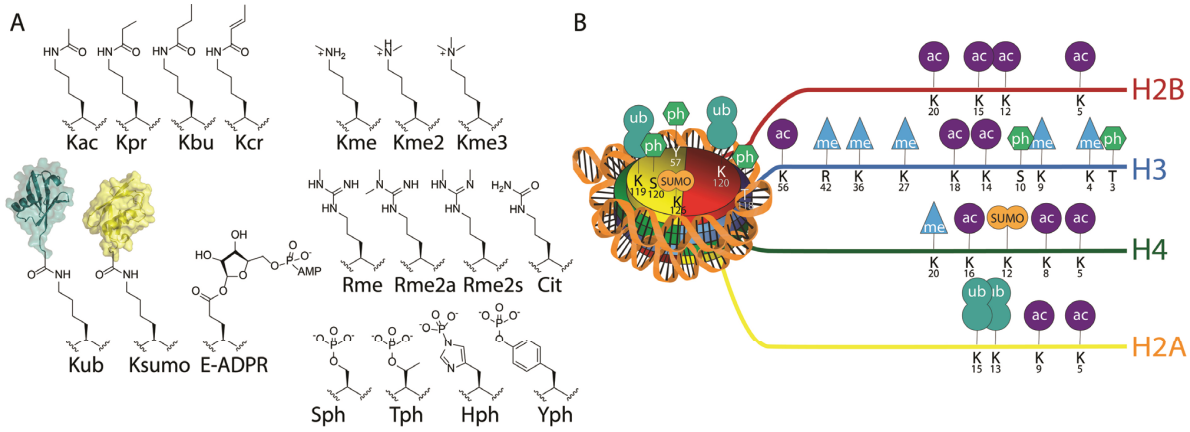

Figure 1.1.4. *Histone PTMs and their nucleosomal distribution.* ***A.*** *Chemical structures of selected histone modifications described for lysine, arginine, glutamate, serine, threonine, tyrosine and histidine.* ***B.*** *Locations of some of the key histone PTMs for which there is knowledge about the functional importance.*

These include well-characterized PTMs like lysine acetylation, mono –, di – and trimethylation, ubiquitylation and SUMOylation, arginine mono – and dimethylation (symmetric and asymmetric) as well as deimination, serine phosphorylation and GlcNAcylation, glutamate ADP-ribosylation and proline isomerization[13]. Less well studied histone PTMs include threonine, tyrosine and histidine phosphorylation [95-97]. Recently also multiple varieties of lysines acylated with longer side-chains have been associated with metabolic regulation[98,99]. Additional modifications without much functional characterization include lysine hydroxyisobutyrylation, formylation, malonylation and glutarylation The discovery of the link between diverse cellular functions and sets of histone PTMs led to the formulation of the histone code hypothesis[100,101]. This proposed that coexisting PTMs function as codes and result in specific transcriptional outcomes of the underlying genes and that PTMs could mutually affect each other through crosstalk. Key to this, was a high correlation between specific

PTMs with transcriptionally silenced heterochromatic regions (H3 K9me2/3, H3 K27me2/3, H2A K119Ub) and active less dense euchromatic regions (H3 K4me2/3, H3 K9ac, H3 K36me3, H3 K27me3, H3 K79me2/3, H4 K5/8/12/16/20ac and H2B K120ub). Refinement of these categories based on associated chromatin proteins and histone PTMs have led to characterization of 5-51 distinct chromatin states dependent on the cell type and analysis method used [102,103]. The PTMs in these different types of chromatin exert their function mainly through two distinct mechanisms. Either they directly alter the interactions between the DNA and the histones or histones between neighbouring nucleosomes, thereby changing the stability of the nucleosome or the conformation of chromatin. Alternatively they function as recruitment platforms for a range of chromatin-associated proteins that then alter the chromatin landscape to control transcription, replication and DNA damage repair.

Direct modulation of nucleosome and chromatin structure by PTMs

The PTMs directly affecting nucleosome or chromatin structure and dynamics are typically either caused by changes in electrostatic charge or relatively large modifications altering the bulk of the histones. PTMs changing the electrostatic charge of side chains are phosphorylations of serine (threonine, tyrosine and histidine) and acetylation of lysine side chains neutralizing a positive charge. Bulky modifications like ubiquitylation and SUMOylation can impact chromatin stability. Phosphorylation of tyrosine 41 near the entry-exit region increases nucleosome DNA unwrapping[104]. Similarly, several lysine acetylations have been directly linked to disruption of histone-DNA contacts resulting in less tight association of the DNA with the nucleosome. These acetylation sites are located at the points of contact between the DNA superhelix and the histone octamer. Dependent on the location, they have distinct influences on nucleosome stability[105,106]. Specifically H3 K56 and H3 K64 contacts the DNA near the entry/exit region of the DNA from the nucleosome. Acetylation of this site results in increased DNA breathing and more frequent unwrapping and a site for invasion of the nucleosome[105-107]. In contrast, the contact points for H3 K115 and H3 K122 are near the nucleosome dyad[108]. Acetylation of these result in internal destabilization of the nucleosome and incoming factors able to unwind the first wrap of DNA would then more easily be able to traverse the entire nucleosome[105]. This is highlighted by H3 K122ac exclusively being found in euchromatic regions and being associated with increased transcription[109]. Histone ubiquitylation has been linked to both stabilizing and destabilizing effects on the nucleosome[110-113]. The location of most ubiquitin sites on the surface of the nucleosome likely don't directly affect the stability of DNA-histone or

histone-histone interactions as demonstrated for H2B K120 and H2A K119, but may rather alter the ability of other factors to bind the nucleosome[114].

As described previously the H4 N-terminal tail is important for inter-nucleosome stacking, and acetylation of this at H4 K16 disrupts interactions between adjacent nucleosomes[55,56]. Similarly ubiquitylation and SUMOylation of H2B and H4 respectively both localized on the nucleosomal surface hinder inter-nucleosome stacking and prevent full compaction of chromatin secondary structure[69,115]. The bulk of histone PTMs however exert their effect by functioning as recognition sites for chromatin-associated effector proteins.

1.1.4. Chromatin-associated effector proteins

Writers and erasers

Chromatin effectors refer to the large group of proteins responsible for installing (writers), removing (erasers) and binding to (readers) histone PTMs[116]. PTMs are attached by designated enzymes either on the individual histones shortly post-translation (e.g. H3 K56ac) or after nuclear entry and introduction into chromatin (e.g. H4 K20me2)[117,118]. Small histone modifications are attached using suitable cofactors. Histone acetyltransferases (HATs) utilize acetyl-CoA for acetylation, lysine methyltransferases (KMTs) S-adenosylmethionine for methylation and kinases use ATP for phosphorylation. Protein modifications SUMO and ubiquitin are attached through an E1-E2-E3 ubiquitin ligase cascade common to this type of PTM. Like most aspects of chromatin regulation, PTMs are dynamically modulated. PTMs are reversibly removed by erasers. Mechanisms for removal of acetylation, phosphorylation and ubiquitin/SUMO were already known from non-histone proteins. No enzymes for removal of methylation were however known and it was thought that this modification was either permanent or could only be removed by dilution through replication[119]. This view has been revised since the identification of several families of histone lysine demethylases (KDMs) and the elucidation of their mechanisms for removal of the methyl mark [120-122]. The enzymatic activities to write and erase histone PTMs are often harbored within proteins or complexes containing also the ability to recognize the written/erased target modification. Thereby this linked recognition and activity provides a feedback mechanism to control the spreading or containment of PTMs[123].

Readers

The chromatin associated effector proteins contain a vast number of protein domains evolved to bind target modified chromatin[124,125]. These distinguish both the type (or the lack) of modification and in most cases also the sequence context of the target modified histone. Domains recognizing almost all types of histone modifications have been discovered, and the structural basis for their distinction between the target and non-target residues elucidated[94,124]. Domains targeting methylated lysines and arginines include ATRX-DNMT3-DNMT3L domains, ankyrin, bromo-adjacent homology, chromo-barrel, chromodomain, double chromodomain, malignant brain tumor, plant homeodomain, PWWP, tandem tudor domains, tudor and WD40[94]. Common to these is that the methylation-state specific readout is facilitated by aromatic residues lining the binding pocket combined with aspartates or glutamates to recognize the lack of methylation [125]. Recognition domains targeting phosphorylation, acetylation, crotonylation, unmodified H3 and H4 as well as ubiquitylation have also been identified[94,126,127]. Structures of motifs bound to ubiquitylated histones have only recently emerged as they seem to require the nucleosomal context and preparing such nucleosomes can be challenging. The typical dissociation constant for most of these reader domains lie in the range of the low micromolar (for methyl-reading domains) to high micromolar (for some acetyl-binding domains), making them relatively low affinity binders on their own.

Multivalent readout of PTMs

The binding modules recognizing and discriminating between different histone PTMs often exist in homo – or heterooligomeric protein complexes with multiple of the same or different recognition modules[128]. This results in favourable thermodynamics by additively increasing the binding enthalpy, while not changing entropy, thereby improving the affinity towards target sets of PTMs. Further, it provides a mechanism whereby patterns of histone PTMs can be recognized in a multivalent fashion to ensure their co-existence[129]. Such binding can occur in multiple different ways either intra-nucleosomal or inter-nucleosomal (**Figure 1.1.5**).

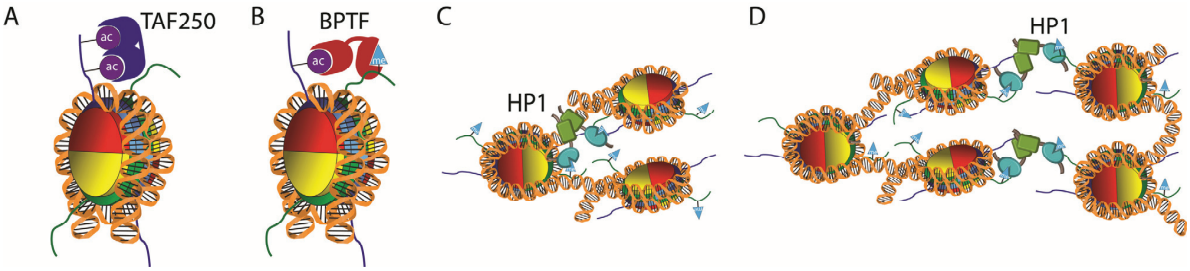

Figure 1.1.5. *Multivalent combinatorial readout of histone modifications.* ***A.*** *Intranucleosomal, cis-histone binding of an effector.* ***B.*** *Intranucleosomal, trans-histone binding of the example BPTF to acetylated H4 and methylated H3.* ***C.*** *Inter-nucleosomal, intra-fiber binding of HP1.* ***D.*** *Inter-nucleosomal, inter-fiber binding of HP1.*

The multiple modifications possible in one tail allows the potential cis-histone binding of two or more PTMs within the same tail. An example of this is the double bromodomain of the TFIID subunit TAF250 capable of recognizing double acetylated H4[130]. In contrast to other types of multivalent binding, cis-histone binding, where two different PTMs in the same histone tail are recognized, can be probed with histone peptide array binding and do not require a nucleosomal template. The target recognition sites can also exist in different histones or within the DNA of the same nucleosome for trans-histone/DNA binding. Prominent examples of this from the past decade include the NURF remodelling complex subunit BPTF recognizing H3 K4me3 and H4 K16ac, UHRF1 recognizing H3 K9me3 and hemimethylated DNA as well as p53 binding protein (53BP1) binding to H4 K20me2, H2A(.X) K15ub and H2A.X S139ph[127,131-133]. The pseudosymmetry of the nucleosome further allows for binding of the same target modification but from different tails in the pair of histones found within the same nucleosome. Finally, effectors can bind different nucleosomes within the same chromatin segment or from different chromatin fibers to compact one fiber or bridge multiple fibers, examples of this include heterochromatin protein 1 (HP1, described in further detail later) and polycomb repressive complex 1[70,134,135]. Although there is a consensus that multivalent engagement of chromatin templates can increase affinity of a chromatin effector towards target chromatin, the kinetic mechanisms underlying this are not well understood. Nuclear proteins are highly mobile, and therefore it is unlikely that prolonged residence times alone explain the advantages gained by multivalent binding. To study the mechanisms of interactions between effectors and modified chromatin, it is necessary with access to modified chromatin. Histones isolated from nuclei or enzymatically modified in vitro often yield heterogenous mixtures, making it challenging to draw conclusions from experiments with such reagents. Instead a number of chemically based methods have been developed to site-specifically modify proteins and provide homogenously modified proteins.

1.2. Protein and chromatin engineering

1.2.1. Introduction of post-translational modifications in (histone) proteins

Modified histone tail peptides generated by solid-phase peptide synthesis (SPPS) in absence of DNA and the histone core have provided important insights into modifications targeted by particular chromatin-associated effectors and their intra-tail cross-talk, using e.g. peptide microarrays[136,137]. Most histone PTMs however have an effect that can only be appreciated in the context of nucleosomes or chromatin arrays due to interplay between DNA and the histones, different histone tails and/or neighbouring nucleosomes[128,138]. Multiple chemistry-based methods have been developed that allow site-specific introduction of modifications, for example to prepare homogenously modified histones [116,139]. Three main approaches are used for this; Amber suppression mutagenesis, cysteine-directed conjugation (**Figure 1.2.1**) as well as native/expressed protein ligation (NCL/EPL). NCL and EPL are described in further detail in section 1.2.2 and onward.

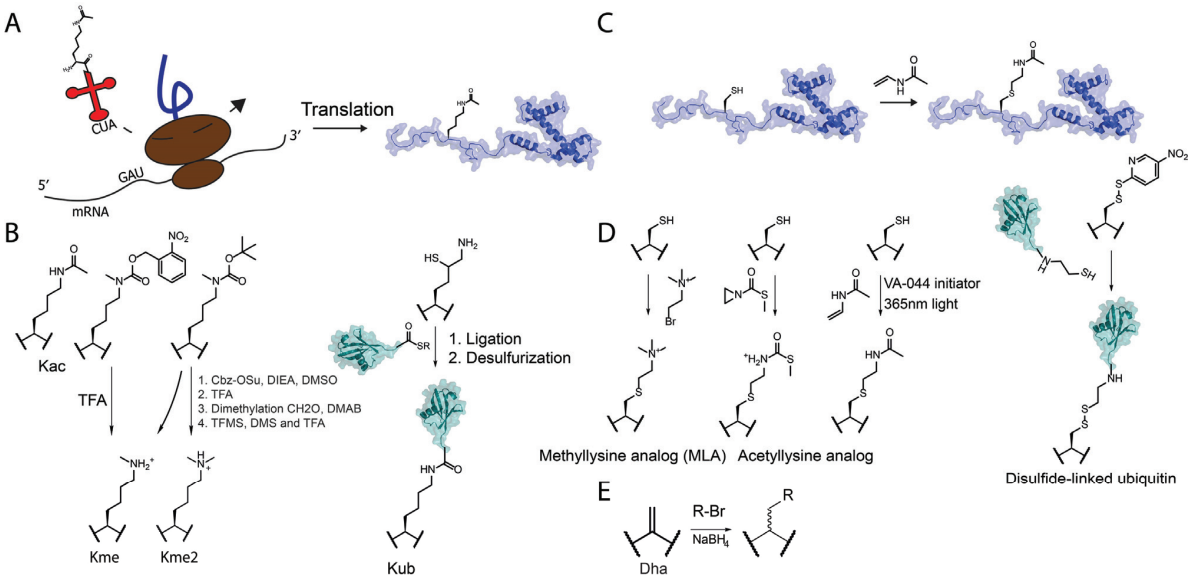

Figure 1.2.1. *Generation of histones containing PTMs or analogs.* ***A.*** *Amber suppression mutagenesis involves the recoding of the ribosome to introduce non-proteogenic amino acids like acetyllysine from an amino acid-tRNA on the amber codon usually signaling stop of translation.* ***B.*** *Examples of this for introduction of histone PTMs include direct translation with acetyllysine, protected methyllysines for subsequent deprotection to produce mono – or dimethyllysine and δ-thiolysine for subsequent ligation to ubiquitin.* ***C.*** *Cysteine conjugation takes advantage of the nucleophilicity of the thiol in cysteine generated in recombinant proteins for reaction with suitable electrophiles to produce covalent analogs of the PTMs of interest.* ***D.*** *Cysteine conjugation has been used to produce methyllysine and acetyllysine analogs where the γ-C in the lysine side chain is replaced with a sulfur, and to produce disulfides with ubiquitin and SUMO.* ***E.*** *Treatment of dehydroalanine with alkyl halides and borohydride in an oxygen-free atmosphere results in modified racemized side-chains.*

Amber suppression mutagenesis involves the usage of the ribosomal machinery to encode a modified amino acid instead of the amber stop codon UAG[140]. Some bacteria use the UAG codon to encode pyrrolysine as a 21st proteogenic amino acid. By evolving pyrrolysyl-tRNA synthetases to recognize

and charge the UAG codon tRNA with differently modified amino acids, the altered amino acids can be incorporated in the target protein recombinantly. This has been used to directly introduce acetylated lysines into histones[107,141]. The subtlety of lysine methylation, has made attempts of evolving tRNA synthetases to directly introduce this modified amino acid unsuccessful. Instead protected versions of mono- and dimethylation have been introduced followed by orthogonal reactions and deprotections after purification of the protein[142,143]. Also, thiol containing lysines have been introduced by amber suppression mutagenesis to produce a handle for subsequent ligations with ubiquitin to yield modified proteins with a native isopeptide bond to glycine 76 in ubiquitin[144]. Thiol conjugation takes advantage of the nucleophilicity of the thiol group in cysteine relative to other side chains in proteogenic amino acids. Alkylation of the cysteine with suitable compounds produces analogs of the modified lysine side chain. Methyl-lysine analogs have been produced by reaction with 2-haloethyl amines to produce mimics of mono-, di- and trimethylated lysine[145]. Symmetric and asymmetric methylation of arginine side chains have been achieved in a similar fashion with unsaturated amidines[146]. Thiol-ene addition with N-vinyl-acetamide has yielded an analog of acetylated lysine[147]. Cysteine side-chains have also been used to generate disulfide conjugates of larger protein modifications like SUMO and ubiquitin that can easily be reversed by reduction [69,115,148]. Finally dehydroalanine can form the basis for the formation of similar thiol-containing side-chain analogs[149]. Dehydroalanines can be produced site-specifically from unique phosphoserines or cysteines incorporated recombinantly[150,151]. Recently, alkylation of these with radicals generated from suitable iodide organic compounds were used to yield proteins containing post-translational 'mutations', with the main drawback being side-chain racemization[152-154].

1.2.2. Native and expressed protein ligation

Solid-phase peptide synthesis

Before the '60s, peptides were typically synthesized in solution with multiple intermediate purifications to ensure purity and avoid side-reactions. Challenges with peptide solubility and limitations in the length of peptides available however prompted the development of solid-phase peptide synthesis in the lab of Bruce Merrifield in the '60s to '70s. In SPPS, the peptide of interest is conjugated to a solid support at the C-terminal and extended one amino acid at a time from the C- to the N-terminal in contrast to ribosome-mediated protein synthesis. Prior to each coupling, the amino acid is activated, typically with hydroxybenzotriazole (HOBt), hydroxyazabenzotriazole (HOAt), their uronium salts (HBTU and HATU) or a combination of these to form active esters. After each

coupling, excess amino acid is washed away to leave the growing peptide on the solid support with the additional amino acid. SPPS relies on orthogonal protection of the amino acid Nα and possible reactive groups on the amino acid side chain. The Nα protection prevents multiple of the same amino acid to be coupled. This needs to be orthogonally removed after each coupling and SPPS approaches are generally divided based on the Nα amine protection group. Tert-butyloxycarbonyl (Boc) or 9-fluorenylmethyloxycarbonyl (Fmoc) is used for this. The protection scheme used for the Nα amine dictates the procedure used for both deprotection and for cleavage from the solid support of the final peptide. The Boc group is deprotected acidolytically with trifluoroacetic acid and the peptide cleaved from the resin along side-chain protection groups with anhydrous hydrofluoric acid (HF)[155]. The harsh conditions necessary for cleavage with Boc SPPS make them incompatible with acid sensitive phosphoryl and glycosyl modifications. Fmoc deprotection is carried out under mild basic conditions typically using piperidine, whereas final removal of side-chain protection groups and cleavage from the resin is done with TFA. Automation of the deprotection, coupling and washes with peptide synthesizers enables synthesis of peptides without much manual intervention. Addition of excess amino acid for each coupling allows reactions to proceed to near completion. With an efficiency of 98% for each coupling however, synthesis of a 40 amino acid peptide would still only provide a theoretical yield of 67%. Therefore SPPS allows routine synthesis of peptides with a length of 40-60 amino acid residues. Most proteins are however longer than this. Joining peptide fragments would allow synthesis of polypeptides resembling protein domains or even small proteins in their full length.

Native chemical ligation

A wide variety of approaches have been developed to join peptide fragments often resulting in non-native bonds including thioester ligation[156], thioether ligation[157], selenocysteine ligation, histidine ligation[158], thiazolidine ligation[159], traceless staudinger ligation[160,161], KAHA ligations[162,163], oxime ligations[164], salicylaldehyde ester-mediated ligations[165] and 3+2 cycloadditions[166]. Few have however gained more traction than native chemical ligation. Native chemical ligation (NCL) provides a chemoselective route for the joining of two unprotected polypeptide fragments in aqueous conditions to form a native amide bond[167-171]. For NCL, The first fragment contains a C-terminal thioester and the second fragment harbors a C-terminal cysteine. Rate-limiting trans-thioesterification of the cysteine thiol to the thioester, followed by an irreversible S to N acyl shift generates the native peptide linkage (**Figure 1.2.2**).

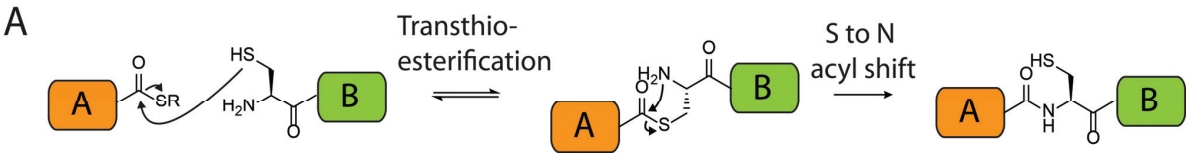

Figure 1.2.2. *Mechanism of the native chemical ligation reaction. The first polypeptide fragment (A) contains a C-terminal thioester and the second an N-terminal thiol. Reversible transthioesterification resulting in leaving of the SR group and formation of a branched intermediate. This intermediate is resolved by an irreversible S to N acyl shift resulting in the formation of a native amide bond between the two fragments.*

This significantly increased the length of synthetically accessible peptides from a range of typically 40-60 to potentially 80-120. Several major barriers still limited the utility of native chemical ligation. First, the generation of the thioester was limited to Boc SPPS, as suitable approaches based on Fmoc SPPS were unavailable. The HF cleavage associated with Boc SPPS limited the number of laboratories that could prepare the required components. Second, the need for a cysteine at the ligation junction combined with the scarcity of cysteines throughout the proteome would necessitate the introduction of site-specific serine to cysteine or similar mutations. Third, the low reaction rate significantly limited the throughput of the reactions. Fourth, unless the target was a natively unstructured or a self-folding protein, refolding would be necessary after synthesis to generate the functional protein. Finally, until the identification of suitable orthogonal groups for the protection and/or masking of the thioester and cysteine thiol, the ligation would be limited to the protein terminals in two fragments and relatively short proteins.

Several important advances in peptide and protein chemistry since then have expanded the scope and utility of NCL to interrogate the structure and function of proteins and their modifications[168,169].

Thioester formation

The C-terminal thioester on one peptide fragment is essential to the NCL reaction. Methods to generate this by Boc SPPS were developed prior to the discovery of the NCL reaction[172]. This was done using either α-thioacids and conventional linkers or modified resin linkers prepared from or 3-mercaptopropionic acid[156,173]. These linkers are compatible with the cycles in standard in situ neutralization Boc SPPS and the final HF cleavage[155,167,171]. Upon cleavage, one approach yields a peptide thioacid convertible to a thioester by reaction with e.g. benzyl bromide or 5,5'-dithio-bis(2-nitrobenzoic acid). Cleavage from the linker prepared with 3-mercaptopropionic acid directly yields a corresponding thioester[174]. The standard deprotection step in Fmoc SPPS involving mild basic treatment with piperidine is incompatible with the thioester functionality leading to hydrolysis.

Although alternative Fmoc deprotection strategies compatible with the thioester functionality have been sought, the complexity and possible lowered efficiency in Fmoc removal for these have prevented widespread use[175,176]. Because of this along with the instrumental requirements for cleavage with Boc SPPS, significant effort has been devoted to developing facile methods for producing thioester peptides synthesized by Fmoc SPPS. Most of these depend on linkers that post-SPPS of the target peptide can be converted on resin or in solution to a thioester[177]. One of the most successful among these is the diamino-benzoic acid (Dbz) linker[178]. Post-SPPS and prior to cleavage, this can be converted with p-nitrophenylchloroformate into the N-acyl-benzimidazolinone (Nbz) followed by TFA cleavage. The peptide-Nbz can either directly be used for ligation or converted into a thioester to accelerate ligation rate. Alternatives take advantage of relative propensities for N→S acyl shifts to take place. Bis(2-sulfanylethyl) amido (SEA), N-sulfanylethylanilide (SEAlide) auxiliaries and N-methyl cysteine can all be triggered through pH adjustments or addition of phosphate to undergo these acyl shifts, that can form the basis for generation of more stable thioesters[179-183]. The general lack of commercial availability of these different linkers along with somewhat complex preparation procedures for many of them have limited their widespread use. One of the most facile and recently realized approaches involves the use of peptide hydrazides as masked thioesters[184]. Simple treatment of chlorotrityl resins with hydrazine converts the resin to produce a peptide hydrazide after finished synthesis instead of a peptide acid[185]. Oxidation of the peptide hydrazide with sodium nitrite to the peptide azide at low pH followed by exchange with a thiol compound yields the desired thioester[186]. The masking of the thioester rather than conversion to thioester on resin has the added advantage of allowing sequential ligations in the N – to C-direction or convergent ligations (Further described in multi-fragment ligations in chapter 5)[187].

Cysteine-free ligation and desulfurization

The residue used as the site for ligation is no longer limited to cysteine. The development of thiol-containing ligation auxiliaries attached close to the second fragment for ligation in place of a cysteine, facilitated ligation in absence of cysteine[188,189]. Post-ligation the auxiliary can be removed by light irradiation at particular wavelengths or additional cleavage reactions to result in glycine at the ligation junction[189-191] (**Figure 1.2.3.A**). The relatively labor-intensive synthesis of the auxiliaries and conditions required to remove some of them however limits their widespread use. Desulfurization provides an alternative means of retaining thiols prior to ligation and removing unprotected thiols throughout the protein afterwards. Initially metal-based protocols involving Raney-Nickel or

Pd/Al_2O_3 was used for this[192,193]. Limitations in protection group compatibility and epimerization of secondary alcohols however, prompted the development of metal-free radical-based desulfurization of ligation products (**Figure 1.2.3.B**)[194].

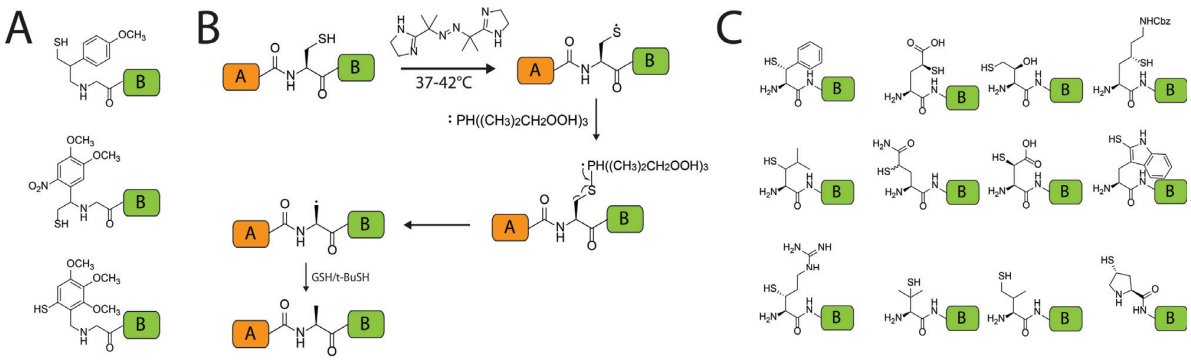

Figure 1.2.3. *Tools for cysteine-free ligation.* ***A.*** *Example ligation auxiliaries to enable ligation in absence of cysteine followed by cleavage or photoremoval.* ***B.*** *Mechanism of radical-initiated desulfurization with VA-044 (2,2'-Azobis[2-(2-imidazolin-2-yl)propane]) and tert-butylthiol or glutathione.* ***C.*** *Thiol-containing amino-acids coupled at the N-terminal of a target peptide fragment. Following ligation and desulfurization these leave phenylalanine, aspartate, threonine, lysine, leucine, glutamine, aspartate, tryptophan, arginine, valine or proline at the ligation junction.*

This reaction depends on a water-soluble radical initiator to first abstract a hydrogen from the thiol, followed by reduction with TCEP to form an alkyl radical intermediate. This is then propagated as a chain reaction by this alkyl radical abstracting a hydrogen from another target peptide thiol or from a buffer compound containing thiols, typically tert-butylthiol or glutathione added in >10-fold excess and the reaction can repeat itself (**Figure 1.2.3.B**)[194,195]. Ultimately this results in complete desulfurization of the target unprotected thiol(s) as it was initially demonstrated for conversion of cysteine to alanine, significantly expanding the number of positions in proteins that can be used for ligation, while maintaining a native protein sequence. Desulfurization to remove thiols combined with synthetic variants of the proteogenic amino acids containing thiols on the β- or γ-carbon of the side chain now allows native chemical ligation followed by desulfurization to yield native phenylalanine[196], glutamine[197], threonine[198], lysine[199], leucine[200,201], glutamate[202], aspartate[203], tryptophan[204], arginine[205], proline[206-208] and valine[195,209] at the ligation junction as well as isopeptide bonds in lysine[210] (**Figure 1.2.3.C**). Similarly, reactions using selenocysteine for ligation followed by desulfurization under aerobic or anaerobic conditions can yield serine or alanine at the ligation junction[211]. Selenocysteine for ligation, has an additional advantage of increased nucleophilicity thereby increasing the ligation rate. Finally it does not strictly require a radical initiator for TCEP-based deselenization, thereby providing orthogonality relative to thiol desulfurization[168,211].

One-pot ligation and desulfurization reactions

Initial NCL reactions were carried out with high pKa alkyl thioesters (e.g. mercaptoethanesulfonate). The sluggish rate of these thioesters in NCL however led to the search for thioesters with accelerated kinetics. The rates of ligation reactions in absence of suitable thiol catalysts and the number of steps required for both ligation and desulfurization with intermediate purifications significantly hampered both the throughput and yield of the reactions. 4-mercaptophenyl acetic acid (MPAA) catalyzes NCL rates such that reactions could typically proceed to 90% completion in 2-3-fold shorter time than with MESNA thioesters alone[212]. Despite this improvement in ligation rate MPAA and other thiophenols simultaneously inhibit radical-initiated desulfurization by functioning as radical-scavengers[213]. Therefore MPAA peptide thioesters either require pre-formation and intermediate purification to allow ligation followed by desulfurization in the presence of 1 equivalent of MPAA or it can be removed by capture or desalting[214,215]. Recently, alternative alkyl thiols have been identified that enable rapid ligation to take place, while still allowing subsequent desulfurization without intermediate purification[216,217]. These include trifluoroethanethiol (TFET) and methylthioglycolate (MTG). Thioesters formed with both of these allow ligation to take place at rates comparable to those of MPAA and also enable desulfurization in the same pot after completion of the ligation reaction without intermediate removal of the thiols. This increases the yield achievable with ligation and desulfurization by skipping an intermediate purification, however often at the cost of requiring more thioester due to competing hydrolysis.

Inteins and EPL

Proteins with a length beyond 100-120 residues if fully synthetic, require ligation of more than two fragments either sequentially or convergently (see chapter 5). Often however the interest in applying the possibilities associated with chemical synthesis (e.g. post-translational modification of side-chains, backbone modifications, fluorescent labeling) is restricted to particular segments in the protein of interest (POI). This led to the development of methods to produce recombinant proteins (With few length restrictions) with C-terminal thioesters or N-terminal cysteines for ligation with SPPS-generated peptides[218-220].

C-terminal thioesters in recombinant proteins are generated using intervening proteins (inteins)[221,222]. Inteins can be considered analogous to introns in RNA; They are elements that are removed to join the two flanking regions and form a product. Whereas RNA splicing proceeds through the action of

the spliceosome, inteins are autocatalytic protein modules that process their own excision and splicing of the flanking external proteins (exteins) in protein transplicing (PTS) (**Figure 1.2.4.A**).

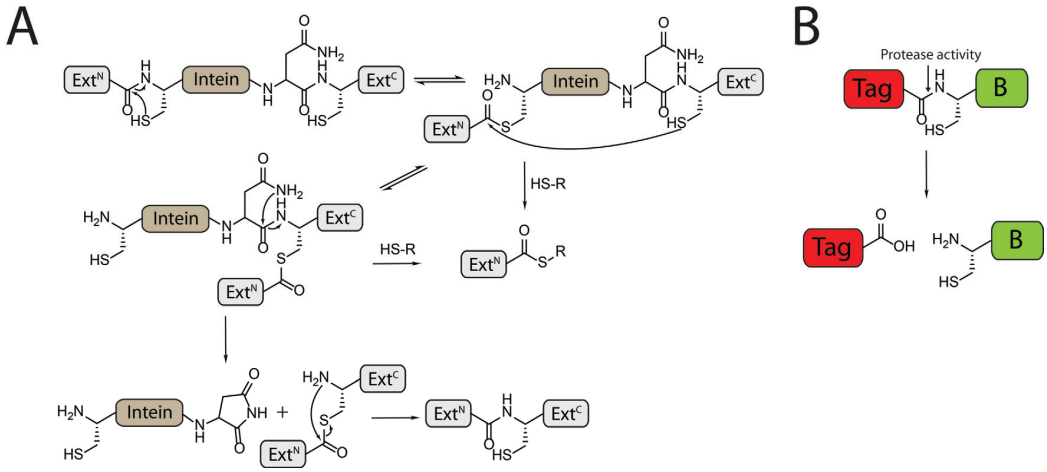

Figure 1.2.4. *Protein splicing and preparation of recombinant components for EPL.* ***A.*** *An intervening protein between two exteins undergoes an N to S acyl shift followed by transthioesterification to form a branched intermediate. If the asparagine at the splice junction is left intact, the branched intermediate is resolved into a succinimide and the native amide bond results from an S to N acyl shift. Intein mutants of asparagine and/or cysteine at the C-terminal of the intein can be thiolyzed with exogenous thiols to produce the recombinant thioester.* ***B.*** *Recombinant fragments with an N-terminal cysteine are produced by expression as a fusion protein followed by cleavage to reveal the free N-terminal cysteine*

Intein splicing starts at the N-terminal of the intein with an N to S acyl shift followed by transthioesterification from the C-terminal of the intein leading to a branched intermediate[223,224]. Resolving of the branched intermediate occurs through succinimide formation by the intein's C-terminal asparagine and a final S to N acyl shift in the spliced exteins lead to the formation of a native amide bond[225,226]. Mutation of the C-1 asparagine to alanine prevents the final steps of intein splicing, whereas mutation of the C+1 cysteine prevents transthioesterification[226,227]. The initial thioester formed after N-to-S acyl shift or after transthioesterification can be thiolysed using small exogenous thiols (e.g. MESNA) for the formation of the desired protein thioester[218,219,228,229]. The thioester can subequently be purified for later usage or immediately employed for ligation reactions with a second fragment containing a thiol and a primary amine. Inteins exist in nature both as contiguous segments resulting in cis-splicing, but also as split inteins where the N- and C-terminal parts of the intein give rise to trans-splicing. Individually, the segments of artificial or natural split inteins are disordered, however upon binding to each other with nanomolar affinity, they form the intein fold allowing them to join the neighbouring extein fragments through the same mechanism as cis-splicing inteins[230-232]. Typically, split inteins catalyze the splicing reaction much faster than cis-acting inteins[233,234]. These traits led to the development of a split intein based column where a double-alanine mutant of the Int^C half of the split intein is covalently immobilized on a solid support[235]. The other half is fused at the

C-terminal to the POI. Incubation of the POI-Int^N fusion from crude lysates or purified intermediates leads to column binding and the initial trans-thioesterification at the POI-Int^N junction. Washing of the column followed by thiolysis releases the pure POI-SR thioester for subsequent purification or ligation. Relative to thiolysis of mutated cis-splicing or contiguous trans-splicing inteins fused to the POI, this procedure avoids premature cleavage due to the dependence of both parts of the split intein, it reduces unwanted premature hydrolysis, is faster, and elutes the purified protein thioester without having to separate this from the intein components as these remain on the column. Importantly the intein column can be re-used after a simple boiling in SDS to dissociate the bound Int^N fragment, whereas the covalently linked Int^C remains on the column.

Recombinant proteins with C-terminal cysteines are typically generated by expressing a fusion protein between the POI linked at the N-terminal to a cleavable tag/protein (**Figure 1.2.4.B**). Protease-mediated cleavage of the N-terminal tag/protein will leave the cysteine at the N-terminal ready for ligation. Examples of proteases and associated tags/proteins previously used for this include methionine amino peptidase to remove an initiator methionine[236], factor Xa protease to remove the target sequence IE/DGR[220], tobacco etch virus (TEV) protease for removal of ENLYFQ[237], Ubiquitin like protease 1 (Ulp1) to cleave off small ubiquitin modifier (SUMO) and inteins for autoproteolysis.

Modification of histones by NCL/EPL

Histone proteins are particularly interesting targets for NCL and EPL owing to two main characteristics; They possess a vast number of functionally important post-translational modifications particularly concentrated at the termini of the proteins and they can be refolded in vitro with other core histones allowing the use of denaturants during reactions. Therefore significant effort has gone into generating full-length histones containing specified PTMs[238-240]. Modified histones are generally obtained by synthesizing PTM-containing peptides by SPPS followed by ligation to the remainder of the histone produced recombinantly.

The first example was generation of H3 phosphorylated at serine 10 (H3S10ph)[241]. This modification placed by Aurora B kinase is important in cell cycle regulation to mobilize proteins as the cell enters mitosis. The synthesis was achieved through ligation of a 1-31 residue peptide containing the modification with the remainder of the protein and acceptance of a T32C mutation as desulfurization was not routine at the time. This was then expanded to other simple modifications like acetylation and methylation[56,242]. Combining recombinant intein-mediated thioester formation, ligation

auxiliaries and desulfurization allowed the traceless semisynthesis of K120 ubiquitylated histone H2B (H2B K120ub)[243,244]. This reagent was used to show how H2B ubiquitylation stimulates methylation of H3K4 and H3K79 and thereby transcription through allosteric activation of RAD6 and DOT1L respectively[244-246].

Since then, NCL and EPL combined with radical-initiated desulfurization has been used to introduce almost all types of post-translational modifications into histones and at a range of different sites with functional importance. Recently, the chemical control granted through these methods combined with reversible linkages have allowed the traceless synthesis of asymmetrically modified nucleosomes thereby directing also macromolecular structure beyond single protein sequences[247]. Although the cycle of synthesis, ligations, purifications, refolding and assembly with DNA still are not particularly high-throughput, recently the largest described library of 54 differently modified nucleosomes was used in combination with DNA barcoding and high-throughput sequencing to probe intra-nucleosomal cross-talk in a multiplexed fashion[248]. The ability to prepare suitable homogenously modified histones subsequently requires quantitative methods for the interrogation of their function. Fluorescence microscopy is particularly well suited for this in being able to provide detailed mechanistic insight.

1.3. Fluorescence and single-molecule methods

1.3.1. Confocal and wide-field fluorescence microscopy

Fluorescence occurs when molecules emit photons upon transition of electrons from an excited singlet state to the ground state (**Figure 1.3.1.A**). Molecules capable of fluorescing typically contain aromatic structures and/or conjugated double bonds (E.g. TAMRA and Cy3 in **Figure 1.3.1.B**). The vibrational loss of energy after excitation prior to returning to the ground state causes the emission to have a lower energy and correspondingly longer wavelength than the light used for emission, resulting in a stokes shift (**Figure 1.3.1.C**).

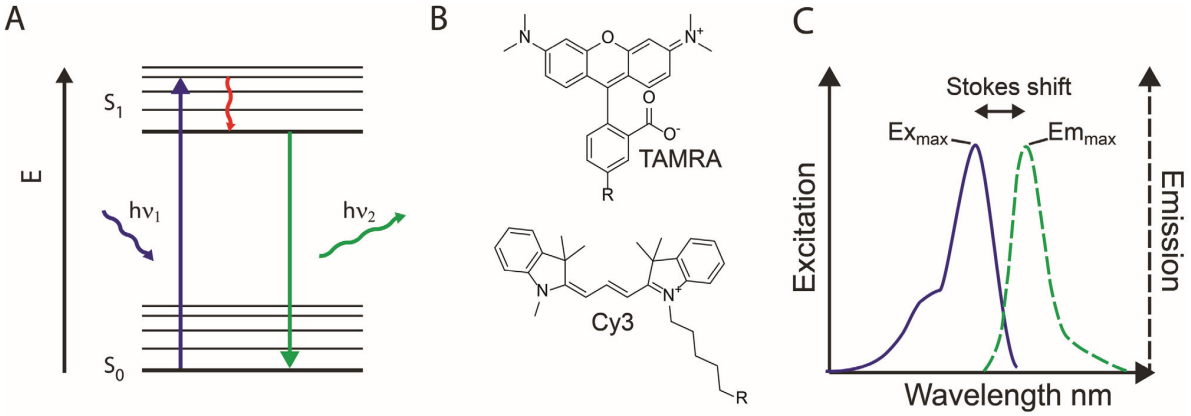

Figure 1.3.1. *Example fluorophore molecules and key characteristics of their excitation and emission.* ***A.*** *Jablonski diagram of excitation by an incoming photon (blue lines), thermal relaxation (red line) and emission of photons upon returning to the ground state (green lines).* ***B.*** *Examples of organic fluorophore molecules displaying the conjugated double-bond systems capable of absorbing and emitting photons.* ***C.*** *Excitation and emission spectra, highlighting the excitation and emission maximum and the stokes shift.*

In addition to the wavelength with maximal excitation and emission, additional important fluorophore characteristics include the life-time, the brightness and the photostability. The brightness is determined by the fluorophores photon absorption capability (the extinction coefficient) combined with the degree to which such absorption results in photon emission (the quantum yield, φ).

Fluorescence microscopy requires both an excitation source and the collection of the light emitted from the sample. Typically, this is carried out by wide-field or confocal microscopy (**Figure 1.3.2**).

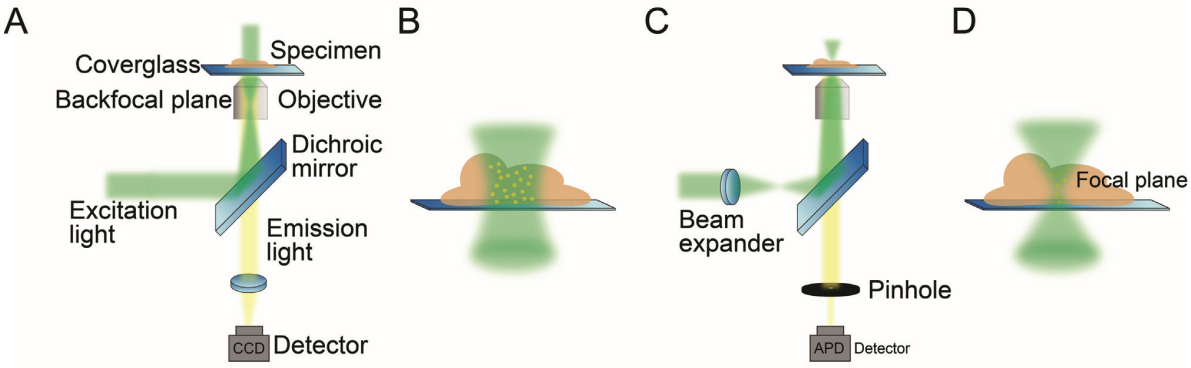

Figure 1.3.2. *Basic schemes of microscope set-ups used for fluorescence microscopy and their focal volumes.* ***A.*** *Outline of light path, excitation and detection in wide-field microscopy, excitation light is focused in a backfocal plane and emission detected from all of the fluorescence in the sample simultaneously with e.g. a CCD camera.* ***B.*** *Scheme of excitation volume in illuminated specimen. Fluorophores are excited in the entire z-volume.* ***C.*** *Schematic of confocal fluorescence excitation and detection. The excitation light is focused in the specimen such as to result in limited out-of-focus excitation. A pinhole blocks out possible out-of-focus fluorescence emission and fluorescence intensity is detected with an avalanche photodiode or similar.* ***D.*** *Volume excited in sample with confocal microscopy resulting in a defined focal plane from which emission is further defined by the pinhole.*

In standard wide-field fluorescence microscopy the entire field of view is illuminated throughout the whole specimen and can concomitantly be visualized at very high detection rates. This is however at a significant cost of both lateral and axial resolution in densely fluorescent samples. The point-spread

function (PSF, see further description below), of the individually fluorescent species is broader than the distance between the molecules and therefore limits the ability to resolve finer details. This resolution ability is further compromised by out-of-focus fluorescence emission, resulting in significant background fluorescence. In confocal microscopy, the excitation focus is formed in the sample and light from background fluorophores emitting out of this focus are further blocked out by a pinhole in front of the detector. Scanning throughout the sample in the x/y and/or z-directions then allows to obtain a resolution that is mainly diffraction-limited, but also limited in temporal resolution by the scanning speed.

1.3.2. Single-molecule fluorescence microscopy

Average molecular behaviour can routinely be obtained from ensemble fluorescence spectroscopy or microscopy. Analysis of single-molecule fluorescence can however provide further key information about molecule stoichiometries, population distributions, binding kinetics and conformational dynamics[249].

Obtaining fluorescence from individual molecules requires that they can be resolved in space or time. The resolution achievable by fluorescence microscopy is constrained by the diffraction limit, d:

$$d = \frac{\lambda}{2 \cdot NA} \qquad (1.1)$$

λ is the wavelength of the light collected and NA is the numerical aperture describing the angles from which a given objective can collect or send emitting light:

$$NA = n \cdot sin\theta \qquad (1.2)$$

Where n is the refractive index of the medium through which the objective is functioning (air = 1.00, water = 1.33, oil = 1.52) and θ is the maximum half-angle between the objective and the point to be imaged. For an objective using oil immersion with an NA of 1.49, this means that it can collect and emit from angles up to 79°. In the case of light with a wavelength of 550nm, the theoretical limit to the lateral resolution with such an objective would be 185nm. The point spread function (PSF) is the gaussian profile resulting from diffraction-limited fluorescence imaging of a single point.

The PSF centre of individually localized fluorophores with sub-diffraction-limit resolution forms the basis for imaging methods photoactivation localization microscopy (PALM) and stochastical optical reconstruction microscopy (STORM) allowing *in vivo* imaging with high resolution[250,251]. In those

types of imaging the density of excited fluorophores at any given time in a wide-field setup is controlled temporally thereby allowing the determination of positions down to 10-50nm lateral resolution.

The most commonly used excitation setups methods for carrying out single-molecule measurements *in vivo* and *in vitro* employ confined excitation either with a confocal setup, using wide-field imaging combined with total internal reflection (TIR) (**Figure 1.3.3**) or by light sheet microscopy (Not explained further here).

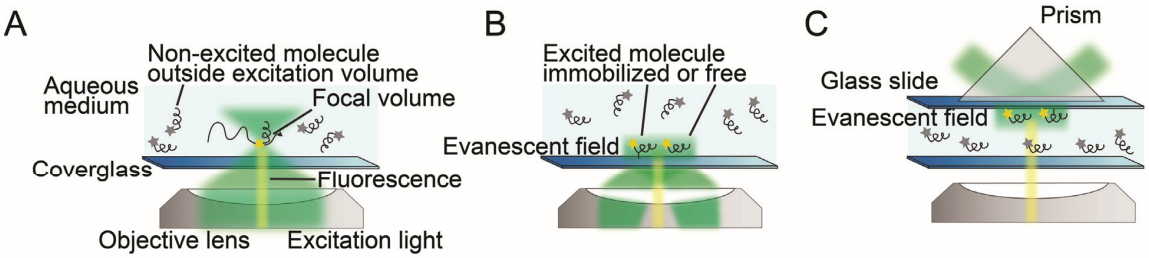

Figure 1.3.3. *Microscopy setups for obtaining single-molecule resolution.* ***A.*** *Confocal microscopy results in a diffraction-limited confocal volume, that combined with the pinhole in front of the detector results in very small detection volumes.* ***B.*** *Objective-type total internal reflection fluorescence (TIRF), where angles above the critical angle for the excitation light is generated through the objective resulting in an evanescent field in the aqueous medium, confining the excitation volume to 100-250nm near the surface.* ***C.*** *Prism-type TIRF, where the angle for TIR is obtained with the use of a prism and fluorescence light is collected from the evanescent wave arising below the glass surface distal to the objective.*

Single-molecule spectroscopy using confocal microscopy takes advantage of the small excitation volume, the pinhole blocking out-of-focus light and fixing the position (no scanning)[252]. This gives an observation volume of just ~1fL. Combined with picomolar concentrations of the labeled species, single molecules diffuse through the confocal volume resulting in fluorescence bursts. Pulsed light sources for excitation and time-correlated single photon counting enables measurement of fluorescence life-times. Further, anisotropy, photon counts and quantum yields can be determined with suitable detectors[253].

Confined excitation for wide-field imaging is done by creating an evanescent wave of light near a surface, thereby exciting only molecules near that surface. This is carried out by total internal reflection; If an incident light beam hits an interface between two media with an angle greater than the critical angle, it will result in total reflection (**Figure 1.3.4**).

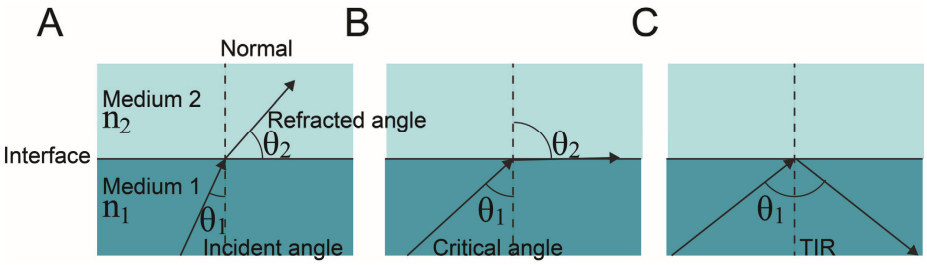

Figure 1.3.4. *Refraction or reflection of light traveling between two media.* ***A.*** *When an incident beam from one medium travels to another medium at an angle lower than the critical angle it will be refracted at the interface in agreement with Snell's law.* ***B.*** *If the incident angle equals the critical angle, the exit angle will be 90°.* ***C.*** *Upon increasing the incident angle beyond the critical angle, the beam is totally internally reflected at the interface.*

The relation between the incident angle in the first medium and the refracted angle in medium 2 is given from Snell's law:

$$n_1 \cdot sin\theta_1 = n_2 \cdot sin\theta_2 \qquad (1.3)$$

n_1 is the refractive index of the first medium, θ_1 the incident beam angle, n_2 is the refractive index of the second medium and θ_2 the refracted beam angle. In the case where no refraction takes place, the refracted beam angle will be orthogonal to the normal of the surface. Therefore θ_2 is 90°, and Snell's law can be rewritten as:

$$\theta_1 = sin^{-1}\left(\frac{n_2}{n_1}\right) \qquad (1.4)$$

This corresponds to the critical angle above which the incident beam will be totally internally reflected. For borosilicate glass (n = 1.47) and water (n = 1.33) the critical angle is ~65°. Upon TIR an evanescent field extending 100-250nm into the medium arises on the side of the second medium opposite to the reflection interface. The intensity of this evanescent wave decays exponentially. The relatively large angle required for the incident beam is generated in microscopy by the use of a prism (prism-type TIRF) or a high numerical aperture oil-immersion objective (objective-type TIRF) (**Figure 1.3.3.B and C**). Through this mode of excitation combined with wide-field imaging, hundreds of fluorescent molecules can be visualized simultaneously at the single-molecule level.

Using these microscopy techniques in suitably designed experiments allows detailed mechanistic investigations into dynamic molecular behavior, that is not easily obtained with other methods.

2. Aim

The aim of the present PhD work was to i) Interrogate recruitment kinetics of HP1 to post-translationally modified chromatin and understand how substrate density, local concentration and multivalency contributes to recruitment. ii) Define the states, equilibrium conformational dynamics and population distributions of chromatin secondary structure and delineate how H4K16ac and HP1 chromatin binding affect these. iii) Devise synthetic routes and prepare histones carrying the trio of histone PTMs recognized by 53BP1.

The bulk of mechanistic studies on chromatin structure and function have been carried out with nucleosomes and/or using ensemble biochemical methods. With these substrates however, the contribution of avidity effects on effector binding to PTMs and the influence of neighbouring nucleosomes on chromatin accessibility are insufficiently accounted for. Additionally, ensemble methods report on molecular averages, making it challenging to draw conclusions about rates, populations distributions and stoichiometries, that can all be obscured by molecular heterogeneities. These shortcomings prompt the development of hybrid approaches combining chromatin engineering with highly sensitive methods capable of detecting molecular behavior at the single-molecule level.

To elucidate the determinants of HP1 effector binding to chromatin, we aimed to establish a method based on co-localization single-molecule microscopy (CoSM) with TIRF using immobilized engineered chromatin and labeled HP1. This allowed us to systematically vary the chromatin substrate and the effector to elucidate the factors influencing effective and dynamic HP1 recruitment.

Study of the conformational dynamics of chromatin secondary structure required methods with high spatiotemporal resolution. Single-molecule FRET (smFRET) allows local nanometer-movements to be detected at time-scales down to microseconds. Distinct sets of FRET positions within chromatin, enabled us to obtain a comprehensive understanding of the subpopulations and their kinetics.

Finally, synthesis of histone H2A.X with PTMs both N – and C-terminal, recognized by 53BP1 required assembly of multiple polypeptide fragments. This hinged on suitable orthogonal protection groups, and the availability of those for long proteogenic fragments. Therefore we sought to use SUMO and the split intein Npu to convergently synthesize H2A.X modified at both ends.

Realization of these aims would shed light on the molecular interplay governing effector recruitment and chromatin conformations, while providing essential materials and tools to further investigate the spatiotemporal dynamics of chromatin regulation.

3. HP1 binding to chromatin by CoSM

3.1. Project introduction and outline

3.1.1. Mechanisms of chromatin effector recruitment and retention

A large fraction of chromatin-associated effector proteins contain linked binding modules allowing multivalent recognition of target chromatin[138]. Despite a resulting increase in affinity, effectors do not stably bind to chromatin for extended time periods, but rather remain highly mobile as revealed by FRAP and FCS in vivo[254,255]. Several mechanisms have been proposed to explain how the balance between efficient targeting and dynamic binding is achieved[256]. First, the presence of multiple target binding sites within three-dimensional nuclear space might allow effectors to remain confined within a region longer by dynamically transferring between binding sites within such regions. The high local density of target sites can thereby function as a kinetic trap, even though the protein is not stably anchored to any one site (**Figure 3.1.1.A**).

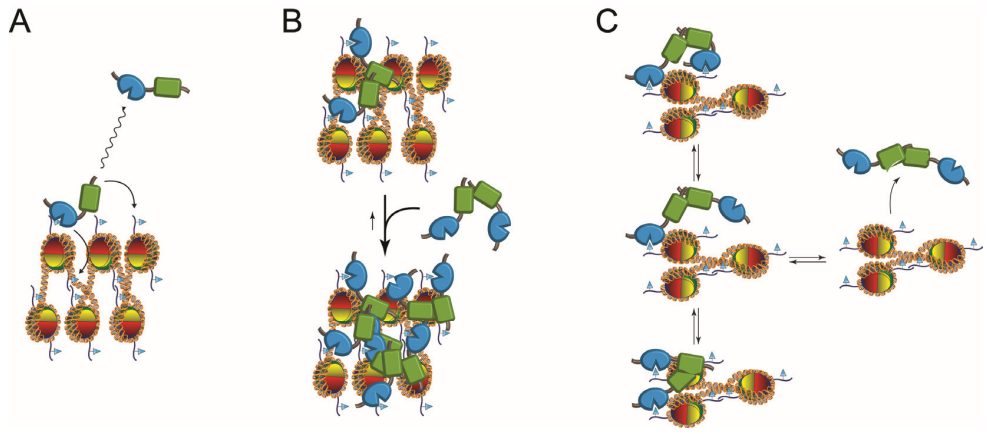

Figure 3.1.1. *Mechanisms governing effector recruitment and retention to chromatin.* ***A.*** *Local substrate density could allow rebinding on the other nearby sites in the same nucleosome or on nearby nucleosomes prior to dissociation.* ***B.*** *On-target oligomerization might stabilize binding.* ***C.*** *Effector multivalency may allow retention in target chromatin despite dissociation of one binding domain.*

Second, work has suggested that effectors can oligomerize on the chromatin template. [135,257,258]. This could lead to the formation of relatively stable oligomeric networks and help chromatin states spread to neighbouring regions (**Figure 3.1.1.B**). Finally, the presence of multiple binding domains in a homo-, or heterooligomeric complexes could potentially influence the binding by increasing the chance of binding to one of the available PTMs in the nucleosome. Dissociation would further require both binding domains to simultaneously disengage (**Figure 3.1.1.C**).

HP1 is a small chromatin-associated effector containing a recognition domain and the ability to dimerize forming a multivalent binder[259]. Further, evidence points towards HP1 capable of intra-, and

inter-nucleosome binding[135]. Therefore HP1 serves well as a model protein for studies on multivalent recognition of chromatin to understand the underlying mechanisms.

3.1.2. HP1 domain organization and function

Heterochromatin, as mentioned previously, is regions of the genome where underlying genes are silenced by both pre – and post-transcriptional mechanisms, including restricted DNA accessibility, gene repression and accelerated mRNA decay. HP1 was identified as a non-histone protein in *Drosophila Melanogaster* specifically associated with heterochromatic regions[260] (**Figure 3.1.2.A**).

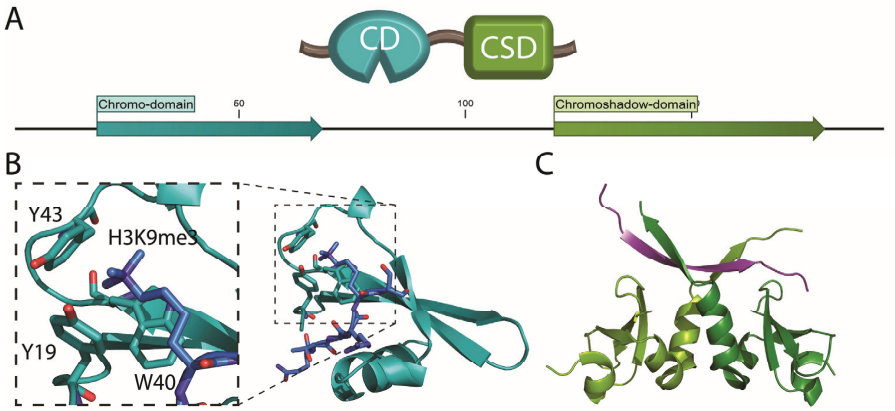

Figure 3.1.2. *Domain architecture and structure of HP1.* ***A.*** *Domain outline of HP1 on the protein sequence.* ***B.*** *Crystal structure of the HP1 chromodomain bound to a trimethylated H3 N-terminal peptide showing the aromatic cage on the zoom-in.* ***C.*** *Crystal structure of chromoshadow-domain dimer with associated PXVXL motif from protein binding partner (purple)*

The N-terminal extension (NTE, Residues 1-20) of HP1 is an acidic unstructured tail involved in DNA binding. A structured chromatin organization modifier (chromo) domain (CD, Residues 21-71), is responsible for recognition of the target modified histone H3 tail (**Figure 3.1.2.B**). This domain is highly similar to several other chromo-domains involved in recognition of methylated N-terminal tails of histones[124]. Towards the C-terminal is a slightly larger chromo shadow domain (CSD, residues 120-179) unique to HP1 (**Figure 3.1.2.C**). Interconnecting the two structured domains is a hinge region involved in non-specific nucleic acid binding.

HP1 is evolutionarily conserved, with variants existing in e.g. *S.pombe* (Swi6 and Chp2), *D. Melanogaster* (HP1α, b, c, d and e) and humans where three isoforms have been identified; HP1α, HP1β and HP1γ[261]. Although conserved in domain architecture, the human isoforms are associated with distinct functions and distributions. HP1α and HP1β both localize primarily to heterochromatin, whereas HP1γ in is found in both heterochromatic and in euchromatic regions at active genes near

RNA pol II[262,263]. HP1α and HP1β is found particularly in constitutively heterochromatic regions that are permanently silenced, including centromeres and telomeres. This is in contrast to facultative heterochromatin controlled by polycomb repressive complexes, often linked to developmental genes that are silenced or activated through differentiation. Heterochromatic states can spread to spatially proximal regions in the genome, thus leading to inactivation of neighbouring genes. A manifestation of this was observed early in *D. Melanogaster*, where repositioning of a gene for white eye color close to heterochromatin in the X chromosome would result in an eye phenotype with variegation, where the eye is mottled with white and red regions[264]. Later, these phenotypic changes were found to be linked to heterochromatin spreading through HP1 functioning in concert with the the methyltransferase Suv39h1(SU(VAR)3-9 in *D. Melanogaster*)[265-268]. This is an example of the feedback loop between a reader and the writer existing in the same complex[265]. These phenotypic changes can further be inherited and persist through mitosis, but also transgenerationally[86,87].

Like histones, HP1 is subject to PTMs. These include serine phosphorylation throughout the different isoforms and at the NTE of HP1α as well as acetylation and SUMOylation[269,270]. Additionally, HP1 functions as an essential scaffold bridging H3 K9me3 with a number of silencing factors, including nucleosome remodelling and deactylation (NuRD) and KRAB-associated protein 1 (KAP1) [271].

The chromo domain

Selective recognition by the HP1 family of proteins of H3K9me2/3 occurs through the N-terminal structured chromo-domain(**Figure 3.1.2.B**)[272,273]. The CD in it's unbound apo-form forms a three-stranded anti-parallel β-sheet neighbouring a short α-helix [274]. The H3 N-terminal tail is bound to the chromo-domain through an induced fit mechanism. Upon binding, the chromo domain distinguishes between the unmodified and methylated lysine side chain through an aromatic cage consisting of highly conserved tyrosine, phenylalanine and tryptophan residues[275,276]. The affinity of the interaction between the chromo-domain and modified N-terminal tail peptides lie in the range of 1-20μM depending on the homologue[275-277]. Additional conserved residues surrounding the aromatic cage ensure that the sequence context of the modified lysine corresponds to that of H3K9me2/3 and not that of other lysines in the N-terminal tails of H3 and H4. Particularly, phosphorylation of H3S10 by aurora kinase functions as a switch to result in dissociation of HP1 upon entry into M-phase[277].

The hinge region and the chromo-shadow domain

The protein fold of the CSD is highly similar to that of the CD with three antiparallel β-sheets and an adjacent α-helix. Additionally, a longer α-helix forms the interface for dimerization of HP1 (**Figure 3.1.2.C**)[278]. The C-terminal tail extending from these α-helices generate a β-sheet shaped platform for protein-protein interaction with a range of partner proteins containing a conserved PXVXL/I motif with an affinity of 0.1-2μM[279-281]. This forms the basis for binding to e.g. the Suv39h1 that as mentioned above is the methyltransferase writing the H3 K9me3 PTM. The dimerization affinity for the chromo-shadow itself domain is around 1-5μM and the higher affinity of the interaction partners induce dimerization[282].

The hinge region interconnects the CD and CSD. Although unstructured and less well conserved, this also harbors functionality for HP1[259]. The basic residues present in the hinge region help bind RNA and DNA non-specifically, and such binding is important for correct heterochromatin localization[283]. Further these residues help capture nascent RNA transcribed from heterochromatic regions for HP1 to shuttle them for degradation and ensure silencing of the underlying genes[284].

Localization and cellular dynamics of HP1

Heterochromatin proteins form stable macro-domains within the cell nucleus that co-localize with densely stained DNA corresponding to centromeric and telomeric heterochromatin. Despite the macroscopically stable domains, microscopic turnover of HP1 as revealed with FRAP and FCS within these regions occur on time-scales of only seconds to minutes, with a significant (~20%) immobile fraction[254,255,285]. During mitosis HP1 proteins are mobilized from chromatin due to the phosphorylation of H3S10 as described above[277]. HP1 localization is further controlled by demethylation of the target H3K9me2/3 by specific demethylases[286-288].

How effective HP1 recruitment is achieved, while maintaining a dynamic protein turnover is not well understood. Distinguishing between the possible mechanisms governing HP1 recruitment to target chromatin requires the establishment of sensitive methods to probe these kinetics. Previously, co-localization single-molecule microscopy (CoSM) with TIRF has been used to probe millisecond interaction dynamics of biomolecules[289,290]. We thus decided to pursue a similar strategy to investigate how HP1 binds to chromatin and delineate the underlying mechanisms.

3.1.3. CoSM

In CoSM a binding substrate e.g. a protein, DNA molecule or a complex thereof is labeled with one dye and immobilized on a glass surface through biotin-neutravidin interactions allowing prolonged single-molecule localization by TIRF (**Figure 3.1.3.A**).

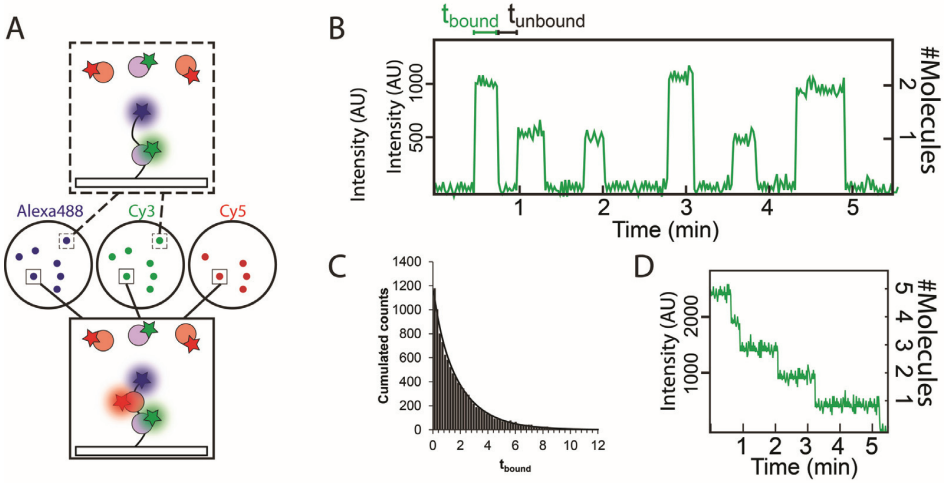

Figure 3.1.3. *CoSM assay and example traces.* ***A.*** *Schematic of CoSM with an immobilized Alexa 488 labeled molecule (Blue) and binding of one (Top) or two (Bottom) factors resulting in co-localizing fluorescence. Based on Larson and Hoskins 2014* ***B.*** *Example trace illustrating the use of CoSM to determine binding, unbinding and binding stoichiometry of a molecule.* ***C.*** *Example histogram of cumulated binding times fitted with a double-exponential function.* ***D.*** *Example trace showing CoSM for determination of stoichiometry by step-wise photobleaching.*

The immobilized molecule is labeled with a fluorophore (e.g. Alexa 488 as in the figure) distinct from that of additional fluorescent molecules added in solution (e.g. Cy3 and Cy5). Molecules staying away from the surface remain non-fluorescent as they are not within the confined excitation volume defined by the evanescent wave.

A critical requirement for the observation of single labeled biomolecules is that they do not stick non-specifically to the surface. Bovine serum albumin (BSA) that by itself sticks non-specifically to glass surfaces, can be effectively used to prevent adsorption of oligonucleotides dsDNA or RNA[291]. In contrast, proteins or nucleoprotein complexes require glass surfaces to be covered with polyethylene glycol (PEG)[292] to prevent non-specific sticking. Incubation of amino-silanized glass with supersaturated clouding point solutions of PEG-NHS esters results in a passivating layer of PEG[293]. Typically, methylene ether capped PEG with a mass of 5000Da is used, sometimes complemented with PEG of smaller size (~500Da) to fill out holes between the large PEG molecules[294].

Upon binding to the surface-immobilized substrate, molecules are excited and visible as diffraction-limited single-molecules with fluorescence increases co-localizing with the labeled substrate (**Figure**

3.1.3.A and B). CoSM is particularly useful in directly determining binding and unbinding kinetics in an equilibrium setting without resorting to rapid-mixing techniques like stopped flow fluorescence spectroscopy. The binding times and times between binding events can be plotted in histograms of the number of cumulated counts that a given binding time is observed (**Figure 3.1.3.C**). Subsequently this can be fitted to single – or double-exponential decay functions (Equation 3.1).

$$\text{Cum counts} = A_1 \cdot e^{-\tau_{bound,1} \cdot t} + A_2 \cdot e^{-\tau_{bound,2} \cdot t} \quad (3.1)$$

In this, A_1 and A_2 are the amplitudes of each phase and $\tau_{bound,1}$ and $\tau_{bound,2}$ are the decay times of each phase. From the decay times, the rates of binding (taking into account the concentration) and unbinding can be determined and subsequently the affinity (Equations 3.2, 3.3, and 3.4).

$$k_{off} = \frac{1}{\tau_{bound}} \quad (3.2)$$

$$k_{on} = \frac{1}{\tau_{unbound} \cdot [Molecule\ in\ solution]} \quad (3.3)$$

$$K_D = \frac{k_{off}}{k_{on}} \quad (3.4)$$

If three or four fluorescently labeled species are present, CoSM can further be employed to determine relative arrival and departure times in complex reaction cascades (**Figure 3.1.3.A**)[295,296].

CoSM is also useful in directly determining stoichiometries in molecular complexes. Either the intensity of each binding event can be correlated with the number of fluorophores resulting in said intensity, or each photobleaching event can be correlated with the presence of one molecule (if fully labeled; One fluorophore per molecule, **Figure 3.1.3.B and D**)[297,298]. CoSM has been increasingly employed to answer biological questions, that are not easily addressed by classic ensemble methods. Interesting recent examples of this include the specificity and ubiquitylation activity of the anaphase promoting complex for cyclins and subsequent proteasome-mediated degradation, spliceosome assembly and processing of mRNA to remove introns, ordered formation of transcription initiation complexes and microRNA-guided targeting of argonaute to mRNA[290,295-300].

3.1.4. Experimental scheme

We conceived a method based on CoSM for detailed investigation of the mechanisms governing HP1 recruitment to chromatin. Designer chromatin containing the target H3 K9me3 modification and a biotin anchor allowed chromatin arrays to be specifically immobilized in a flow chamber onto a PEG

coated surface to prevent non-specific interactions and a fraction of biotin-PEG for specific anchoring. Localization of chromatin arrays at the single-molecule level is made possible by the presence of a fluorophore in the DNA for excitation with TIRF microscopy (**Figure 3.1.4.A**)

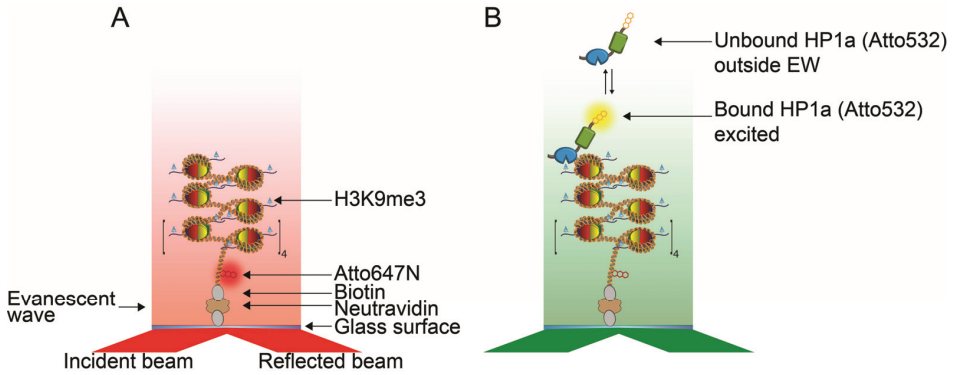

Figure 3.1.4. *Experimental scheme for CoSM using chromatin and the effector HP1α.* ***A.*** *Immobilization of chromatin onto the surface assembled with DNA containing a biotin and an Atto647N fluorophore for localization with smTIRF.* ***B.*** *HP1α labeled with Atto532 enters the excitation volume if in contact with chromatin resulting in fluorescence, whereas unbound will remain non-fluorescing.*

Subsequent introduction of HP1α labeled with another fluorophore permitted binding events to be visualized at the single-molecule level by co-localization with the immobilized chromatin arrays (**Figure 3.1.4.B**). On and off times corresponding to the kinetics of binding and unbinding should then be directly obtainable from analysis of the fluorescence spikes upon HP1α binding. The high degree of chemical control over all the different components allows easy variation of the substrate density in the target chromatin as well as the concentration and oligomeric state of the effector. This equips us to systematically investigate how these different aspects of the chromatin – effector interaction regulates HP1α recruitment to target chromatin. Establishing the system and using it to gain a mechanistic understanding of HP1α binding dynamics required us to:

- Prepare histones and octamers with and without H3 K9me3
- Generate wild-type unlabeled and labeled HP1α and covalently dimerized HP1α
- Prepare labeled biotinylated array DNA and use this with octamers for chromatin assembly
- Establish CoSM assay for chromatin array immobilization and HP1α binding
- Investigate influence of substrate density and total HP1α concentrating on binding
- Examine how multivalency contributes to HP1α association and retention
- Build a model to recapitulate the experimental findings

Where indicated in figure legends, Beat Fierz, Louise Bryan or Andreas Bachmann contributed to this project.

3.2. Preparation of histone octamers with and without H3 K9me3

Semisynthesis of H3 K9me3

Semisynthesis of H3 K9me3 was planned to be done with a synthetic N-terminal peptide hydrazide containing the trimethyllysine to be ligated with the remainder of the protein generated recombinantly. Desulfurization would yield the native protein carrying the desired modification (**Figure 3.2.1**).

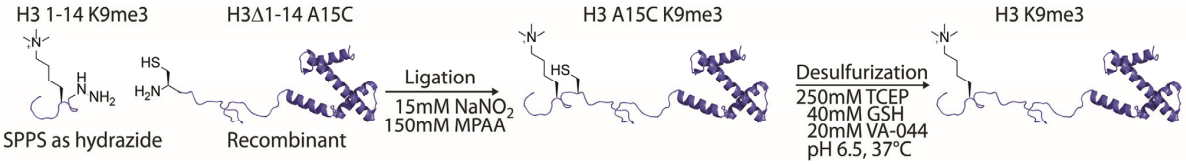

Figure 3.2.1. *Scheme for the planned traceless synthesis of H3 K9me3. The N-terminal tail with K9me3 is synthesized as a peptide hydrazide by SPPS ligated to the core of the protein followed by desulfurization to convert the cysteine to the native alanine.*

The H3 N-terminal 1-14 residue peptide containing the modified trimethylated peptide hydrazide was synthesized by Fmoc SPPS. Conversion of the chlorotrityl resin to the trityl hydrazine resin as well as coupling of the first amino acid (lysine) was done manually. Subsequent couplings including the trimethylated lysine were done by automated Fmoc SPPS. After cleavage from the resin, the crude peptide hydrazide was purified by RP-HPLC to yield the modified histone N-terminal (**Figure 3.2.2**).

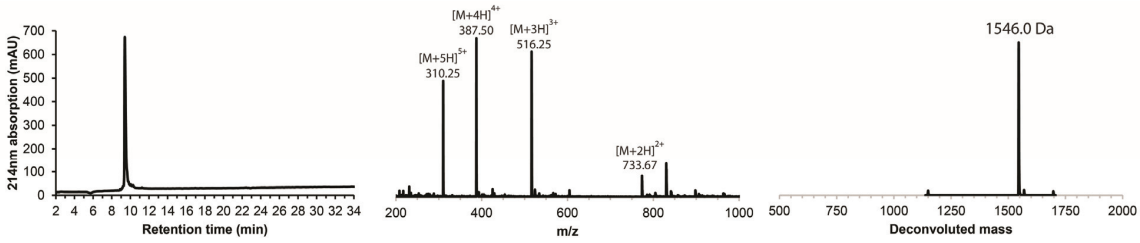

Figure 3.2.2. *Final analysis of H3 K9me3 1-14 peptide by RP-HPLC and ESI-MS. M_{obs} = 1546.0Da. M_{calc} = 1546.8Da.*

The rest of the protein was generated recombinantly by expression of a fusion SUMO-H3Δ1-14 A15C, C110A protein followed by cleavage with Ubiquitin like protease 1 (Ulp1) and RP-HPLC purification to yield the truncated protein (**Figure 3.2.3**).

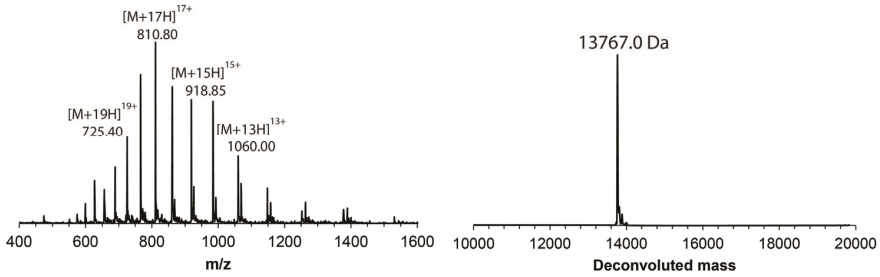

Figure 3.2.3. *ESI-MS analysis of purified truncated H3Δ1-14 A15C protein. M_{obs} = 13767.0Da. M_{calc} = 13769.1Da.*

The peptide hydrazide was oxidized with $NaNO_2$ at pH 3 and converted *in situ* to the MPAA thioester followed by mixture with the remainder of the protein and raising of the pH. The ligation proceeded then to near-completion in about 2h (**Figure 3.2.4**)

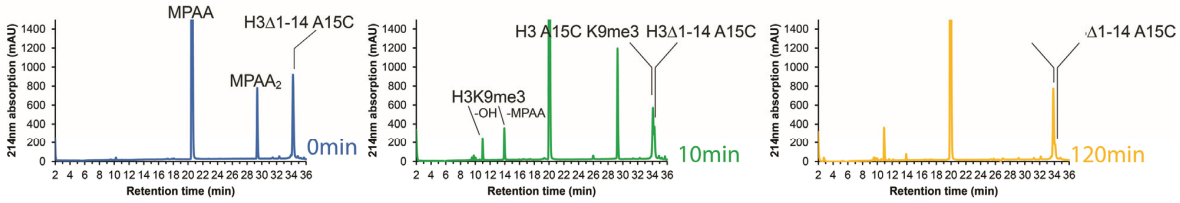

Figure 3.2.4. *Monitoring of the ligation reaction by RP-HPLC after 5min, 30min and 2h respectively.*

The identity of the ligation product was confirmed by ESI-MS followed by purification, final analysis and lyophilization (**Figure 3.2.5**)

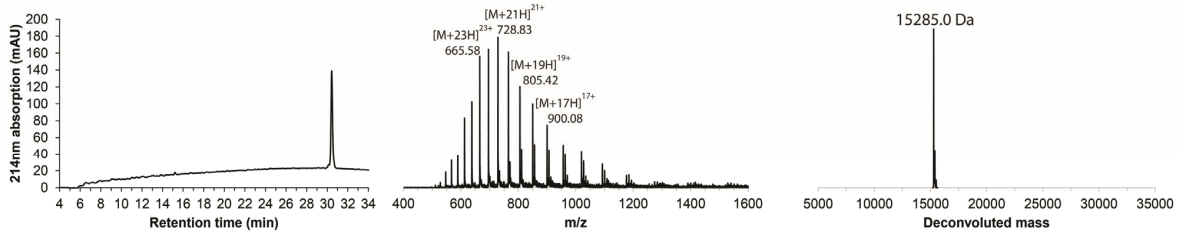

Figure 3.2.5. *Final analysis of H3 K9me3 A15C, C110A after ligation by RP-HPLC and ESI-MS. M_{obs} = 15285.0Da. M_{calc} = 15283.9Da.*

After complete lyophilization it was subject to radical-based desulfurization with TCEP, VA-044, and GSH[195]. This reached completion after 7h followed by another purification to yield the purified H3 K9me3 with the cysteine at the ligation junction converted to the native alanine (**Figure 3.2.6**).

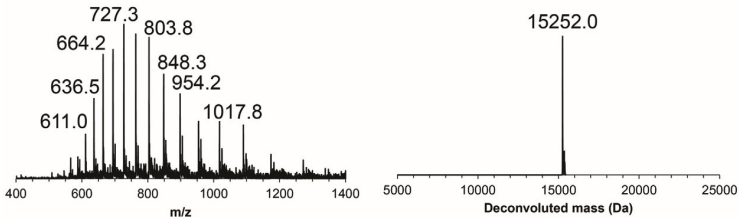

Figure 3.2.6. *Final analysis of H3 K9me3 desulfurization. M_{obs} = 15252.0Da. M_{calc} = 15251.8Da.*

Histone octamer preparation

The other three core human histone proteins H2A, H2B and H4 as well as H3 C110A without the modification were expressed and purified (**Figure 3.2.7.A**). These were refolded by mixture of unfolding buffer followed by dialysis into high salt buffer and gel filtration purification (**Figure**

3.2.7.B). Presence of all 4 histones in the same peak from gel filtration confirmed that the histones had been refolded into octamers (**Figure 3.2.7.C**).

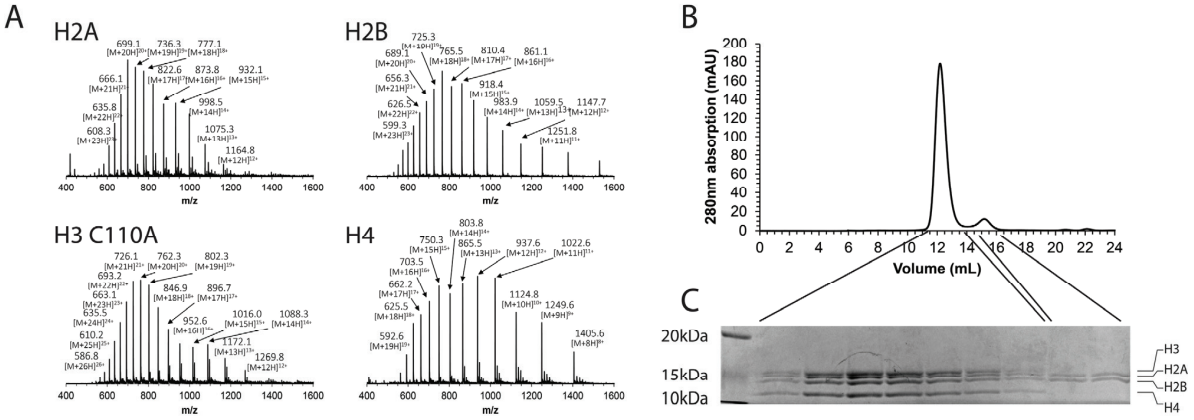

Figure 3.2.7. *Analysis of individual purified histones and refolded histone octamer.* ***A.*** *ESI-MS analysis of histone H2A (M_{obs} = 13965.0, M_{calc} = 13964.3) H2B (M_{obs} = 13761.0, M_{calc} = 13758.91), H3 C110A (M_{obs} = 15227.0, M_{calc} = 15224.7) and H4 (M_{obs} = 11239.0, M_{calc} = 11236.11).* ***B.*** *Chromatogram from gel filtration of refolded histone octamers.* ***C.*** *SDS-PAGE analysis of fractions from gel filtration.*

The octamers either with or without H3 K9me3 were stored for later usage in chromatin reconstitutions.

3.3. 12-mer labeled array DNA and chromatin assembly

To prepare labeled and biotinylated array DNA, a 12x601 NPS encoded in a plasmid was modified at one end to include a BsaI restriction site. The array DNA was isolated from large-scale preparation by EcorV digestion and PEG purifications. Subsequently, the DNA was dephosphorylated to prevent self-ligation and the BsaI site was digested to result in a non-palindromic overhang. To this, a dsDNA piece with one strand containing an internal Atto647, a 5' phosphate and a 3' biotin was ligated (**Figure 3.3.1.A**). Analysis confirmed that the product had been formed (**Figure 3.3.1.B**)

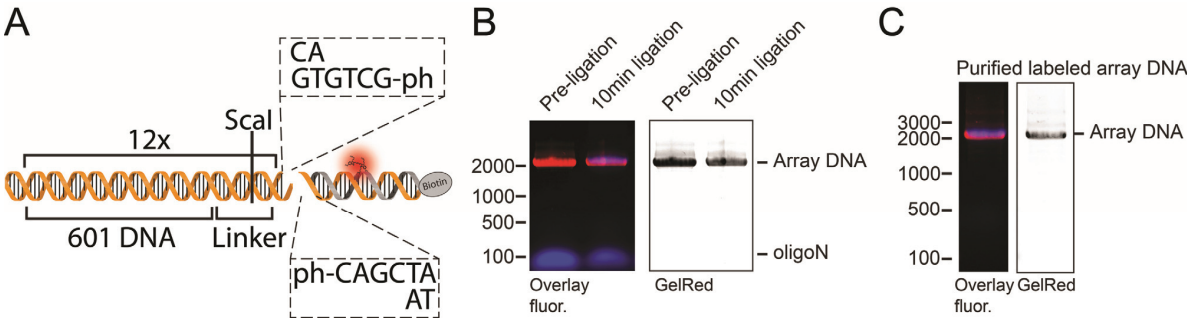

Figure 3.3.1. *Outline of DNA labeling and analysis.* ***A.*** *Scheme of 12x601 array DNA labeling with with biotin modified labeled and phosphorylated dsDNA to non-palindromic BsaI-produced overhang.* ***B.*** *Analysis of ligation of labeled dsDNA to array.* ***C.*** *Analysis of isolated labeled array DNA. Data contributed by Louise Bryan.*

Subsequently the dsDNA piece was removed from the labeled array DNA by PEG precipitation and PCR purification, resulting in the desired labeled and biotinylated array DNA (**Figure 3.3.1.C**).

With the DNA and octamers ready, we proceeded to chromatin assembly for subsequent CoSM. Chromatin samples were assembled by gradual dialysis from high to low salt with excess octamers and in the presence of buffer MMTV DNA with a lower affinity for nucleosome formation. This ensured full or near-full saturation of the 601 array DNA (**Figure 3.3.2.A**)

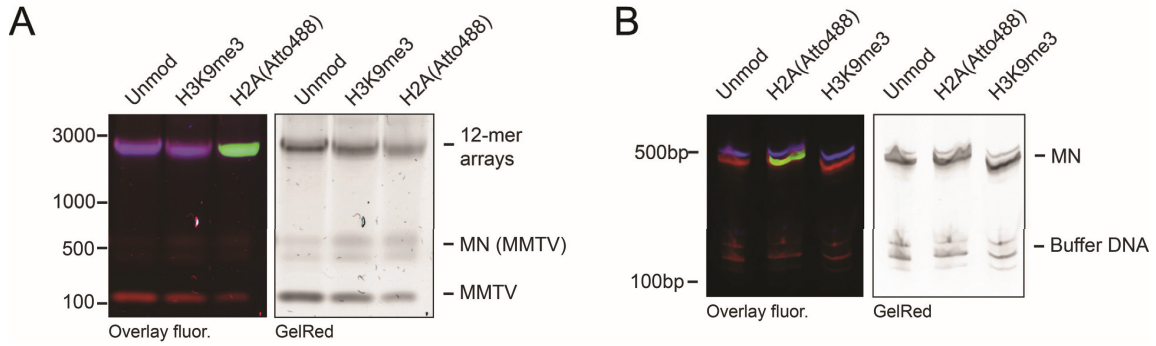

Figure 3.3.2. *Analysis of assembled chromatin.* ***A.*** *Agarose gel analysis of assembled chromatin on labeled DNA with wild-type octamers (unmod), H3 K9me3 octamers, or octamers containing labeled histone H2A (Atto488). MMTV buffer DNA was added to capture excess histone octamers upon full saturation of the array DNA. This results in mononucleosomes (MN) formed on MMTV DNA, while some remains non-nucleosomal.* ***B.*** *ScaI digestion of chromatin ot ensure saturation. Restriction sites are equally distributed in linker DNA between 601 sequences resulting in mononucleosomes upon digestion and little to no free DNA.*

Saturation was further corroborated by digestion with ScaI, for which restriction sites were located in the linker DNA between each nucleosome. This resulted in a band by native PAGE corresponding to single nucleosomes, and not free DNA (**Figure 3.3.2.B**).

3.4. Preparation of labeled and unlabeled HP1

Labeling strategy

With the modified chromatin prepared, the next aim was to prepare labeled HP1α. HP1α is a relatively small protein, which should be taken into account in considering the labeling strategy. Fluorescent proteins and self-labeling tags are relatively large considering the size of HP1α and might adversely affect both diffusivity and ability to bind to target sites in sensitive CoSM applications[301,302]. Although small peptide tags could also be used for labeling, we deemed that the C-terminal of HP1α could be labeled by EPL with a short fluorescently labeled peptide using a recently developed strategy based on split inteins(**Figure 3.4.1**)[235].

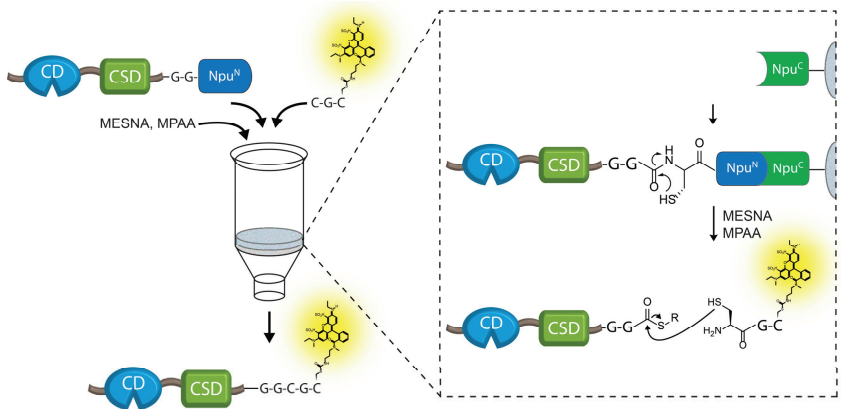

Figure 3.4.1. *Scheme outlining the usage of a split intein column for labeling of HP1α. The protein of interest is fused at the C-terminal to one half of a split intein and the other half immobilized on a solid support. Binding between the two fragments allows the column to work as an affinity column and addition of MESNA results in thioester formation. Further addition of MPAA and a labeled peptide with an N-terminal cysteine allows in situ ligation to take place resulting in elution of the labeled protein.*

As mentioned in the introduction, split inteins are inteins that can carry out transsplicing by bringining together two protein halves. Immobilization of one inactivated half of the C-terminal half of the split intein from Nostoc Punctiforme (Npu^C), allows this to be used as an affinity column for capture of proteins fused to the other half (Npu^N). Simultaneously, this allows thiolysis for the formation of a thioester and ligation with a peptide carrying a thiol and a primary amine. Relative to using linear inteins or fused split inteins, this approach reduces the risk of pre-mature intein hydrolysis and allows ligation with less need for denaturants. Other than the protein of interest the key components needed for this ligation approach to be possible were the labeled tripeptide with an N-terminal cysteine and the inactivated Npu^C peptide for forming the column.

Generation of NpuC-AA-CysOMe peptide for column conjugation

The NpuC-AA peptide, mutated at the C-terminal splice junction to disallow trans-splicing, was generated recombinantly by expression as a fusion with an intervening linker at the C-terminal to the linear intein Mxe GyrA followed by a 6xH-tag for purification. Incubation of the purified protein with cysteine methyl ester resulted in both thiolysis and rearrangement by an S to N acyl shift to form the peptide coupled to CysOMe (**Figure 3.4.2.A**).

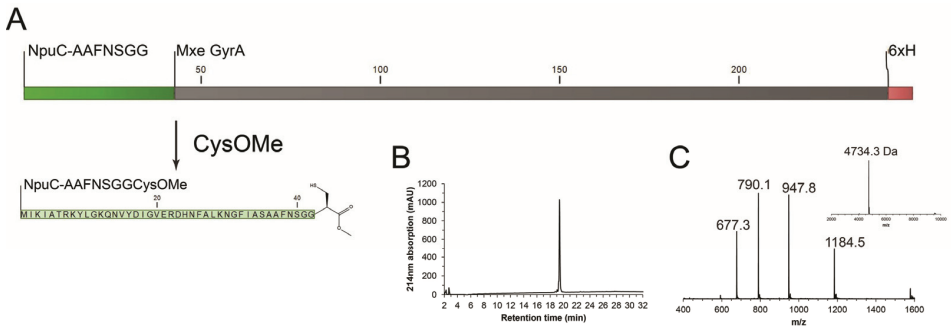

Figure 3.4.2. *Reaction scheme and final analysis of NpuC-AA-CysOMe peptide for the intein column.* ***A****. The split intein was expressed in fusion with a linear intein for thiolysis and acyl shift with cysteine methyl ester.* ***B.*** *Chromatogram from final analysis of NpuC peptide.* ***C.*** *ESI-MS of final purified NpuC-AA-CysOMe peptide.* M_{obs} *= 4734.3,* M_{calc} *= 4733.4.*

Subsequent preparative RP-HPLC yielded the purified peptide that can be coupled to a sulfo-link resin through the cysteine methyl ester (**Figure 3.4.2.B and C**).

Synthesis and labeling of tripeptide

The tripeptide was synthesized by Fmoc SPPS with a sequence of Thz-G-C. The free thiol served as the site for conjugation to Atto532 iodoacetamide. Subsequent methoxyamine treatment at pH 5 opened the thiazolidine ring to free up the free N-terminal cysteine for ligation (**Figure 3.4.3**).

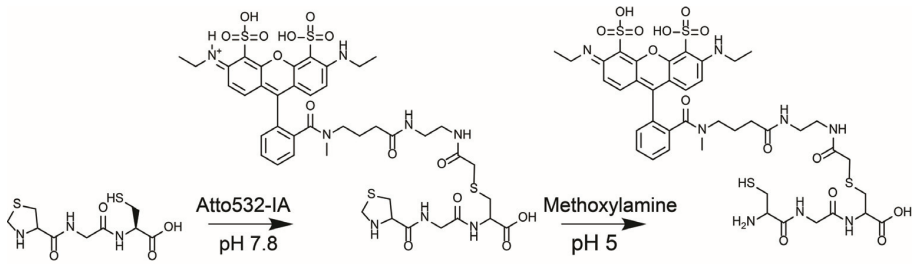

Figure 3.4.3. *Scheme for the modification of tripeptide Thz-G-C to label with Atto532 through an iodoacetamide followed by thiazolidine deprotection with methoxyamine.*

Preparative RP-HPLC yielded the pure peptide for final analysis (**Figure 3.4.4**).

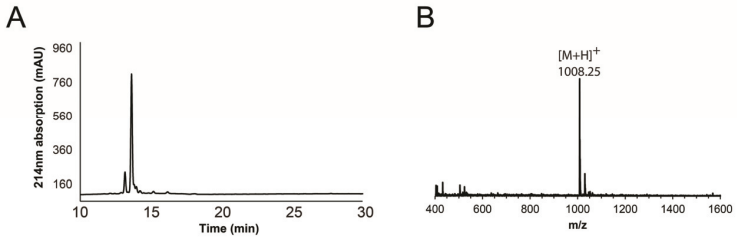

Figure 3.4.4. *Analysis of the purified tripeptide for labeling.* ***A****. Chromatogram from RP-HPLC analysis of the purified peptide.* ***B****. ESI-MS analysis of the peptide* M_{obs} *= 1008.3,* M_{calc} *= 1009.3.*

The peptide was redissolved in water and stored as single-use aliquots for labeling.

Thioester formation and *in situ* EPL to label HP1α

HP1α without any linker at the C-terminal, did not seem to convert to the thioester very efficienctly leading to hydrolysis. This is likely caused by differences in intein efficiency dependent on the N-1 and -2 residue [303]. Insertion of two glycines between HP1α and Npu^N ameliorated this problem. The Npu^C peptide was immobilized on the Sulfo-link resin and HP1α-GG-NpuN was incubated with the resin, leading to complete capture (**Figure 3.4.5.A**). Following washing, the resin slurry with bound HP1α was incubated with MESNA, MPAA and the tripeptide for in situ ligation. After reaction for 10-12h, the labeled protein could be eluted from the column, with a size corresponding to the loss of Npu^N (**Figure 3.4.5.A**).

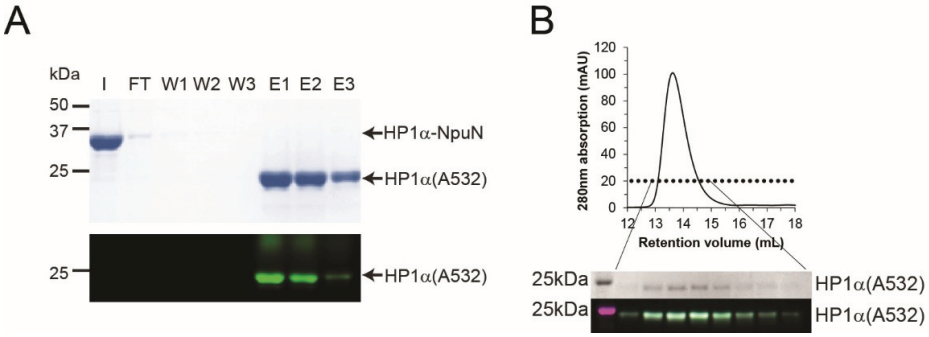

Figure 3.4.5. *Analysis of labeling and purification of HP1α.* ***A.*** *Input, washes and flowthrough from the split intein column-mediated thioester-formation and ligation to yield labeled HP1α.* ***B.*** *Gel filtration purification and fraction analysis of collected labeled HP1α.*

Purification by gel filtration yielded fractions containing the desired labeled HP1α (**Figure 3.4.5.B**)

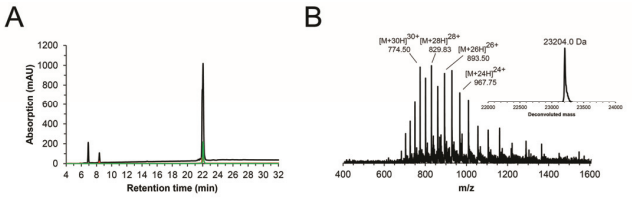

Figure 3.4.6. *Final analysis of labeled HP1α by RP-HPLC and MS. M_{obs} = 23204.0, M_{calc} = 23200.0.*

Final analysis by RP-HPLC and MS confirmed the purity and identity of the product (**Figure 3.4.6**).

HP1α wildtype

HP1α wild-type for experiments to increase the total effector concentration was expressed as a construct fused at the N-terminal to a 6xH tag followed by a thrombin cleavage site. The protein was purified from the soluble fraction by Ni-NTA purification and DNA associated with the protein removed by anion exchange. The his-tag was removed by thrombin cleavage as monitored by HPLC and MS. Finally the protein was purified by size exclusion and the final purified protein analyzed by RP-HPLC and ESI-MS (**Figure 3.4.7**).

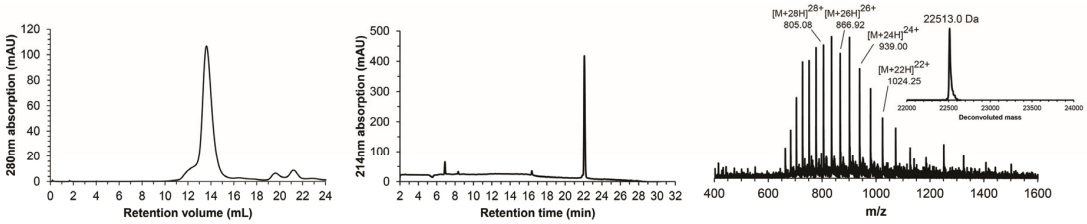

Figure 3.4.7. *Final purification by gel filtration and analysis of HP1α by RP-HPLC and ESI-MS. M_{obs} = 22513.0, M_{calc} = 22506.2).*

Based on this, the protein was pure and like the histone octamers was stored as glycerol stocks in the freezer for later usage.

3.5. Affinity of modified HP1α proteins to H3 K9me3 peptide

To ensure that the HP1α proteins labeled and purified were functional in binding to the target H3 K9me3, affinity measurements were carried out by tryptophan fluorescence and microscale thermophoresis respectively. Binding between the wild-type unlabeled HP1α and the synthesized H3 K9me3 was determined by changes in tryptophan fluorescence. The tryptophan in the aromatic cage binding to the trimethyllysine was expected to change conformation upon binding and thereby changing fluorescence. Indeed, upon titration with the peptide, the fluorescence of the tryptophan

increased. Correction for dilution and subsequent normalization showed that the affinity was 7.6μM in close agreement with published values [304] (**Figure 3.5.1.A**)

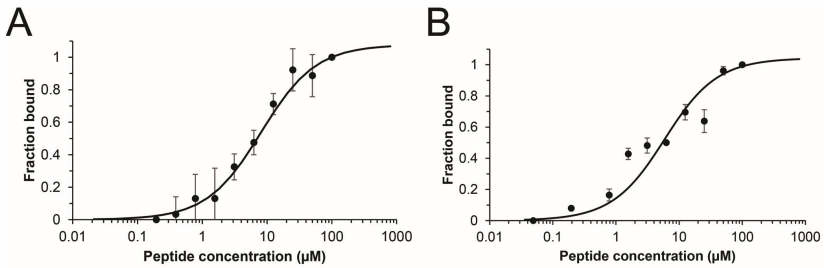

Figure 3.5.1. *Titration curves of unlabeled and labeled HP1α by change in tryptophan fluorescence and MST.* ***A.*** *Wild-type change in tryptophan fluorescence of HP1α upon binding to H3 K9me3 1-14 peptide resulting in an affinity of 7.6±1.4μM .* ***B.*** *Curve from MST with HP1α(Atto532). Affinity determined is 5.4±1.3μM.*

Affinity between the labeled HP1α molecule and H3 K9me3, was determined by microscale thermophoresis (MST). This is a sensitive technique for biomolecular affinity determinations of labeled molecules, requiring low sample amounts and depending on relative mobilities in a heat gradient[305]. The analysis was done with increasing concentration of H3 K9me3 followed by loading into capillaries. Normalization for the start and end point, yielded the binding curves (**Figure 3.5.1.B**). Based on the fitting of the measured points, the affinities were determined based on the curve midpoints. These showed to be 5.4μM. Therefore the affinity between the H3 K9me3 tail display affinities regardless of the labeling that lie within the range of the values that have been previously measured in the literature[277,304].

3.6. CoSM with immobilized chromatin and labeled HP1α

Considerations about fluorophore photophysics

With both the Atto647N labeled chromatin and HP1α labeled with Atto532, we required for our single-molecule experiments that we could observe stable long-lasting fluorescence. Fluorophores can dynamically transition between the ground state and excited states to result in fluorescence, however transitions to triplet states, anionic and cationic states can result in transient fluorophore blinking or promote irreversible bleaching, resulting from distinct transitions in molecular orbital occupancy (**Figure 3.6.1.A**)[306].

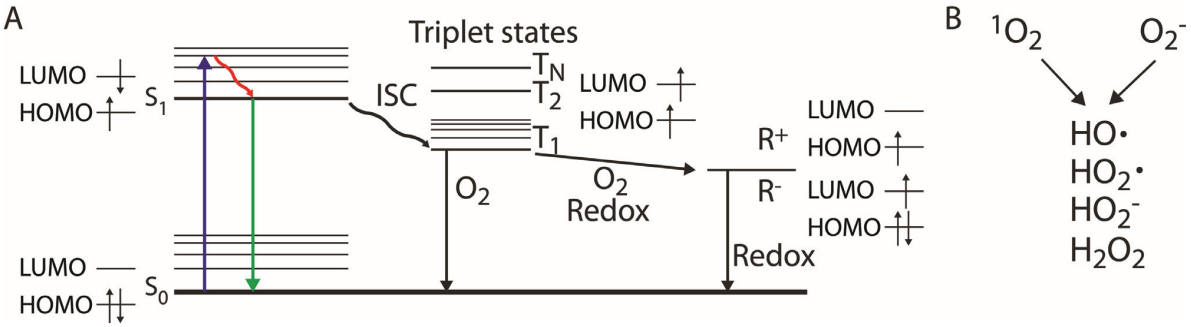

Figure 3.6.1. *Pathways leading to fluorophore inactivation and blinking and molecular oxygen. A. Jablonski diagrams and electron orbital occupation in higher occupied molecular orbitals (HOMO) or lower unoccupied molecular orbitals (LUMO) upon transitions from the ground state S0, to the excited state, inter-system-crossing (ISC) to triplet states and redox reactions for non-fluorescing fluorophores. Molecular oxygen can promote non-fluorescent transitions from triplet states to the ground state and conversion of triplet states to long-lived non-fluorescent anionic or cationic states.* ***B.*** *Possible conversions of molecular oxygen to reactive species. Adapted from Zheng & Blanchard 2014.*

Oxygen can further exacerbate transitions to different non-fluorescent states, through it's ability to form a range of reactive oxygen species (**Figure 3.6.1.B** and oxygen scavenging below)[307]. Although fluorophore blinking is desirable for some applications, in most cases stable, long-lasting fluorescence is the goal[308]. To quench triplet-states the vitamin K analogue Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid) can be employed to significantly reduce both short – and long-lived intensity fluctuations (**Figure 3.6.2.A**)[309,310].

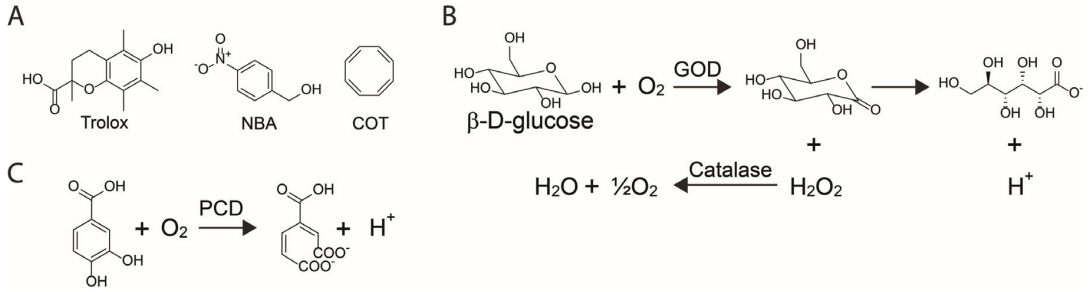

Figure 3.6.2. *Triplet state quenchers and oxygen scavenging.* ***A.*** *Molecular structures of well-characterized triplet-state quenchers for single-molecule photostabilization.* ***B.*** *GODCAT enzymatic oxygen scavenging system. Glucose and oxygen is converted into gluconate and hydrogen peroxide by glucose oxidase, and the hydrogen peroxide is converted by catalase to water and half an equivalent of oxygen.* ***C.*** *Protocatechuate and oxygen are converted by protocatechuate dioxygenase to β-carboxy-cis, cis-muconic acid.*

Further stabilization of fluorescence by both reducing the frequency of intensity fluctuations, shortening the off-times and prolonging the on-times can be achieved by using cyclooctatetraene (COT) and 4-nitrobenzylalchol (NBA) as additional triplet state quenchers (TSQs). This has a particularly profound impact on photostability of fluorophores in the context of nucleoprotein complexes[311]. Ambient oxygen is one of the main reasons for fluorophore blinking and bleaching. Therefore enzymes glucose oxidase (GOD) and protocatechuic acid dehydrogenase (PCD) that utilize oxygen as part of their reaction are employed to scavenge oxygen (**Figure 3.6.2.B and C**) [312,313].

Because of the hydrogen peroxide created by GOD, catalase (CAT) converting this to water is co-employed to avoid build-up of this reactive oxygen species. Each of the enzymatic systems, together with triplet state quenchers have been shown to significantly improve dye photostability[310,314]. Although both according to the reaction scheme should lead to buffer acidification, this is only significant for GODCAT[315].

CoSM with immobilized chromatin and labeled HP1α in solution

Single-molecule microscopy experiments were carried out in microfluidic flow cells with PEGylated borosilicate coverslip surfaces used for TIRF evanescent wave formation. A fraction of the PEG was biotinylated. For photostabilization we used trolox and GODCAT for triplet state quenching and oxygen scavenging as outlined above. Chromatin molecules could be immobilized on the surface in a neutravidin-dependent manner, showing the specificity of surface anchoring, while allowing buffer DNA and nucleosomes formed on this to be washed away (**Figure 3.6.3.A**)

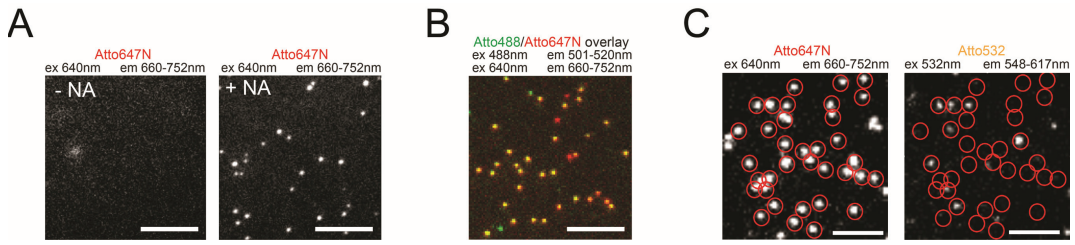

Figure 3.6.3. *Chromatin anchoring and CoSM with HP1α.* ***A.*** *Neutravidin-dependent anchoring of chromatin B. Co-localization of Atto 488 labeled histone octamers with Atto647N labeled DNA indicating chromatin integrity. Green indicates histones and red Atto647N labeled DNA and yellow/orange their co-localization.* ***C.*** *Dynamic co-localization of Atto532 labeled HP1α molecules with Atto647N labeled chromatin. Chromatin molecules are localized with 640nm excitation and 532nm excitation allows localization of HP1α binding. Data contributed by Beat Fierz.*

Tests with chromatin containing histone octamers with H2A labeled with Atto 488 allowed probing of the integrity of chromatin, seeing that histones co-localized with DNA in the majority of cases (**Figure 3.6.3.B**). With chromatin containing H3 K9me3 immobilized in the flow-cells HP1α(Atto532) was added in at a concentration of 1nM. This resulted in regular appearance of HP1α molecules co-localizing with previously localized chromatin molecules (**Figure 3.6.3.C**).

Confirming that we could observe co-localization of HP1α with chromatin, we extracted time-traces for each chromatin peak to monitor the intensity resulting from Atto532 fluorescence upon HP1α binding. The experimental recordings and resulting traces were acquired for periods of up to 10min with 50ms time-resolution. This allowed us to directly probe the binding kinetics of HP1α to modified chromatin under equilibrium conditions at the single-molecule level (**Figure 3.6.4.A**). Binding of

HP1α to chromatin results in an increase in fluorescence intensity until dissociation. Residence times can thereby be directly determined as well as binding rates from the time between binding events (**Figure 3.6.4.A**).

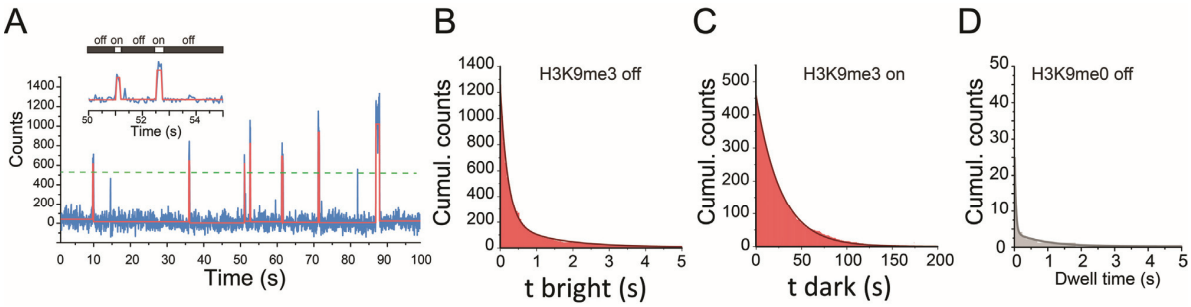

Figure 3.6.4. *Single-molecule binding of HP1α to immobilized chromatin.* ***A.*** *Example binding trace of HP1α to chromatin during 100s showing temporally separated binding events evidenced by fluorescence intensity increases during the time of binding.* ***B.*** *Dissociation kinetics from H3 K9me3: Histogram of cumulated counts of molecules with retention times of x seconds or higher fitted to a double-exponential with $\tau_{off,1}$ = 0.25±0.03s, $\tau_{off,2}$ = 2.26±1.22s. C. Association kinetics to H3 K9me3: Histogram of cumutated counts of times between binding events fitted to a single exponential decay with λ_{on} = 22.9±9.8s. D. Dissociation kinetics from H3K9me0. Very few binding events are observed within the time-resolution of the experiments.*

Evaluation of HP1α binding to more than hundred single chromatin arrays allowed the plotting of histograms for the off and on times respectively (**Figure 3.6.4.B and C**). This showed that HP1α retention times could be fitted well to a bi-exponential function with time constants of ~250ms and 2.26s respectively, suggesting that multiple mechanisms promote retention of HP1α on target chromatin. Conversely the times between binding events could be fitted to a single exponential with a decay rate of 22.9s at the 1nM HP1α concentration used for the experiments. Together this corresponds to an affinity of 1.1μM towards nucleosomes in the context of the 12-mer array. This suggests about 5-fold increased affinity relative to that towards a modified H3 peptide. Incubation with chromatin not containing the target modification resulted in nearly no discernible binding events within the 50ms time resolution of our experiments, highlighting the necessity of the modification for efficient HP1α binding (**Figure 3.6.4.D**).

3.7. Substrate density and facilitated dissociation in HP1α binding

Efficient localization of HP1α requires as seen above, that H3 K9me3 is present. The degree to which possible kinetic trapping of effectors takes place, is not understood. The high substrate density in heterochromatic regions might allow the protein to be retained, by re-binding on adjacent H3 K9me3 at a potentially faster rate than it diffuses away. To test this in our system, we gradually decreased the ratio between wild-type histone octamers and H3 K9me3 used for chromatin assembly. This was done

with 25% intervals, such that we prepared chromatin with 25%, 50% and 75% H3 K9me3 respectively. Employing these arrays in our assay showed that as gaps in the substrate density are present, the retention time decreased by about 24% (from 250ms to 190ms) in going from fully methylated arrays to those with 75% H3 K9me3 (**Figure 3.7.1.A**).

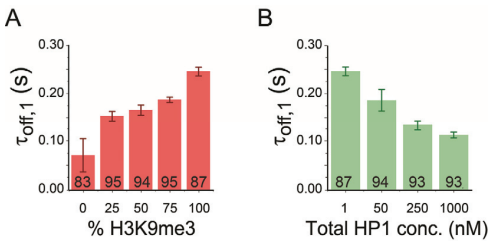

Figure 3.7.1. *Impact of substrate density and facilitated dissociation in HP1α binding to chromatin.* ***A.*** *Retention times of labeled HP1α on chromatin upon changing the density of H3 K9me3 with 25% intervals as controlled in the chromatin assembly.* ***B.*** *Influence of total HP1α concentration on retention of the labeled 1nM HP1α on chromatin. Data contributed by Louise Bryan (A) and Andreas Bachmann (B).*

Further decreasing the substrate density to 50% at 25%, decreased the retention time further by about 10% per step. Finally absence of the mark nearly abolished binding as seen previously. Together, this suggests that HP1α is indeed capable of rebinding on adjacent binding sites closely situated in space.

The oligomerization behavior previously reported for the *S.pombe* orthologue of HP1α, Swi6, could also potentially influence the retention of HP1α. Formation of chromatin-templated oligomers of HP1α would be expected to extend retention of HP1α at concentrations allowing such oligomers to form. Conversely, DNA-binding proteins capable of oligomerization and target binding have been reported to dynamically exchange with non-bound proteins, due to micro-dissociation in a process termed facilitated dissociation[316-318]. To test these possible mechanisms in our single-molecule assay, we systematically increased the concentration of unlabeled HP1α in solution, while maintaining 1nM labeled HP1α (**Figure 3.7.1.B**). In contrast to what would be expected with oligomers retaining HP1α in chromatin longer, the retention time decreased. This shows that competition for binding sites promotes dynamic local HP1α turnover.

3.8. Multivalency in HP1α binding

Strategy for covalent dimerization and peptide synthesis

Nanomolar concentrations of HP1α necessary for characterization of binding with CoSM allows the visualization primarily of HP1α monomers binding to chromatin, since the CSD dimerization constant is ~1μM[282]. Therefore, we required means of ensuring dimers being present even at these

dilute concentrations. Toward this goal, we envisioned a strategy for covalent dimerization, where a peptide motif binding in the cleft between CSDs would be equipped with two cysteines for template-directed dual EPL, thereby linking adjacent HP1α monomers through their C-terminal (**Figure 3.8.1.A**).

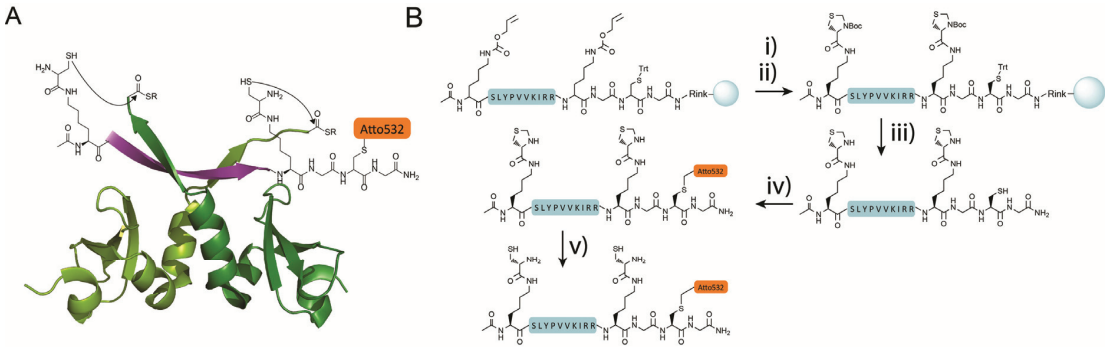

Figure 3.8.1. *HP1α covalent dimerization and labeling peptide.* ***A.*** *Schematic of template-directed dual EPL with peptide in the cleft between adjacent CSDs. B. Synthetic scheme for generation of labeled peptide for dual EPL. i) Alloc-deprotection. ii) Coupling of Boc-Thz. iii) Cleavage from resin and side-chain deprotection with TFA/TIS/H2O. iv) Conjugation of Atto532 iodoacetamide. v) Thz deprotection with methoxyamine.*

The Atto532 peptide with two available cysteines for EPL was synthesized by Fmoc SPPS according to the scheme (**Figure 3.8.1.B**). Labeling of the trityl deprotected cysteine in the backbone followed by Thz deprotection provided the orthogonal control of the cysteines required in the peptide followed by purification and final analysis (**Figure 3.8.2**).

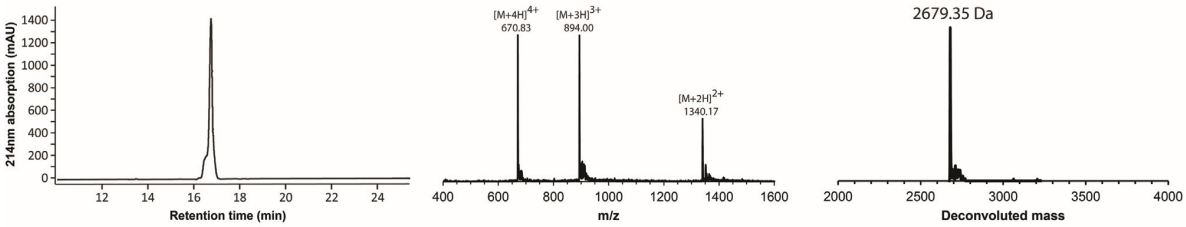

Figure 3.8.2. *Final analysis of labeled peptide for covalent dimerization by RP-HPLC and MS. M_{obs} = 2679.4, M_{calc} = 2680.2. Data contributed by Beat Fierz.*

The final analysis confirms the identity of the produced peptide.

Thioester formation and *in situ* EPL for covalent dimerization and labeling HP1α

HP1α(Δ177-190)-GG-Npu^N was expressed and partially purified by Ni-NTA and AIEX as for the wild-type HP1α. The deletion of the unstructured C-terminal, was to ensure that the termini of the CSD dimers were both within reach for dimerization. In situ EPL and dimerization using the split intein column resulted in the desired covalent dimer, with a smaller fraction of the monomeric labeled

HP1α (**Figure 3.8.3.A**). Subsequent purification by gel filtration yielded fractions for pooling that contained more than 90% of the desired covalently dimerized labeled HP1α (**Figure 3.8.3.B**)

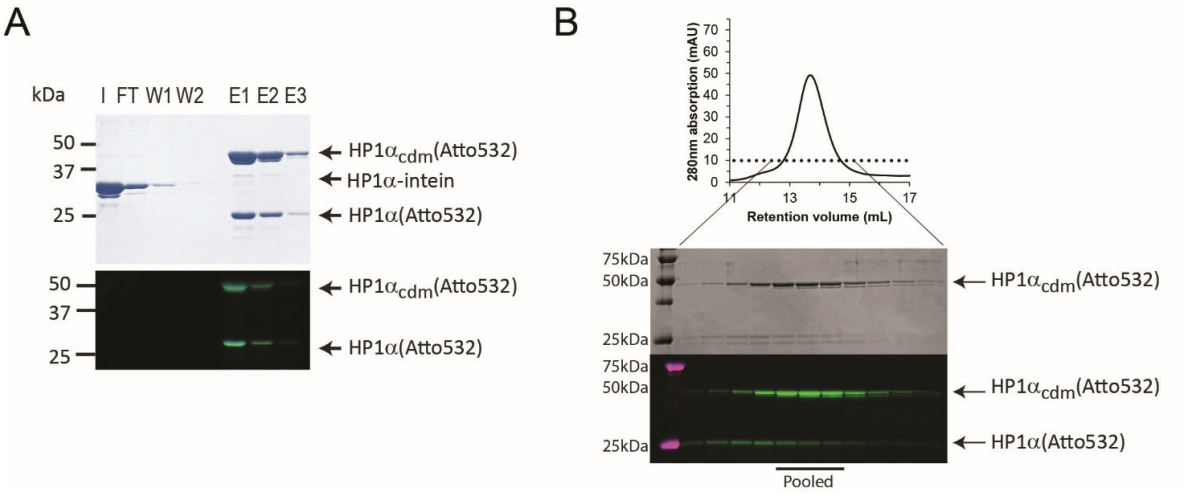

Figure 3.8.3. *Generation and purification of labeled covalently dimerized HP1α. A. SDS-PAGE analysis of input, flowthrough, washes and elutions from intein column resulting in formation of covalently dimerized and labeled HP1α. B. Gel filtration purification of eluted covalently dimerized HP1α with indication of the pooled fractions.*

Final analysis by SDS-PAGE, RP-HPLC and MS showed that the degree of contaminating labeled monomeric HP1α was less than 10% (**Figure 3.8.4**)

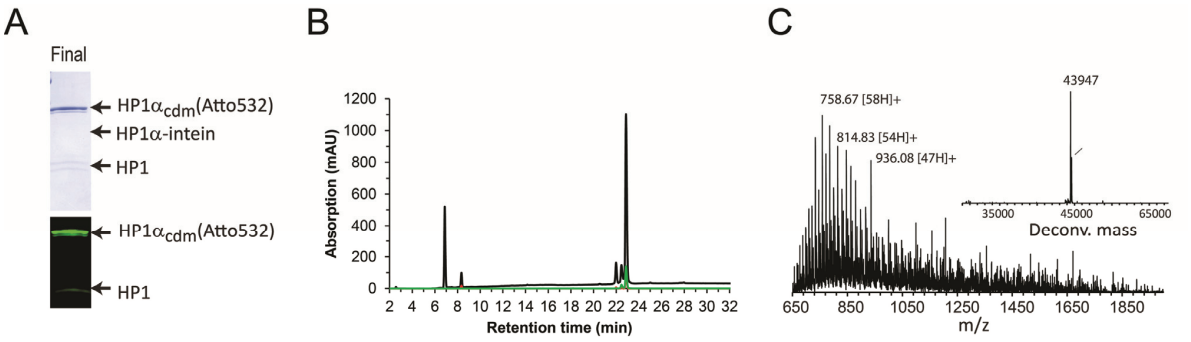

Figure 3.8.4. *Final analysis of labeled covalently dimerized HP1α. A. SDS-PAGE analysis of final purified sample.* ***B.*** *RP-HPLC analysis.* ***C.*** *MS analysis of HP1α$_{cdm}$(Atto532). M_{obs} = 43947, M_{calc} = 43944.6.*

Binding of covalently dimerized HP1α to H3 K9me3 chromatin

Next, to test the possible enhanced binding ability of the HP1α dimer relative to the monomeric form mostly present at 1nM, we employed our covalently dimerized form of the protein. First, we verified by MST that this bound to a H3 K9me3 tail peptide with similar affinity as for the wild-type (Figure **Figure 3.8.5.A**). This was observed at this level to be ~1.7-fold lower (12.4μM vs 5.4μM). Upon

incubation of this in at 0.5nM (Same monomer concentration as for the non-dimerized form), there was a marked prolonging of both the retention times and increase in binding frequency (**Figure 3.8.5**).

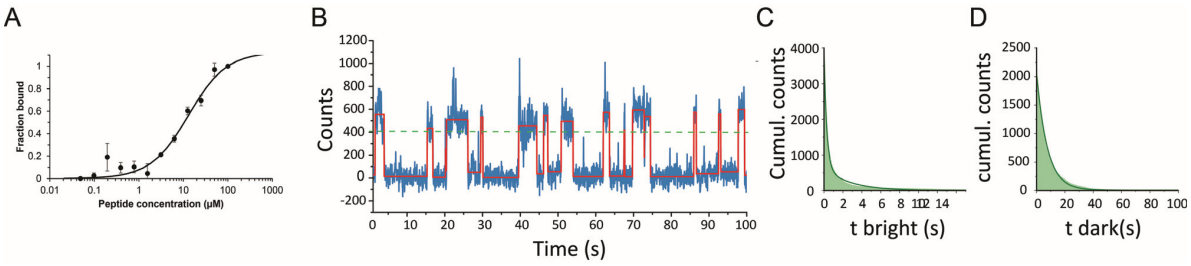

Figure 3.8.5. *Binding of covalently dimerized HP1α to peptide and chromatin arrays.* ***A.*** *Titration curve from MST of covalently dimerized labeled HP1a with H3 K9me3 peptide. Affinity determined is 12.5±0.5μM* ***B.*** *Example trace of HP1α$_{cdm}$ binding to H3 K9me3 modified chromatin.* ***C.*** *Histogram and double-exponential fit for the retention times of covalently dimerized HP1α, $\tau_{off,1}$ = 0.33±0.01s. $\tau_{off,2}$ = 3.40±0.53s* ***D.*** *Histogram and exponential fit for the times between binding events with HP1α$_{cdm}$. λ_{on} = 7.45±1.87s.*

When compared to the binding kinetics of the wild-type, it was clear that the overall binding kinetics had changed towards faster binding times and longer retention times (**Figure 3.8.6.A, B and C**).

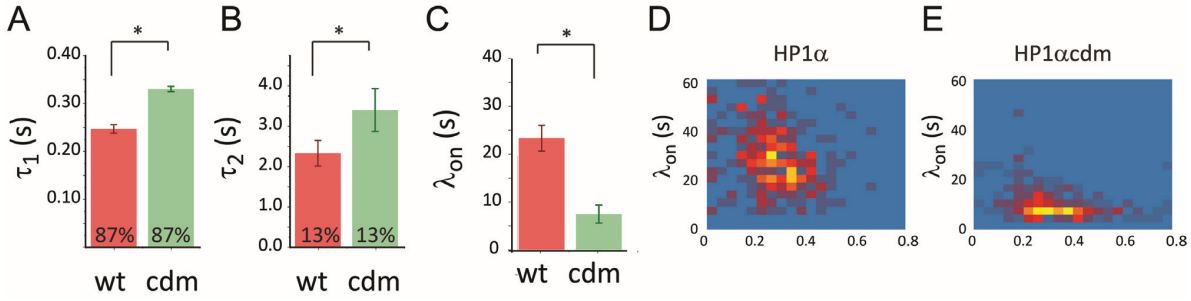

Figure 3.8.6. *Analysis of overall binding times of HP1α$_{cdm}$ and to single chromatin arrays.* ***A.*** *Bar plot of the first decay rate in bi-exponential fit from labeled wild-type (wt) and covalently dimerized (cdm) HP1α. Numbers at the base of the bars indicate the amplitude of the phase in the double-exponential fit* ***B.*** *Bar plot of the second decay rate in bi-exponential fit.* ***C.*** *Bar plot of the association times.* ***D.*** *Heat map plot of HP1α wild-type dissociation (First axis) and association (Second axis) from/to individual chromatin arrays.* ***E.*** *Heat map plot of dissociation and association times for covalently dimerized HP1α to individual chromatin arrays.*

Also when looking at the level of individual chromatin arrays, particularly the distribution of association times had narrowed down to significantly faster binding times (Second Axis in Figure **Figure 3.8.6.D and E**), whereas the prolonged retention (First axis) was less pronounced. These results show that HP1α dimers bind to chromatin at a significantly accelerated rate and also prolongs retention. Compounded, when kinetics and concentrations are calculated into dissociation constants, this corresponds to 0.13μM for the covalent dimer.

Binding of HP1α to H3 K9me3 chromatin in the presence of Shg1 peptide

To test if improve binding as seen above was a result of the artificial covalent dimerization, we further tried to promote dimerization non-covalently by adding in the Shg1 peptide. This peptide bound to

the labeled HP1α with an affinity of 110nM as determined by MST (**Figure 3.8.7.A**). At 500nM together with 1nM of labeled non-dimerized HP1α the CoSM experiments displayed an influence on the binding kinetics largely recapitulating the observations done with the covalently dimerized protein (**Figure 3.8.7.B, C and D**).

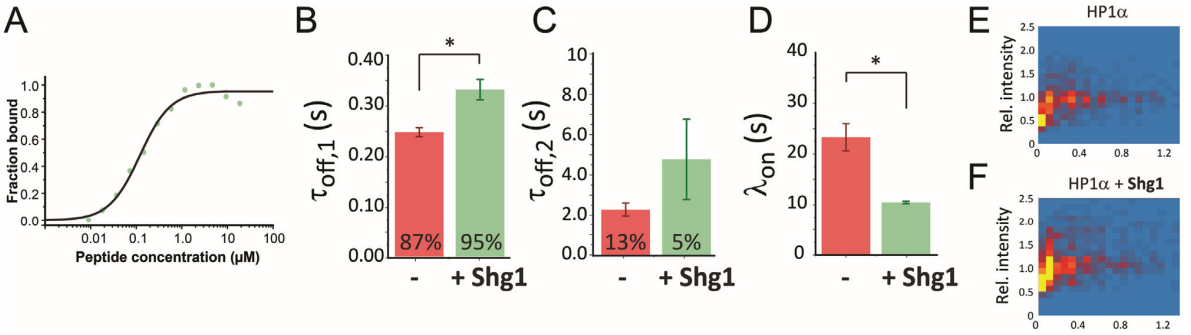

Figure 3.8.7. *Analysis of overall binding times of HP1α wild-type and retention on single chromatin arrays in the presence of the Shg1 peptide.* ***A.*** *Affinity of Shg1 peptide determined by MST. Affinity observed is 110nM.* ***B.*** *Bar plots of first decay rate in bi-exponential fit of HP1α binding to modified chromatin with Shg1 peptide (Green) and without (Red).* ***C.*** *Bar plots of second decay rate in fit similar to A.* ***D.*** *Association times for HP1α wild-type in the absence and presence of Shg1 peptide.* ***E.*** *Heat map plots of both first (primary axis) and second (secondary axis) rate constants in bi-exponential fits for individual chromatin arrays without Shg1 peptide.* ***F.*** *As in E, but with the Shg1 peptide. Data contributed by Andreas Bachmann.*

Since both of the HP1α monomers in this experiment contained a fluorophore, analysis of the fluorescence intensities of binding events should also inform on the protein stoichiometry. In absence of the Shg1 peptide, the relative intensities were typically between 0.5 and 1.0 (Variations likely caused by the location of bound nucleosome(s) relative to surface and evanescent wave) (**Figure 3.8.7.E**). With the peptide added, in conjunction with the increase in residence times, the relative fluorescence intensities also increased to 0.75-1.5 displaying the ability of the peptide to promote dimerization (**Figure 3.8.7.F**).

3.9. FRAP measurements on HP1α mobility in vivo

Finally, to probe the relevance of the multivalency on binding in vivo we carried out confocal microscopy and FRAP. Previously, FRAP measurements have been carried out on HP1 wild-type and with a mutant without the CSD[285]. To probe the dynamics in a cellular context we cloned HP1α constructs fused at the N-terminal to the strictly monomeric fluorescent protein mEos3.2. Together with HP1α mutants I164E and W173A, this would allow us to more finely discriminate between the ability to dimerize and the potential for dimerization to be promoted by partner binding relative to the removal of the entire CSD domain (**Figure 3.9.1.A**). These mutants abolish dimerization (I164E) and disallows binding of partners with PXVXL motifs respectively[282]. All three constructs could be

efficiently transfected into NIH 3T3 mouse embryonic fibroblasts (MEFs). The distribution of fluorescence correlated with that of DAPI staining, showing that the protein correctly localized to heterochromatic foci (**Figure 3.9.1.B**).

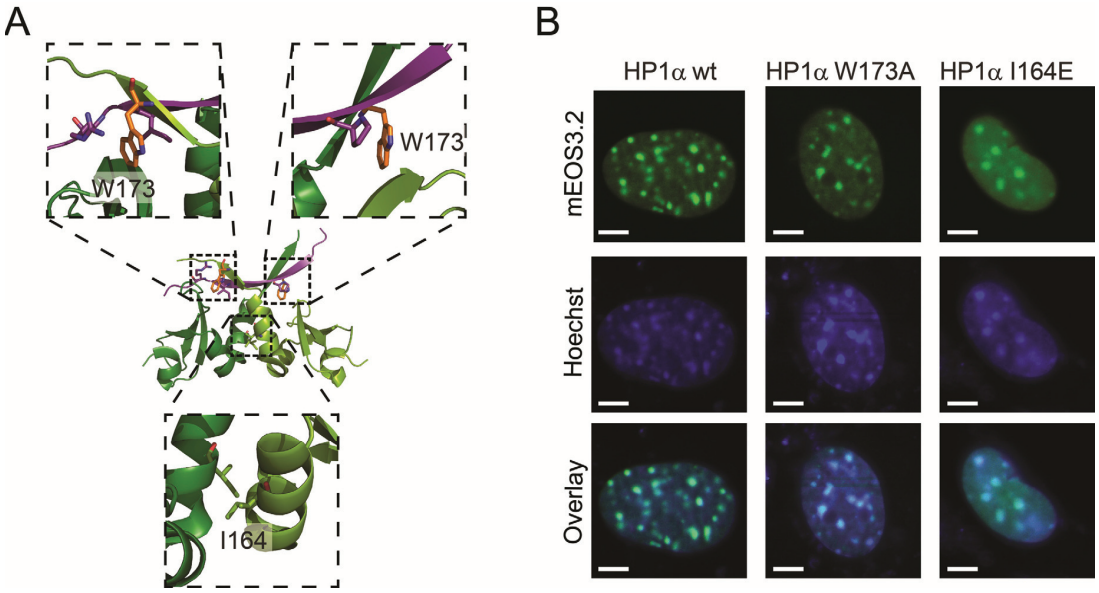

Figure 3.9.1. *Location of residues in CSD and microscopy images of of mEos3.2-HP1α expressing cells.* ***A.*** *Structural outline of HP1α CSD highlightning the residues mutated to modulate dimerization ability (I164) and ability to interact with PXVXL motifs.* ***B.*** *Images from confocal microscopy of NIH 3T3 MEFs expressing mEos3.2-HP1α wild-type and mutants W173A and I164E.*

Upon bleaching of heterochromatic foci, we saw recovery occurring at a time-scale of seconds (**Figure 3.9.2.A**). Plotting of the relative fluorescence intensities before and after bleaching for more than 25 heterochromatic foci with wild-type HP1α resulted in a FRAP curve in close agreement with previously published results [254,285] (**Figure 3.9.2.B**).

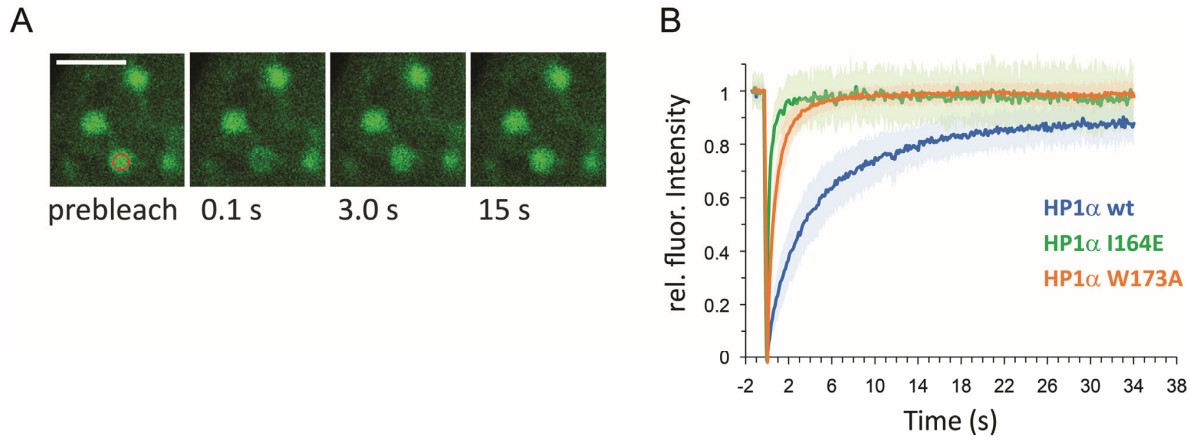

Figure 3.9.2. *FRAP experiments for cellular mobility of HP1α wild-type and mutants.* ***A.*** *Confocal images of cell nucleus with heterochromatic foci containing high concentrations of HP1α before bleaching (prebleach), 0.1s, 3.0s and 15s after photobleaching in red circle.* ***B.*** *FRAP curves obtained from experiments in more than 25 heterochromatic foci in different individual cells using the wild-type, the I164E and the W173A mutants.*

The wild-type displayed half-recovery after 4s and a 15-20% fraction remained immobile. In contrast to this, the mutant I164E incapable of dimerizing fully recovered after less than 2s, and there was no immobile fraction. This displays how monomeric HP1α is unable to efficiently compete with the endogenous HP1 proteins and retain binding in target heterochromatic regions. Along similar lines, the W173A mutant, where dimerization could not be promoted and stabilized by native binding partners lost the ability to retain in heterochromatic foci nearly to the same extent as to that of the protein unable to dimerize at all. This highlights the importance of multivalency for efficient heterochromatin targeting *in vivo*.

3.10. Discussion, conclusions and outlook

Our experiments suggests a model where substrate density, dynamic local competition and effector multivalency all contribute to microscopically dynamic but macroscopically stable binding of HP1. The effect of multivalency on the binding of the protein can be summarized with specific binding states (**Figure 3.10.1.A**).

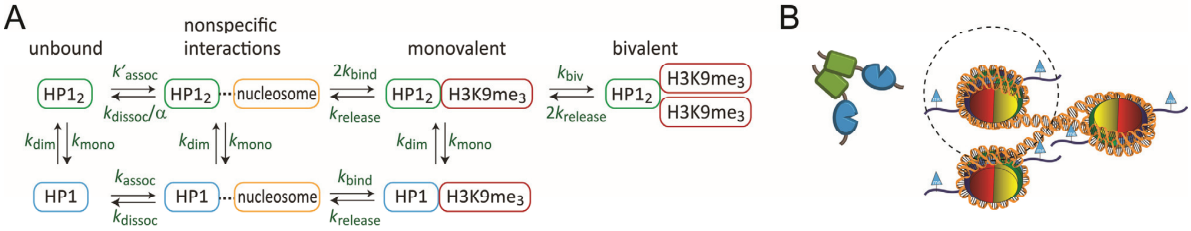

Figure 3.10.1. *Possible binding modes and mechanism of increased association rate.* ***A.*** *States of HP1 in solution and bound to chromatin as a monomer (Lower line) and as a dimer (Upper line). The equilibrium between monomer and dimer is governed by the rates k_{dim} and k_{mono}. Initial non-specific association occurs with a rate of k_{assoc}, and dissociation with k_{dissoc} prolonged with a factor α in the case of the HP1 dimer. Binding and unbinding rates are controlled by k_{bind} and $k_{release}$. Bivalent binding can occur with a rate kbiv and an increased rate of a single release from that state.* ***B.*** *Schematic of a likely reactive radius around the nucleosomes mediated by long-range electrostatic interactions for initial non-specific association. The chance of non-specific association with the larger surface of the dimer is increased as well as that of H3 K9me3 binding doubled with the presence of two reader domains.*

A preliminary non-specific association due to the negatively charged surface of HP1 interacting with the positively histones likely brings HP1 in a position to bind to the target H3 K9me3 and stay long to be monitored during the time-resolution of our measurements. The dimer contributes to the non-specific association, the slowdown of dissociation, the chance of H3 K9me3 binding and possibility of bivalent binding, thereby altogether explaining the increase in both observed increase in association rate and slowdown in dissociation (**Figure 3.10.1.A and B**). In a heterochromatic cellular environment, the concentration of HP1α is ~2μM [254]. Interestingly, the dimerization constant for the CSD is just around 3μM[282]. Together with our measurements *in vitro* and *in vivo*, this highlights the

importance of HP1 multivalency promoted by binding partners. This seems to function as a regulatory platform to allow HP1 to effectively compete for sites and be retained in target heterochromatin. The dynamic micro-dissociation and increased binding capacity of dimers results in rapid eviction of monomers thereby retaining primarily dimers that are bridging the H3 K9me3 mark and HP1 binding partners with the PXVXL motif. The measurements argue against the formation of larger oligomeric network of HP1, but rather emphasizes the dynamic turnover promoted by high local concentrations of HP1.

In addition to the detailed understanding of the mechanisms underlying effective HP1 binding to chromatin, the developed method will allow for the detailed interrogation of other effector protein binding in the context of short chromatin arrays.

Since these measurements were carried out, the methodology was expanded to encompass chromatin with longer linker lengths (50bp instead of 30bp), chromatin with linker histones and chromatin with single recognition sites for transcription factors to probe the targeting of these in the context of chromatin.

The residence time of HP1 measured here is comparable to that of other chromatin effectors, as judged from FRAP and FCS[254,255]. These relatively short <500ms times, challenges the potential of monitoring multi-step reactions with effectors occurring on chromatin in real-time. The time-windows for the co-existence of several factors and the lack of long-term retention of any of them is probably too brief compared to that of e.g. components in the spliceosome or the steps in origin licensing[290,295]. More likely it would be useful simply in monitoring their co-arrival on chromatin vs. the frequencies of them arriving alone.

4. Chromatin conformational dynamics by smFRET

4.1. Project introduction and outline

4.1.1. Chromatin secondary structure and conformational dynamics

Static glimpses of chromatin secondary structure have been obtained by crystallography and cryo-em[58,59,61]. In the cellular context however, factors requiring access to the nucleosomal surface or the DNA need to operate with an ensemble of chromatin conformations and within the timeframes of the specific states required for binding. Little is however known about the hierarchical chromatin folding and their transition kinetics, intrinsic equilibrium fluctuations within chromatin secondary structure and how such conformational dynamics are affected by histone PTMs and effector proteins. HP1 generally induces chromatin states of increased compaction visualized by decreased radius of 12-mer arrays visualized by electron microscopy and by an increased sedimentation coefficient[70,319]. Bridging of nucleosomes by two HP1 molecules upon dimerization is strictly necessary for such compaction to take place[135]. This can occur both within a single chromatin segment or between different chromatin segments[277]. These studies have established that HP1 promotes chromatin condensation, however the spatiotemporal dynamics of such compacted chromatin and how this can be correlated with dynamic HP1 binding is not well understood. Numerous studies using fluorescence microscopy/spectroscopy and force spectroscopy have been devoted to study the kinetics of DNA unwrapping in individual nucleosomes, short di – or trinucleosomes and forces required for their disassembly[60,62,71,72,320]. Although magnetic and optical tweezers have been used to investigate the forces required to stretch chromatin fibers, these experiments do so under invasive conditions. Further, they provide information based on structural rearrangements that manifest themselves as changes in the anchoring points, which are mostly at the end of chromatin, thereby providing mainly global conformational information. In contrast single-molecule FRET (smFRET) is uniquely capable of yielding local information with high spatiotemporal resolution.

4.1.2. smFRET

Instead of fluorophores emitting photons resulting in fluorescence upon return to the ground state, the corresponding energy can be transferred non-radiatively from a higher energy fluorophore (the donor) to a lower energy fluorophore (the acceptor) closely situated in space (10-120Å). This is referred to as förster (/fluorescence) resonance energy transfer (**Figure 4.1.1.A**)[321].

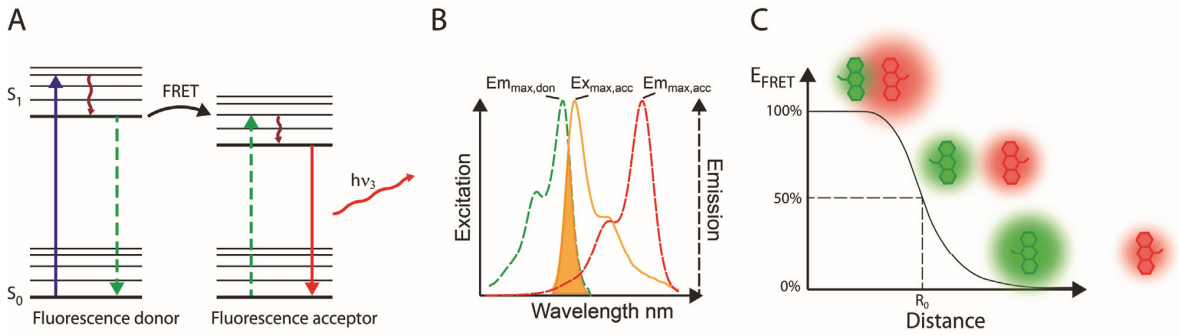

Figure 4.1.1. *Energetic basis and distance-dependence of FRET.* ***A.*** *Jablonski diagram displaying the transfer of energy from a donor to excite an acceptor that upon return to the ground state emits photons with a wavelength longer than that emitted from the donor.* ***B.*** *Emission spectrum of an example donor and the spectral overlap with the excitation spectrum of the acceptor as well as the emission spectrum of the acceptor.* ***C.*** *Schematic of the inter-dye distance dependency of FRET. Within close proximity, almost all energy is transferred. At a distance of* R_0*, half of the energy is transferred. At distances above ~120-150Å no energy is transferred.*

An overlap in the integral between the emission spectra of the fluorescence donor and the excitation spectra of the acceptor is required for FRET to occur (**Figure 4.1.1.B**). The spectral overlap integral is part of determining the Förster radius (R_0), which is the distance between two fluorophores at which 50% of the energy is transferred to the acceptor (**Figure 4.1.1.C** and equation 4.1.).

$$R_0{}^6 = \frac{9Q_D(ln10)\kappa^2 J}{128\pi^5 n^4 N_A} \qquad (4.1)$$

Q_D is the donor quantum yield, k is an orientation factor, J is the spectral overlap integral, n is the refractive index of the medium and N_A is Avogadro's number. Because of the dependence on spectral overlap and quantum yield, the Förster distance is a property of every set of FRET pairs. The extent of FRET is inversely proportional to the distance, r, between the two fluorophores to the power of six (**Figure 4.1.1.C** and equation 4.2).

$$E_{FRET} = \frac{1}{1 + \left({}^{r}/_{R_0}\right)^6} \qquad (4.2)$$

Due to this, FRET can be used as a molecular ruler to probe intra – or intermolecular distances as shown by early experiments of end-labeled peptides with increasing lengths[322,323]. Increasing proximity of the fluorophores further leads to an increased rate of de-excitation of the donor, thereby reducing the donor fluorescence lifetime (equation 4.3).

$$E_{FRET} = 1 - \frac{\tau_{DA}}{\tau_D} \Leftrightarrow \frac{\tau_{DA}}{\tau_D} = 1 - \frac{1}{1 + \left({}^{r}/_{R_0}\right)^6} \qquad (4.3)$$

The characteristics of FRET can also be employed for smFRET either with immobilized molecules or molecules in solution [324,325]. In a single-molecule setting, increases in FRET efficiency are caused

by conformational changes within a single-molecule, thereby allowing a distinction between inter-molecular and intra-molecular changes not possible with ensemble assays. smFRET enables conformational changes to be observed in real-time with nanometer spatial resolution. Finally population distributions resulting from heterogenic conformations can be distinguished and attributed to particular structural states[252,326]. In conjunction with TIRF, simultaneous monitoring of the donor and acceptor fluorescence intensity through splitting of the emission path into donor and acceptor fluorescence intensity allows calculation of the FRET efficiency at any given time (**Figure 4.1.2.A**).

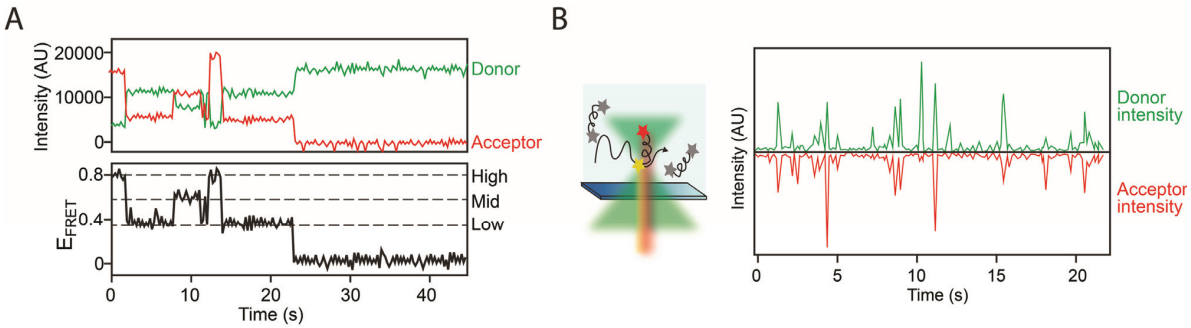

Figure 4.1.2. *Output from smFRET with TIRF and in a confocal volume in solution.* ***A.*** *Example fluorescence intensity and FRET trace of a single molecule undergoing transitions between three distinct FRET states.* ***B.*** *Scheme of FRET labeled molecule traveling through confocal volume and the resulting intensity bursts from experiments with smFRET in solution*

This allows conformational changes occurring on the 10ms-hour time-scale to be observed in real-time. Complementarily smFRET in solution probes conformational changes that can occur on the µs-20ms time-scales. For smFRET in solution, fluorescence emission is recorded as short bursts corresponding to the time of travel through the femtoliter confocal volume (**Figure 4.1.2.B**). Combined with pulsed interleaved excitation (PIE) lasers and time-correlated single-photon counting (TCSPC), the fluorescence lifetime of the donor can be used in conjunction with the fluorescence intensities to interrogate possible fast dynamics.

Nucleosomes have been extensively studied using smFRET both with TIRF and in solution to probe e.g. the stepwise disassembly and unwrapping of DNA from nucleosomes (by labeling different positions within the DNA, or one fluorophore in the DNA and one in the histones). Both in solution and with immobilized samples, this has suggested that DNA in single nucleosomes can unwrap spontaneously modulated by histone PTMs and that disassembly takes places through distinct intermediates (See previously).

Other diverse examples of smFRET employed to go where ensemble methods have proved insufficient, include the bendability of DNA, the resolving of holliday junctions, presynaptic filament target search, rates and transitions in the ribosome during translation, the conformational changes in

ligand/voltage-gated gated channels and transporters, rotation of ATP synthase during synthesis and the conformational changes in the replication fork during progression[327-336].

4.1.3. Experimental outline and goals

smFRET impinges on site-specific FRET pairs. In single repeats of DNA, FRET pairs are easily introduced by PCR with labeled primers. Challenges in preparing samples with multiple 601 DNA repeats containing FRET pairs have limited the use of FRET for investigation of dynamics of nucleosomes in the context of chromatin and inter-nucleosomal dynamics. Therefore we aimed to develop a method to introduce FRET pairs into 12x601 (or longer) nucleosome arrays to investigate inter-nucleosome dynamics in the context of chromatin fibers.

We envisioned a strategy involving ligation of recombinant (containing multiple 601, but not easily amenable for site-specific labeling) and PCR-generated (Containing single 601 with fluorophores) DNA[71,265,337]. Based on this, 5 segments with unique non-palindromic overhangs were designed for ligation (**Figure 4.1.3**).

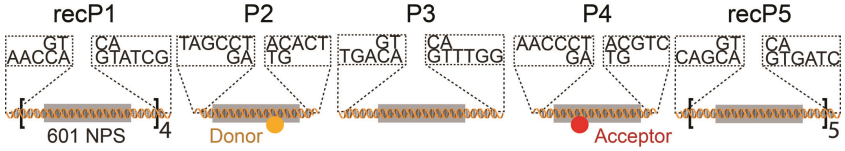

Figure 4.1.3. *Basic scheme for the ligation of 5 pieces of recombinant multi-601 fragments (Pieces 1 and 5) and PCR-generated 1x601 fragments with differently located fluorophores (Pieces 2, 3 and 4) to introduce FRET pairs into 12-mer chromatin arrays.*

Upon succeeding in developing a strategy to introduce these labels, the resulting chromatin templates would allow us to locally probe inter-nucleosome dynamics. An additional requirement is the identification of labeling positions in different nucleosomes resulting in FRET upon compaction.

Based on published trinucleosome experiments, structural modeling with the available tetranucleosome crystal structure and 12-24 nucleosome array cryo-EM structures we chose three distinct FRET pair locations[58,59,71]. These positions, if responsive to compaction by increased FRET, would inform on dyad-distal inter-nucleosomal stacking (**I**nter-**n**ucleosome **FRET** 1: $IN_{FRET}1$), dyad-proximal nucleosome stacking ($IN_{FRET}2$) and movements in the linker DNA ($IN_{FRET}3$) (**Figure 4.1.4**)

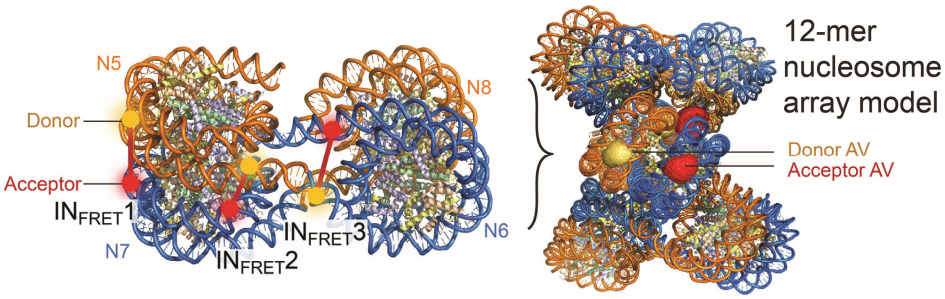

Figure 4.1.4. *Structural modeling and labeling sites for investigation of inter-nucleosome dynamics in chromatin. Labeling sites for reporting of different movements are indicated in the tetranucleosome stack and the corresponding accessible volumes (AV) indicated in the 12-mer nucleosome array.*

With this we wanted to carry out ensemble and single-molecule FRET studies to understand the local structural changes in chromatin conformation upon compaction, the kinetics underlying these changes and how they would be affected by a PTM and by binding of HP1. Long smFRET measurements with 100ms time-resolution would be carried out using TIRF microscopy. For these we planned to use the FRET pair Alexa Fluor 568 and Alexa Fluor 647. These have previously been used for tri-nucleosomes and have one of the longest available R_0 (82Å) allowing to span long molecular distances[71]. Investigation of intrinsic fluctuations in the secondary structure would be carried out with smFRET in solution coupled to multi-parameter fluorescence detection for times ~15ms and lower. Realizing these goals, required us to:

- Clone and isolate recombinant segments for the multi-601 unlabeled parts of the array DNA.
- Label internally modified oligos for PCR and site-specific dye positioning in single 601 DNA.
- Establish near-quantitative ligations of 5-6 DNA pieces and the purification of the product.
- Assemble chromatin with double-labeled DNA.
- Ensemble compaction for possible FRET.
- Establish smFRET with TIRF microscopy and characterize FRET between Alexa 568/647.
- Optimize smFRET for chromatin. Determine distribution and dynamics of FRET populations.
- Investigate the influence of H4K16ac and HP1α binding on chromatin conformation.
- Measure smFRET in solution with MFD to determine fast dynamics.
- Develop models capable of explaining the observed data.

Where indicated in figure legends, Beat Fierz or the lab of Claus Seidel contributed to this project

4.2. Design and cloning of recombinant pieces

Recombinant constructs were designed with four main considerations in mind:

1. Individual constructs should contain digestion sites at the ends to produce unique non-palindromic overhangs for subsequent ligations to produce the assembled DNA. For this we used BsaI/DraIII.
2. The recombinant constructs should encompass the linker DNA to allow labeling sites in PCR generated fragments to reach far into the nucleosomal DNA using primers as short as possible.
3. Cloning of the multi-NPS segments should be modular allowing pieces to be easily transferred between different constructs.
4. Digestion of the final assembled piece with ScaI should result in equidistance space between 177bp nucleosomal repeats for quality to ensure DNA saturation with histone octamers.

Based on this, three recombinant constructs were designed (**Figure 4.2.1**).

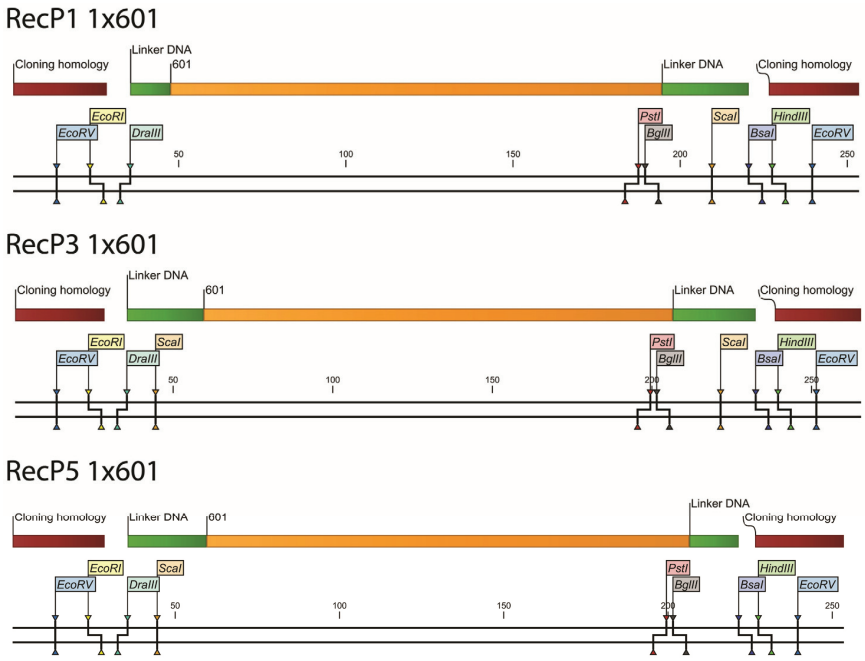

Figure 4.2.1. *Outline of the three recombinant segments indicating restriction sites surrounding the 601 site to facilitate cloning (EcorI, HindIII, PstI and BglII), plasmid excision and ligations (EcorV, BsaI and DraIII) and chromatin analysis (ScaI). Designed in collaboration with Beat Fierz.*

These were cloned by circular polymerase extension cloning into a modified version of the pWM530 vector containing additional EcorV sites exteriorly neighbouring the MCS corresponding to the homology regions in the inserts outlined above[57,338,339] (Named: pWM531). This would ensure that EcorV degradation of the backbone wouldn't result in long fragments next to the BsaI/DraIII sites.

The array in P5 1x601 was then initially extended to 2x601 with an extension piece designed taking into account the main design considerations (**Figure 4.2.2.A**).

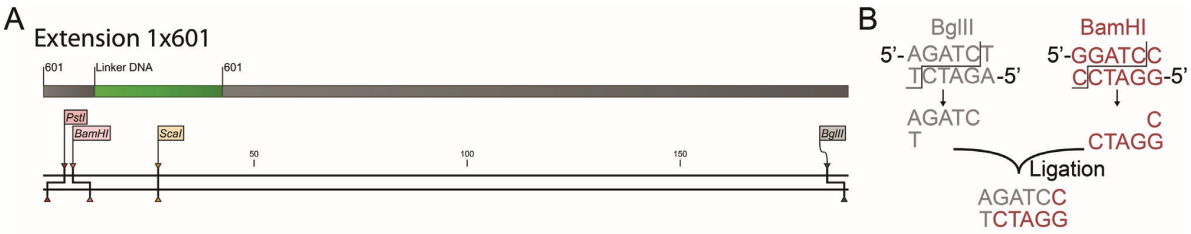

Figure 4.2.2. *Extension piece and BglII/BamHI hybrid.* ***A.*** *Outline of the 1x601 extension piece used for first doubling with sites for doubling (PstI, BamHI and BglII) and chromatin analysis (ScaI).* ***B.*** *The hybrid junctions generated from BamHI/BglII ligation. Designed in collaboration with Beat Fierz.*

3x601, 4x601 and 5x601 were then cloned in P5 through 2 additional rounds of DNA doubling taking advantage of the compatibility of BglII and BamHI digestion overhangs that their ligation results in hybrid junctions not recognized by either of the two (**Figure 4.2.2.B**). In this way, the library of 1-5 nucleosome repeats was cloned (**Figure 4.2.3**).

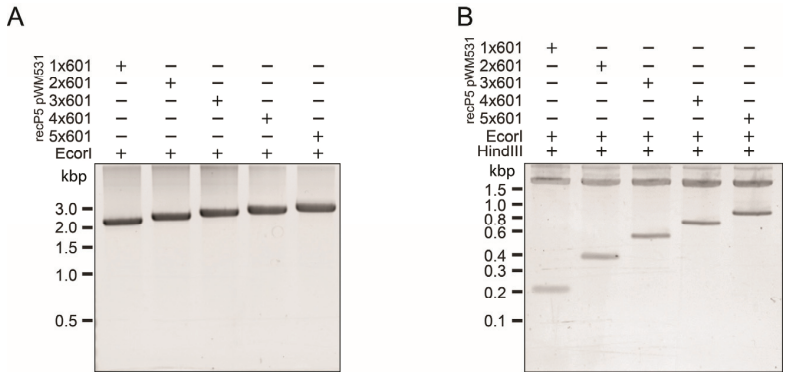

Figure 4.2.3. *Analysis of cloned recP5 pWM531 multi-601 (1-5) constructs upon digestion.* ***A.*** *Digestion with EcorI linearizes the plasmid and display a subtle ladder corresponding to the different lengths.* ***B.*** *Further digestion with HindIII excises the repeats and clearly shows their increasing repeat length rising ~180bp for each increase.*

Finally, the modularity of the fragments allowed the 3x601 extension to be transferred to P1 1x601 array to produce the P1 4x601 construct. Excision of this from the backbone by EcorV confirmed that this had been generated (**Figure 4.2.4**).

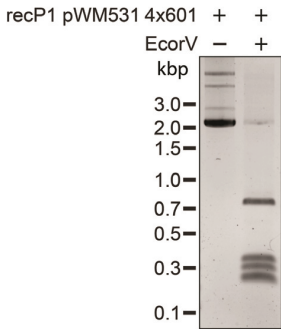

Figure 4.2.4. *Digestion of recP1 4x601 pWM531 with EcorV results in the piece of interest excised from it's fragmented plasmid backbone.*

The pWM531 P1 4x601 and P5 5x601 plasmids were then grown at a large scale to isolate the recombinant fragments of interest.

4.3. Large-scale production and isolation of recombinant pieces

The plasmids grown on a 6L scale were isolated by alkaline lysis and gel filtration by collection in the dead volume from a manually packed 560mL sepharose 6 column (**Figure 4.3.1**).

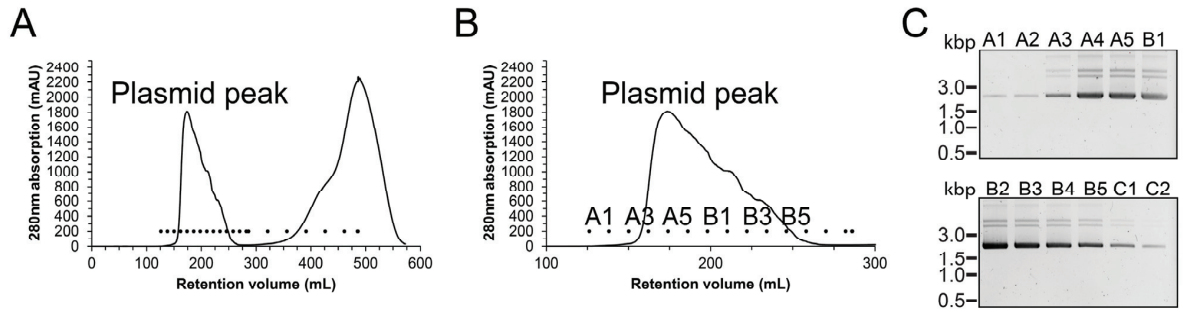

Figure 4.3.1. *Final steps of large-scale plasmid purification by alkaline lysis and gel filtration.* ***A.*** *Full chromatogram of recP1 4x601 plasmid purification by gel filtration in 2M KCl to isolate the plasmid from proteins/RNA/small DNA dissociated from the plasmid at this salt concentration and eluting later.* ***B.*** *Zoom-in on the plasmid peak with indication of numbers of collected fractions.* ***C.*** *Analysis of fractions from purification showing good correspondence between location in the chromatogram and concentration of plasmid.*

The purification typically yielded 10-40mg of pure plasmid.

Isolation of the segment of interest was done by initial digestion with BsaI and DraIII to ensure that the non-palindromic overhangs were fully accessible for ligations. Digestion with EcorV then degraded the backbone and ensured that the pieces of interest had their non-palindromic overhangs completely available. This was done for both recP1 4x601 and recP5 5x601 (**Figure 4.3.2**)

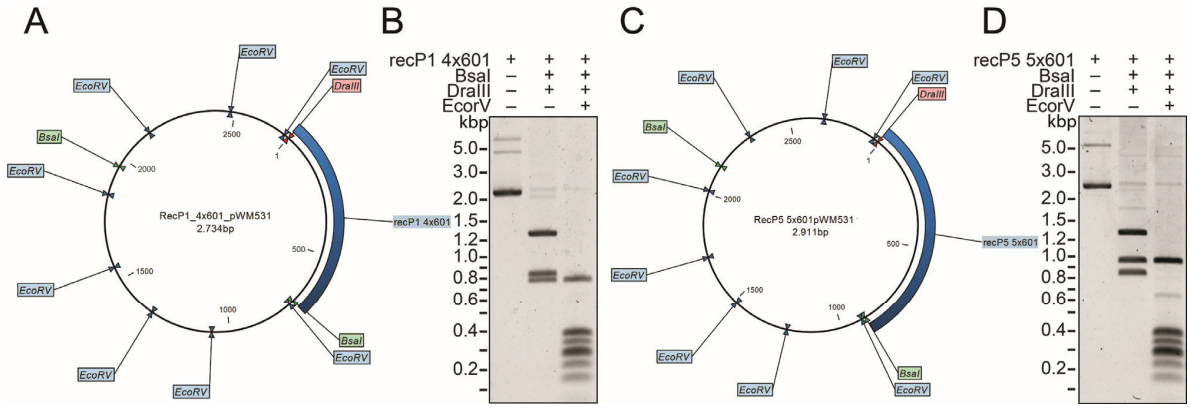

Figure 4.3.2. *Plasmid outlines and digestion patterns ensuring complete digestion of palindromic sites. EcorV digestion sites in blue, BsaI in green and DraIII in red.* ***A.*** *Outline of the recP1 4x601 pWM531 plasmid. Digestion with BsaI and DraIII should yield fragments of sizes 716bp, 1245bp and 773bp.* ***B.*** *Digestion of recP1 4x601 pWM531 with BsaI/DraIII and EcorV resulting in fragments of the expected sizes.* ***C.*** *Outline of recP5 5x601 pWM531 plasmid. Digestion with BsaI and DraIII should yield fragments of sizes 894bp, 1244bp and 773bp.* ***D.*** *Digestion of recP5 5x601 pWM531 with BsaI/DraIII and EcorV resulting in fragments of expected sizes.*

Subsequently, each of the fragments were purified from the plasmid backbone fragments by iterative PEG precipitations (**Figure 4.3.3**).

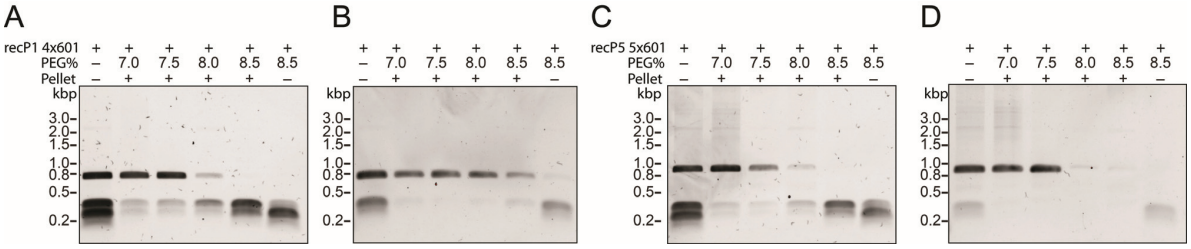

Figure 4.3.3. *PEG precipitations for purification of recP1 4x601 and recP5 5x601. First lane in each gel corresponds to the input for each round* ***A.*** *First round of PEG precipitation for recP1 4x601.* ***B.*** *Second round of PEG precipitation with input from first round yielding the pure recP1 4x601 fragment.* ***C.*** *First round of PEG precipitation for recP5 5x601.* ***D.*** *Second round of Peg precipitation yielding the pure recP5 5x601.*

In this way, the recombinant multi-NPS segments to be ligated were obtained in a yield of approximately 88% from having the plasmid (83pmol per digestion reaction), through multiple digestions, PEG purifications and a final PCR cleanup step (~65-73pmol isolated).

4.4. Generation of labeled single 601 sequences

Fluorescent labels were to be introduced by site-specifically labeling modified oligonucleotide primers and using those for PCR amplification of single 601 repeats with orthogonal non-palindromic overhangs. Oligonucleotides modified with a thymidine containing a carbon-amine linker arm at the desired sites were synthesized by IDT. NHS-ester (Alexa Fluor 568 and Alexa Fluor 647) or TFP-ester (Alexa Fluor 488) were then conjugated to these amine modified oligos with a standard reaction scheme for amines with NHS/TFP esters (**Figure 4.4.1**).

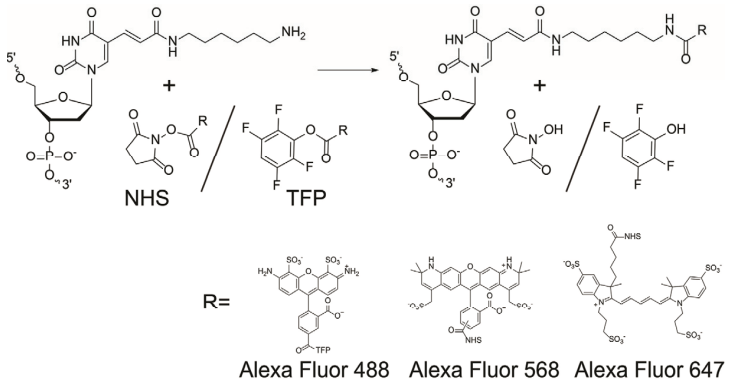

Figure 4.4.1. *Reaction scheme for labeling of thymidine with C6 amine linker to reactive NHS esters. R indicates the structures of the 3 fluorophores Alexa Fluor 488, Alexa Fluor 568 and Alexa Fluor 647 used for labeling.*

3 of the oligonucleotides were labeled with both Alexa488 (For MFD smFRET) and Alexa568 (For TIRF smFRET) as donors and 3 with Alexa647 as acceptor. To monitor labeling progression, purify the labeled oligo and to do the final analysis, oligos were run on neutral pH stable columns using triethylamine/acetic acid as buffer. The labeled oligos could therefore be obtained with high purity as shown by the final analysis (**Figure 4.4.2**).

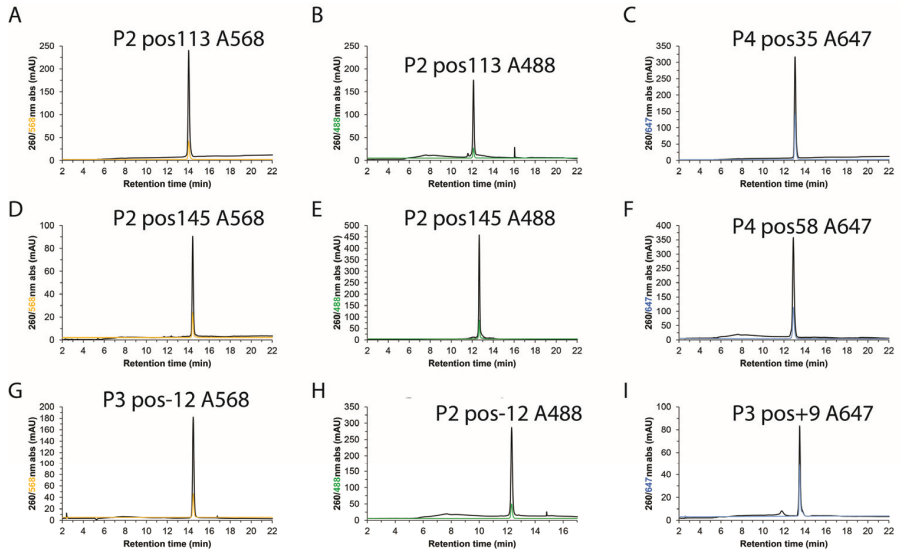

Figure 4.4.2. *Chromatograms from RP-HPLC analysis of final purified oligonucleotides.* ***A,B*** *and* ***C*** *the labeled oligos used for INFRET1.* ***D***, ***E*** *and* ***F*** *the labeled oligos used for IN$_{FRET}$2.* ***G***, ***H*** *and* ***I***, *the labeled oligos used for IN$_{FRET}$3.*

The labeled oligonucleotides were used together with unlabeled partners to produce labeled 1x601 pieces by PCR on a large scale (5mL) with Taq DNA polymerase[340]. In this way or with two unlabeled primers, the pieces listed in table 4.4.1 were generated.

Piece no	Labeling position(s)	Donor fluorophore	Acceptor fluorophore
P2	113	Alexa Fluor 488	-
P2	113	Alexa Fluor 568	-
P2	145	Alexa Fluor 488	-
P2	145	Alexa Fluor 568	-
P2	unlabeled		
P3	-12/+9	Alexa Fluor 488	Alexa Fluor 647
P3	-12/+9	Alexa Fluor 568	Alexa Fluor 647
P3	unlabeled		
P4	35		Alexa Fluor 647
P4	58		Alexa Fluor 647
P4	unlabeled		

Table 4.4.1. *Overview of 1x601 PCR pieces generated from labeled and unlabeled primers*

To generate the non-palindromic overhangs, components to be ligated were purified by standard PCR purification and buffer exchanged from elution buffer to water before digestion with BsaI and DraIII. Analysis of PCR piece digestions ascertained that the ends had been trimmed (**Figure 4.4.3**)

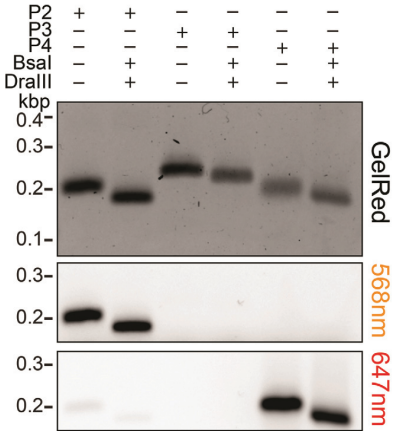

Figure 4.4.3. *Digestion of non-palindromic restriction sites in PCR pieces by BsaI and DraIII visualized by shift in size.*

Components recP1 4x601, P2 1x601 A568, P3 1x601, P4 1x601 A647 and recP5 5x601 for subsequent ligations were then ready.

4.5. Assembly of 12-mer DNA arrays from components

Considerations about ligation strategy

Convergent ligations of two 6x601 halves of the DNA arrays would provide significant advantages relative to sequential ligations. First, it would require less material since there would be less of a requirement in adding increasing amounts of each additional piece conjugated to ensure each reaction going to completion. Second, convergent assembly would disallow the formation of any constructs

with both fluorophores than the full-length 12x array. Finally, it would facilitate purification by PEG precipitation due to as big of a size difference as possible between the desired final product and the intermediates. Therefore the 6x intermediates (P1-3 and P4-5-anchor) were to be prepared prior to assembly of the full-length 12x601 (**Figure 4.5.1**)

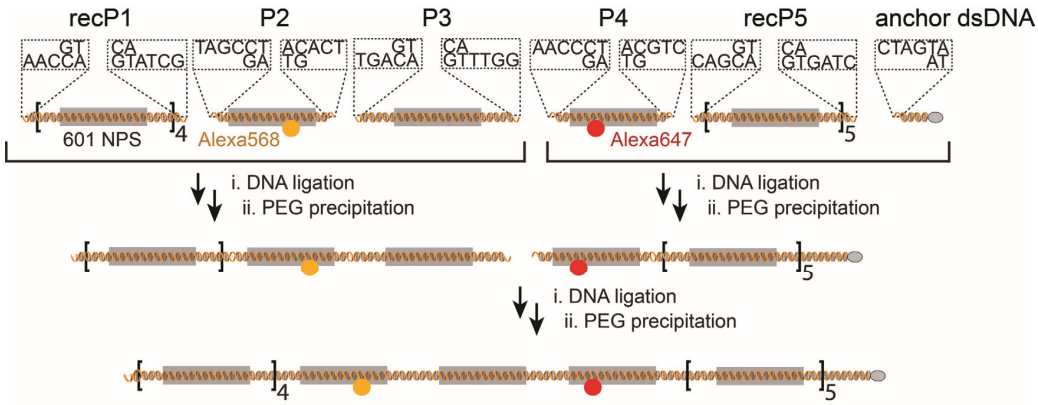

Figure 4.5.1. *Convergent assembly of DNA pieces. To facilitate assembly, purification and diversification, intermediate 6x601 pieces are first generated from P1-3 and P4-5-biotin anchor respectively, followed by assembly of the full-length 12x601 constructs*

Further, it would have the added benefit of being easier to diversify the full-length piece by ligating 6x intermediates containing libraries of different donor and acceptor positions respectively.

Test ligations to ensure full digestion

To ensure that complete or near-complete ligations could be carried out with the pieces prepared, test ligations were done for P1 with excess P2, P2 with excess P3, P3 with excess P4 and P5 with excess P4 (**Figure 4.5.2**).

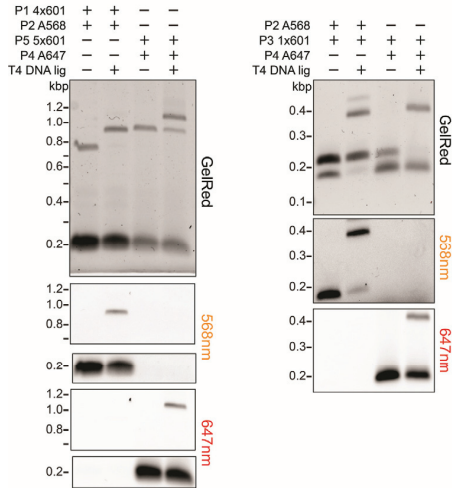

Figure 4.5.2. *Test ligations of key junctions for convergent assembly.* ***A.*** *Near-complete displacement of multi-601 recP1 and recP5 by ligation to excess partners.* ***B.*** *Displacement of P2 by P3 showing complete digestion of P2 on the DraIII junction and P3 by P4 showing complete digestion of the BsaI junction in P3.*

The complete displacement of the piece not in excess showed that this was indeed fully digested. Test ligations were done for each newly generated set of digested pieces to avoid proceeding to large-scale ligations with pieces that were not capable of fully ligating.

Ligations and purifications for 6x intermediates

When the test ligations had showed complete ligation, P1-P2-P3 and P4-P5-anchor were generated by first mixing two components and ligating those for 6-7h, followed by addition of the last component and ligation for another 10-14h. Analysis of samples from each step showed that the ligations were indeed close to completion (**Figure 4.5.3.A**).

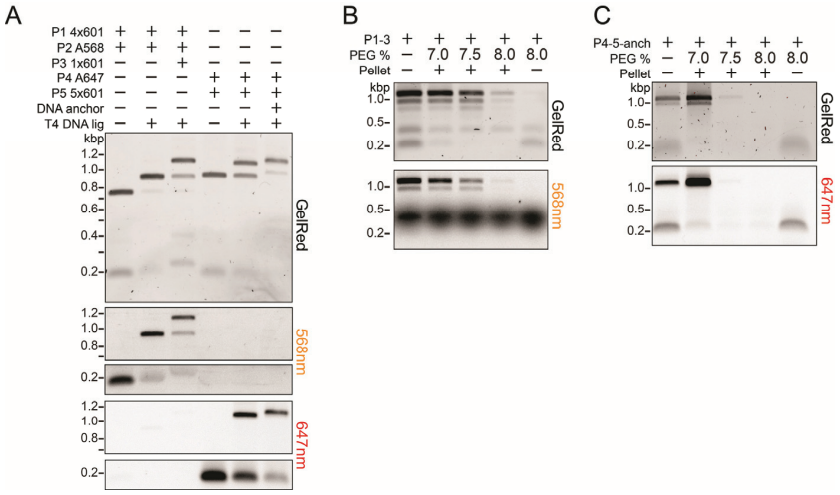

Figure 4.5.3. *Ligations and PEG purifications to generate 6x601 intermediates.* ***A.*** *Stepwise extension by ligation of P1-P2-P3 (first three lanes) and P4-P5-anchor (Last three lanes).* ***B.*** *PEG purification of P1-3 by stepwise increase of PEG concentration and centrifugations.* ***C.*** *PEG purification of P4-5-anch by stepwise increase of PEG concentration and centrifugations.*

Following the preparative scale ligation (~30-40pmol), the 6x intermediates could be effectively separated from the 1x601 pieces by PEG precipitation (**Figure 4.5.3.B and C**). Removal of those 1x601 repeats ensured that there would be no risk of bridging half-complete intermediates in the subsequent ligation. The residual 5x601 pieces appearing as bands under the 6x601 intermediates although they represent a clear loss in yield, they would not interfere with the assembly of the two 6x601 halves and because of the large size difference from the 12x601, they would be removed upon isolation of the final product by PEG precipitation at lower PEG concentrations.

Ligations and purification for 12x arrays

Two 6x601 segments were joined by ligation added slight excess of P4-5-anch (5-10%) to use as much of the P1-2-3 (More difficult to get a lot of) as possible. In this way, only relatively small amounts of the starting materials were left after ligation, and almost all of the functional intermediates were ligated to produce the 12x601 final product (**Figure 4.5.4.A**)

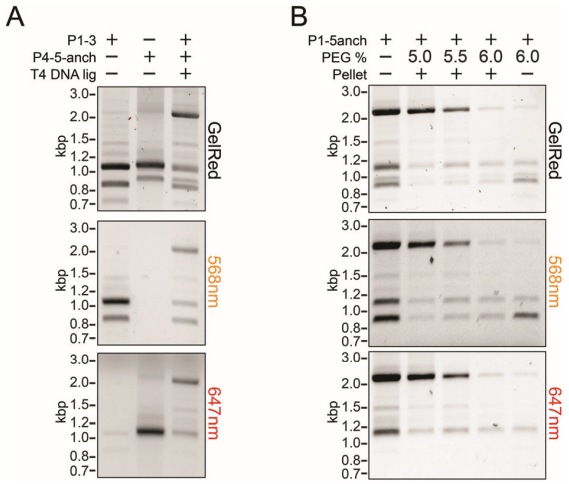

Figure 4.5.4. *Production and isolation of the final 12x601 double-labeled piece.* ***A.*** *The starting 6x601 intermediates with impurities (Lanes 1 and 2) and the two mixed and ligated (Lane 3).* ***B.*** *PEG purification of the 12x601 double-labeled DNA at concentrations from 5.0-6.0% PEG.*

After ligation, the product was isolated by PEG precipitation at lower PEG concentrations than used for the intermediates. In this way, there was almost complete separation from the intermediates. A final PCR purification step was done to remove residual PEG, ligase and other components from the preparation. Importantly, at the single molecule level we would only analyze components containing both the donor and the acceptor and the intermediates with P2 and P4 labeled would not contain both. In the case of $IN_{FRET}3$, where P3 contains both labels, this would not be anchored in smTIRF experiments, since the biotin anchor is on the other intermediate. Typically starting with about 35pmol of the recombinant pieces as the lowest amount used of the components for generating the intermediates, ligation with all of the generated 6x601 intermediates would result in about 12-14pmol of the final double-labeled 12x601 DNA (~30-40% yield). This amounts to about 20-22µg. For each chromatin assembly for measurements at the single-molecule level, typically 0.65pmol (~900µg) DNA was used. For ensemble experiments about 2pmol was needed. Therefore a preparation of the DNA would be enough for 20-25 chromatin assemblies for single-molecule experiments. Several different constructs were prepared for the different measurements, as summarized in table 4.5.1.

Construct	Donor	Acceptor
$IN_{FRET}\underline{1}$	Alexa Fluor 488	Alexa Fluor 647
$IN_{FRET}\underline{1}$	Alexa Fluor 568	Alexa Fluor 647
$IN_{FRET}\underline{1}$	Alexa Fluor 488	unlabeled
$IN_{FRET}\underline{1}$	Alexa Fluor 568	unlabeled
$IN_{FRET}\underline{1}$	unlabeled	Alexa Fluor 647

Construct	Donor	Acceptor
$IN_{FRET}\underline{2}$	Alexa Fluor 488	Alexa Fluor 647
$IN_{FRET}\underline{2}$	Alexa Fluor 568	Alexa Fluor 647
$IN_{FRET}\underline{2}$	Alexa Fluor 488	Alexa Fluor 647
$IN_{FRET}\underline{3}$	Alexa Fluor 488	Alexa Fluor 647
$IN_{FRET}\underline{3}$	Alexa Fluor 568	Alexa Fluor 647

Table 4.5.1. *Overview of constructs generated with different combinations of fluorophores*

The convergent assembly greatly facilitated the speed of generating all of these different constructs.

4.6. Chromatin assembly with double-labeled DNA arrays

Chromatin was assembled by salt dialysis from high (2M) to low (10mM) salt after mixing of double-labeled DNA, histone octamers and MMTV DNA in the same way as for the end-labeled DNA in the previous project. For ensemble FRET experiments, the buffer DNA and the buffer nucleosomes needed to be removed, as they might interfere with chromatin compaction at the ensemble level. Therefore the chromatin was precipitated with magnesium (4mM), the supernatant removed and the chromatin redissolved in TE. To test saturation of the chromatin array, they were digested with ScaI similarly to the first project, with ScaI sites present at equal distances between the 12 nucleosomes. Samples from the DNA alone, the chromatin assembly before and after magnesium precipitation and after ScaI digestion were analyzed by gel (**Figure 4.6.1**).

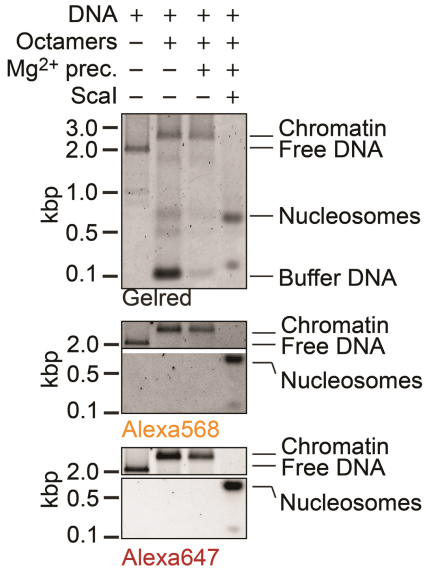

Figure 4.6.1. *Analysis of DNA, chromatin assembly before and after magnesium precipitation and following ScaI digestion*

The gel shows that the DNA shifts upwards upon assembly into chromatin due to the size increase (Lane 1 vs 2). Further, the assembly of chromatin yields nucleosomes on the buffer DNA, suggesting that the chromatin array DNA is fully saturated. Still, there is free DNA that could have buffered more histones. Upon precipitation, the chromatin is specifically recovered, whereas the buffer DNA and nucleosomes are removed. Digestion with ScaI shows freeing up of nucleosomes and only little DNA (Free DNA is stained more significantly than nucleosomal DNA). Together this shows that chromatin could be assembled on the double-labeled DNA, isolated from other assembly components and that the chromatin was nearly fully saturated.

4.7. Ensemble analysis of chromatin compaction

To probe whether FRET increased upon compaction, and how the different sets of positions were differentially sensitive to compaction, we carried out ensemble fluorescence measurements on the spectra encompassing the donor and the acceptor. Ensemble fluorescence measurements were carried out in 60µL volumes with sample from the chromatin preparation (55µL) mixed with magnesium and buffer to the final volume and concentration of magnesium desired. Concentrations of magnesium in the range 0-4mM results in chromatin compaction. The different sets of fluorophores, should provide unique vantage points for possible intermediate steps in this compaction. Compaction was assessed by measuring the spectra at 0, 0.5, 1.0 and 4.0mM for chromatin assembled with each of the 3 constructs (**Figure 4.7.1.A, E and I**). Similar was measured for DNA alone, to ensure that FRET would only result upon formation of chromatin (**Figure 4.7.1.C, G and K**) . The anisotropy of the donor was also measured in these different samples, to ensure that the fluorophore was not trapped in a position disallowing free rotation (**Figure 4.7.1. B, D, F, H, J and L**).

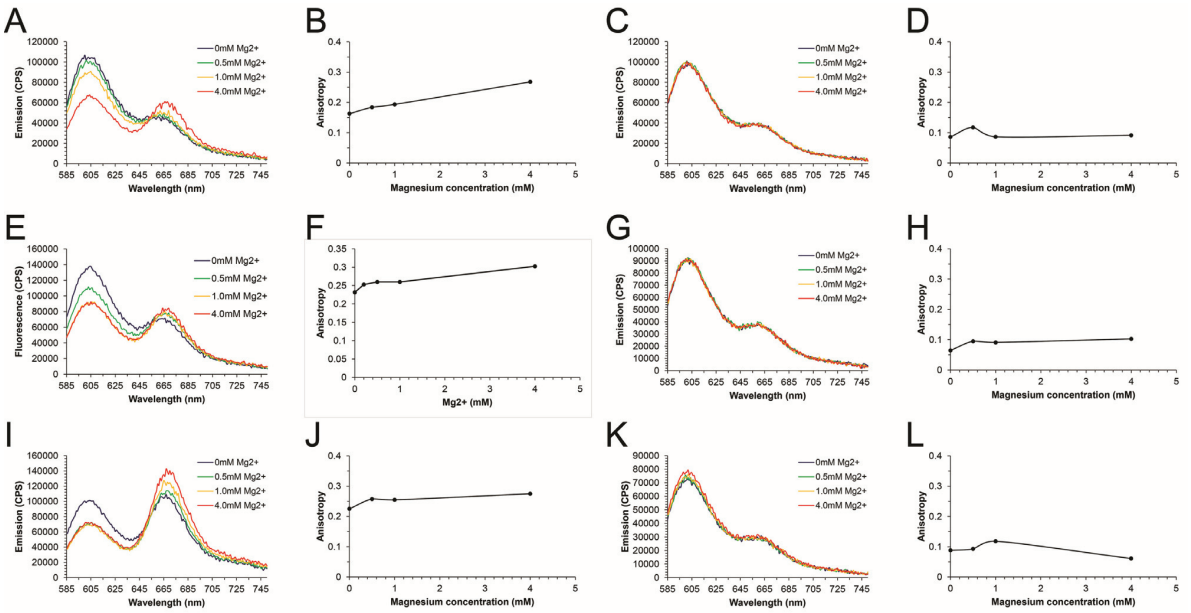

Figure 4.7.1. *Spectra and anisotropies from analysis of local conformational changes by ensemble FRET. **A** and **B**. Spectra and donor anisotropy for chromatin assembled on IN$_{FRET}$**1** at increasing Mg^{2+} concentration. **C** and **D**. Spectra and anisotropy for IN$_{FRET}$**1** DNA alone at increasing Mg^{2+} concentration. **E** and **F** Spectra and donor anisotropy for chromatin assembled on IN$_{FRET}$**2** at increasing Mg^{2+} concentration. **G** and **H**. Spectra and anisotropy for IN$_{FRET}$**2** DNA alone at increasing Mg^{2+} concentration. **I** and **J**. Spectra and donor anisotropy for chromatin assembled on IN$_{FRET}$**3** at increasing Mg^{2+} concentration. **K** and **L**. Spectra and anisotropy for IN$_{FRET}$**3** DNA alone at increasing Mg^{2+} concentration.*

The spectra consistently show, that in the absence of nucleosomes to position the fluorophores, they are not in a proximity to result in FRET and there is no change in the spectra upon increasing Mg^{2+}. The residual peak corresponding to acceptor emission is caused by some degree of direct excitation

of the acceptor at the wavelength of 575nm used for acquiring the spectra. The anisotropy consistently stays low in the DNA regardless of the magnesium concentration. The three different constructs in the context of chromatin result in changes in FRET at distinct points in the range of magnesium concentrations used. Based on the expected locations of the fluorophores and the ensemble spectra, overall conformational changes upon compaction can be outlined (**Figure 4.7.2**).

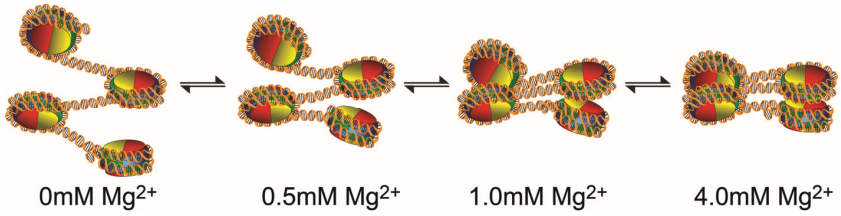

Figure 4.7.2. *Chromatin folding pathway based on ensemble FRET measurements. Involving first increased proximity in the linkers and near the dyad from 0-1mM Mg^{2+} followed by complete stacking at 4mM Mg^{2+}.*

From 0 to 0.5mM Mg^{2+}, the most increase in FRET occurs in $IN_{FRET}3$ and to some degree in 2, showing that a rearrangement in this region represent first steps in chromatin compaction. Beyond 0.5mM Mg^{2+} not much further change is observed with $IN_{FRET}3$. Up to 1.0mM Mg^{2+} additional increase is observed with $IN_{FRET}2$, suggesting that the dyad proximal region moves even closer together. Finally upon inceasing to 4.0 Mg^{2+}, the largest change is observed with INFRET1 showing that the nucleosomes only fully stack at relatively high concentrations of magnesium. The anisotropies for all three constructs slightly increase at increasing concentrations, showing that the rotational movement of the dyes become increasingly restricted upon chromatin compaction. Together, the multiple perspectives combined with ensemble FRET provide local information about the structural changes occurring upon compaction. The ensemble experiments however reveals little about the possible subpopulations that may be present and the possible kinetics of change between different states. This could however be done with smFRET measurements. For initial tests and characterization of photophysics, smFRET was done with double-labeled dsDNA.

4.8. Tests of smFRET with short double-labeled dsDNA

Measurements were carried out on a micro-mirror objective-type TIRF system. In this a small mirror reflects the incident laser beam into the objective at the angle necessary to result in TIRF[289]. Initial smFRET measurements were carried out with two pieces of dsDNA having each strand labeled with donor and acceptor respectively at two different distances (**Figure 4.8.1.A**). The individual labeled strands showed no FRET (**Figure 4.8.1.B**). The dsDNA spectra from ensemble measurements

showed two degrees of FRET, and their mixture an average of the two (**Figure 4.8.1.C**). The mixture was used to probe for photophysics of the FRET pair at the single-molecule level (Alexa Fluor 568 donor and Alexa Fluor 647 acceptor) and settings required on the microscope and in the buffer for acquisition of useful signals. Optimization of these conditions allowed the monitoring of FRET for durations of up to minutes using trolox and GODCAT for triplet state quenching and oxygen scavenging[309,310,341]. They further allowed us to build preliminary versions of the software necessary for data extraction and trace analysis (**Figure 4.8.1.D and C**).

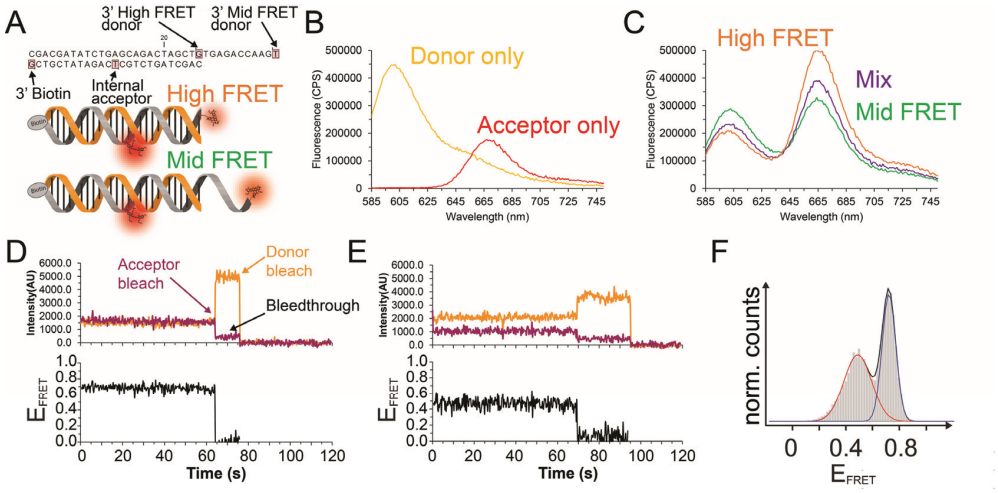

Figure 4.8.1. *smFRET with dsDNA in ensemble and single-molecule measurements.* ***A.*** *Scheme of dsDNA constructs for high FRET and mid FRET, with location of biotin anchor, acceptor and donors respectively.* ***B.*** *Spectra from donor and acceptor labeled oligonucleotides alone.* ***C.*** *Spectra from high FRET and mid FRET DNA as well as their mixture being an average of the two.* ***D.*** *Example smFRET trace from high FRET dsDNA highlighting the anticorrelated donor intensity increase upon acceptor bleaching, the bleedthrough from the donor to the acceptor channel and the bleaching of the donor.* ***E.*** *smFRET trace for mid FRET dsDNA.* ***F.*** *FRET efficiency histogram showing the distribution for the mixture.*

FRET efficiencies shown along each trace and used for subsequent plotting of histograms were calculated taking into account the donor bleedthrough to the acceptor channel β and the efficiency of detection of the donor versus the acceptor γ (Equation 4.4) [342]:

$$E_{FRET} = \frac{I_A - \beta I_D}{\left((I_A - \beta I_D) + \gamma I_D\right)} \qquad (4.4)$$

The values of β and γ respectively, were determined based on traces with photobleaching events with the anticorrelated increase in donor signal upon acceptor photobleaching (for γ) and the residual red emission signal after acceptor photobleaching (for β). With our set of filters and the used FRET pair, these were determined to $\beta = 0.141$ and $\gamma = 0.468$. Plotting of FRET efficiency histograms using more than 80 individual dsDNA molecules in the mixture, showed that we could efficiently distinguish between subpopulations in a mixture (**Figure 4.8.1.F**). The difference in width of the populations are

likely because one is double-stranded to the point of dye anchoring (High FRET). The other is extending as single stranded DNA beyond the annealed segment (Mid FRET), thereby showing potentially more flexibility. Based on the 24 and 13 bases separating the FRET pairs in each construct, this would be expected to result in FRET efficiencies of 0.51 and 0.97 respectively. The value observed for the Mid FRET construct is accurate, resulting in a center of the peak at 0.5. In the high FRET construct, the observed FRET efficiency is however lower, resulting in about 0.77, suggesting likely that orientation might bring them further apart than dictated by simple bp distance. Still however, the tests with dsDNA showed that we could observe smFRET and were able to distinguish between different populations in a heterogenous ensemble.

4.9. TIRF smFRET chromatin conformation upon compaction

To further optimize conditions, further measurements were done uniquely with chromatin arrays, since these were the target samples and examples in the literature show differences in photophysics depending on the biomolecular context, i.e. if the dyes are on DNA alone, proteins alone or in nucleoprotein complexes[311]. Initial experiments would be to establish that chromatin arrays could be immobilized, also resulted in inter-nucleosomal FRET at the single-molecule level and that the FRET efficiency would change upon compaction (**Figure 4.9.1.A**).

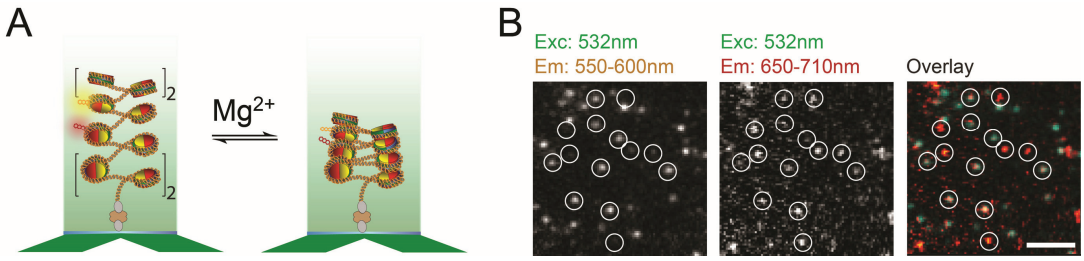

Figure 4.9.1. *Single-molecule monitoring of chromatin conformational dynamics.* ***A.*** *Scheme outlining the anchoring of the chromatin arrays on the coverslip surface and the conformational change between open and closed states which should result in increased FRET.* ***B.*** *Example image from single-molecule measurements of IN$_{FRET}$3 showing the colocalized emission from both donor and acceptor upon excitation of only the donor.*

Measurements with IN$_{FRET}$3 where the dyes were expected and shown to be in close proximity in the chromatin array (**Figure 4.1.4 and Figure 4.7.1**) showed clear emission from the acceptor upon donor excitation (**Figure 4.9.1.B**). This corroborated that our chromatin arrays could result in FRET at the single-molecule level. The duration that the fluorophores remained stable with GODCAT and trolox was relatively short and resulted in often rapid photobleaching. Therefore two additional TSQs 4-nitrobenzyl alcohol (NBA) and cyclooctatetraene (COT) were used. Together with trolox these have

been reported to synergistically improve dye photostability for nucleoprotein complexes in particular[311]. Indeed when these were used together, the photostability was enhanced, with excitation at laser powers needed to result in sufficient signal-to-noise at 100ms acquisition. To further prolong the time possible for acquisition in an area prior to bleaching, we included in the software for the microscope a scheme to program an acquisition sequence involving shifting to a new area with the stage and changing between the excitation lasers. Therefore, all subsequent acqusitions were done by moving to a new area with 647nm excitation, taking 10 images of acceptor emission followed by change to donor excitation and acquisition for 2-3min followed by another 40-100 images of the acceptor to correlate possible photobleaching behavior during the donor excitation with the absence/presence of the acceptor towards the end. In this way, the challenges with photostability were overcome making it feasible to carry out inter-nucleosomal smFRET with chromatin. smFRET was done at 100ms acquisition rate for all the three constructs at magnesium concentrations 0, 0.5, 1.0 and 4.0mM Mg^{2+}. Further, acetylation of histone H4 is expected to disrupt aspects of chromatin compaction. Therefore measurements were also carried out with the three constructs at 4mM Mg^{2+} using octamers harbouring a mimic of this modification prepared by cysteine alkylation[147]. In contrast to the ensemble experiments done at 10mM KCl, all single-molecule experiments were done at 40mM for the possible necessity of salt for HP1α to bind in later experiments. Experiments with mixed salt conditions have shown that chromatin is only slightly compacted at this salt concentration[343], and thus there should be a possibility to monitor further compaction.

Traces were acquired for unmodified histones at least and most magnesium (0mm and 4mM) and for H4Ks16ac at most magnesium (**Figure 4.9.2**).

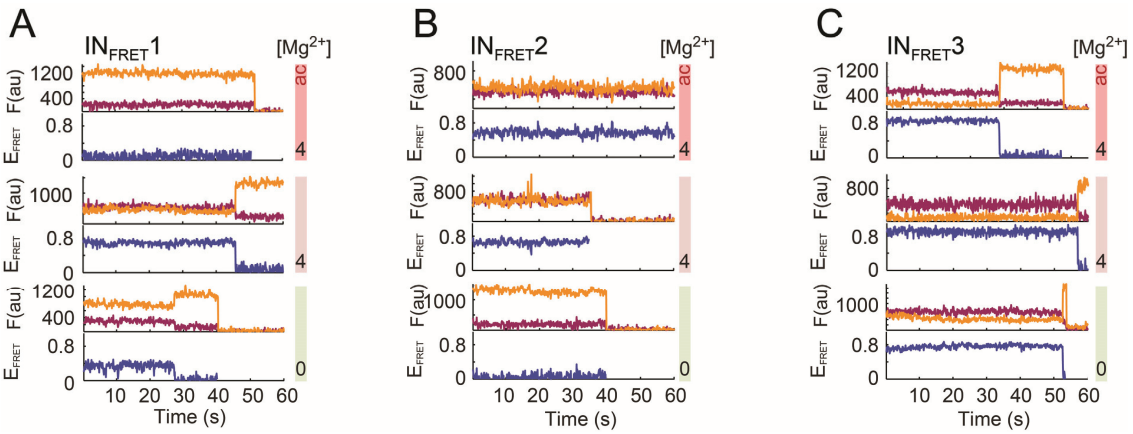

Figure 4.9.2. *Example traces from smFRET measurements with different constructs and magnesium concentrations.* ***A.*** *Example smFRET traces from IN$_{FRET}$1 at 0mM Mg2+ (bottom), 4mM Mg2+ and 4mM Mg2+ with H4K$_S$16ac octamers.* ***B.*** *Example smFRET traces for IN$_{FRET}$2 with similar conditions as in A.* ***C.*** *Example smFRET traces for IN$_{FRET}$3 at similar conditions as in A.*

The example traces show that increase of magnesium leads to compaction at the single-molecule level with all three of our constructs and acetylation of H4 mostly impacting the contacts reported on by IN$_{FRET}$1. This is more clearly seen from the histograms for the different conditions. More than hundred traces (See *n* for each condition in figure legend) were analyzed in this way and histograms of the FRET efficiencies plotted to reveal the population distribution and changes in these upon increase of magnesium concentration and in the presence of the acetyl modification (**Figure 4.9.3**)

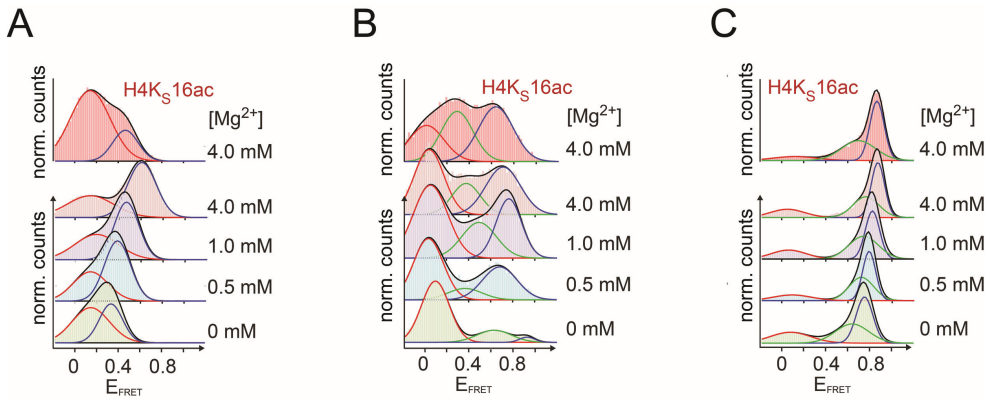

Figure 4.9.3. *FRET efficiency histograms from smFRET measurements using TIRF.* ***A.*** *Histograms for IN$_{FRET}$1 at 0mM (n = 448), 0.5mM (n = 205), 1.0mM (n = 156) and 4.0mM Mg^{2+} (n = 790) on unmodified arrays and with H4K$_S$16ac and 4mM Mg^{2+} (n = 146).* ***B.*** *Histograms for IN$_{FRET}$2 at 0mM (n = 174), 0.5mM (n = 133), 1.0mM (n = 103) and 4.0mM Mg^{2+} (n = 138) on unmodified arrays and with H4K$_S$16ac and 4mM Mg^{2+} (n = 113).* ***C.*** *Histograms for IN$_{FRET}$3 at 0mM (n = 228), 0.5mM (n = 165), 1.0mM (n = 155) and 4.0mM Mg^{2+} (n = 145) on unmodified arrays and with H4K$_S$16ac and 4mM Mg^{2+} (n = 130).*

IN$_{FRET}$1 shows a gradual increase towards a medium FRET efficiency state (~0.65), with a smaller population continuously present that does not result in significant FRET. The most pronounced increase occurs in the change from 1mM to 4mM magnesium in agreement with the ensemble measurements. Interestingly, the contacts reported on by IN$_{FRET}$1 are almost eliminated upon acetylation of H4 at 4mM magnesium, with most molecules appearing in a near-zero FRET state.

The distribution of FRET efficiencies with IN$_{FRET}$2 is relatively heterogenous, requiring 3 gaussians for fitting. This may likely reflect a lack of inter-nucleosomal anchoring points stabilizing these interactions like how the interaction between helices in adjacent nucleosomes likely stabilize populations observed with IN$_{FRET}$1.

IN$_{FRET}$3 shows a gradual change towards higher FRET efficiency, with one main population. This construct verifies that most of the DNA was indeed covered by nucleosomes, otherwise we would not see stable FRET efficiency with fluorophores that are about 200bp away from each. There is only a small population of zero-FRET molecules which may be DNA lacking nucleosomes around the

FRET pair. The structural rearrangements in the linker DNA do not seem to be adversely affected by the acetylation of H4K16.

At the 100ms time-scales measured here, there are no clearly defined fluctuations between distinct states that could justify dwell-time analysis with hidden Markov modeling. Instead fluctuations seem to be more rapid or not visible with this time-scale, however they are also relatively difficult to distinguish from noise because of the low γ value observed with our selected FRET pair (0.468) resulting in relatively small anticorrelated changes in acceptor intensity upon changes in proximity of the FRET pair. Instead we used cross-correlation analysis between the donor and the acceptor to visualize possible dynamics (**Figure 4.9.4**).

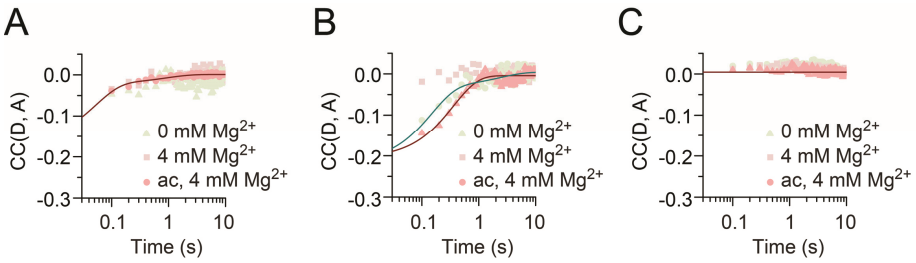

Figure 4.9.4. *Cross-correlation analysis of accumulated traces from smFRET with TIRF for different constructs.* ***A.*** *Cross-correlation analysis between donor and acceptor for IN$_{FRET}$1 at 0mM Mg2+, 4mM Mg2+ and 4mM with H4KS16ac.* ***B.*** *Cross-correlation as in A, but for IN$_{FRET}$2.* ***C.*** *Cross-correlation analysis as in A, but for IN$_{FRET}$3.*

The cross-correlation shows that there are short conformational fluctuations observed for IN$_{FRET}$1 and IN$_{FRET}$2, whereas the linkers monitored with IN$_{FRET}$3 stay in stable conformations on the time-scales investigated here. The rate of fluctuations do not change significantly for IN$_{FRET}$1 under different compactional states. For IN$_{FRET}$2 the compacted state displays increased dynamics regardless of the acetylation state, whereas the decompacted state does not.

4.10. Real-time monitoring of chromatin conformational changes

The immobilization of chromatin molecules in our flow chambers for TIRF smFRET would potentially allow us to change the buffer conditions and monitor induced conformational changes in real-time as has previously been done with the ribosome and DNA curtain stretching and cohesin movement on such curtains and their dependence on salt[330,344,345]. To test this idea, we added a functionality to the interface for control of our microscope for timed triggering of a pump connected to our flow chamber to change buffer conditions. With this added functionality, we set up an acquisition program where the pump would start 5s after starting image acquisition. Injection of 4mM

Mg^{2+} into chambers containing $IN_{FRET}1$ chromatin arrays resulted in a marked change in FRET efficiency for a large number of the arrays (**Figure 4.10.1.A**). Heat map plotting of the FRET efficiencies for more than hundreds of traces showed that the conformational change from decompacted to fully compacted occurred in less than 0.5s (Likely limited mostly by injection rate) (**Figure 4.10.1.B**)

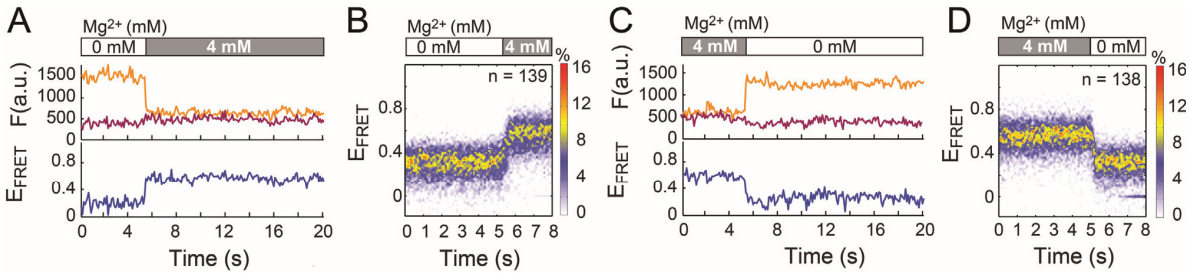

Figure 4.10.1. *Real-time monitoring of chromatin conformational change by change of buffer.* ***A.*** *Example trace of the conformational change upon change from 0mM to 4mM* Mg^{2+} *for* $IN_{FRET}1$. ***B.*** *Heat map plot for multiple chromatin arrays displaying the conformational change after 5s.* ***C.*** *Example trace of the change from 4mM to 0mM at 5s.* ***D.*** *Heat map plot for multiple chromatin arrays displaying the conformational change upon decompaction.*

Importantly, chromatin conformational changes are expected to be reversible. Therefore to completely remove magnesium from the flow channel, we tried timed injection of 5mM EDTA. This resulted in a corresponding decrease in FRET efficiency observed in the time traces (**Figure 4.10.1.C**). The corresponding heat map displayed that decompaction occurred at a similar rate as compaction (**Figure 4.10.1.D**). Together this shows that the immobilized chromatin arrays can be used to monitor conformational changes induced by buffers and possibly also proteins in real-time at the single-molecule level. Also, it shows that conformational changes can occur at sub-second rates. Since we observed little intrinsic dynamics at the 100ms time-resolution of our TIRF experiments, we continued to smFRET in solution.

4.11. smFRET in solution by multi-parameter fluorescence detection

To obtain information of the equilibrium dynamics on time-scales that were not accessible in our TIRF set-up, we carried out smFRET with a confocal setup coupled to MFD in solution. PIE was done to ensure that only measurements with both the donor and acceptor were used for analysis. These measurements were done with the constructs containing the fluorophores Alexa Fluor 488 and Alexa Fluor 647 to avoid the challenges associated with spectral cross-talk from donor emission in the acceptor channel. Possible inter-fiber oligomerization in the experiments are limited due to the low picomolar concentrations needed for the experiments, and in case of their occurrence, these can be

sorted by subsequent analysis of donor/acceptor stoichiometry. The output of the experiments are burst-wise recordings resulting from the excitation of a molecule passing through the confocal volume (**Figure 4.1.2.B**). The fluorescence life-time of the donor constitutes one dimension in the resulting plots. The intensities of the donor and acceptor for each burst are used to calculate the FRET efficiency as the second dimension (**Figure 4.11.1**).

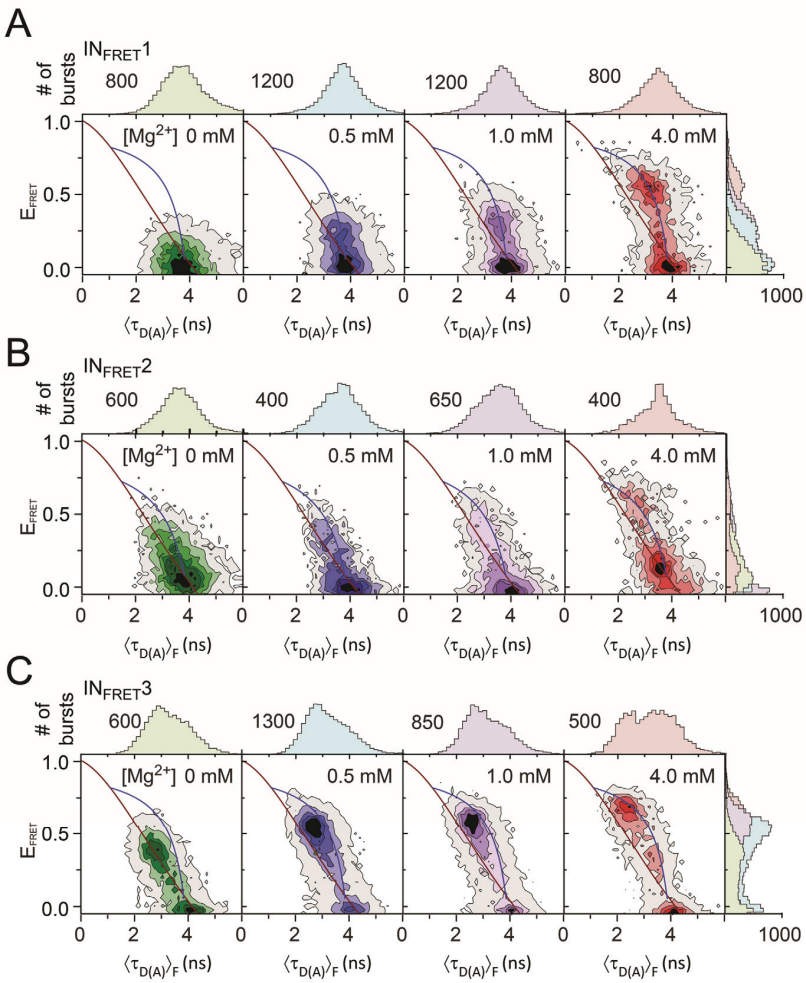

Figure 4.11.1. *Results from multiparameter fluorescence detection in solution. Contour plots of different constructs upon increasing magnesium concentration plotted as the FRET efficiency versus the donor fluorescence llife-time in nanoseconds. On each plot, the static FRET life is indicated in red and the dynamic FRET line in blue. Data contributed by the lab of Claus Seidel.*

In cases where intrinsic FRET motions occur faster than the dwell-time of the molecule in the confocal volume, this will give rise to populations deviating from the static FRET line otherwise dictated by the equations (Equations 4.2. and 4.3). This is due to averaging per species for the FRET efficiency and averaging per brightness for the donor life-time[253]. In the plots we almost consistently see that the FRET populations follow the dynamic FRET line showing dynamics faster than the dwell-

time of ~15ms[253]. For INFRET2, there is again a broad distribution of the conformations, whereas INFRET1 shows a gradual change towards higher FRET with a remaining low FRET population. INFRET3 results in increased FRET efficiency for nearly a single population. Together these measurements align with the population distributions seen in ensemble and smTIRF experiments, but also informs us that intrinsic dynamics take place at time-scales below the observed passage time through the confocal volume of ~15-20ms. Based on INFRET1 data, we carry out photon distribution analysis to estimate the populations and rates of interconversions between different states, using the distance constraints given from the FRET measurements[346]. The corresponding structural states were based on accessible volume calculations[347], indicating that nucleosome stacking in a tetranucleosome would result in a FRET distance of ~63Å, whereas those between tetranucleosomal twists (I.e between adjacent tetranucleosome stacks, see figure) would have a distance of ~44Å (**Figure 4.11.2.A**).

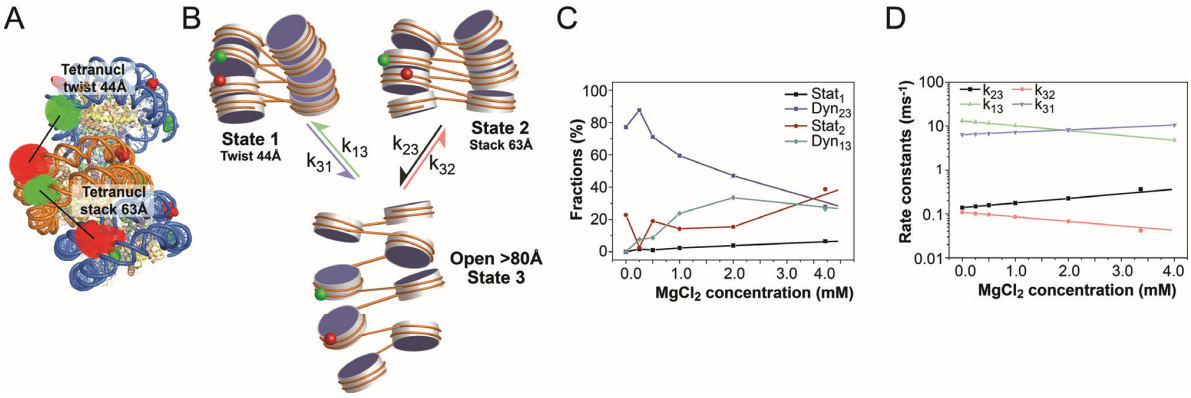

Figure 4.11.2. *Conformational states and kinetic modeling based on INFRET1.* ***A.*** *Accessible dye volumes indicated for the acceptor and the donor, upon positioning in a tetranucleosomal stack (63Å) or a twist (44Å).* ***B.*** *Model used for global fit to structural states based on MFD measurements with state 1 representing the tetranucleosomal twist, state 2 the tetranucleosome stack, and state 3 an open state and their rates of interconversion.* ***C.*** *Population distribution of static and dynamic states in the global fit* ***D.*** *Rate constants for conversion between structural states. Fitting contributed by the lab of Claus Seidel.*

The data could be fitted to a model with three states corresponding to the two described above, as well as an open one (**Figure 4.11.2.B**). Additionally, the compacted states could occupy either a dynamic state on the time-scales of the experiment or a static state, with possible slower movements than the travel time through the confocal volume (>15-20ms). Finally the logarithm of the rate constants were constrained to have linear dependence on the magnesium concentration. Based on this, the model showed that the states mainly responsible for chromatin compaction are static intra-tetranucleosome stacks, and dynamic inter-tetranucleosome stacks (**Figure 4.11.2.C**). Correspondingly, the global fit shows that the transitions to and from tetranucleosome stacks occur

at ~0.1ms^{-1}, whereas the inter-tetranucleosomal stacks convert with the open form at a faster rate of ~10ms^{-1}(**Figure 4.11.2.D**).

Calculations of the dye accessible volumes for IN$_{FRET}$2, showed that the dyes would be similarly close as for IN$_{FRET}$1 in inter-tetranucleosome twists (~46Å). Within a tetranucleosome they would be closer than for IN$_{FRET}$1 (~57Å) (**Figure 4.11.3.A**). Fitting of the MFD results for this set of positions resulted in a model with distances matching those seen in the structure, however with an additional state at a distance of ~72Å (**Figure 4.11.3.B**). This state we attribute to torsional movement near the nucleosome dyads, while maintaining contact near the positions of IN$_{FRET}$1, where helix interactions contribute to stabilization (**Figure 4.11.3.C**). Linker histone H1 could stabilize the conformational fluctuations at this site.

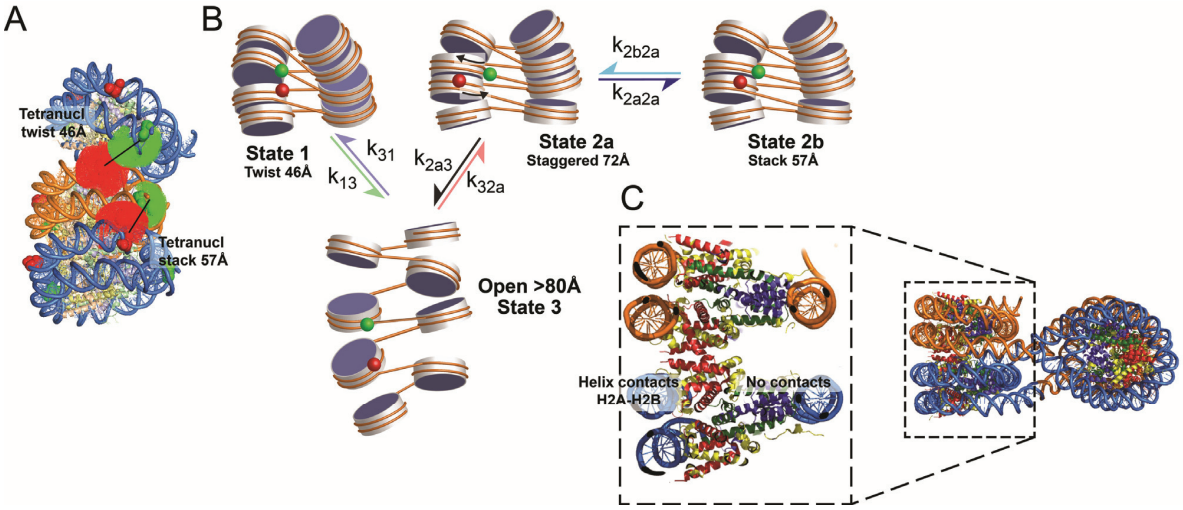

Figure 4.11.3. *Conformational states and kinetics modelling based on IN$_{FRET}$2. A. Accessible dye volume calculations based on positions within a tetranucleosome stack (57Å) or between tetranucleosome stacks (46Å).* ***B.*** *States modeled based on FRET distances including those modeled based on INFRET1 and a further partioning of state 2 into two different torsional states resulting in distances of 57Å and 72Å.* ***C.*** *Display of the stacking between H2A and H2B helices on adjacent nucleosomes, stabilizing INFRET1, while lack of stabilizing near the dyad allows torsional movement. Fitting contributed by the lab of Claus Seidel.*

The fits were further split into static and dynamic states as done for IN$_{FRET}$1 (**Figure 4.11.4.A and B**). These show again that the states that get increasingly populated upon compaction are the static state and to a lesser extent state 1, whereas the dynamic states decrease.

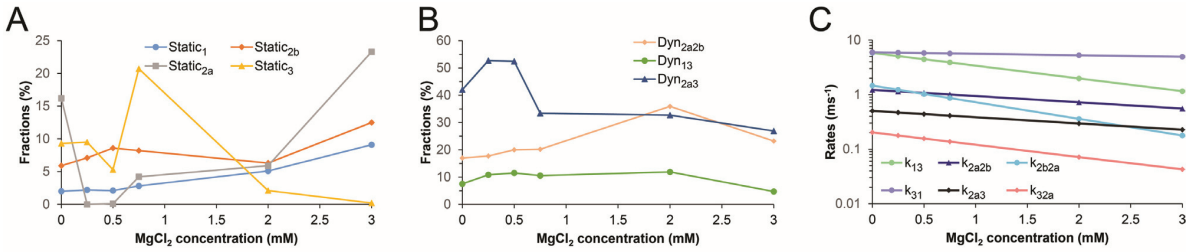

Figure 4.11.4. *Results from global fits of INFRET2 MFD data.* ***A.*** *Population of static states 1, 2a, 2b and 3 showing increased static compacted states and decreased static state 3.* ***B.*** *Populations of dynamics states interconversions with changes mainly occurring in the fraction of molecules dynamically changing between different torsional states.* ***C.*** *Rate constants for conversion between structural states. Fitting contributed by the lab of Claus Seidel.*

Similarly to the results for $IN_{FRET}1$, the conversions between states 1 and 3 occur fast at rates of 4-8ms^{-1}, whereas the switch between the open and tetranucleosomal states occur at ~0.07-0.12ms^{-1} (**Figure 4.11.4.C**). The rotation resulting in the different torsional states 2a and 2b occur according to the fit at rates of ~1ms^{-1}.

Together this displays that chromatin fibers dynamically interconvert between different conformations at times of 0.1-10ms and that global compaction is driven by increased stability of tetranucleosomal stacks. The fits for the data with $IN_{FRET}3$ remain to be done. Effectors have previously been shown to affect chromatin conformations in static measurements, however not in a dynamic context. Therefore we continued to probe the influence of HP1a on the compaction and movements in the arrays.

4.12. HP1α induced conformational changes by TIRF smFRET

HP1α is localized in compact heterochromatic regions particularly at the centromeres and telomeres. HP1α binding to H3 K9me3 can induce chromatin compaction, however the local conformational changes, the degree of compaction and the underlying dynamics are not well understood. Therefore we wanted to use our compaction sensitive FRET arrays to investigate how HP1α results in chromatin compaction at the single-molecule level. For this, we assembled chromatin with histone octamers with or without the H3 K9me3 modification onto the $IN_{FRET}1$ and $IN_{FRET}2$ DNA. Arrays were immobilized in flow chambers and FRET monitored with 1µM HP1α (**Figure 4.12.1**).

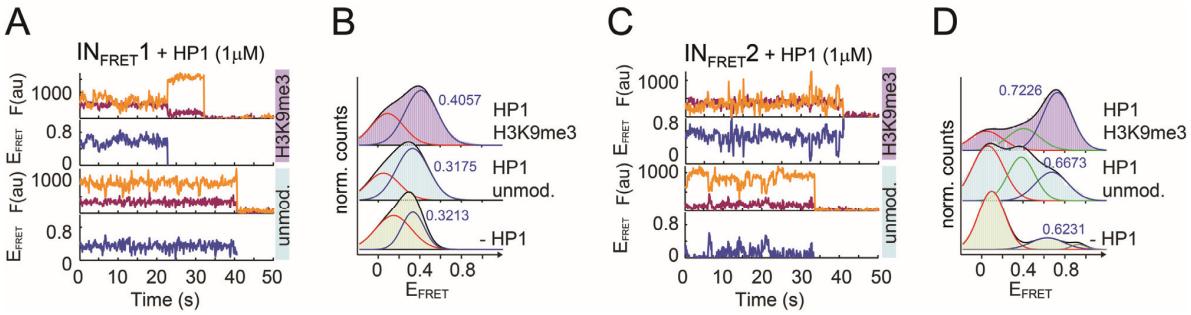

Figure 4.12.1. *Example traces and histograms from analysis of INFRET1 and INFRET2 containing unmodified and H3 K9me3 modified chromatin incubated with HP1α.*

The measurements show that HP1α without H3 K9me3 recognition results in slight compaction measured by both INFRET1 and INFRET2 and significant broadening of the distribution for INFRET2. The recognition of H3 K9me3 results particularly in nucleosome compaction in the linker proximal region as measured by INFRET2, whereas the effect on FRET efficiency with INFRET1 is relatively small (0.32 to 0.4). The traces also suggest that that HP1α promotes dynamic conformational fluctuations. When this was analyzed by cross-correlation between the donor and the acceptor, this seemed to be the case mainly for INFRET1(**Figure 4.12.2.A and B**).

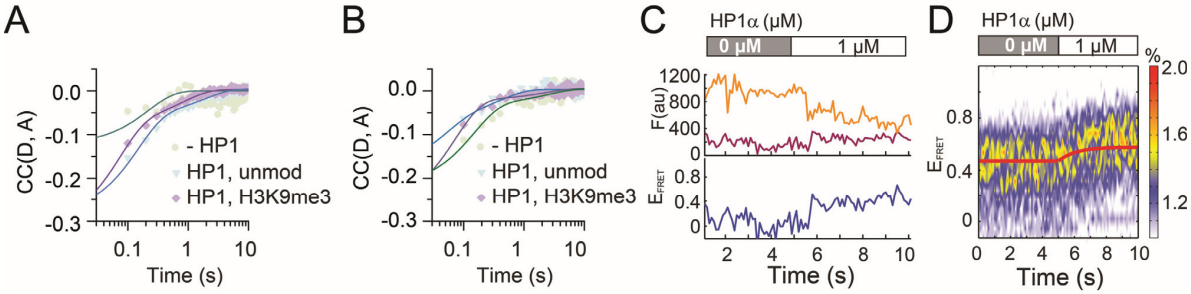

Figure 4.12.2. *Analysis of conformational dynamics in the presence of HP1α.* ***A.*** *Cross-correlation analysis of INFRET1 in the presence of 1μM HP1α, with or without H3 K9me3.* ***B.*** *Cross-correlation analysis of INFRET2 in the presence of 1μM HP1α with or without H3 K9me3.* ***C.*** *Example trace of injection of HP1α into flow-channel after 5s.* ***D.*** *Heat map plot of experiment with injection of HP1α.*

To possibly get at the rate of chromatin compaction monitored by INFRET2 upon HP1α incubation, we injected 1μM HP1α into a chamber with immobilized INFRET2 chromatin and monitored the conformation change in real-time (**Figure 4.12.2.C**). Plotting of this in a heat map for about 50 chromatin molecules, showed that there was a relatively large distribution both prior to and after injection (As seen previously) (**Figure 4.12.2.D**). Still however, we were able to see a clear change towards increased compaction. Fitting of these showed that it occurred at a rate of $0.86s^{-1}$. Together this shows that HP1α compacts chromatin toward an intermediately compacted state and that the stacking of nucleosomes in this state remains dynamic, with increased speed of fluctuations in the lateral stacking.

4.13. Discussion, conclusions and outlook

Here, we monitored chromatin inter-nucleosomal dynamics at the single-molecule level from multiple perspectives and on multiple time-scales. This allowed us to obtain detailed spatiotemporal information on the local conformational changes occurring upon chromatin folding, the rates and motions between nucleosomes, and how these are affected by the PTM H4K16ac and binding of HP1α respectively. (**Figure 4.13.1**).

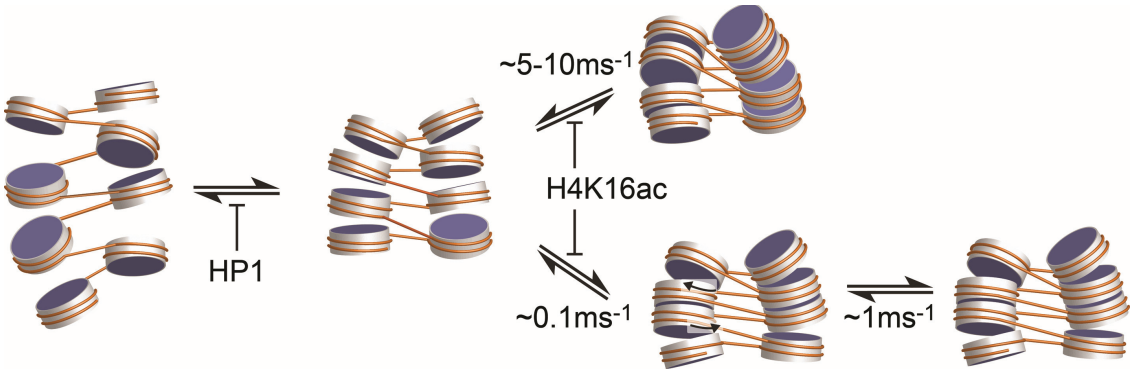

Figure 4.13.1. *Combined model based on ensemble, smTIRF and MFD measurements. Starting from an open state with linkers pointing in separate directions, the first step towards compaction involves rearrangement for the formation of a pre-tetranucleosome stacked state. This is stabilized by HP1, slowing down the reversion to the completely open state. From this state, compacted arrays can form in different registers. Tetranucleosome stacks form and open at rates of ~0.1ms^{-1} and undergo fluctuations between different staggered states at rates of ~1ms^{-1}. The twists between tetranucleosome stacks occur slower, but are key in gaining fully compacted chromatin. H4K16ac inhibits close-proximity of nucleosome stacks as seen both within tetranucleosomes and in the twist between adjacent tetranucleosomes.*

The results suggest an initial conformational change from an extended conformation to one prior to the formation of tetranucleosomes. This change we primarily see with IN$_{FRET}$3, in the ensemble experiments because of differences in the monovalent salt concentration in these experiments relative to all the single molecule experiments. The fitting of the IN$_{FRET}$3 MFD data will likely reveal further information about the interconversion between these open subpopulations. The action of HP1 seems to prevent reversion to the fully open form as evidenced from increased populations of the intermediately stacked and fully stacked conformations.

From the intermediately open state, different registers of tetranucleosomes can form for nucleosome N to stack with either nucleosome N+2 or N-2. Regardless of the register, H4K16ac reduces the ability to stably form these interactions. Stacking in tetranucleosomes further can result in staggered conformations, with an offset near the dyads interchanging at rates of ~1ms^{-1}. Stable tetranucleosome stacks appear to be the main driver of the global increase in chromatin compaction. Finally twists

between tetranucleosome stacks interchange at faster rates of 5-10ms^{-1} with the intermediately open state.

Combined, the results shed light on some of the apparent dichotomies on chromatin structure that have come up in the literature during the past decade. The dynamic interchange between different stacking registers explain why results from sequencing-based genome-wide chromatin conformation capture have seemed in discord with long stretches of stable 30nm fibers[68]. Still however, recent nucleosome-resolution conformation capture methods confirmed that nucleosomes do stack in two-start conformations in a cellular environment[64,78]. The measurements elucidate why 4-12 nucleosomes forming local clutches can be observed with high-resolution imaging methods in vivo, whereas longer stretches cannot[65,348].

It has also remained elusive how HP1 can promote compaction, while remaining dynamically associated with chromatin[70,135,254,255,285,349]. The results here suggest that the compacted chromatin state imparted by HP1 remains highly dynamic with conformational fluctuations on time-scales of 100-500ms, similar to the residence times we typically found in our CoSM assay. HP1 only seems to stabilize an intermediately compacted state such that the chromatin can be further directed towards full compaction or open states depending on the PTM landscape and auxiliary chromatin factors. This potential for dynamic interchange between different structural states endows chromatin with the plasticity required for regulating access for various DNA-templated processes.

It was previously observed that some of the early steps towards chromatin compaction had the most profound effect on global structure[69]. The local information provided by the FRET pairs here suggest that this is caused by rearrangement in the linker DNA interconnecting the nucleosomes.

In contrast to what many studies on single nucleosomes have found, we find little evidence suggesting spontaneous DNA unwrapping from either side of the nucleosome, unless the linker arms move synchronously to allow increased access on one side, while tightening on the opposite side. Recent hybrid smFRET/force spectroscopy experiments on single nucleosomes have suggested that unravelling of the DNA on one side simultaneously stabilizes the opposite side[320].

Biological outlook

The methods developed here will be useful in interrogating the detailed inter-nucleosomal dynamics of other chromatin-associated effectors and the unwrapping dynamics induced by binding. This would complement the plethora of studies investigating such dynamics in the context of individual

nucleosomes, where possible shielding effects and restricted movements in the context of chromatin may be insufficiently taken into account. The most obvious next process to look at, is the compaction of chromatin induced by histone H1. Histone H1 binds at the entry-exit position of the nucleosomes and thereby force them into more compact conformations, however also doing so in a dynamic manner[6,36]. Conversely, the binding of transcription factor recognition sites introduced through PCR of the middle pieces would likely result in chromatin decondensation, however there might be different effects dependent on the positions examined. Chromatin remodelling factors and histone chaperones are known to result in changes in mononucleosome DNA wrapping, therefore for example Nap1, FACT and ACF and their effect on chromatin conformation could also be interesting to study. Homologous recombination is another process that requires access to chromatinized DNA. Although it is known that Rad51 ssDNA filaments are able to bind targets within chromatin, the resulting effect on chromatin unwrapping is not known[350].

During the experiments, a simple modification to the ligation scheme was tested to allow labeling of nucleosomes neighbouring in sequence for possible N/N+1 contacts. This could be done by simply changing P3 with a synthetic annealed phosphorylated dsDNA linker DNA piece with overhangs matching those of P2 and P4. So far however no sets of dye positions resulted in FRET in the context of chromatin for this labeling configuration, and it would be interesting to systematically explore this to gain another dimension to the conformational changes.

Availability of recombinant pieces with different linker lengths than 177bp could extend the methodology to see the dynamics in chromatin arrays with longer links. Towards this we designed and cloned initial 197bp 1x601 and 2x601 constructs of recP1, recP3 and recP5 that would be immediately compatible with the PCR pieces already available (Linker DNA is encompassed in the cloned fragments). These are being extended to 3x, 4x, 5x and possibly further by a colleague.

Technical outlook

Other donors than Alexa Fluor 568 should be tested with the TIRF setup, since the detection efficiency of the two fluorophores in our measurements is quite different, whereas Cy3/Cy5 or Alexa Fluor 555/Alexa Fluor 647 most frequently used in the literature typically have the same detection efficiency. This would manifest itself as changes in the donor intensity resulting in equally sized anticorrelated changes in the acceptor intensity when the dyes are within distance for FRET and thus

more ease of resolving conformational changes from signal noise. These however have a lower R_0 that limit the distances that could be studied.

Usage of prism-type TIRF instead of objective-type TIRF would likely also benefit the measurements by improving signal-to-noise ratios, particularly in the red channel. Although objective-type TIRF is easy to adjust, the presence of both excitation path and emission collection path through the objective is known to result in stray light that can interfere with signal collection and increase background noise[351-353]. Further, the borosilicate surface of typical glass coverslips also likely contribute to background relative to the optical purity of quartz slides used as the surface for evanescent wave generation in prism-type TIRF.

A higher degree of automation in the trace sorting and analysis would also greatly increase the throughput of the experiments, since this along with DNA preparation is the most time-consuming.

Finally, both the photostability and the ability to do the experiments by real-time monitoring of conformation upon change of conditions might benefit from a redesign of the inlets and outlets of the flow chambers. Closing the system to disallow buffers to mix with ambient air would likely reduce the amount of oxygen entering the flow channel thereby reducing the rate of bleaching. Further, this could allow stopped-flow type of injections to take place to ensure that changes in environment are synchronized for real-time conformational changes.

5. Semisynthetic double-strand break nucleosomes

5.1. Project introduction and outline

5.1.1. The DNA damage response and 53BP1

DNA stability is continuously compromised by light-exposure, chemical reactants and the molecular machineries involved in transcription and replication[354]. This can result in DNA lesions including DNA double-strand breaks (DSBs) requiring signaling through the DNA damage response (DDR) to ensure their timely repair and guard genome integrity[355]. DDR signaling in eukaryotes occurs in the context of the chromatin template. In the DDR, the choice of repair pathway between error-prone non-homologous end joining (NHEJ) or high-fidelity homologous recombination (HR) is in part mediated by the chromatin-effector protein p53 binding protein (53BP1)[356]. HR involves exonuclease-mediated DNA resection to result in single-stranded overhangs that through the action of ATP-dependent recombinases pair with and use the sister chromatid as repair template[357]. In contrast, NHEJ results in joining of blunt or micro-resected DNA ends and therefore has a higher risk of mutations or deletions[358]. 53BP1 functions in part by governing the recruitment of downstream factors to prevent end resection and thereby promote NHEJ, although it also has a role in facilitating HR[359]. Dimers of 53BP1 are recruited specifically to chromatin with a distinct pattern of histone PTMs including H2A.X S139ph, H4K20me2, and H2A(.X) K15ub (**Figure 5.1.1.A, B and C**) [131,360-362].

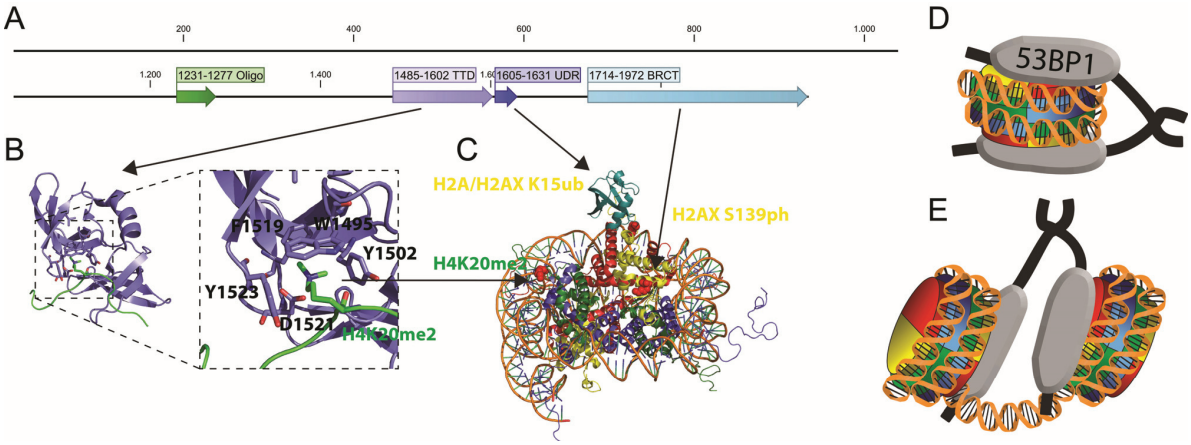

Figure 5.1.1. *Outline and histone targeting domains of 53BP1.* ***A.*** *Full-length 53BP1 encompasses 1972 residues, with long stretches of unknown function. Residues 1231-1277 allows 53BP1 dimerization to take place for effective 53BP1 recruitment to take place. The tandem tudor (TTD), ubiquitin dependent region (UDR) and BRCT targets 53BP1 to nucleosomes.* ***B.*** *The tandem tudor domain has an aromatic cage for H4K20me2 specific readout.* ***C.*** *The locations of the three target binding sites on the nucleosome H4K20me2, H2A(.X) K15ub and H2AX S139ph.* ***D.*** *Wheel-clamp model for 53BP1 binding to individual nucleosomes.* ***E.*** *Cross-linker model for 53BP1 binding to spatially proximal nucleosomes.*

Little function has been annotated to the first 1200 residues of 53BP1, other than numerous serines being the target for phosphorylation[363]. The tandem tudor domain (TTD, residues 1485-1602) binds to H4 K20me2 through an aromatic cage with an affinity of ~20μM (53μM to Kme1 and >1mM to me0 and me3)[364]. The BRCT domain (residues 1714-1972) binds S139 phosphorylated H2A.X with an affinity of ~3μM[131]. Although K15 ubiquitylation has a role in facilitating 53BP1 recruitment through the ubiquitin dependent region (UDR), the degree of affinity increase is not known[127]. Similar to HP1, 53BP1 depends on dimer-formation for effective binding, recognition of modified histone proteins for chromatin targeting and accumulates in macroscopically discernible foci at sites of DSBs with microscopic protein turnover[361,362].

The three histone modifications recognized by 53BP1 however give it even more potential anchoring points on the nucleosome than HP1. Minimally required for 53BP1 recruitment is the focus forming region encompassing the protein from the oligomerization domain until the end of the UDR[362]. A recent cryo-EM structure and biochemical studies shed light on the molecular details of 53BP1 binding to H4 K20me2 and H2A K15ub in the context of the nucleosome[127]. Although this minimal motif has been structurally characterized, little is known about how H2AX S139ph contributes to 53BP1 binding, how oligomerization contributes to recruitment and retention and whether 53BP1 dimers bind to two sides of the same nucleosome or are capable of bridging nucleosomes proximal in space (**Figure 5.1.1.D and E**). Further, 53BP1 binding could contribute to protection of DNA near DSBs from end resection by enveloping the nucleosome or may alternatively help prying chromatin open to allow access for down-stream repair factors. The lacking mechanistic insights into this are in part due to challenges in obtaining histone substrates for systematic investigations.

5.1.2. Multifragment ligations and synthetic schemes

We were interested in elucidating the binding mechanism and conformational dynamics imposed by 53BP1 upon binding using CoSM and smFRET as described in the two previous projects. Towards this however, suitable homogenously modified histones are required.

To our knowledge, there are no examples in the literature of H2A.X with modifications at either the C-terminal or N-terminal prepared by EPL. Instead histone PTMs for DSB signalling have been generated either enzymatically (E1-2-3 ligases, methyltransferases), through cysteine alkylation or di-sulfide directed ubiquitylation. Generation of H2A.X containing both the N-terminal ubiquitin and the C-terminal phosphoserine requires strategies for the modification of each site. Cysteine alkylation

does not easily lend itself to such challenges due to lack of orthogonal control and amber suppression mutagenesis often results in low yields even with single repurposed stop codons[365]. Site-specific introduction of modified or unnatural amino acids with NCL or EPL remains one of the most feasible choices to introduce C – or N-terminal modifications. Orthogonal protection/masking of the C-terminal thioester or the N-terminal cysteine or kinetic control of reaction rates allows multiple fragments to be ligated thereby enabling modification of both ends. N-terminal cysteines are typically protected with thiazolidine or acm[366,367]. C-terminal thioesters can be formed *in situ* from peptide hydrazides by oxidation with $NaNO_2$ and exchange with thiol catalysts allowing ligations in the N – to C-terminal direction[184,186,187]. Alternatively, synthesis and purification of thioester intermediates with faster ligation reactivity e.g. MPAA or TFET, followed by addition of exogenous thiol catalysts after the first reaction, allows three peptide segments to be ligated in a kinetically controlled fashion from the N – to C-terminal[217,368]. Ligations can therefore be done sequentially from one end to the other, or convergently by generating intermediates prior to assembly of the full protein[369]. Although these strategies are useful for convergent or sequential assembly, they are restricted to polypeptides generated by SPPS[369]. Often however, middle segments in such ligations do not contain non-proteogenic residues and do not strictly need to be produced by SPPS. Therefore similar orthogonal activation/deprotection schemes in the context of recombinant proteins would be highly useful. Proteogenic middle segments longer than 50-60 amino acid residues would particularly benefit from strategies to produce these recombinantly. We envisioned that SUMO and the split intein Npu could be used similarly as orthogonal recombinant protection groups/masked thioesters respectively.

Histone H4 containing K20me2 has previously been synthesized, however with position 29 used as ligation junction[118]. After ligation, the cysteine at the junction was left to result in a I29C mutation in the modified protein. Instead, synthesis of acetylated H4 have previously employed A38 as the ligation junction, to result in a tracelessly modified histone after ligation and desulfurization[69,248].

To overcome these shortcomings in histone accessibility, we developed synthetic schemes to prepare these using EPL and desulfurization (**Figure 5.1.2**). The H2A.X sequence provided useful ligation junctions between R20 and A21 for the N-terminal and between K134 and A135 at the C-terminal (**Figure 5.1.2.A**).

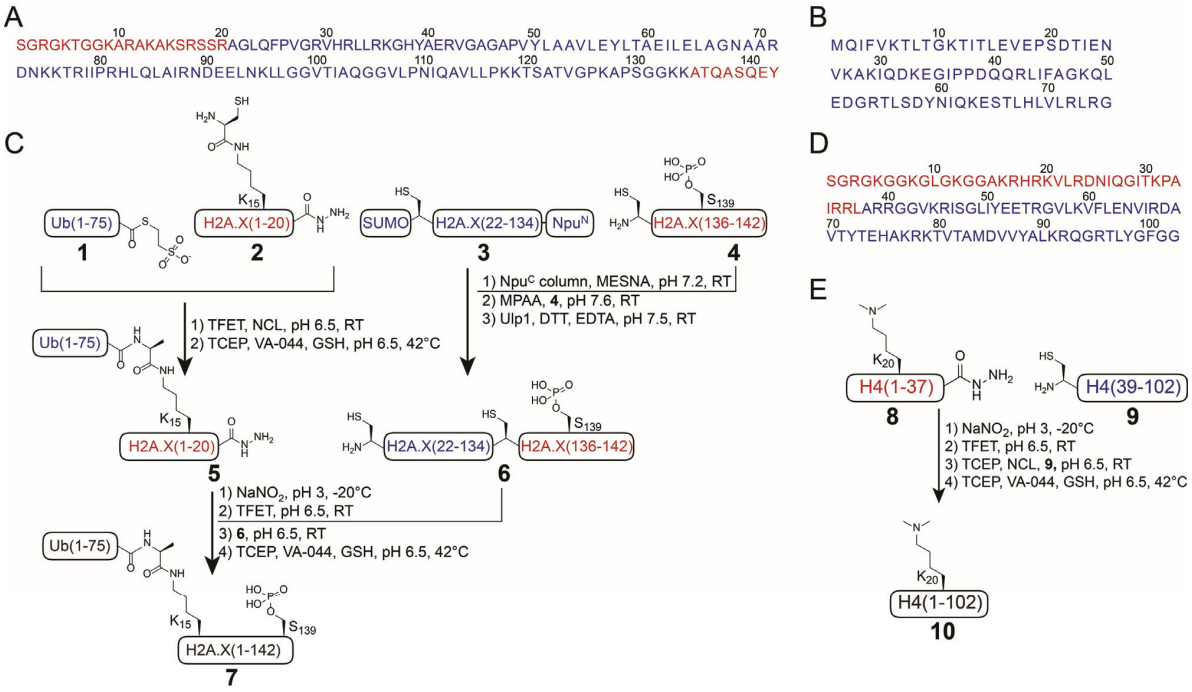

Figure 5.1.2. *Amino acid sequences and planned semisynthetic routes for the differently modified histones. Fragments generated by SPPS are in red and recombinant segments in blue.* ***A.*** *Amino acid sequence of histone H2A.X.* ***B.*** *Amino acid sequence ubiquitin without G76.* ***C.*** *Planned convergent semisynthesis of H2A.X modified at both ends with K15ub(G76A) and S139ph (7). This was to be done with recombinant UbΔ76-MES thioester (**1**), SPPS synthesized H2A.X 1-20 K15-ε–Cys (**2**), recombinant SUMO-H2A.X(Δ1-20,135-142, A21C)-NpuN (**3**) and SPPS synthesized H2A.X(135-142, C135A). Intermediate ubiquitylated N-terminal peptide (**5**) and the H2A.X core phosphorylated at the C-terminal (**6**) would be generated by ligations (and desulfurizations). These two intermediates would then be ligated and desulfurized to obtain 7.* ***D.*** *Amino acid sequence for histone H4.* ***E.*** *Synthetic scheme for generation of tracelessly modified H4 K20me2 (**10**) from H4 K20 1-37 peptide hydrazide (**8**) and truncated H4 Δ1-37 (**9**).*

H2A.X carrying both N-terminal ubiquitylation and C-terminal phosphorylation was synthesized convergently from two recombinant fragments (**1** and **3**) and two SPPS generated fragments (**2** and **4**). Initial ligations of **1** with **2** and **3** with **4** would yield the intermediates **5** and **6** respectively. Final ligation then resulted in the double-modified histone, followed by desulfurization to make it near-traceless with only G76A remaining from the synthesis (**7**). Histone H4 with H4K20me2 was synthesized from an SPPS-generated N-terminal 1-37 peptide (**8**) with the target modification ligated to a truncated H4 (**9**) followed by desulfurization to yield the tracelessly modified protein (**10**) (**Figure 5.1.2.E**). The steps necessary for generating the modified histones were:

- Synthesis of four peptides (**2a** (H2A), **2b** (H2A.X), **4** and **7**) by SPPS
- Cloning, expression and purification for recombinant pieces (**1**, **3**, **6**(no Sph), **3**(no ΔN))
- EPL and desulfurization for individual modifications in **7**
- EPL and desulfurization for **7** with both modifications
- EPL and desulfurization for **10**

5.2. Synthetic modified peptides

The three required modified N-terminal peptide histone tails were synthesized on 2-chlorotrityl hydrazine resins to produce peptides hydrazides after cleavage and the C-terminal tail on a pre-loaded Fmoc-Tyr Wang resin. The first amino acid was coupled manually, followed by automated SPPS. The ε,-ε-dimethylated Fmoc-lysine was synthesized manually using reductive amination with formaldehyde[370]. Fmoc-S(OPOBzlOH)-OH was used for phosphorylated serine. The ε-lysine to be coupled with cysteine was protected with Alloc, followed by on-resin deprotection and coupling with Boc-Cys(Trt)-OH. The peptides were all synthesized on a 0.1mmol scale by Carolin Lechner. The final analysis and yields of all four peptides were determined (**Figures 5.2.1-5.2.4**).

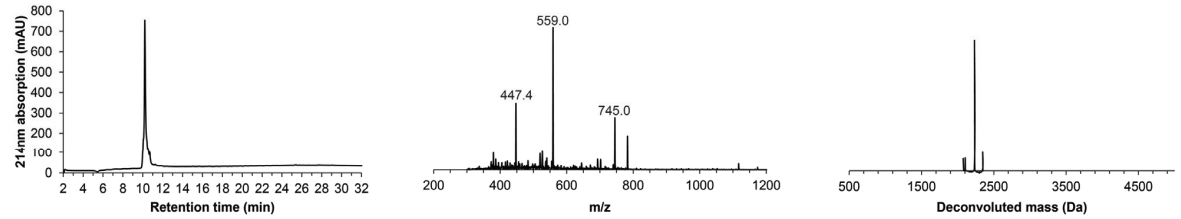

Figure 5.2.1. *Final analysis of H2A 1-20 K15ε-Cys hydrazide (**2a**). RP-HPLC and MS. M_{obs} =2232.0Da. M_{calc} =2232.5Da. Yield 9%*

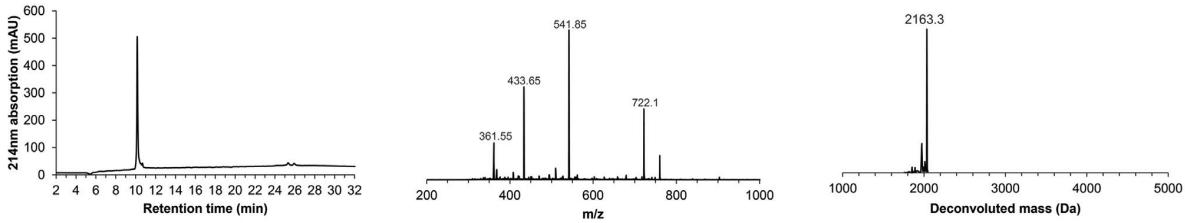

Figure 5.2.2. *Final analysis of H2A.X 1-20 K15ε-Cys hydrazide (2**b**). RP-HPLC and MS. M_{obs}=2163.3Da. M_{calc}=2163.5Da. Yield 8%*

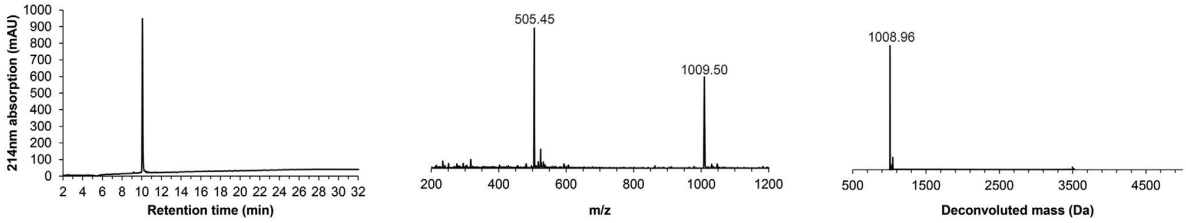

Figure 5.2.3. *Final analysis of H2A.X. 135-142 S139ph (**4**) by RP-HPLC and MS. M_{obs} = 1009.0Da. M_{calc} = 1008.3Da. Yield ~47%.*

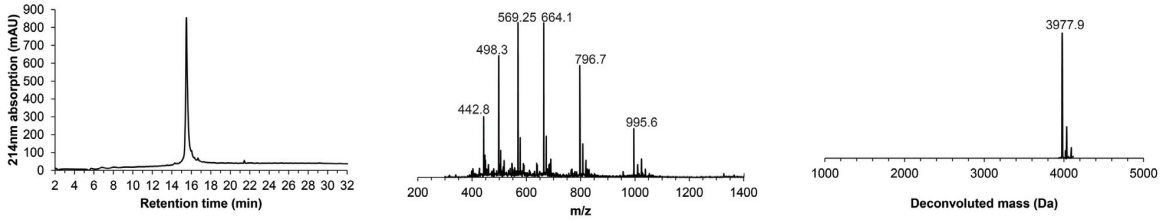

Figure 5.2.4. *Final analysis of H4 1-37 K20me2 hydrazide (**8**). RP-HPLC and MS. M_{obs} =3977.9Da. M_{calc} =3979.7Da. Yield 31%.*

5.3. Recombinant components

Constructs were cloned by overlap extension PCR (H2A.X wild-type), mutation (UbΔ75-NpuN-NpuC-6xH) or by Gibson assembly into linearized vectors for 6xH-SUMO and/or NpuN-6xH fusions (Four remaining constructs). They were expressed and purified by Ni-NTA followed either by MESNA thiolysis, Ulp1 cleavage to result in the free N-terminal cysteine and preparative RP-HPLC, or directly by RP-HPLC. The final analysis was done for all six components (Figures 5.3.1-5.3.6)

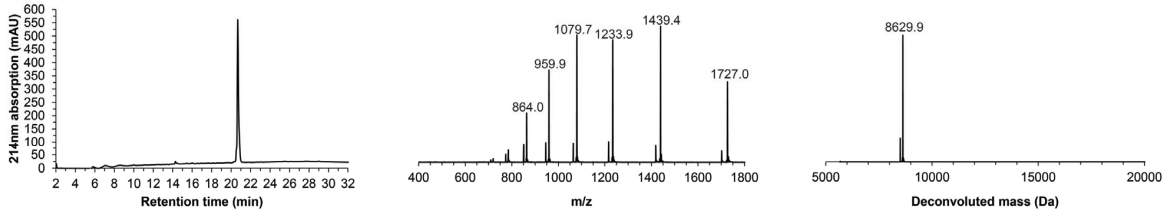

Figure 5.3.1. *Final analysis of Ub(Δ76)-MES (**1**) by RP-HPLC and MS. M_{obs} = 8629.9Da. M_{calc} = 8630.9Da. Yield ~50mg/6L.*

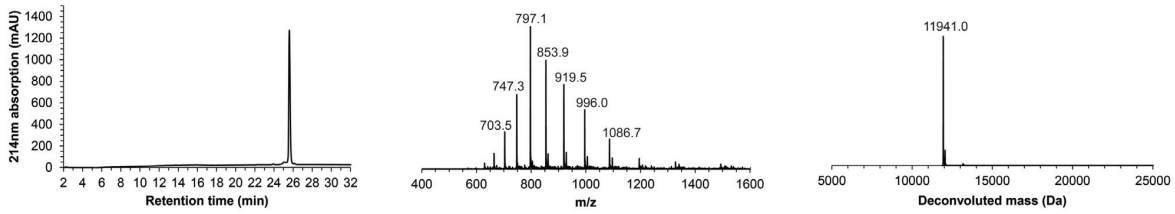

Figure 5.3.2. *Final analysis of H2A Δ1-20 A21C (**6 without S139ph**) by RP-HPLC and MS. M_{obs} = 11941.0Da. M_{calc} = 11941.0Da.*

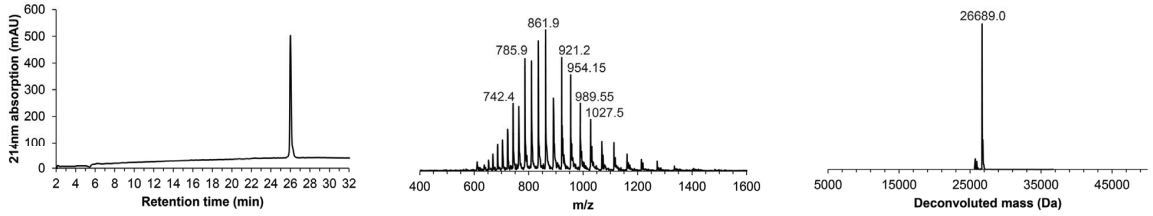

Figure 5.3.3. *Final analysis of H2A.X(Δ135-142)-Npu^N-6xH (**3 without ΔN**) by RP-HPLC and MS. M_{obs} = 26689.0Da. M_{calc} = 26681.6 Da.*

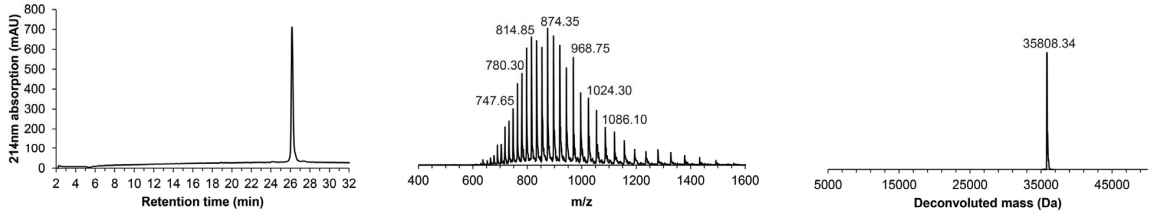

Figure 5.3.4. *Final analysis of SUMO-H2A.X(Δ 1-20, 135-142, A21C)-Npu^N-6xH (**3**) by RP-HPLC and MS. M_{obs} = 35808.3Da. M_{calc} = 35797.8Da.*

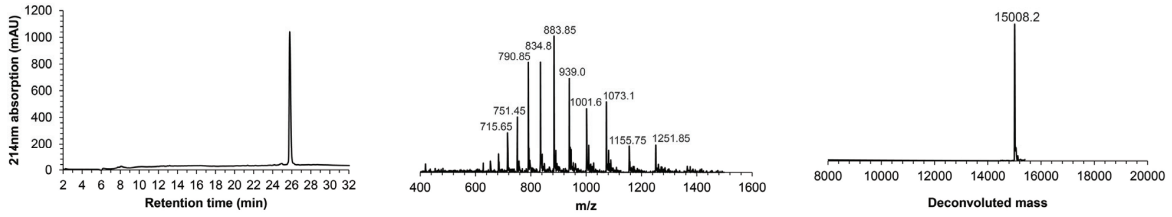

Figure 5.3.5. *Final analysis of H2A.X wild-type by RP-HPLC and MS. M_{obs} = 15008.2Da. M_{calc} = 15013.3Da.*

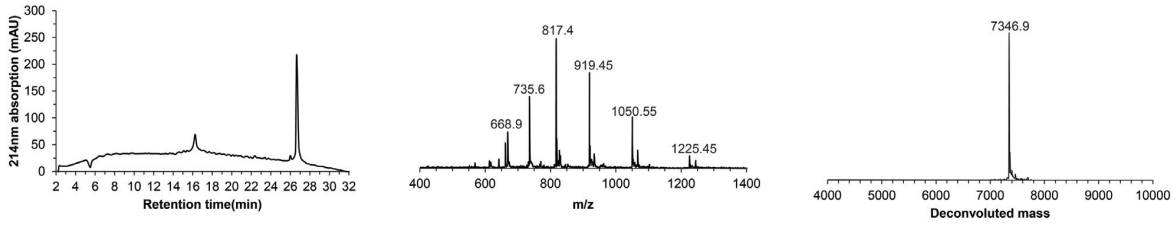

Figure 5.3.6. *Final analysis of H4 Δ1-37 A38 (**9**) by RP-HPLC and MS. M_{obs} = 7346.9Da. M_{calc} = 7348.5Da. Yield ~3mg/4L*

5.4. EPL and desulfurization for H2A K15ub(G76A)

Preparation of ubiquitylated N-terminal peptide without desulfurization

EPL strategies for ubiquitylated histones have focused on C-terminal ubiquitin modifications, using thiazolidine or penicillamine as N-terminal protection groups of peptides bearing the ubiquitin[114,246]. Likely the ubiquitylation of the histone N-terminal with orthogonal protection of the C-terminal would result in different synthetic challenges. The ligation reaction between **1** and **2** would likely be straightforward, as the ubiquitin was already a thioester. The reaction should therefore just require simple mixing of the components along with a thiol catalyst. For this, TFET was used. The reaction proceeded, however with significant hydrolysis within the 4h of reaction (**Figure 5.4.1**). Also, after 1h, more of the H2A peptide was added.

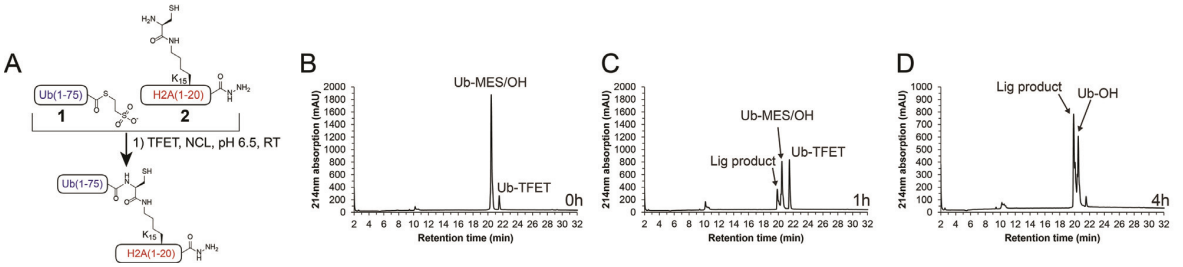

Figure 5.4.1. *Ligation reaction of Ub-MES with H2AK15εCys 1-20 peptide. **A.** Scheme of reaction between 1 and 2 for generation of ubiquitylated peptide **B.** Reaction immediately after mixing. **C**. Reaction after 1h and addition of more peptide. **D.** Reaction after 4h.*

After 4h, almost all of the TFET thioester had either ligated to form the product, or hydrolyzed thus being non-functional for ligation. Simultaneously, most of the peptide, which was the main limiting factor, seemed to have been used. Therefore the modified peptide was purified by RP-HPLC and the collected main fraction analyzed (**Figure 5.4.2**).

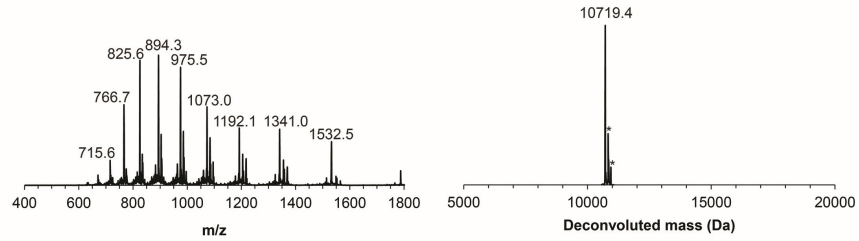

Figure 5.4.2. *Final analysis of H2AK15ub(G76A) MS. M_{obs} = 10719.4Da. M_{calc} = 10722.8Da.* in deconvoluted MS is TFA adducts*

The analysis shows, that the desired ubiquitylated peptide with the cysteine at the ligation junction remaining, was obtained in good purity.

Ligation and desulfurization for H2A K15ub (G76A)

For reaction to generate H2A K15ub(G76A), the peptide hydrazide was activated with sodium nitrite followed by addition of TFET for thioester formation and mixing with the truncated histone core. The reaction was monitored after 1h, 4h and 21h, however no significant change in the retention times of the peaks was seen (**Figure 5.4.3.A**)

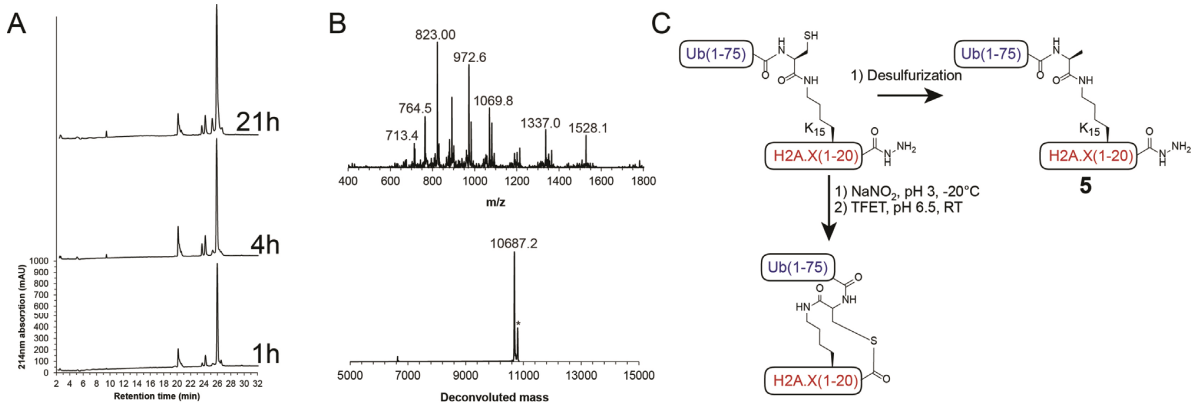

Figure 5.4.3. *Analysis of ligation reaction progress for ubiquitylated N-terminal peptide to the H2A core. A. Analytical RP-HPLC monitoring of reaction progress. B. MS showing the loss of mass corresponding to intra-molecular cyclization. M_{obs} = 10687.2Da. M_{calc} = 10690.7Da.* in deconvoluted MS is a TFA adduct. C. Scheme displaying the intra-molecular cyclization to form a thiolactone.*

When the peak for ubiquitin was analyzed by MS, the observed mass was found to be -32Da, which was smaller than the hydrolysis product of -14Da (**Figure 5.4.3.B**). Based on literature, this likely

corresponded to intra-molecular cyclization between the peptide C-terminal, and the cysteine thiol remaining at the ligation junction from the first ligation forming a thiolactone (**Figure 5.4.3.C**) [371,372]. Therefore, the cysteine needed to be removed to allow the ligation reaction to proceed. This was done by desulfurization to yield **5**. Ligation of this with the truncated H2A core proceeded such that a new peak resulted on the HPLC after 1h (**Figure 5.4.4.A**).

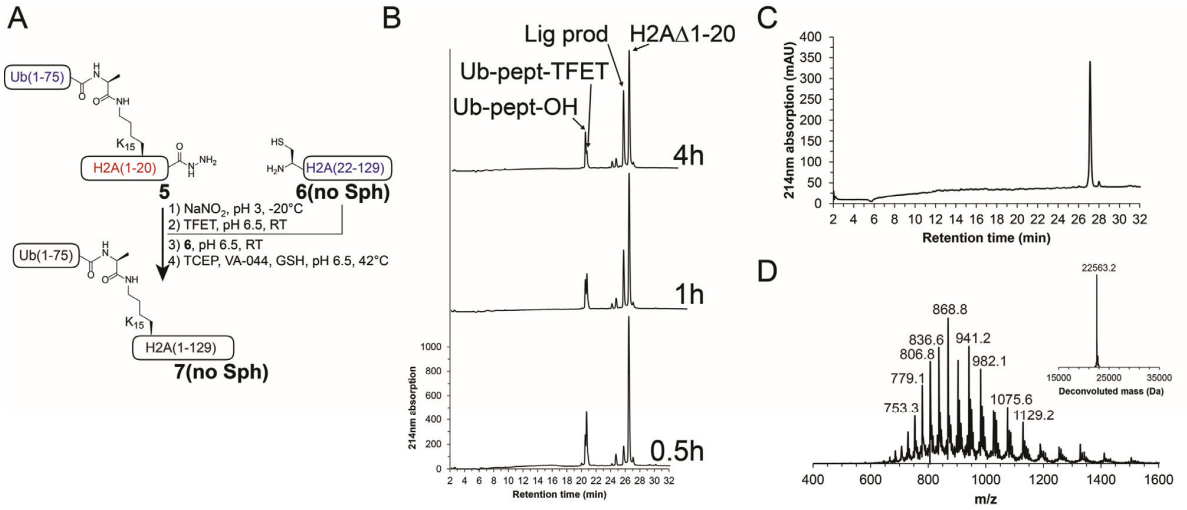

Figure 5.4.4. *Analysis of ligation reaction progress for ubiquitylated N-terminal peptide to the H2A core after desulfurization of the N-terminal.* ***A.*** *Scheme of reaction between ubiquitylated N-terminal peptide and H2A truncated core.* ***B***. *Analytical RP-HPLC monitoring of reaction progress.* ***C***. *RP-HPLC of the final purified product.* ***D***. *MS of the final purified product.* M_{obs} *= 22563.2Da.* M_{calc} *= 22567.0Da.*

The TFET thioester seemed to be hydrolyzed after 4h and therefore desulfurization was started in the same pot and allowed to proceed for 18h followed by purification and final analysis, showing that the desired product had been obtained (**Figure 5.4.4.B and C**), however in relatively low yield of only 15%. Improvements like usage of MTG as thiol catalyst, the timing of mixing with the truncated histone and adjustment of pH could however likely increase the yield.

5.5. EPL and desulfurization for H2A.X. S139ph

To produce the H2A.X thioester, the fusion with C-terminal Npu^N was incubated with the Npu^C column in the presence of 2M urea, leading to almost complete capture (**Figure 5.5.1**).

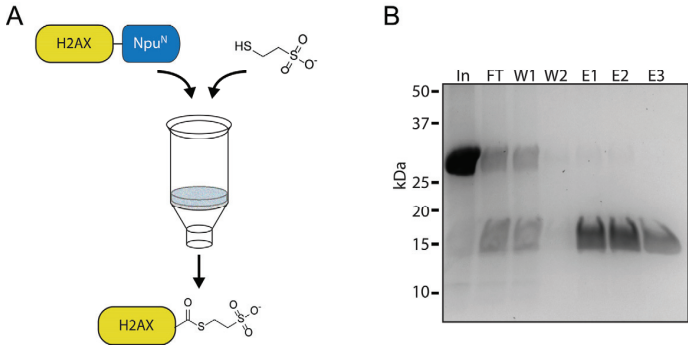

Figure 5.5.1. *Scheme and gel analysis of thioester formation with split intein column. A. Schematic of thioester formation with the intein column. B. SDS-PAGE analysis of input, flowthrough, washes and elution from the split intein column.*

Subsequent incubation for 14h with MESNA resulted in clean conversion to the corresponding thioester and this could be collected from the elution of the column (**Figure 5.5.1.B and Figure 5.5.2**)

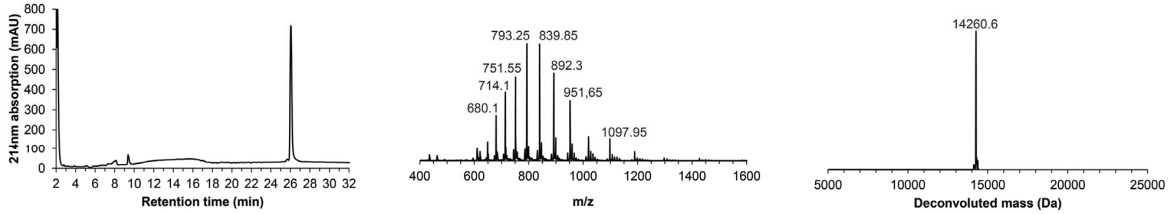

Figure 5.5.2. *Analysis of eluted H2A.XΔ135-142 MESNA thioester by RP-HPLC and MS. M_{obs} = 14260.6Da. M_{calc} = 14257.61Da.*

Addition of MPAA to 50mM along with the C-terminal phosphopeptide to the eluted protein without intermediate purification resulted in the ligation product within 2h (No change in HPLC retention time) for subsequent purification of the ligation product (**Figure 5.5.3**).

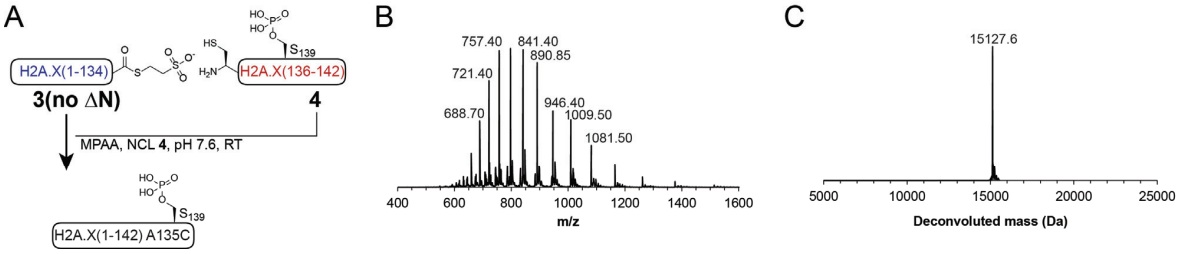

Figure 5.5.3. *Scheme for ligation and analysis of ligation product H2A.X S139ph A135C. **A.** Scheme for the ligation of the thioester with the C-terminal peptide. **B.** Analysis by ESI-MS. **C.** Deconvolution of ESI-MS. M_{obs} = 15127.6Da. M_{calc} = 15125.4Da.*

This showed, that the split intein column could be used for thioester formation followed directly by ligation between the eluted thioester and the target C-terminal peptide. Desulfurization for this is yet to be done.

5.6. EPL and desulfurization for H2A.X K15ub(G76A) S139ph

Semisynthesis of truncated H2A.X with C-terminal phosphorylation (9)

Towards the synthesis of H2A.X modified at both ends, we planned to combine the methods for the synthesis of the protein containing the individual modifications (**Figure 5.1.2**). This would require convergent assembly of first the N-terminal ubiquitylated peptide hydrazide **5** as well as the C-terminal phosphoprotein **6** with a deprotected N-terminal cysteine available for ligation. We prepared the ubiquitylated H2A.X peptide by ligation and desulfurization, similarly to what was described previously (5.5). For the C-terminal part **6**, we aimed to test if the number of steps carried out in one pot after elution of the thioester from the intein column could be extended to also include the Ulp1-mediated removal of SUMO after ligation of the C-terminal phosphopeptide (**Figure 5.1.2**). Initial thioester formation proceeded as smoothly as it had done for the protein without N-terminal SUMO. The eluted thioester peptide was highly pure, with a mass corresponding with the calculated (**Figure 5.6.1**).

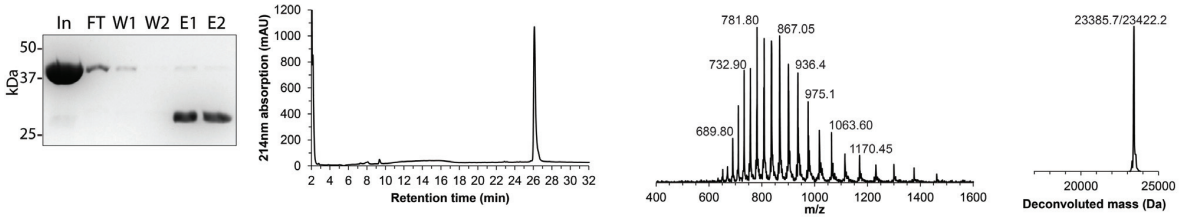

Figure 5.6.1. *Analysis of eluted SUMO-H2A.X (Δ1-20,135-142, A21C) by SDS-PAGE, RP-HPLC and MS. M_{obs} = 23385.7Da. M_{calc} = 23373.8Da.*

The additional peak with a larger mass (~37Da) is difficult to attribute to a specific adduct due to the relatively large inaccuracy of the determined MW in the first place. It is however likely that it is a result of the NaCl in the buffer, since the change in mass is too small to be MESNA, urea or phosphate, which are the only other compounds present. The splitting of peaks disappeared at a later stage (After SUMO cleavage). Addition of MPAA and the octapeptide resulted in rapid (~1h) conversion to the ligation product (**Figure 5.6.2**)

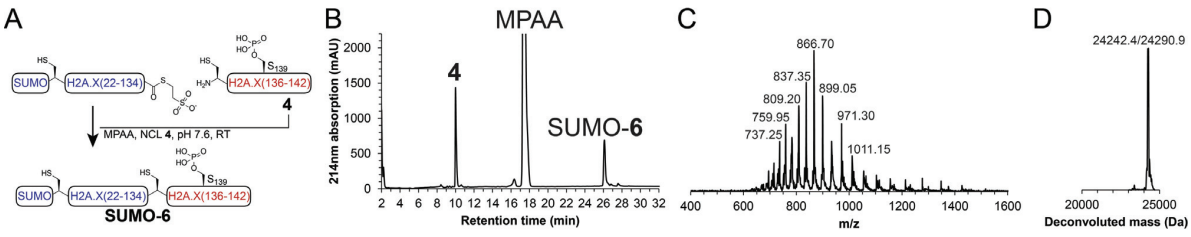

Figure 5.6.2. *Scheme and analysis of ligated SUMO-H2A.X (Δ1-20, A21,135C, S139ph).* ***A.*** *Scheme for ligation. B. RP-HPLC analysis of ligation after 1h.* ***C.*** *ESI- MS of ligation product.* ***D.*** *Deconvoluted mass of product. M_{obs} = 24242.4Da. M_{calc} = 24241.6Da.*

The analysis shows, that the desired modified peptide was generated. With the complete ligation at the C-terminal, the N-terminal was deprotected by addition of Ulp1, EDTA and DTT. The reaction was monitored after 5min, 2.5h and 12h (**Figure 5.6.3**).

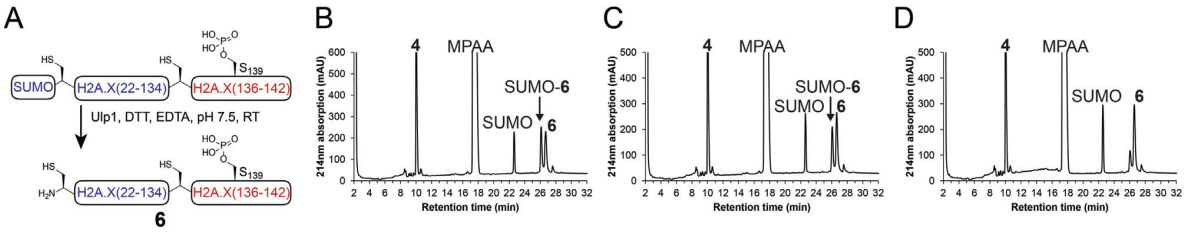

Figure 5.6.3. *Analysis of progress of enzymatic deprotection to yield H2A.X (Δ1-20, A21,135C, S139ph).* ***A.*** *Scheme of reaction.* ***B.*** *Reaction after 5min.* ***C.*** *Reaction after 2.5h.* ***D.*** *Reaction after 12h.*

The addition of Ulp1 (retention time 27.8min) resulted as hoped in N-terminal deprotection to yield **9** (retention time 26.8min), still in the same pot. After 12h, this was complete, and the intermediate was purified and analyzed (**Figure 5.6.4**).

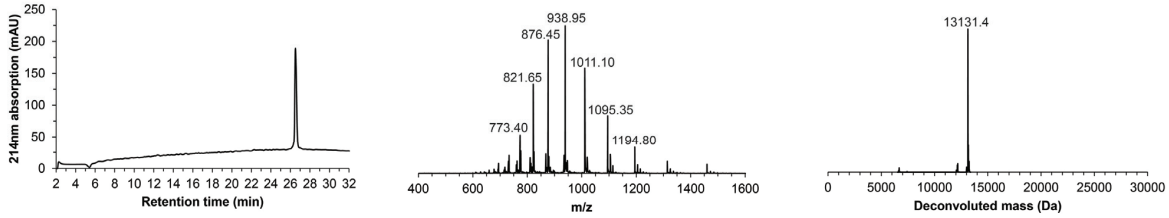

Figure 5.6.4. *Final analysis of H2A.X (Δ1-20, A21,135C, S139ph) (**6**) by RP-HPLC and MS.* M_{obs} *= 13131.4Da.* M_{calc} *= 13129.1Da*

The analysis showed the purity of the obtained fragment, and importantly only a single mass was observed corresponding with the expected size.

Semisynthesis of γH2A.X K15ub (10)

For the final assembly of **5** and **6**, **5** was activated with $NaNO_2$ and initially attempted converted to the MTG thioester, however too little MTG was added for this single attempt. Therefore the first amount of **5** used, had completely hydrolyzed after ~1h. Instead more of **5** was activated, converted to the TFET thioester and the reaction monitored, seeing that ligation took place (**Figure 5.6.5**).

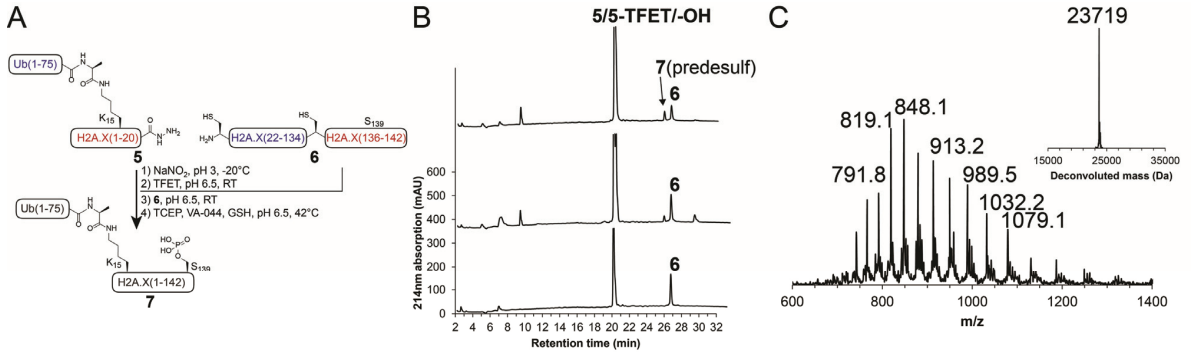

Figure 5.6.5. *Ligation of 8 and 9 and analysis of product.* ***A.*** *Scheme of ligation.* ***B.*** *Monitoring of the ligation reaction after 1, 2 and 5h.* ***C.*** *ESI-MS spectrum of ligation product and Deconvoluted mass of ligation product.* M_{obs} *= 23718.9Da.* M_{calc} *= 23718.3Da*

We accepted that the reaction would not proceed at this time further than the observed 30-40% and therefore proceeded to desulfurization in the same pot. This proceeded efficiently converting both of the residual cysteines into alanine within 6h followed by semiprep purification and final analysis (**Figure 5.6.6**)

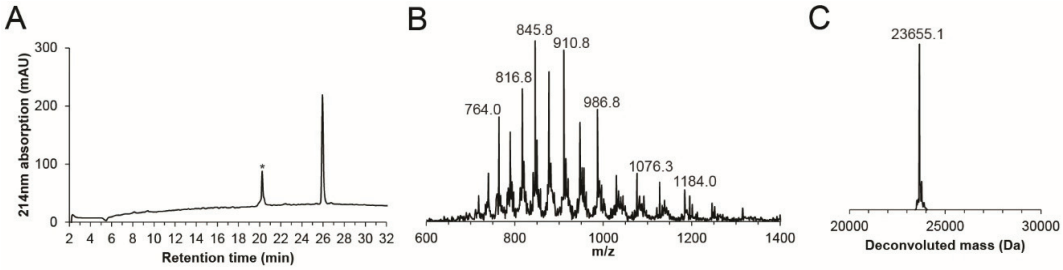

Figure 5.6.6. *Final analysis of* ***10****. A. RP-HPLC chromatogram from final analysis (* indicates peak that is likely a residue from large amounts of ubiquitin injected during monitoring of the ligation reaction, no trace of this is in the MS).* ***B.*** *ESI-MS spectrum of ligation product.* ***C.*** *Deconvoluted mass spec of ligation product.* M_{obs} *= 23655.1Da.* M_{calc} *= 23654.2Da*

This showed that the double-end modified H2A.X could be successfully synthesized. The isolated yield of **7** was 27% relative to the starting amount of **6**, and therefore still requires further optimization, particularly in getting the ligation reaction closer to completion.

5.7. EPL and desulfurization for H4 K20me2

The synthesized **8** and recombinant **9** fragments formed the basis for one-pot ligation with TFET as catalyst and desulfurization to generate the tracelessly modified histone according to the planned scheme (**Figure 5.7.1**).

Activation of the peptide hydrazide at low pH followed by incubation with TFET effectively resulted in the desired TFET thioester with a retention time of 18.8min (From 15.6min for the hydrazide)

(Figure 5.4.1.A). Mixing with the truncated H4 (Retention time 26min), resulted after 2h in more than half being converted to the desired ligation product (Retention time 24.7min) (Figure 5.4.1.B)

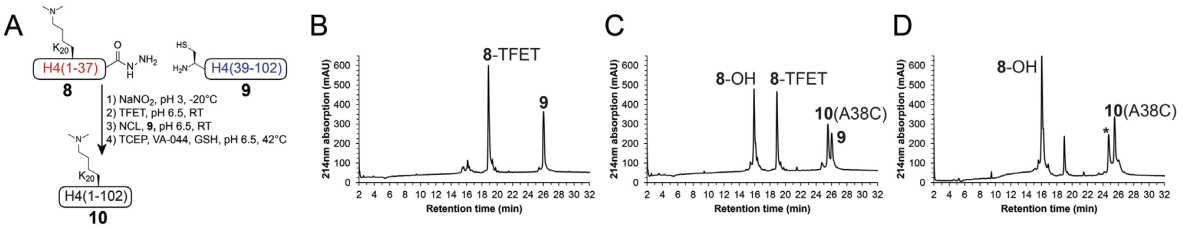

Figure 5.7.1. *Ligation progress for the generation of H4K20me2 A38C.* ***A.*** *Reaction after activation of N-terminal peptide and mixing with truncated protein.* ***B.*** *Reaction after 2h.* ***C.*** *Reaction after 16h progress. * indicates an unidentified side-product.*

Although the ligation product appeared after 2h, about half of the TFET thioester had also simultaneously hydrolyzed as evidenced from reversion back to the retention time at 15.6min. The single time that this experiment was done so far, the reaction was left to continue overnight and only monitored again after 16h. At this time, the starting material had been completely used, and the major peak was the desired ligation product (**Figure 5.7.1.C**). There was however an additional non-identified product appearing with a retention time shorter than that of the product. The peak could not be removed by reduction with TCEP, and could also not be identified from mass spectrometry. This calls for closer monitoring of the reaction next time it is done. After overnight reaction, GSH, additional TCEP and VA-044 were added for desulfurization in the same pot. 6h of radical-initiated reaction completely removed the thiol at the ligation junction and the desired modified protein was purified by semiprep followed by final analysis (**Figure 5.7.2**).

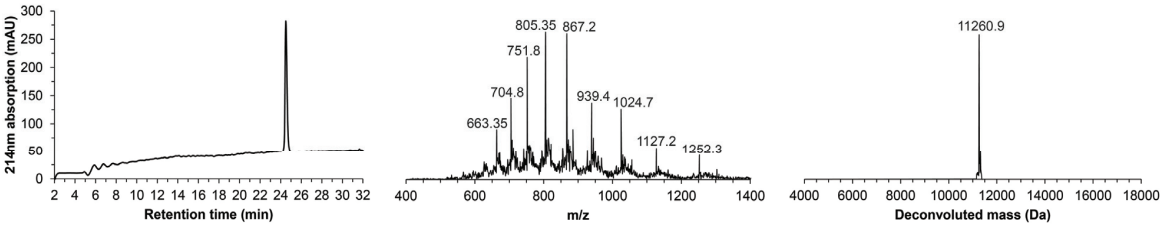

Figure 5.7.2. *Final analysis of H4K20me2 (**10**) by RP-HPLC and MS. M_{obs} = 11260.9Da. M_{calc} = 11264.2Da.*

The single peak, and the mass corresponding with the expected calculated, shows that the desired product could be successfully obtained. This was done with a yield of approximately 38%. This could however likely be increased, if the side product resulting from the ligation reaction was avoided.

5.8. Discussion, conclusion and outlook

The SPPS-synthesized and PTM-containing H4 and H2A(.X) N– and C-terminal peptides along with recombinant truncated proteins containing either N-terminal SUMO or C-terminal Npu formed the basis for the synthesis of H4K20me2, H2A.X S139ph, H2A K15ub and H2A.X containing both N-terminal ubiquitylation and C-terminal phosphorylation. The multivalent effector protein 53BP1 recognizes the PTM signature on these histones and resultingly contributes to the repair pathway chosen for DSBs. Although most of the synthetic routes were only tested once or twice, the experiments constitute proof-of-principle, that they are all possible and result in these non-optimized settings in yields of 15-45%. The often low ligation efficiencies point to issues with TFET ester hydrolysis. Instead likely MTG should be used. Most significantly the experiments demonstrate that SUMO and Npu^N respectively can function as orthogonal recombinant protection groups similarly to the synthetic thiazolidine/acetamidomethyl for N-terminal cysteines and C-terminal hydrazides respectively. This provides an alternative route to access relatively long proteinogenic middle pieces in ligation strategies, where such fragments would otherwise require assembly of at least 2 synthetic segments. This would likely facilitate synthetic access to other proteins including histones, that contain modifications at both ends e.g. H2B S14ph/K120ub.

Only recently were 53BP1 reader functionalities uncovered beyond the TTD for H4K20me2 and the oligomerization domain for dimerization[362,364]. An additional motif next to the TTD was identified as a UDR and the BRCT realized to bind to H2A.X S139ph[131,361]. Nucleosomes with H4K20me2 (as MLA) and enzymatically ubiquitylated H2A have provided structural insight into the TTD and UDR positioning on the nucleosome. Still however, the individual contributions of the PTMs to 53BP1 recruitment as well as their concerted effect lack detailed elucidation. Further, the structural impact on nucleosomes and the chromatin fiber of the PTMs might be influential in helping to provide access to DSB-affected chromatin regions, as both histone phosphorylation and ubiquitylation at different sites have previously been found to affect nucleosome stability and chromatin compaction respectively [69,104,373]. The material generated through the synthetic routes devised here along with the single-molecule based methods described in the two previous projects should prove useful in interrogating the detailed mechanisms of the PTMs in regulating chromatin structure and recruiting 53BP1.

6. Materials

Chemicals

Chemical	Provider
1-hydroxybenzotriazole hydrate	Sigma Aldrich
2,2,2-trifluoroethanethiol	Sigma Aldrich
2,2'-Azobis[2-(2-imidazolin-2-yl)propane] Dihydrochloride	Wako
2-mercaptoethanol	Sigma Aldrich
30% acrylamide/bis (29:1)	BioRad
3-aminopropyltriethoxysilane, 99%	Acros organics
4-mercaptophenylacetic acid	Sigma Aldrich
4-nitrobenzyl alchohol	Acros organics
5kDa biotin-PEG-SC	Laysan Bio
5kDa biotin-PEG-SVA	Laysan Bio
5kDa mPEG-SC	Laysan Bio
5kDa mPEG-SVA	Laysan Bio
6-hydroxy-2,5,7,8-tetramethylchroman-2-c	Sigma Aldrich
Acetone 99.6	Sigma Aldrich
Acetonitrile HPLC gradient grade	Avantor
Alexa Fluor 488 TFP ester	Thermo Fisher
Alexa Fluor 568 NHS ester	Thermo Fisher
Alexa Fluor 647 NHS ester	Thermo Fisher
Ammonium persulfate	BioRad
Ampicillin sodium salt	Applichem
Argon gas	Carbagas
Atto488 iodoacetamide	Atto-TEC
Atto532 iodoacetamide	Atto-TEC
Atto647N NHS ester	Atto-TEC
Boric acid	Sigma Aldrich
Brilliant Blue R-250	Fisher Scientific
Bromophenol blue	Fisher Scientific
Certified molecular biology agarose	BioRad
Chloramphenicol	Applichem
Cyclooctatetraene	Sigma Aldrich
Dextrose anhydrous	Acros
Dichloromethane	Acros Organics
Diethylether	Sigma Aldrich
Diisopropylethylamine	Acros Organics

Dimethyl sulfoxide	Sigma Aldrich
Dimethylformamide	Acros Organics
Dithiothreitol	Sigma Aldrich
D-Sucrose	Fisher Scientific
EDTA	Applichem
Ethanol	Fluka/Sigma
GelRed	Biotium
Glacial acetic acid	Merck
Glycerol	Fisher Scientific
Glycine electrophoresis grade	Applichem
Guanidine hydrochloride pure	Applichem
Guanidine hydrochloride ultrapure	MP Bio
HATU	Chempep
HBTU	PTI technologies
Hydrazine monohydrate	VWR
Hydrochloric acid 2M	Fisher Scientific
Hydrochloric acid 37% fuming	Merck
Hydrogen peroxide	Reactolab
Imidazole, extra pure	Fisher Scientific
Isopropanol	Fisher Scientific
Isopropyl-beta-D-thiogalactopyranoside	Applichem
Kanamycin sulfate	Roth
L-Arginine hydrochloride	Applichem
LB agar	Fisher Scientific
LB broth	Fisher Scientific
L-Cysteine hydrochloride monohydrate	Fluka Biochemika
L-cysteine methyl ester hydrochloride	Fluka
Magnesium sulfate heptahydrate	Merck
Manganese(ii) chloride tetrahydrate	Applichem
Methanol	Thommen Furler
Methoxyamine hydrochloride	Sigma Aldrich
Methylthioglycolate	Sigma Aldrich
NAD	Applichem
Nitrogen gas	Carbagas
Phenylmethyl sulfonyl fluoride	Applichem
Phenylsilane	Sigma Aldrich
piperazine-N,N′-bis(2-ethanesulfonic acid)	Applichem

Piperidine 99%, extra pure	Acros Organics
Polyethylene glycol	Sigma Aldrich
Potassium acetate	Applichem
Potassium chloride	Sigma Aldrich
Potassium phosphate dibasic	Fisher Scientific
Protocatechuate 3,4-dioxygenase	Sigma Aldrich
Sodium acetate anhydrous	Fisher Scientific
Sodium bicarbonate	Acros Organics
Sodium chloride	Applichem
Sodium dodecyl sulfate	Applichem
Sodium hydroxide 2M	Fisher Scientific
Sodium hydroxide pellets	Sigma/Fluka
Sodium hypochlorite 13/14%	Reactolab
Sodium phosphate dibasic heptahydrate	Fisher Scientific
Sodium phosphate monobasic monohydrate	Fisher Scientific
Sodium tetraborate decahydrate	Sigma Aldrich
Sulfuric acid	Merck
TEMED	Applichem
Tetrakis(Triphenylphosphine)palladium(0)	Sigma Aldrich
Triethylamine for HPLC	Roth
Trifluoroacetic acid, HPLC grade	Thermo Fisher
Triisopropylsilane	Sigma Aldrich
Tris ultrapure	Applichem
Tris(carboxyethyl)phosphine	Sigma Aldrich
Triton X-100	Sigma Aldrich
Tween 20	Applichem
Urea crystalline	Applichem
Water	Romil

Various components

Dulbeccos PBS w/o Ca/Mg	BioConcept
DMEM/F-12	Thermo Fisher
Newborn calf serum	Life Technologies
OptiMEM-1	Life Technologies
Neutravidin	Thermo Fisher
Tetraspeck microspheres	Thermo Fisher
SA-S-1L-SecureSeal 0.12mm doublesided tape	Grace Biolabs

DNA clean and concentrator 100	Zymo
QiaPrep miniprep kits	Qiagen
QiaQuick DNA purification kit	Qiagen
Ni-NTA resin	Qiagen
Complete EDTA-free protease inhibitor tablets	Sigma Aldrich
Microtube PTFE 0.56x1.07mm	Fisher Science
DAPI	BioConcept (NEB)
Miracloth	VWR
RNAse A	Qiagen
Yeast extract	Applichem
Tryptone	Sigma/Fluka
Slide-A-Lyzer MINI Dialysis Device, 10K MWCO, 0.1 mL	Thermo Fisher
Slide-A-Lyzer Dialysis Cassettes, 7K MWCO, 3 mL	Thermo Fisher
Slide-A-Lyzer Dialysis Cassettes, 7K MWCO, 0.5mL (10 cassettes)	Thermo Fisher
Spectra/Por 3 Dialysis Membranes 3,500 MWCO dia 29mm	Socochim
Vivaspin 500 centrifugal concentrators, MWCO 10,000 Da	Sartorius Stedim
Amicon Ultra-0.5, Ultracel-3 Membrane, 3 kDa	Merck
5 min epoxy	Swiss composite
5% Criterion TBE gels	BioRad

Peptide synthesis reagents

Reagent	**Supplier**
Fmoc-Ala-OH	Novabiochem
Boc-Cys(Trt)-OH	Novabiochem
Fmoc-Asp(OtBu)-OH	Novabiochem
Fmoc-Glu(OtBu)-OH	Novabiochem
Fmoc-Phe-OH	Novabiochem
Fmoc-Gly-OH	Novabiochem
Fmoc-His(Trt)-OH	Novabiochem
Fmoc-Ile-OH	Novabiochem
Fmoc-Lys(Boc)-OH	Novabiochem
Fmoc-Lys(OH)-HCl	FluoroChem
Fmoc-Lys(Me3Cl)-OH	Novabiochem
Fmoc-Lys(Alloc)-OH	GL Biochem
Fmoc-Leu-OH	Novabiochem
Fmoc-Met-OH	Novabiochem
Fmoc-Asn(Trt)-OH	Novabiochem

Fmoc-Pro-OH	Novabiochem
Fmoc-Gln(Trt)-oh	Novabiochem
Fmoc-Arg(Pbf)-OH	Novabiochem
Fmoc-Ser(tBu)-oh	Novabiochem
Fmoc-Ser(PO(PBzl)OH)-OH	Novabiochem
Fmoc-Thr(tbu)-OH	Novabiochem
Fmoc-Tyr(tbu)-OH	Novabiochem
Fmoc-Val-OH	Novabiochem
Fmoc-Trp(Boc)-OH	Novabiochem
2-Chlorotrityl chloride resin (100-200 mesh), Novabiochem	Novabiochem
Fmoc-Tyr(tBtu)- Wang Resin (100-200 mesh)	Novabiochem
Rink Amide resin (100-200 mesh)	Novabiochem

Enzymes and molecular biology reagents

Enzyme/reagent	**Supplier**
2-log ladder	BioConcept (NEB)
Antarctic phosphatase	BioConcept (NEB)
BamHI-HF	BioConcept (NEB)
BglII	BioConcept (NEB)
BsaI-HF	BioConcept (NEB)
Catalase	Sigma Aldrich
dNTP set	BioConcept(NEB)
DpnI	BioConcept (NEB)
DraIII-HF	BioConcept (NEB)
EcorI-HF	BioConcept (NEB)
EcorV-HF	BioConcept (NEB)
Effectene transfection reagent	Qiagen
Glucose oxidase	Sigma Aldrich
HindIII-HF	BioConcept (NEB)
Klenow fragment	BioConcept (NEB)
Pfu Turbo DNA polymerase	Agilent
Phusion High-fidelity DNA polymerase	BioConcept (NEB)
Precision plus protein all blue standard	BioRad
Precision plus protein dual-color standard	BioRad
PstI	BioConcept (NEB)
PstI-HF	BioConcept (NEB)
Q5 DNA polymerase	BioConcept (NEB)

QuikChange XL-II site-directed mutagenesis kit	Agilent
Recombinant shrimp alkaline phosphatase	BioConcept (NEB)
RNAse A	Qiagen
ScaI-HF	BioConcept (NEB)
Shrimp alkaline phosphatase	BioConcept (NEB)
T4 DNA ligase	BioConcept (NEB)
T4 DNA ligase	BioConcept (NEB)
T5 DNA exonuclease	BioConcept (NEB)
Taq DNA ligase	BioConcept (NEB)
Taq DNA polymerase	BioConcept (NEB)
Taq DNA polymerase	BioConcept (NEB)
Thrombin high activity	Calbiochem

Bacterial strains

Strain name	Genotype	Provider
NEB5a *E.Coli*	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	*New England Biolabs*
BL21DE3 *E.Coli*	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	*New England Biolabs*
BL21DE3(plysS) *E.Coli*	F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pLysS (Cam^R)	Calbiochem (Merck)
Rosetta2(BL21) *E.Coli*	F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pRARE2 (Cam^R)	Novagen (Merck)

Plasmid vectors

Name	Purpose	Resistance
Pet3a	T7 RNA pol II mediated expression without tags	Ampicillin
Pet30	6xH N-term SUMO for N-term truncations	Kanamycin
Pet15b	6xH N-term with thrombin cleavage site	Ampicillin
pUC19	General purpose cloning vector used for individual 601 sequences	Ampicillin
pWM530	Backbone digestable plasmid for array DNA excision	Ampicillin
pTXB	Plasmid used for –NpuN-6xH, NpuC-GyrA and NpuN-NpuC-6xH	Ampicillin
pCDNA3.1+	Mammalian expression vector	Ampicillin

Laboratory equipment

Instrument	Manufacturer	Application
Agilent 8453	Agilent	UV-Vis spectrophotometer
Chemidoc MP imaging system	Bio-Rad	Gel imager
Multitron standard	HT Infors AG	Incubator
Avanti J-20 XPI	Beckman Coulter	Centrifuge
Allegra X-15R	Beckmann Coulter	Table-top centrifuge
Äkta Pure	GE Healthcare	Protein and plasimd purification
5424R	Eppendorf	Microcentrifuge
Seven Compact pH/Ion S220	Mettler Toledo	pH-meter
InLab Science Pro-ISM	Mettler Toledo	pH electrode
Lyoquest	Telstar	Lyophilizer
Nanodrop 2000	Thermo Scientific	Small volume spectrophotometer
T100	Bio-Rad	Thermal Cycler
1260 Infinity with Quat pump and DAD	Agilent	Analytical HPLC system
1260 Infinity with dual pump and MWD	Agilent	Preparative HPLC system
Zorbax 300SB 7µm 21.2x250mm	Agilent	C18 preparative column
Zorbax 300SB 5µm 9.4x150mm	Agilent	C18 semipreparative column
Zorbax 300SB 5µm 4.6x150mm	Agilent	C18 analytical column
LCQFleet Ion trap	Thermo Fisher	Mass spectrometer
Sonifier 450	Branson	Cell sonicator
S200 10/300 GL	GE Healthcare	Analytical gel filtration column
HiTrap SP HP	GE Healthcare	Cation exchange trap
HiTrap Q FF	GE Healthcare	Anion exchange trap
PowerPac™ HC	Bio-Rad	High current power supply
PowerPac™ Basic	Bio-Rad	Power supply
XK 50/30 with Sepharose 6 FF	GE Healthcare	Plasmid gel filtration column
TI-E microscope	Nikon	CoSM
iXon	Andor	Image acquisition
Micro-mirror TIRF microscope	Mad city labs	smFRET
OBIS lasers 532nm/647nm	Coherent	Exciation smFRET
iXon Ultra	Andor	Image acquisition
SP210IWZ syringe pump	World Precision Instruments	Flow channel injection

7. Methods

7.1. DNA preparation and purification methods

PCR for cloning

Q5 DNA polymerase or Phusion DNA polymerase were used for PCR amplifications for cloning. With final volumes given in parentheses, template (0.02ng/µL reaction), forward and reverse primers (0.25µM) and dNTP mixture (0.2mM) were mixed with reaction buffer (1x) and water were mixed to the desired final volume before the polymerase was added (1U/µL). PCRs were carried out by initial denaturation at 98°C for 30s followed by 25-35 cycles of 10s 98°C denaturation, 30s annealing at 55-72°C, 30s/kb target extension at 72°C and ended with 2min extension at 72°C.

Agarose gel analysis of DNA

Gels were prepared by boiling 0.5-2.0% (w/v) solid agarose in 25mL 1x TBE with excess water in a microwave until agarose was fully dissolved and excess water had evaporated following by casting in 11x8cm trays with 8-well combs and cooling until solid. Samples were prepared in 6µL with 1x gel loading dye with SDS and the gels were run with 1x TBE as running buffer at 110V for 30-70min. Gels were stained by post-staining using 3x GelRed in 50mL followed by imaging using the GelRed channel on the imager.

PCR purification of DNA

The DNA to be purified was mixed with 5 volumes of PB followed by loading and centrifugation for 1min at 15000g through a QiaQuick column. The column was washed with 750µL buffer PE and centrifuged for 1min at 15000g, followed by flowthrough removal and centrifugation for another 5min. Finally the column was placed in a clean tube and plasmid DNA was eluted with 30-50µL buffer EB followed by centrifugation for 1min

Gibson assembly cloning

Gibson assembly mix for isothermal cloning was prepared as described in the publications describing its development[374,375]. 150µL 5x isothermal buffer (0.5M Tris, 1mM each dNTP, 50mM $MgCl_2$, 50mM DTT, 5mM NAD, 0.025% (w/v) PEG6000), 167µL water, 3.0U T5 DNA exonuclease, 18.5U Phusion DNA polymerase and 2000U Taq DNA ligase followed by distribution into 15µL aliquots, flash freezing and storage at -20°C. For cloning, 50-100ng of the linearized vector backbone along with equimolar amounts of all inserts were mixed to 10µL with MQ H_2O. 5µL of this was added to 15µL Gibson assembly mixture and 5µL to MQ H_2O as a negative control followed by 60min incubation at 50°C. 5µL from each reaction was analyzed by agarose gel and 5µL was transformed.

Circular polymerase extension cloning

Approximately 100ng of the linearized vector backbone was mixed with the insert pieces at 2-3 molar equivalents. DMSO (3% v/v), dNTPs (400µM), 5x HF Phusion reaction buffer (1x), Phusion HF polymerase (0.04U/µL) were added for the indicated final concentrations in 25µL, followed by dilution to this volume with water. Similar mock reactions without polymerase were prepared. The reactions were put for thermal cycling with 3min denaturation at 98° and 3-10 cycles 30s denaturation at 98°C, 30s annealing at 55°C and extension for 15s/kb at 72°. Samples were analyzed by gel and transformed.

Transformation

1-100ng of plasmid DNA or cloning product was mixed with 50µL well-suspended chemically competent cells. These were incubated 10min on ice. After 10min, the lower part of the tubes were immersed into a water bath at 42°C and kept there for 35s, followed by putting them them back on ice and leaving them there 2min. Then 900µL SOC medium was added to each transformation reaction and incubated shaking at 37°C for 1 hour. Depending on the expected transformation efficiency 10 or 100µL were plated on agar plates containing appropriate antibiotics. If the efficiency was expected to be low all of the recovered cells were transformed; Centrifugation at 21000g for 1min, removal of 875µL of the supernatant, resuspension of the pellet in the remaining supernatant and plating out all of the suspension. Agar plates were incubated for 10-18h at 37°C.

Plasmid miniprep and purification

Small scale plasmid growth and purification was carried out largely in accordance with the manufacturers protocol; 5mL 2xTY broth with suitable antibiotics was inoculated with a single colony from an agar plate and grown for 10-20h shaking at 37°C. The cell pellet was centrifuged at 4000g for 10min followed by removal of the medium and resuspension of the pellet in 250µL buffer P1 supplemented with RNAse A. The cells were lysed with 250µL P2, followed by neutralization with 350µL N3 and centrifugation for 10min at 21000g to remove chromosomal DNA. The supernatant was loaded onto a QiaPrep column and centrifuged at 15000g for 1min. Then it was washed with 500µL buffer PB and centrifuged for 1min, followed by washing with 750µL buffer PE, centrifugation for 1min and removal of flowthrough followed by another 5min of centrifugation. Finally the column was placed in a clean tube and plasmid DNA was eluted with 30-50µL buffer EB followed by centrifugation for 1min.

DNA array doubling

The first doubling was done by ligation of an 1x601 extension piece digested with BglII and Pst into the similarly digested 1x601 P5 in pWM531, transformation and isolation of the resulting 2x601 plasmid. Subsequent doublings were done by ligation of the isolated insert obtained by digestion of the construct with PstI and BglII into the isolated vector digested with PstI and BamHI. This was done in two rounds to get the 3x601 in the first round and the 4x601 and 5x601 in the second round. The 3x insert was transferred from P5 4x601 to P1 1x601 similarly by digestion of both with PstI and BglII respectively followed by ligation, transformation, plasmid isolation and sequencing.

Large-scale DNA purification by alkaline lysis and gel filtration

Plasmids were generated on large scale in DH5α cells grown in 6L 2xTY medium and isolated by alkaline lysis followed by preparative gel filtration. A pre-culture of 100mL was prepared from a glycerol stock of cells verified to harbor the desired plasmid and this pre-culture was grown for approximately 6h at 37°C. 15mL from this was used to inoculate each of 6x1L 2xTY medium and this was grown for the subsequent 18-20h at 37°C. Centrifugations and alkaline lysis until filtering were done at 4°C. Cells were harvested by centrifugation, resuspended and homogenized in 80mL alkaline lysis solution I (50mM glucose, 25mM Tris, 10mM EDTA, pH 8.0). The homogenate was diluted to a total volume of 120mL with the same solution. 240mL alkaline lysis solution II (0.2M NaOH, 1% SDS) was added and mixed. 240mL alkaline lysis solution III (4M KAc, 2M Acetic acid) was added to neutralize the solution followed by rapid, but gentle mixing and subsequent standing for 15min. The supernatant was recovered by centrifugation and filtered through miracloth into a volumetric cylinder. After transfer to a centrifuge tube, 0.52 volumes of isopropanol was added and it was allowed to

stand for 20min at RT, followed by centrifugation at 11000g for 20min at RT. The supernatant was removed and the pellet redissolved in 30mL TE 10/50 and transferred to falcon tube. 100 units of RNAse A was added and allowed to digest 2h at 37°C. Solid KCl was added to a final concentration of 2.0M and volume adjusted to 35-40mL. The sample was centrifuged and the supernatant loaded onto a 50mL superloop. This was injected into a 550mL sepharose 6 XK 50/30 column and the pure plasmid collected in the dead volume. The purified plasmid was precipitated with 0.5 volumes of isopropanol, redissolved in 5-20mL TE 10/0.1 and stored at -20°C for later use.

Generation of labeled and biotinylated chromatin array DNA

A plasmid containing a 12x601 NPS separated by 30-bp linker DNA segments (containing ScaI restriction sites) and a unique BsaI restriction site at the end was generated recombinantly, released from the vector backbone by EcorV digestion and purified by PEG purification. The plasmid was dephosphorylated to prevent self-ligation with Antarctic phosphate followed by inactivation. BsaI digestion then released the non-palindromic overhang. For labelling, 5 nmol of a synthetic oligonucleotide (5'-ph-CAGCTAGTCTGCT(amine-linker)CAGATATCGTCG-3'-Biotin), containing an internal amine attached to a dT and a C-terminal biotin, was mixed with 3mM Atto647N-NHS ester (from a 30mM stock in DMSO) in 0.1M sodium tetraborate and incubated overnight at 4°C. Subsequently, the labeled fragment was ethanol precipitated and the pellet was dissolved in 10 ml of deionized MQ H_2O and quantified. Then, the labelled oligonucleotide was annealed to its complementary DNA strand (5'-CGACGATATCTGAGCAGACTA-3') in T4 DNA ligase reaction buffer (NEB). The labelled dsDNA (5 eq. excess) was ligated to 12x601 array DNA using T4 DNA ligase for 1 h at room temperature. The excess oligonucleotide was removed by purification with QIAquick spin columns (Qiagen), followed by PEG precipitation with 9% PEG 6000. Mononucleosome DNA was prepared using the same methodology, but employing a 1x601 NPS sequence. Concentrations were determined by absorption spectrometry[349].

Isolation of recombinant fragments for FRET arrays

75-85pmol of plasmid DNA buffer exchanged by ethanol precipitation to MQ H_2O was mixed with CutSmart and water to a final volume of 200µL. 50 units of BsaI-HF and 50 units of DraIII-HF were added for digestion for 8-10h, then another 20 units of each enzyme was added to get the digestion to completion (additional 20 units added if not complete). 60 units of EcorV-HF were added and digestion allowed to continue for another 6-10h. Two rounds of stepwise PEG precipitation was done to separate the excised fragment of interest from the plasmid backbone fragments using concentrations of PEG from 7.0-8.5%. After two rounds, a final DNA cleanup step was done using a Zymo Clean and Concentrator 100 column according to the manufacturers manual. Completion of digestion was checked by small-scale test ligation to excess of a neighbouring fragment and verification of complete displacement of the isolated recombinant fragment to a larger species on an agarose gel. The ligation ready pieces were stored at -20°C for later use.

Oligonucleotide labeling and purification for FRET

To remove residual amines from synthesis and purification, oligonucleotides were washed by ethanol precipitation. Oligonucleotides were dissolved in 100µL MQ H_2O and precipitated by addition of 10µL 3M NaCl and 275µL cold ethanol. They were pelleted by 30min centrifugation at 21000g and 4°C. The supernatant was removed and the pellet allowed to air dry for 15min. The redissolving, precipitation and centrifugation were repeated. The pellet was redissolved in a volume of MQ H_2O expected to result in a concentration of 1mM. For each labeling reaction, 5-10nmol of the washed oligo was diluted with oligo labeling buffer (0.1 M sodium tetraborate, pH 8.5 (9.25 for TFP ester labeling) to a final volume of 30µL. A 0.6µL sample was taken and diluted with 50µL RP-HPLC oligo solvent A (95% 0.1M TEAA, 5%

ACN). 40µL of this was injected for analysis by RP-HPLC on an InertSustain 3µm, 4.6x150mm GL sciences C18 analytical column using a gradient of 0-100% oligo RP-HPLC solvent B (70% 0.1M TEAA, 30% ACN). 5µL of 5mM NHS-ester dye in DMSO was added and the reaction allowed to proceed for 4-8h at room temperature. The progression of the reaction was monitored by RP-HPLC analysis as done for the sample prior to starting the labeling reaction, more stocks of 5µL 5mM NHS ester were added until reaching >50% labeled oligo. The oligo was precipitated twice with ethanol using similar ratios as done for the initial washing prior to labeling. It was redissolved in 30µL MQ H_2O and diluted with 70µL oligo RP-HPLC solvent A. The labeled oligo was purified by RP-HPLC using the same gradient and column as above and collected manually followed by ethanol precipitation. The purified labeled oligonucleotide was redissolved in a volume of MQ H_2O to give an approximate concentration of 2.5µM.

PCR for 1x601 labeled/unlabeled fragments

The labeled P2 1x601, P3 1x601 and labeled P4 1x601 segments were generated by PCR by mixing Thermopol (1x), template (0.02ng/µL), forward primer (labeled or unlabeled) (0.250µM), reverse primer (labeled or unlabeled) (0.250µM) and dNTPs (0.2mM each) with water in Nx50µL to the final concentrations given in the parentheses. Nx1.25 units Taq DNA polymerase was added to the mixture and gently mixed followed by aliquoting 50µL into each of N PCR tubes. PCR was carried out with a PCR program consisting of 12s initial denaturation at 94°C followed by 12 cycles of 12s denaturation at 94°C, 12s annealing at 60-65°C and 12s extension at 72°C. Final extension was done for 12s at 72°C. PCR product from the N tubes were pooled and stored at -20°C for later use.

Digestion and test ligations for PCR fragments

Prior to digestion, 400-500µL of the PCR mixture for each PCR-generated fragment was purified by PCR purification using 3 QiaQuick columns according to the manufacturers instruction eluted with TE and then exchanged by ethanol precipitation to water. After purification the pieces were digested by mixing 75-85pmol of each piece in 200µL with 10x CutSmart to a final concentration of 1x and a sample taken. The pieces were digested by addition of 50 units BsaI-HF and DraIII-HF. The digestion was allowed to proceed for 6-8h. Another 20 units of each enzyme was added and the digestion allowed to proceed overnight. Samples were taken and analyzed on a 2% agarose gel alongside the undigested samples. The digestion reactions were purified with QiaQuick PCR purification columns and the concentration determined. Small scale test ligations in 20µL were done with T4 DNA ligase using 0.5-4µL of these pieces as well as P1 and P5 to ensure that P1 could be completely displaced with excess P2, P2 displaced with P3, P3 displaced with P4 and P5 displaced with P4. Upon verification of this, the pieces were either immediately used for convergent ligations to generate the 6xNPS intermediates, or stored until later use.

Convergent assembly and purification of 12x FRET DNA arrays

30-40pmol recP1 4x601 was mixed with 15% excess P2 1x601 in T4 DNA ligase buffer to 1x with 600U T4 DNA ligase followed by ligation for 6-8h. Then P3 1x601 was added in 15% excess relative to P2 along with T4 DNA ligase buffer to retain 1x total concentration and another 600U T4 DNA ligase and ligation for 12-16h. Similar was done with 30-40 recP5 5x601 with 15% excess P4 1x601 in T4 DNA ligase buffer and with ligase. After 6-8h, the P5 anchor was added at about 60pmol followed by more ligase buffer and ligase for another 12-16h of ligation. Samples from all stages of ligations were analyzed by agarose gel to confirm that the ligations had proceeded to a satisfactory degree. Then the 6x601 intermediates were purified by step-wise PEG precipitation using PEG in the range of 7.0-8.0%. Pellets were

redissolved in TE(10/0.1) and analyzed along with the final supernatant to identify and pool fractions with the desired intermediates. 15-35pmol of the 6x601 intermediates were mixed using 5-10% excess P4-P5-anchor in 1x T4 DNA ligase buffer with 600U of ligase and left to ligate for 10-16h. The formation of the product was confirmed by agarose gel and purified by step-wise PEG precipitation in the range 5.0-6.0%. The pellets were redissolved in TE(10/0.1) and analyzed by gel to pool the final double-labeled array DNA.

7.2. Solid-phase peptide synthesis

Conversion of chlorotrityl to trityl hydrazine resin and manual coupling

0.85mmol chlorotrityl resin was swollen in 3mL DMF in a scintillation vial and the scintillation vial transferred to an ice bath at 0°C. 2.55mmol (444µL) DIPEA, 2.6mmol hydrazine monohydrate (82µL) were dissolved in DMF to 1000µL (474µL). This mixture was added dropwise to the resin stirring at 0°C. The ice bath was removed and the suspension kept stirring at RT for 1h. To cap the resin, 100µL methanol was added and the resin allowed to stir for another 10min at RT. The resin was transferred to a SPPS reaction vessel and washed 1min with DMF.

The loading of the resin was assumed to be half of the initial amount (0.425mmol) due to possible hydrolysis. 5 equivalents (2.125mmol) of the first amino acid was weighed into a scintillation vial (1g for Fmoc-Lys(Boc)-OH). In 8mL DMF, 0.5M HBTU solution was prepared (4mmol, 1.5g). Half of this was used to dissolve the amino acid and DIPEA added (4.3mmol, 740µmol) followed by 1min incubation at RT. The solution was then added to the drained peptide reaction vessel containing the resin. Coupling was allowed to proceed 30min with stirring every 3-5min. Coupling was repeated with 2 equivalent of amino acid and correspondingly smaller volumes of HBTU and DIPEA. The double-coupled resin was flow washed with DMF, DCM and methanol before drying in a dessicator. Resin loading was determined by deprotection of the Fmoc with 20% piperidine and measurement at 310nm in triplicate samples.

Automated solid-phase peptide synthesis and cleavage

Automated peptide synthesis was done on a 0.1mmol scale using a Tribute peptide synthesizer without microwave heating. 5 equivalents of each amino acid to be coupled were weighed into a tube for the synthesizer. β-branched amino acids and arginine were double-coupled. Deprotection was done with 20 % (v/v) piperidine in DMF. Amino acids were dissolved by the synthesizer with 4.8eq HBTU and 10eq DIPEA followed by addition to the resin and coupling for 30min. The resin was washed with DMF between coupling and deprotection steps. Following completion of the synthesis, resins were washed with DMF, DCM and MeOH and dried. Peptides were cleaved from the resin by incubation with 10mL/g of a mixture of TFA/TIS/H_2O (95/2.5/2.5%) for 3h at RT and filtered through a frit. Crude peptides were precipitated by addition of cold diethyl ether, followed by centrifugation and dissolving of the pellet in 50% ACN. The crude peptide was lyophilized for later purification.

Alloc-deprotection and cysteine coupling through isopeptide bond

Resin (~0.1mmol) was washed in DCM and allowed to swell for 30min. 1mL dry DCM and 2.4mmol (294µL PhSiH3 was added followed by stirring and addition of 3mL DCM with 0.025mmol Pd(PPh3)4) and treatment for 30min to deprotect. The resin was washed with DCM and the treatment with PhSiH3 and Pd(PPh3)4) repeated twice. The resin was then washed with 0.5% (v/v) DIEA in DMF, 0.5%(w/v) sodium diethyldithio carbamate in DMF, DCM/DMF (1:1)

and 0.5% (w/v) HOBt in DMF. 5equivalent of the Boc(Cys(Trt)-OH amino acid was activated with 4.76eq of HATU from a 0.5M stock in DMF for 2min and 10eq DIEA for 1min followed by addition to the resin and incubation with stirring for 30min. This was repeated once and the resin washed with DMF.

Peptide labeling and thiazolidine deprotection

~0.5µmol of the peptide was dissolved in labeling buffer (200mM phosphate, pH 7.3, 6M GuaHCl) to a concentration of 10mM and 1mg Atto532 iodoacetamide was added in 50µL DMSO and the reaction monitored by HPLC. After completion, the reaction was quenched by addition of 2mM βME. 0.5M methoxyamine was added and pH adjusted to 5 with 2M NaOH. Opening of the thiazolidine was followed by RP-HPLC and ESI and after completion the peptide was purified with semipreparative RP-HPLC on a gradient of 0-70%B in 45min.

7.3. Protein expression, purification and analysis

SDS-PAGE casting and analysis

Gels were cast manually. Resolving gel was prepared by mixing MQ H_2O (24-41%v/v), 30% acrylamide (29:1 bis) (10-15%w/v), 1.5M Tris pH 8.8 (0.375M) and 10% SDS (0.1%w/v) to the values indicated in parentheses followed by addition of TEMED (0.05%v/v) and 10% APS (0.05%w/v) to the final volumes indicated and deposition of 4.5-5.0mL between spacer and glass plates for the Mini-PROTEAN system. After polymerization for 30-45min, the water overaly was removed and stacking gel prepared; MQ H_2O (57%), 30% acrylamide (5%w/v), 0.5M Tris (pH 6.8) (0.125M) and 10% SDS (0.1%w/v) along with TEMED (0.1%v/v) and 10% APS (0.05%w/v) followed by deposition on the resolving gel, insertion of a comb and polymerization for 30-45min. Gels were stored in 1x Tris-glycine running buffer at 4°C for up to 14days before use or immediately used. Samples were loaded mixed with loading dye up to 50:50 in volumes of 10-15µL and run at 200V for 45-55min followed by staining with coomassie staining solution and destaining with destaining solution.

Wild-type histone expression and purification

Human *wild-type* histones were expressed in BL21 DE3 plysS cells from genes inserted into a pet3a plasmid. Cell cultures were grown at 37°C in LB media supplemented with 100µg/mL ampicillin and 35µg/mL chloramphenicol until an OD600 of 0.6. Protein expression was induced by addition of IPTG to a final concentration of 0.5mM. Expression of the protein was allowed to continue until 3h post-induction. Cells were harvested, cell pellets were resuspended in lysis buffer (20mM Tris, 1mM EDTA, 200mM NaCl, 1mM βME, 1 protease inhibitor tablet/50mL, pH 7.5). Cells were lysed by freeze-thawing and sonication. The inclusion body pellet was washed twice with 7.5mL of histone lysis buffer with 1% triton and once without triton. Histones were resolubilized in histone resolubilization buffer (6M GdmCl, 20mM Tris, 1mM EDTA, 1mM βME, pH 7.5), dialyzed into urea buffer (7M urea, 10mM Tris, 1mM EDTA, 0.1M NaCl, 5mM βME, pH 7.5) followed by purification by cation exchange (using a HiTrap SP HP 5mL column). Collected fractions were analyzed by SDS-PAGE, followed by final purification using preparative RP-HPLC. Collected fractions were characterized by analytical RP-HPLC and ESI-MS, lyophilized and stored at -20°C until use in octamer refolding[349].

Purification of N-terminally truncated histones

6xH-SUMO-Histone(ΔN) constructs (H4, H3, H2A, H2AX) were expressed and inclusion bodies treated as for the wild-type histones in either BL21DE or Rosetta 2 cells. Similarly Ulp1 was expressed and washed in inclusion bodies. Inclusion bodies were resolubilized in imidazole resolubilization buffer (6M GuaHCl, 50mM Tris, 100mM NaCl, 5mM imidazole, pH 7.5) and the insoluble cell debris removed by centrifugation. The supernatant was loaded onto pre-equilibrated Ni-NTA resin (~2mL). The resin was allowed to bind for 45-60min and the flowthrough collected. The resin was washed with 3x5CV of resolubilization buffer with 25mM imidazole, followed by elution with 4x2CV of the same buffer in 300mM imidazole. Ulp1 was purified similarly however with only ~1mL resin. Protease refolding and cleavage was done by dialysis of the eluted proteins against urea cleavage buffer (1M urea, 75mM Tris, 150mM NaCl, 5mM DTT, 25mM L-arginine, 5mM L-cysteine, pH 7.5) for 16h. The dialysis mixture was dialyzed against 1L MQ H2O with 1%(v/v/) acetic acid for 3x3-4h and lyophilized. The crude cleavage mixture was purified by RP-HPLC using a gradient of 30-70%B in 45min. The final purified truncated histones were lyophilized and stored for later reactions.

Expression and purification for Npu^{C}-AA-CysOMe peptide

NpuC-AA-GyrA-6xH was expressed in BL21 DE3 plysS cells. From a pre-culture, 4x1L of LB medium with chloramphenicol and ampicillin were inoculated and grown at 37°C until reaching an OD600 of 0.5-0.9. Protein expression was induced by addition of IPTG to a final concentration of 0.5mM followed by incubation for another 3-4h. Cells were harvested by centrifugation, homogenized with Ni-NTA lysis buffer (50mM Pi, 300mM NaCl, 5mM imidazole, pH 8.0) and flash frozen. Cells were lysed by thawing and sonication followed by collection of the supernatant by centrifugation. The supernatant was purified by Ni-NTA purification using 2-3mL of resin. The supernatant was bound to the resin for 30-45min and the flowthrough collected. The resin was washed with lysis buffer containing 20mM imidazole and the protein eluted with the same buffer at 300mM imidazole. Thiolysis was initiated by addition of CysOMe and TCEP to final concentrations of 100mM and 5mM respectively, followed by reaction for the next 16h. The pellet was collected, redissolved in resolubilization buffer and purified by preparative RP-HPLC using a gradient of 20-50% B in 45min. Fractions containing the purified protein were pooled and lyophilized for later usage.

HP1α expression

From a 100mL pre-culture grown at 37°C for 12-16h, 10-15mL was transferred to each 1L of LB medium containing appropriate antibiotics. The flasks were incubated shaking at 180-220rpm and 37°C until reaching an OD600 of 0.5-0.9 after 1.5-3h. Then culture flasks were transferred to 4°C for 10-15min and the temperature on the incubator reduced to 18°C. After 15min at cold room temperature, IPTG was added to a final concentration of 0.25mM and incubation done for the next 16-20h at 18°C shaking at the same frequency. Cells were centrifuged at 10000g for 10min at 4°C and cells resuspended in 15-18mL of Ni-NTA lysis buffer for each liter of culture. The suspesnsion was homogenized by passage through a needle and 10-20mL (10mL for -NpuN-6xH fusion constructs and 20mL for 6xH-Thr- constructs) of the homogenized cells were aliquoted in 50mL falcon tubes, flash frozen and stored at -80°C until use.

HP1α Ni-NTA and AIEX purification

An aliquot of the HP1 cells were thawed and sonicated at 30% amplitude with 10 cycles of 15s on and 45s off, with the tube submerged in an ice bath. A sample was taken from the lysate before centrifuging the cells at 15000g for 10min at 4°C and a sample taken from the supernatant. The rest of the supernatant was transferred to 2mL pre-equilibrated and

drained Ni-NTA resin followed by binding for 15-30min. The flowthrough was collected and the resin washed with 2x5CV of Ni-NTA wash buffer. Elution was done with 2x1.5CV Ni-NTA elution buffer and samples from lysate, supernatant, flowthrough, washes and elution were analyzed by SDS-PAGE. Elutions were collected and injected into a 5mL for purification with the FPLC and a 1mL AIEX column with the following steps: loading (0% B, 15CV), wash (0% B, 3CV), gradient(0-100%B, 20CV), final wash (100%B, 3CV). 1mL fractions were collected in a 96-well plate. Fractions with 280nm absorption were analyzed by SDS-PAGE and fractions with HP1α protein pooled.

7.4. Expressed protein ligations and desulfurization

Preparation of split intein column

Lyophilized stocks of NpuC-peptide were dissolved in water to a concentration of 2mM, as quantified by UV spectroscopy ($\varepsilon_{280nm,NpuC\text{-}CysOMe}$ = 2980 M-1cm-1). 125 µL stocks of this were flash frozen and stored at -80°C for later use. For the preparation of NpuC-resin, 500 µL of SulfoLink resin slurry (Pierce) was put into a small fritted column and washed with water. A 125 µL stock of the NpuC-CysOMe peptide was mixed with 125 µL of 2x coupling buffer (100 mM Tris, 10 mM EDTA, pH 8.5), TCEP was added to a final concentration of 25 mM and the pH was adjusted to 8.5. The peptide solution was then added to the SulfoLink resin in the capped fritted column and incubated for 45 min (15 min included agitation by nutation), then the column was drained and washed with 2 x 1 CV of 1x coupling buffer. Residual unconjugated iodoacetamide groups on the resin were capped by treatment with cysteine methylester (CysOMe).

Synthesis of HP1α(Atto532) and HP1α_{cdm}(Atto532)

500µL of HP1α-GG-NpuN-6xHis, HP1α(1-176)-GG-NpuN-6xHis or HP1α(I164E)-GG-NpuN-6xH at a concentration of 50 µM was added to 125µL of the NpuC resin in a small fritted column and incubated on a nutator for 5min. The flowthrough was collected and the resin washed with 4 CV of binding buffer (100mM phosphate, 1 mM EDTA, 1 mM TCEP, pH 7.2) with 500 mM NaCl, 4 CV of binding buffer with 300 mM NaCl and 4CV of binding buffer with 150 mM NaCl. Then 1 CV of *in situ* EPL buffer (50 mM MPAA, 200 mM MESNa, 100 mM phosphate, 150 mM NaCl, 10 mM TCEP, 1 mM EDTA, pH 7.7-7.9) containing 1 mM of peptide P1 or P2 was added to the resin and the column was capped and incubated on a nutator for 16-18h. The eluate was collected and the column further washed with 3 x 1 CV of intein column elution buffer (200 mM MESNa, 100 mM phosphate, 150 mM NaCl, 10 mM TCEP, 1 mM EDTA, pH 7.2). All fractions were analyzed by SDS-PAGE. Elution fractions were pooled and purified by gel filtration using a Superdex S200 10/300GL column using HP1 gel filtration buffer

Ligation and desulfurization for H3 K9me3

In a typical ligation, 3 mmol of H3 1–14 (K9me3) peptide hydrazide (5.0 mg) was dissolved in ligation buffer (6M GdmCl, 200mM phosphate, pH 3.0 at RT) to a final concentration of 5mM and the pH readjusted to 3.0. The peptide solution was kept at -10°C in an ice/KCl bath and $NaNO_2$ was added dropwise from a 200mM stock solution to a final concentration of 15mM. After mixing, the peptide was left at -10°C for 20 min. A total of 0.75 mmol truncated protein was dissolved in MPAA ligation buffer (6M GdmCl, 200mM phosphate, 300mM MPAA, pH 8.0) to a final concentration of 1.2 mM. After 20 min of activation, the peptide was brought to room temperature and the solution of protein in MPAA ligation buffer mixed with the activated peptide. pH was readjusted to 7.7–7.9 and ligation allowed to proceed until

completion as determined by RP-HPLC. When complete, TCEP was added, the reaction was mixed with two volumes of RP-HPLC solvent 20% B and purified by semipreparative RP-HPLC with a gradient of 0–70% B over 45 min and lyophilized. The ligation product was dissolved in TCEP desulfurization buffer (6M GuaHCl, 200mM phosphate, 250mM TCEP, pH 6.5) to a final concentration of 0.5 mM. GSH in TCEP desulfurization buffer and VA-044 radical starter were added to give final concentrations of 40 and 20 mM, respectively. pH was readjusted to 6.5 and desulfurization proceeded at 37°C. The reaction was continuously monitored by RP-HPLC and ESI-MS. When complete, two volumes of RP-HPLC solvent (20% B) were added and the product purified by semipreparative RP-HPLC using a gradient of 0–70% B over 45 min[349].

One-pot ligation and desulfurization for H4 K20me2

Approximately 2mg peptide was dissolved in ligation buffer (6M GuaHCl, 200mM Pi, pH 3) and cooled to -10°C. The hydrazide was activated by addition of $NaNO_2$ to a final concentration of 15mM and kept at -10°C for 10min. Then 100eq TFET was added and thioester formation proceeded 10min at room temperature. The TFET thioester was transferred to ~1mg of the lyophilized truncated H4 Δ1-37 A38C and TCEP added to 0.5mM. The reaction was incubated 25°C and progress monitored by RP-HPLC and ESI-MS. The reaction was allowed to proceed for ~16h until deemed complete. For desulfurization in the same pot, TCEP desulfurization buffer (500mM TCEP, 6M GuaHCl, Pi 200mM, pH 7) was added to equal volume along with GSH and VA-044 to 40mM and 20mM respectively. The reaction was incubated at 42 °C under agitation. The reaction was monitored by RP-HPLC and ESI-MS until it was seen to complete after about 6h. Then it was purified by RP-HPLC using a gradient of 0-70B in 45min and purified fractions lyophilized for later usage.

One-pot ligation and desulfurization for H2A.X K15ub 1-20

~8.5mg (~1µmol) Ub-MES was dissolved in 130 µl ligation buffer (6M GuaHCl, 0.2M Pi, pH 7.0)to get a protein concentration of 7.5 mM. The solution with the Ub-MES to was transferred to a tube containing 1.4mg lyophilized peptide along with 0.88 µl MTG (resulting in 75 mM, 10 eq over Ub-MES) were added. Finally, 9 µl of 0.5 M TCEP buffer was added to ensure reducing conditions. The ligation solution was incubated at RT, 200 rpm and monitored by RP-HPLC and ESI-MS. Upon reaching completion, the product was desulfurized in the same pot by addition of TCEP desulfuriztion buffer to equal volume along with GSH and VA-044 to final concentration of 40mM and 20mM. Desulfurization was monitored by RP-HPLC and ESI-MS. When complete, the product was purified by semipreparative RP-HPLC using a gradient of 0-70%B in 45min. Fractions with the pure product were lyophilized for later use.

Thioester formation, ligation and SUMO cleavage for H2A.X S139ph Δ1-20 A21C

~6mg (0.16µmol) of the lyophilized SUMO-H2A.X-NpuN protein was dissolved in 60µL resolubilization buffer (6M guanidine, 200mM Pi, pH 7.5). This was transferred slowly to 700µL of NpuC column binding buffer with 2M urea. A 2µL sample was taken from this mixture before loading 380µL of the resulting mixture was loaded onto each of the two NpuC columns from 250µL resin. The protein was incubated with the column material on a nutator for 3h before collecting the flowthrough and washing the column with 300µL high salt wash buffer and 300µL wash buffer. Then 250µL elution buffer with MESNA was incubated with the column 12-16h. The elution buffer was collected and another 200µL of elution buffer used to wash the column. Then 250µL elution buffer was added to each column and incubated nutating again for another 2h. The samples from the input, flowthrough, wash and elution fractions were analyzed by SDS-PAGE and the elution also by RP-HPLC and ESI-MS. The combined elution fractions were concentrated to a volume

of about 700µL with a centrifugal concentrator. Equal volume of MPAA at 100mM in 2M urea binding buffer was used to dissolve ~2mg (2µmol) of peptide C-terminal and this was added to the protein thioester. Ligation was monitored by RP-HPLC and ESI-MS and completed after about 1h. 100µL Ulp1 at a concentration of about 2µM was added along with DTT and L-arginine for SUMO removal. The reaction was monitored by RP-HPLC and ESI-MS and appeared complete after about 6h at which point the reaction mixture was purified by semipreparative RP-HPLC with a gradient of 0-70%B in 45min. Purified fractions were lyophilized.

One-pot ligation and desulfurization for H2A.X K15ub S139ph

Approximately 0.6mg (~0.05µmol) of the truncated phosphorylated N-terminal was weighed into a tube. The N-terminal ubiquitylated peptide was activated by in situ oxidation on ice as described above, followed by conversion with TFET thioester using ~100eq and incubation 10min at room temperature. Subsequently the components were mixed and TCEP added to 0.5mM. The reaction was monitored by RP-HPLC and ESI-MS. Upon reaching an acceptable level of ligation, the desulfurization reaction was initiated by addition of TCEP buffer, VA-044 and GSH as described above. The reaction was monitored by RP-HPLC and ESI-MS and upon reaching completion, the product was purified by semipreparative RP-HPLC to yield the desired product which was lyophilized for later use in octamer refolding.

7.5. Affinity titrations

Tryptophan fluorescence for affinity of unlabeled HP1α wild-type with H3 K9me3 peptide

A sample of 300nM unmodified HP1α was loaded as 80µl into an Ultra-micro fluorescence cuvette. Increasing concentrations of H3 K9me3 1–14 peptide to cover the range 0.2–100 mM was titrated in as 1.5–7µl per addition. Single-point measurements of the intrinsic tryptophan fluorescence were done for each titration point using 280nm as excitation wavelength and 3 nm as the excitation slit width. Emission was measured at 348 nm with a slit width of 3.0nm and an integration time of 0.5 s. The typical number of counts obtained was 50–70 k. The number of trials allowed per point was 5 or until a maximum s.d. of 0.5% was reached. The fluorescence counts measured were corrected for the dilution upon titration of the peptide and then normalized between the start and the end point. The normalized data were fitted with a quadratic binding equation using Origin[349].

Microscale thermophoresis for affinity of labeled HP1α wild-type with H3 K9me3 peptide

Measurements were carried out on a MonoLith NT.115 equipped with blue/green filters. H3 K9me3 1–14 peptide from a main stock was serially diluted in titration buffer (150mM NaCl, 50mM HEPES, 0.05% Tween, 1mM DTT, pH 7.5) to cover the concentration range 0.1–100 mM and result in 10µl samples of each. From a 300nM stock (based on Atto532 absorption) of HP1α(A532) or HP1αcdm(A532), 10µl was added to the serially diluted peptide. Samples were loaded into hydrophilic-coated capillaries and cap scanning done at 100% LED power. Thermophoresis was done at 80% IR laser power. Data were normalized between the start and the end point and fitted with a quadratic binding equation using Origin[349].

7.6. Chromatin array reconstitution and analysis

Histone octamer refolding

In a typical octamer refolding reaction, 0.4-1.5mg of each of the pure lyophilized human histones were dissolved in unfolding buffer (6M GuaHCl, 10mM Tris, 5mM DTT, pH 7.5). The exact concentration was determined by UV spectroscopy, using the following extinction coefficients: $\varepsilon_{280nm,H2A}$=4470$M^{-1}cm^{-1}$, $\varepsilon_{280nm,H2B}$=7450$M^{-1}cm^{-1}$, $\varepsilon_{280nm,,H3}$ =4470$M^{-1}cm^{-1}$, $\varepsilon_{280nm,H4}$=5960$M^{-1}cm^{-1}$. Equimolar amounts of H3 and H4 were then mixed along with 1.05 equivalents of H2A and H2B at 1 mg ml^{-1} and octamers were refolded by dialysis against refolding buffer (2M NaCl, 10mM Tris, 1mM EDTA, 5mM DTT, pH 7.5). The refolded octamers were subsequently purified by gel filtration on a Superdex S200 10/300GL column. Collected fractions were analysed by SDS-PAGE), and octamer containing fractions were pooled and concentrated to ~50µM octamer concentration. Finally, glycerol was added to a final concentration of 50%, concentrations were determined by UV spectroscopy and octamer stocks were stored at -20°C[349].

Chromatin array reconstitution on 12-mer DNA fragments and analysis

Chromatin arrays were reconstitured on a scale of 6.5-30pmol (Calculated based on nucleosome positioning sequences). 12xNPS array DNA was mixed with equimolar MMTV buffer DNA, NaCl to a final concentration of 2M and water, followed by mixing and addition of 2-2.4 molar equivalents of histone octamers with or without modifications. The mixture was transferred to a micro-dialysis tube and dialyzed with a linear gradient from TEK2000 to TEK10 over 16-18h. The dialysis tube was transferred to 200-600mL TEK10 for another 1h of dialysis. Chromatin assembly mixture was taken out of the dialysis tube and centrifuged at 21000g for 10min followed by transfer of the supernatant to another tube. The concentration and volume of the mixture was determined. Gel analysis was done with 0.25-0.50pmol of chromatin assembly sample (Calculated based on the total 260nm absorption and the extinction coefficient for each nucleosomal repeat) mixed to 10µL with TEK10 and 25% sucrose further added to a final concentration of 5-7%. This was run in 0.7% agarose gels made with 0.25x TB and using the same as running buffer at 90V for 90-100min.

For ensemble FRET analysis requiring removal of MMTV DNA and nucleosomes, 5-10% of the volume was taken aside for comparison and the remainder mixed with an equal volume of 6mM Mg^{2+} for precipitation on ice for 10min followed by 10min centrifugation at 21000g. The supernatant was transferred to another tube and the chromatin pellet redissolved in a similar volume of TEK10 as present prior to addition of magnesium. Similar volumes as taken for chromatin assembly analysis were used for subsequent analysis of the recovered chromatin. For ScaI-digestion, again a similar volume of sample was mixed with TEK10 and CutSmart buffer to 1x and 10units of ScaI-HF followed by digestion for 5-7h. Samples of chromatin before and after precipitation and after ScaI digestion were analyzed as described above. Alternatively ScaI digests were analyzed by native PAGE using 0.5% TBE gels. Gels were visualized first in fluorescence channels and then stained with GelRed for visualization of both fluorescing and non-fluorescent DNA and nucleosomal/chromatin bands.

7.7. Ensemble FRET measurements on chromatin

The settings used on the fluorometer for ensemble analysis of the spectra were: 575nm excitation with 4nm slit width and measuring emission in the range 585-700nm with 5nm slit width. For anisotropy measurements the emission slit width was opened to 10nm and measuring uniquely at 592nm. Chromatin samples isolated after magnesium precipitation were diluted to a final volume of typically 220-250µL resulting in a concentration giving a spectral count of around 90000-

130000 for the donor prior to compaction. The sample was then split in 4x50µL volumes. TEK10 and Mg^{2+} from stocks of 10mM or 50mM was added along with TEK10 to a final volume of 55µL 5min prior to measuring. After standing 5min, the sample was transferred to the fluorescence micro-cuvette for measurement of the spectra twice first followed by the anisotropy twice. This was done for all the samples in the range 0-4mM Mg^{2+}.

7.8. Live-cell culturing and confocal fluorescence microscopy

Culturing and transient transfection of NIH 3T3 fibroblasts

Cell culture. NIH/3T3 mouse fibroblasts were cultured in 75 cm^2 tissue culture flasks in DMEM/F-12 supplemented with 10% Newborn calf serum at 37 _C in a water-saturated 5% CO_2 atmosphere. The day before transfection 1.5–2.0·10^5 cells were seeded onto 25-mm round coverslips in six-well plates and grown overnight. On the day of transfection with a 70–90% level of confluency, the cells were transfected with 0.8 mg of plasmid DNA per well using Effectene according to the manufacturer's protocol[349].

Confocal microscopy for cell imaging

Microscopy was performed 24–32 h posttransfection using an inverted LSM 700 confocal microscope and a Plan apochromat x63/1.4 NA objective. The solid-state lasers at 405 and 488 nm were used for excitation. A short pass filter below 490nm and a long pass filter beyond 490nm were used for DAPI and mEos3.2 fluorescence, respectively. Images were acquired in 512_512 pixels with a 0.07 mm pixel size, 12-bit grey-scale depth, line averaging of 4 and a pixel dwell time of 6.30 ms. The laser power was set at 1.5–3.0%, the master gain at 750–800 and the digital gain at 1.5 using a pinhole size of 60.5 mm[349].

Confocal microscopy for FRAP imaging

FRAP bleaching and time-series images were acquired in 128–128 pixels with 0.07 mm pixel size, 12-bit grey-scale depth and a pixel dwell time of 1.58 ms (scan time: 121 ms). Master gain and digital gain were 750–800 and 1.5, respectively, with the pinhole set at 201 mm. A circular spot of 14 pixels (1.04 mm) in diameter was used for bleaching. Twenty prebleach images were acquired before 10 iterations of a bleaching pulse at 80% laser power used and images were acquired for the subsequent 34 s. Photobleaching during the time-series was corrected using the intensity in the bleach region relative to the entire acquisition region. The time-intensity acquisitions were normalized to the pre-bleach intensity and the first image after the bleach pulse. Results are averaged over 30–40 individual FRAP curves for the wild type and the mutants[349].

7.9. Single-molecule TIRFM methods

Cleaning and silanization of borosilicate coverslips and slides

Borosilicate glass slides with 2 rows of 4 holes and borosilicate coverslips were sonicated standing upright in glass containers for 20min in MQ H2O, then in acetone and then in ethanol. They were cleaned in piranha solution (25%v/v H_2O_2 and 75%v/v H_2SO_4) in the same glass containers for 1h followed by washing with MQ H2O until reaching neutral pH again. A 500mL Erlenmeyer flask was cleaned in the same way. The erlenmyer flask, coverslips and slides were all sonicated in acetone for 10min. 3%v/v aminopropyltriethylsilane was prepared in acetone in the Erlenmeyer flask and

used to immerse the microscopy glass and incubate 20min. The aminosilane was disposed, the slides washed in water and dried with N_2. Flow-chambers were assembled from one glass slide and one coverslip separated by double-sided 0.12mm tape positioned between each hole in the glass slide. The ends were sealed with epoxy glue and the silanized slides stored under vacuum in the freezer until use.

Preparation of microfluidic channels and surface PEGylation

Silanized glass flow chambers stored in the freezer were allowed to cool off for 20-30min. Then pipette tip inlet reservoir and outlet sources were neatly fitted in each of the 2x4 holes on each side of the flow chamber and glued in place with epoxy. This was allowed to dry for 30-40min. Subsequently 350µL of 0.1M tetraborate buffer at pH 8.5 was used to dissolve ~1mg of biotin-mPEG(5000kDa)-SVA(SC for CoSM project), and 175µL from this was transferred to 20mg mPEG(5000kDa)-SVA(SC) to generate a transparent clouding-point solution after 10s centrifugation. This was mixed to homogeneity with a pipette and centrifuged again for 10s before 40-45µL aliquots were loaded into each of the four channels in the flow chamber. The PEGylation reaction was allowed to continue for the next 2½-4h whereafter this was washed out with degassed ultra-pure water.

CoSM with Nikon Ti-E microscope

The channels were extensively washed using water and buffer T50 (10mM Tris, 50mM KCl) buffer before proceeding to experiments. For chromatin immobilization, 0.2 mg ml_1 neutravidin solution was infused using a high-precision syringe pump and incubated for 5min, followed by extensive washes with T50. Then, 500 pM chromatin arrays in T50 buffer were injected into the neutravidin treated flow chamber for 5min, followed by a wash with T50 and imaging buffer (50mM HEPES, 130mM KCl, 10% glycerol, 2mM trolox, 0.005% tween-20, 3.2% glucose, 1x GODCAT). Chromatin coverage was observed in the TIRF microscope by fluorescent emission in the far-red channel upon excitation by a 640-nm laser line. Dynamic experiments were initiated by infusion of 1 nM HP1α(Atto532) or 0.5nM HP1α_{cdm} in imaging buffer. All proteins were freshly diluted from a 100nM stock and immediately injected to avoid changes in concentration due to adsorption to tube walls. HP1α dynamics were observed in the yellow/orange channel using a 530 nm laser line for excitation at 20 W/cm^2 using an EMCCD camera (Andor iXon) at 20 frames/s for a 25x50 mm area at a resolution of 160 nm/pixels. Every 200 frames, an image of the chromatin positions in the far-red channel was recorded for drift correction[349].

CoSM data analysis

Movies of HP1α-chromatin interactions were recorded and converted into individual traces using a custom-made semi-automated Matlab (Mathworks) script: in short, after a global baseline-correction and drift-correction (using the far-red images interspersed at 200 frame intervals), individual chromatin array positions were determined using a peak-finding algorithm from the far-red images. Fluorescence intensity traces for each chromatin position were obtained by integrating over a circle of 2 pixel radius. Individual HP1 fluorescence peaks were included based on point-spread-function (PSF) and distance cutoffs. Kinetics were extracted from fluorescence traces using a semi-automated thresholding algorithm. Cumulative histograms were constructed from dark and bright intervals and fitted to sums of exponential functions in Origin[349].

smFRET measurements with micro-mirror TIRF microscope

smFRET measurements were carried out with a micro-mirror TIRF system. The axial focus position was coupled to TIRF evanescent wave formation and maintenance through the reflected beam targeted on a 4-quadrant photodiode. Buffers

with/without biomolecules were deposited in the inlet reservoir and drawn into the chamber with tubing connected from the outlet to a 1mL syringe operated manually or with a motor-driven syringe pump.

After deposition of immersion oil, the slide was loaded onto the microscope and the focus identified by adjustment of the axial stage position until reaching the point of TIR. Calibration of z-stage movement with the axial position allowed subsequent locking of the TIRF position. For each experiment, the channel to be used was washed with 200-300µL T50, followed by incubation with 50µL 0.2mg/mL neutravidin for 5min. This was washed out with another 400-500µL T50. Then 0.5-2µL of chromatin assembly reaction at a concentration of 5-40ng/µL (Measured on total DNA concentration) was loaded into the chamber while monitoring acceptor emission, to assess chromatin coverage. If the density was too low, more chromatin was loaded until reaching 150-400 chromatin arrays in a 25x50µm field of view. Excess chromatin was washed out with T50 followed by exchange to imaging buffer (40mM KCl, 50mM Tris, 2mM Trolox, 2mM NBA, 2mM COT, 10% glycerol and 3.2% glucose) supplemented with 100x GODCAT (165U/mL glucose oxidase, 2170U/mL catalase) to 1x. Before proceeding to imaging, the focus was re-adjusted and the lateral position changed to a new region.

The programmed sequence to be used for imaging was specified in the labview program and the number of repeats for this set to 1. Typically, this entailed moving to a new position that would be used for the entire image acquisition. In this area, 640nm excitation was done for 4000ms with the TIRF lock on to ensure the focus. Then the camera was triggered to acquire 1300-2000 frames with 100ms time-resolution, the TIRF lock turned off and images acquired for 1000ms with 640nm excitation. With the TIRF lock remaining off, excitation was changed to 532nm followed by continued acquisition for 12500-18500ms followed by a final change to 640nm excitation acquired until the end of the kinetic acquisition series programmed by the camera. For sequences requiring timed programmed injection, an additional 5000ms segment was interjected prior to the donor acquisition, followed by automated triggering of the pump and manual stopping after injection of at least 80µL.

For each image acquisition, the file name and file destination was changed to describe the video taken. The live monitoring of the field of view was turned on and the view changed to a previously non-imaged region similar to the new position specified with the programmed sequence above. After changing to the new region, it was checked that the axial position recovered to the setting locked on to with the 4-quadrant photodiode and shortly thereafter image acquisition was started along with the programmed sequence described above.

For experiments with magnesium and HP1α, the mixture with the desired concentration was prepared and loaded into the inlet reservoir followed by injection into the channel and imaging as described above.

Image processing, selection of peaks and trace selection for smFRET

The acquired movies were initially processed in ImageJ. The stack of donor excitation, donor emission images were duplicated once for the raw images and once to generate the background with a rolling ball radius of 20 pixel. Stacks for the acceptor were generated from concatenation of a single image from 640nm excitation prior to 532nm excitation, followed by the long stack upon 532nm excitation and finally 5-20 images of 640nm excitation after the long stack. Background for these were generated as for the donor.

Trace extraction and analysis was done in custom-written matlab software. Donor and acceptor image stacks and corresponding backgrounds from ImageJ were loaded and background subtracted from the donor and the acceptor

channels. The donor and the acceptor images were aligned using 8-10 sets of peaks appearing in both the donor and the acceptor channels. Peaks were automatically detected in the initial acceptor image prior to donor excitation and the same peaks were selected in the donor channel. Peaks tightly clustered, close to the edges or above a set intensity threshold in either the donor or the acceptor channels indicating aggregation were removed from analysis. The analysis was then limited to the peaks appearing in both the donor and the acceptor channel and these traces were extracted for further analysis.

Traces were selected based on the following criteria: 1) An initial total fluorescence of the donor and the β-corrected acceptor of >600. 2) At least 5s prior to bleaching of acceptor or donor. 3) Single catastrophic event for donor or acceptor. i) if acceptor first; leads to anticorrelated increase in donor to same total fluorescence level as prior to bleaching. ii) if donor first; acceptor is still present in final images upon direct excitation at end of stack. 4) Bleaching of the donor during the 120s of acquisition for final determination of background levels.

8. References

1 Andrews, A. J. & Luger, K. Nucleosome structure(s) and stability: variations on a theme. *Annual review of biophysics* **40**, 99-117, (2011).

2 Kornberg, R. D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-871, (1974).

3 Kornberg, R. D. & Thomas, J. O. Chromatin structure; oligomers of the histones. *Science* **184**, 865-868, (1974).

4 Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285-294, (1999).

5 Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. Structure of the nucleosome core particle at 7 A resolution. *Nature* **311**, 532-537, (1984).

6 Zhou, B. R. *et al.* Structural Mechanisms of Nucleosome Recognition by Linker Histones. *Molecular cell* **59**, 628-638, (2015).

7 DeLange, R. J., Fambrough, D. M., Smith, E. L. & Bonner, J. Calf and pea histone IV. 3. Complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone. *The Journal of biological chemistry* **244**, 5669-5679, (1969).

8 Thatcher, T. H. & Gorovsky, M. A. Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic acids research* **22**, 174-179, (1994).

9 Luger, K., Dechassa, M. L. & Tremethick, D. J. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nature reviews. Molecular cell biology* **13**, 436-447, (2012).

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

11 Bonisch, C. & Hake, S. B. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic acids research* **40**, 10719-10741, (2012).

12 Wiedemann, S. M. *et al.* Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *The Journal of cell biology* **190**, 777-791, (2010).

13 Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705, (2007).

14 Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *Journal of molecular biology* **319**, 1097-1113, (2002).

15 Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423**, 145-150, (2003).

16 Vasudevan, D., Chua, E. Y. & Davey, C. A. Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *Journal of molecular biology* **403**, 1-10, (2010).

17 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).

18 Thastrom, A. *et al.* Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *Journal of molecular biology* **288**, 213-229, (1999).

19 Visnapuu, M. L. & Greene, E. C. Single-molecule imaging of DNA curtains reveals intrinsic energy landscapes for nucleosome deposition. *Nature structural & molecular biology* **16**, 1056-1062, (2009).

20 Simpson, R. T. & Stafford, D. W. Structural features of a phased nucleosome core particle. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 51-55, (1983).
21 Segal, E. *et al.* A genomic code for nucleosome positioning. *Nature* **442**, 772-778, (2006).
22 Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362-366, (2009).
23 Alabert, C. & Groth, A. Chromatin replication and epigenome maintenance. *Nature reviews. Molecular cell biology* **13**, 153-167, (2012).
24 Duro, E. *et al.* Identification of the MMS22L-TONSL complex that promotes homologous recombination. *Molecular cell* **40**, 632-644, (2010).
25 Natsume, R. *et al.* Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature* **446**, 338-341, (2007).
26 Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**, 51-61, (2004).
27 Lewis, P. W., Elsaesser, S. J., Noh, K. M., Stadler, S. C. & Allis, C. D. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14075-14080, (2010).
28 Elsasser, S. J. *et al.* DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition. *Nature* **491**, 560-565, (2012).
29 Belotserkovskaya, R. *et al.* FACT facilitates transcription-dependent nucleosome alteration. *Science* **301**, 1090-1093, (2003).
30 Hondele, M. *et al.* Structural basis of histone H2A-H2B recognition by the essential chaperone FACT. *Nature* **499**, 111-114, (2013).
31 Aguilar-Gurrieri, C. *et al.* Structural evidence for Nap1-dependent H2A-H2B deposition and nucleosome assembly. *The EMBO journal* **35**, 1465-1482, (2016).
32 Park, Y. J., Chodaparambil, J. V., Bao, Y., McBryant, S. J. & Luger, K. Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. *The Journal of biological chemistry* **280**, 1817-1825, (2005).
33 Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annual review of biochemistry* **78**, 273-304, (2009).
34 Krietenstein, N. *et al.* Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* **167**, 709-721 e712, (2016).
35 Phair, R. D. & Misteli, T. High mobility of proteins in the mammalian cell nucleus. *Nature* **404**, 604-609, (2000).
36 Misteli, T., Gunjan, A., Hock, R., Bustin, M. & Brown, D. T. Dynamic binding of histone H1 to chromatin in living cells. *Nature* **408**, 877-881, (2000).
37 Li, G. & Widom, J. Nucleosomes facilitate their own invasion. *Nature structural & molecular biology* **11**, 763-769, (2004).
38 Li, G., Levitus, M., Bustamante, C. & Widom, J. Rapid spontaneous accessibility of nucleosomal DNA. *Nature structural & molecular biology* **12**, 46-53, (2005).
39 Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M. & Richmond, T. J. Characterization of nucleosome core particles containing histone proteins made in bacteria. *Journal of molecular biology* **272**, 301-311, (1997).
40 Park, Y. J., Dyer, P. N., Tremethick, D. J. & Luger, K. A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome. *The Journal of biological chemistry* **279**, 24274-24282, (2004).

41 Dyer, P. N. *et al.* Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods in enzymology* **375**, 23-44, (2004).
42 Thastrom, A., Gottesfeld, J. M., Luger, K. & Widom, J. Histone-DNA binding free energy cannot be measured in dilution-driven dissociation experiments. *Biochemistry* **43**, 736-741, (2004).
43 Bohm, V. *et al.* Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic acids research* **39**, 3093-3102, (2011).
44 Gansen, A. *et al.* Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 15308-15313, (2009).
45 Risca, V. I. & Greenleaf, W. J. Unraveling the 3D genome: genomics tools for multiscale exploration. *Trends in genetics : TIG* **31**, 357-372, (2015).
46 Wu, T. P. *et al.* DNA methylation on N(6)-adenine in mammalian embryonic stem cells. *Nature* **532**, 329-333, (2016).
47 Olins, A. L. & Olins, D. E. Spheroid chromatin units (v bodies). *Science* **183**, 330-332, (1974).
48 Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75-82, (2012).
49 Finch, J. T. & Klug, A. Solenoidal model for superstructure in chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 1897-1901, (1976).
50 Thoma, F., Koller, T. & Klug, A. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *The Journal of cell biology* **83**, 403-427, (1979).
51 Worcel, A., Strogatz, S. & Riley, D. Structure of chromatin and the linking number of DNA. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 1461-1465, (1981).
52 Simpson, R. T., Thoma, F. & Brubaker, J. M. Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: a model system for study of higher order structure. *Cell* **42**, 799-808, (1985).
53 Hansen, J. C., Ausio, J., Stanik, V. H. & van Holde, K. E. Homogeneous reconstituted oligonucleosomes, evidence for salt-dependent folding in the absence of histone H1. *Biochemistry* **28**, 9129-9136, (1989).
54 Hansen, J. C. Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annual review of biophysics and biomolecular structure* **31**, 361-392, (2002).
55 Dorigo, B., Schalch, T., Bystricky, K. & Richmond, T. J. Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *Journal of molecular biology* **327**, 85-96, (2003).
56 Shogren-Knaak, M. *et al.* Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844-847, (2006).
57 Dorigo, B. *et al.* Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571-1573, (2004).
58 Schalch, T., Duda, S., Sargent, D. F. & Richmond, T. J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**, 138-141, (2005).
59 Song, F. *et al.* Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **344**, 376-380, (2014).
60 Li, W. *et al.* FACT Remodels the Tetranucleosomal Unit of Chromatin Fibers for Gene Transcription. *Molecular cell*, (2016).

61 Robinson, P. J., Fairall, L., Huynh, V. A. & Rhodes, D. EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6506-6511, (2006).
62 Kruithof, M. *et al.* Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nature structural & molecular biology* **16**, 534-540, (2009).
63 Grigoryev, S. A., Arya, G., Correll, S., Woodcock, C. L. & Schlick, T. Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13317-13322, (2009).
64 Hsieh, T. H. *et al.* Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* **162**, 108-119, (2015).
65 Ricci, M. A., Manzo, C., Garcia-Parajo, M. F., Lakadamyali, M. & Cosma, M. P. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* **160**, 1145-1158, (2015).
66 Grigoryev, S. A. *et al.* Hierarchical looping of zigzag nucleosome chains in metaphase chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 1238-1243, (2016).
67 Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293, (2009).
68 Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature reviews. Genetics* **14**, 390-403, (2013).
69 Fierz, B. *et al.* Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nature chemical biology* **7**, 113-119, (2011).
70 Azzaz, A. M. *et al.* Human heterochromatin protein 1alpha promotes nucleosome associations that drive chromatin condensation. *The Journal of biological chemistry* **289**, 6850-6861, (2014).
71 Poirier, M. G., Oh, E., Tims, H. S. & Widom, J. Dynamics and function of compact nucleosome arrays. *Nature structural & molecular biology* **16**, 938-944, (2009).
72 Buning, R., Kropff, W., Martens, K. & van Noort, J. spFRET reveals changes in nucleosome breathing by neighboring nucleosomes. *Journal of physics. Condensed matter : an Institute of Physics journal* **27**, 064103, (2015).
73 Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306-1311, (2002).
74 Dekker, J. The three 'C' s of chromosome conformation capture: controls, controls, controls. *Nature methods* **3**, 17-21, (2006).
75 Zhao, Z. *et al.* Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature genetics* **38**, 1341-1347, (2006).
76 Simonis, M. *et al.* Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature genetics* **38**, 1348-1354, (2006).
77 Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research* **16**, 1299-1309, (2006).
78 Hsieh, T. S., Fudenberg, G., Goloborodko, A. & Rando, O. J. Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome. *Nature methods* **13**, 1009-1011, (2016).

79 Dekker, J. & Misteli, T. Long-Range Chromatin Interactions. *Cold Spring Harbor perspectives in biology* **7**, a019356, (2015).
80 Goloborodko, A., Marko, J. F. & Mirny, L. A. Chromosome Compaction by Active Loop Extrusion. *Biophysical journal* **110**, 2162-2168, (2016).
81 Goloborodko, A., Imakaev, M. V., Marko, J. F. & Mirny, L. Compaction and segregation of sister chromatids via active loop extrusion. *eLife* **5**, (2016).
82 Lupianez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012-1025, (2015).
83 Franke, M. *et al.* Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265-269, (2016).
84 Boyle, S. *et al.* The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Human molecular genetics* **10**, 211-219, (2001).
85 Heard, E. & Martienssen, R. A. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* **157**, 95-109, (2014).
86 Audergon, P. N. *et al.* Epigenetics. Restricted epigenetic inheritance of H3K9 methylation. *Science* **348**, 132-135, (2015).
87 Ragunathan, K., Jih, G. & Moazed, D. Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* **348**, 1258699, (2015).
88 Roseboom, T., de Rooij, S. & Painter, R. The Dutch famine and its long-term consequences for adult health. *Early human development* **82**, 485-491, (2006).
89 Bell, J. T. & Spector, T. D. A twin approach to unraveling epigenetics. *Trends in genetics : TIG* **27**, 116-125, (2011).
90 Brownell, J. E. *et al.* Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843-851, (1996).
91 Kuo, M. H. *et al.* Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* **383**, 269-272, (1996).
92 Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641-643, (1996).
93 Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408-411, (1996).
94 Andrews, F. H., Strahl, B. D. & Kutateladze, T. G. Insights into newly discovered marks and readers of epigenetic information. *Nature chemical biology* **12**, 662-668, (2016).
95 Kee, J. M., Oslund, R. C., Perlman, D. H. & Muir, T. W. A pan-specific antibody for direct detection of protein histidine phosphorylation. *Nature chemical biology* **9**, 416-421, (2013).
96 Jbara, M., Maity, S. K., Morgan, M., Wolberger, C. & Brik, A. Chemical Synthesis of Phosphorylated Histone H2A at Tyr57 Reveals Insight into the Inhibition Mode of the SAGA Deubiquitinating Module. *Angewandte Chemie* **55**, 4972-4976, (2016).
97 Wike, C. L. *et al.* Aurora-A mediated histone H3 phosphorylation of threonine 118 controls condensin I and cohesin occupancy in mitosis. *eLife* **5**, e11402, (2016).
98 Tan, M. *et al.* Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016-1028, (2011).
99 Sabari, B. R. *et al.* Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Molecular cell* **58**, 203-215, (2015).
100 Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41-45, (2000).
101 Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074-1080, (2001).
102 Filion, G. J. *et al.* Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **143**, 212-224, (2010).

103 Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature biotechnology* **28**, 817-825, (2010).
104 Brehove, M. *et al.* Histone core phosphorylation regulates DNA accessibility. *The Journal of biological chemistry* **290**, 22612-22621, (2015).
105 Simon, M. *et al.* Histone fold modifications control nucleosome unwrapping and disassembly. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 12711-12716, (2011).
106 Di Cerbo, V. *et al.* Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. *eLife* **3**, e01632, (2014).
107 Neumann, H. *et al.* A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Molecular cell* **36**, 153-163, (2009).
108 Manohar, M. *et al.* Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. *The Journal of biological chemistry* **284**, 23312-23321, (2009).
109 Tropberger, P. *et al.* Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* **152**, 859-872, (2013).
110 Li, W., Nagaraja, S., Delcuve, G. P., Hendzel, M. J. & Davie, J. R. Effects of histone acetylation, ubiquitination and variants on nucleosome stability. *The Biochemical journal* **296 (Pt 3)**, 737-744, (1993).
111 Davies, N. & Lindsey, G. G. Histone H2B (and H2A) ubiquitination allows normal histone octamer and core particle reconstitution. *Biochimica et biophysica acta* **1218**, 187-193, (1994).
112 Chandrasekharan, M. B., Huang, F. & Sun, Z. W. Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 16686-16691, (2009).
113 Lan, L. *et al.* Monoubiquitinated histone H2A destabilizes photolesion-containing nucleosomes with concomitant release of UV-damaged DNA-binding protein E3 ligase. *The Journal of biological chemistry* **287**, 12036-12049, (2012).
114 Fierz, B., Kilic, S., Hieb, A. R., Luger, K. & Muir, T. W. Stability of nucleosomes containing homogenously ubiquitylated H2A and H2B prepared using semisynthesis. *Journal of the American Chemical Society* **134**, 19548-19551, (2012).
115 Dhall, A. *et al.* Sumoylated human histone H4 prevents chromatin compaction by inhibiting long-range internucleosomal interactions. *The Journal of biological chemistry* **289**, 33827-33837, (2014).
116 Fierz, B. & Muir, T. W. Chromatin as an expansive canvas for chemical biology. *Nature chemical biology* **8**, 417-427, (2012).
117 Masumoto, H., Hawke, D., Kobayashi, R. & Verreault, A. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* **436**, 294-298, (2005).
118 Saredi, G. *et al.* H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex. *Nature* **534**, 714-718, (2016).
119 Kubicek, S. & Jenuwein, T. A crack in histone lysine methylation. *Cell* **119**, 903-906, (2004).
120 Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-953, (2004).
121 Tsukada, Y. *et al.* Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439**, 811-816, (2006).
122 Cloos, P. A. *et al.* The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* **442**, 307-311, (2006).
123 Bannister, A. J. & Kouzarides, T. Reversing histone methylation. *Nature* **436**, 1103-1106, (2005).

124 Musselman, C. A., Lalonde, M. E., Cote, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nature structural & molecular biology* **19**, 1218-1227, (2012).
125 Patel, D. J. & Wang, Z. Readout of epigenetic modifications. *Annual review of biochemistry* **82**, 81-118, (2013).
126 Morgan, M. T. *et al.* Structural basis for histone H2B deubiquitination by the SAGA DUB module. *Science* **351**, 725-728, (2016).
127 Wilson, M. D. *et al.* The structural basis of modified nucleosome recognition by 53BP1. *Nature* **536**, 100-103, (2016).
128 Ruthenburg, A. J., Li, H., Patel, D. J. & Allis, C. D. Multivalent engagement of chromatin modifications by linked binding modules. *Nature reviews. Molecular cell biology* **8**, 983-994, (2007).
129 Su, Z. & Denu, J. M. Reading the Combinatorial Histone Language. *ACS chemical biology* **11**, 564-574, (2016).
130 Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422-1425, (2000).
131 Kleiner, R. E., Verma, P., Molloy, K. R., Chait, B. T. & Kapoor, T. M. Chemical proteomics reveals a gammaH2AX-53BP1 interaction in the DNA damage response. *Nature chemical biology* **11**, 807-814, (2015).
132 Ruthenburg, A. J. *et al.* Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145**, 692-706, (2011).
133 Harrison, J. S. *et al.* Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *eLife* **5**, (2016).
134 Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574-1577, (2004).
135 Hiragami-Hamada, K. *et al.* Dynamic and flexible H3K9me3 bridging via HP1beta dimerization establishes a plastic state of condensed chromatin. *Nature communications* **7**, 11310, (2016).
136 Cornett, E. M. *et al.* Substrate Specificity Profiling of Histone-Modifying Enzymes by Peptide Microarray. *Methods in enzymology* **574**, 31-52, (2016).
137 Fuchs, S. M., Krajewski, K., Baker, R. W., Miller, V. L. & Strahl, B. D. Influence of combinatorial histone modifications on antibody and effector protein recognition. *Current biology : CB* **21**, 53-58, (2011).
138 Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature structural & molecular biology* **14**, 1025-1040, (2007).
139 Pick, H., Kilic, S. & Fierz, B. Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function. *Biochimica et biophysica acta* **1839**, 644-656, (2014).
140 Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *Annual review of biochemistry* **79**, 413-444, (2010).
141 Neumann, H., Peak-Chew, S. Y. & Chin, J. W. Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nature chemical biology* **4**, 232-234, (2008).
142 Nguyen, D. P., Garcia Alai, M. M., Kapadnis, P. B., Neumann, H. & Chin, J. W. Genetically encoding N(epsilon)-methyl-L-lysine in recombinant histones. *Journal of the American Chemical Society* **131**, 14194-14195, (2009).
143 Nguyen, D. P., Garcia Alai, M. M., Virdee, S. & Chin, J. W. Genetically directing varepsilon-N, N-dimethyl-L-lysine in recombinant histones. *Chemistry & biology* **17**, 1072-1076, (2010).

144 Virdee, S. *et al.* Traceless and site-specific ubiquitination of recombinant proteins. *Journal of the American Chemical Society* **133**, 10708-10711, (2011).
145 Simon, M. D. *et al.* The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **128**, 1003-1012, (2007).
146 Le, D. D., Cortesi, A. T., Myers, S. A., Burlingame, A. L. & Fujimori, D. G. Site-specific and regiospecific installation of methylarginine analogues into recombinant histones and insights into effector protein binding. *Journal of the American Chemical Society* **135**, 2879-2882, (2013).
147 Li, F. *et al.* A direct method for site-specific protein acetylation. *Angewandte Chemie* **50**, 9611-9614, (2011).
148 Chatterjee, C., McGinty, R. K., Fierz, B. & Muir, T. W. Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation. *Nature chemical biology* **6**, 267-269, (2010).
149 Chalker, J. M., Lercher, L., Rose, N. R., Schofield, C. J. & Davis, B. G. Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications. *Angewandte Chemie* **51**, 1835-1839, (2012).
150 Bernardes, G. J., Chalker, J. M., Errey, J. C. & Davis, B. G. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins. *Journal of the American Chemical Society* **130**, 5052-5053, (2008).
151 Park, H. S. *et al.* Expanding the genetic code of Escherichia coli with phosphoserine. *Science* **333**, 1151-1154, (2011).
152 Wright, T. H. *et al.* Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. *Science* **354**, (2016).
153 Wright, T. H., Vallee, M. R. & Davis, B. G. From Chemical Mutagenesis to Post-Expression Mutagenesis: A 50 Year Odyssey. *Angewandte Chemie* **55**, 5896-5903, (2016).
154 Yang, A. *et al.* A chemical biology route to site-specific authentic protein modifications. *Science* **354**, 623-626, (2016).
155 Schnolzer, M., Alewood, P., Jones, A., Alewood, D. & Kent, S. B. In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *International journal of peptide and protein research* **40**, 180-193, (1992).
156 Schnolzer, M. & Kent, S. B. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* **256**, 221-225, (1992).
157 Englebretsen, D. R., Garnham, B. C., Bergman, D. A. & Alewood, P. F. A novel thioether linker: Chemical synthesis of a HIV-1 protease analogue by thioether ligation. *Tetrahedron letters* **36**, 8871-8874, (1995).
158 Zhang, L. & Tam, J. P. Orthogonai Coupling of Unprotected Peptide Segments through Histidyl Amino Terminus. *Tetrahedron letters* **38**, 3-6, (1997).
159 Liu, C. F. & Tam, J. P. Chemical Ligation Approach To Form a Peptide Bond between Unprotected Peptide Segments. Concept and Model Study. *Journal of the American Chemical Society* **116**, 4149-4153, (1994).
160 Nilsson, B. L., Kiessling, L. L. & Raines, R. T. Staudinger ligation: a peptide from a thioester and azide. *Organic letters* **2**, 1939-1941, (2000).
161 Saxon, E., Armstrong, J. I. & Bertozzi, C. R. A "traceless" Staudinger ligation for the chemoselective synthesis of amide bonds. *Organic letters* **2**, 2141-2143, (2000).
162 Bode, J. W., Fox, R. M. & Baucom, K. D. Chemoselective amide ligations by decarboxylative condensations of N-alkylhydroxylamines and alpha-ketoacids. *Angewandte Chemie* **45**, 1248-1252, (2006).

163 Pusterla, I. & Bode, J. W. An oxazetidine amino acid for chemical protein synthesis by rapid, serine-forming ligations. *Nature chemistry* **7**, 668-672, (2015).
164 Rose, K. Facile Synthesis of Homogeneous Artificial Proteins. *Journal of the American Chemical Society* **116**, 30-33, (1994).
165 Zhang, Y., Xu, C., Lam, H. Y., Lee, C. L. & Li, X. Protein chemical synthesis by serine and threonine ligation. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6657-6662, (2013).
166 Tornoe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(i)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *The Journal of organic chemistry* **67**, 3057-3064, (2002).
167 Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779, (1994).
168 Hackenberger, C. P. & Schwarzer, D. Chemoselective ligation and modification strategies for peptides and proteins. *Angewandte Chemie* **47**, 10030-10074, (2008).
169 Bondalapati, S., Jbara, M. & Brik, A. Expanding the chemical toolbox for the synthesis of large and uniquely modified proteins. *Nature chemistry* **8**, 407-418, (2016).
170 Dawson, P. E. & Kent, S. B. Synthesis of native proteins by chemical ligation. *Annual review of biochemistry* **69**, 923-960, (2000).
171 Tam, J. P., Lu, Y. A., Liu, C. F. & Shao, J. Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 12485-12489, (1995).
172 Kimmerlin, T. & Seebach, D. '100 years of peptide synthesis': ligation methods for peptide and protein synthesis with applications to beta-peptide assemblies. *The journal of peptide research : official journal of the American Peptide Society* **65**, 229-260, (2005).
173 Hojo, H. & Aimoto, S. Polypeptide synthesis using the S-alkyl thioester of a partially protected peptide segment. Synthesis of the DNA-binding domain of c-Myb protein (142-193)-NH2. *Bull Chem Soc Jpn* **64**, 111-117, (1991).
174 Camarero, J. A., Adeva, A. & Muir, T. W. 3-Thiopropionic acid as a highly versatile multidetachable thioester resin linker. *Lett Pept Sci* **7**, 17-21, (2000).
175 Clippingdale, A. B., Barrow, C. J. & Wade, J. D. Peptide thioester preparation by Fmoc solid phase peptide synthesis for use in native chemical ligation. *Journal of peptide science : an official publication of the European Peptide Society* **6**, 225-234, (2000).
176 Li, X., Kawakami, T. & Aimoto, S. Direct preparation of peptide thioesters using an Fmoc solid-phase method. *Tetrahedron letters* **39**, 8669-8672, (1998).
177 Muralidharan, V. & Muir, T. W. Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nature methods* **3**, 429-438, (2006).
178 Blanco-Canosa, J. B. & Dawson, P. E. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angewandte Chemie* **47**, 6851-6855, (2008).
179 Hojo, H., Onuma, Y., Akimoto, Y., Nakahara, Y. & Nakahara, Y. N-alkyl cysteine-assisted thioesterification of peptides. *Tetrahedron letters* **48**, 25-28, (2007).
180 Tsuda, S., Shigenaga, A., Bando, K. & Otaka, A. N-->S acyl-transfer-mediated synthesis of peptide thioesters using anilide derivatives. *Organic letters* **11**, 823-826, (2009).
181 Erlich, L. A., Kumar, K. S., Haj-Yahya, M., Dawson, P. E. & Brik, A. N-methylcysteine-mediated total chemical synthesis of ubiquitin thioester. *Organic & biomolecular chemistry* **8**, 2392-2396, (2010).
182 Hou, W., Zhang, X. H., Li, F. P. & Liu, C. F. Peptidyl N,N-bis(2-mercaptoethyl)-amides as thioester precursors for native chemical ligation. *Organic letters* **13**, 386-389, (2011).

183 Ollivier, N. *et al.* Tidbits for the synthesis of bis(2-sulfanylethyl)amido (SEA) polystyrene resin, SEA peptides and peptide thioesters. *Journal of peptide science : an official publication of the European Peptide Society* **20**, 92-97, (2014).
184 Fang, G. M. *et al.* Protein chemical synthesis by ligation of peptide hydrazides. *Angewandte Chemie* **50**, 7645-7649, (2011).
185 Stavropoulos, G., Gatos, D., Magafa, V. & Barlos, K. Preparation of polymer-bound trityl-hydrazines and their application in the solid phase synthesis of partially protected peptide hydrazides. *Lett Pept Sci* **2**, 315-318, (1997).
186 Zheng, J. S., Tang, S., Qi, Y. K., Wang, Z. P. & Liu, L. Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. *Nature protocols* **8**, 2483-2495, (2013).
187 Fang, G. M., Wang, J. X. & Liu, L. Convergent chemical synthesis of proteins by ligation of peptide hydrazides. *Angewandte Chemie* **51**, 10347-10350, (2012).
188 Low, D. W., Hill, M. G., Carrasco, M. R., Kent, S. B. & Botti, P. Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 6554-6559, (2001).
189 Offer, J., Boddy, C. N. & Dawson, P. E. Extending synthetic access to proteins with a removable acyl transfer auxiliary. *Journal of the American Chemical Society* **124**, 4642-4646, (2002).
190 Marinzi, C., Offer, J., Longhi, R. & Dawson, P. E. An o-nitrobenzyl scaffold for peptide ligation: synthesis and applications. *Bioorganic & medicinal chemistry* **12**, 2749-2757, (2004).
191 Kawakami, T. & Aimoto, S. A photoremovable ligation auxiliary for use in polypeptide synthesis. *Tetrahedron letters* **44**, 6059-6061, (2003).
192 Yan, L. Z. & Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *Journal of the American Chemical Society* **123**, 526-533, (2001).
193 Pentelute, B. L. & Kent, S. B. Selective desulfurization of cysteine in the presence of Cys(Acm) in polypeptides obtained by native chemical ligation. *Organic letters* **9**, 687-690, (2007).
194 Wan, Q. & Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angewandte Chemie* **46**, 9248-9252, (2007).
195 Haase, C., Rohde, H. & Seitz, O. Native chemical ligation at valine. *Angewandte Chemie* **47**, 6807-6810, (2008).
196 Crich, D. & Banerjee, A. Native chemical ligation at phenylalanine. *Journal of the American Chemical Society* **129**, 10064-10065, (2007).
197 Siman, P., Karthikeyan, S. V. & Brik, A. Native chemical ligation at glutamine. *Organic letters* **14**, 1520-1523, (2012).
198 Chen, J., Wang, P., Zhu, J., Wan, Q. & Danishefsky, S. J. A program for ligation at threonine sites: application to the controlled total synthesis of glycopeptides. *Tetrahedron* **66**, 2277-2283, (2010).
199 Yang, R., Pasunooti, K. K., Li, F., Liu, X. W. & Liu, C. F. Dual native chemical ligation at lysine. *Journal of the American Chemical Society* **131**, 13592-13593, (2009).
200 Harpaz, Z., Siman, P., Kumar, K. S. & Brik, A. Protein synthesis assisted by native chemical ligation at leucine. *Chembiochem : a European journal of chemical biology* **11**, 1232-1235, (2010).
201 Tan, Z., Shang, S. & Danishefsky, S. J. Insights into the finer issues of native chemical ligation: an approach to cascade ligations. *Angewandte Chemie* **49**, 9500-9503, (2010).

202 Cergol, K. M., Thompson, R. E., Malins, L. R., Turner, P. & Payne, R. J. One-pot peptide ligation-desulfurization at glutamate. *Organic letters* **16**, 290-293, (2014).
203 Thompson, R. E., Chan, B., Radom, L., Jolliffe, K. A. & Payne, R. J. Chemoselective peptide ligation-desulfurization at aspartate. *Angewandte Chemie* **52**, 9723-9727, (2013).
204 Malins, L. R., Cergol, K. M. & Payne, R. J. Chemoselective sulfenylation and peptide ligation at tryptophan. *Chemical science* **5**, 266, (2014).
205 Malins, L. R., Cergol, K. M. & Payne, R. J. Peptide ligation-desulfurization chemistry at arginine. *Chembiochem : a European journal of chemical biology* **14**, 559-563, (2013).
206 Shang, S., Tan, Z., Dong, S. & Danishefsky, S. J. An advance in proline ligation. *Journal of the American Chemical Society* **133**, 10784-10786, (2011).
207 Townsend, S. D. *et al.* Advances in proline ligation. *Journal of the American Chemical Society* **134**, 3912-3916, (2012).
208 Ding, H., Shigenaga, A., Sato, K., Morishita, K. & Otaka, A. Dual kinetically controlled native chemical ligation using a combination of sulfanylproline and sulfanylethylanilide peptide. *Organic letters* **13**, 5588-5591, (2011).
209 Chen, J., Wan, Q., Yuan, Y., Zhu, J. & Danishefsky, S. J. Native chemical ligation at valine: a contribution to peptide and glycopeptide synthesis. *Angewandte Chemie* **47**, 8521-8524, (2008).
210 Ajish Kumar, K. S., Haj-Yahya, M., Olschewski, D., Lashuel, H. A. & Brik, A. Highly efficient and chemoselective peptide ubiquitylation. *Angewandte Chemie* **48**, 8090-8094, (2009).
211 Dery, S. *et al.* Insights into the deselinization of selenocysteine into alanine and serine. *Chemical science* **6**, 6207-6212, (2015).
212 Johnson, E. C. & Kent, S. B. Insights into the mechanism and catalysis of the native chemical ligation reaction. *Journal of the American Chemical Society* **128**, 6640-6646, (2006).
213 Siman, P. *et al.* Chemical synthesis and expression of the HIV-1 Rev protein. *Chembiochem : a European journal of chemical biology* **12**, 1097-1104, (2011).
214 Moyal, T., Hemantha, H. P., Siman, P., Refua, M. & Brik, A. Highly efficient one-pot ligation and desulfurization. *Chemical science* **4**, 2496-2501, (2013).
215 Fauvet, B., Butterfield, S. M., Fuks, J., Brik, A. & Lashuel, H. A. One-pot total chemical synthesis of human alpha-synuclein. *Chemical communications* **49**, 9254-9256, (2013).
216 Huang, Y. C. *et al.* Synthesis of l- and d-Ubiquitin by One-Pot Ligation and Metal-Free Desulfurization. *Chemistry* **22**, 7623-7628, (2016).
217 Thompson, R. E. *et al.* Trifluoroethanethiol: an additive for efficient one-pot peptide ligation-desulfurization chemistry. *Journal of the American Chemical Society* **136**, 8161-8164, (2014).
218 Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: a general method for protein engineering. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 6705-6710, (1998).
219 Severinov, K. & Muir, T. W. Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *The Journal of biological chemistry* **273**, 16205-16209, (1998).
220 Erlanson, D. A., Chytil, M. & Verdine, G. L. The leucine zipper domain controls the orientation of AP-1 in the NFAT.AP-1.DNA complex. *Chemistry & biology* **3**, 981-991, (1996).
221 Shah, N. H. & Muir, T. W. Inteins: Nature's Gift to Protein Chemists. *Chemical science* **5**, 446-461, (2014).
222 Mootz, H. D. Split inteins as versatile tools for protein semisynthesis. *Chembiochem : a European journal of chemical biology* **10**, 2579-2589, (2009).

223 Frutos, S., Goger, M., Giovani, B., Cowburn, D. & Muir, T. W. Branched intermediate formation stimulates peptide bond cleavage in protein splicing. *Nature chemical biology* **6**, 527-533, (2010).
224 Liu, Z. *et al.* Structure of the branched intermediate in protein splicing. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8422-8427, (2014).
225 Vila-Perello, M. & Muir, T. W. Biological applications of protein splicing. *Cell* **143**, 191-200, (2010).
226 Xu, M. Q. & Perler, F. B. The mechanism of protein splicing and its modulation by mutation. *The EMBO journal* **15**, 5146-5153, (1996).
227 Chong, S. *et al.* Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271-281, (1997).
228 Muir, T. W. Semisynthesis of proteins by expressed protein ligation. *Annual review of biochemistry* **72**, 249-289, (2003).
229 Evans, T. C., Jr., Benner, J. & Xu, M. Q. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein science : a publication of the Protein Society* **7**, 2256-2264, (1998).
230 Shah, N. H., Eryilmaz, E., Cowburn, D. & Muir, T. W. Naturally split inteins assemble through a "capture and collapse" mechanism. *Journal of the American Chemical Society* **135**, 18673-18681, (2013).
231 Klabunde, T., Sharma, S., Telenti, A., Jacobs, W. R., Jr. & Sacchettini, J. C. Crystal structure of GyrA intein from Mycobacterium xenopi reveals structural basis of protein splicing. *Nature structural biology* **5**, 31-36, (1998).
232 Shah, N. H., Vila-Perello, M. & Muir, T. W. Kinetic control of one-pot trans-splicing reactions by using a wild-type and designed split intein. *Angewandte Chemie* **50**, 6511-6515, (2011).
233 Shah, N. H., Dann, G. P., Vila-Perello, M., Liu, Z. & Muir, T. W. Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. *Journal of the American Chemical Society* **134**, 11338-11341, (2012).
234 Zettler, J., Schutz, V. & Mootz, H. D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS letters* **583**, 909-914, (2009).
235 Vila-Perello, M. *et al.* Streamlined expressed protein ligation using split inteins. *Journal of the American Chemical Society* **135**, 286-292, (2013).
236 Iwai, H. & Pluckthun, A. Circular beta-lactamase: stability enhancement by cyclizing the backbone. *FEBS letters* **459**, 166-172, (1999).
237 Tolbert, T. J. & Wong, C. H. New methods for proteomic research: preparation of proteins with N-terminal cysteines for labeling and conjugation. *Angewandte Chemie* **41**, 2171-2174, (2002).
238 Holt, M. & Muir, T. Application of the protein semisynthesis strategy to the generation of modified chromatin. *Annual review of biochemistry* **84**, 265-290, (2015).
239 Muller, M. M. & Muir, T. W. Histones: at the crossroads of peptide and protein chemistry. *Chemical reviews* **115**, 2296-2349, (2015).
240 Maity, S. K., Jbara, M. & Brik, A. Chemical and semisynthesis of modified histones. *Journal of peptide science : an official publication of the European Peptide Society* **22**, 252-259, (2016).
241 Shogren-Knaak, M. A., Fry, C. J. & Peterson, C. L. A native peptide ligation strategy for deciphering nucleosomal histone modifications. *The Journal of biological chemistry* **278**, 15744-15748, (2003).

242 He, S. *et al.* Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12033-12038, (2003).
243 Chatterjee, C., McGinty, R. K., Pellois, J. P. & Muir, T. W. Auxiliary-mediated site-specific peptide ubiquitylation. *Angewandte Chemie* **46**, 2814-2818, (2007).
244 McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G. & Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453**, 812-816, (2008).
245 Kim, J. *et al.* RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* **137**, 459-471, (2009).
246 McGinty, R. K. *et al.* Structure-activity analysis of semisynthetic nucleosomes: mechanistic insights into the stimulation of Dot1L by ubiquitylated histone H2B. *ACS chemical biology* **4**, 958-968, (2009).
247 Lechner, C. C., Agashe, N. D. & Fierz, B. Traceless Synthesis of Asymmetrically Modified Bivalent Nucleosomes. *Angewandte Chemie* **55**, 2903-2906, (2016).
248 Nguyen, U. T. *et al.* Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nature methods* **11**, 834-840, (2014).
249 Stratmann, S. A. & van Oijen, A. M. DNA replication at the single-molecule level. *Chemical Society reviews* **43**, 1201-1220, (2014).
250 Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642-1645, (2006).
251 Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods* **3**, 793-795, (2006).
252 Schuler, B. Single-molecule FRET of protein structure and dynamics - a primer. *Journal of nanobiotechnology* **11 Suppl 1**, S2, (2013).
253 Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods in enzymology* **475**, 455-514, (2010).
254 Muller, K. P. *et al.* Multiscale analysis of dynamics and interactions of heterochromatin protein 1 by fluorescence fluctuation microscopy. *Biophysical journal* **97**, 2876-2885, (2009).
255 Muller-Ott, K. *et al.* Specificity, propagation, and memory of pericentric heterochromatin. *Molecular systems biology* **10**, 746, (2014).
256 Steffen, P. A., Fonseca, J. P. & Ringrose, L. Epigenetics meets mathematics: towards a quantitative understanding of chromatin biology. *BioEssays : news and reviews in molecular, cellular and developmental biology* **34**, 901-913, (2012).
257 Canzio, D. *et al.* Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Molecular cell* **41**, 67-81, (2011).
258 Canzio, D. *et al.* A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature* **496**, 377-381, (2013).
259 Nishibuchi, G. & Nakayama, J. Biochemical and structural properties of heterochromatin protein 1: understanding its role in chromatin assembly. *Journal of biochemistry* **156**, 11-20, (2014).
260 James, T. C. & Elgin, S. C. Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila melanogaster and its gene. *Molecular and cellular biology* **6**, 3862-3872, (1986).
261 Canzio, D., Larson, A. & Narlikar, G. J. Mechanisms of functional promiscuity by HP1 proteins. *Trends in cell biology* **24**, 377-386, (2014).

262 Vakoc, C. R., Mandat, S. A., Olenchock, B. A. & Blobel, G. A. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Molecular cell* **19**, 381-391, (2005).
263 Hayakawa, T., Haraguchi, T., Masumoto, H. & Hiraoka, Y. Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. *Journal of cell science* **116**, 3327-3338, (2003).
264 Muller, H. J. Types of visible variations induced by X-rays in Drosophila. *J Genet* **22**, 299-334, (1930).
265 Muller, M. M., Fierz, B., Bittova, L., Liszczak, G. & Muir, T. W. A two-state activation mechanism controls the histone methyltransferase Suv39h1. *Nature chemical biology* **12**, 188-193, (2016).
266 Al-Sady, B., Madhani, H. D. & Narlikar, G. J. Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. *Molecular cell* **51**, 80-91, (2013).
267 Elgin, S. C. & Reuter, G. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harbor perspectives in biology* **5**, a017780, (2013).
268 Schotta, G. *et al.* Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *The EMBO journal* **21**, 1121-1131, (2002).
269 LeRoy, G. *et al.* Heterochromatin protein 1 is extensively decorated with histone code-like post-translational modifications. *Molecular & cellular proteomics : MCP* **8**, 2432-2442, (2009).
270 Maison, C. *et al.* SUMOylation promotes de novo targeting of HP1alpha to pericentric heterochromatin. *Nature genetics* **43**, 220-227, (2011).
271 Iyengar, S. & Farnham, P. J. KAP1 protein: an enigmatic master regulator of the genome. *The Journal of biological chemistry* **286**, 26267-26276, (2011).
272 Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124, (2001).
273 Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-120, (2001).
274 Ball, L. J. *et al.* Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. *The EMBO journal* **16**, 2473-2481, (1997).
275 Jacobs, S. A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080-2083, (2002).
276 Nielsen, P. R. *et al.* Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**, 103-107, (2002).
277 Fischle, W. *et al.* Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116-1122, (2005).
278 Cowieson, N. P., Partridge, J. F., Allshire, R. C. & McLaughlin, P. J. Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Current biology : CB* **10**, 517-525, (2000).
279 Thiru, A. *et al.* Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *The EMBO journal* **23**, 489-499, (2004).
280 Brasher, S. V. *et al.* The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *The EMBO journal* **19**, 1587-1597, (2000).
281 Smothers, J. F. & Henikoff, S. The HP1 chromo shadow domain binds a consensus peptide pentamer. *Current biology : CB* **10**, 27-30, (2000).

282 Mendez, D. L. *et al.* The HP1a disordered C terminus and chromo shadow domain cooperate to select target peptide partners. *Chembiochem : a European journal of chemical biology* **12**, 1084-1096, (2011).
283 Maison, C. *et al.* Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nature genetics* **30**, 329-334, (2002).
284 Keller, C. *et al.* HP1(Swi6) mediates the recognition and destruction of heterochromatic RNA transcripts. *Molecular cell* **47**, 215-227, (2012).
285 Cheutin, T. *et al.* Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**, 721-725, (2003).
286 Klose, R. J. *et al.* The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* **442**, 312-316, (2006).
287 Yamane, K. *et al.* JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* **125**, 483-495, (2006).
288 Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436-439, (2005).
289 Friedman, L. J., Chung, J. & Gelles, J. Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. *Biophysical journal* **91**, 1023-1031, (2006).
290 Hoskins, A. A. *et al.* Ordered and dynamic assembly of single spliceosomes. *Science* **331**, 1289-1295, (2011).
291 Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nature methods* **5**, 507-516, (2008).
292 Sofia, S. J., Premnath, V. V. & Merrill, E. W. Poly(ethylene oxide) Grafted to Silicon Surfaces: Grafting Density and Protein Adsorption. *Macromolecules* **31**, 5059-5070, (1998).
293 Kingshott, P., Thissen, H. & Griesser, H. J. Effects of cloud-point grafting, chain length, and density of PEG layers on competitive adsorption of ocular proteins. *Biomaterials* **23**, 2043-2056, (2002).
294 Chandradoss, S. D. *et al.* Surface passivation for single-molecule protein studies. *Journal of visualized experiments : JoVE*, (2014).
295 Ticau, S., Friedman, L. J., Ivica, N. A., Gelles, J. & Bell, S. P. Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell* **161**, 513-525, (2015).
296 Larson, J. D., Rodgers, M. L. & Hoskins, A. A. Visualizing cellular machines with colocalization single molecule microscopy. *Chemical Society reviews* **43**, 1189-1200, (2014).
297 Lu, Y., Wang, W. & Kirschner, M. W. Specificity of the anaphase-promoting complex: a single-molecule study. *Science* **348**, 1248737, (2015).
298 Lu, Y., Lee, B. H., King, R. W., Finley, D. & Kirschner, M. W. Substrate degradation by the proteasome: a single-molecule kinetic analysis. *Science* **348**, 1250834, (2015).
299 Revyakin, A. *et al.* Transcription initiation by human RNA polymerase II visualized at single-molecule resolution. *Genes & development* **26**, 1691-1702, (2012).
300 Chandradoss, S. D., Schirle, N. T., Szczepaniak, M., MacRae, I. J. & Joo, C. A Dynamic Search Process Underlies MicroRNA Targeting. *Cell* **162**, 96-107, (2015).
301 Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **111**, 229-233, (1992).
302 Keppler, A. *et al.* A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature biotechnology* **21**, 86-89, (2003).

303 Amitai, G., Callahan, B. P., Stanger, M. J., Belfort, G. & Belfort, M. Modulation of intein activity by its neighboring extein substrates. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 11005-11010, (2009).
304 Fischle, W. *et al.* Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes & development* **17**, 1870-1881, (2003).
305 Seidel, S. A. *et al.* Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods* **59**, 301-315, (2013).
306 Zheng, Q., Jockusch, S., Zhou, Z. & Blanchard, S. C. The contribution of reactive oxygen species to the photobleaching of organic fluorophores. *Photochemistry and photobiology* **90**, 448-454, (2014).
307 Zheng, Q. *et al.* Ultra-stable organic fluorophores for single-molecule research. *Chemical Society reviews* **43**, 1044-1056, (2014).
308 Ha, T. & Tinnefeld, P. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. *Annual review of physical chemistry* **63**, 595-617, (2012).
309 Cordes, T., Vogelsang, J. & Tinnefeld, P. On the mechanism of Trolox as antiblinking and antibleaching reagent. *Journal of the American Chemical Society* **131**, 5018-5019, (2009).
310 Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nature methods* **3**, 891-893, (2006).
311 Dave, R., Terry, D. S., Munro, J. B. & Blanchard, S. C. Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging. *Biophysical journal* **96**, 2371-2381, (2009).
312 Benesch, R. E. & Benesch, R. Enzymatic removal of oxygen for polarography and related methods. *Science* **118**, 447-448, (1953).
313 Gibson, D. T. Microbial degradation of aromatic compounds. *Science* **161**, 1093-1097, (1968).
314 Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophysical journal* **94**, 1826-1835, (2008).
315 Shi, X., Lim, J. & Ha, T. Acidification of the oxygen scavenging system in single-molecule fluorescence studies: in situ sensing with a ratiometric dual-emission probe. *Analytical chemistry* **82**, 6132-6138, (2010).
316 Ha, T. Single-molecule approaches embrace molecular cohorts. *Cell* **154**, 723-726, (2013).
317 Graham, J. S., Johnson, R. C. & Marko, J. F. Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. *Nucleic acids research* **39**, 2249-2259, (2011).
318 Sing, C. E., Olvera de la Cruz, M. & Marko, J. F. Multiple-binding-site mechanism explains concentration-dependent unbinding rates of DNA-binding proteins. *Nucleic acids research* **42**, 3783-3791, (2014).
319 Fan, J. Y., Rangasamy, D., Luger, K. & Tremethick, D. J. H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. *Molecular cell* **16**, 655-661, (2004).
320 Ngo, T. T., Zhang, Q., Zhou, R., Yodh, J. G. & Ha, T. Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell* **160**, 1135-1144, (2015).
321 Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* **437**, 55-75, (1948).
322 Stryer, L. & Haugland, R. P. Energy transfer: a spectroscopic ruler. *Proceedings of the National Academy of Sciences of the United States of America* **58**, 719-726, (1967).

323 Stryer, L. Fluorescence energy transfer as a spectroscopic ruler. *Annual review of biochemistry* **47**, 819-846, (1978).
324 Ha, T. *et al.* Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6264-6268, (1996).
325 Schuler, B., Lipman, E. A. & Eaton, W. A. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* **419**, 743-747, (2002).
326 Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in single-molecule fluorescence methods for molecular biology. *Annual review of biochemistry* **77**, 51-76, (2008).
327 Blanchard, S. C., Gonzalez, R. L., Kim, H. D., Chu, S. & Puglisi, J. D. tRNA selection and kinetic proofreading in translation. *Nature structural & molecular biology* **11**, 1008-1014, (2004).
328 Akyuz, N., Altman, R. B., Blanchard, S. C. & Boudker, O. Transport dynamics in a glutamate transporter homologue. *Nature* **502**, 114-118, (2013).
329 Akyuz, N. *et al.* Transport domain unlocking sets the uptake rate of an aspartate transporter. *Nature* **518**, 68-73, (2015).
330 Wasserman, M. R., Alejo, J. L., Altman, R. B. & Blanchard, S. C. Multiperspective smFRET reveals rate-determining late intermediates of ribosomal translocation. *Nature structural & molecular biology* **23**, 333-341, (2016).
331 Hamdan, S. M., Loparo, J. J., Takahashi, M., Richardson, C. C. & van Oijen, A. M. Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* **457**, 336-339, (2009).
332 McKinney, S. A., Declais, A. C., Lilley, D. M. & Ha, T. Structural dynamics of individual Holliday junctions. *Nature structural biology* **10**, 93-97, (2003).
333 Diez, M. *et al.* Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase. *Nature structural & molecular biology* **11**, 135-141, (2004).
334 Wang, S., Vafabakhsh, R., Borschel, W. F., Ha, T. & Nichols, C. G. Structural dynamics of potassium-channel gating revealed by single-molecule FRET. *Nature structural & molecular biology* **23**, 31-36, (2016).
335 Ragunathan, K., Liu, C. & Ha, T. RecA filament sliding on DNA facilitates homology search. *eLife* **1**, e00067, (2012).
336 Vafabakhsh, R. & Ha, T. Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science* **337**, 1097-1101, (2012).
337 Blacketer, M. J., Feely, S. J. & Shogren-Knaak, M. A. Nucleosome interactions and stability in an ordered nucleosome array model system. *The Journal of biological chemistry* **285**, 34597-34607, (2010).
338 Mandecki, W., Hayden, M. A., Shallcross, M. A. & Stotland, E. A totally synthetic plasmid for general cloning, gene expression and mutagenesis in Escherichia coli. *Gene* **94**, 103-107, (1990).
339 Quan, J. & Tian, J. Circular polymerase extension cloning of complex gene libraries and pathways. *PloS one* **4**, e6441, (2009).
340 Tims, H. S. & Widom, J. Stopped-flow fluorescence resonance energy transfer for analysis of nucleosome dynamics. *Methods* **41**, 296-303, (2007).
341 Blanchard, S. C., Kim, H. D., Gonzalez, R. L., Jr., Puglisi, J. D. & Chu, S. tRNA dynamics on the ribosome during translation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12893-12898, (2004).

342 McCann, J. J., Choi, U. B., Zheng, L., Weninger, K. & Bowen, M. E. Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. *Biophysical journal* **99**, 961-970, (2010).
343 Allahverdi, A., Chen, Q., Korolev, N. & Nordenskiold, L. Chromatin compaction under mixed salt conditions: opposite effects of sodium and potassium ions on nucleosome array folding. *Scientific reports* **5**, 8512, (2015).
344 Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62-67, (2014).
345 Stigler, J., Camdere, G. O., Koshland, D. E. & Greene, E. C. Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. *Cell reports* **15**, 988-998, (2016).
346 Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. *The journal of physical chemistry. B* **114**, 7983-7995, (2010).
347 Muschielok, A. *et al.* A nano-positioning system for macromolecular structural analysis. *Nature methods* **5**, 965-971, (2008).
348 Fussner, E. *et al.* Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO reports* **13**, 992-996, (2012).
349 Kilic, S., Bachmann, A. L., Bryan, L. C. & Fierz, B. Multivalency governs HP1alpha association dynamics with the silent chromatin state. *Nature communications* **6**, 7313, (2015).
350 Sinha, M. & Peterson, C. L. A Rad51 presynaptic filament is sufficient to capture nucleosomal homology during recombinational repair of a DNA double-strand break. *Molecular cell* **30**, 803-810, (2008).
351 Brunstein, M., Herault, K. & Oheim, M. Eliminating unwanted far-field excitation in objective-type TIRF. Part II. combined evanescent-wave excitation and supercritical-angle fluorescence detection improves optical sectioning. *Biophysical journal* **106**, 1044-1056, (2014).
352 Brunstein, M., Teremetz, M., Herault, K., Tourain, C. & Oheim, M. Eliminating unwanted far-field excitation in objective-type TIRF. Part I. identifying sources of nonevanescent excitation light. *Biophysical journal* **106**, 1020-1032, (2014).
353 Burghardt, T. P., Hipp, A. D. & Ajtai, K. Around-the-objective total internal reflection fluorescence microscopy. *Applied optics* **48**, 6120-6131, (2009).
354 Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071-1078, (2009).
355 Bunting, S. F. & Nussenzweig, A. End-joining, translocations and cancer. *Nature reviews. Cancer* **13**, 443-454, (2013).
356 Ochs, F. *et al.* 53BP1 fosters fidelity of homology-directed DNA repair. *Nature structural & molecular biology* **23**, 714-721, (2016).
357 Qi, Z. *et al.* DNA sequence alignment by microhomology sampling during homologous recombination. *Cell* **160**, 856-869, (2015).
358 Ceccaldi, R., Rondinelli, B. & D'Andrea, A. D. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in cell biology* **26**, 52-64, (2016).
359 Zimmermann, M., Lottersberger, F., Buonomo, S. B., Sfeir, A. & de Lange, T. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science* **339**, 700-704, (2013).
360 Sanders, S. L. *et al.* Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell* **119**, 603-614, (2004).
361 Fradet-Turcotte, A. *et al.* 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature* **499**, 50-54, (2013).

362 Zgheib, O., Pataky, K., Brugger, J. & Halazonetis, T. D. An oligomerized 53BP1 tudor domain suffices for recognition of DNA double-strand breaks. *Molecular and cellular biology* **29**, 1050-1058, (2009).
363 Callen, E. *et al.* 53BP1 mediates productive and mutagenic DNA repair through distinct phosphoprotein interactions. *Cell* **153**, 1266-1280, (2013).
364 Botuyan, M. V. *et al.* Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell* **127**, 1361-1373, (2006).
365 Neumann, H., Wang, K., Davis, L., Garcia-Alai, M. & Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **464**, 441-444, (2010).
366 Bang, D. & Kent, S. B. A one-pot total synthesis of crambin. *Angewandte Chemie* **43**, 2534-2538, (2004).
367 Kochendoerfer, G. G. *et al.* Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein. *Science* **299**, 884-887, (2003).
368 Bang, D., Pentelute, B. L. & Kent, S. B. Kinetically controlled ligation for the convergent chemical synthesis of proteins. *Angewandte Chemie* **45**, 3985-3988, (2006).
369 Seenaiah, M., Jbara, M., Mali, S. M. & Brik, A. Convergent versus sequential protein synthesis: the case of ubiquitinated and glycosylated H2B. *Angewandte Chemie* **54**, 12374-12378, (2015).
370 Wieneke, R. *et al.* Silica precipitation with synthetic silaffin peptides. *Organic & biomolecular chemistry* **9**, 5482-5486, (2011).
371 Camarero, J. A. & Muir, T. W. Chemoselective backbone cyclization of unprotected peptides. *Chemical communications*, 1369-1370, (1997).
372 Camarero, J. A. *et al.* Rescuing a destabilized protein fold through backbone cyclization. *Journal of molecular biology* **308**, 1045-1062, (2001).
373 Debelouchina, G. T., Gerecht, K. & Muir, T. W. Ubiquitin utilizes an acidic surface patch to alter chromatin structure. *Nature chemical biology*, (2016).
374 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **6**, 343-345, (2009).
375 Gibson, D. G., Smith, H. O., Hutchison, C. A., 3rd, Venter, J. C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nature methods* **7**, 901-903, (2010).

9. CV

Sinan Kilic

sinan16_8@msn.com +41 078 723 14 56 St-Sulpice, Switzerland

Education

PhD (Chemical biology)
Ecole polytechnique federale Lausanne (EPFL)
Nov 2012 - Jan 2017 Lausanne, Switzerland

M.Sc Pharmaceutical Sciences
University of Copenhagen
Sept 2009 - Sept 2012 Copenhagen, Denmark

B.Sc Pharmaceutical Sciences
University of Copenhagen
Sept 2006 - June 2009 Copenhagen, Denmark

Experience

Doctoral assistant
Ecole polytechnique federale Lausanne (EPFL)
Nov 2012 - Jan 2017 Lausanne, Switzerland
- Chemical biology for chromatin engineering
- Single-molecule microscopy for chromatin binding and dynamics
- Supervised 6 master's thesis or project students

Visiting student research collaborator
Princeton University
Sept 2011 - June 2012 Princeton, NJ, USA
- Fluorescence and FRET studies of nucleosome stability

Scholarhip assistant
University of Copenhagen
Feb 2011 - July 2011 Copenhagen, Denmark
- Peptide synthesis and ligation for a glutamate binding domain

Assistant in medical department
Novo Nordisk Scandinavia
Sept 2010 - Dec 2010 Copenhagen, Denmark
- Compiled data from clinical studies for reimbursement evaluations

Pharmacist intern
Herning Løve Pharmacy
Feb 2010 - July 2010 Herning, Denmark
- Prepared information leaflet for customers about generic medication
- Project on health insurance card as customer ID for safety and speed

Assistant in sales, marketing, education
Novo Nordisk Scandinavia
Jan 2008 - Jan 2010 Copenhagen, Denmark
- Organized education material for courses to teach doctors and nurses

Publications

Kilic, S., Doroshenko, O., Felekyan, S., Boichenko, I., Arya, G., Seidel, C.A.M., Fierz, B. Inter-nucleosome smFRET in chromatin fibers reveals multiscale dynamics modulated by HP1. ***Manuscript in preparation***

Kilic, S., Lechner, M., Tobler, M., Fierz, B. Semisynthesis of histones signaling DNA double-strand breaks. ***Manuscript in preparation.***

Kilic, S., Bachmann, A., Bryan, L.C., Fierz, B. Multivalency governs HP1a association dynamics with the silent chromatin state. ***Nat Comm,* 2015**

Pick, H., **Kilic, S.**, Fierz, B. Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function**.** ***Biochim Biophys Acta*, 2014.**

Fierz, B., **Kilic, S**., Hieb, A.R., Luger, K., Muir, T.W. Stability of nucleosomes containing homogenously ubiquitylated H2A and H2B prepared using semisynthesis. **J Am Chem Soc**, **2012.**

Awards and funding

Best student talk
21st PicoQuant International workshop on single-molecule spectroscopy and super-resolution microscopy in life-sciences
Sept 2015

Boehringer Ingelheim PhD fellowship
Funding for salary and travel during 2 years of PhD studies
Dec 2013 - Nov 2015

Novo scholarship for master's students
~12.000USD
Funding for master's thesis work
Sept 2011 - Aug 2012

Further funding as Princeton VSRC
~12.000USD
Additional funds applied for and granted for stay abroad
Sept 2011 - Aug 2012