

Aus dem
Lehrstuhl für Molekularbiologie, Biomedizinisches Centrum (BMC)
Institut der Ludwig-Maximilians-Universität München
Vorstand: Prof. Dr. Peter Becker

**Challenging chromatin assembly in vitro:
Leveraging and expanding the use of the *Drosophila* embryonic extract
system to study chromatin structure and function**

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

vorgelegt von
Vera Kleene

aus
Bückeburg

Jahr
2024

Mit Genehmigung der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

Betreuer: Prof. Dr. Axel Imhof

Zweitgutachter: Prof. Dr. Stefan Stricker

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 25.10.2024



Affidavit

Kleene, Vera

Surname, first name

I hereby declare, that the submitted thesis entitled

Challenging chromatin assembly in vitro: Leveraging and expanding the use of the Drosophila embryonic extract system to study chromatin structure and function

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

München, 28.10.2024

Place, Date

Vera Kleene

Signature doctoral candidate

Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make because they lead little by little to the truth.

-Jules Verne

Remember, kids, the only difference between screwing around and science is writing it down.

-Mythbusters

Parts of this thesis have been published in:

Kleene, V., Corvaglia, V., Chacin, E., Forne, I., Konrad, D.B., Khosravani, P., Douat, C., Kurat, C.F., Huc, I. and Imhof, A. (2023) DNA mimic foldamers affect chromatin composition and disturb cell cycle progression. *Nucleic Acids Research*, 51, 9629-9642, <https://doi.org/10.1093/nar/gkad681>.

Used excerpts of the publication within this thesis are marked in *italics*.

Teile dieser Arbeit sind veröffentlicht in:

Kleene, V., Corvaglia, V., Chacin, E., Forne, I., Konrad, D.B., Khosravani, P., Douat, C., Kurat, C.F., Huc, I. and Imhof, A. (2023) DNA mimic foldamers affect chromatin composition and disturb cell cycle progression. *Nucleic Acids Research*, 51, 9629-9642, <https://doi.org/10.1093/nar/gkad681>.

Verwendete Auszüge aus der Publikation in dieser Dissertation sind *kursiv* markiert.

Table of Contents

List of Figures	5
List of Main Tables	7
Summary	8
Zusammenfassung.....	9
1 Introduction.....	11
1.1 Chromatin.....	11
1.2 Chromatin assembly.....	13
1.2.1 Methods to investigate chromatin assembly.....	14
1.2.2 Proteome and Protein binding modes in chromatin assembly.....	17
1.2.3 Histone modifications and chromatin assembly.....	18
1.2.4 Bottlenecks of further investigation of protein binding modes during chromatin assembly.....	19
1.3 Metabolism and its link to chromatin assembly.....	21
1.3.1 Metabolism and histone modifications	21
1.3.2 Acetylation.....	21
1.3.3 Methylation.....	22
2 Aims of this Study.....	24
3 Project 1: Chromatin assembly and protein binding	25
3.1 Graphical abstract.....	25
3.2 Results	26
3.2.1 Foldamer interferes with in vitro chromatin-bound proteome.....	26
3.2.2 The foldamer interactome.....	30
3.2.3 The protein subset competed off chromatin fiber by binding to foldamers includes ORC	32
3.2.4 Foldamers also disturb the chromatin-bound proteome in <i>Drosophila</i> S2 cells.....	33
3.2.5 Foldamer treatment interferes with cell cycle progression in <i>Drosophila</i> S2 cells.....	36
3.3 Discussion.....	40
3.3.1 The interactome supports and extends insights into foldamer-mimicking capabilities	40
3.3.2 Foldamer does not disturb basic processes like nucleosomal integration spacing during chromatin assembly in vitro.....	41
3.3.3 A subset of proteins remains unaffected by foldamer in their binding to chromatin in vitro due to their binding mode	41

3.3.4	Proteins are differentially disrupted in their binding to chromatin by foldamer in vitro, depending on their binding mode to DNA/chromatin	43
3.3.5	S2 cells allow for functional in vivo experiments with foldamer	45
3.3.6	In vivo study reveals transferability of in vitro results, validating foldamer-sensitivity-based protein clusters	45
3.3.7	In vivo cell cycle experiments reveal the physiological impact of foldamer.....	46
3.3.8	The difference between foldamers and DNA in interference assays	46
3.3.9	Future perspectives and recommendations for foldamer design and modifications	47
3.3.10	ORC and the foldamer, a proof-of-concept: Interaction, interference, and functional disablement leading to physiological consequences	49
3.3.11	DREX assembly confirmed as an exploratory assay to test small molecules in a complex chromatin context	50
4	Project 2: Chromatin assembly and metabolism	51
4.1	Graphical abstract.....	51
4.2	Results	52
4.2.1	The DREX proteome.....	52
4.2.2	Investigation of the DREX fundamental requirements for chromatin assembly capacity.....	55
4.2.3	The activity of the proteasome in <i>Drosophila</i> embryo extract.....	57
4.2.4	Inhibition of proteasome does not affect nucleosome integration and spacing.....	58
4.2.5	Proteasome inhibition causes the accumulation of proteins on chromatin over time..	59
4.2.6	Inhibition of proteasome leads to differential changes in chromatin proteome	60
4.2.7	Benchmarking metabolomics studies in DREX extract	62
4.2.8	Dialysis of DREX extract removes most metabolites	63
4.2.9	Histone modifications upon metabolite depletion	65
4.2.10	Supplementation of isotopically labeled Methionine leads to incorporation of isotopically labeled methyl to H4K20me1	66
4.2.11	SAH excess traps Pr-Set7 on and removes AHCY from chromatin fiber	67
4.2.12	Metabolites reappear/metabolome slowly recovers over time in DREX after dialysis	69
4.3	Discussion.....	72
4.3.1	The DREX proteome.....	72
4.3.2	What makes the DREX functional: concentration or composition?	73
4.3.3	Proteasomal activity in DREX	73
4.3.4	Inhibition of proteasome leads to overall accumulation and differential changes in chromatome	74

4.3.5	Metabolite profiling in DREX.....	75
4.3.6	DREX exhibits complex metabolic activity	77
4.3.7	SAH excess traps Pr-Set7 fiber and effectively inhibits H4K20me1	80
4.3.8	Lack of diversity in histone methylations in DREX assembly	82
4.3.9	A model for metabolic coupling of H4K20me1 during chromatin assembly	84
5	Final Conclusions.....	86
6	Materials and Methods	87
6.1	Materials	87
6.1.1	Bacteria strains	87
6.1.2	Chemicals, peptides, and recombinant proteins	87
6.1.3	DNA: Plasmids and oligos.....	88
6.1.4	Software and algorithms.....	89
6.1.5	Standard buffers	90
6.2	Molecular biology methods.....	91
6.2.1	E. Coli DNA transformation.....	91
6.2.2	Growing of bacteria and plasmid DNA purification	91
6.2.3	DNA linearization and biotinylation	91
6.2.4	DNA precipitation	91
6.2.5	Agarose gel electrophoresis.....	92
6.2.6	Annealing of short DNA oligos.....	92
6.2.7	Preparation of preblastoderm <i>Drosophila</i> embryo extract [DREX]	92
6.3	Biochemical methods	93
6.3.1	SDS-Polyacrylamide gel electrophoresis (SDS-PAGE).....	93
6.3.2	Coomassie staining.....	93
6.3.3	Proteasomal activity assay	93
6.3.4	Dialysis of DREX	93
6.3.5	Protein concentration measurement with BCA.....	94
6.4	In vitro chromatin methods	95
6.4.1	In vitro chromatin assembly in DREX.....	95
6.4.2	Pulldown with biotinylated foldamer in DREX	95
6.4.3	Chromatin accessibility assay by Micrococcal nuclease digestion	95
6.5	Cell-based methods	96
6.5.1	Cell culture.....	96

6.5.2	Treatment of cells with DNA mimic foldamer.....	96
6.5.3	Subcellular fractionation.....	96
6.5.4	Flow cytometry	96
6.6	Mass spectrometry methods	97
6.6.1	Sample preparation for total proteome analysis via LC-MS/MS.....	97
6.6.2	Sample preparation with on-beads digest for Proteome via LC-MS/MS.....	97
6.6.3	Sample preparation for histone modification analysis	97
6.6.4	Sample preparation for metabolomics by liquid-liquid extraction.....	98
6.6.5	LC-MS/MS.....	98
6.6.6	CESI-MS.....	99
6.7	Mass spectrometry data processing and analysis.....	100
6.7.1	Processing and analysis of LC-MS/MS proteome data.....	100
6.7.2	Processing and analysis of LC-MS/MS histone modification data.....	100
6.7.3	Processing and analysis of CESI-MS metabolome data	100
7	References	101
8	Acknowledgments.....	114
9	List of publications	116
10	Annex.....	117
10.1	Abbreviations	117
10.2	Access credentials for the datasets and corresponding analysis scripts (available online) ..	118

List of Figures

Figure 1 Chromatin basic structure	11
Figure 2 Potential mechanisms and functions of combinatorial histone PTMs as outlined in the original ‘histone code’ hypothesis	12
Figure 3 Cellular processes involving chromatin assembly	13
Figure 4 Chromatin assembly in preblastoderm <i>Drosophila</i> embryo extract (DREX)	16
Figure 5 Histone modification overview	19
Figure 6 Chemical structures of DNA mimic foldamers	20
Figure 7 Scheme of methionine cycle in <i>Drosophila</i>	23
Figure 8 Graphical abstract of project 1	25
Figure 9 In vitro chromatin assemblies in the presence of foldamer	26
Figure 10 Schematic diagram of experimental design for <i>Drosophila</i> embryo extract assisted in vitro chromatin assembly.....	27
Figure 11 “Chromatin binders” and their susceptibility to interference in binding by foldamer.....	29
Figure 12 Susceptibility to interference in the binding of ORC1-5 proteins to foldamer or DNA control	30
Figure 13 Foldamer interactome experiment	31
Figure 14 Foldamer interactome analysis.....	32
Figure 15 Subsetting for chromatin proteins affected by foldamer.....	33
Figure 16 Subcellular fractionation experiment	34
Figure 17 Protein distribution in all subcellular fractions	35
Figure 18 Change in protein intensities of chromatin binders in a chromatin-bound fraction.....	36
Figure 19 Effect of foldamer treatment on cell cycle progression.....	38
Figure 20 Scatterplot of proteins in pulldown with “8 bp” or “16 bp” foldamers.	48
Figure 21 Graphical abstract of project 2	51
Figure 22 Distribution of iBAQ intensities of proteins identified in the DREX extract	53
Figure 23 Complete <i>Drosophila</i> embryo extract (DREX) proteomic analysis	54
Figure 24 Proteomic analysis of top 100 most abundant proteins in <i>Drosophila</i> embryo extract (DREX)	55
Figure 25 Difference in proteome for DREX with and without chromatin assembling capacity	57
Figure 26 Proteasomal subunits detected on DREX-assembled chromatin	57
Figure 27 Inhibition of proteasomal activity in DREX.....	58
Figure 28 In vitro chromatin assemblies in the presence of proteasome inhibitor epoxomicin.....	59
Figure 29 Effect of proteasome inhibition by epoxomicin on overall protein abundance on chromatin fiber over time.....	60
Figure 30 Effect of proteasome inhibition with 20 μ M Epoxomicin on abundance of specific proteins on chromatin fiber after 4h.....	61
Figure 31 Concentration of metabolites involved in methionine cycle in DREX.....	63
Figure 32 Dialysis of DREX depletes small molecules from the extract.....	64
Figure 33 Dialysis of DREX depletes or reduces metabolites involved in the methionine cycle from the extract	65
Figure 34 Effect of DREX dialysis and SAH addition on the histone modification H4K20monomethylation	66
Figure 35 DREX incorporates methyl group from isotopically labeled methionine on H4 histones .	67

Figure 36 Histone methyl transferases accumulate on chromatin during assembly upon SAH addition 68

Figure 37 DNA dependence of recovery of metabolites involved in methionine cycle over time after dialysis of DREX..... 70

Figure 38 ATP dependence of recovery of metabolites involved in methionine cycle over time after dialysis of DREX..... 71

Figure 39 Scheme of working hypothesis for metabolic coupling of H4K20me1 in DREX..... 85

List of Main Tables

Table 1 Representative DREX batches and their concentrations.....	56
---	----

Summary

Challenging chromatin assembly in vitro: Leveraging and expanding the use of the *Drosophila* embryonic extract system to study chromatin structure and function

In eukaryotes, the genome is packaged into chromatin to compact but also regulate the genome. Chromatin consists of the DNA, histones, and non-histone proteins that interact with the chromatin fiber. The structure of chromatin fiber is highly dynamic and depends on multiple modulators. The question what factors influence the protein association with chromatin is therefore of critical importance. The influencing factors range from the binding modes that contribute to fiber association to the question whether the chromatinome is regulated by its metabolic environment. The challenge in studying these questions is the requirement of sufficiently complex systems that allow the detection of transferable and physiologically relevant effects while still being easy-to-manipulate and well-characterized to allow deconvolution of the factors. The preblastoderm *Drosophila* embryo extract (DREX) chromatin assembly system is currently the best-characterized method to investigate the assembling chromatin fiber. However, it still lacks thorough characterization regarding total protein composition and metabolic state and activity.

This study characterizes previously unknown key aspects of the DREX in vitro system including a full proteome analysis with a hitherto unreached depth and the first metabolomics profiling of the extract. Notably, this study also proves that the extract is metabolically active and able to degrade proteins and replenish metabolites. These findings significantly advance the understanding of the extract's protein composition and metabolic state. The obtained data provide a crucial baseline reference that can be used as a tool for future research and facilitates the adaption of the DREX as a model with significant potential for advancing the understanding of metabolic processes.

Secondly, the metabolic processes detected within the DREX extract are leveraged to study the link between metabolism and chromatin structure. The DREX chromatin assembly model is challenged using the supplementation of isotopically labeled metabolites. Subsequent mass spectrometry analysis reveals changes in histone modifications, suggesting a strong coupling of metabolism and histone modifications in the extract.

Thirdly, this study addresses the effect of synthetic DNA mimic foldamers that mimic DNA shape on the composition of chromatin. Changes in protein composition revealed insights into the partial contributions of different binding modes to the chromatin association of selected proteins. By using flow cytometry and fractionation coupled with proteomics, similar effects and disruption cell cycle progression in vivo were observed. In addition to the mechanistic insights, this provides a proof-of-concept for the DREX assembly system to be used as a method to screen small molecules for pharmacological intervention with chromatin assembly in the future.

In conclusion, in this study the in vitro chromatin assembly system DREX is characterized in depth, proteomically and metabolically, to facilitate future research and findings are leveraged to challenge the system using DNA mimic foldamers, metabolite depletion, and supplementation. Novel insights into the contribution of chromatin protein binding modes and coupling of metabolism and histone modifications during chromatin assembly are gained. This research characterizes and leverages the in vitro system DREX chromatin assembly, expanding the model for adaptation in metabolomics research and establishes it as a pharmacological screening assay with good transferability to in vivo setting.

Zusammenfassung

Erforschung des Chromatinaufbaus in vitro durch molekulare Herausforderungen: Charakterisierung, Nutzung und Ausweitung der Verwendung des *Drosophila*-Embryonenextraktsystems zur Untersuchung von Chromatinstruktur und -funktion

In Eukaryoten ist das Genom in Chromatin verpackt, um es zu verdichten, aber auch um es zu regulieren. Chromatin besteht aus der DNA, Histonen und Nicht-Histon-Proteinen, welche mit der Chromatinfaser interagieren. Die Struktur der Chromatinfasern ist hochdynamisch und hängt von zahlreichen Modulatoren ab. Daher ist die Frage, welche Faktoren die Assoziation von Proteinen mit Chromatin beeinflussen, von entscheidender Bedeutung. Die Einflussfaktoren reichen von den verschiedenen Bindungsmodi, die zur Faserassoziation beitragen, bis hin zu der Frage, ob das Chromatin durch seine metabolische Umgebung reguliert wird. Die Herausforderung bei der Untersuchung dieser Fragen besteht darin, dass hinreichend komplexe Systeme benötigt werden, die den Nachweis übertragbarer und physiologisch relevanter Effekte ermöglichen, aber dennoch leicht zu manipulieren und gut charakterisiert sind, um eine Entschlüsselung der Faktoren zu ermöglichen. Der Chromatinaufbau im präblastoderme *Drosophila*-Embryoextrakt (DREX)-System ist derzeit die am besten charakterisierte Methode zur Untersuchung der sich zusammensetzenden Chromatinfaser. Dennoch fehlt es bisher an einer gründlichen Charakterisierung der gesamten Proteinzusammensetzung sowie des Stoffwechselzustandes des Extrakts.

Diese Studie charakterisiert bisher unbekannte Schlüsselaspekte des DREX in vitro Systems, einschließlich einer vollständigen Proteomanalyse, in bisher nicht erreichten Tiefe, und der ersten Metabolomik-Profilierung des Extrakts. Zudem beweist diese Arbeit auch, dass der Extrakt metabolisch aktiv ist und in der Lage ist, Proteine abzubauen und Metaboliten zu erzeugen. Diese Erkenntnisse tragen wesentlich zum Verständnis der Proteinzusammensetzung und des Stoffwechselzustands bei. Die gewonnenen Daten stellen eine wichtige Basisreferenz dar, die als Werkzeug für künftige Forschungen genutzt werden kann und die Etablierung des DREX als Modell für das Verständnis von Stoffwechselprozessen ermöglicht.

Des Weiteren werden die im DREX-Extrakt nachgewiesenen Stoffwechselprozesse genutzt, um den Zusammenhang zwischen Stoffwechsel und Chromatinstruktur zu untersuchen. Das Modell des DREX-Chromatinaufbaus wird durch die Zugabe von isotopisch markierten Metaboliten moduliert. Die anschließende massenspektrometrische Analyse zeigt Veränderungen in den Histonmodifikationen, was auf eine starke Kopplung von Stoffwechsel und Histonmodifikationen in dem Extrakt hindeutet.

Darüber hinaus befasst sich diese Studie mit den Auswirkungen von synthetischen, DNA Mimikri Foldameren, welche die DNA-Form nachahmen, auf die Zusammensetzung des Chromatins. Veränderungen in der Proteinzusammensetzung geben Aufschluss über die Teilbeiträge verschiedener Bindungsarten zur Chromatinassoziation ausgewählter Proteine. Durch den Einsatz von Durchflusszytometrie und Fraktionierung in Verbindung mit Proteomik werden ähnliche Effekte und eine Unterbrechung der Zellzyklusprogression in vivo beobachtet. Zusätzlich zu den mechanistischen Erkenntnissen liefert dies einen konzeptionellen Beweis dafür, dass das DREX-Assemblierungssystem als Methode zum Screening kleiner Moleküle für pharmakologische Intervention in der Chromatinassemblierung geeignet ist.

Zusammenfassend lässt sich sagen, dass in dieser Studie das in vitro Chromatinaufbausystem DREX eingehend proteomisch und metabolomisch charakterisiert wird und dass die Ergebnisse genutzt

werden, um das System mit Hilfe von DNA-nachahmenden Foldamern, Metabolitenentfernung und -ergänzung zu testen. Es werden neue Erkenntnisse über den Beitrag der Chromatin-Proteinbindungsmodi und die Kopplung von Metabolismus und Histonmodifikationen während des Chromatinaufbaus gewonnen. Diese Arbeit charakterisiert und nutzt das *in vitro* System für Chromatinaufbau in DREX, erweitert das Modell für die Anpassung in der Metabolomforschung und etabliert es als pharmakologischen Screening-Assay mit guter Übertragbarkeit auf *in vivo* Szenarien.

1 Introduction

1.1 Chromatin

In eukaryotic cells, nuclear genomic DNA is packaged in a dynamic complex with RNA, histones, and non-histone proteins (Figure 1). This dynamic complex called chromatin not only enables compaction of the ca. 2 m long DNA molecule to fit in a cell nucleus of approximately 10 μM diameter but also plays a fundamental role in gene expression and maintaining genomic stability.

Chromatin structure is pivotal to allow for the functionality and control of all DNA-templated processes. Therefore, different levels of organization and regulation exist in chromatin. The basic unit of chromatin is the nucleosome, comprised of 146bp of DNA wrapped left-handed around an octamer of core histones (2 histones of each H2A, H2B, H3, and H4 histones) (1). In line, these units form a beads-on-a-string structure, the chromatin fiber. The accessibility of the DNA for factors involved in DNA-templated processes like transcription or DNA repair is crucial in fine-tuning and controlling gene expression and maintenance. Chromatin structure is shaped on multiple levels, by incorporation of histone variants, nucleosome positioning, and chemical modifications of chromatin components.

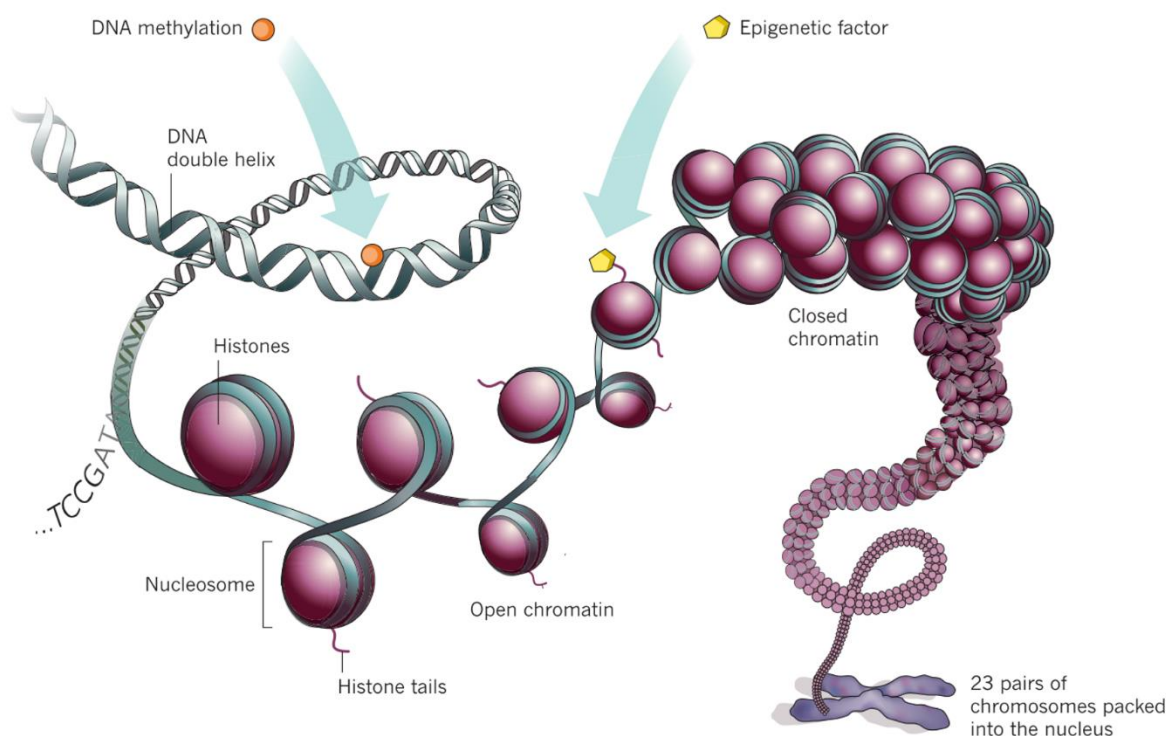


Figure 1 Chromatin basic structure

DNA is packed into chromatin. DNA double helix wraps around nucleosomes, which consist of histones. DNA and histones can be modified by epigenetic modifications to influence packaging and regulate accessibility and readout (modified from Marx et al. (2), with permission from SNCSC).

The chemical modifications that alter chromatin structure can be on either DNA or histone proteins. DNA itself can be methylated on cytosines leading to 5-methylcytosines in most eukaryotes excluding

low eukaryotes like the yeast species. DNA methylation regulates gene expression by serving as a recruiter platform for proteins involved in gene repression. DNA methylation is therefore seen as a repressive chromatin modification (3). Additionally, posttranslational modification (PTM) of histones is a major mediator for chromatin structure and function (Figure 2). One option is the integration of histone variants, e.g. integration of H2A.v instead of H2A into nucleosomes close to transcription origins, thereby altering DNA accessibility (4). More importantly for this work, posttranslational modifications are set on the N-terminal tails of histones. Histone tails are posttranslationally modified in all organisms that possess histones. Some of the modifications can be posttranslationally set before nucleosome assembly. An example of this is H4 K5 and K12 acetylation, which are present on free histones but erased upon integration of the histone into nucleosomes (5). The majority of modifications are set after the integration of histones into nucleosomes. The two main modifications are methylation and acetylation. Depending on the nature of the modification, the position of the modified amino acid on the histone tail as well as the presence of other modifications in proximity, the histone modifications alter chromatin structure and DNA accessibility. This theory is called the “histone code” hypothesis (6,7).

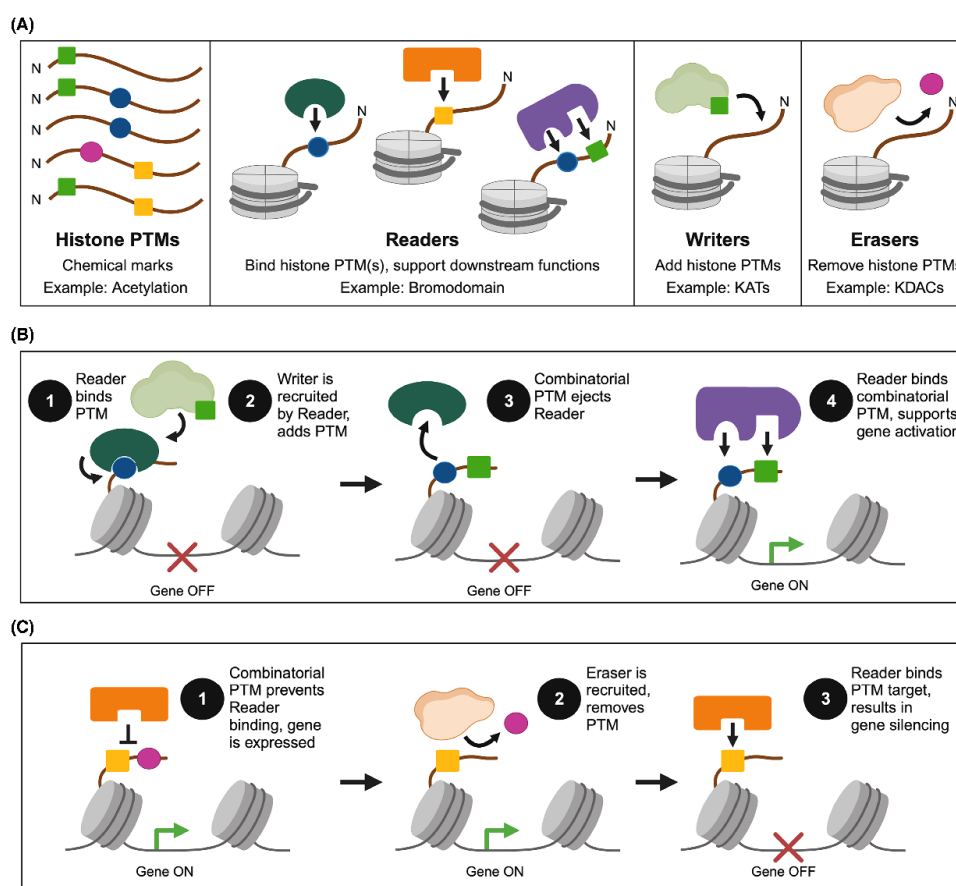


Figure 2 Potential mechanisms and functions of combinatorial histone PTMs as outlined in the original ‘histone code’ hypothesis

(A) Key elements of the language. (B, C) Examples of readout include where sequential histone PTMs facilitate gene activation (B), or combinatorial histone PTMs regulate gene activation (C). Figure created with BioRender. Figure and legend reproduced from Weinzapfel et al. (8) under open access Creative Commons Attribution License 4.0 (CC BY).

The chemical modulation of chromatin structure is mediated by enzymes whose functions can be categorized into “readers”, “writers” and “erasers”, alternatively as “modifiers” and “binders”

depending on preferred nomenclature (Figure 2) (6,8). “Readers”, “reader domains” or “binders” can bind certain domains or modifications on the chromatin fiber. An example would be the reader domains on proteins involved in gene repression that bind 5-methylcytosine on DNA. “Writers”, “writer domains” as well as “erasers” are “modifiers” that can set or remove chromatin marks. For example, the histone methyltransferase Pr-Set7 can methylate the lysine in position 20 of the histone tail of H4 using the small metabolite SAM as a cofactor while histone demethylases remove methyl groups from histones. It is important to note that one enzyme can often carry and combine different functions. This is to say that a molecule can e.g. be a “binder” and a “modifier”.

1.2 Chromatin assembly

An essential element and process shaping chromatin structure and function is chromatin (re-) assembly. To understand chromatin structure, the understanding of the process of chromatin assembly is crucial. It is when the structure is at its most vulnerable and most important determinants are set. The process is highly dynamic. Chromatin assembly is part of replication but also non-replication-dependent processes like DNA repair and transcription (Figure 3). This requires high fidelity and control. While the basics of e.g. DNA replication and DNA repair are fairly well understood, the bigger picture of these cellular processes, especially the (re-)assembly of chromatin still lacks understanding regarding e.g. chaperone recycling and correct histone modification setting behind replication forks for epigenetic stability (9-11). Identification and in-depth understanding of all factors involved in these complex processes as well as their crosstalk to cellular and metabolic environments remains understudied.

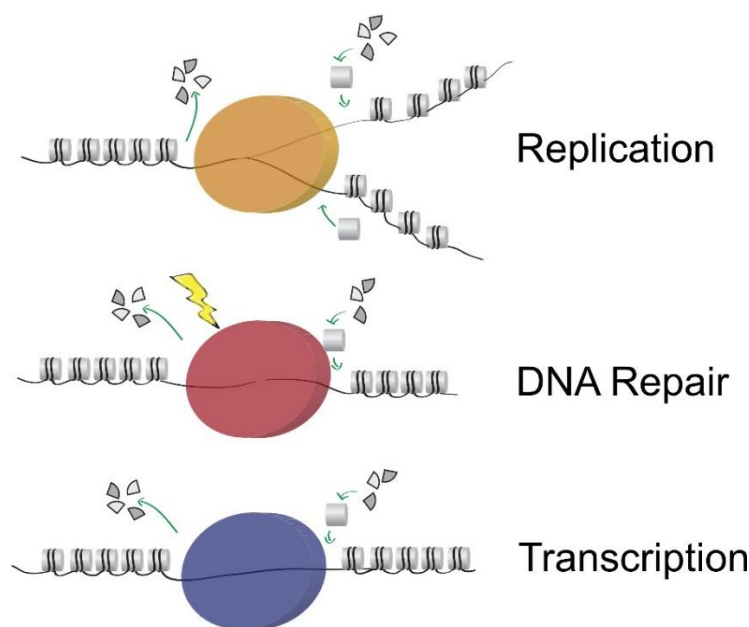


Figure 3 Cellular processes involving chromatin assembly

Replication, DNA Repair, and Transcription happen in the context of chromatin and can only be completed successfully if chromatin (re-)assembly is achieved.

It has been shown that understanding these basic DNA-templated processes without their context of chromatin highly underestimates complexity and overlooks or even negates physiologically crucial steps and factors (8). Since all these processes are essential for cell survival and cell cycle progression, further investigation regarding what proteins associate with the chromatin fiber when during

chromatin assembly, how they recognize the structure, and how they respond to environmental challenges is necessary.

1.2.1 Methods to investigate chromatin assembly

Chromatin assembly is a highly complex process that is tightly regulated and subject to many quality control mechanisms. It is essential to the proliferation and survival of cells, making it an interesting subject to investigate the many engaged mechanisms for control that guide the process. Additionally, chromatin assembly is an exceptional target for pharmacological interference. To understand the influence of complex changes in the chromatin environment either by changes in metabolome or by interference with small molecules, different experimental approaches have been developed *in vitro* and *in vivo*.

1.2.1.1 *In vivo*

Multiple methods have been developed to track chromatin assembly and the associated changes on the nascent chromatin fiber *in vivo*, each with its advantages and limitations. These methods have increasingly focused on dynamics, investigating the changes in proteome on the fiber concerning replication or DNA damage control timing. However, these methods could also be applied to look at changes to the proteome provoked by other agents disrupting processes involved in chromatin assembly, e.g. small molecule interference or metabolic changes.

Most early *in vivo* methods relied on immunofluorescence imaging to e.g. detect protein presence or accumulation at DNA damage sites or replication forks in cells (12). The limitations here are resolution as well as antibody availability, specificity, and sensitivity. New methods like AMPL-MS (Antibody-mediated proximity labeling mass spectrometry) overcome imaging limitations (13). Additionally, other techniques allow for spatial and temporal resolution while circumventing antibody limitations. These methods use the incorporation of tagged nuclear bases for the isolation of chromatin and interactors, which are subsequently analyzed by mass spectrometry. Examples of these are nascent chromatin capture (NCC), which uses biotin-deoxyuridine triphosphate (dUTP) to label and isolate replicative chromatin (9) as well as other systems –called iPOND and Dm-ChP - based on labeling DNA with 5-ethynyl-2'-deoxyuridine (EdU) that biotin can be attached to via “click chemistry” (14-17). All *in vivo* methods share the inherent limitation; if the effect of manipulation affects multiple processes or the addition of any intrinsic or external factor is cytotoxic, no mechanistic investigation is possible. Cell or organism death, too many confounding factors, and failure to separate effects on different processes often hinder pinpointing and studying any local or specific role or mechanism.

1.2.1.2 *In vitro*

In vitro systems have the power to allow for increased control of the experimenter over the environment, isolating specific factors and processes while at the same time reaching near-physiological complexity if chosen. Additionally, major limitations accounted for in cell-based assays including delivery/membrane permeability and cell toxicity are overcome in *in vitro* assays. Chromatin assembly systems *in vitro* are mainly based on 2 approaches, either buffer-based with only purified components or embryo-extract-based assembly.

Buffer-based chromatin assembly happens in a 100% defined, precise to anatomic-level, environment. Histones are either recombinantly expressed in bacteria or isolated from organisms like yeast or flies (18,19). In salt-gradient dialysis (SGD) chromatin assembly, the salt concentration in a buffer containing DNA with a sequence and purified histones is gradually decreased, thereby favoring nucleosome formation. During this process and afterward, the system can then be manipulated in a highly controlled manner. Conditions like temperature and buffer salinity can be tweaked but also addition specific inhibitors, other molecules, recombinant or purified proteins, or isolated metabolites can be added to the system to investigate their effect on the system. This system has mostly been used to investigate the effect of remodeling factors in chromatin on nucleosomal spacing (20,21). Recently, a complete chromatin replication system has been reconstituted with only highly purified factors, demonstrating mechanistic insights into the progression of the replisome through chromatin (22,23). While these systems with their great transparency allow eliminating many confounding factors and features for mechanistic studies, they also carry the inherent drawback of being far away from the physiological reality of a highly complex and crowded environment as present in vivo.

Embryo extract-based chromatin assembly is established for extracts from *Xenopus laevis* and *Drosophila melanogaster* (24,25). Here, the eggs of the species contain sufficient maternally deposited proteins and factors to undergo cell cycles before the transcriptional activation (26,27). Embryo extracts are obtained from homogenized embryos of these specific stages of embryonic development, before the start of genome activation. These extracts were discovered to be competent to assemble chromatin in vitro, making them a model system ever since. In contrast to SGD, extract-based chromatin assembly offers a complex environment much closer to physiological conditions than fully recombinant systems. These cell-free systems mimic many key aspects of chromatin assembly in vivo (24,28). In vitro systems shine with their susceptibility for manipulation, where changing conditions, e.g. by inhibitors or reagents, do not risk unspecific side effects leading to systemic failures that make readout impossible.

DREX-assisted chromatin assembly

In this thesis, chromatin assembly with DREX takes a central role. The knowledge and characterization of the central component of the method, DREX, is significantly expanded. Additionally, the method is applied to investigate mechanisms and sensitivities of chromatin assembly while in vivo cell experiments and reductionist in vitro assays in collaboration with the Kurat lab orthogonally complement the method and demonstrate the transferability of findings.

Drosophila melanogaster, commonly known as the fruit fly, has played a pivotal role in genetic research, particularly in advancing our understanding of chromatin and epigenetics. Modern historical *Drosophila* research dates back to the early 20th century, led by Thomas Hunt Morgan and colleagues, and laid the foundation for fundamental concepts in chromosomal inheritance, including sex-linked traits and genetic linkage (29). Leveraging its well-characterized genome, advanced genetic tools, well-established cell culture models, and various mutant strains, *Drosophila* remains a primary model organism for studying conserved chromatin-related processes shared with mammals, such as histone modifications and transcriptional regulation. It is a particularly valuable model organism, not only due to its striking similarity to mammals in terms of chromatin organization and gene, but also because of its short generation time, ease of manipulation, and cost-effectiveness (30). This makes it an ideal system for investigating chromatin structure and epigenetic modifications conserved across evolutionary lineages (31).

The preparation of DREX was first described by Becker et al and DREX is still prepared following the basic principles of the first protocol with minor adjustments (Figure 4 A) (24). Embryos of wild-type *Drosophila melanogaster* fruit flies are collected in 90 min intervals and separated from contaminations. The collection time point is chosen so all embryos collected are in stages 1-4, having undergone a maximum of 9 syncytial divisions and before cellularization of the embryo and major wave of zygotic genome activation (ZGA) around nuclear cycle (NC) 14 (32). These preblastoderm embryos have a particularly high prevalence of histone chaperones and other factors necessary for chromatin assembly because the embryos of these stage cycle very fast, replicating their genome every 8 min (in cycles 1-10 and then progressively slower), only passing through S-Phase and mitosis (26). The collected embryos are subjected to dechoronation, homogenization, and multiple ultracentrifugation steps to remove lipids and cell debris, leaving a pale yellow, clear, soluble protein-rich extract –DREX- that is flash frozen for storage. Due to the nature of the preblastoderm embryos and the preparation protocol, the extract is neither a cytoplasmic nor a nuclear extract. Most intact nuclei are removed by centrifugation and no measurable amount of DNA is present in DREX.

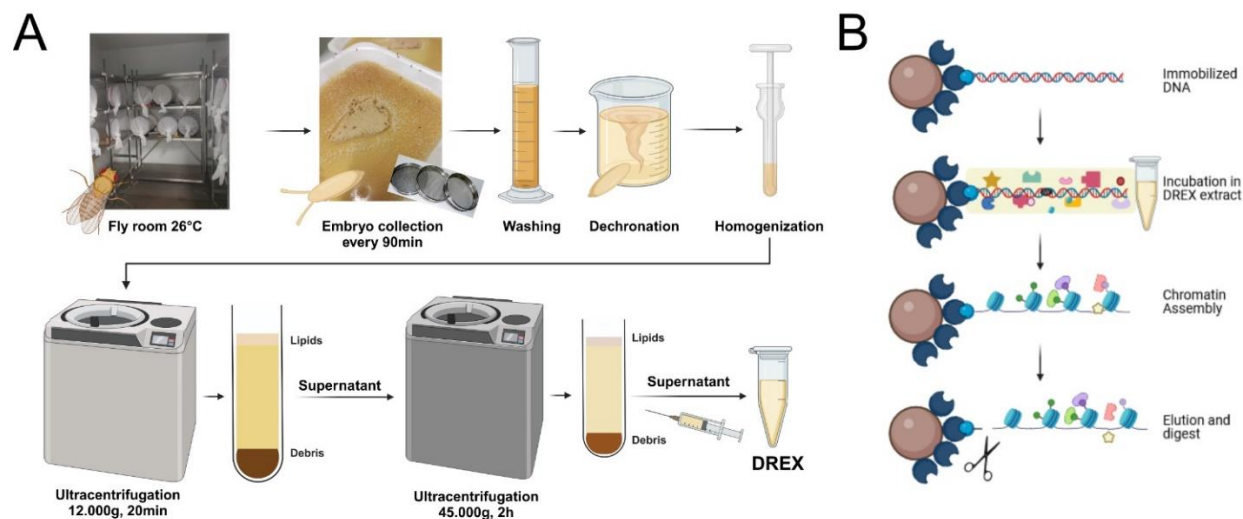


Figure 4 Chromatin assembly in preblastoderm *Drosophila* embryo extract (DREX)

A Scheme of preparation steps for DREX extract. Fly husbandry in custom-made Plexiglas “cages” in a controlled 26 °C environment. Collection of embryos every 90 min, Washing, dechoronation with Hypochlorite, and homogenization. Initial ultracentrifugation at 12.000g for 20min, then subsequent ultracentrifugation of resulting supernatant at 45.00g for 2h. Finally, extraction of supernatant by syringe from plastic centrifugation tubes and snap freeze, storage at -80 °C until usage. B Scheme of chromatin assembly in DREX extract with immobilized DNA. DNA is incubated in DREX for a defined time. DNA is packaged into chromatin, chromatin assembly factors associated with the fiber. Beads allow for the pulldown of chromatin fiber. Subsequent Elution and digest can be chosen based on readout assay e.g. Mnase digest for DNA analysis Agarose gel or Mnase-ChIP-Seq, or trypsin digest for mass spectrometry read-out. Figure created using Biorender.com, photos by Vera Kleene.

Since its establishment, the DREX chromatin assembly system has been used to study numerous processes including DNA repair, replication, transcription factor binding on chromatin, and histone modification dependencies (5,28,33-35). For this, its unique properties, including the lack of some factors, e.g. CLAMP and the thereby resulting opportunity to titrate them in, and the ubiquitous presence of others e.g. histone chaperones allowing for the assembly, were leveraged. While the system does not reach in vivo complexity, it is not as reductionist as other in vitro systems, making it ideal for the observation of factors in complex environments and systemic responses to challenges while staying

in a controlled environment compared to *in vivo*. For chromatin assembly, DNA, e.g. immobilized on beads, is incubated in DREX and then isolated and analyzed (Figure 4 B). Many techniques including protein detection via western blot but also dynamic proteomics of the assembled chromatin fiber, ChIP-Seq, and histone modification mass spectrometry have been established in the extract. This provides an extensive toolkit for the system. Nevertheless, the nature of the extract remained vastly understudied in terms of a comprehensive overview of the proteins present and completely uncharacterized regarding its metabolome until this study.

1.2.2 Proteome and Protein binding modes in chromatin assembly

Chromatin structure, maintenance, and assembly are determined and fine-tuned by different features like the spatial segmentation into nuclear compartments, the incorporation of histone variants, the presence and activity of nucleosome remodeling factors, and the association of non-histone chromatin proteins – like histone modification modifiers- with the chromatin fiber (36,37). Especially the chromatin-associated proteome – chromatome- gives particular insight into the processes happening on the fiber, their dynamics, regulation, and capacity to react to molecular challenges or physiological metabolic fluctuations.

The chromatin-bound proteome consists of a plethora of proteins associated with the DNA fiber. The nature of their interaction is manifold and yet to be fully understood, ranging from strong, highly specific protein-protein or DNA-protein interactions (38) to weaker and more indirect interactions (39,40). The association of proteins with DNA in the cell can happen directly, by amino acid side chains of proteins interacting directly with DNA bases but also indirectly by proteins associating with the DNA shape and being fine-tuned by its alterations (41). Additionally, proteins can associate with the fiber by interacting with other protein-binding proteins. A distinction and decoupling of features is important to distinguish the influence of the different contributions to protein binding.

In DREX assembled chromatin, proteomic studies have revealed that DNA replication factors, DNA repair factors but also all proteasomal subunits reproducibly bind assembling chromatin fibers (28,35). The studies also showed dynamics over the observed time, with different factors, like PCNA, associating with early (15 min assembly) chromatin and others, like Ku80, binding more mature (4 h assembly) chromatin. These findings highly correlated to NCC *in vivo* data, confirming the DREX assembly system as a good proxy for *in vivo* observations. This has established the DREX chromatin assembly method as a stellar candidate to investigate the response of the chromatin assembly proteome to molecular challenges like small molecule inhibition and metabolic challenges like nutrient shortage.

1.2.2.1 Methods to investigate the proteins on the fiber

Different methods are available to study proteins binding to the chromatin fiber. Different systems, *in vitro* and *in vivo* can be used to assemble chromatin. Then, antibody detection, antibody-based proximity labeling, or nuclear base tagging allow for readout via immunofluorescence, sequencing for ChIP-Seq experiments, or mass spectrometry to detect and identify proteins on the fiber as well as their dynamics.

1.2.3 Histone modifications and chromatin assembly

1.2.3.1 Methods to investigate histone modification

Historically, histone modifications have been detected and analyzed by antibodies. However, this still widely used method has extensive limitations. These include a high prevalence of cross-reactivity caused by the similar chemical structure of modifications (e.g. mono- or di-methylation of an amino acid) or strongly resembling flanking sequences (42). Additionally, proximity to other modifications in combinatorial histone modifications can influence antibody recognition and binding (43). As many methods, like Chromatin immunoprecipitation + sequencing (ChIP-Seq), CUT&RUN, and CUT&Tag rely strongly on antibodies for analytic readout, these limitations need to be kept in mind when interpreting results, especially in complex environments (44,45).

As an orthogonal technique to analyze histone modifications, mass spectrometry has emerged. Detection and quantification of specific peptides allows for detection as well as relative or even absolute, when combined with isotopically labeled spike-in controls, quantification of histone modifications (5,46). Nevertheless, this method has its limitations including its inability to determine the position of the modification of DNA sequence but also to other modifications. Mass spectrometry has led to the discovery of many new histone modifications.

1.2.3.2 The diversity of histone methylations

Histone methylation and acetylation are the most common modifications and occur mostly on defined lysine sites of the histone tails (Figure 5). Phosphorylation of serines (S), threonines (T), and tyrosines (Y) and Ubiquitination of lysines (K) are also well established as modifications (47). Additional modifications including citrullination (48,49), biotinylation (50,51), sumoylation (52,53), ADP-ribosylation (54,55), isomerization (56), lactylation (57), propionylation (58), butyrylation (59) and crotonylation (60) have been suggested more recently, enriching the diversity of histone modifications. The list of histone modifications is constantly growing and knowledge about the structural and functional impact of the modifications is expanding, even suggesting further modifications (61). Several amino acid sites can be modified with two or more different modifications.

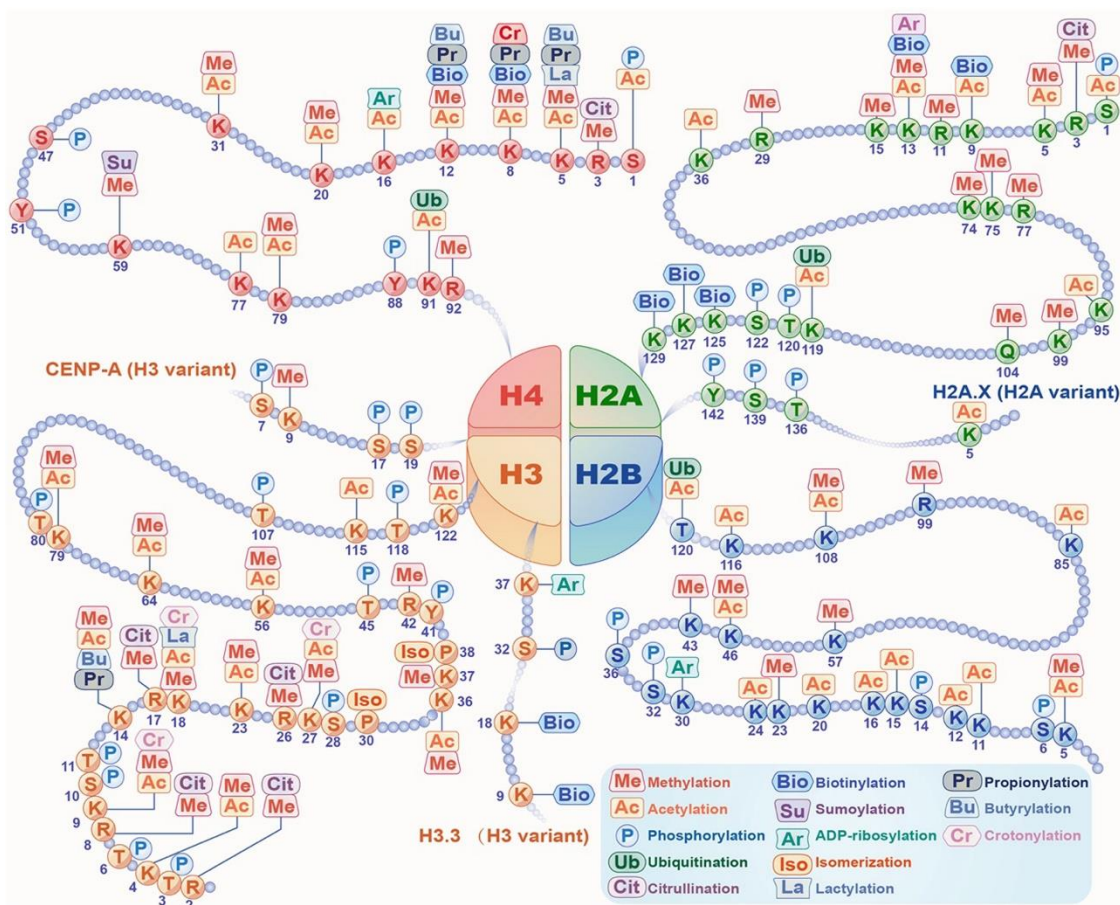


Figure 5 Histone modification overview

Schematic overview of histone modifications on canonical and histone variants. Various modifications can occur on the N-terminal tails of the histones. This figure is reprinted from Yang et al. (47) under the terms of the Creative Commons Attribution-Noncommercial-No Derivatives License (CC BY NC ND).

To sum it up, different methods to investigate histone modifications have been established most prominently antibody-based techniques and proteomics. Combined with different methods for chromatin assembly and other analysis methods like Next-Generation Sequencing (NGS), the readout can e.g. be done in a time, interaction, and condition-based manner. Histone modifications exhibit a great diversity, underlining their multifaceted and fine-tuning roles as well as their close link to metabolism.

1.2.4 Bottlenecks of further investigation of protein binding modes during chromatin assembly

Historically, it has proven difficult to decouple the influences of sequence binding from the recognition of shape for DNA-protein interactions experimentally (62,63). The association with DNA is mediated by sequence-specific binding in the major groove and by less specific interaction via the minor groove, recognizing the overall shape and electrostatic surface of B-DNA (64). Sequence-specific binding is mediated by hydrogen bonding between the side chains of the amino acids of proteins and the base pairs, as each base pair combination has a unique pattern of hydrogen donors and acceptors exposed in the major groove (65). Interestingly, the binding of protein in the major groove is primarily enthalpy

driven, favored by the formation of hydrogen bonds, while the binding of proteins to the minor groove is not favored with regards to enthalpy, however energetically overall favored due to the compensation by entropic forces resulting from the displacement of water (66). Very often, both modes, sequence, and shape-dependent, have a contribution to the final binding of a protein. To overcome the main bottleneck to study the influence of different features on a protein's chromatin binding, the deconvolution of the effects of features, we used foldamer, a DNA mimic that mimics only the shape of DNA without sequence information to investigate binding modes of different proteins on the chromatin fiber.

1.2.4.1 Foldamer, a stable DNA mimic

Due to the unique feature of the synthetic mimic foldamer, mimicking only shape but not carrying any sequence information were able to selectively investigate and interfere with only one binding mode. This allows us to broaden the knowledge base to decipher the contribution different binding modes have for specific proteins, unraveling how they might bind to the chromatin fiber. Additionally, foldamer is highly stable and insensitive to DNases, RNases, and proteinases.

The DNA mimic “foldamer” is a novel compound mimicking the shape of B-DNA while not carrying any sequence information. It was designed, synthesized, and previously described by the Huc group (67). The molecule is an oligoamide, based on alternating monomers $^m\text{Q}^{\text{Pho}}$ and Q^{Pho} (Figure 6 A). The size of the foldamer used in this project is a 16x repeat of the $^m\text{Q}^{\text{Pho}}\text{-Q}^{\text{Pho}}$ dimer, corresponding to a 16bp DNA molecule. The polymer adopts a B-DNA-like double helical structure (Figure 6 B) with an electronegative charge surface resembling that of double-stranded DNA. Due to their structure, the foldamer has been proposed to bind and interfere with the function of DNA-binding proteins, specifically those that recognize the DNA shape.

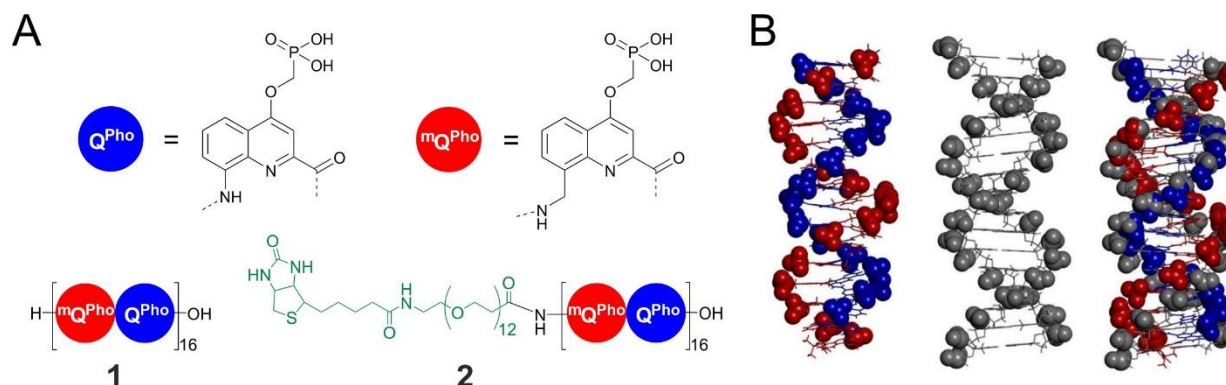


Figure 6 Chemical structures of DNA mimic foldamers

(A) Formulae of amino acid monomers Q^{Pho} (blue) and $^m\text{Q}^{\text{Pho}}$ (red) used to produce the foldamer sequences 1 and 2. Foldamer 2 is functionalized at the N-terminus with biotin (green) for pull-down experiments. (B) Crystal structure of the side chain protected $(^m\text{Q}^{\text{Pho}}\text{-Q}^{\text{Pho}})_{16}$ DNA mimic overlaid with the 16 bp B-DNA duplex $\text{d}(\text{ACTGAACGGCTACGTA})_2$ shown in grey (67). Figure by Dr. Valentina Corvaglia, reprinted with permission.

The foldamer DNA mimics combine the idea of synthetically mimicking DNA shape with an orthogonal synthetic molecule and interfering with DNA-based processes using oligoamides. On the one hand, naturally occurring DNA mimic proteins, that also mimic the DNA shape, are known to bind and thereby interfere with DNA binding proteins (68,69). Additionally, certain oligoamides, that

do not mimic DNA shape, have been proposed to bind to DNA and to interfere with gene expression *in vivo* (70-72).

Promising preliminary findings support the hypothesis for the DNA mimic foldamer functioning as a process-interfering agent: binding and interference with purified HIV-1 and Top1 proteins were shown *in vitro* (67,73). None of the foldamers have been studied in a complex yet systematic approach.

1.3 Metabolism and its link to chromatin assembly

1.3.1 Metabolism and histone modifications

Metabolism is linked to chromatin structure and function in many ways, most prominently via histone modification. Many of the recently discovered histone modifications have the structure of products and intermediate products of metabolic pathways (47). This displays the tight connection between metabolism and the modification of histones. The impact of metabolite availability on the rate of chromatin modification depends on the kinetic and thermodynamic parameters of the corresponding modifying enzyme, e.g. K_m value (74). The influence appears to be reciprocal with not only metabolite availability having a major impact on histone modification states but also histone modifications serving as landing platforms for metabolic enzymes like methylases but also the proteasome, facilitating the formation of metabolic microenvironments (75,76). Additionally, chromatin with its size and high mass of histones in the cell has been proposed to serve as a methyl sink and acetate reservoir (77).

Often, the effect of metabolic changes to chromatin is investigated by treatment with inhibitors and subsequent readout of the system's response by studying the impact on histone modifications, protein binding, or transcription, e.g. by MS, ChIP-Seq, and RNA-Seq (78,79). To determine metabolite levels and flux, thereby properly grasping the metabolic state of a system, MS and less frequently NMR are employed (80-82).

1.3.2 Acetylation

Histone acetylations are set on lysine residues of the N-terminal histone tail. They are the result of a dynamic equilibrium facilitated by acetylation, catalyzed by enzymes called histone acetyltransferases (HATs), and deacetylation by histone deacetylases (HDACs). Acetylation marks are considered highly dynamic with HATs and HDACs facilitating constant turnover. The acetylation reaction requires the cofactor acetyl-CoA. Acetylation of lysines neutralizes their positive charge, thereby weakening the interaction of the corresponding histone with the negatively charged DNA and making the chromatin more accessible (83-85). Histone acetylation mainly is involved in opening chromatin to set a background for transcriptional regulation in bulk, as well as setting local, targeted activating marks at enhancer and promotor elements (86,87).

HATs are often discovered for their ability to acetylate histones but are also able to acetylate a plethora of non-histone proteins. Thousands of proteins are acetylated in the cell with acetylation being a major regulatory and signaling modification in metabolism. Therefore, it is not surprising that histone acetylation and metabolism are intimately linked. Additionally, acetyl-CoA has a dual role as a central metabolite and substrate for histone modification. It has been proposed that acetyl-CoA might act as a rheostat for nutrient availability, partially by affecting histone acetylation (88,89).

1.3.3 Methylation

Histone methylation can happen at lysine and arginine residues and does not change the charge of the histone. Lysines can be mono-, di- and tri-methylated on the free amino group while arginine can be mono or symmetrically or asymmetrically di-methylated at its side chain amino groups (90,91). The methyl groups are added enzymatically in the cell by histone methyltransferases (HMTs) which use their cofactor S-adenosylmethionine (SAM) as a methyl donor, yielding SAH as a byproduct of the methylation reaction. While methylations are considered a stable mark, another group of specialized enzymes, demethylases, can remove methylation from histones (92).

Histone methylation changes chromatin accessibility and gene expression. Which amino acid site on the histone is methylated and to what degree is critical for its effect. While methylation of H3 at K4, K36, and K79 is associated with active chromatin, H3 K9 and K27 are generally associated with repressed chromatin. H4 K20 methylation is even ambiguous in its effect on gene expression depending on context (93,94). Thereby, histone methylation greatly influences different stages of transcription, thereby strongly influencing cell physiology and development (95,96).

1.3.3.1 The Methionine Cycle

While histone methylation is a well-known modification with extensively studied influence on chromatin accessibility and gene expression, its link to metabolism and role as an indicator of metabolic states of a cell or organism has only been studied more thoroughly in the last decade (93). Methylations happen in all cell compartments and histone methylation is a subclass of all methylation reactions, sharing their common essential cofactor, SAM. Due to SAM's involvement in numerous metabolic pathways, likely, histone methylation is tightly connected to a system's metabolic state.

All histone methylases use SAM as a cofactor. SAM is produced and regulated by the one-carbon metabolic pathway, which uses nutrients as substrates (Figure 7). Therefore, in cells, SAM levels are influenced by dietary intake or other means of nutrient availability. SAM generation from methionine and ATP is conserved across all species and catalyzed by methionine adenosyltransferase enzymes (MATs) (97). In *Drosophila*, only one MAT exists, called S-Adenosyl methionine synthetase (SAM-S) (98). HMT but also other methylation transferases use SAM as the donor for the activated methyl group. During this, SAM is hydrolyzed, yielding the methylated species and SAH.

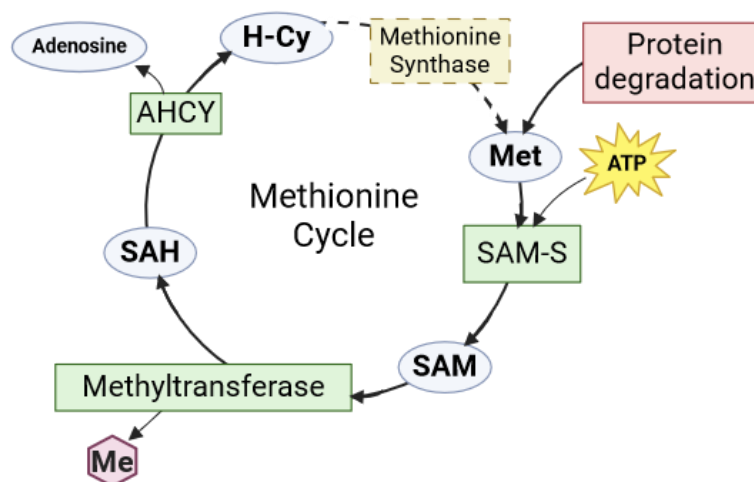


Figure 7 Scheme of methionine cycle in *Drosophila*

Methionine is mainly supplied by the degradation of methionine-containing proteins in *Drosophila*. S-adenosyl methionine (SAM) is catalyzed by SAM-S (S-Adenosyl methionine synthetase) from Adenosine triphosphate (ATP) and Methionine (Met). SAM is a Cofactor and methyl donor for Methyltransferases. Methyltransferase reactions yield a methylated species and S-Adenosyl homocysteine (SAH). SAH is degraded by Adenosyl homocysteine lyase (AHCY) into adenosine and Homocysteine (H-Cy). H-Cy might be recycled by a methionine synthase in *Drosophila* to generate methionine. Figure created with biorender.com

SAH is an inhibitor of the methylation reaction. It binds to the SAM binding pockets of MTases with similar affinity as SAM, creating a negative feedback loop. Due to this competition, the SAM/SAH ratio is considered an indicator of the methylation capacity of an environment or cell (99,100). SAH is removed from the cycle by being converted to homocysteine and adenosine by the enzyme Adenosylhomocysteinase (AHCY), this reaction requires NAD (101,102). In some species, methionine can be recycled from SAH. Other species lack these pathways. In *Drosophila*, these pathways are so far under investigated, revealing potential candidates by homology but not any definitive proof that these pathways exist in the species (103,104).

Intracellular concentrations of SAM are estimated at around 10 μM but fluctuate greatly (10-100 fold) under normal physiological conditions (77). Furthermore, a significant nuclear heterogeneity in methylation capacity is expected due to increased local fluxes of SAH caused by methylations e.g. DNA methylation at replication sites and co-transcriptional mRNA cap methylation (105-107). Additionally, the prevalence of liquid-liquid phase separated, membrane-less compartments contributes to heterogeneity in metabolite levels (75). AHCY is recruited to specific sites on chromatin, potentially to regulate SAH levels and thereby sustain methylation efficiency but still lacks mechanistic or protein-protein interaction-based insights (108-110).

2 Aims of this Study

General Aim: Challenging chromatin assembly in vitro to investigate chromatin structure and function.

Sub Aims:

- G.1** Dissection of the impact of changes in the molecular environment on the different levels of chromatin assembly in the preblastoderm *Drosophila* embryo extract (DREX) system.
- G.2** Investigation of the effect on chromatin structure and function on multiple levels using mass spectrometry, including the incorporation of nucleosomes, chromatin-bound proteome, histone modification, and metabolite concentrations and dynamics.
- G.3** Establishment and benchmarking of new methods and protocols necessary for the further characterization of the extract and observation of the changes caused by the challenges.
- G.4** Establishment of in vivo controls to further validate in vitro data.

Specific Aim Project 1 (Chromatin assembly and protein binding):

Challenging protein binding dynamics during chromatin assembly with a small molecule DNA mimic.

Sub Aims:

- 1.1.** Investigation of the potential of foldamers to interfere with chromatin-bound proteome in vitro using DREX extract.
- 1.2.** Identification of specific proteins affected by foldamer treatment in vitro.
- 1.3.** Identification of foldamer interactome in the complex in vitro.
- 1.4.** Validation of effect of foldamer on chromatin-bound proteome by in vivo study using subcellular fractionation.
- 1.5.** Investigation of physiological implications of protein binding interference by foldamer via flow cytometry-based cell cycle analysis in vivo.

Specific Aim Project 2 (Chromatin assembly and metabolism):

Investigation and challenging of metabolic coupling of histone 4 lysine 20 monomethylation (H4K20me1) during chromatin assembly in vitro

Sub Aims:

- 2.1.** Determination of the DREX proteome.
- 2.2.** Establishment of mass spectrometry-based metabolome analysis in DREX and determination of DREX metabolome.
- 2.3.** Identification of metabolic proteins present on assembled chromatin fibers in vitro.
- 2.4.** Proof of metabolic activity in DREX extract using isotopic labeling and histone modification analysis.
- 2.5.** Identification of changes in histone modifications upon change in metabolite levels in DREX.
- 2.6.** Identification of changes in chromatin-bound metabolic proteome upon changes in metabolite levels in DREX.

3 Project 1: Chromatin assembly and protein binding

3.1 Graphical abstract

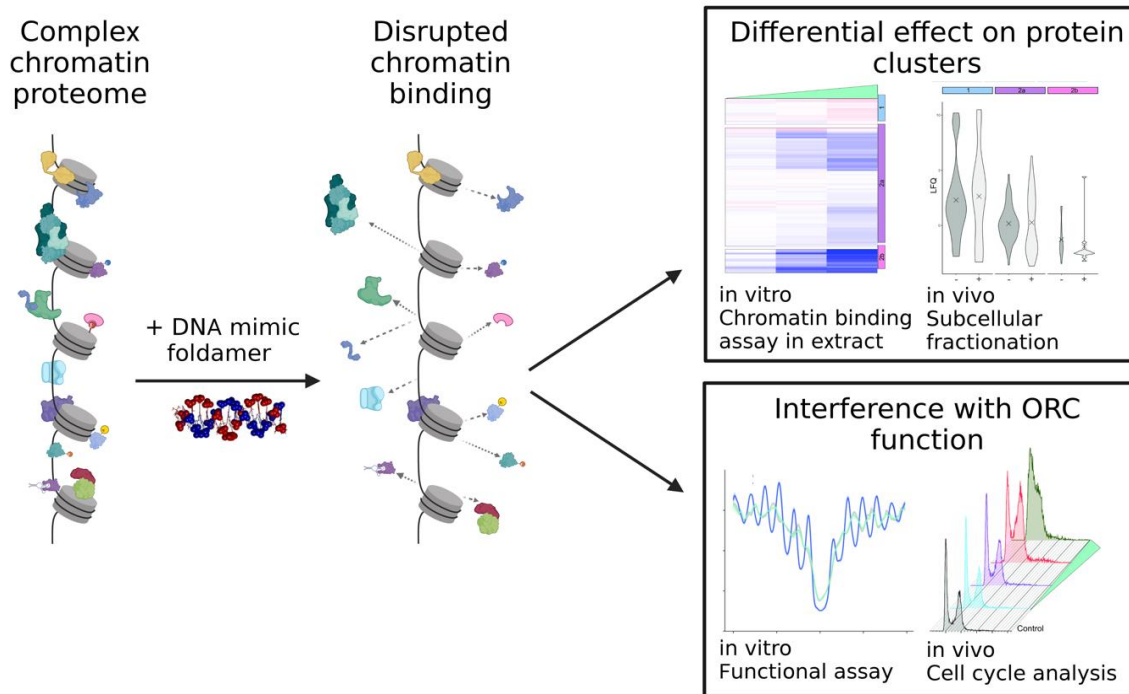


Figure 8 Graphical abstract of project 1

DNA mimic foldamers interfere with chromatin assembly and reveal insights into protein binding modes. Graphical abstract from Kleene et al. (111) by permission of Oxford University Press.

“The use of synthetic chemicals to selectively interfere with chromatin and the chromatin-bound proteome represents a great opportunity for pharmacological intervention. Recently, synthetic foldamers that mimic the charge surface of double-stranded DNA have been shown to interfere with selected protein–DNA interactions. However, to better understand their pharmacological potential and to improve their specificity and selectivity, the effect of these molecules on complex chromatin needs to be investigated. We therefore systematically studied the influence of the DNA mimic foldamers on the chromatin-bound proteome using an in vitro chromatin assembly extract. Our studies show that the foldamer efficiently interferes with the chromatin association of the origin recognition complex in vitro and in vivo, which leads to a disturbance of the cell cycle in cells treated with foldamers (Figure 8). This effect is mediated by a strong direct interaction between the foldamers and the origin recognition complex and results in a failure of the complex to organize chromatin around replication origins. Foldamers that mimic double-stranded nucleic acids thus emerge as a powerful tool with designable features to alter chromatin assembly and selectively interfere with biological mechanisms.”

- Direct citation of the Abstract of Kleene et al (111) that is based on the findings described in this chapter

3.2 Results

3.2.1 Foldamer interferes with in vitro chromatin-bound proteome

We aimed to investigate whether the foldamer can interfere with chromatin formation using the in vitro DREX-mediated chromatin assembly. First, we tested whether the addition of foldamer interferes with the generation of nucleosomal ladders, which indicates the integration of histones and regular spacing of nucleosomes during the DREX chromatin assembly reaction. As visible in the Agarose gel of Mnase-digested chromatin, we detected mostly no change in array formation upon foldamer presence (Figure 9). Only for high foldamer concentrations, we observed a slight reduction in array regularity. Based on the findings we assumed that general aspects of chromatin assembly were not disturbed by foldamer, allowing us to proceed to look at potential changes on the assembling chromatin fiber in more detail.

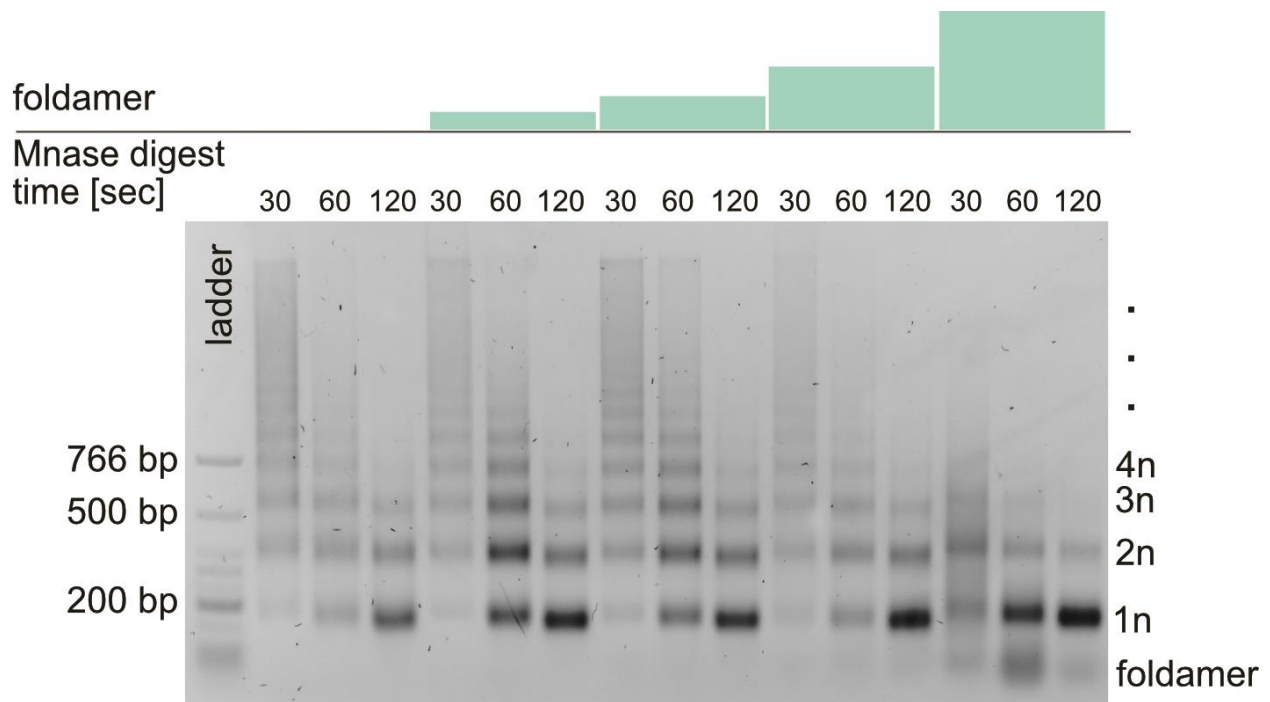


Figure 9 In vitro chromatin assemblies in the presence of foldamer

Agarose gel of Mnase digested chromatin assemblies in the absence of foldamer or the presence of one, two, four, or eight equivalents of foldamer compared to DNA. Labeling reflects a nucleosomal ladder containing a regular array of nucleosomes (1n/2n/3n/4n...) as well as a band of residual free foldamer.

Next, we wanted to look at the process in more detail, leveraging our ability to quantify the chromatin-bound proteome using mass spectrometry. Therefore, we investigated whether increasing concentrations of foldamer would differentially interfere with the proteome associated with the fiber during in vitro chromatin assembly. Hence, we incubated linearized and immobilized DNA with a chromatin assembly extract made from early *Drosophila* embryos in the absence of or presence of different concentrations of foldamer, ranging from a 1:1 to 1:8 ratio of DNA: free foldamer (Figure 10).

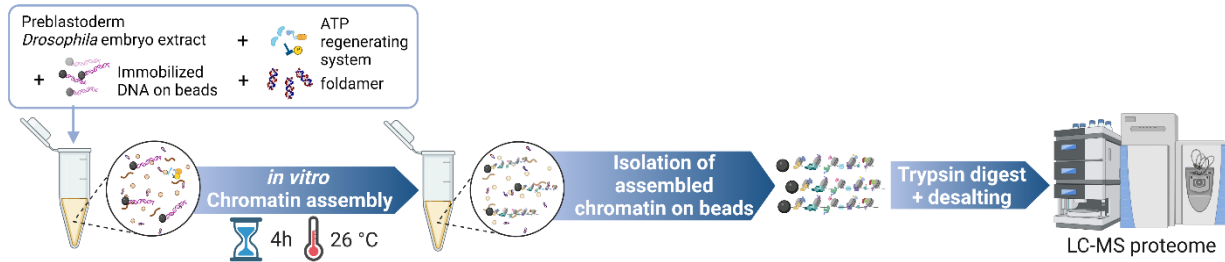


Figure 10 Schematic diagram of experimental design for *Drosophila* embryo extract assisted *in vitro* chromatin assembly.

Immobilized DNA on streptavidin beads, ATP generating system, and foldamer were added to DREX extract, incubated for 4 h at 26 °C, and assembled chromatin was isolated, washed, and prepared for mass spectrometry measurement. Modified from Kleene et al (111) by permission of Oxford University Press.

To differentially assess the effect of foldamer on specific proteins, we first leveraged our ability to identify and quantify the proteins bound to chromatin. After rigorous filtering, requiring at least 2 out of 3 valid values for the LFQ protein intensities per condition (each DNA: foldamer ratio is one condition), we observed 1993 proteins total. We then defined the 168 proteins significantly enriched (FDR=0.05) on the isolated chromatin fiber over a beads-only control in the absence of foldamer condition as “chromatin binders” (Figure 11 left). This is a novel approach, introducing a threshold when most analyses before looked at the proteins bound to chromatin without a strict cutoff ((28) and analysis of chromatin-bound proteome upon proteasomal inhibition (4.2.6) in this thesis). The approach used here allowed us to reduce the complexity of the dataset and follow a discrete set of proteins downstream, to identify the specific proteins majorly affected by foldamer.

After the definition of the “chromatin binder” protein set, we performed a chromatin assembly experiment in DREX, where we could quantify the foldamers' effects on chromatin composition. Here, we added increasing amounts of foldamer, ranging from a 1:1 to a 1:8 mass ratio of immobilized DNA to foldamer to the standard DREX assembly reaction. Intensities of all proteins were normalized to their intensities in the 1:1 DNA: foldamer ratio condition. We focused our analysis on the relevant “chromatin binders” subset. By unsupervised clustering based on Pearson correlation, we identified 2 main clusters, separated by their susceptibility to the presence of foldamer (Figure 11, heatmap on the right, Cluster 1 and Cluster 2).

Cluster 1 contains 33 proteins that were not affected in their chromatin binding in the presence of the foldamer. Proteins found in this cluster include, as expected, the core histones, and a protein that can substitute for the absent linker histone H1 in preblastoderm embryos in *Drosophila*, the HMG-D protein (112,113). Additionally, the histone chaperone Caf-1 and all subunits of the MRN DNA damage repair complex (mre11, rad50, and nbs (114)) and the complete heterotrimeric RPA complex (RPA3/Rpa14, RPA2/Rpa32, and RPA1/RpA-70) remained stably bound to the chromatin fiber, independent of foldamer concentration. Intriguingly, two major *Drosophila* topoisomerases (Top3 α and Top2) were also identified to be unaffected in their chromatin binding behavior by the addition of foldamer. The protein intensities for Top1 suggest an inhibited binding of Top1. However, in our dataset, Top1 does not appear in the cluster as it was not defined as a “chromatin binder” initially due to abundant missing values and thereby eliminated by filtering. In conclusion, Cluster 1 demonstrates that not all proteins are affected by foldamer at the concentrations tested in this assay.

Cluster 2 contains the 136 proteins affected in their binding to chromatin in a foldamer-concentration-dependent manner. The higher the concentration of foldamer, the lower the intensity of those proteins on the chromatin fiber. To differentiate the effect further, we built sub-clusters via the Euclidian method using unsupervised clustering. This yielded 2 distinct clusters, separating the proteins by the intensity of the effect of foldamer on their binding. The Clusters contain either proteins that were mildly (Cluster 2a) or strongly (Cluster 2b) affected by the increasing concentration of foldamer (Figure 11). The 113 proteins in Cluster 2a that were mildly affected by foldamer contain numerous known chromatin-associated interactors, ranging from structural maintenance of chromosomes proteins, including SMC1 and SMC2, to subunits of the condensin and cohesion complex and multiple other specialized proteins involved in different forms of DNA repair (35,115). Intriguingly, the 23 proteins found in Cluster 2b, strongly affected by foldamer, included transcription and DNA repair factors, several subunits of the Sin3a transcriptional repressor complex, and, importantly, all subunits of the ORC.

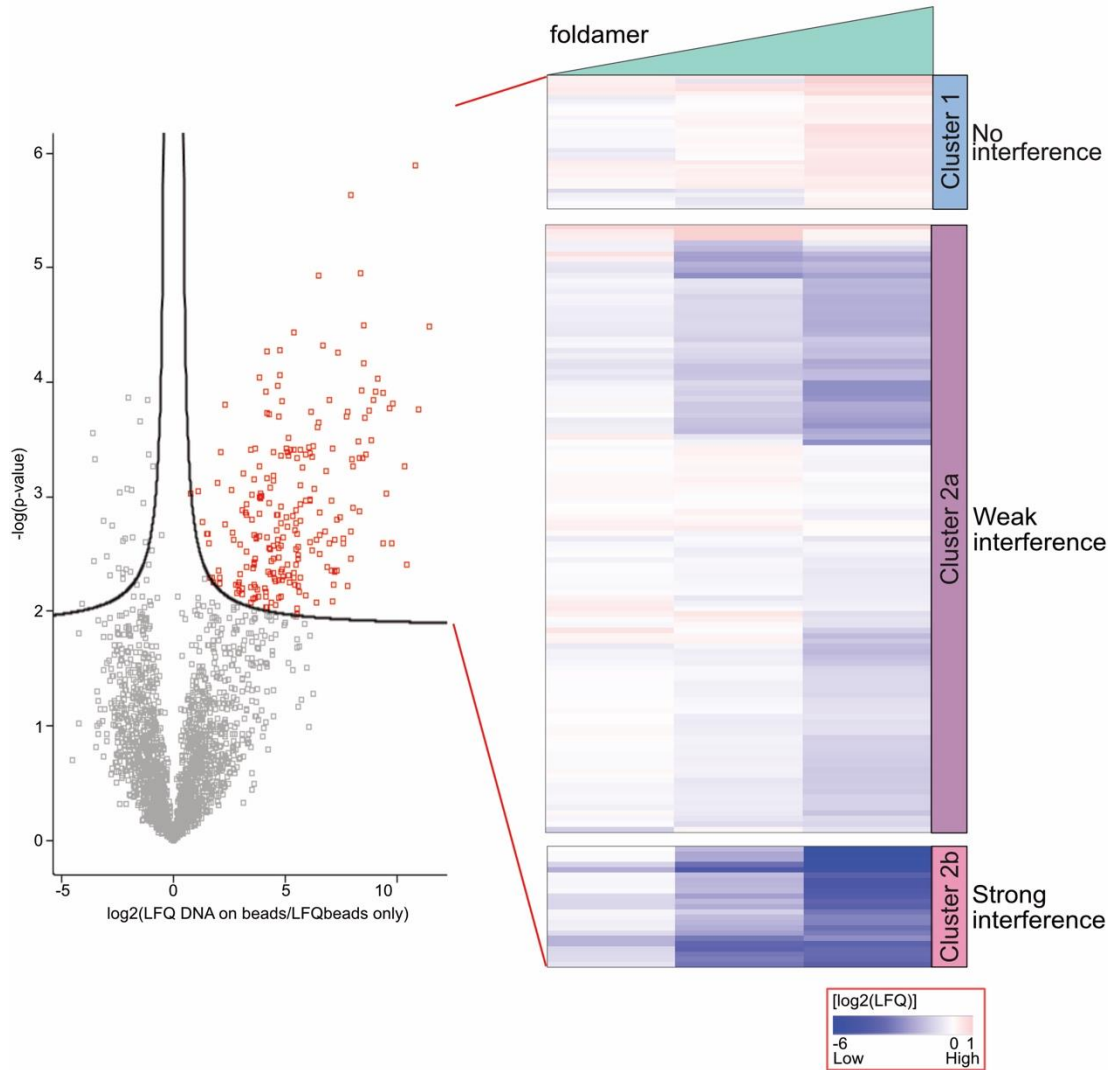


Figure 11 “Chromatin binders” and their susceptibility to interference in binding by foldamer

(Left) Volcano plot for proteins on biotinylated DNA against beads only control after 4 h incubation in DREX extract in the absence of foldamer 1. Proteins significantly enriched on DNA over beads were defined as “chromatin binders” in the assay. $N=3$ FDR=0.05. (Right) Heatmap of proteins binding to DNA fiber during chromatin assembly (rows) against their mean intensities at different concentrations of free foldamer present (columns). Unbiased Pearson clustering results in 2 groups: proteins whose binding is not interfered with by foldamer 1: “No interference (Cluster1)” (blue) and interfered by foldamer 1 “Interference (Cluster2)”. Subsequent Euclidian clustering separates Cluster 2 into Cluster 2a “weak interference” (dark purple) and Cluster 2b “strong interference” (pink). The DNA: foldamer 1 weight ratio ranged from 1:1 to 1:8. $N=3$. Modified from Kleene et al (111) by permission of Oxford University Press. Data based on Datasets 2 and 3.

While the foldamer mimics double-stranded DNA, we also wanted to investigate the difference in interference potential between the two molecules, foldamer, and DNA, in a complex biological setting. To compare the effects, we performed the same assembly reactions while using a 16bp control DNA as a free competitor instead of a foldamer. We found no effect on chromatin binding for all proteins when the 16bp control DNA is used as a competitor to chromatin binding. The difference in the foldamer effect is most striking for those proteins being strongly affected by the foldamer, like the ORC complex (Figure 12).

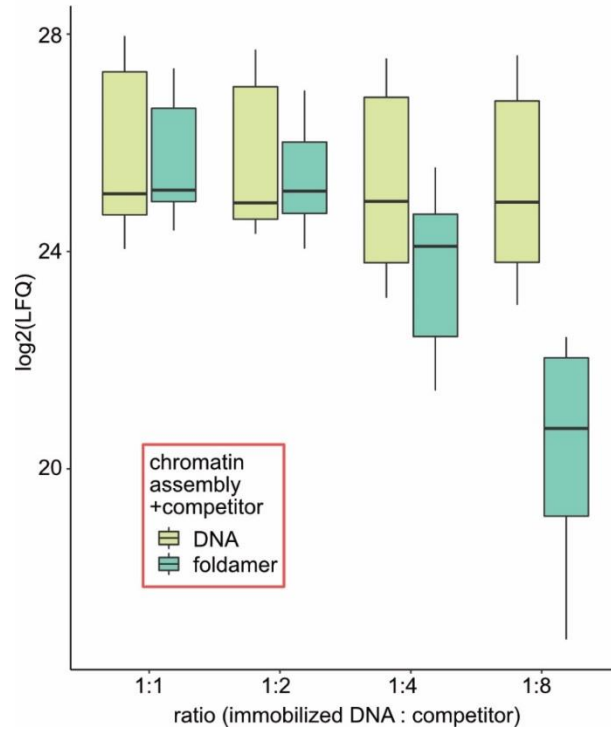


Figure 12 Susceptibility to interference in the binding of ORC1-5 proteins to foldamer or DNA control

Boxplot of chromatin binding of proteins belonging to the origin recognition complex ORC 1-5, all located in Cluster 2b “strong interference”. Titration of chromatin binding using the competitor's foldamer **1** (dark green) or free 16 bp DNA control (lime green). Each box consists of the N=3 technical replicates of all 5 proteins. Modified from Kleene et al (111) by permission of Oxford University Press. Data based on Dataset 3.

To sum it up, we tested the sensitivity of the chromatin-bound proteome to foldamer during assembly in the DREX system. Here, we identified differential effects on the proteins that can be clustered according to their susceptibility to foldamer interference. Additionally, we report a control experiment showing that a double-stranded DNA control of similar length does not affect protein binding as observed for the foldamer. Overall, this allows for interesting insights into the protein binding modes to chromatin during assembly in a complex system by an initial categorization and characterization of proteins based on their sensitivity to foldamer.

3.2.2 The foldamer interactome

To identify the proteins and proteins complexes binding to the foldamer in complex mixtures, we performed a pull-down with biotinylated foldamer immobilized on magnetic beads in DREX and analyzed the bound proteins using LC-MS (Figure 13 A). We revealed 640 significant proteins binding to the biotinylated foldamer when compared to a background control of a beads-only pulldown (Figure 13 B). This number is higher than the amount of different proteins enriched on immobilized DNA in our previous experiment (Figure 11).

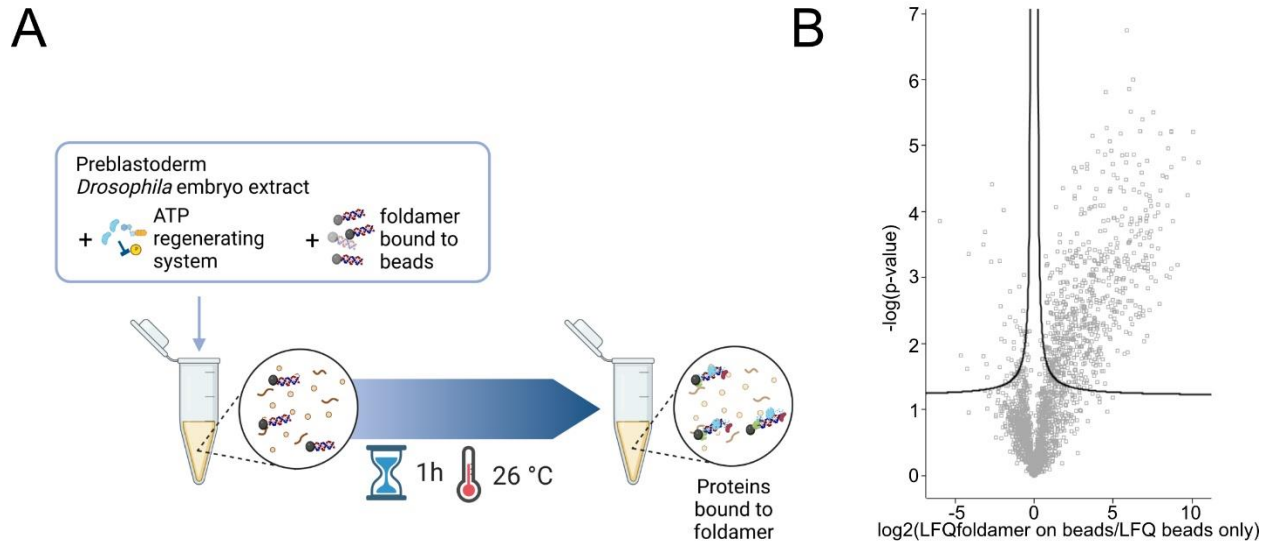


Figure 13 Foldamer interactome experiment

(A) Schematic diagram of the experimental flow of foldamer pulldown from preblastoderm *Drosophila* embryo extract (DREX). (B) Volcano plot for Pulldown of proteins with 1 µg biotinylated foldamer from DREX. N=3 FDR=0.05. Modified from Kleene et al (111) by permission of Oxford University Press. Data based on Dataset 4.

To characterize our set of foldamer-bound proteins, we performed network as well as GO term analysis. First, network analysis was executed using the String Plugin in the Cytoscape software with the following settings for easier visualization: only experimental evidence as input for interaction analysis, only protein links with the highest confidence interaction score of at least 0.9 and filtering out of all proteins with no interaction. (Figure 14 A). The obtained subset of the foldamer interactome included rRNA-associated complexes and processes, comprised of proteins from the ribosome complex itself to proteins involved in rRNA processing. Furthermore, we detected numerous proteins associated with mRNA, including the spliceosome and other networks of proteins involved in mRNA splicing and mRNA surveillance. Finally, we found many proteins involved DNA associated processes. Here, topoisomerase 1 and the entire ORC complex are among the specific binders. Additionally, proteins of transcription regulation, including the basal transcription machinery as well as transcription factors containing Zn-finger domains and replication factors are among the interactors.

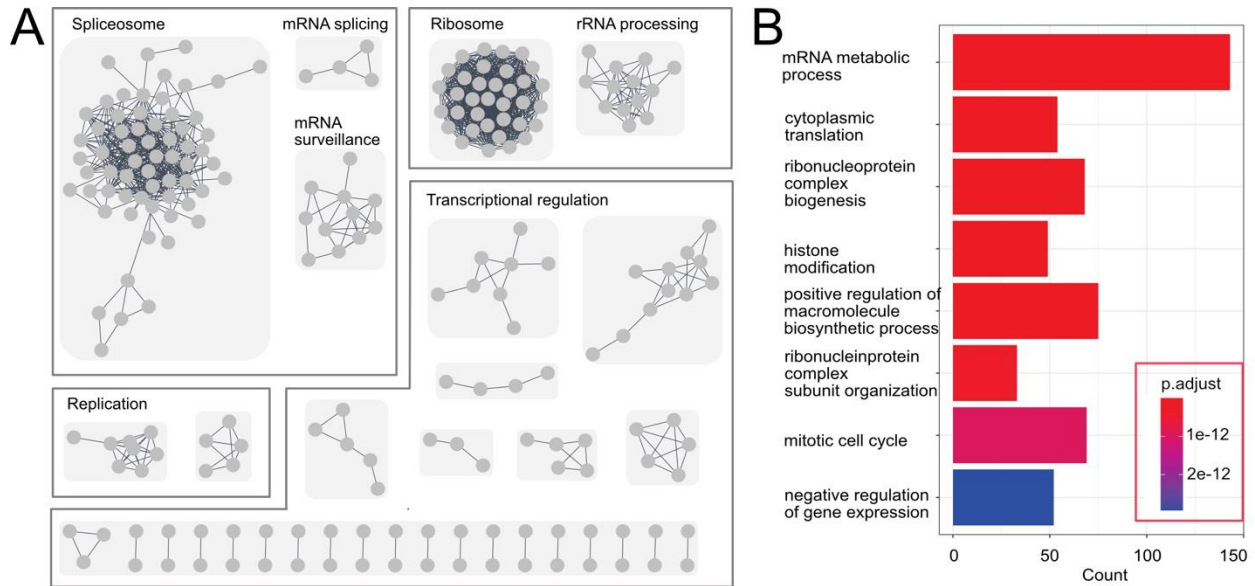


Figure 14 Foldamer interactome analysis

(A) Cytoscape network plot after string analysis (only experimental links, highest confidence interaction score 0.9) for proteins binding to biotinylated foldamer in DREX. (B) Top 8 GO Terms by p-value for all 640 proteins specifically binding to biotinylated foldamer in DREX. Modified from Kleene et al (111) by permission of Oxford University Press. Data based on Dataset 4.

Next, the GO term analysis of all foldamer binders against a whole genome *Drosophila* background was performed with R Studio using the Cluster profiler package. Enriched GO terms were corrected for semantic redundancy using the integrated algorithm with a cutoff of 0.6. Finally, only the top eight GO terms were plotted here with corresponding p-adjusted values (Figure 14 B). This analysis also showed a significant GO term enrichment of processes involving RNA and DNA molecules. Overall, we were able to conclude that the foldamer interactome contains a lot of proteins and protein complexes that bind or are associated with nucleic acids.

3.2.3 The protein subset competed off chromatin fiber by binding to foldamers includes ORC

Having investigated the binding of proteins to foldamer as well as foldamer interference with the binding of proteins to chromatin, we were interested in the overlap of these two sets of proteins (Figure 15). The identified set contained 15 proteins which interference with binding to the chromatin is potentially mediated by binding of the proteins to foldamer.

The proteins in the overlap set were primarily involved in transcriptional regulation, DNA repair, and replication. They include the members of the Sin3a histone deacetylase complex that is involved in response to cell stress (116) and haywire/TFIIH and mrn, two proteins involved in DNA repair pathways (117,118). Importantly, all subunits of the ORC were present in the subset.

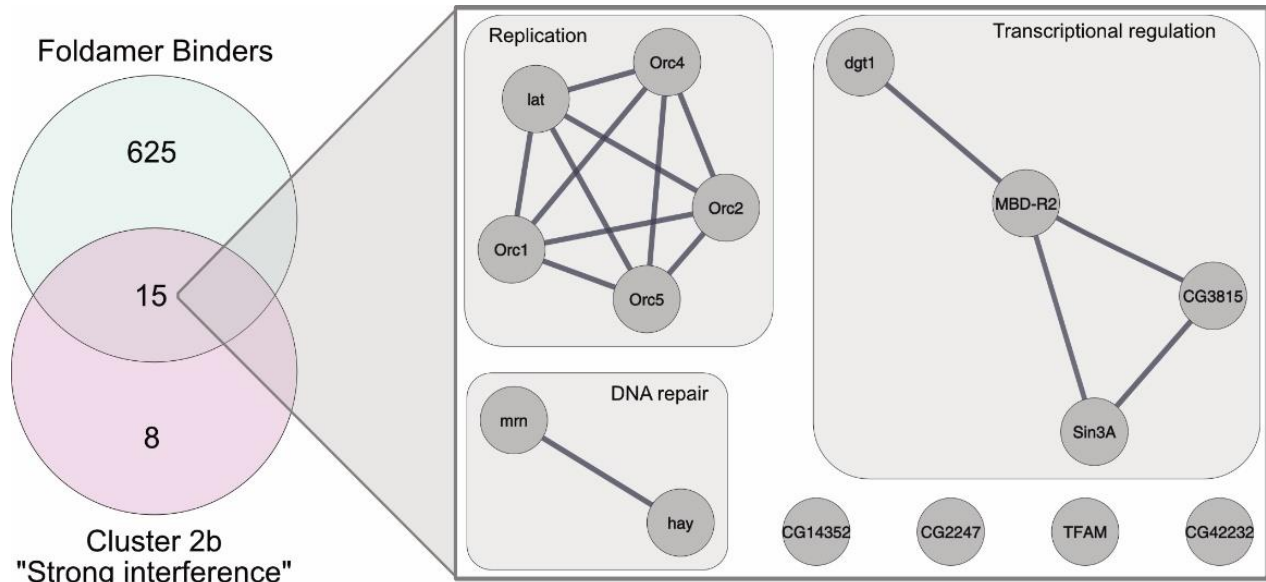


Figure 15 Subsetting for chromatin proteins affected by foldamer

Venn diagram of proteins in the overlap of Foldamer binders set and strong interference Cluster 2b with a box depicting Cytoscape network plot after string analysis (only experimental links, highest confidence interaction score 0.9) for proteins in subset defined by overlap. Modified from Kleene et al (111) by permission of Oxford University Press. Data based on Datasets 3 and 4.

3.2.4 Foldamers also disturb the chromatin-bound proteome in *Drosophila* S2 cells

The promising results for the effect of foldamers across different setups in vitro prompted us to extend our tests to *Drosophila* cells to validate and thereby expand the results in vivo. Not only had we shown the interference of foldamer with chromatin binding for a specific subgroup, but also confirmed the disturbance of Orc function in an in vitro remodeling assay in collaboration with Erica Chacin and Christoph Kurat (111). For our first vivo experiments, to validate the displacement of defined clusters of proteins from chromatin upon foldamer treatment, S2 *Drosophila* cells were treated with 10 μ M foldamer in the medium for 48h to confirm the differential effect on distinct clusters of chromatin binding proteins. After harvest, the cells were subjected to subcellular fractionation, using the Subcellular Protein Fractionation Kit for Cultured Cells (Thermo), resulting in 6 fractions as depicted in Figure 16. We observed that all fractions collected appeared yellow in color, indicating that foldamer (yellow) reached all fractions. All fraction samples were prepared using the Preomics IST kit and then subjected to proteomic measurement and analysis, allowing us to follow protein distribution over fractions upon treatment.

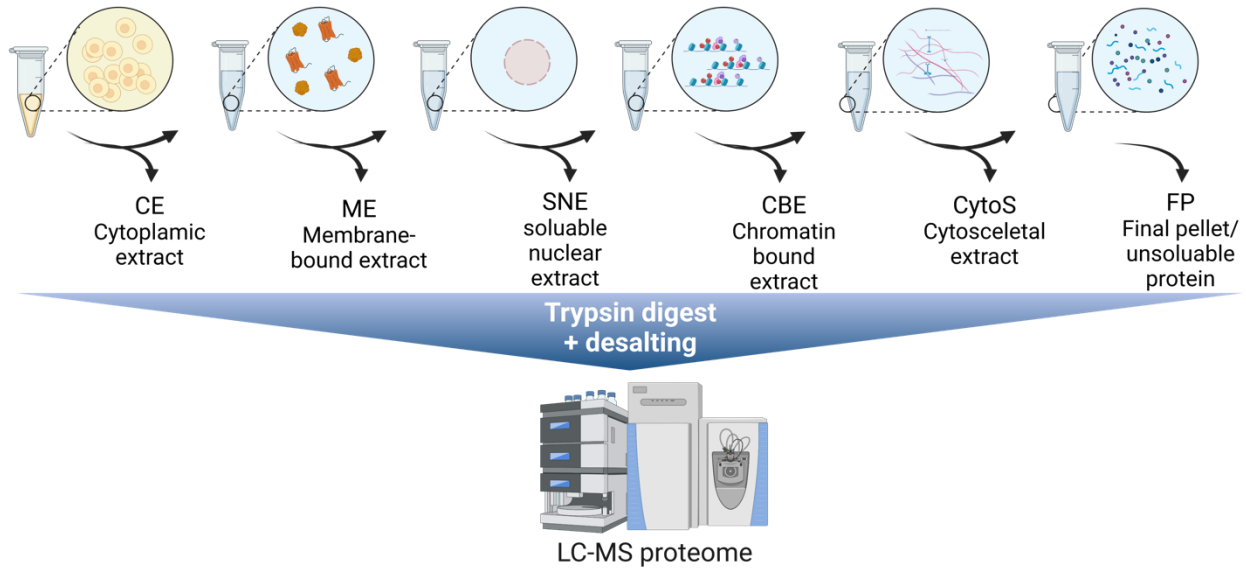


Figure 16 Subcellular fractionation experiment

Schematic diagram of the experimental flow of subcellular fractionation experiment. Modified from Kleene et al (111) by permission of Oxford University Press.

To assess general changes in protein distribution across subcellular fractions by foldamer treatment, the intensities of proteins across all fractions were determined via LC-MS and analyzed. First, we filtered the data to select only proteins with at least 2/3 valid values in at least one fraction across both conditions (treated or untreated) before standard imputation. The subsequent normalization of each protein intensity to the overall mean protein intensity in the condition (control/foldamer) allowed us to compare protein distribution across fractions upon foldamer treatment (Figure 17). We observed no drastic overall change. However, we detected an upshift trend in total protein intensities of the cytoplasmic fraction (CE) as well as of the final pellet (FP) and a slight downshift in the other fractions, most strongly in the chromatin-bound fraction (CBE).

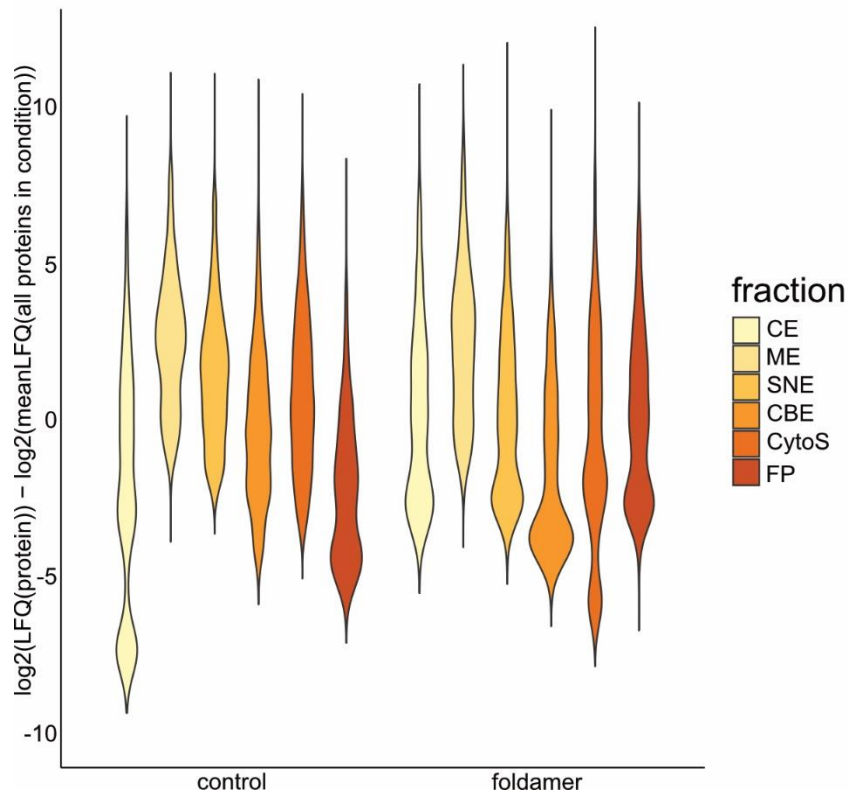


Figure 17 Protein distribution in all subcellular fractions

Violin plot of mean protein intensities in the subcellular fractions in control and upon treatment with 10 μ M foldamer 1. Number of proteins = 3765. N=3. Modified from Kleene et al (111) by permission of Oxford University Press. Data based on Dataset 5.

To specifically address the impact of foldamer treatment on the chromatin-bound proteome (CBE fraction) in vivo, we analyzed the changes in protein intensity for the same proteins we had defined as chromatin binders before in vitro (3.2.1 Figure 11) and followed in the in vitro analysis. Not all proteins detected in vitro could be detected with sufficient consistency across replicates in vivo and were therefore not included in the analysis of the in vivo experiment. Filtering was performed to only keep proteins with 2/3 valid values in at least one condition (control/foldamer) in the chromatin-bound fraction (CBE) to prevent false confidence by imputation. We then grouped the “chromatin binding” proteins according to their clustering in vitro. Next, we compared the shift in intensity of the proteins of each predefined cluster in the fraction upon foldamer treatment (Figure 18 A). Cluster 1 proteins, maintaining their chromatin binding upon foldamer treatment in vitro, did the same in vivo. Not only looking at the delta mean of the whole cluster but also at single proteins in the cluster, this trend is confirmed, e.g. by the histone proteins, where no significant change in protein intensity was measured upon foldamer treatment in cells (Figure 18 B). Just as Cluster 1, so too Cluster 2a being mildly and Cluster 2b being strongly affected by foldamer in vitro, showed the same susceptibility in vivo. Again, looking at specific proteins, the members of the ORC complex confidently detected in vivo showed a significant decrease in the chromatin-bound proteome upon treatment. Cluster 2b significantly separated itself from the other clusters when comparing their change in mean between treated and untreated samples (Figure 18 A). This validates that the cluster assignments of the proteins we yielded

from our in vitro analysis also categorize the chromatin binding proteins' susceptibility to foldamer well in vivo. Altogether, we were able to confirm and validate the cluster-specific effect of foldamer treatment on chromatin-binding proteins in vivo.

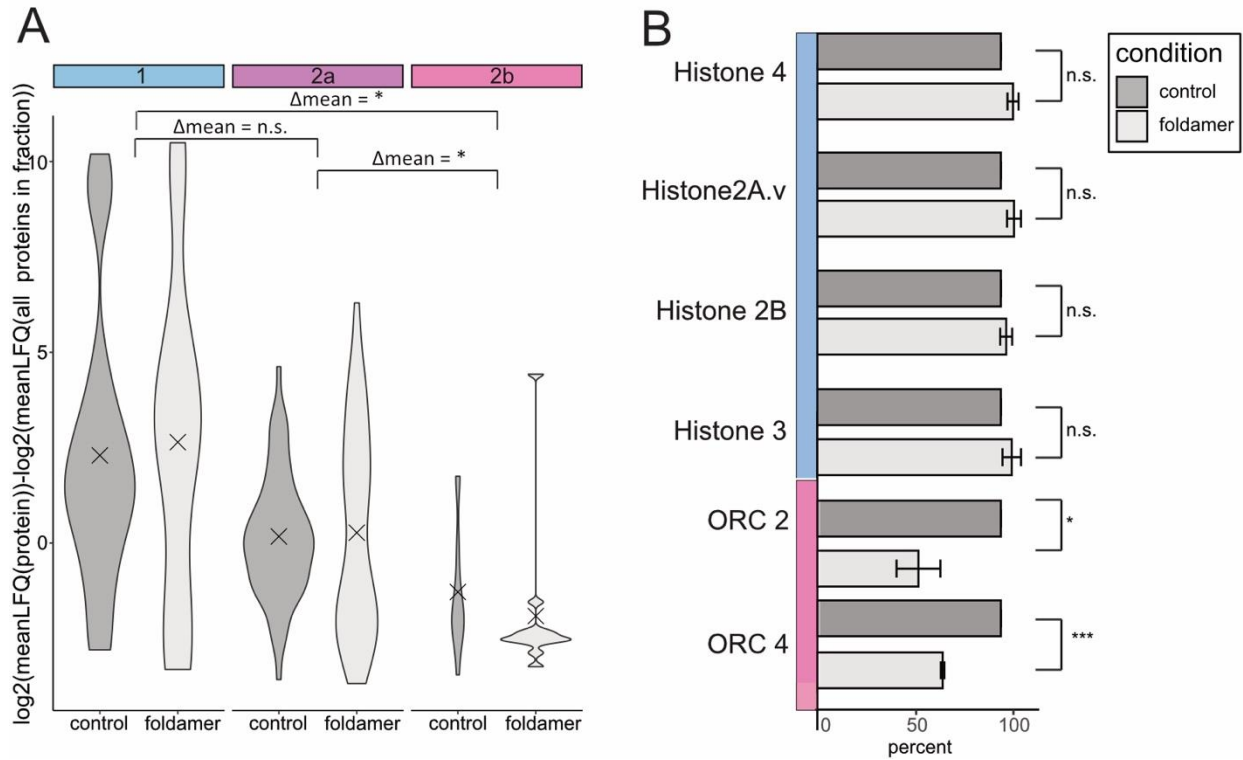


Figure 18 Change in protein intensities of chromatin binders in a chromatin-bound fraction

(A) Violin plot of means of protein intensities in chromatin-bound fraction (CBE) of the proteins appearing in clusters 1, 2a, and 2b as defined by their sensitivity to foldamer in vitro. Cross marks indicate the mean value. Statistics describe the comparison of the mean differences of all proteins in each group with N=3 replicates * = $p < 0.05$, 1way ANOVA + Tukey's test. (B) Bar graph of foldamer effect on protein intensity of selected proteins in chromatin-bound Fraction (CBE). Statistics describe the comparison of normalized mean LFQs upon foldamer treatment in comparison to control N=3 replicates ** = $p < 0.01$, *** = $p < 0.001$, t-test. Modified from Kleene et al (111) by permission of Oxford University Press. Data based on Dataset 5.

3.2.5 Foldamer treatment interferes with cell cycle progression in *Drosophila* S2 cells

Next, we wanted to investigate whether foldamer has an impact on cell cycle progression. Before, we had already shown that Orc binding to chromatin as well as its function of setting up regular nucleosomal arrays is interfered with by foldamer treatment in vitro (111). The Orc complex is essential for successful replication and thereby cell cycle progression. Hence, we tested whether foldamer would interfere with cell cycle progression in vivo.

For this in vivo study, we treated *Drosophila* S2 cells with different concentrations of foldamer in medium (0, 0.01 μ M, 0.1 μ M, 1 μ M, 10 μ M) for 4 h, 24 h, or 48 h (Figure 19 A). We subsequently performed Flow Cytometry after propidium iodide staining to identify the cell cycle stages in the population of living cells in a time and concentration-dependent manner (Figure 19 B). Through quantification of the populations by further gating, we determined the percentage of cells in the S-

Phase (Figure 19 C). Interestingly, upon treatment with 10 μ M foldamer, the proportion of cells in the S-Phase increased to 15% compared to 10% in the untreated control after 24h. For longer incubation, we saw an even more drastic and significant increase to 33% of cells in the S-Phase (11% in untreated control). A similar but weaker trend was observed for treatment with 1 μ M foldamer. Altogether, we were able to successfully show a cell cycle physiological response to foldamer treatment in cells.

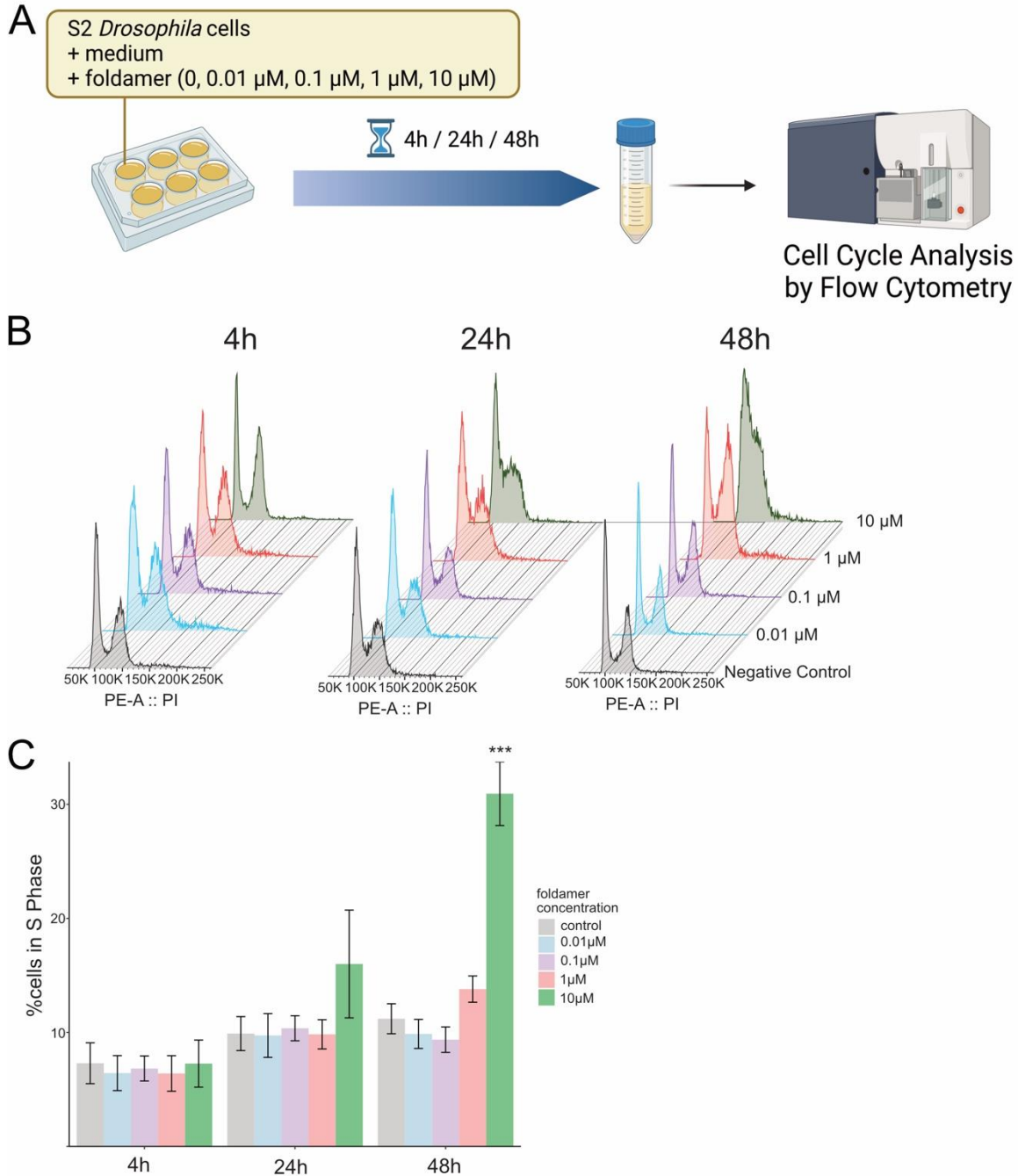


Figure 19 Effect of foldamer treatment on cell cycle progression

(A) Schematic diagram of the experimental flow of cell cycle analysis experiments. (B) Representative cell cycle profiles determined by flow cytometry with PI stain. Time points: 4 h, 24 h, and 48 h of treatment and Foldamer concentrations: 0 (control), 0.01 μ M, 0.1 μ M, 1 μ M, 10 μ M of foldamer 1 in medium, (C) Bar graph representing the percentage of cells in S-Phase after foldamer treatment with indicated concentration after the indicated time. Error bars represent standard deviation, N=3 replicates***= $p < 0.001$ against all other values in the group, 1way ANOVA + Tukey's test. Modified from Kleene et al (111) by permission of Oxford University Press. *Flow cytometry measurement, gating, quantification, and cell cycle profile visualization by Pardis Khosravani, Flow Cytometry Core Facility.*

To sum it up, in this project we were able to show that foldamers can interfere with the complex chromatin composition in vitro and in vivo in a specific, differential, and concentration-dependent

manner. The DNA mimic foldamer that mimics the shape, but not the sequence of double-stranded nucleic acids, allowed us to investigate the binding modes of proteins to chromatin and revealed potential for interference with basic processes resulting in physiological changes. Not only do we show the differential disturbance protein of binding to chromatin in a complex in vitro assay, but we also present a comprehensive foldamer interactome confirming the oligomer's affinity to double nucleic acid binders. Furthermore, we show that the foldamers' interference with specific proteins translates to in vivo experiments in *Drosophila* S2 cells. This beautifully resulted in cell cycle disturbance. Overall, the synthetic mimic's potential was thoroughly investigated and analyzed using a myriad of techniques, yielding promising results that allow for exciting conclusions in this proof-of-concept and future-directing study.

3.3 Discussion

This project allowed us to explore the impact of a synthetic mimic of a small biological molecule, foldamer, on the chromatin proteome. Studies on in vitro assembled chromatin were followed up by cell culture experiments to validate the data in vivo and thereby increase the scope of the findings. This provided novel clues towards the description and understanding of basic principles of chromatin binding of proteins and data on the physiological in vivo impact of foldamers as well as new indications for the potential to pharmaceutically leverage the synthetic mimic in the future.

3.3.1 The interactome supports and extends insights into foldamer-mimicking capabilities

Our identification of the foldamer interactome in DREX confirmed and expanded our knowledge of the foldamer as a double-stranded nucleic acid mimic. To investigate which proteins interact with the molecule we immobilized a biotinylated (covalent functionalization) foldamer on streptavidin beads and performed a pulldown from DREX, which was subsequently analyzed by LC-MS, revealing a diverse interactome of 640 proteins. Not all proteins were bound directly to foldamer but were also sometimes indirectly bound by protein-protein interaction to direct foldamer interactors, e.g. in complexes.

The foldamer was conceived as a DNA mimic. Indeed, we find many DNA-binding proteins and proteins involved in DNA processes in the foldamer interactome, attesting that many proteins recognize and bind the mimic as DNA. Interestingly, we detected more proteins as interactors than for a comparable interactome for a longer DNA molecule. Chromatin structure in eukaryotes has evolved to limit and fine-tune interaction between proteins and DNA (119). This finding attests to the hypothesis of the protective and regulating role of chromatin structure for the polyanionic DNA. Foldamers, putatively because of their very short length, cannot assemble into such structures, making them a more promiscuous binder. Ultimately, when synthesis allows, it would be interesting to investigate the ability of longer foldamer polymers to assemble chromatin-like structures. The 16bp length of the foldamer seems to have been sufficient to attract many transcription factors, which is per literature estimating the DNA sequence recognized and bound by transcription factors to mostly 5-20 base pairs (120,121). Overall, the function of foldamer as a DNA mimic was confirmed in a complex biological context.

Additionally, we also detect a variety of RNA binding proteins and proteins involved in RNA processes in the foldamer interactome. In our foldamer pulldown, we find a plethora of spliceosomal and ribosomal factors and other proteins with RNA binding domains. Due to the high prevalence of such molecules in the set, we hypothesize that foldamer, despite its conception as a DNA mimic, also resembles short stretches double-stranded RNA as present e.g. in mRNA stem-loops.

Interestingly, we find Zn-finger proteins among the foldamer interactors. Zn-finger proteins are known to bind DNA (and other nucleic acids) mostly via a sequence-specific interaction (122). Nevertheless, the foldamer does not have sequence-specific features that should be recognized by the motif. This undermines that even sequence-selective DNA binding proteins have a contribution to confer some of their binding affinity by recognition of and affinity to the framework the sequence

always comes packaged in the general structure comprised of the anionic double helix spaced in a certain range as mimicked by foldamer.

No quantifiable interaction studies were performed in this project, making it impossible to directly quantify the contributions of certain protein motifs or sites to binding or compare the binding strength or rates between different proteins. However, this is feasible in a follow-up study e.g. by using already available GFP tagged foldamer(s) and tagged proteins for FRET studies (123).

To sum it up, the foldamer interactome revealed that the foldamer does not only resemble DNA, as conceived but also interacts with RNA binding proteins. Since the foldamer is designed to expose the same electrostatic charge surface as DNA, a double helix with an anionic backbone, this finding is not entirely surprising. The molecule seems to be promiscuous enough to be recognized by binders of both nucleic acids, DNA, and RNA. Based on our data, we believe the foldamer is not only a DNA mimic but a double-stranded nucleic acid mimic.

3.3.2 Foldamer does not disturb basic processes like nucleosomal integration spacing during chromatin assembly in vitro

We showed that up to the ratio of 1:8 DNA: foldamer, the formation of regularly spaced nucleosomal arrays on DNA is not interfered with by foldamer in our DREX assembly as visualized on an agarose gel of the Mnase digested chromatin (Figure 9). Chromatin structure is not affected on this macroscopic level in vitro, meaning that foldamer does not impede general aspects of chromatin assembly including histone integration and nucleosome spacing. The observation that at higher foldamer concentration (1:8 ratio), the nucleosome spacing is fuzzier might reflect a change in protein composition on the fiber that we showed in the following more in-depth proteomic analysis (Figure 11).

3.3.3 A subset of proteins remains unaffected by foldamer in their binding to chromatin in vitro due to their binding mode

Proteins whose binding to chromatin during assembly was unaffected by foldamer are very stable chromatin binders that associate with the DNA directly or indirectly very strongly or multivalently. The influence of foldamer is not enough to disrupt either their initial binding during chromatin assembly, foldamer is added to DREX extract at the same time as DNA, nor to displace the proteins after they have bound.

Among the proteins in this cluster were the histones, known for their high affinity and multivalent affinity to DNA. Due to their large contact surface with DNA, it is not surprising that the very short foldamer is not able to compete with the much longer polymer, especially once binding is established and DNA occupies the possible binding areas.

We additionally find other DNA binders like Caf-1 unaffected. Caf-1's binding to chromatin is mediated indirectly by PCNA as well as directly through an alpha helix-like domain and a winged helix domain (124,125). From the foldamer interactome, we know Caf-1 does not bind foldamer. While PCNA did not crystallize out as a "chromatin binder" in this dataset, it has reliably been shown as a chromatin binder in other DREX chromatin assembly datasets (28,35). Hence, a stabilization of binding through PCNA is likely, a protein-protein interaction that would not be disturbed by the

“distraction” of one of the proteins’ binding sites by foldamer. The alpha-helix-like domain mediates sequence-specific DNA binding and should therefore not bind to foldamer. Interestingly, the winged helix domain, which is also present in Caf-1, is a general, sequence-independent nucleic acid binding domain known to also bind stem-loops in mRNA (126,127). Thereby this domain is theoretically susceptible to foldamer binding. Overall, Caf-1 does not seem to have binding sites that stably bind foldamer or if any of the sites binds it is too unstable to allow for a pulldown and its overall contribution to binding to chromatin is too small to change behavior when interrupted e.g. by foldamer.

The diversity and redundancy of binding sites might protect proteins from interference even if one of them is theoretically able to bind foldamer. This also introduces the difficulty of dissecting the influence of each factor in the multimeditated binding of proteins to DNA/chromatin and generally the challenge to disentangle the complexity of any biological system, even if simplified, from experimental observations. It does however underpin the value of observing proteins in a more complex environment to include contributions of bridging and stabilizing proteins into binding events.

Among the unaffected proteins, we also discover the whole MRN DNA repair complex and RPA complex. The presence of these repair proteins makes sense as DREX assembly works with a linearized plasmid, thus mimicking a double-strand break to the system. The recruitment of repair factors and DREX assembly as a great model system to investigate DNA repair mechanisms has been reported before (35). In addition to this finding, we observe the detection of and resistance to foldamer of all these complexes as a whole, including all proteins. This might point towards the protein-protein interaction, as expected, not being disturbed by foldamer and the “DNA binding” interaction site(s) of the complexes protected, strong/multivalent, or not attracted to foldamer. An especially interesting observation in this context is that the RPA complex, involved in all types of DNA repair and with an affinity specifically for single-stranded DNA, is not affected in its binding (128-133). It is intriguing to think that because foldamer mimics a double helix, not a singular strand, RPA would not bind it and thereby not be competed off DNA.

Overall, after considering the proteins in cluster 1, unaffected by foldamer in their DNA/chromatin binding, we find that 4 main phenomena protect proteins from being displaced from DNA: First, proteins associated with DNA by binding through protein-protein interaction e.g. in a complex to a DNA binder that is not displaced are protected. For higher foldamer, concentrations crowding effects that influence complex stability and formation are thinkable. Secondly, proteins that bind to DNA over a very large surface and multivalence, like histones, are not affected because the short foldamer cannot compete for the accumulated binding affinity. Additionally, it is likely that even if foldamer were to be produced in sufficient length to wrap a nucleosome it might be too stiff to do so and nestle in all the binding pockets. This remains to be tested once a foldamer of 146bp equivalent length is attainable. Thirdly, any DNA binders that bind DNA features that are not at all mimicked by the foldamer seem to be protected from its interference. While sequence specificity does not seem to be enough, affinity to single-strandedness seems to be an exclusive enough feature. Finally, we see that multiple diverse binding sites are used to stabilize the binding of certain proteins, like Caf-1, to chromatin. Only interfering with one or not all can be insufficient to disrupt the binding. To sum it up, we observe that certain proteins are not affected in their binding to chromatin, putatively because of stronger overall affinity to DNA than to foldamer in their DNA binding domains or because their binding is (partially) mediated or stabilized by foldamer-resistant protein-protein interactions.

3.3.4 Proteins are differentially disrupted in their binding to chromatin by foldamer in vitro, depending on their binding mode to DNA/chromatin

In our interference experiments, we investigated the effect of foldamer presence on the chromatin binding of proteins. We observed that foldamer differentially affects proteins where some remain unaffected (3.3.3) while others change their chromatin binding. Within those disrupted, we find 2 differential clusters of proteins; either mild (Cluster2a) or strong (Cluster 2b) displacement from the chromatin fiber in a foldamer-concentration-dependent manner compared to the control. Furthermore, we correlated the interference dataset with the foldamer interactome, allowing us to gain additional insights into the different modes in which protein binding is disrupted.

3.3.4.1 The strength of disruption in binding depends on proteins' chromatin binding mode

On the one hand, the proteins of cluster 2a are mildly affected by foldamer in a concentration-dependent manner, binding less to chromatin when more foldamer is present. These proteins include several factors involved in the structural maintenance of chromosomes and DNA repair proteins. For these factors, part of their association with chromatin is likely mediated by sequence-independent DNA-protein. However, the binding is also stabilized by another mediator that is not affected by the foldamer that structurally resembles a double-stranded nucleic acid but carries no sequence information, e.g. a secondary sequence independent DNA-protein binding site or a protein-protein binding-based anchor to chromatin.

Interestingly, among the mildly foldamer-sensitive proteins we find cohesion, which is indicated to have a binding preference for supercoiled DNA, compared to linear DNA (134,135). The intriguing question is whether the binding of cohesion to the immobilized DNA is simply weak and thereby easier to disrupt by foldamer or if foldamer mimics the structure of a double DNA with increased torsional stress. No data are available on this yet; a comparison of interactomes or binding assays with next-generation foldamers with adjustments to mimic the small but significant changes in the secondary structure of DNA upon supercoiling and relaxing might elucidate this question (136,137).

On the other hand, looking at cluster 2b, those proteins affected strongly in their binding to chromatin by foldamer, we identify DNA repair factors, the Sin3a complex, and the whole origin recognition complex (ORC). All proteins, or at least one subunit of each complex, in this cluster putatively rely strongly on their recognition of the general shape of double-stranded DNA to mediate binding to chromatin.

Additionally, the observation that the different topoisomerases have different sensitivities towards the foldamer further supports previous findings of some structural specificity of the foldamer. Purified Top1 was shown to be inhibited by foldamer in vitro (67,73). In our more complex study, we faced the limitation of many missing values for the protein and therefore an exclusion of analysis during filtering. However, looking at the remaining values, we do observe a confirming trend of Top1 being displaced from the chromatin fiber in a foldamer concentration-dependent manner also in our more complex setup. This serves as a valuable reminder that some proteins, especially when displacement from chromatin is highly efficient, might not have been reported here due to missing values for their

abundance on the chromatin fiber in the presence of high amounts of foldamer. Simplifying the dataset means accepting the tradeoff that some candidates or information might be missed. In the future, an additional analysis, identifying proteins with the exclusivity of data e.g. only in control might help include those candidates in the analysis (138).

3.3.4.2 Correlation of foldamer interactome and interference strength reveals insights into foldamer binding mechanisms

Finally, correlating the interference sensitivity and foldamer binding capacity of proteins allowed us to derive a conclusion of how proteins recognize and bind DNA and chromatin, or associate with them indirectly. Combining our data from the interactome and interference study, we identified 3 different modes in which foldamer can illicit disruption in binding:

First, we look at proteins whose binding is strongly interfered with by foldamer and that we identify as foldamer-interactors in our pull-down (Figure 15). This group contains proteins of the Orc complex as well as other proteins involved in DNA repair (117,118) and cell stress (116) response pathways. In this group, proteins are putatively disturbed directly in their interaction with the chromatin fiber by competitive binding to foldamer.

Second, we look at the subset of proteins that bind to foldamer but did not qualify as chromatin binders in our assay or are not disrupted in their binding to chromatin by foldamer. This phenomenon can appear because the protein was associating with foldamer in the pulldown as part of a complex, not binding foldamer directly but through protein-protein interaction to a foldamer binding protein, which might not associate with DNA but e.g. RNA. On the other hand, this difference might be due to the protection of DNA from unspecific interactions by chromatin structure that foldamer is lacking, thereby rendering it a more promiscuous binder. To sum it up, these proteins most likely bind to other nucleic-acid-binding proteins, are RNA binders or less specific DNA binders.

Third, we identified proteins that are interfered with in their binding to chromatin when foldamer is present but do not appear foldamer in our pulldowns. This group curiously has to be indirectly interfered with. Most likely, the inability to bind is due to an essential bridging interactor or activator not binding anymore, thereby interrupting recruitment to the fiber.

To sum it up, we discovered that foldamer disrupts the binding of proteins to chromatin in a concentration-dependent manner but in varying strength and manner, depending on how the protein associates with the chromatin fiber. Sequence-independent binding is competed with foldamer most efficiently and putatively directly by competition, as foldamer is a mimic of the general shape of DNA but without sequence information. However, even sequence-dependent binding is interfered with, suggesting that for sequence-specific binding, recognition and binding of the framework – the nucleic acid helix- contributes to the binding. Otherwise, foldamer would not be able to bind to and interfere with proteins that mediate their binding to DNA in such a manner. We also observe that some proteins that do not have any nucleic acid recognition sites can be affected indirectly. Ultimately, these interference and binding assays with foldamer emphasize the importance of studying the effects of foldamer not only with isolated proteins but also in the context of a complex chromatin environment. This allows one to be able to also include physiologically highly relevant processes like indirect effects including inhibitor/activator displacement or loss of binding because a whole complex is displaced.

3.3.5 S2 cells allow for functional in vivo experiments with foldamer

This project offers the first results of the effect of foldamer on cells without the need to use transfection agents or other transfection methods like electroporation. Due to their polyanionic nature, foldamers had before been characterized to have difficulties readily entering human cells (67). Using *Drosophila* S2 cells instead of human cells, we were able to overcome this limitation and observe the foldamer's effects as it was absorbed by the cells even without the addition of a carrier. We deduce the uptake from the visual observation of the characteristic yellow coloring of foldamer in all fractions during subcellular fractionation. This uptake was anticipated due to the known characteristic of *Drosophila* S2 cells to allow the uptake of polyanions (139).

For in vivo experiments, higher amounts of foldamers in the milligram range were necessary thereby putting a high production load on the chemists producing the foldamer. However, advancements in the different fields' steps and methods involved in these studies might allow for streamlining of the experiments. Firstly, the establishment of a solid phase synthesis of foldamers now allows for an easier upscaling of the reaction (140). Secondly, *Drosophila* cell culture as well as subcellular fractionation protocols could be implemented, optimized, and adapted to pipetting platforms, thereby allowing smaller volumes. The main challenges here would be separating pellet and supernatant cleanly and reliably in small volumes in both protocols. Thirdly, while the required amount for cell cycle analysis via Flow Cytometry is somewhat set (141), constant improvements low volume sample preparation as well as in sensitivity in mass spectrometry already allow for a significant reduction in input with the development towards even lower input requirements (142).

Overall, this study is the first report of an assay to study the effect of foldamer on whole cells while eliminating the confounding feature of transfection. The results obtained here can be viewed as exclusively caused by foldamer and thereby form an exceptional basis for mechanistic elucidation

3.3.6 In vivo study reveals transferability of in vitro results, validating foldamer-sensitivity-based protein clusters

To validate the transfer of the in vitro findings to an in vivo setting, we treated *Drosophila* S2 cells with foldamer for 48h to subsequently fractionate the cells and investigate the impact of protein distribution overall but also the changes in the chromatin-bound proteome upon treatment specifically.

Looking at all fractions, we observe that the overall protein intensity distribution remains generally the same (Figure 17), allowing us to proceed with a more in-depth analysis of the fractions of interest. We observed an overall decrease in protein intensity in the chromatin-bound fraction; this is per the in vitro data. The fractionation showed not only a validation of the clusters of proteins according to their response to foldamer in vitro but also that the proteins are similarly affected in cells.

Furthermore, our results in vivo suggest an aggregation of foldamer with certain factors. In our subcellular fractionation (Figure 17), we observe an overall downshift for all fractions except for the cytoplasmic and final pellet fraction. The increase of protein in the cytoplasm points towards an effect of foldamer on protein homeostasis in the cytoplasm but was not investigated further here. The accumulation of proteins in the final pellet points towards aggregation of proteins. This also fits with our practical observation of increased turbidity noticeable by the eye in any protein-containing solution when higher amounts of protein were added. We hypothesize that some proteins become more

aggregation-prone in the presence of foldamer. With its double helical structure, it might mimic double-stranded stretches (hairpins) of RNA, implicated in phase separation and aggregate formation (143-146). To avoid unspecific crowding issues in this project, only foldamer concentrations were used which resulted in a linear effect on the observed proteins.

Overall, the fractionation of foldamer-treated cells shows that the observations made in vitro also hold up in vivo, confirming the DREX assay as an excellent fusion between simplicity of handling and biological complexity generating biologically relevant insights that provide good predictions for in vivo outcomes.

3.3.7 In vivo cell cycle experiments reveal the physiological impact of foldamer

We wanted to not only show the molecular effect of foldamer in cells but also investigate its physiological relevance and impact. Therefore, we performed cell cycle stage analysis via flow cytometry with cells after 4h, 24h, and 48h of treatment with different foldamer concentrations. Here, we report a concentration- and time-dependent effect of foldamer on cell cycle progression.

This interference with the cell cycle is putatively caused by direct interference with proteins binding to chromatin. Nevertheless, other pathways in which foldamer could affect cell functions, like immune response (147) or disturbance of compartmentalization (148), that have so far not been investigated might also contribute to the effect observed. However, we not only observed a change in chromatin binding for ORC but also showed interference with its function of setting up chromatin structure in collaboration with the Kurat lab (111). S-Phase arrest to a similar extent to what we report has been observed before when this ORC function is impaired (23), hinting that the physiological effect observed indeed is caused by the mechanism of ORC depletion from chromatin and thereby impairment of function.

We observe the accumulation of cells in the S-Phase over time, most significantly for high foldamer concentrations. Since foldamer is protease- and nuclease-resistant, the clearance of the small mimic is putatively not possible for the cells. To establish foldamer as a clinical agent, further investigation and optimization would be necessary. This includes the study of in-cell off-target effects, potentially with interactomes of foldamer from isolated, different fractions or cellular localization studies (3.3.9). Additionally, better targeting and higher specificity for interference with specific proteins realized by chemical adjustments to foldamer structure are thinkable (3.3.9). Overall, the experiments performed here give a first, valuable insight into what is physiologically possible should these limitations be overcome.

3.3.8 The difference between foldamers and DNA in interference assays

In our experiments, we find that DNA of a similar length is not able to cause interference with chromatin structure and function in the DREX chromatin assembly assay as the foldamer does (Figure 12). This can be explained in 2 ways: Firstly, the Huc group has shown that for some applications and contexts, foldamer can bind stronger to certain proteins than DNA itself (149). Secondly, the foldamer is resistant to nucleases and proteases that in a complex proteomic environment, like an extract or even a cell, would readily degrade DNA. This also results in great pharmacological potential by allowing

the orthogonal compound to persevere and intervene in cells when similar DNA-based therapies cannot overcome internal control mechanisms like DNA degradation.

3.3.9 Future perspectives and recommendations for foldamer design and modifications

The beauty of synthetic mimics lies not only in their stability but also in the many ways they can be manipulated, changed, and adapted in the future. It has many designable modules that will allow future studies to investigate the impact of certain features in detail. This includes but is not limited to the length of the foldamer, functionalization, and adjustment of shape features by chemical design.

Firstly, in our project, we tested 16mer and 32mer foldamers, corresponding to 8bp/16bp length DNA double helices. When comparing the sets of proteins compromising the interactome of the foldamers of different lengths, we found that the overarching GO terms were the same. When we compared the protein intensities in their pulldown, we discovered the same proteins. However, the abundance of said proteins was higher for the 32mer, probably due to higher avidity (Figure 20). Overall, we yielded an extensive interactome even for the short foldamers we tested, showing that even short stretches of DNA shape were able to bind a significant number of proteins. In the future, even longer foldamers will be available due to the recent development of foldamer solid phase synthesis (140), promising exciting results. We propose testing longer foldamers, when available, in similar setups to test their similarities and differences in their interactomes but also their differential ability for interference with the complex chromatin proteome. This comparison of foldamers with 2 different lengths represents a proof of concept set up for the comparison of different versions of the foldamers in complex environments in the future.

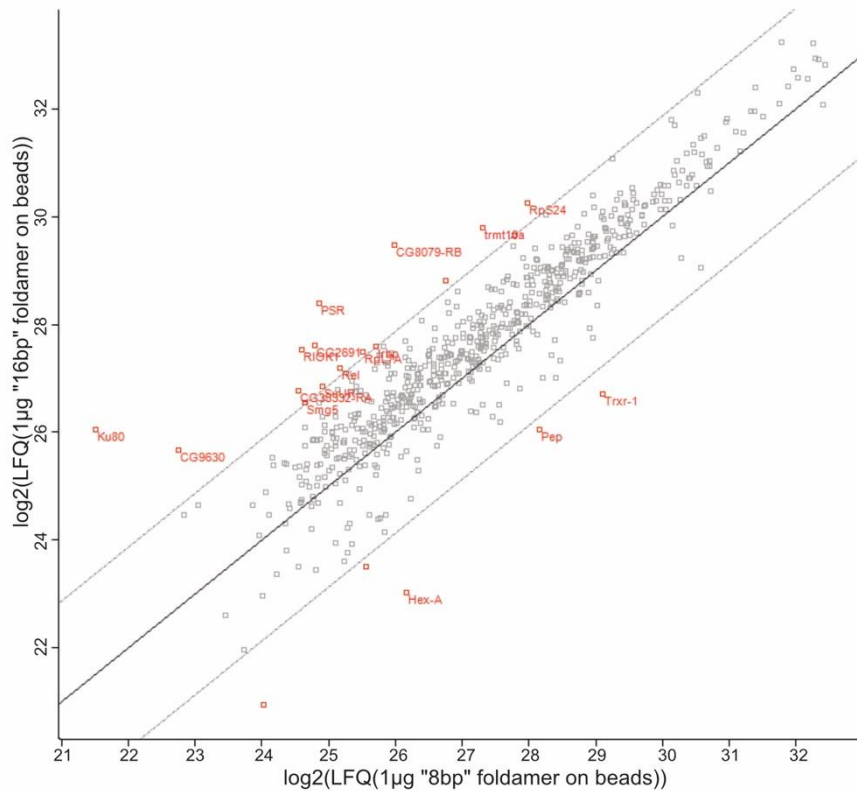


Figure 20 Scatterplot of proteins in pulldown with “8 bp” or “16 bp” foldamers.

Mean intensities of proteins in Pulldown from DREX with 1µg biotinylated foldamer of “8 base pair” or “16 base pair” (“16 bp” in all other experiments of this study) length. Solid line= $y=x$, Dash-dotted lines= $y=x+sd$, and $y=x-sd$. $N=3$. Data based on Dataset 4.

Secondly, the functionalization of the foldamer offers exciting perspectives for interaction studies, detection, and visualization but also for future clinical applications. We used a biotinylated form of the foldamer to perform pulldown experiments using streptavidin-coated magnetic beads. This biotin functionalization would also allow for imaging and detection of foldamer using streptavidin antibodies in cells, however, with limitations due strong biotin background naturally occurring in cells (150). The Huc group has already managed to attach a GFP tag to the foldamer. Nevertheless, due to its disproportional size in comparison to the foldamer, a changed behavior of the molecule has to be anticipated regarding cellular uptake, localization, and interaction within cells. Therefore, other functionalizations for visualization in the cell might be desirable. We propose functionalizations that are smaller like the orthogonal SNAP and CLIP tags (151,152). Additionally, bioconjugation of foldamer to clickable nanobodies (153) or a Protein A tag, detectable by nanobodies and thereby allowing for proximity biotinylation (reference Rupam paper) are thinkable. Furthermore, a direct functionalization of the foldamer with a clickable azide rest would have even less risk of steric hindrance and can be activated the visualization or other purposes in the cell (154-156). These alternatives should allow for easier delivery compared to GFP-tagged and less background than for a biotin-tagged foldamer.

Finally, leveraging the polymer structure of foldamers, small adjustments in the building blocks allow for changes in the overall shape of the resulting double helix. Comparison of interactome as well as

interference potential of slightly altered foldamers will shed light on the differential impact of those features. The options include changing major and minor groove sizes as well as mimicking DNA methylation (157,158). Additionally, long term, the mimicking of DNA bending, kinks, or other DNA shape features might be aimed for (62,64,159).

Clinically, foldamers might long-term be interesting as cell cycle interference agents as demonstrated in this study, potentially reaching even higher specificity due to further aforementioned adjustments. While aptamers are already tested for various clinical treatments including cancer treatment, the use of foldamers as an orthogonal, chemically, and metabolically more stable alternative is conceivable (160,161). A major challenge here will be the targeting and tissue penetration while mitigating in vivo off-target effects and potential safety issues, proving safe especially also in immunogenicity which has not been investigated in this study. To faithfully target foldamers in the future, the delivery by a new promising therapeutic system for nanomedicine, extracellular vesicles, could be explored (162-167). Leveraging this intrinsic form of cell-cell communication transfer of biomolecules as a tool, the first advances to target specific cancer cells for delivery of DNA for gene therapy have been shown (168). Loading of the vesicles with foldamer could be explored in the future to overcome delivery but also targeting challenges (169).

To sum it up, future advancements in foldamer design should take into account the parameters and biological features that should be mimicked and tested or could aim to generate molecules with distinct differences from their molecular twin to leverage orthogonal biology for interference with biological processes. We propose targeting to generate longer foldamers to explore their binding affinities as well as elucidating their structures in combination with different binding partners. Additionally, we recommend leveraging the option to functionalize the foldamer for biochemical and in cellulo assays and to alter features like groove size to elucidate the influence on binding behavior. Ultimately, we believe the synthetic mimic foldamer has a lot of exciting, yet-to-be-discovered and exploited potential and the newly established solid phase synthesis (140) will make rapid advancements possible.

3.3.10 ORC and the foldamer, a proof-of-concept: Interaction, interference, and functional disablement leading to physiological consequences

This study presents a plethora of information that can be exploited to gain new indications and impulses regarding the elucidation of fundamental binding events and mechanisms. We have led a proof-of-concept by investigating ORC binding and interference thereof, expanding the study to in vivo experiments to highlight transferability but also display the physiological implications of the knowledge gained. By interfering with ORC binding we were able to not only confirm that *Drosophila* ORC binds at least proportionally sequence independently, but also that this binding is essential for cell cycle progression and can be interfered with by foldamer.

In eukaryotic cell division, ORCs canonical function is binding replication origins and subsequently loading the replication helicase (Cdt1-MCM) during G1-Phase (170). In yeast, ORC is known to strictly recognize and bind an ARS consensus sequence (ACS) to recognize replication origins. Metazoan origin recognition is also guided by ORC, however, despite high structural conservation between the protein complexes, metazoan ORC seems to recognize origins not by sequence but in a not yet fully elucidated mechanism involving histone modifications and variants as well as chromatin structure and

accessory proteins (171-176). While in yeast ORC, stays bound throughout the cell cycle, human ORC assembles stepwise in the S-Phase with first Orc1 and then Orc 2-5 binding (177-179). Disassembly of human ORC then happens in S-Phase by ubiquitination and subsequent degradation of Orc1 and detachment of ORC2-5 by phosphorylation, preserving this part of the complex in an auto-inhibited state until the next G1 Phase (180,181).

In our in vitro experiments, we observed that *Drosophila* ORC (dsOrC) can bind directly to foldamer (Figure 15). This can be explained by the ORC subunits 1-5 each bearing an α -helix winged helix domain. Winged helix domains can mediate binding by having one of the helices make direct sequence-dependent interactions while the wings mediate general binding to the backbone (182,183). Interestingly, find Orc1-5 but not Orc 6 as foldamer interactors and proteins suffering interference from foldamer in their chromatin binding. This can have two causes: First: Since only those subunits (1-5) inherently bind DNA, ORCs foldamer-binding subunits could each bind individually, not as a complex. This is less likely because, for other complexes, we see non-DNA binding subunits pulled down as a complex. More likely, the *Drosophila* ORC is built like a human ORC. There, ORC2-5 forms a stable subcomplex while ORC1 and ORC6, essential for helicase activation, are more loosely attached. This assumption is feasible by looking at the dsORC structure (184).

We believe that the direct interaction of foldamer makes the molecule a competitive binder at DNA binding sites of ORC subunits. It thereby causes interference with ORC's binding to chromatin in our in vitro assays (Figure 12). Additionally, an in vitro functional assay by the Kurat lab revealed that ORCs function to organize nucleosomes at replication origins in yeast was impaired in the presence of foldamer (111).

Finally, we confirmed in S2 *Drosophila* cells that ORC association with chromatin is affected in the same manner in vivo as in our in vitro assays (Figure 17). Furthermore, we were able to observe a defect in the cell cycle progression of cells treated with 10 μ M foldamer for 48h. We see an accumulation of cells in the population in the S-Phase, which we suggest is caused by the premature dissociation of ORC from the chromatin fiber in S-Phase when foldamer is present.

3.3.11 DREX assembly confirmed as an exploratory assay to test small molecules in a complex chromatin context

Our study provides a proof of concept for investigating the impact of synthetic mimic molecules on chromatin assembly. Our findings underline the importance of studying the impact of these molecules in complex environments instead of with purified proteins only. This is the first specific yet complex investigation of foldamer interaction with proteins. Before, only studies on isolated DNA-binding proteins had been performed (67,73). Like in the cell, the extracts provide a highly dynamic and physiological environment in which a plethora of different proteins compete for interaction. Additionally, the extract allows us to circumvent issues of delivery encountered in many cell-based experiments. Finally, our study confirmed that the findings were transferrable to in vivo systems. We, therefore, propose the DREX-assisted chromatin assembly for the initial study of small molecules' potential on chromatin assembly.

4 Project 2: Chromatin assembly and metabolism

4.1 Graphical abstract

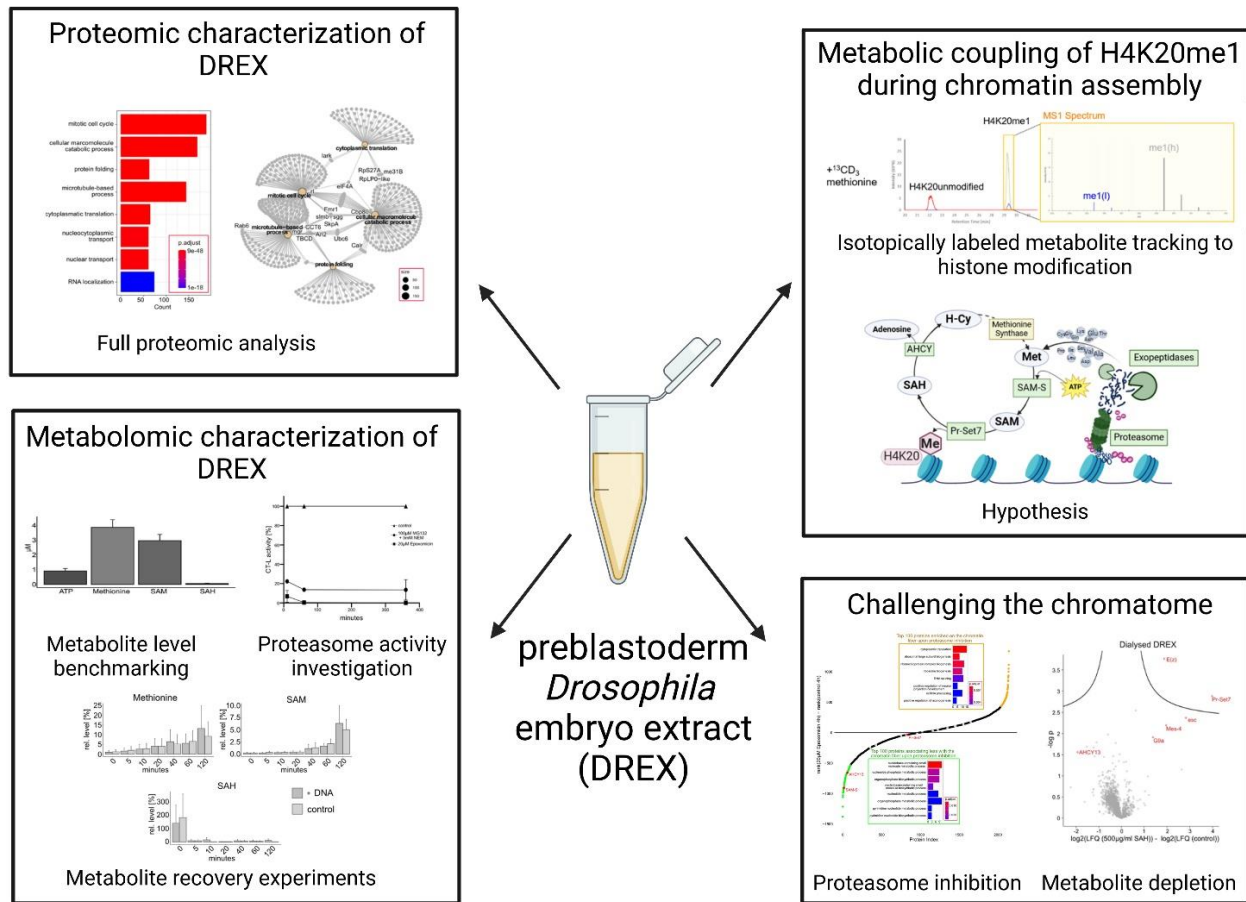


Figure 21 Graphical abstract of project 2

Graphical abstract depicting an overview of project 2, created with Biorender.com.

While the DREX extract is a well-established system for studying chromatin assembly and composition, it has not yet been metabolically characterized. Here, we use mass spectrometry to analyze the metabolome of the extract, test the extract's metabolic activity, and its response to challenges to the system (Figure 21). The metabolomic characterization of the DREX extract reveals a great diversity of metabolites, an activity of metabolic pathways, including the methionine cycle and proteasomal degradation, and possibilities to interfere with the processes. Finally, the metabolic coupling of H4K20me1 during chromatin assembly in the extract is shown.

4.2 Results

To understand, characterize, and explore the metabolic capabilities, flexibilities, and limitations of the DREX (*Drosophila* embryo extract), we employed a wide range of methods and approaches and applied the findings to gather knowledge to address different biological questions. We performed mass spectrometry to measure whole-extract proteomes and whole-extract metabolomes. Additionally, we created proteomes reflecting the abundance of proteins enriched on an assembled chromatin fiber. Furthermore, we performed regular as well as isotopically labeled histone modification analysis of said fibers. The easy handling and manipulation allowed us to introduce different challenges to the established assembly system while getting a clean readout through our benchmarked methods and controls.

4.2.1 The DREX proteome

DREX is rich in protein whereas DNA, lipids, and insoluble molecules are separated from the extract during its preparation (24). Through Western Blot and mass spectrometry analysis of proteins bound to DREX-assisted in vitro assembled chromatin, the presence of certain proteins in the extract had already been shown, including transcription factors, histones, and histone chaperones (5,28,35). Due to its origin from embryos, the presence of other proteins in DREX had been hypothesized, assumed, or observed as contaminations in other experiments. Yet, an untargeted and in-depth analysis of the whole spectrum of proteins in DREX without any manipulations or artificial enrichments had been lacking. To get a better insight into which proteins are present in the DREX and how abundant they are compared to each other, a proteome of DREX was prepared and analyzed via LC-MS.

In a total proteome analysis of 4 separately prepared DREX, measured in 4 technical replicates, 1650 proteins with varying intensities were detected and identified (Figure 22). The most abundant protein group were the Yolk proteins: Vitellogenin -1 Vitellogenin -2, and Vitellogenin-3.

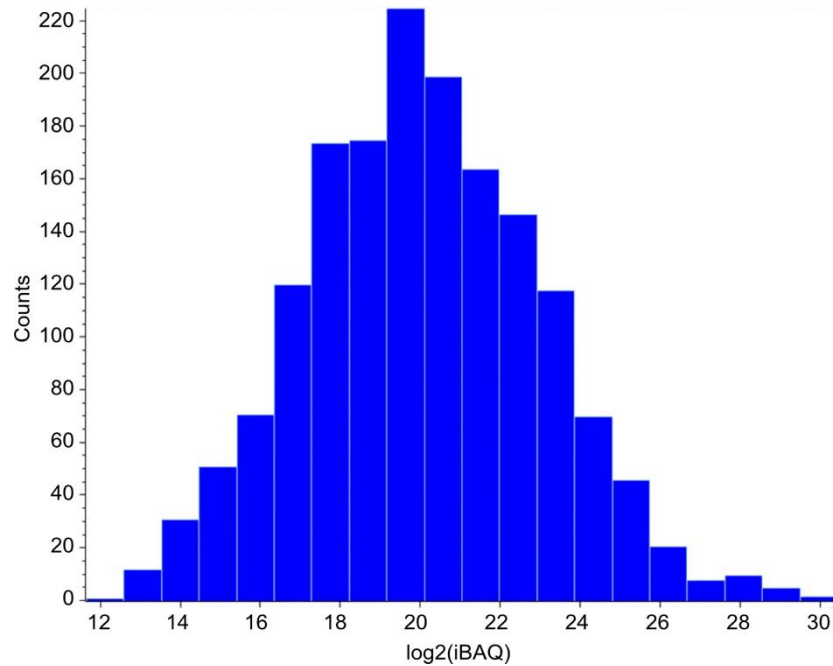


Figure 22 Distribution of iBAQ intensities of proteins identified in the DREX extract

Barplot depicting distribution of $\log_2(\text{iBAQ})$ intensities of proteins in DREX extract. Data based on Dataset 1.

Network analysis of all proteins in the DREX proteome shows a GO term enrichment for Biological processes revolving around catabolic processes, protein folding, and mitotic cell cycle over a whole genome background. (Figure 23). For the visualization, the 8 most significant (“top”) GO terms were identified by GO Term analysis and plotted using the R package ‘ClusterProfiler’ and the ‘org.Dm.eg.db’ database. The identified terms were filtered for semantic redundancy to allow for a broader overview. The analysis shows high gene counts for the GO terms detected (Figure 23 A). A network plot of the top 5 GO terms visualizes the frequent overlaps between the different processes (Figure 23 B). The biological processes with the best match were the “mitotic cell cycle”, “cellular macromolecule catabolic process”, “protein folding”, “microtubule-based process” and “cytoplasmic transport”.

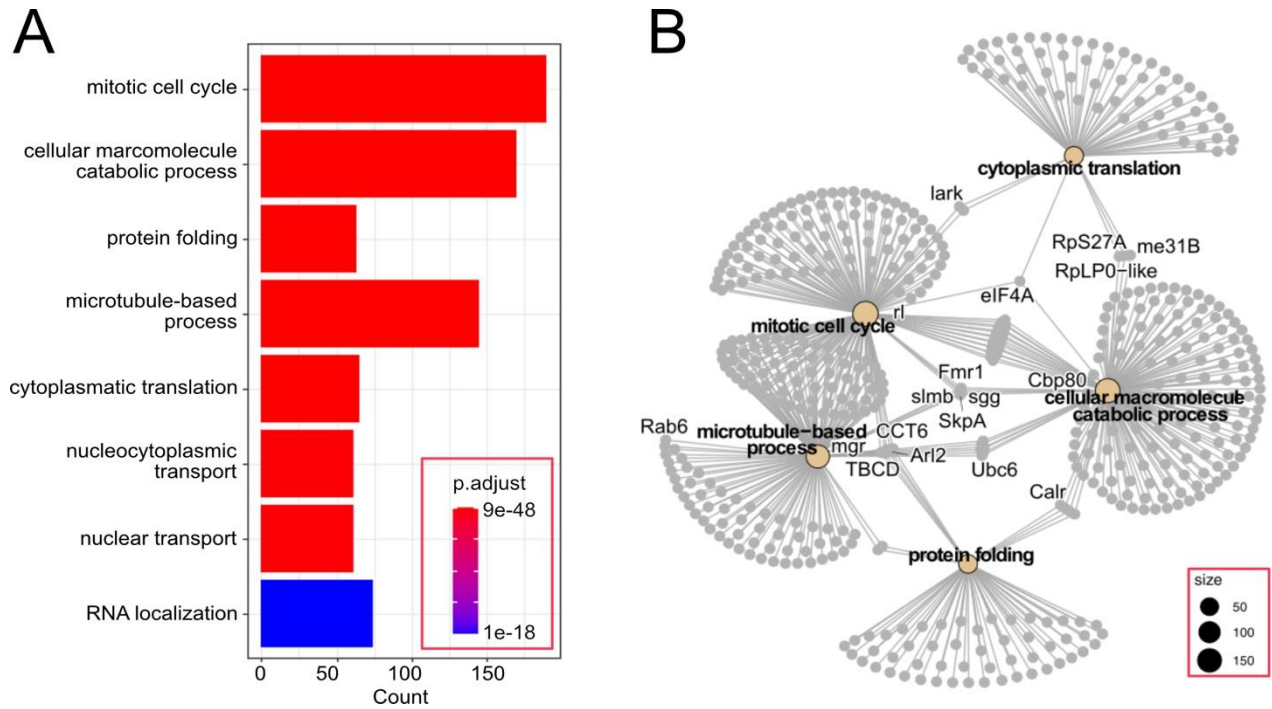


Figure 23 Complete *Drosophila* embryo extract (DREX) proteomic analysis

(A) Top 8 GO Terms by p-value for all proteins detected in DREX. GO terms were filtered for semantic redundancy. (B) Cytoscape network plot for the top 5 GO terms for all proteins detected in DREX. GO terms were filtered for semantic redundancy. Beige nodes with fat letter labels represent GO terms, size of these nodes represents the Count of proteins detected for this cluster. grey nodes represent proteins, and connecting proteins are labeled in grey. Data based on Dataset 1. Visualization with adapted R script from Anuroop Venkatasubramani.

Additionally, a “biological process” GO term enrichment analysis of only the 100 most abundant proteins in DREX against a *Drosophila* whole genome background was performed (Figure 24 A). In this TOP100 analysis, two main networks emerged (below, Figure 24 B). On the one hand, a network of proteins that are involved in “protein folding”, “protein maturation” and “chaperone-mediated protein folding”. This subset was characterized by diverse chaperones, including T-complex protein Ring Complex (TriC) subunits and an array of heat shock proteins. On the other hand, a network surrounding the biological processes of modification-dependent protein and macromolecule catabolic processes. Proteins involved in catabolic processes include the subunits of the proteasome as well as proteins involved in the ubiquitination pathway. Overall, the analysis of the 100 most abundant proteins revealed that proteins involved in protein homeostasis were enriched in this subset.

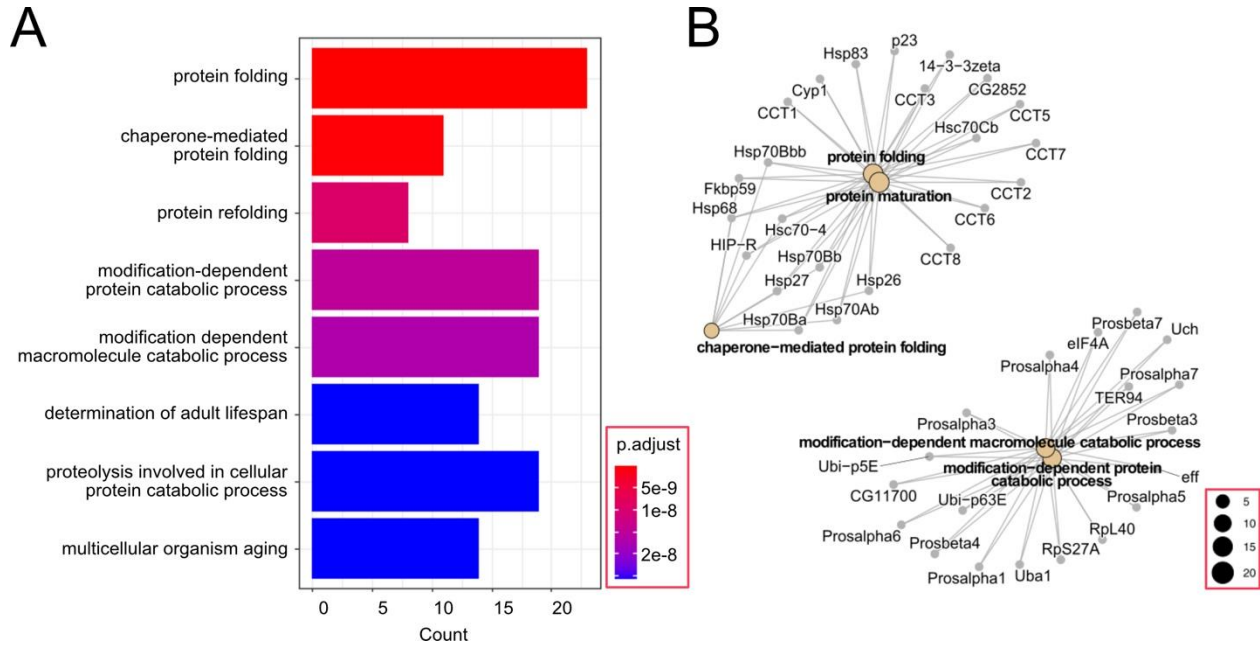


Figure 24 Proteomic analysis of top 100 most abundant proteins in *Drosophila* embryo extract (DREX)

(A) Top 8 GO Terms by p-value for the 100 most abundant proteins in DREX (B) Cytoscape network plot for the top 5 GO terms for the 100 most abundant proteins detected in DREX. Beige nodes with fat letter labels represent GO terms, size of these nodes represents the Count of proteins detected for this cluster. grey nodes represent proteins, and connecting proteins are labeled in grey. Data based on Dataset 1. *Visualization with adapted R script from Anuroop Venkatasubramani.*

4.2.2 Investigation of the DREX fundamental requirements for chromatin assembly capacity

We explored which characteristics determine whether a batch of DREX can assemble chromatin. First, we tested different batches for their ability to assemble chromatin on a plasmid containing an inset of a 5S sea urchin nucleosome positioning sequence. 3 DREX batches assembled chromatin successfully using the standard assembly protocol using 40µl-80µl DREX per 1µ DNA and 1 batch did not assemble chromatin in the setup given. The success of assembly was determined using Mnase digest and Agarose gel electrophoresis, yielding a ladder in case of successful assembly (not shown). To explore what is needed for a DREX to be able to assemble chromatin, we compared the protein concentration and composition of the DREX extract batches tested.

The total protein concentrations of the DREX batches were determined by BCA assay against a BSA standard curve (Table 1). Due to the complexity of the sample, determination of protein concentration by other methods like DeNovix detection or via Bradford assay yielded discordant concentrations (data not shown). We therefore only compared differences in concentration between the batches using one technique to circumvent technical error by measurement assay. However, the limitation of accounting for the complexity of the sample compared to the BSA standard remains and the concentrations should therefore be seen as approximates. Concentrations ranged from 48µg/µl to 95µg/µl We observed that the higher the determined protein concentration of a DREX batch, the lower the volume of DREX necessary to assemble 1µg of DNA into chromatin (Table 1). The lowest concentrated DREX was not able to assemble chromatin, as determined by the Mnase digest and

Agarose gel. Overall, we observe that total protein concentration had a crucial impact on the chromatin assembly capacity of DREX.

Table 1 Representative DREX batches and their concentrations

DREX batch name	Concentration determined (<i>measured</i> <i>by</i> <i>Gromadskiy</i>)	[$\mu\text{g}/\mu\text{l}$] by BCA <i>Mikhail</i>	μl of DREX needed for chromatin assembly with 1 μg DNA
Quadrat	95		60
Triangle	57		80
Star	54		80
Dot	48		Does not work (tested with 40,60,80)

Additionally, we compared the proteomes of the 3 assembling DREX with that of the non-assembling DREX (Figure 25). Here, we plot the difference of protein abundance in z-scored iBAQ of each protein between assembling and non-assembling DREX against the $-\log(\text{p-value})$. Interestingly, only 9 (of 1650) proteins appear significantly enriched over the threshold of FDR 0.05. Only one protein is slightly enriched in the non-assembling DREX: CG9125/dRai1, an automatically annotated protein by the Uniprot/Trembl database that is a putative decapping nuclease for NAS-capped RNAs thereby playing a role as a mRNA regulator (185). The 8 proteins that are enriched in the assembling DREX compared to the non-assembling batch are: 3 Translation Initiation factors (eIF-3p40, eIF3-S10, eIF3-S8), 3 ribosomal proteins (RpS20, RpS3, RpS6), Fmr1; an RNA binding protein associated with fragile X syndrome in mammals (186), and Krishah (kri); a uracil phosphoribosyltransferase (187). None of the proteins significantly enriched showed any functional relation to chromatin assembly or structure.

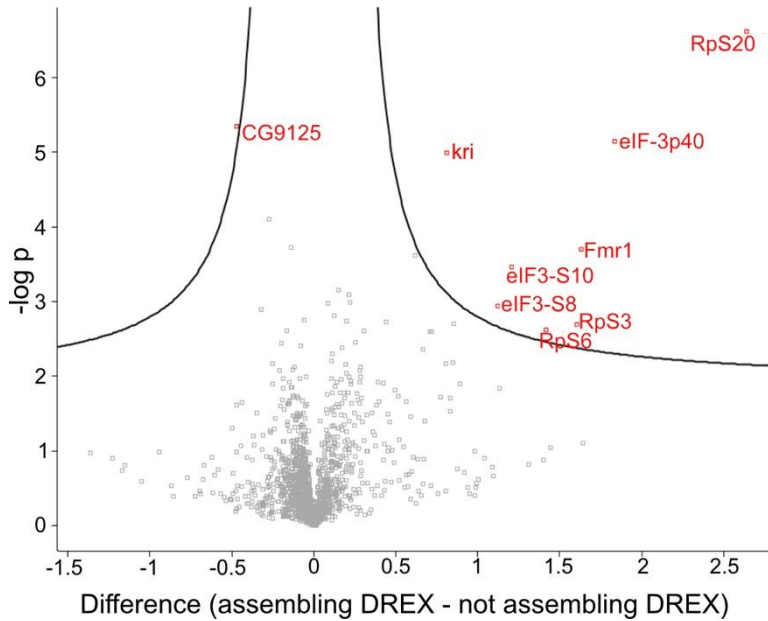


Figure 25 Difference in proteome for DREX with and without chromatin assembling capacity
Volcano plot for proteins detected in DREX that can assemble chromatin on naked DNA (assembling DREX) against DREX that has no capacity for assembly (not assembling DREX), based on z-scored abundances based on iBAQ. N=3 (assembling) +1(not assembling), FDR=0.05. Data based on Dataset 1.

4.2.3 The activity of the proteasome in *Drosophila* embryo extract

The DREX proteome and a proteome on the assembled chromatin fiber indicate the presence of proteins of the catabolic processes in DREX in high abundance, notably including almost all subunits of the canonical proteasome and its regulatory particles in *Drosophila* (Figure 26). Therefore, we wanted to investigate whether the proteasome subunits are not only present but functionally assembled to an active proteasome in the DREX, which then, like in the embryo can break down proteins and peptides to supply amino acids for metabolic activity.

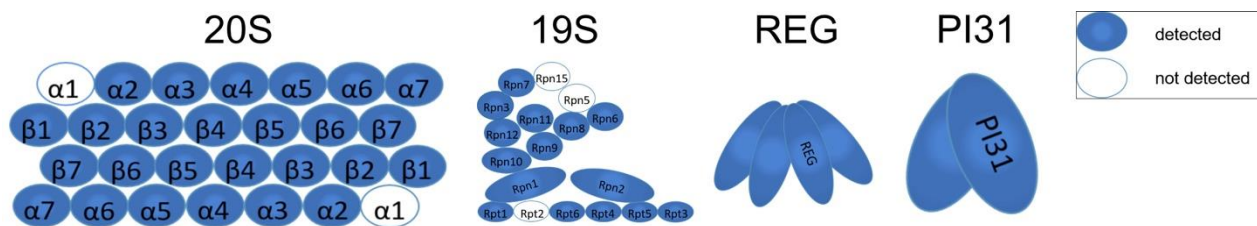


Figure 26 Proteasomal subunits detected on DREX-assembled chromatin

Schematic representation of detection of all subunits of the 20S proteasome core as well as the regulatory particles 19S, REG, and PI31. Filled symbols indicate the detection of proteins on assembling chromatin fiber in DREX; empty symbols depict not detected proteins. N=3. Data based on Dataset 2.

A luciferase-based activity assay (Figure 27 A) shows that the proteasome is active in the DREX extract (with the supply of an ATP regenerating system). CT-L activity was monitored as a proxy for total proteasome activity. Therefore, a Suc-LLVY-Aminoluciferase was added to the sample. This peptide was specifically degraded by an active proteasome, releasing Aminoluciferin. The Luciferase then

catalyzes the oxidation of the Aminoluciferin substrate, a reaction which emits a proportional emission of a photon light to quantify the amount of substrate processed and thereby the degree of proteasome activity.

We tested and benchmarked the inhibition of the proteasome with different inhibitors in the DREX extract (Figure 27 B). Again, CT-L activity was determined as a proxy for proteasomal activity and inhibition. The activity of the proteasome in DREX was successfully reduced within 30 min and was stably inhibited using 20 μ M Epoxomicin as well as a combination of the inhibitors 100 μ M MG132 and 5 mM NEM. The inhibitor combination Epx+NEM still exhibited leaky activity of about 15%. Meanwhile, Epoxomicin led to full inhibition and allowed for reduced solvent use, avoiding effects on protein solubility caused by the solvent ethanol that was used for the MG132+NEM inhibition experiments. Additionally, MG132+NEM has off-target effects because NEM is a broad-spectrum cysteine protease inhibitor (188,189) and MG132 also inhibits N-kB (190). Meanwhile, Epoxomicin is a potent selective proteasomal inhibitor (191-193). Due to its quick, complete, and specific inhibitory effect, 20 μ M Epoxomicin was used for inhibition of proteasomal activity in all following experiments.

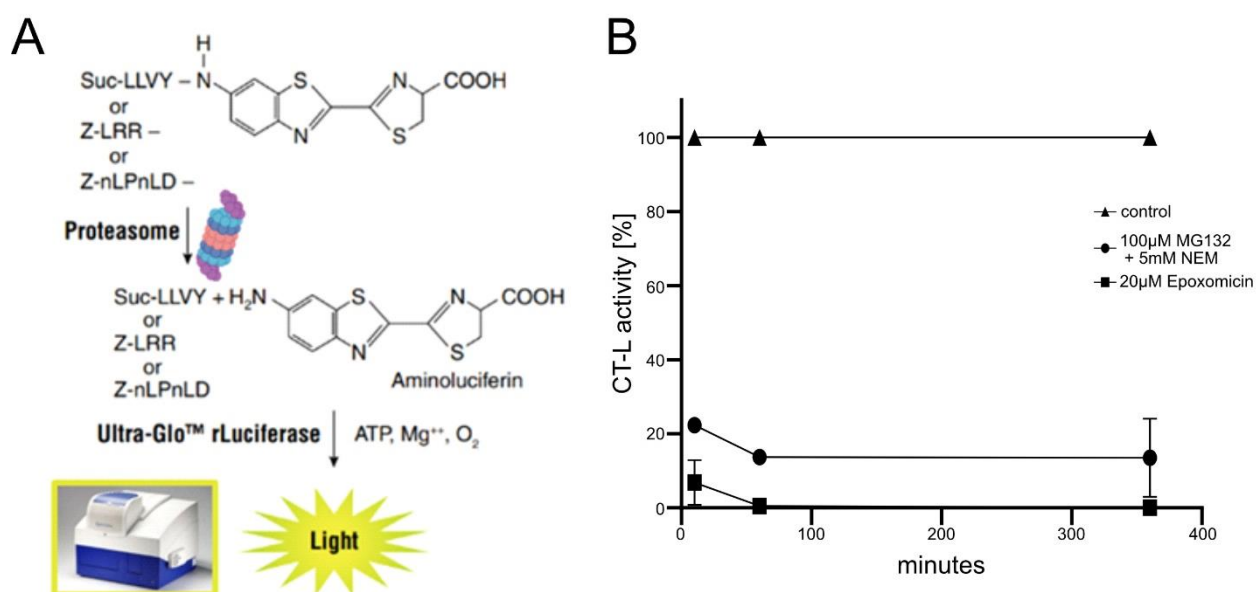


Figure 27 Inhibition of proteasomal activity in DREX

The activity was determined by testing CT-L activity using the Proteasome-Glow™ Assay. (A) Schematic depiction of a chemical reaction during Luciferase assay and photo of plate reader used to detect luminescent output. (B) CT-L activity of the proteasome in the DREX over time upon the addition of different proteasome inhibition agents. Figure A modified from Promega promotional content. *Experiments in collaboration with the Meiners lab.*

4.2.4 Inhibition of proteasome does not affect nucleosome integration and spacing

Having established the proteasomal inhibition in DREX, we wanted to investigate the impact of proteasomal inhibition on chromatin assembly. We tested chromatin assembly with our standard protocol, using a DREX, a nucleosome positioning repeat sequence DNA, and an ATP regenerating system. Epoxomicin was added to DREX 30 min before the addition of DNA for assembly reaction to ensure full inhibition of the proteasome throughout the process. After assembly, the DNA was digested using Mnase for different time intervals and then visualized on an agarose gel (Figure 28).

Comparing a standard chromatin assembly control to a chromatin assembly in DREX treated with 20 μ M Epoxomicin, we observe Mnase ladder formation in both conditions. This shows that proteasome inhibition does not interfere with general aspects of chromatin assembly like histone binding and nucleosomal spacing.

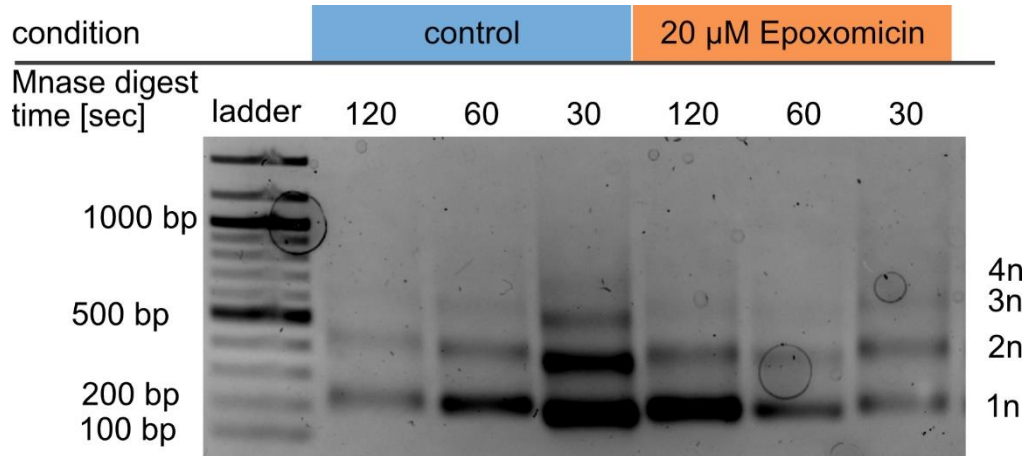


Figure 28 In vitro chromatin assemblies in the presence of proteasome inhibitor epoxomicin

Agarose gel of Mnase digested chromatin assembled in DREX under standard conditions (control) or after treatment with a proteasome inhibitor (20 μ M Epoxomicin). Labeling reflects a nucleosomal ladder containing a regular array of nucleosomes (1n/2n/3n/4n).

4.2.5 Proteasome inhibition causes the accumulation of proteins on chromatin over time

While proteasomal inhibition did not challenge chromatin assembly on a macroscopic level, we set out to investigate the roles of the proteasome in chromatin assembly on a more differentiated proteomic level. Again, we leveraged our knowledge of the presence and activity of the proteasome in DREX as well as our benchmarking and establishment of its fast and complete inhibition.

Therefore, we inhibited the proteasomal activity using Epoxomicin (Epx) and analyzed the proteome on the chromatin fiber during the early (after 15 min) and late/mature (after 4 hours) stages of assembly. We detected 2788 Proteins when comparing 8 conditions (conditions: were all combinations of the 3 variables: beads only or immobilized DNA; Epx or control treatment; 15min or 4h assembly) across 4 biological replicates. The iBAQ data were log2-transformed and subsequently filtered to only include proteins with 6 out of 8 valid values per condition. The data were further processed by subtracting log2(iBAQ) values of samples with beads only control of each protein from their corresponding with immobilized DNA to remove background. This yielded normalized log2(iBAQ) values for each protein indicating their enrichment on the chromatin fiber, dependent on assembly time as well as proteasomal activity. Importantly, no cutoff was set. When plotting the time-dependent intensities of all proteins enriched on the chromatin fiber by treatment (“control” or “20 μ M Epoxomicin”), proteins that bind early and then dissociate have positive values, while proteins that accumulate over time have positive values (Figure 29). Interestingly, we observed that inhibition of proteasome changes the protein dynamics and leads to the overall accumulation of proteins on fiber over time.

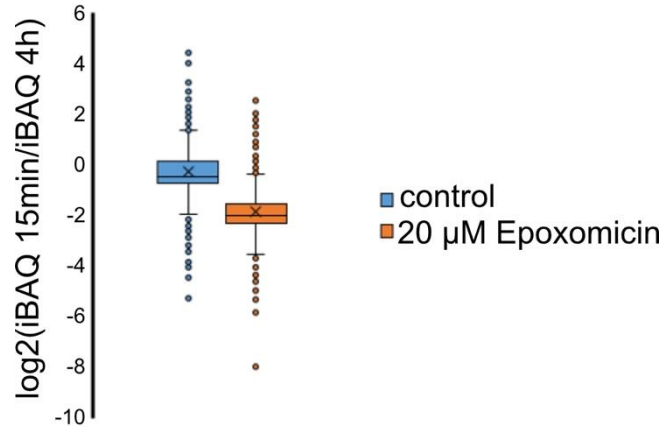


Figure 29 Effect of proteasome inhibition by epoxomicin on overall protein abundance on chromatin fiber over time

Boxplot of protein intensities of all proteins significantly enriched on the chromatin fiber after 15min and 4h under standard conditions (control, blue) and after treatment with 20 μM Epoxomicin (orange). Each box consists of the difference of mean iBAQ after 15min and 4h of each significant protein from N=3 biological replicates. Data based on Dataset 6.

4.2.6 Inhibition of proteasome leads to differential changes in chromatin proteome

After assessing the general overall effect of proteasome inhibition on chromatin assembly, we decided to look more in-depth and investigate the proteins differentially. Therefore, we took the processed dataset (see chapter 4.2.5 above) and ranked the proteins according to their abundance in control and 20 μM Epoxomicin treated for 4h samples. Then, we subtracted the rank each protein had in the control sample from their respective rank in the Epoxomicin-treated sample. This way we were able to calculate and visualize which proteins were enriched or associated less with the chromatin fiber upon proteasome inhibition (Figure 30). Overall, we observed that not all proteins are affected in the same way, neither by strength nor by direction of effect by proteasome inhibition.

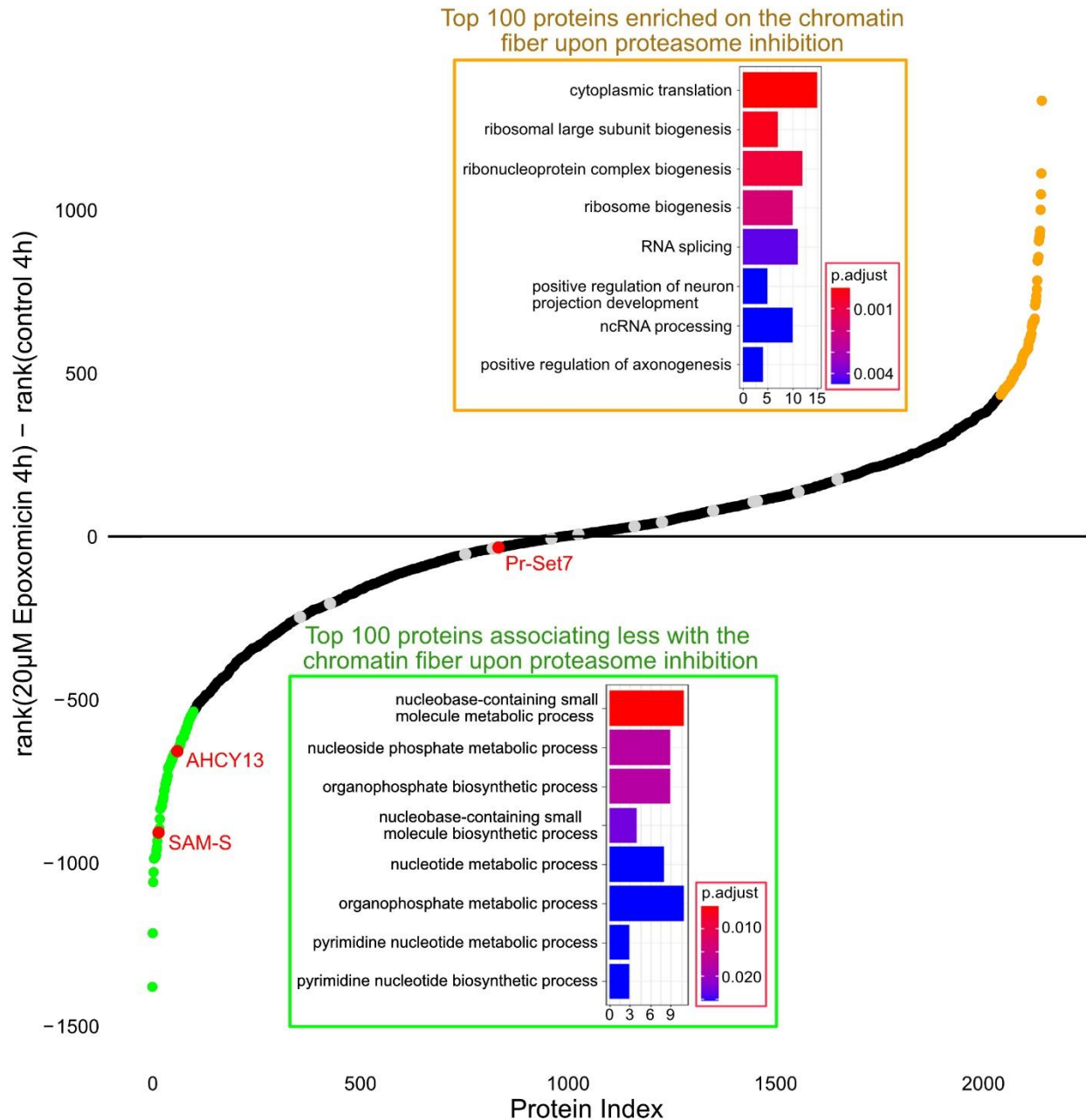


Figure 30 Effect of proteasome inhibition with 20 μM Epoxomicin on abundance of specific proteins on chromatin fiber after 4h

Waterfall plot visualizing proteins on the chromatin fiber concerning their change in abundance (by rank) between control conditions and after Epoxomicin treatment. Proteins above Zero are enriched on fiber upon inhibition, and proteins below 0 are less abundant on fiber upon inhibition. The inset at the top shows the top 8 GO terms by p-value for the top 100 most enriched proteins upon inhibition (orange). The inset at the bottom shows the top 8 GO terms for the 100 most depleted proteins upon inhibition (green). Red-marked proteins Pr-Set7 and AHCY are part of H4K20monomethylation metabolism. Highlighted in grey are the subunits of the 20S proteasome. Data based on Dataset 6.

To analyze the proteins that were enriched on the fiber upon proteasomal inhibition, we performed a GO Term enrichment analysis based on biological processes of the 100 proteins with the highest difference in rank between treatment and control. For these “Top100 proteins enriched on the chromatin fiber upon proteasome inhibition” (Figure 30 top orange inlay box), we found that they

were mostly involved in translational activity, ribosomal biogenesis, and different pathways of RNA processing.

Following that, we assessed the Top 100 proteins associating less with the chromatin fiber upon proteasome inhibition, again using biological processing GO term analysis (Figure 30 bottom green inlay box). Here, we discovered that these proteins were mostly involved in different metabolic processes. Notably, proteins involved in the methionine cycle and H4K20monomethylation, such as SAM synthetase (SAM-S), S-adenosyl homocysteine (AHCY), and the histone methyltransferase Pr-Set7 (Figure 30 all marked in red), were strongly or mildly depleted from the fiber upon proteasomal inhibition.

To sum it up, our data suggests that metabolic enzymes moonlight to produce metabolites for posttranslational histone modifications in situ on assembling chromatin fibers. Therefore, we set out to investigate whether in situ metabolism happens on the chromatin fiber in the system we chose, i.e. DREX-assisted chromatin assembly. The main histone methylation mark present on DREX assembled chromatin is H4K20me1((5,46) and unpublished data by Beyza Bozdağ) Consequently, we chose the metabolic pathway of the methionine cycle and H4K20me1 cycle, consisting of the proteins SAM-S, AHCY and PrSet-7, the only H4K20monomethyl transferase in *Drosophila* (194), as our proof-of-concept study.

4.2.7 Benchmarking metabolomics studies in DREX extract

In collaboration with Marco Borsò and Beyza Bozdağ, I established the detection of metabolite levels in DREX. Metabolites were obtained from the samples via liquid-liquid extraction before being measured via Capillary electrophoresis mass spectrometry (CEMS). Peak selection and analysis were performed by Marco Borsò, using the corresponding m/z and fragmentation patterns of the metabolites.

To benchmark our system, we detected a broad spectrum of metabolites in DREX. These ranged from different amino acids to energy storage and transfer metabolites like ATP and NADH (Dataset 9). We were able to quantify the most important metabolites involved in the methionine cycle in DREX (Figure 31). For this, stable isotopically labeled methionine and ATP were used as spike-ins (Merck), allowing us to directly normalize those metabolites to the spike-in. However, due to the lack of commercial SAM and SAH isotopically labeled standards, the quantification of these two analytes was based on the methionine spike-in, considering their similar retention times and therefore assuming the same ionization efficiency. The magnitude, nevertheless, is correct with high confidence. For higher confidence and analytical exactness, specific spike-ins for SAM and SAH should be used. We observed that ATP and Methionine occur in a concentration of 0.92 μM and 3.86 μM , while SAM and SAH were estimated to be 2.97 μM and 0.06 μM correspondingly in DREX. To sum it up, we were able to benchmark our new technique in the extract we utilized for our studies and thereby laid the groundwork for the investigation of metabolism on the chromatin fiber by challenging the system.

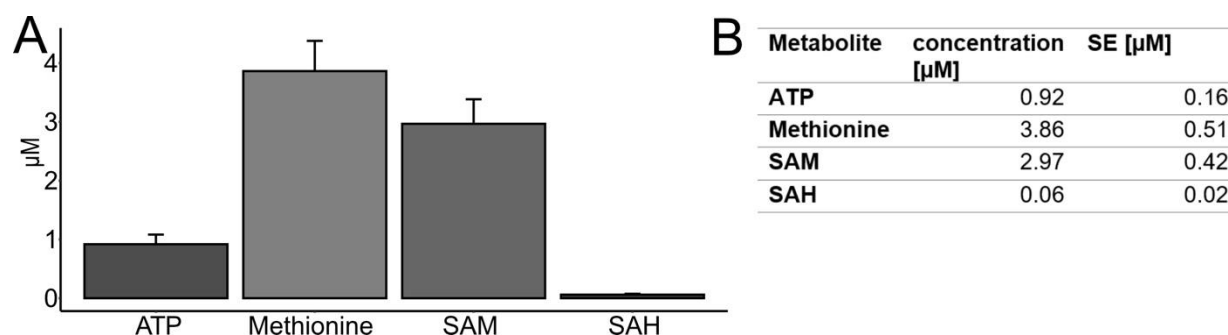


Figure 31 Concentration of metabolites involved in methionine cycle in DREX

(A) Bar plots with metabolite concentrations in DREX. Absolute concentrations were calculated from isotopically labeled spike-in controls. Error bars=SEM, (B) Table of metabolite concentrations and standard errors. N=4. Data based on Dataset 9. *Experiment and visualization in collaboration with Marco Borsò and Beyza Bozdağ.*

4.2.8 Dialysis of DREX extract removes most metabolites

The first challenge we employed to the metabolic system in DREX was the removal of all metabolites via dialysis. We aimed to remove all small molecules while keeping the proteome intact. Several different dialysis methods were tested. Only overnight dialysis in a semi-permeable membrane tube with a molecular weight cut-off of 3.5kDA in a 1:500 sample: dialysate ratio with one buffer change proved to be sufficient for effective removal of the metabolite to the minimal level individually possible (Figure 32 A). Swimmer-based dialysis proved inefficient for the complete removal of metabolites (data not shown). CEMS metabolomics analysis of the DREX extract showed visible differences in the number of masses detected before and after dialysis (Figure 32 B, C). Here, we first visualize the data in an intuitive overview plot that displays retention time against m/z. This visualization does not allow for the assignment of masses to specific metabolites (yet) but gives a good impression of the overall dialysis success.

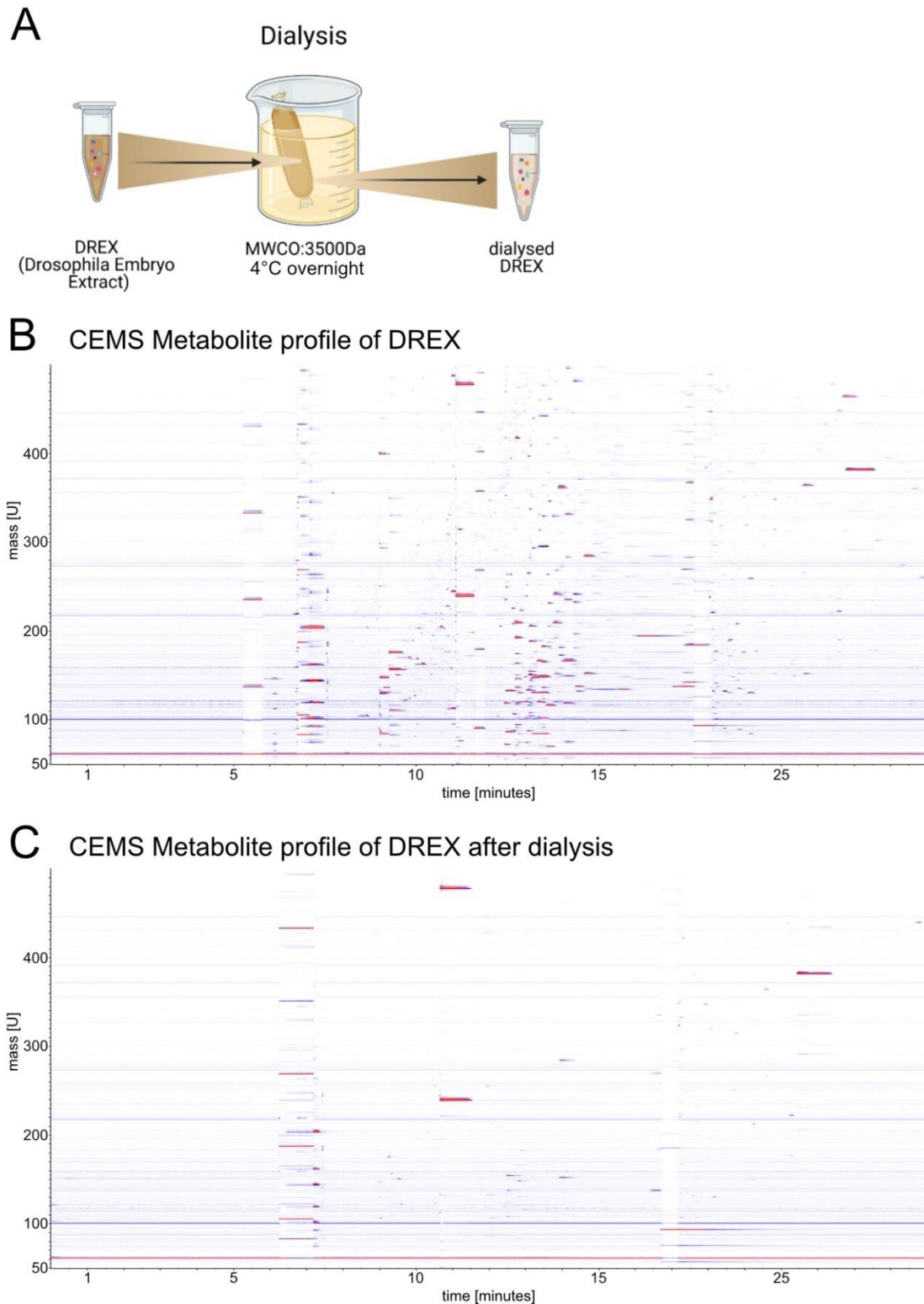


Figure 32 Dialysis of DREX depletes small molecules from the extract

(A) Schematic diagram of the experimental flow of dialysis of DREX extract overnight at 4 °C through 3.5kDa molecular weight cut-off (MWCO) membrane. (B) Overview plot of masses detected over retention time from a sample of undialyzed DREX. (C) Overview plot of masses detected over retention time from a sample of DREX after dialysis. Data based on Dataset 9. *Experiment and visualization in collaboration with Marco Borsò.*

While the overview plots allowed us to grasp the success of dialysis, we were also able to reliably assign and quantify the levels of specific metabolites in DREX before and after dialysis. As a non-dialyzed control, the same volume of extract was put on a shaker in an Eppendorf tube at 4 °C for the same time as the corresponding dialyzed sample was undergoing dialysis. This allowed us to look specifically at the effect of dialysis, excluding any aging/storage-specific differences. Most amino acids and a great variety of other small molecule metabolites were detected and isotopically labeled spike-ins allowed the quantification of a subset, including ATP and methionine. Representatively, we look at metabolites relevant in the context of the methionine cycle in detail: methionine, S-Adenosylmethionine (SAM), and S-Adenosylhomocysteine (SAH) (Figure 33). The pathway and metabolites were of particular interest due to the enzymes found enriched on the chromatin fiber upon proteasomal inhibition before (4.2.6). The metabolome of the dialysis samples revealed that methionine and SAM were efficiently removed from the extract via dialysis (from 3.865 μM to 0.021 μM and 2.967 μM to 0.025 μM correspondingly). Contrastingly, only roughly half of the SAH was dialyzed out from 0.060 μM to a final concentration of 0.028 μM . Overall, we observed a reduction in all metabolites by dialysis to low concentrations but depending on the original concentration, a variety of relative reduction from only 50% down to 99.5% depletion.

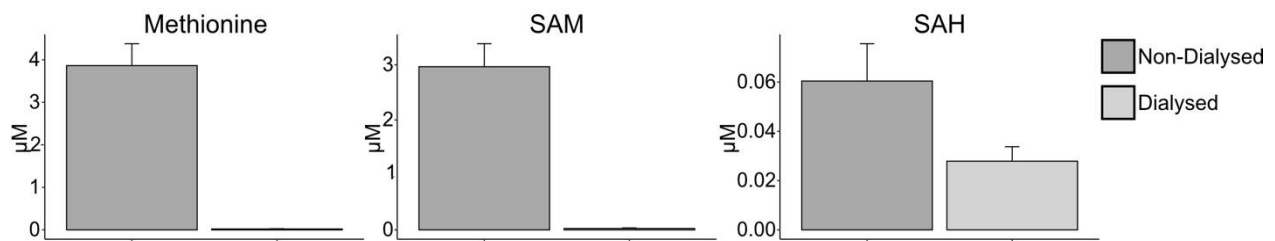


Figure 33 Dialysis of DREX depletes or reduces metabolites involved in the methionine cycle from the extract

Bar plots with metabolite concentrations before dialysis and after dialysis. Absolute concentrations were calculated from heavily labeled spike-in controls. Error bars=SEM. N=4. Data based on Dataset 9. *Experiment and visualization in collaboration with Marco Borsò and Beyza Bozdağ.*

4.2.9 Histone modifications upon metabolite depletion

After setting up and benchmarking our methods and analysis, we ran our investigation of the metabolic coupling of histone modifications. For this, we chose the H4K20me1 modification in our DREX system, as we already showed the association of all involved enzymes on the chromatin fiber during chromatin assembly and the presence of all necessary metabolites in DREX. Furthermore, we know from the literature (5,46), as well as our data, that this histone mark is reliably set during chromatin assembly *in vivo* as well as in our *in vitro* system.

First, we wanted to challenge the system by depleting our extract of all free small metabolites. Hence, we dialyzed the extract as described before and subsequently added the ATP regenerating system as well as the beads-immobilized linear DNA to the dialyzed or non-dialyzed extract. After 1h of assembly, we isolated the chromatin and analyzed the histone modifications. Looking at the H4K20 peptides, we observed that in our control, undialyzed sample about 40% of the H4K20 were monomethylated while about 60% remained unmethylated (Figure 34 1st of bars). We detected no di- or tri-methylation on lysine 20. Surprisingly, we did not see a reduction, but an increase of over 60% of monomethylation when the assembly was performed in dialyzed DREX (Figure 34, 2nd set of bars).

In addition, we wanted to look into the effect of the addition of 500 µg/µl SAH on the H4K20methylation in the assembly system. SAH is a byproduct and competitive inhibitor of the H4K20 monomethylation. Indeed, we saw that H4K20 monomethylation is reduced to below 20% when SAH is added to the standard assembly with undialyzed DREX (Figure 34 3rd set of bars). If the assembly is performed in dialyzed DREX, this effect is even more pronounced with only below 5% of H4K20 peptides getting methylated (Figure 34, last set of bars).

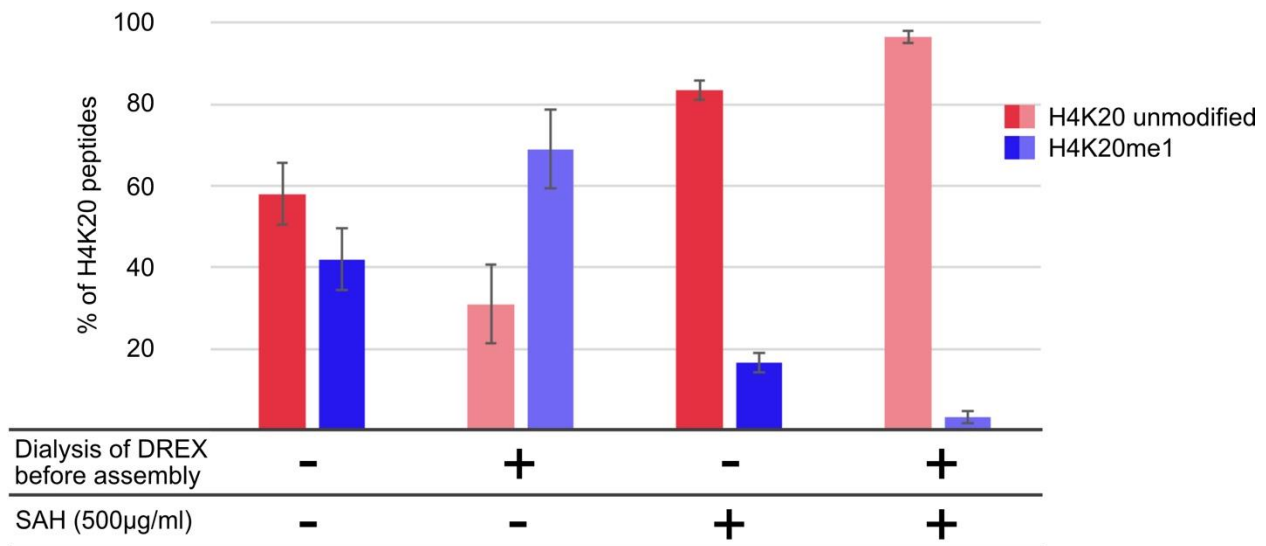


Figure 34 Effect of DREX dialysis and SAH addition on the histone modification H4K20monomethylation

Bar plot quantifying the percentage of unmodified H4K20 (red) as well as H4K20me1 (blue) modification after 4h chromatin assembly in DREX extract dependent on variables of dialysis (cutoff 3.5kDa) of DREX before assembly and addition of 500µg/ml SAH during assembly. Error bars=SD. N=3. Data based on Dataset 8.

Overall, we were able to successfully quantify and thereby observe histone H4K20 monomethylation upon challenging the metabolic system in DREX. While only the depletion of metabolites caused surprisingly little change in H4K20me1, the treatment with SAH caused a strong response and showed sensitivity to dialysis. To sum it up, we were able to gain general knowledge about the DREX metabolic activity, including the metabolic coupling of H4K20me1 during chromatin assembly in the DREX system.

4.2.10 Supplementation of isotopically labeled Methionine leads to incorporation of isotopically labeled methyl to H4K20me1

To test the metabolic activity of the DREX extract, we decided to track the methionine cycle and its link to histone methylation. For that, we added 0.5 µM of isotopically labeled methionine (¹³C, D₃) to the dialyzed extract during chromatin assembly. This is about a tenth of the original methionine concentration in DREX of 3.86 µM as quantified by us during metabolism benchmarking. After 1h of assembly, we isolated the chromatin and analyzed the histone modifications via LC-MS.

We looked at H4K20monomethylation to look for potential incorporation of isotopically labeled methyl. In the control condition, we, again, detected approximately 60% monomethylation (Figure 35,

control) for the assembly in dialyzed DREX, similar to our earlier experiments (see Chapter 4.2.9 and Figure 34 above). In the MS1 spectrum, we were able to identify the isotopic envelope of the peptide (Figure 35). When we added isotopically labeled methionine to the reaction, we observed about the same level of monomethylation of the H4K20 peptide as in control (Figure 35, $+^{13}\text{CD}_3$). Conveniently, we can separate the peaks of the light and heavy monomethylated histone peptides by integration of the MS1 spectrum. Here, we see that about 90% of the monomethylation detected is heavy (across 3 replicates, Figure only shows one representative spectrum). Curiously, we did not detect an increase in H4K20me1 nor an increase in the ratio of isotopically labeled methylation compared to light methylation even when we increased the concentration of isotopically labeled methionine from 0.5 μM stepwise up via 5 μM and 50 μM to 500 μM (data not shown). To sum it up, we successfully showed that methyl groups from a heavily labeled methionine reappear as histone modifications after 1h or chromatin assembly in DREX.

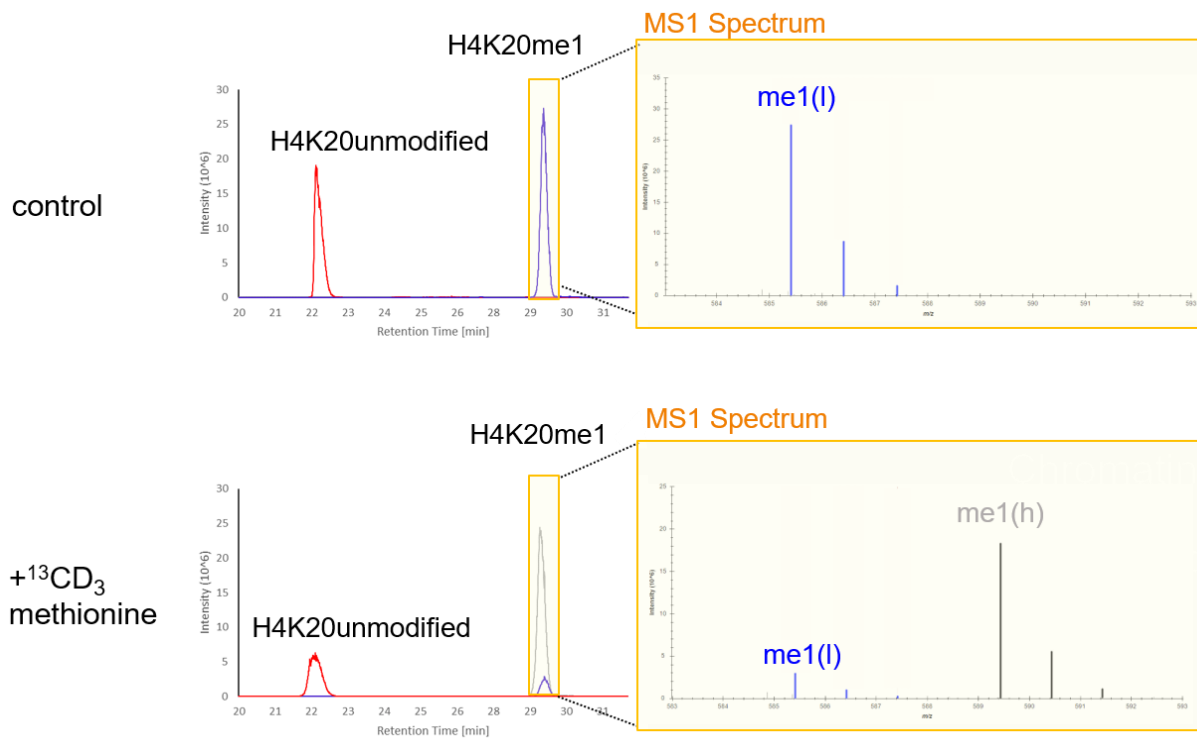


Figure 35 DREX incorporates methyl group from isotopically labeled methionine on H4 histones

Chromatogram and MS1 spectrum of H4K20 unmodified peptide: K[+56]VLR and peptides with H4K20 monomethylation peptides: K[+70]VLR for light $^{12}\text{CH}_3$ methyl group (me(l)) and K[+74.1]VLR for isotopically labeled $^{13}\text{CD}_3$ methyl group (me(h)). Data based on Dataset 8.

4.2.11 SAH excess traps Pr-Set7 on and removes AHCY from chromatin fiber

To follow the change in proteome on the chromatin fiber upon inhibition of the H4K20monomethylation reaction, we added 500 $\mu\text{g}/\text{ml}$ (1300 μM) of the inhibitor SAH to the DREX assembly reaction (baseline concentration of SAH in DREX is 0.06 μM). This experiment was conducted in dialyzed and non-dialyzed DREX.

After the pulldown of the chromatin fiber, the associated proteome was measured by LC-MS. A total of 3206 proteins were detected across the 3 biological replicates. The LFQ values indicating the protein abundance were log2 transformed. The data were filtered to only keep proteins with valid values for at least 3 out of the 4 conditions (non-dialyzed control, non-dialyzed +SAH, dialyzed control, and dialyzed + SAH) per biological replicate. Subsequently, all missing values were imputed from a normal distribution with a width of 0.3 and a downshift of 1.8, separately for each column.

We compared the proteins associated with the chromatin fiber in the control and in the treated sample as visualized in the volcano plots (Figure 36). We observed a difference in protein abundance (LFQ) for multiple proteins, some of which were significant. This effect was visible for assembly in pre-dialyzed as well as non-dialyzed DREX (Figure 36 A, B), albeit more pronounced in the dialyzed DREX. The analysis showed that multiple histone methyltransferases associated more with the chromatin fiber upon the addition of SAH to the reaction, including Pr-Set7, E(z), Mes-4, G9a, and esc. Most notably, Pr-Set7, and under dialyzed conditions also E(z), is significantly enriched on the fiber. Contrastingly, we observed the SAH hydrolyzing enzyme AHCY13 associating less with the fiber upon SAH addition.

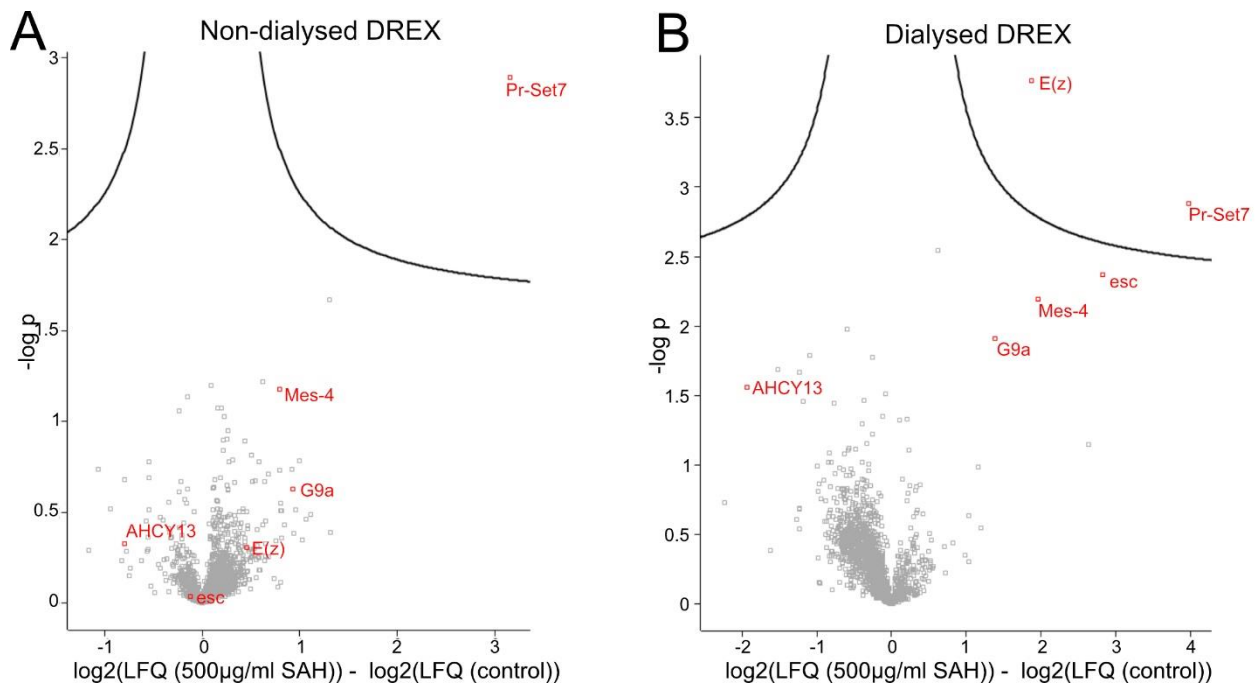


Figure 36 Histone methyl transferases accumulate on chromatin during assembly upon SAH addition

(A) Volcano plot comparing proteome of standard chromatin assembly and assembly with the addition of 500 µg/ml SAH in non-dialyzed DREX. (B) Volcano plot comparing proteome of standard chromatin assembly and assembly with the addition of 500 µg/ml SAH in non-dialyzed DREX. Histone methyltransferases and AHCY are marked and labeled in red. $n=3$ FDR=0.05. Data based on Dataset 7.

4.2.12 Metabolites reappear/metabolome slowly recovers over time in DREX after dialysis

After establishing that metabolism is active in DREX, we wanted to determine if at all and at what rate and with which dependencies the metabolites recover after they are dialyzed from DREX. Again, we used the set of metabolites of the methionine cycle as a proof of concept. Here, we used dialyzed DREX and subsequently incubated it at 26 °C. For the different experimental setups, we added either none or the described molecules and took samples after the indicated times. The metabolites from these samples were immediately liquid-liquid extracted and then measured via CE-MS.

First, we wanted to check whether the recovery of metabolites happens in DREX, and if yes, its dependency on DNA. Therefore, one sample of DREX had DNA, in the form of the plasmid also used for assemblies, while the other one had no DNA. Importantly, both samples were always supplied with an ATP regenerating system. The metabolites methionine, SAM, and SAH were quantified in relative levels compared to the non-dialyzed DREX as a control. We observed changes in metabolite levels over time (Figure 37). Interestingly, we discovered that methionine and SAM only recovered up to barely 10% and 5% of their original levels, correspondingly over 120 min tested (Dataset 9). Additionally, we saw that any SAH remaining after dialysis was reduced to about 9% of the concentration in non-dialyzed DREX within the first 5 min and afterward SAH levels stayed consistently low (between 0.3% and 16%) over the whole time tested. Overall, we found a time-dependent change in metabolite levels in DREX but is not dependent on DNA presence.

The SAH concentration at time point 0min, the SAM concentration at 120min as well as all methionine measurements showed substantial variance. We had high variance because of outlier replicates that we decided to still include to truthfully show the variation in the metabolic experiments. We do, however, not know whether this outlier is of a technical or biological nature but believe that this time course shows a dynamic change that is sensitive to even small delays in sample collection.

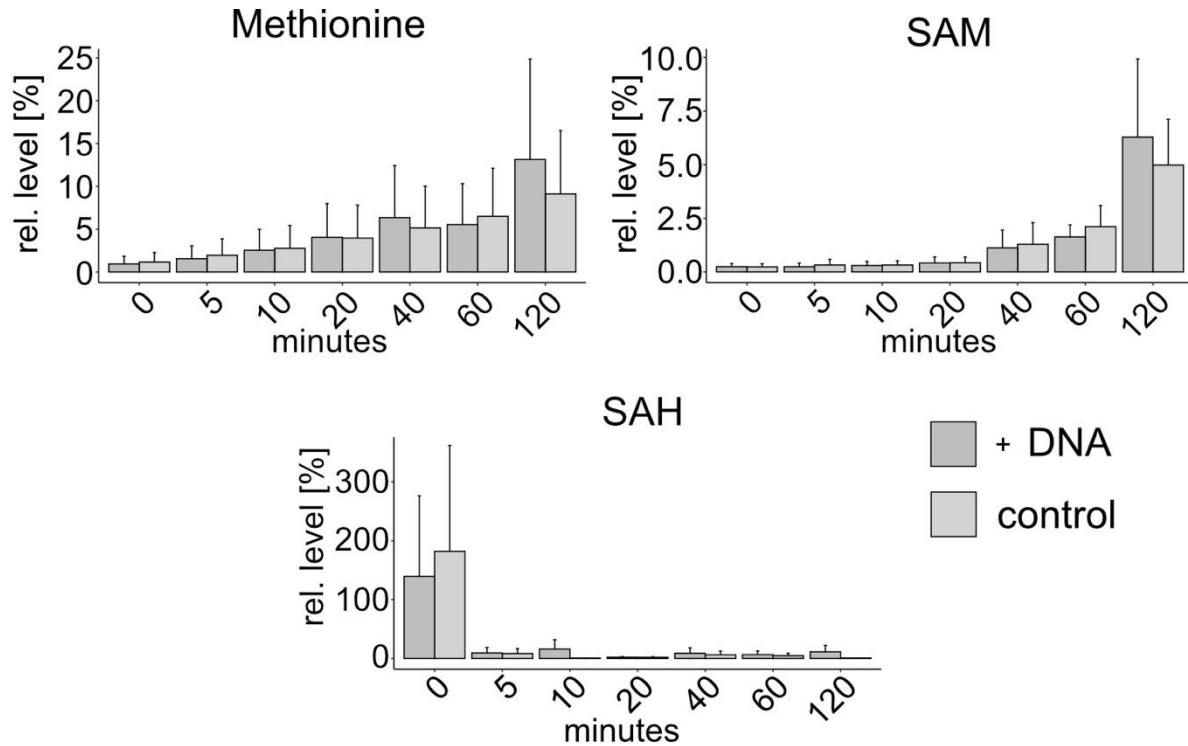


Figure 37 DNA dependence of recovery of metabolites involved in methionine cycle over time after dialysis of DREX

Bar plots with metabolite levels after different time points after dialysis. Relative levels in % are normalized to corresponding levels in the non-dialyzed DREX sample. Error bars=SEM. N=3. Data based on Dataset 9. *Experiment and visualization in collaboration with Beyza Bozdağ and Marco Borsò.*

Next, we investigated whether the observed metabolite recovery was dependent on the addition of the ATP regenerating system (Figure 38). Therefore, we had one DREX sample with no additional substances added, while the other one was supplied with the ATP regenerating system (consisting of ATP, creatine kinase, and creatine phosphate).

We observed that methionine recovered slightly faster when ATP is visible but, independent of the condition, methionine only recovers to about 2% of its concentration in original DREX (3.86 μM as measured in our metabolite benchmarking in chapter 4.2.7). On the other hand, SAM recovery shows a high dependency on ATP and additionally recovers up to roughly 10% of its concentration in the original DREX (2.97 μM as measured in our metabolite benchmarking in chapter 4.2.7). SAH shows yet another pattern: not increasing in concentration over time but leveling out. This level is constant and depends on ATP. In the presence of ATP, SAH disappears below 1% (0.001 μM), while it levels out at around 0.02 μM (30% of its level in undialyzed DREX) over time when ATP is not supplied. Overall, the variance was lower than in our first dataset but still showed up, most likely due to sample-collection-time-dependent effects.

Finally, we also measured ATP concentration itself as the addition of the ATP regenerating system had so far only been qualitatively observed to be sufficient to support chromatin assembly. However, no quantification has been successful so far and the longevity of the system has not been tested before. Here, we observe that directly after the addition of the ATP regeneration system, a very high ATP concentration of around 70 μM levels out at around 20 μM and stays constant in the time frame tested

(Figure 38). This concentration was much higher than the “natural” ATP concentration of $0.5\ \mu\text{M}$ in undialyzed DREX (as measured in our metabolite benchmarking in chapter 4.2.7).

Overall, all metabolites we investigated in detail showed a dependency in their concentrations on ATP availability. However, there is a great difference in the strength of that dependency ranging from mild to total. Additionally, we saw that not all metabolites recover but some level out over time.

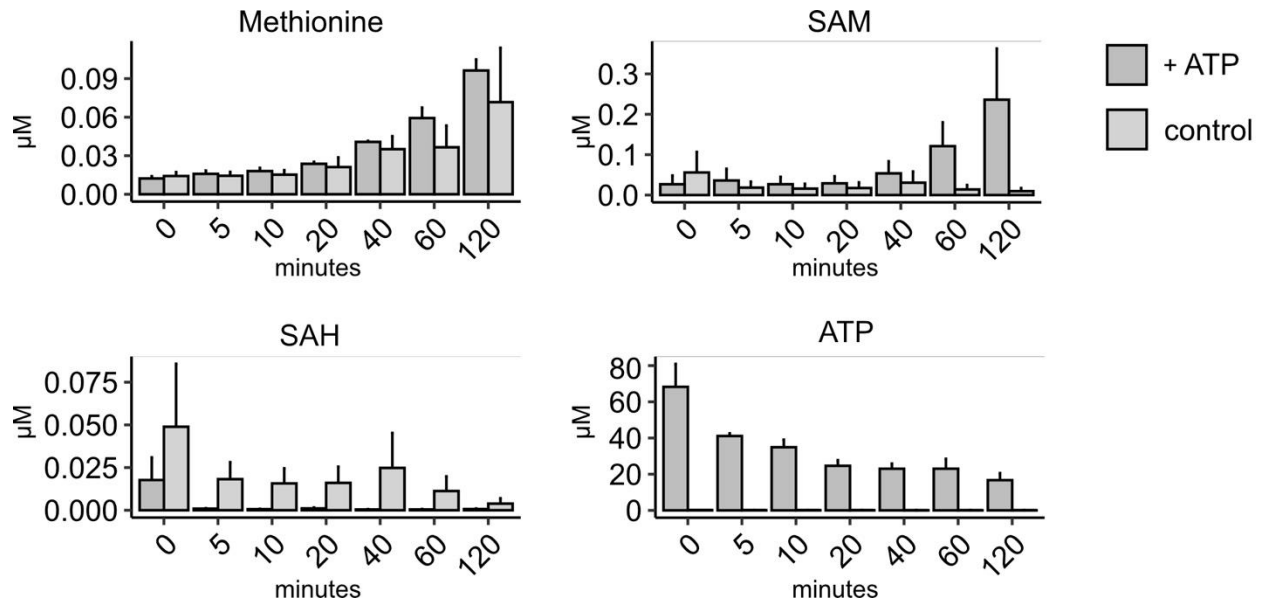


Figure 38 ATP dependence of recovery of metabolites involved in methionine cycle over time after dialysis of DREX

Bar plots with metabolite levels after different time points after dialysis. Absolute concentrations were calculated from heavily labeled spike-in controls. Error bars=SEM. N=3. Data based on Dataset 9. *Experiment and visualization in collaboration with Beyza Bozdağ and Marco Borsò.*

4.3 Discussion

4.3.1 The DREX proteome

While the presence of certain proteins had been implied or directly reported via proteomic pull-down analysis and immunochemistry (5,24,28,33-35,46), no global semi-quantitative analysis of DREX had ever been performed up to date. Overall, the identification of 1650 proteins is much deeper than any unbiased/non-enriched analysis in the DREX composition run before. However, the proportionally very high abundance of a few factors poses the challenge of a substantial dynamic range, thereby limiting sensitivity by obscuring low abundant proteins (195). To identify more low-abundant proteins, this limitation could be overcome by fractionation of the extract or by removal of the most highly abundant proteins e.g. by antibody depletion (196) or by antibody-targeted specific degradation (197). Despite its limitations, this dataset compromises a wealth of information and novel insights.

The high abundance of the 3 Yolk proteins is expected and nicely fits their role to serve as metabolite precursors which are broken down to provide embryos with nutrients essential for growth and development (198). These Vitellogenins are essential maternally deposited factors produced in the fat body and secreted for uptake by maturing oocytes. Fittingly, biological process-based GO term analysis revealed an enrichment of proteins involved in “catabolic processes” in all proteins but particularly in the 100 most abundant ones. The detected proteins in this GO term included proteasomal subunits and other factors involved in the ubiquitin-proteasomal system (UPS). The proteasome has been reported to be present in the early *Drosophila* embryo (199). Overall, these findings nicely tie together, creating an image suggesting the role of the main degradation machinery of the cell, the proteasome, as an essential factor in the generation of metabolites for histone modifications.

Additionally, our analysis showed an enrichment of factors involved in the mitotic cell cycle and microtubule-based processes. In the preblastoderm embryo, the genome is replicated and packaged into chromatin every 9 min, requiring a high abundance of maternally deposited histones and assembly factors as was already suggested by Becker et al in 1992 (24). However, here it is shown proteomically and for as many proteins differentially for the first time. Interestingly, we also found proteins involved in translation that were highly enriched in DREX. While transcription is not active in these embryos (200), translation of maternally deposited mRNAs occurs in the embryo before it switches to zygotic control (201). The syncytial state of the embryos at collection, making the extract neither exclusively nuclear nor cytoplasmic extract explains this coappearance of nuclear and translational factors.

With its pre-transcriptional timing of the embryo, many transcription factors are absent from the extract while others are detectable, making the extract an excellent model for transcription factor targeting (33,34,202). Together with the comprehensive proteomic data on assembling chromatin fiber in DREX, this resource will aid future insights into how and which proteins are targeted to the fiber.

Overall, the produced proteome gives further insight into the nature and composition of the versatile Preblastoderm embryo extract (DREX). A first unbiased view already revealed the many applications of this resource. We were able to back long-standing hypotheses about the protein composition with data and tie our findings into the biological contexts of embryo development as well as provide a comprehensive protein list including relative abundances for downstream analysis. This DREX proteome is a long-awaited tool and resource that will be wielded by researchers using the extract to explore different biological concepts, pathways, and questions.

4.3.2 What makes the DREX functional: concentration or composition?

When preparing different batches of DREX, researchers are confronted with the challenge that while all are produced with the same protocol, not all extracts produced are capable of assembling chromatin within the specifications provided. Therefore, a multi-day test assembly and titration has to be performed with each new DREX batch. While we know the presence of histones, histone chaperones and other chromatin assembly factors in high concentrations are vital, it remained elusive whether one single protein or a collection might be a predictive determinant of which batch would assemble chromatin successfully. We investigated two different features of DREX batches, total protein concentration, and proteome composition to determine the correlation of these factors to assembly capability.

Our data suggest that the ability of a DREX batch to assemble chromatin is determined by total protein concentration, but that at least small differences in protein composition do not affect assembly. The data were collected across 4 replicates, including 3 DREX that were able to assemble chromatin and 1 DREX that was not. We show that only very few proteins differ significantly in their abundance between assembly-competent or incompetent extracts. The few proteins that do differ have no functional link to chromatin assembly, function, or structure. They are most likely an artifact caused by the limited amount of non-assembling DREX batches available and analyzed. Since all tested DREX had approximately the same protein composition but one did not assemble chromatin, the right composition is necessary but not sufficient to allow for chromatin assembly.

Protein concentration is the main denominator, within the variance we find in prepared DREX, for chromatin assembly potential in DREX. The higher the total protein concentration, the lower the extract volume per μg DNA is necessary to assemble chromatin. Additionally, below a certain threshold, which appears to lie around $50\mu\text{g}/\mu\text{l}$ (measured by BCA with a BSA standard), the extract is not competent to assemble chromatin anymore. This is explained by the involvement of a great number of different proteins to assemble chromatin. These necessary proteins seem to have to be present at a high enough abundance.

We conclude that both, composition and concentration are necessary to make DREX a functional extract. Total protein concentration, is a good proxy for whether the DREX batch is capable of assembling chromatin and how much DREX is necessary per unit of DNA. To finally prove that the dilution factor is a relevant difference, a dilution of an assembly-capable DREX to the concentration of the DREX categorized as “non-assembling” and subsequent determination of whether the extract does consequently become non-functional would be an option. Identification of essential factors in protein composition that might render the extract non-assembly competent might be explored by removing certain factors by immunoprecipitation or degradation (197).

4.3.3 Proteasomal activity in DREX

We successfully showed that the proteasomal subunits are not only present and abundant in DREX (Figure 26) but also assemble into a fully enzymatically functional proteasome complex (Figure 27). Additionally, we were able to test and benchmark different proteasome inhibitors about their efficacy and efficiency.

Unfortunately, the inhibitor combination of MG132 + NEM used earlier in the project proved to be unsuitable for further analysis. Along the off-target effects (188-190), an observed disturbance of clumping of beads upon the addition of the inhibitors but also in a vehicle-only control (data not shown) makes this inhibitor combination unreliable. The clumping is most likely caused by the limited solubility of the NEM inhibitor, necessitating the addition of a high volume of solvent ethanol to the reaction. This leads to a disturbance of the process biophysically, not chemically/biologically. Epoxomicin proved to be an effective and efficient inhibitor that could be applied with minimal use of solvent.

This study reports the first direct detection of proteasomal enzymatic activity in DREX going beyond reports of the presence of subunits (199). This finding was also confirmed outside of this project by preliminary experiments applying the proteasome-dependent TRIM away system (197) in DREX by Nikolas Eggers in the Becker lab (data not shown). The proteasomal activity in DREX also represents our first indication of an active and complex metabolic system in DREX:

4.3.4 Inhibition of proteasome leads to overall accumulation and differential changes in chromatome

While the inhibition of the proteasome does not interfere with basic processes like histone integration and spacing, we do find differential changes in the chromatin-associated proteasome. From our Mnase-based studies, we conclude that chromatin assembly at its most basic macroscopic level, histone placement, is still achieved in DREX extract upon proteasome inhibition with Epoxomicin (Figure 28). Our method, however, only compromises a very zoomed-out and general view, allowing us to make sure that assembly still happens and we can compare the epoxomicin-treated samples to control and get meaningful data. An approach for the future to take a closer look at the integrity of histones and their positioning might be to investigate the role of the proteasome in histone degradation and chromatin plasticity (203). To examine this compelling role in the DREX chromatin system, the use of longer DNA templates and different readouts like DNA hypermethylation coupled to nanopore sequencing to investigate nucleosome density might be employed (204).

In our proteomic study of the assembling chromatin fiber, we detect more proteins than in the proteome of the whole DREX. This increased sensitivity is caused by the smaller dynamic range (195). Overall, we reproducibly found that proteins accumulate on the chromatin fiber over time when the proteasome is inhibited (Figure 29). However, we also showed that proteins are differentially affected in their binding to chromatin by proteasome inhibition (Figure 30). Not all protein groups accumulate; some subsets also bind less.

The proteins accumulating on the fiber upon proteasome inhibition are proteins involved in ribosome biogenesis, translation, and transcription as well as RNA processing. The regulation of transcriptional elongation and termination might both be under proteasomal control (205,206). The inhibition of proteasome activity leading to the reduction of proteins of the RNA-related processes can therefore be explained by failed recruitment of the factors when the proteasome is inactive or the degrading enzymes' failure to remove inhibitory proteins. Proteins of ribosome biogenesis are associated with chromatin in adult cells because early ribosome biogenesis happens in the nucleolus at the genomic rDNA tandem where the four ribosomal RNAs (rRNAs) are transcribed and then dissociate from the fiber (207). However, in the preblastoderm DREX, transcription is not active (200) and ribosomal

subunits associate less with the chromatin fiber under control conditions (not visualized here). Their accumulation upon proteasomal inhibition suggests that their removal or non-association under control conditions is mediated by the proteasome directly or indirectly. This finding of proteasome-mediated degradation of proteins involved in translation is consistent with literature describing co-translational degradation of elongation factors (208). Overall, our result suggests that the proteasome directly or indirectly removes proteins from the chromatin that do not, at this point in development, fulfill a role for chromatin structure of function and fails to do so when its proteolytic function is impaired.

Most interestingly, the subset of proteins that bind less to the fiber upon proteasome inhibition contains mainly proteins involved in chromatin-proximal metabolism. This contains but is not limited to, nucleotide/nucleoside metabolism. The 26S proteasome requires nucleotide - ideally ATP - hydrolysis for the degradation of ubiquitinated proteins (209). As the proteolytic activity of the proteasome is inhibited by Epoxomicin, the factors of this process seem to detach from chromatin, and potentially so do proteins of other processes that are interlinked to this activity by the dependency on the common nucleotide ATP.

Curiously, the proteins of the methionine cycle, SAM-S, and AHCY, are also less abundant on the fiber when the proteasome is inhibited. Methionine, together with ATP, constitutes the essential precursor of SAM, the methyl donor for histone methylation (210). This modification can be found in DREX mainly in the form of H4 lysine 20 monomethylation (H4K20me1) (5,46). We hypothesize that proteasomal activity produces the metabolite methionine (and potentially other metabolites as well) through proteolytic degradation of other proteins. Under normal conditions, the proteasome thereby creates a methylation-favorable environment that is associated with the presence of the enzymes on the fiber. This environment is disrupted upon inhibition of the proteasome.

The abundance of subunits of the core 20S proteasome itself on the chromatin fiber remains somewhat unaffected by the proteolytic inhibition. The proteasome appears to be present still, simply not functional or active. This suggests that proteolytic function does not seem to be necessary for the recruitment or association of the proteasome on the chromatin fiber.

4.3.5 Metabolite profiling in DREX

The initial metabolomic profiling of DREX in collaboration with Marco Borsò and Beyza Bozdağ now allows a more holistic approach to investigate metabolic pathways in DREX by expanding the mass spectrometry tool kit of proteomics and histone modification analysis with a state-of-the-art metabolomics approach.

Our benchmarking showed a successful application of metabolite detection via CEMS for the DREX extract. The intuitive retention time vs m/z overview plot might be used in the future using pattern recognition algorithms to readily identify and track metabolite levels dynamically in a time stack/over time as has been explored for NMR-based metabolomics (211). This can be done over different time points or by administering different manipulations and challenges to the extract to then read out the systematic metabolic reaction of the extract. Additionally, this method could be applied to different systems. Hereby, DREX constitutes an ideal proof-of-concept model due to its easy manipulation and high yet limited complexity.

A great variety of metabolites was identified and a subset was quantified using isotopically labeled spike-ins. In the future, Marco Borsò will expand the portfolio of quantifiable metabolites using isotopically labeled spike-ins for a greater variety of metabolites. Additionally, differential metabolomes of experimental fractions e.g. DREX-assembled chromatin and remaining supernatant could be established to investigate the formation of microenvironments and enrichments.

Biologically, the first metabolomic snapshot of a steady-state DREX already yielded physiologically relevant insights. We observed remarkably reproducible metabolite concentrations among biological replicates despite the complexity of the extract (Figure 31). The absolute levels of methionine (3.86 μM) in our extract appear to be at only 15% of that measured via LC-MS in *Drosophila* embryos by others. However, SAM and SAH concentrations in DREX (2.97 μM , and 0.06 μM correspondingly) are in the same dimensions as those measured in *Drosophila* embryos in the same study (1.72 μM SAM and 0.15 μM SAH) ((212) supplement 5). Overall, this confirms the potential of the extract for transferable metabolic studies.

Moreover, our data allowed us to calculate the methylation capacity of our extract, which is defined by the SAM/SAH ratio, a determinant relevant to enzymatic methylation reactions. The ratio calculated from our data was 50 and equaled the one obtained by Marco Borsò for *Drosophila* cells (data not shown), revealing that any difference seen in methyltransferase activity between the in vitro system and in vivo reports does not result from a difference in the metabolite ratio.

Dialysis of the DREX extract and subsequent metabolomics snapshot analysis allowed us to gain even more insights into our system. In our study of the metabolites of the methionine cycle, we find that SAH is not dialyzed out as efficiently as the other metabolites, such as SAM and methionine (Figure 33). SAH has a molecular weight of 384 DA, therefore with a membrane size cutoff of 3.5 kDa, the size of the molecule is not the reason it was not dialyzed completely. Dialysis method and time are also unlikely to cause an issue as determined by the successful dialysis of the other molecules of the category. However, we observe that all metabolites are dialyzed out to a final concentration of around 0.02 μM . This might also simply mark the concentration limit to which the metabolites can be reduced using the dialysis method chosen. While this reduction to 0.02 μM comprises a major relative reduction for the naturally more abundant metabolites, it is a smaller decrease for the less abundant SAH. Nevertheless, we know we have not reached our detection limit and therefore suspect a biological reason for the phenomenon observed. We hypothesize that the SAH pool is partially bound to enzymes or protein complexes and thereby protected from dialysis. To evade dialysis, the metabolite has to be quite stably bound to the protein(s), as it has to stay protected over the whole dialysis time of roughly 18h. The treatment of the samples before the analysis releases the metabolite from the proteins(s) allowing us to detect bound as well as unbound metabolites. Due to additional data suggesting the trapping of Pr-Set7 on the fiber, we believe the enzyme is bound to AHCY or Pr-Set7. Therefore, we propose further binding studies to verify this theory of metabolite-protein binding by e.g. fluorescence or mass spectrometry-based assays (213,214).

To sum it up, metabolomics was established and employed on different levels from intuitive overview plots with exciting future development possibilities to in-depth identification and quantification of specific peptides. To our knowledge, we report the first successful absolute quantification of metabolites in *Drosophila* embryo-derived extracts. The biological insights gained allow for a deeper understanding of the DREX model system and its new application as a metabolic study system. Finally, we collected several valuable indications, including the metabolite concentrations and ratio of the

metabolites involved and SAH partial protection from dialysis, which contributed to our hypothesis of metabolic coupling of H4K20me1 during chromatin assembly.

4.3.6 DREX exhibits complex metabolic activity

Our first indication for metabolic activity in the extract was discovered early in the project when we successfully detected and inhibited proteasomal activity in the DREX 4.2.3. We have shown the degradation capacity of DREX. This prompted us to dive deeper into the potentials of active metabolic pathways in DREX to establish the extract to study metabolism. We wanted to know if there was more metabolic complexity in DREX, allowing proteins and metabolites not only to be degraded but also recycled, made anew, or transferred. We identified metabolic proteins in the DREX (4.2.1, 4.3.1), characterized their behavior on the chromatin fiber in response to the metabolic challenge of proteasomal inhibition (4.2.6, 4.3.4), and showed the presence of metabolites in DREX (4.2.7, 4.3.5). After this, we now investigated, characterized, and challenged the C1 metabolism coupled to H4K20me1 modification in the extract (4.2.9, 4.2.10, 4.2.12). Thereby, we collected multiple indications as well as definite proof of the metabolic activity of the DREX beyond an active protein-degrading proteasome.

To sum up all subchapters below, we conclude that a great variety of metabolism is active in DREX. We showed active metabolism for the proteasomal degradation pathway as well as for the methionine cycle linked to H4K20me1 histone modification. We not only find metabolic enzymes associated with these two pathways but also with others, for example, nucleotide metabolism (Figure 30). Additionally, our metabolome profiling as well as recovery after dialysis studies also showed the presence and dynamic, ATP-dependent, levels of not only metabolites involved in the methionine cycle but also many other metabolites including other amino acids and nucleotides (4.2.7). The proof-of-concept studies surrounding the proteasome and methionine cycle in combination with histone methylation successfully showed their activity in the extract. In combination with the finding of the presence of a wealth of other metabolic enzymes and metabolites in DREX, the data strongly suggests the activity of many other pathways in the extract, making DREX an excellent system to investigate chromatin-related metabolism *in vitro*.

4.3.6.1 H4K20monomethylation is persistent after dialysis

We observed that even upon the partial or full depletion of metabolites from DREX by dialysis, H4K20monomethylation is still reliably set. We could exclude unsuccessful dialysis of the key metabolites, methionine and SAM due to our comprehensive metabolome studies showing the absence of these directly after dialysis (Figure 33). Therefore, we conclude that SAM and methionine generated a new in the DREX. The necessary enzymes for this, the proteasome as well as SAM Synthetase, are associated specifically with the chromatin fiber in our proteomic studies (control in Figure 30, Figure 36). Potentially, this generation of metabolites even happens directly on the chromatin fiber where the metabolites are needed for histone methylation e.g. by the histone methyltransferase Pr-Set7.

Comparing our values of the control condition (no dialysis) to those previously reported via Western Blot (5), we notice that monomethylation is detected to a higher degree in our proteomic experiments after the same time. Interestingly, we detect low variation among biological replicates regarding the grade of methylation after 1h. Variance in methylation speed had been observed in the lab before (not published). The cause for this difference is not known but might result from differences in detection

method. Additionally, an experimenter-dependent bias might be introduced because DNA to DREX ratios for assemblies are not determined via protein or metabolite concentration but by volume and visual judgment of assembly efficiency by Mnase-digested chromatin on Agarose gel.

4.3.6.2 SAM: SAH ratio influences methylation efficiency of Pr-Set7

We found that the addition of SAH blocks H4K20monomethylation. SAH is a product of the methylation reactions with SAM a methyl donor and is known to act as a competitive inhibitor of the reaction. The inhibition of monomethylation by SAH had been shown in an isolated histone methyltransferase assay (5) and proved transferable to the DREX system, highlighting the DREX as a useful system to study this modification and the sensitivity of histone modification setup in crosstalk with its metabolic environment. The SAH concentration of 500 μM chosen is the same used in literature where SAH was first shown to efficiently inhibit H4K20monomethylation by PR-Set7 in DREX (5). Comparing the concentration of our now-known baseline SAH concentration of 0.06 μM in DREX as measured by us, it represents a great excess addition.

The SAM/SAH ratio determines the methylation capacity of a system. In non-manipulated DREX, the ratio is about 50 (2.97 μM /0.06 μM), according to our metabolite benchmarking (Figure 31). The methylation capacity of *Drosophila* S2 cells is at a comparable level, around 55-65 (data not shown, Marco Borsò). By adding an excess of 500 μM SAH, the methylation capacity drops to 0.006 (2.97 μM /500 μM), reducing the efficiency of H4K20monomethylation by Pr-Set7 in the extract. This inhibition of methylation is even stronger when the DREX is dialyzed. Here, we not only increase SAH concentration but also lower the SAM level by dialysis. The methylation capacity is then at a bare 0.00005 (0.024 μM /500 μM), thus explaining that blocking of methylation is even more efficient in dialyzed DREX.

These experiments are an indication that the metabolic steps of the SAH competitive feedback loop for H4K20monomethylation are functional in DREX, pointing towards not only present but also active metabolic enzymes, in this specific case Pr-Set7, on the chromatin fiber. While this enzymatic activity and inhibition has been reported in the literature, the embedding in this context as well as the acquirement of more insight into the underlying metabolite dynamics guiding the setting of the H4K20me1 mark are novel.

4.3.6.3 The extract metabolizes isotopically labeled methionine to SAM and isotopically labeled histone methylation marks

Our reported incorporation of the heavily labeled methyl group of a supplied methionine as histone methylation is direct proof of complex metabolic activity in DREX (Figure 35). The detection of an isotopically labeled mono-methylated histone peptide proves that in the extract, at least two more steps of the methionine cycle are happening. The added isotopically labeled methionine can be converted into SAM by SAM synthetase (SAM-S) and then this SAM is used to methylate nucleosomes incorporated into the chromatin fiber, in this case H4K20me1 by PR-Set7. Intriguingly, we detect the enzymes SAM-S and Pr-Set7 not only in the DREX proteome but also enriched on the chromatin fiber compared to a beads-only control (Figure 36). Taken together, these findings lead us to hypothesize that the conversion of metabolites might happen in situ on the fiber, creating a local microenvironment favoring methylation.

4.3.6.4 Like in vivo, excess of methionine does not translate to an increase of H4K20me1 in DREX

In the context of our studies on the incorporation of heavily labeled methionine building blocks as histone methylations, we also explore the effect of an excess of methionine. Interestingly, the addition of high concentrations of isotopically labeled methionine did not alter the methylation levels compared to the addition of physiological levels. This is true for the overall percentage of H4K20me1 as well as for the ratio of heavy to light methylation. This result reflects findings by others that report a similar failure to increase methylation levels above a certain threshold in human cell lines, even with high levels of methionine supplied in the medium (215). Additionally, they discovered that SAM concentration is not increased through this methionine oversupply. This suggests that SAM synthesis from methionine is most likely the metabolic bottleneck for this pathway. To test this hypothesis in DREX, a future experiment with the supply and tracking of different concentrations of isotopically labeled SAM instead of methionine during chromatin assembly is recommended. Taken together, these findings support the hypothesis of SAM-S creating a metabolite bottleneck and provide an example of how the DREX in vitro system successfully mimics in vivo situations.

4.3.6.5 Active metabolism replenishes dialysis-depleted metabolites in DREX

In this study, we showed that metabolites that have been removed from DREX via dialysis recover at 26 °C, independent of the presence of DNA but dependent on ATP availability (Figure 37 and Figure 38). This is only possible because the enzymes for the metabolic pathways necessary to produce these metabolites are present and active in DREX. Here, we looked specifically at the metabolites involved in the methionine cycle.

Interestingly, methionine showed a trend for dependence on ATP but also recovered to a slower yet substantial level without it after dialysis (Figure 38). Methionine is an essential amino acid in *Drosophila* and can be obtained by the degradation of methionine-containing proteins. The regeneration by remethylation from homocysteine in *Drosophila* has been suggested in a knockdown (103). The putative betaine-dependent methionine synthetase CG10623 is tremble annotated in the Uniprot database but had not been at the time of analysis of the dataset, not allowing us to gain any insight on the protein's presence and whereabouts in our system (104). While more data on the existence and role of potential regeneration pathways are necessary to fully exclude its role in replenishing the methionine pool, we believe that the generation by degeneration of methionine-containing proteins is the main source. We hypothesize that methionine is replenished by the degradation of methionine-containing proteins by the proteasome and subsequently proteases. We have shown before that the subunits of the proteasome core particle as well as several regulatory particles are present on the chromatin fiber and active in DREX. Thus, a quick generation of a metabolite-rich environment is made possible in proximity to the chromatin fiber. Proteasomal degradation can happen in ATP-dependent and ATP-independent manner, explaining the replenishing of methionine without ATP and a faster recovery when ATP is available. Inhibiting the proteasome and/or exopeptidases and then tracking the recovery of methionine could provide valuable direct insight into the dependence of methionine recovery on the proteasome.

Contrastingly, SAM shows a full dependency on the nucleotide ATP (Figure 38). This is to be expected, as ATP is a building block of SAM, which is synthesized from methionine and ATP in an enzymatic

reaction that is catalyzed by SAM-Synthetase. Thus, we also see a delay in recovery compared to methionine due to the same dependency on the availability of its second building block, methionine.

Curiously, SAH concentration does not change over time and stays consistent after dialysis (Figure 38). The concentration it stays consistent at, however, depends on ATP presence. When no ATP is supplied, the concentration simply stays constant at where it was after dialysis. SAH concentration is very low when ATP is added to the extract, although its degradation by AHCY is dependent on NADH, not ATP. Taking our hypothesis of the inability to fully dialyze SAH out because its pool is partially bound to an enzyme, we suggest it stays bound when no ATP is present but requires ATP to be released and subsequently degraded by AHCY. If the release is ATP-dependent, it is tempting to argue that SAH release from e.g. Pr-Set7 is dependent on SAM levels, which in turn depend on ATP. We have established that the methylation capacity (SAM/SAH concentration ratio) is relevant for the processivity of the PR-Set7 enzyme. This suggests that SAM, at a certain excess threshold, displaces SAH in the Pr-set7 binding pocket. One way this hypothesis can be further explored was applied in this project by testing the effect of an excess of SAH on the binding of the enzymes on the chromatin fiber. Further possibilities, including in vitro fluorescence-based binding dynamic assays with purified proteins and isolated metabolites on salt gradient assembled chromatin might be explored in the future.

Overall, we observed a recovery of metabolites to a maximum of 20% of their level in DREX before any manipulation after 120 min. To investigate whether the metabolites recover further and how long the ATP regenerating system provides energy, an investigation of the phenomenon for an extended period with a minimum of 4h, the time needed for assembly of mature chromatin in DREX, will be conducted in the future.

Finally, we would like to remark that our results suggest experimenters need to keep in mind that after dialysis, samples of DREX need to be stored at low temperatures and brought to experimental temperature on point depending on the nature of the experiment. We have shown that dialyzed DREX is highly dynamic. The metabolite concentration changes over time at 26 °C should be factored into any experimental setup.

4.3.7 SAH excess traps Pr-Set7 fiber and effectively inhibits H4K20me1

We investigated the effect of a block or inhibition of the directly and indirectly observed and shown metabolism, leveraging the versatility of our DREX system. Therefore, we used the competitive inhibitor SAH that had been shown to affect H4K20me1 in our system before (5). In our study, we were able to generate even more complex and clean readouts for the chromatin-bound proteome as well as a histone modification analysis than ever before.

We observed an inhibition in histone H4K20 monomethylation by Pr-Set7 after 1h of chromatin assembly in DREX as previously reported ((5) and Figure 34). Additionally, we were able to quantify the histone H4K20me1 levels in this experiment for the first time via mass spectrometry analysis. In our analysis, we observed a slightly weaker inhibition of the methylation reaction than previously reported by western blot. This difference might stem from the difference between the detection methods, antibody specificity, its limit in quantification accuracy, and the overall variability of the system. The trend of inhibition of H4K20monomethylation by SAH, however, is consistent. In our study, we added a new layer of complexity by also quantifying the inhibition of the H4K20monomethylation upon treatment with SAH in dialyzed and thereby metabolite-depleted

DREX. Here, we observed the inhibition to be much stronger most likely due to a lower methylation capacity in the DREX (4.3.6.2).

When analyzing the chromatin-bound proteome in DREX upon SAH treatment, we observed an enrichment of different histone methyltransferases on the fiber compared to the untreated control (Figure 36). Interestingly, these methyltransferases catalyze different histone methylations. The *Drosophila* orthologue of G9a, which catalyzes H3K9 mono- and dimethylation in vivo (216) and shows indication to methylate H4K12 or 16 in vitro (217). The enzymes extra sex combs (esc) and enhancer of zeste E(z) form the Esc/E(z) complex, which methylates H3K9 and H3K27, especially in early *Drosophila* embryos to create polycomb islands (218-221). Mes-4, a protein belonging to the NSD (nuclear receptor SET domain) methyltransferases (222) mediates H3K36 mono- and demethylation (223). However, despite the presence of these enzymes on the chromatin fiber, we do not detect these histone methylations on chromatin assembled in the DREX in vitro system. The only histone methylation detected at a proportion above 1% on DREX-assembled chromatin is H4K20me1, catalyzed by Pr-Set7, which is also among the methyltransferases we detect as enriched on the chromatin fiber upon SAH treatment.

In the context of our study of the link of methionine metabolism to H4k20me1, the finding that Pr-Set7 is stuck on the chromatin fiber upon treatment with SAH provided us with valuable new mechanistic insights. Based on this result, we hypothesize that Pr-Set7 cannot release from its product, the methylated histone when the SAM/SAH is too low. Unpublished results from the Kurumizaka lab confirm that structures of complexes of the Pr-Set7 orthologue SET8 with nucleosomes on DNA are more stable when SAH is present, as determined by EMSA. The binding of the methyl transferase to nucleosomes on DNA also happens without SAH. Therefore, they suggest that the acidic patch-binding mode of the methyl transferase comprises the main binding mode in vitro, while the SET domain H4 tail interaction might serve as an auxiliary binding mode to increase the binding to the nucleosome (Personal correspondence, Hitoshi Kurumizaka). Overall, this validates our observation that the methyl transferase is bound more or longer to chromatin in the presence of high concentrations of SAH.

The release of the methyltransferase from the nucleosome most likely has one of two underlying mechanisms. Possibly, the enzyme sits on the nucleosome on the chromatin fiber after methylation with SAH bound until SAM replaces SAH in the co-substrate binding pocket and thereby releases the enzyme from the fiber. This hypothesis is supported by the correlation of the amount of Pr-Set7 bound to chromatin to the SAM/SAH ratio. Alternatively, the cooperation of Pr-Set7 and AHCY to remove the methyltransferase from the fiber is thinkable. Supporting this stream of thought, we observe the opposite effect of SAH excess on the protein AHCY compared to how Pr-Set7 was affected. The enzyme AHCY disappears from the chromatin fiber when excess SAH is present, hypothetically because, under high SAH conditions, it binds to free-floating SAH instead of processing SAH from the histone methylation reaction. A direct handover of the SAH molecule from Pr-Set7 to AHCY upon completion of the methylation for processing is thinkable. This could be tested by investigating the direct binding of the two enzymes, e.g. by FRET.

When comparing chromatin assembly in the presence of SAH in dialyzed vs. non-dialyzed DREX, we find a much stronger effect of SAH in the experiments where DREX was dialyzed before assembly to remove all metabolites. This is true for the chromatome analysis as well as for the histone modification analysis (Figure 34 and Figure 36). Because of the addition of a very high concentration of SAH, the

methylation potential is very low at 0.002 in the non-dialyzed sample. However, in the dialyzed sample, it is even lower at 1.5×10^{-5} , explaining the even stronger effect (4.3.6.2).

In the future, the effect of inhibition of AHCY by Dz-nep on the chromatin proteome could be investigated (224). This might give further insight into the interplay of AHCY and Pr-Set7, the influence of SAH on the histone methylation percentage, and a chance for differentiation of the effect of different players on the on-off rates of Pr-Set7 on nucleosomes on DNA. Furthermore, the role of Pr-Set7 as a methylation potential sensor could be investigated. We recommend expanding these studies by in vitro experiments with recombinant proteins and SGD experiments. SGD-based titration studies or a combination of the recombinant in vitro system with FRET or surface immobilization of one of the substrates would allow the generation of quantifiable data on enzymatic turnover and on/off rates.

Overall, we observe that SAH leads to a reduction in H4K20monomethylation and traps Pr-Set7 and other methyltransferases on the chromatin fiber. We hypothesize the processivity of the Pr-Set7 enzyme goes down as it either cannot be unloaded from the fiber anymore due to metabolite ratio or hindered interaction with the enzyme AHCY. Dialysis of small metabolites enhances the effect, indicating a high dependence of Pr-Set7 on the methylation potential. We propose further studies with inhibitors in the DREX system but also in a recombinant system for additional mechanistic insights.

4.3.8 Lack of diversity in histone methylations in DREX assembly

In our project, we validated that in the in vitro system of DREX chromatin assembly, H4K20monomethylation is by far the most abundant histone methylation modification. All other histone methylations are detected below 1% abundance (analysis in cooperation with Beyza Bozdağ, data not shown). There are multiple hypotheses, yet only limited data on the reason(s) for this phenomenon. This can be interrogated further in the future utilizing and expanding the methods, datasets, and results produced in this project.

The first idea of a lack of abundance or recruitment of other methyltransferases on the fiber could be rejected based on our data. In our proteome, we find not only Pr-Set7 but also other methyltransferases, including G9a, Mes4, and the esc/E(z) complex bound specifically and high abundantly to the chromatin fiber (Figure 36).

A low concentration of available metabolites is another try at explaining the lack of more diverse methylation in the in vitro system. This is unlikely due to 3 pieces of evidence: First., Pr-set7 has the highest K_m value of the detected histone methyltransferases and should thereby be the most affected by a lack of SAM or a low SAM/SAH ratio (225). Second, we benchmarked the availability of SAM and SAH in DREX. We determined by metabolomics analysis that the concentration and ratio are high enough based on the enzymes' K_m values and additionally similar to those observed in cells where more histone methylation variation is set (4.3.5). The metabolite concentrations most likely reflect the free pool of the metabolites of this pathway within a 12-hour time window. The exception is the SAH pool, which we believe to be 50% bound by proteins due to our dialysis experiments. Lastly, dialysis to remove metabolites before assembly did not lead to an abolishment of the H4K20me1 mark. This suggests that the amount of metabolites left was high enough, at least locally, to support the methylation reaction. Lastly, we do not believe that metabolite availability restricts methylation mark diversity because we observe supplementation of additional isotopically labeled methionine over the

physiological level does not recover other histone methylations either (not shown, data analysis in cooperation with Beyza Bozdağ). However, the limitation of this third point lies in the question of whether the SAM-S creates a bottleneck by enzymatic rate. This could easily be examined by the addition of isotopically labeled SAM and tracking the methyl group in the DREX chromatin assembly reaction. Overall, we believe that there is no metabolite shortage in DREX that would explain the setting of only one type of histone methylation.

DNA sequence dependency represents another option for the lack of histone modification diversity in our DREX assemblies. In this study, only one DNA sequence based on a nucleosome positioning sequence repeat has been used for chromatin assemblies. To investigate sequence effects in the future, we recommend DREX studies using genomic DNA or different DNA sequences for chromatin assemblies. Whole genome assemblies can be used to track HMT placement by ChIP-Seq (33,34), while comparative analysis of histone modifications on assemblies with different defined sequences might reveal the influence of different DNA features on methylation efficiency. Proteomic studies on different DNA sequences during DREX chromatin assembly revealed no difference in the chromatin-bound proteome between different DNA sequences in the past (28).

Furthermore, the influence of inhibitors and activators might shape the activity of histone modification enzymes. These activators and inhibitors can come in 3 different forms; they might be other proteins, pre-existing histone marks, or small molecules: First, more often than not, histone marks are set by enzymes that are influenced in their activity by other enzymes. The HMT Set1, for example, is part of a bigger complex, interacting with the PAF complex and RNA polymerase II. Additionally, the recruitment is dependent on many transcription factors and transcriptional co-activators (226). The complexity of the mechanism of recruitment and activity of the HMTs ensures a high level of organization and control in the setting of marks and allows for additional crosstalk with the cellular environment. In our system, the effect of other enzymes on the HMT activity could be tested by analyzing histone modifications on DREX-assembled chromatin upon removal of candidate interactors from the DREX using PROTACs or the TRIM away system (197,227). Secondly, some HMTs, like PRC2, which sets H3K27trimethylation, require activating pre-deposited H3K27methylation to bind and spread the modification (228). In vitro HMT assays, based on recombinant naïve oligonucleosomes and recombinant PRC2, show that only the addition of a stimulatory H3K27me3 peptide stimulates H3K27me3 on the nucleotides (229). The addition of the stimulating H3K27me3 peptide to the DREX assembly might increase this mark and will generate valuable insights on whether allosteric activation of HMTs is lacking at least. Ultimately, the inhibitor or activator could also be other small molecules of yet unknown nature, including metabolites or cofactors. Because we see a different methylation type diversity in embryos, it is tempting to hypothesize that a common activator might be lost in the preparation of DREX from embryos. This factor can be imagined to be of different natures, including salts, buffer systems, lipids, insoluble proteins, and other factors separated off or diluted out during DREX preparation. This broader idea of a yet unknown player is beyond the scope of the project but sets an example of the complexity yet to be discovered in the regulation and dynamics of histone modification in the extract and beyond. To sum it up, all three domains, interacting proteins, histone marks, and small molecules represent feasible regulators that have not yet been but should easily be explored in this context in the DREX system in the future.

Finally, another intriguing perspective is not simply the differential and specific inhibition or activation of certain HMTs by different factors but the idea of a master regulator. In the extract, no histone

methyl modifications except the H4K20me1 were detected. This might be caused by an overall protection of histone tails in DREX by an unknown agent or mechanism. In this line of thought, the H4K20me1 modification would represent an escapee of sorts. Pr-Set7 might then be the only methyltransferase able to circumvent or overwrite this signal and set the H4K20me1 modification. The literature points towards an activated repair signaling pathway in the DREX-assisted chromatin assembly when performed with a linearized DNA template, as done in the histone modification studies in this project (35). This signal could represent the foundation of the mechanism for the H4K20me1 modification to specifically overcome the general inhibition. The extract is lacking one major feature that embryos have, cell cycle. Despite its reduction to S and M phases in the preblastoderm embryo, many factors still cycle every round. Much like the sensing of nuclear to cytoplasmic ratio in *Drosophila* embryo triggers ZGA, it is conceivable that a mayor signal that is dependent on cell cycle factors regulates the modifications (230). This makes the factors that do cycle in embryos interesting candidates for investigations.

The current study provides new insights and methods to investigate the mystery of the lack of diversity of histone modifications in the DREX. We believe that further investigation of the different dimensions and elucidation of the mechanisms that shape the modification landscape in vitro will give valuable insights into the complex system of histone modification regulation in the embryo and beyond in vivo.

4.3.9 A model for metabolic coupling of H4K20me1 during chromatin assembly

Finally, we have come up with a working model for the metabolic coupling of H4K20me1 during chromatin assembly, exploiting the various methods established and benchmarked in this project in DREX. The many indications or direct proofs combined allow us to paint a picture of how a methylation-favoring environment is created on the chromatin fiber.

Our working hypothesis starts with the proteasome, which we detect enriched on the chromatin fiber via proteomic analysis (Figure 26). The proteasome recognizes and degrades marked-for-degradation proteins on the chromatin fiber. The resulting peptides are digested into amino acids. Based on our metabolomics results, we believe that this degradation happens in ATP dependent and independent manner (Figure 38). The presence of the exopeptidases and their activity should be further investigated in the future. The essential metabolite, methionine, is putatively generated directly on the fiber to attain a high local concentration and thereby a methylation-favoring environment. Here, we recommend a metabolomics analysis of an isolated chromatin-assembled-in-DREX fraction compared to a general DREX fraction in the future to validate and quantify the enrichment directly. Additionally, we detect all other relevant proteins of the methionine cycle and the necessary histone methyl transferase (HMT) Pr-Set7 on the chromatin fiber (Figure 36). The presence of these proteins not only in DREX (Figure 23) but specifically enriched on the fiber is a further implication of a microenvironment for metabolic coupling. We have shown directly that methionine can be converted into SAM in the extract and subsequently used as a cofactor by the HMT to methylate histones (Figure 35). Pr-Set7 is then released from the fiber again when new SAM binds its catalytic center or via interaction with the AHCY, the lyase of its product SAH. The mechanics of this process have yet to be dissected e.g. by protein-protein binding assays and nucleosome-protein binding assays in vitro under different metabolite concentrations. Overall, we believe the evidence collected points towards a model where metabolites

as well as enzymatically active proteins create a local metabolite-rich environment that is generated, favoring histone methylation.

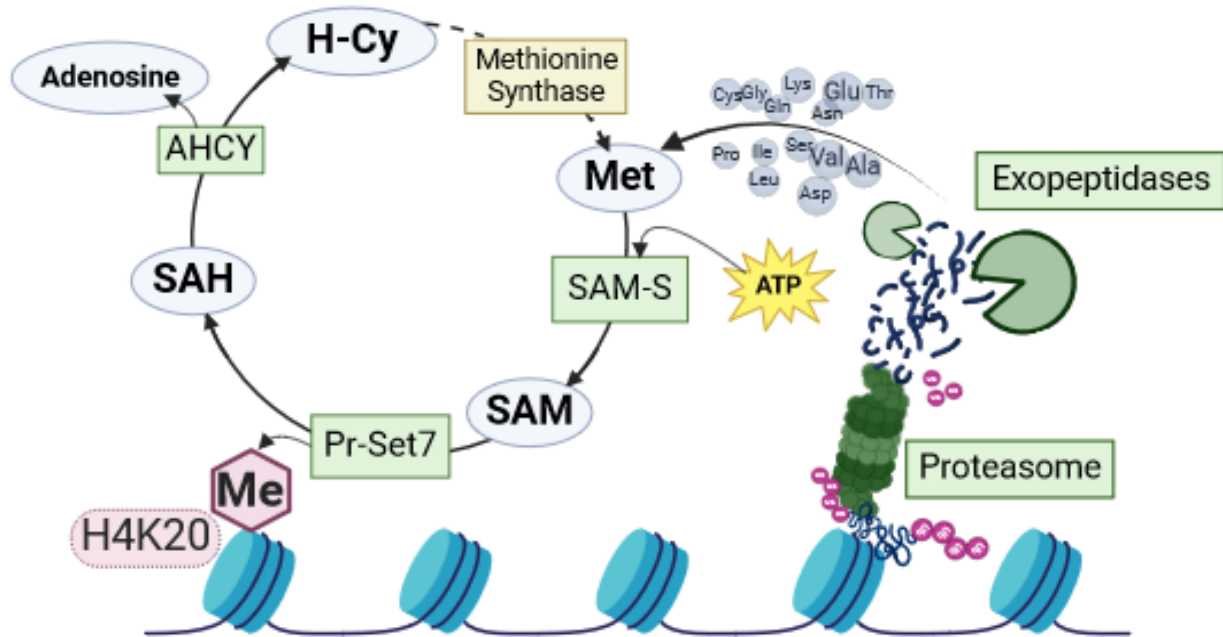


Figure 39 Scheme of working hypothesis for metabolic coupling of H4K20me1 in DREX

To depict the hypothesis, enzymes are depicted in green, metabolites in blue. In this hypothesis, proteins on the chromatin fiber get degraded by the proteasome and resulting peptides are cleaved into amino acids by exopeptidases. These amino acids are metabolites and can be used by the proteins present in the abstract. Here, methionine is used, together with ATP, by SAM-S to produce SAM. This SAM is then used as a cofactor by Pr-Set7 to methylate H4K20 lysine. The resulting SAH is cleaved by AHCY into adenosine and homocysteine. Homocysteine might be converted into methionine synthase, however this link is not established so the enzyme is depicted in yellow and pathways in dotted lined only. Created with Biorender.com.

5 Final Conclusions

In this thesis, the characterization DREX system was advanced, and new insights into the process of chromatin assembly were won by challenging the system.

For the first time, a comprehensive metabolomics and proteomic analysis of the extract was performed. Valuable insight into which proteins are present and at what relative abundance in the extract was gained. It was shown that for fundamental metabolites, their levels are similar to in vivo levels and that metabolic processes in DREX are functional and working, establishing it as a new method to investigate metabolic processes in complex environments but without a living organism.

We report that multiple learnings from the experiments including novel insights on the mimic molecule foldamer but, most importantly, a more thorough understanding of the dynamics of chromatin structure and function that we unearthed by disturbance of the system using foldamer. This study shows the impact of general DNA shape recognition on binding events on chromatin. By leveraging the biologically orthogonal compound “foldamer”, different aspects of binding events could be deconvoluted, revealing that binding, even that of sequence-specific binders, have at least a partial contribution of DNA shape recognition binding to their total binding capacity. Furthermore, this study recommends DREX as a future model to study metabolism and its link to chromatin, elucidating fundamentally required knowledge for usage of the system while already elucidating new key aspects of a specific process, the connection of the methionine cycle and H4K20monomethylation.

Finally, our experiments show that the findings made in the DREX in vitro system were highly transferable to in vivo studies but also following mechanistic in vitro findings. Titration experiments that would have been impossible in an in vivo setting due to systemic effects were elegantly leveraged here while exploiting the DREX system’s complexity to also include indirect effects caused by crowding, known and unknown competitors. This study expands the knowledge and tools surrounding the DREX chromatin assembly system while delivering key insights into the mechanisms of protein binding to chromatin and the crosstalk between metabolism and chromatin structure. Ultimately, this represents a proof of concept of the use of the system and its versatile outputs to elucidate mechanisms involving chromatin, chromatin assembly, and metabolism.

6 Materials and Methods

6.1 Materials

6.1.1 Bacteria strains

Name	Supplier	Remarks
DH5alpha competent cells	Thermo Fisher Scientific	High insert stability due to recA1 mutation High yield and quality of DNA due to endA mutation High transformation efficiency

6.1.2 Chemicals, peptides, and recombinant proteins

DNA mimics/Foldamers

“Foldamers” and biotinylated “foldamers” were obtained from Valentina Corvaglia from the Ivan Huc laboratory Ludwig-Maximilians-Universität München, Butenandtstraße 5-13, D-81377 Munich – Germany. Corresponding synthesis experimental numbers of the compounds used in this thesis were VC391 and VC392.

Antibiotics

Name	Concentration of stock solution	Working concentration	Supplier
Ampicillin	100 mg/ml in H ₂ O	100 µg/ml	Roth

Spike-ins for metabolomics

Name	Stock concentration	Supplier, Cat. Nr.
L-Methionine-(methyl-13C, d3)	6.53 mM in ms grade H ₂ O	Merck, 299154
Adenosine-15N5 5'-Triphosphate	100 mM (in 5mM Tris HCl / H ₂ O), ≥98 atom % 15N, ≥95% (CP)	Merck, 707783
L-Lysin-13C6,15N2 – hydrochloride	32.4 mM in ms grade H ₂ O	Merck, 608041

Protease and proteasome inhibitors, reducing agents, isotopically labeled metabolites, and other modulators

Name	Stock concentration	Standard dilution	Supplier
Aprotinin	1 mg/ml in ddH ₂ O	1:1000	Genaxxon
Leupeptin	1 mg/ml in ddH ₂ O	1:1000	Genaxxon

Pepstatin	0.7 mg/ml in ethanol	1:1000	Genaxxon
PMSF (Phenylmethylsulfonylfluorid)	0.2 M in Isopropanol	1:1000	Sigma
DTT (Dithiothreitol)	1 M ddH ₂ O	1:1000	Roth
Epoxomicin	10 mM in DMSO	According to protocol	ApexBio
NEM (N-ethylmaleimide)	100 mM in ethanol	1:20	
MG132	4 mM in DMSO	1:400	MedChemExpress (Kind gift from the Meiners group)
SAH (S-adenosyl homocysteine)	26 mM in ddH ₂ O	1:20	Merck (Kind gift from Daan Verhagen)

6.1.3 DNA: Plasmids and oligos

Plasmids

Plasmid	Insert	Application	Remark
pAI61 (pBS SK (-))	13 repeats of <i>L. variegatus</i> 5S rRNA	Bacterial expression	As published in (28)

Customized oligos

Name	Sequence	Restriction site added	Supplier	Application
16bp_1_mimic_ctrl	5'-ATCTAGATCGAGCTACA-3'	XbaI	Sigma Aldrich	Ligation with complimentary oligo for interference experiment (Chapter 3.2.1)
16bp_2_mimic_ctrl	5'-TGAGCTCGATCTAGAT-3'	XbaI	Sigma Aldrich	Ligation with complimentary oligo for interference experiment (Chapter 3.2.1)

6.1.4 Software and algorithms

Affinity Designer - software version 1.10.4.1198 was used to assemble figures.

Affinity photo - software version 1.10.4.1198 was used to visualize, crop, and adjust DNA agarose gels documented with the Chemidoc Imaging Touch system (Biorad). Only non-distortive adjustments were made using brightness and contrast.

Biorender.com - was used to generate schemes where indicated.

Cytoscape - software version 3.8.2 and 3.9.1 with a string database plug-in version 11.5 was used for protein clustering and visualization of protein networks. Parameters specified in figure legends were used.

DeepL.com web application - was used to facilitate translation of “Summary” of this thesis in english to “Zusammenfassung” in german.

GraphPad Prism 5 - was used for the analysis and visualization of data from Proteasome-Glow Assay.

MaxQuant - was used for processing LC- mass spectrometry raw data.

Microsoft Excel 2016 – was used during metabolite and histone modification analysis after skyline, and for visualization.

Microsoft Word 2016 – was used to compile text documents, including this thesis.

Microsoft Powerpoint 2016 - was used for visualization of proteasomal subunits.

Perseus – software version versions 1.6.7.0. and 1.6.15.0 were used for filtering, analysis, and visualization of data in the form of volcano plots, heatmaps, scatterplots, and intensity distribution bar plots.

R Studio – software version 4.0.3 was used for GO term analysis, simple normalizations and calculation of p.adjust values, and visualization of data in the form of box plots, bar plots, violin plots, network plots, and waterfall plots.

Skyline - version 21.10.1.146 and 2.22 were used for analysis of metabolites and histone modifications and visualization of integration of isotopically labeled modifications on histones.

Sunrise with Magellan 7.2 - software by Tecan Group Ltd., Männedorf; Switzerland Luminescence was used for data acquisition from the plate reader.

6.1.5 Standard buffers

Luria-Bertani (LB) medium

1.0% (w/v) Bacto-Tryptone
1.0%(w/v) NaCl
0.5% (w/v) Bacto-Yeast extract
in deionized H₂O

EX100 buffer

10 mM Hepes pH 7.6
100 mM NaCl
1.5 mM MgCL₂
0.5 mM EGTA
10% (v/v) glycerol
10 mM DTT (add fresh)
0.2 mM PMSF (add fresh)

1x TE buffer

10 mM Tris pH 8.0
1 mM EDTA
in deionized H₂O

1x TBE buffer

90 mM Tris base
90 mM Boric acid
2 mM EDTA (pH 8.0)
in deionized H₂O

1x PBS buffer

137 mM NaCl
2.7 mM KCl
10 mM Na₂HPO₄
1.8 mM KH₂PO₄
in deionized H₂O
adjust pH to 7.4
for cell culture purposes, the buffer was sterilized
by autoclaving

6.2 Molecular biology methods

6.2.1 E. Coli DNA transformation

100 µl competent cells thawed on ice and 5 µl plasmid DNA were added. The bacteria suspension was incubated for 20 min and then heatshocked at 42 °C for 90 s, then incubated on ice for 2 min. Next, 900 µl Luria-Bertani (LB) medium was added and cells were incubated in a shaking incubator at 750 rpm at 37 °C for 1 h. Then, cells were centrifuged at room temperature at 800 g for 3-5 min. The supernatant was removed and the pellet was resuspended in 100 µl LB medium. Since all plasmids used carried an ampicillin (amp) resistance gene for selection, the 100 µl bacteria suspension was then plated on LB-Amp agar plates (amp 100 µg/ml) and incubated at 37 °C overnight.

6.2.2 Growing of bacteria and plasmid DNA purification

Single colonies were picked from overnight plates from E.coli transformation and inoculated into 5 ml/500 ml liquid LB medium containing the resistance gene in the plasmid: ampicillin (100 µg/ml) for miniprep and maxiprep, respectively. These bacteria suspensions were incubated in a shaking incubator at 140 rpm at 37 °C overnight. The following isolation of plasmid DNA was performed using the Macherey-Nagel Kits according to the manufacturer's protocols. DNA was reconstituted in 1xTE buffer and stored at -20 °C. DNA concentration was determined by a Spectrophotometer (Ds-11, DeNovix, Wilmington, USA).

6.2.3 DNA linearization and biotinylation

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

Biotinylation of DNA was performed as previously described (28). In short, 500 µg of the pAI61 plasmid was linearized by SacI and XbaI digestion and DNA was precipitated. Restriction enzymes were purchased from New England Biolabs (NEB) and conditions were chosen according to manufacturer's protocol. Subsequently, one end of the DNA was biotinylated by incubation of the linearized DNA with 80 mM dCTP and dGTP, 3 mM biotinylated dUTP and dATP as well as the Klenow Polymerase. The biotinylated DNA was then purified using G50 Sepharose columns (Roche) according to the manufacturer's protocol. Finally, DNA concentration was determined by Spectrophotometer (Ds-11, DeNovix, Wilmington, USA) and adjusted to 200 ng/µl in 1xTE buffer and stored at -20 °C.

6.2.4 DNA precipitation

In all cases, whether after restriction or MNase digests, DNA was precipitated with a final concentration of 1 M ammonium acetate and 1.5 volumes of ethanol. After incubation at -20 °C for 2°C at full speed (21130 xg) for 1 h. The supernatant was removed, samples washed with 70% ethanol, and centrifuged at °C at full speed (21130 xg) for 30 min. The supernatant was removed with a pipet and samples were left at room temperature for the remaining ethanol to evaporate. Finally, DNA was resuspended in 1x TE buffer.

6.2.5 Agarose gel electrophoresis

Agarose gel electrophoresis was performed to separate DNA fragments by size and confirmation to analyze results from restriction digests, binding assays, and MNase digests. Agarose percentage was chosen based on DNA fragment size to analyze, 0.8% was chosen for restriction reaction testing, and 1.5% Agarose was chosen for gels for Mnase digest reactions. Agarose powder was weighed and dissolved in corresponding volumes of 1x TBE buffer by heating in the microwave until the solution was clear without any remaining visible particles. The solution was poured into the gel tray and left to gel with combs in place to mold sample loading wells. When the gel was stable, the chamber was filled with 1x TBE buffer. DNA samples and DNA ladders were mixed with Midori Green Direct (Nippon Genetics) at a 1:10 dye: sample ratio. Samples were loaded into the sample loading wells in the gel, DNA ladders in a separate lane as a size standard. Electrophoreses were performed at 50-150V until separation was sufficient to determine necessary distinctions. DNA on the gels was visualized and analyzed by radiation, with UV light, and documented using the Chemidoc Imaging Touch System (Biorad).

6.2.6 Annealing of short DNA oligos

Each oligo was dissolved in freshly prepared 1x annealing buffer (10 mM Tris-HCl pH 7.8, 50 mM NaCl, 1 mM EDTA) separately, leading to a concentration of 10 µg/µl (2 mM) per oligo. Then, identical volumes of each of the solutions of complimentary oligos were mixed (ratio 1:1) and absorption at 260 nm (dsDNA) was measured by Spectrophotometer (Ds-11, DeNovix, Wilmington, USA). Then, the mixture was incubated at 95 °C in a shaking heat block at 500 rpm for 5 min. Afterward, the sample remained in the heat block that was gradually cooled down to 16 °C over 42 min. Next, the sample was transferred to 4 °C and cooled for another 20 min. Subsequent measurement of the sample's absorption at 260 nm confirmed that the raw absorption value at this wavelength increased, showing successful annealing. Finally, the DNA was precipitated and taken up in 1x TE buffer.

6.2.7 Preparation of preblastoderm *Drosophila* embryo extract [DREX]

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

Preblastoderm Drosophila embryo extract (DREX) was prepared as previously described (28), with minor adjustments. In short, Drosophila melanogaster embryos were collected 0–90 min after egg laying and subsequently dechorionated using 3% hypochlorite. The dechorionated embryos were washed in 0.7% NaCl, resuspended in extract buffer (10 mM HEPES [pH 7.6], 10 mM KCl, 1.5 mM MgCl₂, 0.5 mM EGTA, 10% glycerol, 10 mM 3-glycero-phosphate; 1 mM dithiothreitol [DTT], and 0.2 mM phenylmethylsulfonyl fluoride [PMSF], added freshly) at 4 °C and homogenized using a tight pestle connected to a drill press. The homogenate was supplemented with MgCl₂ to a final MgCl₂ concentration of 5 mM and centrifuged for 10 min at 10,000 rpm in an SS34 rotor (Sorvall, Thermo-Fisher Scientific, Waltham, USA). The supernatant was centrifuged again for 2 h at 45,000 rpm at 4 °C in a SW 56 rotor (Beckman-Coulter, Germany). The clear extract was isolated with a syringe, avoiding the top layer of lipids. Extract aliquots were frozen in liquid nitrogen. Protein concentration was determined by Spectrophotometer (Ds-11, DeNovix, Wilmington, USA) measurement and titration with chromatin assembly experiments.

6.3 Biochemical methods

6.3.1 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

This method is used to separate proteins according to their electrophoretic mobility in an electrical field. First, the sample is heated to 95 °C in Laemmli buffer (200 mM Tris [pH 6.8], 8% (w/v) SDS, 40%(v/v) glycerol, 0.2% (w/v) bromphenol blue) to denature the proteins and bind anions to the hydrophobic sites of the proteins to add size dependent negative charge. Next, proteins were loaded into loading pockets of gels with an acrylamide gradient of 4-20% (Serva) while Protein markers (peqGOLD Protein Marker IV and V by Peqlab) were used as molecular weight standards. Then the gel was subjected to Electrophoresis at 24mA per gel in a chamber filled with SDS running buffer (25 mM Tris, 190 mM glycine, 0.1% (w/v) SDS), smaller proteins migrate faster through the gel mesh than bigger ones. Afterwards, the gel was stained with Coomassie.

6.3.2 Coomassie staining

The polyacrylamide gels were washed with ddH₂O after SDS-PAGE and subsequently incubated fully submerged in InstantBlue solution (Expedeon) for 60min at room temperature. Protein bands were thereby stained blue. Gels were washed once in ddH₂O and then imaged in the Chrmidoc Imaging Touch system (Biorad).

6.3.3 Proteasomal activity assay

The proteasomal activity was determined by measuring chymotrypsin-like activity using the Proteasome-Glo™ Assay (Promega) according to the manufacturer's protocol. 40µl of DREX were used for each sample, and 3 biological samples were prepared for each time point and each condition. ATP generating system McNap (3 mM ATP, 30 mM creatine phosphate, 10 µg creatine kinase/ml, 3 mM MgCl₂, and 1 mM DTT) was added to the DREX. Then, proteasome inhibitors or corresponding solvent controls were added. The addition of inhibitors was set as timepoint 0 min. After 10 min, 50 min, or 6h, the sample was diluted 1:500 in EX100 buffer. Then, 20µl of this sample was mixed with 20 µl reaction buffer (provided with the kit, containing CT-L substrate and defrosted in the dark 30 min ahead of use) in a white flat bottom 96 well plate (Berthold Technologies, Bad Wildbad, Germany). Cleavage of the Suc-LLVY-aminoluciferin by the proteasome releases aminoluciferin which is transformed into a luminescent signal by the luciferase in the reaction buffer. Luminescence was measured and quantified at time points 30 min, 70 min, and 6:20 h after time point 0 min (addition of inhibitors or control solvent) using a TriStar LB 941 plate reader (Berthold Technologies) and the Sunrise with Magellan 7.2 software (Tecan Group Ltd., Männedorf; Switzerland) when the values reached a plateau of the signal. Raw data extraction and analysis by normalization to controls was performed with Excel, and statistics and plotting were performed with GraphPad Prism 5.

6.3.4 Dialysis of DREX

Dialysis of DREX was performed at 4°C overnight with a 2x buffer change against the EX100 buffer including protease inhibitors. 200 µl – 2 ml of DREX were transferred with a pipet into a dialysis tube Spectra/Por®3 Dialysis Membrane standard RC Tubing with a molecular weight cut off (MWCO) of

3.5 kDa cut off (flat width 18 mm, Spectrumlabs, Inc.) which had been incubated in ddH₂O at room temperature for 3 h before. The tube was closed with clips (Spectrum) and placed in a beaker with a magnetic stirrer filled with the cold dialysate of 500 times the volume of the sample. The dialysate buffer was EX100 with the Aprotinin, Leupeptin, Pepstatin, DTT, and PMSF freshly added. The sample was left in the dialysate on the stirrer for 2 h at 4 °C. The dialysate was replaced with fresh dialysate of the same nature and the sample was left in the stirred solution at 4 °C overnight. On the next day, the sample was removed from the tube using a pipet. As a control, a DREX sample of the same volume was transferred to an Eppendorf tube and incubated at 4 °C and slow movement (50 rpm) on a shaking plate overnight. The samples were flash frozen in liquid nitrogen directly after dialysis/incubation and stored at -20 °C short term (up to 2 weeks) or -80 °C long term (up to 6 months).

6.3.5 Protein concentration measurement with BCA

Protein concentration was performed as indicated using either the nanodrop application of the Spectrophotometer (Ds-11, DeNovix, Wilmington, USA) in “ProteinA280” setting, which estimated protein concentration dependent on the absorption of the analyte at 280 nm, or using the PierceTM Microplate BCA Protein Assay Kit according to the manufacturer’s protocol (Thermo Fisher) using Pre-Diluted Protein Assay Standard: Bovine Serum Albumin (BSA) Set (Thermo Scientific, #23208). Blanks/controls were always the corresponding sample buffers.

6.4 In vitro chromatin methods

6.4.1 In vitro chromatin assembly in DREX

Part of the description marked in *italics* is a direct citation from (111), authored by Vera Kleene.

In vitro chromatin assembly was performed in a total volume of 240 μ l, containing 2 μ g of DNA, 80-160 μ l DREX, an ATP regenerating system (3 mM ATP, 30 mM creatine phosphate, 10 μ g creatine kinase/ml, 3 mM MgCl₂, and 1 mM DTT). Additionally, depending on the experiment, different competitors or modulations were added to the standard setup, including the usage of dialyzed DREX instead of standard DREX or the addition of free foldamer, isotopically labeled methionine, SAH, or proteasome inhibitors to the assembly reaction. EX100 buffer as a background buffer to adjust volume. The assembly reaction was incubated in a turning wheel at 26 °C for 1h or 4h.

Assemblies were performed with plasmid DNA but also with beads-immobilized DNA. For experiments with immobilized DNA, 2 μ g DNA *was immobilized on 60 μ l M280 paramagnetic streptavidin beads (Invitrogen) in Dynabash buffer (10 mM Tris-HCl [pH 8], 1 M NaCl, 1 mM EDTA) for 1 h. Beads were blocked with BSA (1.75 g/l) for 30 min in EX100, then washed in EX-NP40 (10 mM Hepes pH 7.6, 1.5 mM MgCl₂, 0.5 mM EGTA, 10% (v/v) glycerol, 0.05% NP-40). Subsequently, beads were resuspended in the standard assembly mix specified above.*

After incubation and *two wash steps with EX200 (10 mM HEPES [pH 7.6], 200 mM NaCl, 1.5 mM MgCl₂, 0.5 mM EGTA, 10% [vol/vol] glycerol; 0.2 mM PMSF, 1 mM DTT, 0.7 μ g/ml Pepstatin, 1 μ g/ml Aprotinin, 1 μ g/ml Leupeptin added fresh), beads were prepared for Proteomic Analysis or Micrococcal Nuclease Digestion.*

6.4.2 Pulldown with biotinylated foldamer in DREX

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

1 μ g foldamer was immobilized on 30 μ l M280 paramagnetic streptavidin beads (Invitrogen) in Dynabash buffer (10 mM Tris-HCl [pH 8], 1 M NaCl, 1 mM EDTA) for 1 h. Beads were blocked with BSA for 30 min (1.75 g/l) in EX100 (10 mM HEPES [pH 7.6], 200 mM NaCl, 1.5 mM MgCl₂, 0.5 mM EGTA, 10% [vol/vol] glycerol), washed in EX-NP40 (10 mM Hepes pH 7.6, 1.5 mM MgCl₂, 0.5 mM EGTA, 10% (v/v) glycerol, 0.05% NP-40) and resuspended in a total volume of 120 μ l containing 40-80 μ l DREX, EX100 buffer, and ATP regenerating system (3 mM ATP, 30 mM creatine phosphate, 10 μ g creatine kinase/ml, 3 mM MgCl₂, and 1 mM DTT). The reaction was incubated at 26 °C for 1 h, then the beads were prepared for Proteomic Analysis.

6.4.3 Chromatin accessibility assay by Micrococcal nuclease digestion

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

Chromatin from 2 μ g circular DNA assembled for 4 h was resuspended in EX100 containing 5 mM CaCl₂ and 100 units/ μ l of MNase (Sigma). After incubation at room temperature for 30 s, 60 s, and 120 s, respectively, a 110 μ l fraction of the digestion was stopped by adding 40 μ l MNase stop solution (100 mM EDTA [pH 8.0]). The DNA was precipitated and separated with a 1.5% agarose gel upon RNase A and proteinase K treatment.

6.5 Cell-based methods

6.5.1 Cell culture

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

Drosophila L2–4 cells (36) were grown in Schneider medium supplemented with 10% fetal calf serum, penicillin, and streptomycin at 26 °C.

6.5.2 Treatment of cells with DNA mimic foldamer

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

Drosophila L2-4 cells were seeded at 1 mio cells/ml in 6 well plates (Sarstedt), in medium (Gibco, Schneider Drosophila medium) with 10 µM foldamer in medium or without foldamer as control and harvested after 48 h.

6.5.3 Subcellular fractionation

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

Harvested cells were prepared by spinning down of cells at 1000 × g for 4 min and washing the pellet twice with 2 ml ice-cold PBS each. Then, the cells were fractionated using the Subcellular Protein Fractionation Kit for Cultured Cells (Thermo) following the manufacturer's instructions. Adjustments were included for Drosophila cell size, treating the pellet of each well of cells as 5µl packed cell volume and adjusting buffer volumes accordingly, while also doubling the amount of enzyme and tripling incubation time in the MNase digest step. All fractions were stored at -20 °C until they were prepared for Proteomic Analysis.

6.5.4 Flow cytometry

This description in *italics* is a direct citation from (111), authored by Vera Kleene. Flow cytometry measurement, gating, quantification, and cell cycle profile visualization were performed by Pardis Khosravani, Flow Cytometry Core Facility.

Drosophila L2-4 cells were seeded at 1 mio cells/ml in 6 well plates (Sarstedt, Ref: 83.3920), in medium (gibco, Schneider Drosophila medium, Ref: 21720-024) with different concentrations of foldamer (0, 0.01 µM, 0.1 µM, 1 µM, 10 µM) and harvested after 4 h, 24 h, or 48 h. Harvested cells were prepared by spinning down of cells at 1000 × g, 4 min and dissolving the pellet in 1 ml PBS. 2.7 ml ice-cold ethanol was added to cell suspension while vortexing. Alcohol-fixed cells were stored stably at 4 °C for up to 1 week. On the day of measurement, ethanol was removed by centrifuging cells at 1000 × g for 4 min at 4 °C. The supernatant was removed, and cells were resuspended in 1 ml PBS + 1% FBS. Then, cells were counted and 0.5×10⁶ cells were taken up in 500 µl FACS buffer (PBS + 1% FBS). Finally, 5 µl 100×RNaseA solution in PBS was added to a final concentration of 20 µg/µl. The suspension was incubated for 15 min at 37 °C, then 50 µl PI stain (10 mg/ml Sigma 1002755458) was added and incubation of 30 min at RT was allowed before measurement. Stained cells were measured using BD LSRFortessa (equipped with 405,488,561,633 nm lasers; BD Bioscience), and FlowJo™ v10.8.1 software was used to analyze data.

6.6 Mass spectrometry methods

6.6.1 Sample preparation for total proteome analysis via LC-MS/MS

Cell lysates or extracts were prepared for LC-MS/MS analysis using the PreOmics iST kit (Preomics). For cell fraction analysis SP3 add-on (Preomics) was added. In all cases, 100 µg of samples were processed according to the manufacturer's protocol. The final sample was eluted in 15 µl sample loading buffer and stored at -20 °C until further processing.

6.6.2 Sample preparation with on-beads digest for Proteome via LC-MS/MS

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

Assembled chromatin and foldamer pulldowns were subjected to mass spectrometry analysis. The beads-bound fraction was separated by a magnet from the supernatant and washed three times with EX100 and 4 times with fresh 50 mM NH₄HCO₃ to remove detergents and unspecific binders. Tryptic digestion was performed on beads by incubation with 100 µl of 10 ng/µL trypsin (Promega, Ref: V511 in 1 M urea 50 mM NH₄HCO₃ for 30 min at 25 °C. Beads were separated by a magnet, the supernatant was transferred into a fresh tube, beads were washed twice with 50 mM NH₄HCO₃, and supernatants pooled into the same tube. The supernatant pool was adjusted to a final concentration of 1 mM DTT by addition of DTT and digestion was completed overnight at 25°C. Next, the tryptic peptide mixture sample was incubated for 30 min in the dark at 25 °C with iodoacetamide at a final concentration of 35 mM to carbamidomethylate sulfhydryl groups of free cysteine. Subsequently, DTT was added to a final concentration of 50 mM and the sample was left to incubate for 10 min at 25 °C. Then, the sample was acidified using trifluoroacetic acid (TFA), followed by desalting using SDB-RPS (Styrenedivinylbenzene - Reversed Phase Sulfonate, 3M Empore) before mass spectrometry analyses, and redissolved in 15µl MS loading buffer (Preomics) and stored at -20 °C until further processing.

6.6.3 Sample preparation for histone modification analysis

Assembled chromatin fractions were separated by SDS-PAGE (4-20% gradient gel, Serva) and stained with Coomassie without methanol (Brilliant blue G-250). Per sample, a single band was excised encompassing the weight range of histones (~15-23kDa). The gel slices were destained by washing with 50% acetonitrile (ACN) in 100 mM NH₄HCO₃ until clear. After two washes with mass spectrometry grade (ms grade) H₂O the gel pieces were dehydrated by treatment with 100% ACN. Next, propionic anhydride (Sigma) was added to the gel pieces, then 100 mM NH₄HCO₃, and finally 1 mM NH₄HCO₃ to a final concentration of 2.5% propionic anhydride. Then, the sample was incubated for 45 min at 37 °C. This step leads to the propionylation of lysine residues, preventing tryptic cleavage at those sites. Afterward, five washes with 100 mM NH₄HCO₃, 5 washes with ms grade H₂O, and subsequent dehydration with 100% ACN washes until gel pieces turned white were performed. Finally, the proteins were digested by incubation of the gel pieces with 200 ng trypsin (Promega) in 100 mM NH₄HCO₃ at 37 °C overnight. Peptides were extracted with 50%ACN 0,25% trifluoroacetic acid (TFA). The peptide solution was then desalted using carbon SDB-RPS tips (Styrenedivinylbenzene - Reversed Phase Sulfonate, 3M Empore), reconstituted in 0.1% formic acid (FA), and stored at 20 °C until further processing.

6.6.4 Sample preparation for metabolomics by liquid-liquid extraction

For extraction of polar metabolites, 50 to 100 μL of the sample were added to ice-cold $\text{CHCl}_3/\text{MeOH}/\text{H}_2\text{O}$ (200 μL /250 μL /350 μL) (ms grade, Sigma), strongly vortexed for 1 min and then rested for 30 min on ice. If spike-ins are used for quantification, they are added to the sample before extraction. After subsequent centrifugation at 14000 $\times$ g at 4 °C for 15 min, 500 μL of the upper polar phase were taken and transferred to Amicon centrifugal filters with 3KDa cutoff which were then centrifuged for 2h at 10.000 $\times$ g at 4 °C (Filters need to be rinsed before use: Add 500 μL of H_2O (ms grade, Sigma) and centrifuge 10min at 14.000 $\times$ g in a waste Amicon vial; flip the filters and spin for 3 min at 1000 $\times$ g). Finally, the collected filtrated samples were dried in the SpeedVac at 45 °C (~1.5-2h), resuspended in 30 μL of H_2O , shaken for 10 min at 500 rpm, and transferred to Nanovial for measurement. This method is based on Chetwynd et al. (231) and Zhang et al. (232).

6.6.5 LC-MS/MS

6.6.5.1 For all proteomes but the whole DREX proteome

This description in *italics* is a direct citation from (111), authored by Vera Kleene.

LC-MS/MS measurements were performed by Ignasi Forné at the Protein Analysis Unit, LMU.

For LC-MS purposes, desalted peptides were injected in an Ultimate 3000 RSLCnano system (Thermo) and separated in a 25 cm analytical column (75 μm ID, 1.6 μm C18, IonOpticks) with a 50 min gradient from 2 to 37% acetonitrile in 0.1% formic acid. The effluent from the HPLC was directly electrosprayed into a Qexactive HF (Thermo) or an Orbitrap Exploris 480 (Thermo) both operated in data-dependent mode to automatically switch between full scan MS and MS/MS acquisition. For Qexactive HF measurements, survey full scan MS spectra (from m/z 375–1600) were acquired with resolution $R = 60000$ at m/z 400 (AGC target of 3×10^6). The 10 most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of 1×10^5 , and fragmented at 27% normalized collision energy. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250 °C; ion selection threshold, 33000 counts. For Orbitrap Exploris 480 measurements, survey full scan MS spectra (from m/z 350 to 1200) were acquired with resolution $R = 60000$ at m/z 400 (AGC target of 3×10^6). The 20 most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of 1×10^5 and fragmented at 30% normalized collision energy. Typical mass spectrometric conditions were as follows: spray voltage, 1.5 kV; heated capillary temperature, 275 °C; ion selection threshold, 33000 counts.

6.6.5.2 For the whole DREX proteome

LC-MS/MS measurements of DREX proteomes were performed by Teresa Barth at the Protein Analysis Unit, LMU.

For LC-MS purposes, desalted peptides were injected in an Evosep One system (Evosep) in 30 samples per day (spd) setting. The effluent was directly electrosprayed with a nanospray ion source (spray voltage, 2 kV) into an Orbitrap Exploris 480 (Thermo) operated in data-dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 375 to 1500) were acquired with resolution $R = 120000$ and a normalized AGC target of 3×10^6 .

(300%) The 20 most intense peptide ions with charge states between 2 and 6 were sequentially isolated to a target value of 1×10^5 (100%) and fragmented at 30% normalized collision energy.

6.6.5.3 For histone modification analysis

LC-MS/MS measurements were performed by Ignasi Forné at the Protein Analysis Unit, LMU.

For LC-MS/MS purposes, desalted peptides were injected in an Ultimate 3000 RSLCnano system (Thermo) and separated in a 25-cm analytical column (75 μ m ID, 1.6 μ m C18, IonOpticks) with a 50-min gradient from 2 to 37% acetonitrile in 0.1% formic acid. The effluent from the HPLC was directly electrosprayed into a QExactive HF (Thermo) operated in data-dependent mode to automatically switch between full scan MS and MS/MS acquisition with the following parameters: survey full scan MS spectra (from m/z 250–900) were acquired with resolution $R=60,000$ at m/z 400 (AGC target of 3×10^6). The 10 most intense peptide ions with charge states between 2 and 3 were sequentially isolated to a target value of 1×10^5 , and fragmented at 27% normalized collision energy. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250°C; ion selection threshold, 33,000 counts.

6.6.6 CESI-MS

CE-MS measurements were performed by Marco Borsó at the Protein Analysis Unit, LMU.

Electrophoretic separation of analytes was carried out using a CESI 8000 (Sciex) equipped with a sheathless OptiMS CESI cartridge (30 μ m ID x 91 cm bare fused silica capillary) maintained at 25 °C coupled to a 6600 TTOF (Sciex) through a NanoSpray III source. Samples were kept in a thermostated tray at 8 °C, and injected hydrodynamically into the capillary by using a pressure of 1 psi for 60sec which roughly corresponded to an injection volume of 10nL. Metabolites were separated into acetic acid 10% (pH 2.2) buffer using a 30kV voltage in normal polarity for 24min. Between injections, the capillary was rinsed with 0.1N NaOH (ms grade, Sigma) and 0.1N HCl (ms grade, Sigma) at 100psi for 2min each followed by H₂O (ms grade, Sigma) at 100psi for 2min and finally by BGE at 100psi for 3min. The optimal position of the porous tip of the capillary concerning the MS inlet was achieved by moving the XYZ stage to get a stable electrospray (ESI) and the highest total ion current (TIC) signal. The values for gas 1 (GS1), gas 2 (GS2), and temperature (TEM) were set at 0, 0 and 50 °C, respectively. An accumulation time of 250 ms was used and full scan MS data was recorded in positive TOF-MS mode using an Ion Spray Voltage Floating (ISVF) ranging from 1500V to 1700V. System control and data acquisition were performed using ABSciex 32 Karat (v 10.3) and Analyst® (v 1.8.1) software. This method is based on Chetwynd et al. (231) and Zhang et al. (232).

6.7 Mass spectrometry data processing and analysis

6.7.1 Processing and analysis of LC-MS/MS proteome data

Proteomics raw data were processed using MaxQuant software using the most recent UniProt proteome library file and standard parameters *MS tol, 10 ppm; MS/MS tol, 20 ppm Da; Peptide FDR, 0.1; Protein FDR, 0.01 min; Peptide Length, 7; Variable modifications, Oxidation (M), Acetyl (Protein N-term); Fixed modifications, Carbamidomethyl (C); Peptides for protein quantitation, razor and unique; Min. peptides, 1; Min. ratio count, 2*. The exact MaxQuant software version and exact library file for each separate dataset are specified in each dataset parameter.txt file as linked to the dataset.

Downstream, identified proteins and intensities from MaxQuant's "proteingroups" output file were then handled with Perseus software (version 1.6.7.0. and 1.6.15.0) (233), then Microsoft Excel and R Studio (version 4.0.3). For analysis in Perseus, the output protein_groups.txt file from MaxQuant processing was imported, then protein hits associated with the reversed database, only identified by site and common contaminants were filtered out. Then further analysis was performed as specified in the corresponding results sections.

Data and analysis scripts are available via ProteomeXchange with identifiers specified Chapter "10.2 Access credentials for the datasets and corresponding analysis scripts (available online)".

6.7.2 Processing and analysis of LC-MS/MS histone modification data

Data processing and analysis were performed with Skyline (version 21.10.1.146 and 2.22) by using doubly and triply charged peptide masses for extracted ion chromatograms. Automatic selection of peaks was manually curated based on the relative retention times and fragmentation spectra with results from Proteome Discoverer 1.4. Peak selection was curated by Axel Imhof and Beyza Bozdağ. Integrated peak values were exported to Excel for further calculations and visualization. The relative abundance of an observed modified peptide was calculated as the percentage of the overall peptide.

Data and analysis scripts are available via ProteomeXchange with identifiers specified Chapter "10.2 Access credentials for the datasets and corresponding analysis scripts (available online)".

6.7.3 Processing and analysis of CESI-MS metabolome data

MS1 peaks integration was performed using Skyline software (version 22.2). Peak integration was performed by Marco Borsó. A freshly prepared mixture of selected metabolites was injected with samples and used for m/z and migration time matching. Then, intensities were exported to Excel and normalized and analyzed there. Finally, data were visualized in Excel or R.

Data and skyline analysis files are available via the lrz link provided in Chapter "10.2 Access credentials for the datasets and corresponding analysis scripts (available online)".

7 References

1. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251-260.
2. Marx, V. (2012) Reading the second genomic code. *Nature*, **491**, 143-147.
3. Moore, L.D., Le, T. and Fan, G. (2013) DNA methylation and its basic function. *Neuropsychopharmacology*, **38**, 23-38.
4. Zhang, Z. and Pugh, B.F. (2011) Genomic organization of H2Av containing nucleosomes in *Drosophila* heterochromatin. *PLoS One*, **6**, e20511.
5. Scharf, A.N., Meier, K., Seitz, V., Kremmer, E., Brehm, A. and Imhof, A. (2009) Monomethylation of lysine 20 on histone H4 facilitates chromatin maturation. *Mol Cell Biol*, **29**, 57-67.
6. Henikoff, S. and Shilatifard, A. (2011) Histone modification: cause or cog? *Trends Genet*, **27**, 389-396.
7. Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41-45.
8. Weinzapfel, E.N., Fedder-Semmes, K.N., Sun, Z.-W. and Keogh, M.-C. (2024) Beyond the tail: the consequence of context in histone post-translational modification and chromatin research. *Biochemical Journal*, **481**, 219-244.
9. Alabert, C., Bukowski-Wills, J.C., Lee, S.B., Kustatscher, G., Nakamura, K., de Lima Alves, F., Menard, P., Mejlvang, J., Rappsilber, J. and Groth, A. (2014) Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nat Cell Biol*, **16**, 281-293.
10. Nakamura, K., Kustatscher, G., Alabert, C., Hödl, M., Forne, I., Völker-Albert, M., Satpathy, S., Beyer, T.E., Mailand, N., Choudhary, C. *et al.* (2021) Proteome dynamics at broken replication forks reveal a distinct ATM-directed repair response suppressing DNA double-strand break ubiquitination. *Mol. Cell*, **81**, 1084-1099.
11. Reveron-Gomez, N., Gonzalez-Aguilera, C., Stewart-Morgan, K.R., Petryk, N., Flury, V., Graziano, S., Johansen, J.V., Jakobsen, J.S., Alabert, C. and Groth, A. (2018) Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. *Mol Cell*, **72**, 239-249.e235.
12. Berkovich, E., Monnat, R.J., Jr. and Kastan, M.B. (2008) Assessment of protein dynamics and DNA repair following generation of DNA double-strand breaks at defined genomic sites. *Nat Protoc*, **3**, 915-922.
13. Choudhury, R., Venkateswaran Venkatasubramani, A., Hua, J., Borsò, M., Franconi, C., Kinkley, S., Forné, I. and Imhof, A. (2024) The role of RNA in the maintenance of chromatin domains as revealed by antibody mediated proximity labelling coupled to mass spectrometry. *eLife*, **13**, e95718.
14. Dungrawala, H., Rose, K.L., Bhat, K.P., Mohni, K.N., Glick, G.G., Couch, F.B. and Cortez, D. (2015) The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability. *Mol Cell*, **59**, 998-1010.
15. Sirbu, B.M., Couch, F.B. and Cortez, D. (2012) Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA. *Nat Protoc*, **7**, 594-605.
16. Sirbu, B.M., McDonald, W.H., Dungrawala, H., Badu-Nkansah, A., Kavanaugh, G.M., Chen, Y., Tabb, D.L. and Cortez, D. (2013) Identification of proteins at active, stalled, and collapsed replication forks using isolation of proteins on nascent DNA (iPOND) coupled with mass spectrometry. *J Biol Chem*, **288**, 31458-31467.

17. Kliszczak, A.E., Rainey, M.D., Harhen, B., Boisvert, F.M. and Santocanale, C. (2011) DNA mediated chromatin pull-down for the study of chromatin replication. *Scientific reports*, **1**, 95.
18. Krietenstein, N., Wippo, C.J., Lieleg, C. and Korber, P. (2012) In Wu, C. and Allis, C. D. (eds.), *Methods Enzymol.* Academic Press, Vol. 513, pp. 205-232.
19. Wippo, C.J., Israel, L., Watanabe, S., Hochheimer, A., Peterson, C.L. and Korber, P. (2011) The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. *Embo j*, **30**, 1277-1288.
20. Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C.L., Pugh, B.F. and Korber, P. (2016) Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell*, **167**, 709--721.e712.
21. Oberbeckmann, E., Niebauer, V., Watanabe, S., Farnung, L., Moldt, M., Schmid, A., Cramer, P., Peterson, C.L., Eustermann, S., Hopfner, K.-P. *et al.* (2021) Ruler elements in chromatin remodelers set nucleosome array spacing and phasing. *Nat. Commun.*, **12**, 3232.
22. Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A. and Diffley, J.F.X. (2017) Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. *Mol Cell*, **65**, 117-130.
23. Chacin, E., Reuswig, K.-U., Furtmeier, J., Bansal, P., Karl, L.A., Pfander, B., Straub, T., Korber, P. and Kurat, C.F. (2023) Establishment and function of chromatin organization at replication origins. *Nature*, **616**, 836--842.
24. Becker, P.B. and Wu, C. (1992) Cell-free system for assembly of transcriptionally repressed chromatin from *Drosophila* embryos. *Mol Cell Biol*, **12**, 2241-2249.
25. Wang, W.L., Onikubo, T. and Shechter, D. (2019) Chromatin Characterization in *Xenopus laevis* Cell-Free Egg Extracts and Embryos. *Cold Spring Harb Protoc*, **2019**, pdb.prot099879.
26. Gilbert, S.F. (2000), *Developmental Biology*. 6th ed. Sinauer Associates, Sunderland, Massachusetts.
27. Newport, J. and Kirschner, M. (1982) A major developmental transition in early *Xenopus* embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell*, **30**, 675-686.
28. Völker-Albert, M.C., Pusch, M.C., Fedisch, A., Schilcher, P., Schmidt, A. and Imhof, A. (2016) A Quantitative Proteomic Analysis of In Vitro Assembled Chromatin. *Mol. Cell Proteomics*, **15**, 945 -- 959.
29. Jennings, B.H. (2011) *Drosophila* – a versatile model in biology & medicine. *Materials Today*, **14**, 190-195.
30. Tolwinski, N.S. (2017) Introduction: *Drosophila*-A Model System for Developmental Biology. *J Dev Biol*, **5**.
31. Becker, P.B. (2024) Cell-free genomics: transcription factor interactions in reconstituted naïve embryonic chromatin. *Biochemical Society Transactions*, **52**, 423-429.
32. Colonna, M.M., Schedl, P. and Deshpande, G. (2023) Germline/soma distinction in *Drosophila* embryos requires regulators of zygotic genome activation. *eLife*, **12**, e78188.
33. Eggers, N. and Becker, P.B. (2021) Cell-free genomics reveal intrinsic, cooperative and competitive determinants of chromatin interactions. *Nucleic Acids Res.*, **49**, 7602-7617.
34. Eggers, N., Gkoutromichos, F., Krause, S., Campos-Sparr, A. and Becker, Peter B. (2023) Physical interaction between MSL2 and CLAMP assures direct cooperativity and prevents competition at composite binding sites. *Nucleic Acids Res.*
35. Harpprecht, L., Baldi, S., Schauer, T., Schmidt, A., Bange, T., Robles, M.S., Kremmer, E., Imhof, A. and Becker, P.B. (2019) A *Drosophila* cell-free system that senses DNA breaks and triggers phosphorylation signalling. *Nucleic Acids Res.*, **47**, 7444 -- 7459.
36. Imhof, A. (2006) Epigenetic regulators and histone modification. *Brief Funct Genomic Proteomic*, **5**, 222-227.

37. Harris, H.L., Gu, H., Olshansky, M., Wang, A., Farabella, I., Eliaz, Y., Kalluchi, A., Krishna, A., Jacobs, M., Cauer, G. *et al.* (2023) Chromatin alternates between A and B compartments at kilobase scale for subgenic organization. *Nat. Commun.*, **14**, 3303.
38. Campos, E.I. and Reinberg, D. (2009) Histones: annotating chromatin. *Annu. Rev. Genet.*, **43**, 559 – 599.
39. Lyon, A.S., Peeples, W.B. and Rosen, M.K. (2021) A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.*, **22**, 215--235.
40. Gibson, B.A., Doolittle, L.K., Jensen, L.E., Gamarra, N., Redding, S. and Rosen, M.K. (2019) Organization and Regulation of Chromatin by Liquid-Liquid Phase Separation. *Cell*, **179**, 470-484.e421.
41. Yesudhas, D., Batool, M., Anwar, M.A., Panneerselvam, S. and Choi, S. (2017) Proteins Recognizing DNA: Structural Uniqueness and Versatility of DNA-Binding Domains in Stem Cell Transcription Factors. *Genes (Basel)*, **8**.
42. Peach, S.E., Rudomin, E.L., Udeshi, N.D., Carr, S.A. and Jaffe, J.D. (2012) Quantitative assessment of chromatin immunoprecipitation grade antibodies directed against histone modifications reveals patterns of co-occurring marks on histone protein molecules. *Mol Cell Proteomics*, **11**, 128-137.
43. Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L. and Strahl, B.D. (2011) Influence of combinatorial histone modifications on antibody and effector protein recognition. *Curr Biol*, **21**, 53-58.
44. Skene, P.J. and Henikoff, S. (2017) An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife*, **6**.
45. Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K. and Henikoff, S. (2019) CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.*, **10**, 1930.
46. Scharf, A.N., Barth, T.K. and Imhof, A. (2009) Establishment of histone modifications after chromatin assembly. *Nucleic Acids Res*, **37**, 5032-5040.
47. Yang, Y., Zhang, M. and Wang, Y. (2022) The roles of histone modifications in tumorigenesis and associated inhibitors in cancer therapy. *Journal of the National Cancer Center*, **2**, 277-290.
48. Thålin, C., Lundström, S., Seignez, C., Daleskog, M., Lundström, A., Henriksson, P., Helleday, T., Phillipson, M., Wallén, H. and Demers, M. (2018) Citrullinated histone H3 as a novel prognostic blood marker in patients with advanced cancer. *PLoS One*, **13**, e0191231.
49. Tanikawa, C., Espinosa, M., Suzuki, A., Masuda, K., Yamamoto, K., Tsuchiya, E., Ueda, K., Daigo, Y., Nakamura, Y. and Matsuda, K. (2012) Regulation of histone modification and chromatin structure by the p53–PADI4 pathway. *Nat. Commun.*, **3**, 676.
50. Stanley, J.S., Griffin, J.B. and Zemleni, J. (2001) Biotinylation of histones in human cells *European journal of biochemistry*, **268**, 5424-5429.
51. Kothapalli, N., Camporeale, G., Kueh, A., Chew, Y.C., Oommen, A.M., Griffin, J.B. and Zemleni, J. (2005) Biological functions of biotinylated histones. *J Nutr Biochem*, **16**, 446-448.
52. Ryu, H.Y. and Hochstrasser, M. (2021) Histone sumoylation and chromatin dynamics. *Nucleic Acids Res*, **49**, 6043-6052.
53. Galisson, F., Mahrouche, L., Courcelles, M., Bonneil, E., Meloche, S., Chelbi-Alix, M.K. and Thibault, P. (2011) A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. *Mol Cell Proteomics*, **10**, M110.004796.
54. Messner, S., Altmeyer, M., Zhao, H., Pozivil, A., Roschitzki, B., Gehrig, P., Rutishauser, D., Huang, D., Caflisch, A. and Hottiger, M.O. (2010) PARP1 ADP-ribosylates lysine residues of the core histone tails. *Nucleic Acids Res*, **38**, 6350-6362.

55. Karch, K.R., Langelier, M.F., Pascal, J.M. and Garcia, B.A. (2017) The nucleosomal surface is the main target of histone ADP-ribosylation in response to DNA damage. *Mol Biosyst*, **13**, 2660-2671.
56. Nelson, C.J., Santos-Rosa, H. and Kouzarides, T. (2006) Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell*, **126**, 905-916.
57. Zhang, D., Tang, Z., Huang, H., Zhou, G., Cui, C., Weng, Y., Liu, W., Kim, S., Lee, S., Perez-Neut, M. *et al.* (2019) Metabolic regulation of gene expression by histone lactylation. *Nature*, **574**, 575-580.
58. Kebede, A.F., Nieborak, A., Shahidian, L.Z., Le Gras, S., Richter, F., Gómez, D.A., Baltissen, M.P., Meszaros, G., Magliarelli, H.d.F., Taudt, A. *et al.* (2017) Histone propionylation is a mark of active chromatin. *Nature structural & molecular biology*, **24**, 1048-1056.
59. Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S.C., Falck, J.R., Peng, J., Gu, W. and Zhao, Y. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics*, **6**, 812-819.
60. Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N. *et al.* (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, **146**, 1016-1028.
61. Cavalieri, V. (2021) The Expanding Constellation of Histone Post-Translational Modifications in the Epigenetic Landscape. *Genes*, **12**, 1596.
62. Chiu, T.-P., Rao, S., Mann, R.S., Honig, B. and Rohs, R. (2017) Genome-wide prediction of minor-groove electrostatic potential enables biophysical modeling of protein–DNA binding. *Nucleic Acids Res.*, **45**, 12565-12576.
63. Afek, A., Schipper, J.L., Horton, J., Gordân, R. and Lukatsky, D.B. (2014) Protein–DNA binding in the absence of specific base-pair recognition. *Proceedings of the National Academy of Sciences*, **111**, 17140-17145.
64. Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S. and Honig, B. (2009) The role of DNA shape in protein-DNA recognition. *Nature*, **461**, 1248-1253.
65. Harrison, S.C. and Aggarwal, A.K. (1990) DNA recognition by proteins with the helix-turn-helix motif. *Annu Rev Biochem*, **59**, 933-969.
66. Privalov, P.L., Dragan, A.I., Crane-Robinson, C., Breslauer, K.J., Remeta, D.P. and Minetti, C.A. (2007) What drives proteins into the major or minor grooves of DNA? *J Mol Biol*, **365**, 1-9.
67. Ziach, K., Chollet, C., Parissi, V., Prabhakaran, P., Marchivie, M., Corvaglia, V., Bose, P.P., Laxmi-Reddy, K., Godde, F., Schmitter, J.-M. *et al.* (2018) Single helically folded aromatic oligoamides that mimic the charge surface of double-stranded B-DNA. *Nat. Chem.*, **10**, 511--518.
68. Wang, H.-C., Ho, C.-H., Hsu, K.-C., Yang, J.-M. and Wang, A.H.J. (2014) DNA Mimic Proteins: Functions, Structures, and Bioinformatic Analysis. *Biochemistry*, **53**, 2865--2874.
69. Wang, H.C., Chou, C.C., Hsu, K.C., Lee, C.H. and Wang, A.H.J. (2019) New paradigm of functional regulation by DNA mimic proteins: Recent updates. *IUBMB Life*, **71**, 539--548.
70. Bando, T. and Sugiyama, H. (2006) Synthesis and Biological Properties of Sequence-Specific DNA-Alkylating Pyrrole–Imidazole Polyamides. *Acc. Chem. Res.*, **39**, 935--944.
71. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.*, **13**, 284--299.
72. Chenoweth, D.M., Poposki, J.A., Marques, M.A. and Dervan, P.B. (2007) Programmable oligomers targeting 5'-GGGG-3' in the minor groove of DNA and NF- κ B binding inhibition. *Bioorg. Med. Chem.*, **15**, 759--770.
73. Corvaglia, V., Carbajo, D., Prabhakaran, P., Ziach, K., Mandal, P.K., Santos, V.D., Legeay, C., Vogel, R., Parissi, V., Pourquier, P. *et al.* (2019) Carboxylate-functionalized foldamer inhibitors

- of HIV-1 integrase and Topoisomerase 1: artificial analogues of DNA mimic proteins. *Nucleic Acids Res.*, **47**, 5511--5521.
74. Dai, Z., Ramesh, V. and Locasale, J.W. (2020) The evolving metabolic landscape of chromatin biology and epigenetics. *Nat. Rev. Genet.*, **21**, 737--753.
 75. Erdel, F. and Rippe, K. (2018) Formation of Chromatin Subcompartments by Phase Separation. *Biophysical journal*, **114**, 2262-2270.
 76. Campit, S.E., Meliki, A., Youngson, N.A. and Chandrasekaran, S. (2020) Nutrient Sensing by Histone Marks: Reading the Metabolic Histone Code Using Tracing, Omics, and Modeling. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **42**, 2000083.
 77. Ye, C. and Tu, B.P. (2018) Sink into the Epigenome: Histones as Repositories That Influence Cellular Metabolism. *Trends Endocrinol Metab*, **29**, 626-637.
 78. Alexander, D.C., Corman, T., Mendoza, M., Glass, A., Belity, T., Wu, R., Campbell, R.R., Han, J., Keiser, A.A., Winkler, J. *et al.* (2022) Targeting acetyl-CoA metabolism attenuates the formation of fear memories through reduced activity-dependent histone acetylation. *Proceedings of the National Academy of Sciences*, **119**, e2114758119.
 79. Mews, P. and Berger, S.L. (2016) In Marmorstein, R. (ed.), *Methods Enzymol.* Academic Press, Vol. 574, pp. 311-329.
 80. Shlomi, T., Fan, J., Tang, B., Kruger, W.D. and Rabinowitz, J.D. (2014) Quantitation of cellular metabolic fluxes of methionine. *Anal Chem*, **86**, 1583-1591.
 81. Danzi, F., Pacchiana, R., Mafficini, A., Scupoli, M.T., Scarpa, A., Donadelli, M. and Fiore, A. (2023) To metabolomics and beyond: a technological portfolio to investigate cancer metabolism. *Signal Transduction and Targeted Therapy*, **8**, 137.
 82. Rakusanova, S., Fiehn, O. and Cajka, T. (2023) Toward building mass spectrometry-based metabolomics and lipidomics atlases for biological and clinical research. *TrAC Trends in Analytical Chemistry*, **158**, 116825.
 83. Görisch, S.M., Wachsmuth, M., Tóth, K.F., Lichter, P. and Rippe, K. (2005) Histone acetylation increases chromatin accessibility. *J Cell Sci*, **118**, 5825-5834.
 84. Hebbes, T.R., Clayton, A.L., Thorne, A.W. and Crane-Robinson, C. (1994) Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *Embo j*, **13**, 1823-1830.
 85. Krajewski, W.A. and Becker, P.B. (1998) Reconstitution of hyperacetylated, DNase I-sensitive chromatin characterized by high conformational flexibility of nucleosomal DNA. *Proceedings of the National Academy of Sciences*, **95**, 1540-1545.
 86. Vogelauer, M., Wu, J., Suka, N. and Grunstein, M. (2000) Global histone acetylation and deacetylation in yeast. *Nature*, **408**, 495-498.
 87. Bone, J.R., Lavender, J., Richman, R., Palmer, M.J., Turner, B.M. and Kuroda, M.I. (1994) Acetylated histone H4 on the male X chromosome is associated with dosage compensation in *Drosophila*. *Genes Dev*, **8**, 96-104.
 88. Trefely, S., Doan, M.T. and Snyder, N.W. (2019) Crosstalk between cellular metabolism and histone acetylation. *Methods Enzymol*, **626**, 1-21.
 89. Peleg, S., Feller, C., Ladurner, A.G. and Imhof, A. (2016) The Metabolic Impact on Histone Acetylation and Transcription in Ageing. *Trends in biochemical sciences*, **41**, 700-711.
 90. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. *et al.* (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, **406**, 593-599.
 91. Bedford, M.T. and Richard, S. (2005) Arginine Methylation: An Emerging Regulator of Protein Function. *Mol. Cell*, **18**, 263-272.
 92. Kooistra, S.M. and Helin, K. (2012) Molecular mechanisms and potential functions of histone demethylases. *Nat. Rev. Mol. Cell Biol.*, **13**, 297-311.

93. Serefidou, M., Venkatasubramani, A.V. and Imhof, A. (2019) The Impact of One Carbon Metabolism on Histone Methylation. *Front Genet*, **10**, 764.
94. Shoaib, M., Chen, Q., Shi, X., Nair, N., Prasanna, C., Yang, R., Walter, D., Frederiksen, K.S., Einarsson, H., Svensson, J.P. *et al.* (2021) Histone H4 lysine 20 mono-methylation directly facilitates chromatin openness and promotes transcription of housekeeping genes. *Nat Commun*, **12**, 4800.
95. Pedersen, M.T. and Helin, K. (2010) Histone demethylases in development and disease. *Trends Cell Biol*, **20**, 662-671.
96. Greer, E.L. and Shi, Y. (2012) Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet*, **13**, 343-357.
97. Cantoni, G.L. (1953) S-Adenosylmethionine; a new intermediate formed enzymatically from L-methionine and adenosinetriphosphate. *J Biol Chem*, **204**, 403-416.
98. Larsson, J., Zhang, J. and Rasmuson-Lestander, A. (1996) Mutations in the *Drosophila melanogaster* gene encoding S-adenosylmethionine synthetase [corrected] suppress position-effect variegation. *Genetics*, **143**, 887-896.
99. Petrossian, T.C. and Clarke, S.G. (2011) Uncovering the human methyltransferasome. *Mol Cell Proteomics*, **10**, M110.000976.
100. Caudill, M.A., Wang, J.C., Melnyk, S., Pogribny, I.P., Jernigan, S., Collins, M.D., Santos-Guzman, J., Swendseid, M.E., Cogger, E.A. and James, S.J. (2001) Intracellular S-adenosylhomocysteine concentrations predict global DNA hypomethylation in tissues of methyl-deficient cystathionine beta-synthase heterozygous mice. *J Nutr*, **131**, 2811-2818.
101. Palmer, J.L. and Abeles, R.H. (1979) The mechanism of action of S-adenosylhomocysteinase. *Journal of Biological Chemistry*, **254**, 1217-1226.
102. Kusakabe, Y., Ishihara, M., Umeda, T., Kuroda, D., Nakanishi, M., Kitade, Y., Gouda, H., Nakamura, K.T. and Tanaka, N. (2015) Structural insights into the reaction mechanism of S-adenosyl-L-homocysteine hydrolase. *Scientific reports*, **5**, 16641.
103. Liu, M., Barnes, V.L. and Pile, L.A. (2016) Disruption of Methionine Metabolism in *Drosophila melanogaster* Impacts Histone Methylation and Results in Loss of Viability. *G3 Genes | Genomes | Genetics*, **6**, 121-132.
104. Fan, J., Krautkramer, K.A., Feldman, J.L. and Denu, J.M. (2015) Metabolic regulation of histone post-translational modifications. *ACS Chem Biol*, **10**, 95-108.
105. Leonhardt, H., Page, A.W., Weier, H.U. and Bestor, T.H. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell*, **71**, 865-873.
106. Hermann, A., Goyal, R. and Jeltsch, A. (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem*, **279**, 48350-48359.
107. Cowling, V.H. (2009) Regulation of mRNA cap methylation. *Biochem J*, **425**, 295-302.
108. Aranda, S., Alcaine-Colet, A., Blanco, E., Borràs, E., Caillot, C., Sabido, E. and Croce, L.D. (2019) Chromatin capture links the metabolic enzyme AHCY to stem cell proliferation. *Sci Adv*, **5**, eaav2448.
109. Greco, C.M., Cervantes, M., Fustin, J.M., Ito, K., Ceglia, N., Samad, M., Shi, J., Koronowski, K.B., Forne, I., Ranjit, S. *et al.* (2020) S-adenosyl-L-homocysteine hydrolase links methionine metabolism to the circadian clock and chromatin remodeling. *Sci Adv*, **6**.
110. Vizán, P., Di Croce, L. and Aranda, S. (2021) Functional and Pathological Roles of AHCY. *Front Cell Dev Biol*, **9**, 654344.
111. Kleene, V., Corvaglia, V., Chacin, E., Forne, I., Konrad, D.B., Khosravani, P., Douat, C., Kurat, C.F., Huc, I. and Imhof, A. (2023) DNA mimic foldamers affect chromatin composition and disturb cell cycle progression. *Nucleic Acids Res.*, **51**, 9629-9642.

112. Ner, S.S., Blank, T., Pérez-Paralle, M.L., Grigliatti, T.A., Becker, P.B. and Travers, A.A. (2001) HMG-D and histone H1 interplay during chromatin assembly and early embryogenesis. *J Biol Chem*, **276**, 37569-37576.
113. Ner, S.S. and Travers, A.A. (1994) HMG-D, the *Drosophila melanogaster* homologue of HMG 1 protein, is associated with early embryonic chromatin in the absence of histone H1. *EMBO J.*, **13**, 1817--1822.
114. Syed, A. and Tainer, J.A. (2018) The MRE11–RAD50–NBS1 Complex Conducts the Orchestration of Damage Signaling and Outcomes to Stress in DNA Replication and Repair. *Annual Review of Biochemistry*, **87**, 263-294.
115. Oldenkamp, R. and Rowland, B.D. (2022) A walk through the SMC cycle: From catching DNAs to shaping the genome. *Mol. Cell*, **82**, 1616--1632.
116. Tiana, M., Acosta-Iborra, B., Puente-Santamaría, L., Hernansanz-Agustin, P., Worsley-Hunt, R., Masson, N., García-Rio, F., Mole, D., Ratcliffe, P., Wasserman, W.W. *et al.* (2018) The SIN3A histone deacetylase complex is required for a complete transcriptional response to hypoxia. *Nucleic Acids Res*, **46**, 120-133.
117. McCarthy-Leo, C., Darwiche, F. and Tainsky, M.A. (2022) DNA Repair Mechanisms, Protein Interactions and Therapeutic Targeting of the MRN Complex. *Cancers (Basel)*, **14**.
118. Mounkes, L.C. and Fuller, M.T. (1999) Molecular characterization of mutant alleles of the DNA repair/basal transcription factor haywire/ERCC3 in *Drosophila*. *Genetics*, **152**, 291-297.
119. Grau-Bové, X., Navarrete, C., Chiva, C., Pribasnig, T., Antó, M., Torruella, G., Galindo, L.J., Lang, B.F., Moreira, D., López-Garcia, P. *et al.* (2022) A phylogenetic and proteomic reconstruction of eukaryotic chromatin evolution. *Nat. Ecol. Evol.*, **6**, 1007--1023.
120. Donald, J.E., Chen, W.W. and Shakhnovich, E.I. (2007) Energetics of protein-DNA interactions. *Nucleic Acids Res*, **35**, 1039-1047.
121. Segal, E. and Widom, J. (2009) From DNA sequence to transcriptional behaviour: a quantitative approach. *Nat Rev Genet*, **10**, 443-456.
122. Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G. and Raschella, G. (2017) Zinc-finger proteins in health and disease. *Cell Death Discovery*, **3**, 17071.
123. Margineanu, A., Chan, J.J., Kelly, D.J., Warren, S.C., Flatters, D., Kumar, S., Katan, M., Dunsby, C.W. and French, P.M.W. (2016) Screening for protein-protein interactions using Förster resonance energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM). *Scientific reports*, **6**, 28186.
124. Rosas, R., Aguilar, R.R., Arslanovic, N., Seck, A., Smith, D.J., Tyler, J.K. and Churchill, M.E.A. (2023) A novel single alpha-helix DNA-binding domain in CAF-1 promotes gene silencing and DNA damage survival through tetrasome-length DNA selectivity and spacer function. *eLife*, **12**.
125. Zhang, K., Gao, Y., Li, J., Burgess, R., Han, J., Liang, H., Zhang, Z. and Liu, Y. (2016) A DNA binding winged helix domain in CAF-1 functions with PCNA to stabilize CAF-1 at replication forks. *Nucleic Acids Res*, **44**, 5083-5094.
126. Harami, G.M., Gyimesi, M. and Kovács, M. (2013) From keys to bulldozers: expanding roles for winged helix domains in nucleic-acid-binding proteins. *Trends in biochemical sciences*, **38**, 364-371.
127. Schuetz, A., Murakawa, Y., Rosenbaum, E., Landthaler, M. and Heinemann, U. (2014) Roquin binding to target mRNAs involves a winged helix-turn-helix motif. *Nat. Commun.*, **5**, 5701.
128. Zou, Y., Liu, Y., Wu, X. and Shell, S.M. (2006) Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. *J Cell Physiol*, **208**, 267-273.

129. Krasikova, Y.S., Rechkunova, N.I. and Lavrik, O.I. (2016) [Replication protein A as a major eukaryotic single-stranded DNA-binding protein and its role in DNA repair]. *Mol Biol (Mosk)*, **50**, 735-750.
130. Iftode, C., Daniely, Y. and Borowiec, J.A. (1999) Replication protein A (RPA): the eukaryotic SSB. *Crit Rev Biochem Mol Biol*, **34**, 141-180.
131. Wold, M.S. (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem*, **66**, 61-92.
132. de Laat, W.L., Appeldoorn, E., Sugawara, K., Weterings, E., Jaspers, N.G. and Hoeijmakers, J.H. (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev*, **12**, 2598-2609.
133. Bochkarev, A. and Bochkareva, E. (2004) From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol*, **14**, 36-42.
134. Sun, M., Nishino, T. and Marko, J.F. (2013) The SMC1-SMC3 cohesin heterodimer structures DNA through supercoiling-dependent loop formation. *Nucleic Acids Res*, **41**, 6149-6160.
135. Kim, E., Gonzalez, A.M., Pradhan, B., van der Torre, J. and Dekker, C. (2022) Condensin-driven loop extrusion on supercoiled DNA. *Nature structural & molecular biology*, **29**, 719-727.
136. Serban, D., Benevides, J.M. and Thomas, G.J., Jr. (2002) DNA secondary structure and Raman markers of supercoiling in Escherichia coli plasmid pUC19. *Biochemistry*, **41**, 847-853.
137. King, G.A., Burla, F., Peterman, E.J.G. and Wuite, G.J.L. (2019) Supercoiling DNA optically. *Proceedings of the National Academy of Sciences*, **116**, 26534-26539.
138. Aftab, W., Lahiri, S. and Imhof, A. (2022) ImShot: An Open-Source Software for Probabilistic Identification of Proteins *In Situ* and Visualization of Proteomics Data. *Mol. Cell Proteomics*, **21**.
139. Rogers, S.L. and Rogers, G.C. (2008) Culture of Drosophila S2 cells and their use for RNAi-mediated loss-of-function studies and immunofluorescence microscopy. *Nat. Protoc.*, **3**, 606--611.
140. Corvaglia, V., Sanchez, F., Menke, F.S., Douat, C. and Huc, I. (2023) Optimization and Automation of Helical Aromatic Oligoamide Foldamer Solid-Phase Synthesis. *Chem. Eur. J.*, **29**, e202300898.
141. Mura, M., Chaudhury, S., Farooq, F., Duncan, E.H., Beck, K. and Bergmann-Leitner, E.S. (2020) Optimized flow cytometric protocol for the detection of functional subsets of low frequency antigen-specific CD4(+) and CD8(+) T cells. *MethodsX*, **7**, 101005.
142. Bennett, H.M., Stephenson, W., Rose, C.M. and Darmanis, S. (2023) Single-cell proteomics enabled by next-generation sequencing or mass spectrometry. *Nat. Methods*, **20**, 363-374.
143. Cid-Samper, F., Gelabert-Baldrich, M., Lang, B., Lorenzo-Gotor, N., Ponti, R.D., Severijnen, L.-A.W.F.M., Bolognesi, B., Gelpi, E., Hukema, R.K., Botta-Orfila, T. *et al.* (2018) An Integrative Study of Protein-RNA Condensates Identifies Scaffolding RNAs and Reveals Players in Fragile X-Associated Tremor/Ataxia Syndrome. *Cell Reports*, **25**, 3422-3434.e3427.
144. Sanchez de Groot, N., Armaos, A., Graña-Montes, R., Alriquet, M., Calloni, G., Vabulas, R.M. and Tartaglia, G.G. (2019) RNA structure drives interaction with proteins. *Nat. Commun.*, **10**, 3246.
145. Kovachev, P.S., Gomes, M.P.B., Cordeiro, Y., Ferreira, N.C., Valadão, L.P.F., Ascari, L.M., Rangel, L.P., Silva, J.L. and Sanyal, S. (2019) RNA modulates aggregation of the recombinant mammalian prion protein by direct interaction. *Scientific reports*, **9**, 12406.
146. Louka, A., Zacco, E., Temussi, P.A., Tartaglia, G.G. and Pastore, A. (2020) RNA as the stone guest of protein aggregation. *Nucleic Acids Res.*, **48**, 11880-11889.
147. Pohar, J., Lainšček, D., Ivičak-Kocjan, K., Cajnko, M.M., Jerala, R. and Benčina, M. (2017) Short single-stranded DNA degradation products augment the activation of Toll-like receptor 9. *Nat Commun*, **8**, 15363.

148. Quinodoz, S.A., Jachowicz, J.W., Bhat, P., Ollikainen, N., Banerjee, A.K., Goronzy, I.N., Blanco, M.R., Chovanec, P., Chow, A., Markaki, Y. *et al.* (2021) RNA promotes the formation of spatial compartments in the nucleus. *Cell*, **184**, 5775-5790.e5730.
149. Deepak D, C.V., Wu J, Allmendinger L, Huc I. (2023) DNA Mimic Foldamer Recognition of a Chromosomal Protein. *ChemRxiv*.
150. Ahmed, R., Spikings, E., Zhou, S., Thompson, A. and Zhang, T. (2014) Pre-hybridisation: An efficient way of suppressing endogenous biotin-binding activity inherent to biotin–streptavidin detection system. *Journal of Immunological Methods*, **406**, 143-147.
151. Keppler, A., Gendrezig, S., Gronemeyer, T., Pick, H., Vogel, H. and Johnsson, K. (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat Biotechnol*, **21**, 86-89.
152. Gautier, A., Juillerat, A., Heinis, C., Corrêa, I.R., Jr., Kindermann, M., Beauflis, F. and Johnsson, K. (2008) An engineered protein tag for multiprotein labeling in living cells. *Chem Biol*, **15**, 128-136.
153. Teodori, L., Omer, M., Märcher, A., Skaanning, M.K., Andersen, V.L., Nielsen, J.S., Oldenburg, E., Lin, Y., Gothelf, K.V. and Kjems, J. (2022) Site-specific nanobody-oligonucleotide conjugation for super-resolution imaging. *J Biol Methods*, **9**, e159.
154. Li, S., Wang, L., Yu, F., Zhu, Z., Shobaki, D., Chen, H., Wang, M., Wang, J., Qin, G., Erasquin, U.J. *et al.* (2017) Copper-catalyzed click reaction on/in live cells. *Chemical Science*, **8**, 2107-2114.
155. Ishizuka, T., Liu, H.S., Ito, K. and Xu, Y. (2016) Fluorescence imaging of chromosomal DNA using click chemistry. *Scientific reports*, **6**, 33217.
156. Baskin, J.M., Prescher, J.A., Laughlin, S.T., Agard, N.J., Chang, P.V., Miller, I.A., Lo, A., Codelli, J.A. and Bertozzi, C.R. (2007) Copper-free click chemistry for dynamic *in vivo* imaging. *Proceedings of the National Academy of Sciences*, **104**, 16793-16797.
157. Hirokazu, S. (2012) In Anica, D. (ed.), *Methylation*. IntechOpen, Rijeka, pp. Ch. 9.
158. Morriss-Andrews, A., Rottler, J. and Plotkin, S. (2010) A systematically coarse-grained model for DNA and its predictions for persistence length, stacking, twist, and chirality. *The Journal of chemical physics*, **132**, 035105.
159. Li, J., Sagendorf, J.M., Chiu, T.-P., Pasi, M., Perez, A. and Rohs, R. (2017) Expanding the repertoire of DNA shape features for genome-scale studies of transcription factor binding. *Nucleic Acids Res.*, **45**, 12877-12887.
160. He, J., Duan, Q., Ran, C., Fu, T., Liu, Y. and Tan, W. (2023) Recent progress of aptamer–drug conjugates in cancer therapy. *Acta Pharm Sin B*, **13**, 1358-1370.
161. Soldevilla, M.M., Meraviglia-Crivelli de Caso, D., Menon, A.P. and Pastor, F. (2018) Aptamer-iRNAs as Therapeutics for Cancer Treatment. *Pharmaceuticals (Basel)*, **11**.
162. Kanada, M., Bachmann, M.H., Hardy, J.W., Frimannson, D.O., Bronsart, L., Wang, A., Sylvester, M.D., Schmidt, T.L., Kaspar, R.L., Butte, M.J. *et al.* (2015) Differential fates of biomolecules delivered to target cells via extracellular vesicles. *Proceedings of the National Academy of Sciences*, **112**, E1433-E1442.
163. Liu, C., He, D., Cen, H., Chen, H., Li, L., Nie, G., Zhong, Z., He, Q., Yang, X., Guo, S. *et al.* (2022) Nucleic acid functionalized extracellular vesicles as promising therapeutic systems for nanomedicine. *Extracellular Vesicles and Circulating Nucleic Acids*, **3**, 14-30.
164. Ghanam, J., Chetty, V.K., Barthel, L., Reinhardt, D., Hoyer, P.-F. and Thakur, B.K. (2022) DNA in extracellular vesicles: from evolution to its current application in health and disease. *Cell & Bioscience*, **12**, 37.
165. Cecchin, R., Troyer, Z., Witwer, K. and Morris, K.V. (2023) Extracellular vesicles: The next generation in gene therapy delivery. *Molecular Therapy*, **31**, 1225-1230.

166. O'Brien, K., Breyne, K., Ughetto, S., Laurent, L.C. and Breakefield, X.O. (2020) RNA delivery by extracellular vesicles in mammalian cells and its applications. *Nat. Rev. Mol. Cell Biol.*, **21**, 585-606.
167. Du, R., Wang, C., Zhu, L. and Yang, Y. (2022) Extracellular Vesicles as Delivery Vehicles for Therapeutic Nucleic Acids in Cancer Gene Therapy: Progress and Challenges. *Pharmaceutics*, **14**.
168. Zhuang, J., Tan, J., Wu, C., Zhang, J., Liu, T., Fan, C., Li, J. and Zhang, Y. (2020) Extracellular vesicles engineered with valency-controlled DNA nanostructures deliver CRISPR/Cas9 system for gene therapy. *Nucleic Acids Res.*, **48**, 8870-8882.
169. Orefice, N. (2024).
170. Chesnokov, I.N. (2007) Multiple functions of the origin recognition complex. *Int Rev Cytol*, **256**, 69-109.
171. Kuo, A.J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J.K., Patel, D.J. and Gozani, O. (2012) The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature*, **484**, 115--119.
172. Cayrou, C., Ballester, B., Peiffer, I., Fenouil, R., Coulombe, P., Andrau, J.C., van Helden, J. and Méchali, M. (2015) The chromatin environment shapes DNA replication origin organization and defines origin classes. *Genome research*, **25**, 1873-1885.
173. Li, S., Yang, Z., Du, X., Liu, R., Wilkinson, A.W., Gozani, O., Jacobsen, S.E., Patel, D.J. and Du, J. (2016) Structural Basis for the Unique Multivalent Readout of Unmodified H3 Tail by Arabidopsis ORC1b BAH-PHD Cassette. *Structure*, **24**, 486-494.
174. Long, H., Zhang, L., Lv, M., Wen, Z., Zhang, W., Chen, X., Zhang, P., Li, T., Chang, L., Jin, C. *et al.* (2020) H2A.Z facilitates licensing and activation of early replication origins. *Nature*, **577**, 576-581.
175. Pak, D.T., Pflumm, M., Chesnokov, I., Huang, D.W., Kellum, R., Marr, J., Romanowski, P. and Botchan, M.R. (1997) Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell*, **91**, 311-323.
176. Bosco, G., Du, W. and Orr-Weaver, T.L. (2001) DNA replication control through interaction of E2F-RB and the origin recognition complex. *Nat Cell Biol*, **3**, 289-295.
177. Fujita, M., Hori, Y., Shirahige, K., Tsurimoto, T., Yoshikawa, H. and Obuse, C. (1998) Cell cycle dependent topological changes of chromosomal replication origins in *Saccharomyces cerevisiae*. *Genes Cells*, **3**, 737-749.
178. Kara, N., Hossain, M., Prasanth, S.G. and Stillman, B. (2015) Orc1 Binding to Mitotic Chromosomes Precedes Spatial Patterning during G1 Phase and Assembly of the Origin Recognition Complex in Human Cells. *J Biol Chem*, **290**, 12355-12369.
179. Ohta, S., Bukowski-Wills, J.C., Sanchez-Pulido, L., Alves Fde, L., Wood, L., Chen, Z.A., Platani, M., Fischer, L., Hudson, D.F., Ponting, C.P. *et al.* (2010) The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics. *Cell*, **142**, 810-821.
180. Lee, K.Y., Bang, S.W., Yoon, S.W., Lee, S.H., Yoon, J.B. and Hwang, D.S. (2012) Phosphorylation of ORC2 protein dissociates origin recognition complex from chromatin and replication origins. *J Biol Chem*, **287**, 11891-11898.
181. Kreitz, S., Ritz, M., Baack, M. and Knippers, R. (2001) The human origin recognition complex protein 1 dissociates from chromatin during S phase in HeLa cells. *J Biol Chem*, **276**, 6337-6342.
182. Bravieri, R., Shiyanova, T., Chen, T.-H., Overdier, D. and Liao, X. (1997) Different DNA contact schemes are used by two winged helix proteins to recognize a DNA binding sequence. *Nucleic Acids Res.*, **25**, 2888-2896.
183. Clark, K.L., Halay, E.D., Lai, E. and Burley, S.K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, **364**, 412-420.

184. Schmidt, J.M. and Bleichert, F. (2020) Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. *Nat. Commun.*, **11**, 4263.
185. Chen, Y.A., Stuwe, E., Luo, Y., Ninova, M., Le Thomas, A., Rozhavskaia, E., Li, S., Vempati, S., Laver, J.D., Patel, D.J. *et al.* (2016) Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Mol Cell*, **63**, 97-109.
186. Wan, L., Dockendorff, T.C., Jongens, T.A. and Dreyfuss, G. (2000) Characterization of dFMR1, a *Drosophila melanogaster* Homolog of the Fragile X Mental Retardation Protein. *Mol. Cell. Biol.*, **20**, 8536-8547.
187. Ghosh, A.C., Shimell, M., Leof, E.R., Haley, M.J. and O'Connor, M.B. (2015) UPRT, a suicide-gene therapy candidate in higher eukaryotes, is required for *Drosophila* larval growth and normal adult lifespan. *Scientific reports*, **5**, 13176.
188. Tian, X., Isamiddinova, N.S., Peroutka, R.J., Goldenberg, S.J., Mattern, M.R., Nicholson, B. and Leach, C. (2011) Characterization of selective ubiquitin and ubiquitin-like protease inhibitors using a fluorescence-based multiplex assay format. *Assay Drug Dev Technol*, **9**, 165-173.
189. Ostrowska, H., Wojcik, C., Omura, S. and Worowski, K. (1997) Lactacystin, a specific inhibitor of the proteasome, inhibits human platelet lysosomal cathepsin A-like enzyme. *Biochemical and biophysical research communications*, **234**, 729-732.
190. GUO, N. and PENG, Z. (2013) MG132, a proteasome inhibitor, induces apoptosis in tumor cells. *Asia-Pacific Journal of Clinical Oncology*, **9**, 6-11.
191. Sridevi Gurubaran, I., Hytti, M., Kaarniranta, K. and Kauppinen, A. (2022) Epoxomicin, a Selective Proteasome Inhibitor, Activates AIM2 Inflammasome in Human Retinal Pigment Epithelium Cells. *Antioxidants*, **11**, 1288.
192. Kim, K.B., Myung, J., Sin, N. and Crews, C.M. (1999) Proteasome inhibition by the natural products epoxomicin and dihydroeponemycin: Insights into specificity and potency. *Bioorganic & Medicinal Chemistry Letters*, **9**, 3335-3340.
193. Meng, L., Mohan, R., Kwok, B.H.B., Eloffson, M., Sin, N. and Crews, C.M. (1999) Epoxomicin, a potent and selective proteasome inhibitor, exhibits *in vivo* antiinflammatory activity. *Proceedings of the National Academy of Sciences*, **96**, 10403-10408.
194. Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R. *et al.* (2002) PR-Set7 Is a Nucleosome-Specific Methyltransferase that Modifies Lysine 20 of Histone H4 and Is Associated with Silent Chromatin. *Mol. Cell*, **9**, 1201-1213.
195. Li, C., Chu, S., Tan, S., Yin, X., Jiang, Y., Dai, X., Gong, X., Fang, X. and Tian, D. (2021) Towards Higher Sensitivity of Mass Spectrometry: A Perspective From the Mass Analyzers. *Frontiers in Chemistry*, **9**.
196. Ettorre, A., Rösli, C., Silacci, M., Brack, S., McCombie, G., Knochenmuss, R., Elia, G. and Neri, D. (2006) Recombinant antibodies for the depletion of abundant proteins from human serum. *Proteomics*, **6**, 4496-4505.
197. Clift, D., McEwan, W.A., Labzin, L.I., Konieczny, V., Mogessie, B., James, L.C. and Schuh, M. (2017) A Method for the Acute and Rapid Degradation of Endogenous Proteins. *Cell*, **171**, 1692-1706.e1618.
198. Hoy, M.A. (2019) In Hoy, M. A. (ed.), *Insect Molecular Genetics (Fourth Edition)*. Academic Press, pp. 103-175.
199. Haass, C. and Klotzel, P.M. (1989) The *Drosophila* proteasome undergoes changes in its subunit pattern during development. *Experimental cell research*, **180**, 243-252.

200. Sandaltzopoulos, R., Blank, T. and Becker, P.B. (1994) Transcriptional repression by nucleosomes but not H1 in reconstituted preblastoderm *Drosophila* chromatin. *Embo j*, **13**, 373-379.
201. Qin, X., Ahn, S., Speed, T.P. and Rubin, G.M. (2007) Global analyses of mRNA translational control during early *Drosophila* embryogenesis. *Genome Biol*, **8**, R63.
202. Tikhonova, E., Revel-Muroz, A., Georgiev, P. and Maksimenko, O. (2024) Interaction of MLE with CLAMP zinc finger is involved in proper MSL proteins binding to chromosomes in *Drosophila*. *Open Biol*, **14**, 230270.
203. Shmueli, M.D., Sheban, D., Eisenberg-Lerner, A. and Merbl, Y. (2021) Histone Degradation by the Proteasome Regulates Chromatin and Cellular Plasticity. *Febs j*.
204. Boltengagen, M., Verhagen, D., Wolff, M.R., Oberbeckmann, E., Hanke, M., Gerland, U., Korber, P. and Mueller-Planitz, F. (2023) A single fiber view of the nucleosome organization in eukaryotic chromatin. *Nucleic Acids Res*.
205. McCann, T.S. and Tansey, W.P. (2014) Functions of the Proteasome on Chromatin. *Biomolecules*, **4**, 1026-1044.
206. Ni, J.Q., Liu, L.P., Hess, D., Rietdorf, J. and Sun, F.L. (2006) *Drosophila* ribosomal proteins are associated with linker histone H1 and suppress gene transcription. *Genes Dev*, **20**, 1959-1973.
207. Ugolini, I., Bilokapic, S., Ferrolino, M., Teague, J., Yan, Y., Zhou, X., Deshmukh, A., White, M., Kriwacki, R.W. and Halic, M. (2022) Chromatin localization of nucleophosmin organizes ribosome biogenesis. *Mol. Cell*, **82**, 4443-4457.e4449.
208. Chuang, S.-M., Chen, L., Lambertson, D., Anand, M., Kinzy, T.G. and Madura, K. (2005) Proteasome-Mediated Degradation of Cotranslationally Damaged Proteins Involves Translation Elongation Factor 1A. *Mol. Cell. Biol.*, **25**, 403-413.
209. Hoffman, L. and Rechsteiner, M. (1997) Effects of nucleotides on assembly of the 26S proteasome and degradation of ubiquitin conjugates. *Mol Biol Rep*, **24**, 13-16.
210. Rosa, F. and Osorio, J.S. (2020) Quantitative determination of histone methylation via fluorescence resonance energy transfer (FRET) technology in immortalized bovine mammary alveolar epithelial cells supplemented with methionine. *PLoS One*, **15**, e0244135.
211. Dubey, A., Rangarajan, A., Pal, D. and Atreya, H.S. (2015) Pattern Recognition-Based Approach for Identifying Metabolites in Nuclear Magnetic Resonance-Based Metabolomics. *Analytical Chemistry*, **87**, 7148-7155.
212. Kosakamoto, H., Obata, F., Kuraishi, J., Aikawa, H., Okada, R., Johnstone, J.N., Onuma, T., Piper, M.D.W. and Miura, M. (2023) Early-adult methionine restriction reduces methionine sulfoxide and extends lifespan in *Drosophila*. *Nat. Commun.*, **14**, 7832.
213. Yang, G.X., Li, X. and Snyder, M. (2012) Investigating metabolite-protein interactions: an overview of available techniques. *Methods*, **57**, 459-466.
214. Piazza, I., Kochanowski, K., Cappelletti, V., Fuhrer, T., Noor, E., Sauer, U. and Picotti, P. (2018) A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell*, **172**, 358-372.e323.
215. Mentch, S.J., Mehrmohamadi, M., Huang, L., Liu, X., Gupta, D., Mattocks, D., Gómez Padilla, P., Ables, G., Bamman, M.M., Thalacker-Mercer, A.E. *et al.* (2015) Histone Methylation Dynamics and Gene Regulation Occur through the Sensing of One-Carbon Metabolism. *Cell Metab*, **22**, 861-873.
216. Casciello, F., Windloch, K., Gannon, F. and Lee, J.S. (2015) Functional Role of G9a Histone Methyltransferase in Cancer. *Front Immunol*, **6**, 487.
217. Stabell, M., Eskeland, R., Bjørkmo, M., Larsson, J., Aalen, R.B., Imhof, A. and Lambertsson, A. (2006) The *Drosophila* G9a gene encodes a multi-catalytic histone methyltransferase required for normal development. *Nucleic Acids Res*, **34**, 4609-4621.

218. Tie, F., Furuyama, T. and Harte, P.J. (1998) The Drosophila Polycomb Group proteins ESC and E(Z) bind directly to each other and co-localize at multiple chromosomal sites. *Development*, **125**, 3483-3496.
219. Tie, F., Furuyama, T., Prasad-Sinha, J., Jane, E. and Harte, P.J. (2001) The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. *Development*, **128**, 275-286.
220. Jones, C.A., Ng, J., Peterson, A.J., Morgan, K., Simon, J. and Jones, R.S. (1998) The Drosophila esc and E(z) Proteins Are Direct Partners in Polycomb Group-Mediated Repression. *Mol. Cell. Biol.*, **18**, 2825-2834.
221. Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, **111**, 185-196.
222. Murali, M. and Saloura, V. (2022) Understanding the Roles of the NSD Protein Methyltransferases in Head and Neck Squamous Cell Carcinoma. *Genes (Basel)*, **13**.
223. Bell, O., Wirbelauer, C., Hild, M., Scharf, A.N., Schwaiger, M., MacAlpine, D.M., Zilbermann, F., van Leeuwen, F., Bell, S.P., Imhof, A. *et al.* (2007) Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. *Embo j*, **26**, 4974-4984.
224. Uchiyama, N., Dougan, D.R., Lawson, J.D., Kimura, H., Matsumoto, S.-i., Tanaka, Y. and Kawamoto, T. (2017) Identification of AHCY inhibitors using novel high-throughput mass spectrometry. *Biochemical and biophysical research communications*, **491**, 1-7.
225. Horiuchi, K.Y., Eason, M.M., Ferry, J.J., Planck, J.L., Walsh, C.P., Smith, R.F., Howitz, K.T. and Ma, H. (2013) Assay development for histone methyltransferases. *Assay Drug Dev Technol*, **11**, 227-236.
226. Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res*, **21**, 381-395.
227. Zhao, L., Zhao, J., Zhong, K., Tong, A. and Jia, D. (2022) Targeted protein degradation: mechanisms, strategies and application. *Signal Transduction and Targeted Therapy*, **7**, 113.
228. Niu, Y., Bai, J. and Zheng, S. (2018) The Regulation and Function of Histone Methylation. *Journal of Plant Biology*, **61**, 347-357.
229. Lee, C.H., Yu, J.R., Kumar, S., Jin, Y., LeRoy, G., Bhanu, N., Kaneko, S., Garcia, B.A., Hamilton, A.D. and Reinberg, D. (2018) Allosteric Activation Dictates PRC2 Activity Independent of Its Recruitment to Chromatin. *Mol Cell*, **70**, 422-434.e426.
230. Syed, S., Wilky, H., Raimundo, J., Lim, B. and Amodeo, A.A. (2021) The nuclear to cytoplasmic ratio directly regulates zygotic transcription in *Drosophila* through multiple modalities. *Proceedings of the National Academy of Sciences*, **118**, e2010210118.
231. Chetwynd, A.J., Zhang, W., Faserl, K., Thorn, J.A., Lynch, I., Ramautar, R. and Lindner, H.H. (2020) Capillary Electrophoresis Mass Spectrometry Approaches for Characterization of the Protein and Metabolite Corona Acquired by Nanomaterials. *J Vis Exp*.
232. Zhang, W., Guled, F., Hankemeier, T. and Ramautar, R. (2019) Utility of sheathless capillary electrophoresis-mass spectrometry for metabolic profiling of limited sample amounts. *Journal of Chromatography B*, **1105**, 10-14.
233. Tyanova, S., Temu, T., Sinitsyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods*, **13**, 731-740.

8 Acknowledgments

No researcher is an island. I am grateful for a life full of connection and community that helped me through this journey.

I have many to thank and I want to start with Axel – a role model in creativity, flexibility and enthusiasm. You have taught me so much about science, encouraged my curiosity and given me a lot of room to grow, as a scientist but also as a person. You have shown me how to lead well by being human and compassionate. Thank you for giving me the opportunity for pursuing the projects and my PhD and for supporting me through these challenging years.

Vielen Dank Mama und Papa. Ohne euch wäre das Alles nicht möglich. Vielen Dank für eure Unterstützung und Liebe, wortwörtlich seit Stunde eins. Danke, dass ihr mich immer ermutigt habt mich selbst herauszufordern und immer für mich da wart, wenn ich gestolpert bin.

Thank you Alan, you are my rock and anchor. Thank you for your patience, grounding me and holding me close in the storms. To be loved by you means having a tireless cheerleader and a safe haven to return to.

Thank you to Elena and Tisa - you are my mirrors, you let me see myself through our similarities and differences, you are my fanbase, and my safe place. Thank you for loving and supporting me, wherever I go, whatever I do.

The best support system anyone could wish for, each of you could get your own paragraph, maybe in my next thesis. Thank you for you open ears, arms and minds. Alex, Anka, Daan, David, Elli, Eva, Fionan, Finn, Flo, Gabriel, Gabriele, Juri, Lara, Laura, Linda, Michi, Milena, Nikl, Nisi, Nita, Peter, Ralf, Sinah, Sophia, Steffi, Tobi, Svenja, Viola, Vreni.

Thank you to my direct and indirect family, in Munich and in the north.

Thank you to the MolBio department. The PhD is a time full of challenges but it is a lot easier not to go through it alone. Thank you for every open ear, well-meant feedback, good idea, chocolate, borrowed reagent, experimental tip. Our department is really one of a kind, I could not have asked for a better microcosm to be in. Thank you Peter for facilitating this culture to flourish by good leadership. Thank you to the PhD community, I will miss the international dinners, coffees, mensa lunches, afterwork drinks, yoga sessions and retreats. You really made my PhD time unforgettable and every single one of you contributed to my success. Thank you especially to Anna, Elisa, Fotios, Gizem, Iris, Janet, Jay, Nadia, Nik.

Thank you to Elizabeth for being the best graduate school coordinator I have ever met. You are inspiring and a role model. Thank you for all you have done for me and countless others.

Thank you to my collaborators who contributed directly to my work, Ignasi and his unwavering patience, Teresa and her uncomplicated help, Tom my proteasome brother from another parent, Ivan, Valentina and David with their chemical expertise and crafts(wo)manship, Pardis the cytometry expert, Anuroop and Nik for their patience with my coding, Marco with his metabolomics magic and Beyza, my successor, you have been such a big help.

Thank you to the Imhof lab. I have learned so much from you, coming in as a chromatin and mass spectrometry newbie at the start of my PhD. Thank you for encouraging me, bearing with me and my

questions and supporting me. Special thanks to Andrea, Anuroop, Chandni, Ignasi, Marc, Marco Moritz, Rebecca, Shibo and Teresa. Thank you Beyza, for the era vibes, enthusiasm and new ideas. Sharing the project with you and now handing over has been a privilege, you have been a great mentee,

Thank you to all my teachers. All the people who taught me and allowed me to learn throughout the years, in school, university, and life outside of our academic forts.

Thank you to the future: You give me joy, may you believe in yourself as much as we do, Miriam, Viktoria, Maximilian, and Malin, Martha, Levin, Charlotte, Brunhild, Theo and Charlotte, Leonie, Carla, Loki und Ilvy, Charlotte und Gustav, Aron and Salma.

Thank you to coffee, my bike, dancing, and zoom. For Keeping me going when my brain, my body, or the outside world wanted me to stop moving.

Finally, thank you nature, for allowing us to exist and marvel at the beauty surrounding us—the small and big secrets yet for us to investigate, explore, and try to understand. You don't need us but we need you. May we be better at respecting you.

9 List of publications

Kleene, V., Corvaglia, V., Chacin, E., Forne, I., Konrad, D.B., Khosravani, P., Douat, C., Kurat, C.F., Huc, I. and Imhof, A. (2023) DNA mimic foldamers affect chromatin composition and disturb cell cycle progression. *Nucleic Acids Research*, 51, 9629-9642, <https://doi.org/10.1093/nar/gkad681>.

Welk V., Coux O., **Kleene V.**, Abeza C., Truembach D., Eickelberg O. and Meiners S. (2016) Inhibition of the Proteasome Activity Induces Formation of Alternative Proteasome Complexes. *J. Biol. Chem.* doi: 10.1074/jbc.M116.717652.

10 Annex

10.1 Abbreviations

3D	Three dimensional
AHCY	Adenosylhomocysteinase
AP-MS	Affinity purification coupled with mass spectrometry
ATP	Adenosine triphosphate
bp	base pair(s)
BSA	Bovine serum albumin
CE	Capillary electrophoresis
CT-L	Chymotrypsin-like
ddH ₂ O	Double-distilled water
D.mel, <i>Drosophila</i>	<i>Drosophila melanogaster</i>
DNA	Desoxyribonucleic acid
DREX	Preblastoderm <i>Drosophila</i> Embryo extract
dsDNA	Double-stranded desoxyribonucleic acid
EMSA	Electromobility shift assay
FRET	Förster resonance energy transfer
GC	Gas chromatography
GO term	Gene Ontology term
h	Hour(s)
H2A	Histone 2A
H2B	Histone 2B
H3	Histone 3
H4	Histone 4
H4K20me1	Histone 4 lysine 20 monomethylation
HMT	Histone methyl transferase
LC	Liquid chromatography
LB medium	Luria-Bertani medium
min	minutes
MS	Mass spectrometry
ms grade	Mass spectrometry grade
NC	Nuclear cycle
NMR	Nuclear Magnetic Resonance
ORC	Origin Recognition Complex
PTM	Posttranslational modification
RNA	Ribonucleic acid
SAM	S-Adenosyl methionine
SAM-S	S-Adenosyl methionine synthetase
SAH	S-Adenosyl homocysteine
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SD	Standard deviation
SEM	Standard error of the mean
SGD	Salt gradient dialysis
S-Phase	Synthesis phase
ssDNA	Single-stranded desoxyribonucleic acid
TFA	Trifluoroacetic acid
v/v	Volume per volume
w/v	Weight per volume
ZGA	Zygotic genome activation

10.2 Access credentials for the datasets and corresponding analysis scripts (available online)

Dataset Nr.	Dataset name	Platform	Dataset identifier	Username / Password	Analysis workflow	Figures based on the dataset
Proteomes						
1.	Whole preblastoderm <i>Drosophila</i> embryo extract (DREX) proteome	https://www.ebi.ac.uk/psd/	PXD052592	Username: reviewer_pxd052592@ebi.ac.uk Password: 7cdmyXnhj2w	Raw data RefMG01_MQ2200_MBR_20230307 → MQ → file: "proteingroups" → Perseus → DREX-proteome_Thesis_Vera_RefMG01_MQ2200_MBR_20230307.sps → DREX proteome matrix for downstream R analysis → Excel → "DREX proteome" → R → degGO_analysis_Vera	Figure 22, Figure 23, Figure 24, Figure 25
2.	<i>Drosophila</i> embryo extract assisted <i>in vitro</i> chromatin assembly in the presence of 16mer foldamer	https://www.ebi.ac.uk/psd/	PXD039966	Published, open access		Figure 11, Figure 26
3.	DNA mimicking 32mer foldamers interfere with <i>in vitro</i> chromatin assembly in <i>Drosophila</i> embryo extract	https://www.ebi.ac.uk/psd/	PXD040158	Published, open access		Figure 11, Figure 12, Figure 15
4.	Pulldown with DNA mimicking foldamers in preblastoderm <i>Drosophila</i> embryo extract	https://www.ebi.ac.uk/psd/	PXD040157	Published, open access		Figure 14, Figure 15, Figure 20
5.	Subcellular fractionation of <i>Drosophila</i> S2 cells upon foldamer treatment	https://www.ebi.ac.uk/psd/	PXD042288	Published, open access		Figure 17, Figure 18

Dataset Nr.	Dataset name	Platform	Dataset identifier	Username / Password	Analysis workflow	Figures based on the dataset
6.	DREX chromatin assembly with proteasome inhibition	https://www.ebi.ac.uk/pride/	PXD052215	Username: reviewer_pxd052215@ebi.ac.uk Password: o8F8hoQD	Raw data Ref 2688 →MQ→"proteinGroups_Ref2688_VK_20191113.txt" →Perseus→"Ref2688_VK.sps" →Excel→"Matrix147_complete.txt" →"Matrix147_top100_high_low" →R→ degGO_analysis_Vera_from_Annotation_20102023.R" and "points_waterfall_Plot_the_sis19102023.R"	Figure 29, Figure 30
7.	DREX chromatin assembly +/- dialysis +/- SAH	https://www.ebi.ac.uk/pride/	PXD053067	Username: reviewer_pxd053067@ebi.ac.uk Password: zxuCXTPoBLQ	Raw data Ref6898 and Ref7083 →MQ→"proteinGroups_Ref6898.txt" and "ProteinGroups_Ref7083" →Perseus→ "REF6898_Ref7083_combined_VK98_VK101_102_perseus" →"Matrix52"	Figure 36
Histone modification proteomes						
8.	Histone modifications DREX assembly +/- dialysis +/- SAH +/- C13D3	https://www.ebi.ac.uk/pride/	PXD052593	Username: reviewer_pxd052593@ebi.ac.uk Password: qnEkiHPc1AnP	Raw data Ref4922 →Skyline →"H4_empty_K20heavy_methyl" →Excel→"Peptide Quantification_VK71_VK76_VK77_Axel.xlsx"	Figure 34, Figure 35
Metabolomes						
9.	Metabolite levels in DREX, dialyzed DREX, DREX +/- DNA, and DREX +/- ATP	https://syncandshare.lrz.de/dl/finyUEjjn6wcvSP54RCjEo/.dir	-	Link provides read-only access	Skyline + R Studio	Figure 32, Figure 33, Figure 37, Figure 38