

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**Massenspektrometrische Untersuchung
epigenetischer Proteine**

Leander Simon Runtsch

aus

Heidelberg, Deutschland

2020

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Thomas Carell betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 10.03.2020

Leander Simon Runtsch

Dissertation eingereicht am 30.11.2020

1. Gutachter: Prof. Dr. Thomas Carell
2. Gutachterin: Prof. Dr. Stefanie Kellner

Mündliche Prüfung am 28.01.2021

Publikationsliste

Im Rahmen dieser Arbeit ist folgende Publikation entstanden:

M. Stadlmeier*, L. S. Runtsch*, F. Streshnev, M. Wühr, T. Carell, „A Click-chemistry based enrichable cross-linker for structural and protein interaction analysis by mass spectrometry“, *ChemBioChem* **2020**, *21*, 103–107.

doi: 10.1002/cbic.201900611

* Beitrag zu gleichen Teilen.

„The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.“

Isaac Asimov

*„This is ten percent luck,
twenty percent skill,
fifteen percent concentrated power of will,
five percent pleasure,
fifty percent pain...“*

aus *Remember The Name* von Fort Minor

Danksagung

Diese Arbeit markiert das Ende eines unglaublich spannenden aber manchmal auch steinigen Weges. Daher möchte ich mich auf den nächsten Seiten bei den vielen Menschen bedanken, die mich auf diesem Weg begleitet und unterstützt haben, sodass ich letztendlich am Ziel angekommen bin.

Mein erster Dank gilt meinem Doktorvater *Prof. Dr. Thomas Carell* für die Betreuung dieser Doktorarbeit, die spannenden Projekte und das hervorragende Forschungsumfeld, welches er geschaffen hat und durch seine ständigen Bemühungen, neue Forschungsmittel zu beschaffen, aufrecht erhält.

Ebenso möchte ich *Prof. Dr. Stefanie Kellner* meinen Dank für die Übernahme des Zweitgutachtens sowie ihr Engagement im IRTG 1309-Umfeld aussprechen. Auch den weiteren Mitgliedern der Prüfungskommission danke ich für ihre Bereitschaft, sich mit dieser Arbeit zu beschäftigen.

Dem *Fonds der Chemischen Industrie* danke ich für die Unterstützung dieser Arbeit durch ein zweijähriges Stipendium.

Allen aktuellen und ehemaligen Mitgliedern des AK Carell, welche ich über die Jahre kennenlernen durfte, möchte ich für die schöne gemeinsame Zeit meinen Dank aussprechen, wobei ich im Folgenden einige Personen besonders hervorheben möchte.

Ein großer Dank gilt *Dr. Markus Müller*, der immer ein offenes Ohr für wissenschaftliche oder organisatorische Fragen hatte, viele gute und wichtige Denkanstöße lieferte und der generell schier unermüdlich dafür sorgt, dass diese Gruppe nicht im Alltags-Chaos versinkt. Auch für das Korrekturlesen dieser Arbeit möchte ich mich bei ihm bedanken.

Weiterhin möchte ich mich bei den restlichen mehr oder weniger festen Mitarbeitern des Arbeitskreises bedanken: bei *Slava Gärtner* für die Regelung aller bürokratischen Angelegenheiten, bei *Kerstin Kurz* für die Postabwicklung sowie bei *Kristof Hufnagel* und *Luis de la Osa de la Rosa* für die Unterstützung bei verschiedenen Experimenten.

Dr. David Eisen und *Dr. Silvia Laube* danke ich für die Einführung in die Welt der Massenspektrometrie und dass ich schon früh selbstständig arbeiten durfte. *Dr. Edris Parsa*, *Dr. Iacovos Michaelides*, *Dr. Franziska Traube*, *Dr. Angie Kirchner*, *Dr. Katharina Iwan* sowie *Sabine Oganessian* danke ich für die Einführung in verschiedene Themen und Techniken sowie die Kooperation bei verschiedenen Experimenten. Bei *Bastien Viverge*, *Eva Korytiaková*, *Dávid Baláž*, *Dilara Özdemir*, *Martin Rossa*, *Markus Hillmeier* und *Dr. Christina Schneider* möchte ich mich für die gemeinsamen Boulder- und Kletter-Abende bedanken, durch die ich wunderbar den Kopf frei bekommen konnte. *Eva Korytiaková* danke ich außerdem für die vielen guten Gespräche, Aufmunterungen und dass sie mich (zumindest zeitweise) motiviert hat, wieder Klavier zu spielen.

Ich danke dem gesamten Labor F4.045 und dabei besonders *Dilara Özdemir*, *Dr. Timm Ensfelder* sowie *Dr. Franziska Traube* dafür, dass sie den Labor- und später auch Büroalltag so angenehm und lustig wie möglich gemacht haben, auch wenn unsere unterschiedlichen Musikgeschmäcker leider verhindert haben, dass sich der Party-Freitag dauerhaft etablieren konnte.

Weiterhin danke ich *Aila Rothweiler* für ihre motivierte und zuverlässige Unterstützung im Rahmen ihrer Bachelorarbeit.

Ein besonderer Dank gebührt *Dr. Michael Stadlmeier*, der mich mit seiner Faszination für die Massenspektrometrie angesteckt hat und mir stets ein geduldiger und motivierender Mentor war. Abgesehen davon möchte ich ihm für die gute Zusammenarbeit, unsere hervorragenden Nerd-Gespräche und die vielen lustigen Stunden danken, die mich über so manches Negativergebnis hinweggetröstet haben.

Meinen Eltern *Thomas* und *Regine* danke ich für so viele Dinge, die sich in beinahe 30 Jahren ereignet haben, und besonders für ihre finanzielle und emotionale Unterstützung in den letzten Jahren. Ebenso gilt ein Dank meinen Geschwistern *Antonia* und *Cornelius*, deren spannende Werdegänge mir stets die Vielfältigkeit des Lebens vor Augen führen, wenn ich wieder einmal in meiner Wissenschaftsblase stecke.

Auch *Flo/Mo*, *(Chris)Tina*, *Jannik*, *Alex*, *Markus/Siegi*, *Fabi* und den zahlreichen anderen Freund*innen, die ich während meines Studiums kennengelernt habe, gilt mein Dank. Sie haben das Studium erst zu dem großartigen Erlebnis gemacht, das es letztendlich war.

Zu guter Letzt geht mein Dank an meine wundervolle Freundin und Seelenverwandte *Sarah*, die es schafft, mir in jeder Situation Kraft zu spenden, mich ständig zum Lachen bringt und mein Leben um so vieles bereichert. Ohne sie wäre ich nicht der Mensch, der ich heute bin.

Inhaltsverzeichnis

| | |
|--|----|
| I Zusammenfassung | V |
| II Summary | IX |
| 1 Einleitung | 1 |
| 1.1 Epigenetik | 1 |
| 1.1.1 Modifikation von DNA-Basen | 1 |
| 1.1.2 Nicht-codierende RNA als Inhibitor von Transkription und Translation | 9 |
| 1.1.3 Histone und die Grundlagen für plastisches Chromatin | 10 |
| 1.2 Massenspektrometrie-basierte Proteinanalyse | 14 |
| 1.2.1 Identifikation von Proteinen mittels Massenspektrometrie | 15 |
| 1.2.1.1 Technische Grundlagen der massenspektrometrischen Analyse | 15 |
| 1.2.1.2 Methoden der Protein-Identifikation – zwei Wege, ein Ziel | 19 |
| 1.2.2 Crosslinking-Massenspektrometrie | 23 |
| 2 Zielsetzung | 29 |
| 3 Ergebnisse und Diskussion | 31 |
| 3.1 Untersuchung von Addukten zwischen DNA-Aldehyden und Histonen | 31 |
| 3.1.1 Addukte zwischen einem Peptid und einem einzelsträngigen ODN | 31 |
| 3.1.2 Addukte zwischen dsDNA und Histonen in rekonstituierten Nukleosomen | 35 |
| 3.1.3 Untersuchung einer möglichen Deformylierung von fC in Nukleosomen | 50 |
| 3.2 Untersuchung einer direkten Interaktion von TET3 und GDH | 56 |
| 3.2.1 Etablierung eines Crosslinking-Protokolls mit DSSO | 56 |
| 3.2.2 Crosslinking-Experimente mit TET3 und GDH | 60 |
| 3.3 Etablierung eines neuen, anreicherbaren Protein-Protein-Crosslinkers | 64 |
| 3.3.1 Optimierung der Crosslinking-Reaktion | 65 |
| 3.3.2 Optimierung der Anreicherung von cliXlink-Crosslinks mithilfe von Biotin | 67 |
| 3.3.2.1 CuAAC-Reaktion auf Peptidebene | 68 |
| 3.3.2.2 CuAAC-Reaktion auf Proteinebene | 72 |
| 3.3.3 Publikation | 78 |
| 3.3.3.1 Zusammenfassung | 78 |
| 3.3.3.2 Autorenbeitrag | 78 |

| | | |
|----------|--|-----|
| 4 | Ausblick | 113 |
| 5 | Material und Methoden | 117 |
| 5.1 | Reagenzien | 117 |
| 5.2 | Geräte | 117 |
| 5.3 | Methoden zum Addukt-Nachweis zwischen fC-haltiger DNA und Histonen | 118 |
| 5.3.1 | Oligonukleotid-Peptid-Addukte | 118 |
| 5.3.1.1 | Synthese von fC/fU-haltigen Oligodesoxynukleotiden | 118 |
| 5.3.1.2 | Reaktion von Peptid und ODN | 120 |
| 5.3.1.3 | Harnstoff-PAGE | 120 |
| 5.3.2 | DNA-Nukleosom-Addukte | 121 |
| 5.3.2.1 | Synthese einer überwiegend fC/fU-modifizierten DNA mittels PCR | 121 |
| 5.3.2.2 | Agarose-Gelelektrophorese | 122 |
| 5.3.2.3 | Nukleosom-Rekonstitution | 123 |
| 5.3.2.4 | Native PAGE | 123 |
| 5.3.2.5 | Reduktion der DNA-Protein-Addukte und Hydroxylamin-Behandlung | 123 |
| 5.3.2.6 | SDS-PAGE von DNA-Protein-Addukten | 124 |
| 5.3.2.7 | In-Gel-Verdau | 125 |
| 5.3.2.8 | Verdau der fC-Histon-Addukte | 125 |
| 5.3.2.9 | DNA-Protein-Verdauprotokoll mit einer Positivkontrolle | 126 |
| 5.3.2.10 | Deformylierungs-Experimente | 127 |
| 5.4 | Crosslinking von TET3 und GDH mithilfe von DSSO | 129 |
| 5.4.1 | Crosslinking-Tests mit BSA | 129 |
| 5.4.2 | Methoden für den Nachweis einer Interaktion zwischen TET3 und GDH | 130 |
| 5.4.2.1 | Expression der Proteine in humanen Zellen und Crosslinking in Zellen | 130 |
| 5.4.2.2 | Anreicherung von eGFP-TET3 bzw. GDH-FLAG und Crosslinking in Lysat | 132 |
| 5.4.2.3 | Anreicherung mittels SEC | 133 |
| 5.5 | Methoden zur Etablierung eines neuen Protein-Protein-Crosslinkers | 134 |
| 5.5.1 | Optimierung der Identifikation von Protein-Protein-Crosslinks | 134 |
| 5.5.2 | Untersuchung der CuAAC-Reaktion auf Peptidebene | 135 |
| 5.5.2.1 | Fluoreszenz-basierte Untersuchung mithilfe von Coumarin-Azid | 135 |
| 5.5.2.2 | Anreicherung und massenspektrometrische Analyse | 136 |
| 5.5.3 | CuAAC-Reaktion auf Proteinebene und Anreicherung | 137 |

| | | |
|-------|--|------------|
| 5.6 | Aufreinigung von enzymatisch verdauten Protein-Proben | 138 |
| 5.6.1 | Aufreinigung mithilfe von <i>StageTips</i> | 138 |
| 5.6.2 | Aufreinigung mithilfe von <i>C₁₈-ZipTips</i> | 139 |
| 5.7 | Massenspektrometrische Methoden | 139 |
| 5.7.1 | <i>nano</i> LC-Tandem-Massenspektrometrie | 139 |
| 5.7.2 | Auswertung der massenspektrometrischen Analysen | 140 |
| 6 | Abkürzungsverzeichnis | <u>141</u> |
| 7 | Literaturverzeichnis | <u>145</u> |
| 8 | Anhang | <u>189</u> |
| 8.1 | Sequenzen der verwendeten Fusionproteine | 189 |
| 8.1.1 | GDH-FLAG | 189 |
| 8.1.2 | eGFP-TET3 _{fl} | 189 |
| 8.1.3 | eGFP-TET3 _{cd} | 190 |
| 8.2 | Parameter der LC-MS/MS-Analysen | 190 |
| 8.3 | Einstellungen der Programme zur Auswertung der MS-Analysen | 195 |

I Zusammenfassung

Der erste Teil dieser Doktorarbeit beschäftigte sich mit der Frage, ob 5-Formylcytosin kovalente Addukte mit Histonen eingehen kann. Die Formylgruppe reagiert in diesem Fall mit dem Amin einer Lysin-Seitenkette eines der Histone, wodurch reversibel ein Imin, eine sogenannte Schiff'sche Base, entsteht. Diese Verknüpfung von DNA und Histon könnte die Chromatinstruktur verändern und damit einen Einfluss auf die Transkription bestimmter Gene ausüben.

Um dies zu untersuchen, wurde zunächst mittels DNA-Festphasensynthese ein kurzer DNA-Einzelstrang synthetisiert, der ein 5-Formylcytosin bzw. ein 5-Formyluracil enthielt. Dieser wurde mit einem kurzen, Lysin-haltigen Peptid inkubiert und sich bildende Addukte durch Reduktion zu einem sekundären Amin stabilisiert. Trotz Einsatz eines Katalysators konnte in diesen Experimenten gelelektrophoretisch gezeigt werden, dass eine Reaktion nur für 5-Formyluracil möglich ist. Um näher an den Verhältnissen *in vivo* zu sein, wurden als Nächstes Versuche mit rekonstituierten Mononukleosomen durchgeführt. Dafür wurde mittels PCR eine DNA hergestellt, welche bis auf die in den Primern enthaltenen Cytosine nur 5-Formylcytosine enthielt. Mit dieser DNA und einer Negativkontrolle aus unmodifizierter DNA konnten *in vitro* Mononukleosomen gebildet werden. Mittels denaturierender Polyacrylamid-Gelelektrophorese konnte zudem gezeigt werden, dass sich für 5-Formylcytosin tatsächlich DNA-Protein-Addukte ausbilden. Für eine direkte Identifikation dieser Addukte mittels Massenspektrometrie wurden sie einem kombinierten DNA- und Protein-Verdau unterzogen. Das dadurch an Peptiden zurückbleibende fdC-Nukleosid konnte jedoch nicht identifiziert werden. Durch die Publikation der Ergebnisse zweier anderer Forschungsgruppen während der Durchführung dieser Experimente konnten die gelelektrophoretisch erhaltenen Ergebnisse bestätigt werden.

Da die Anzahl der zu einem Zeitpunkt vorliegenden Addukte in der Zelle sehr niedrig ist, wurde in einem nächsten Schritt untersucht, ob die Formylcytosin-Histon-Addukte als Intermediat an der aktiven Demethylierung beteiligt sind. Die Demethylierung nach Oxidation zu 5-Formylcytosin oder 5-Carboxylcytosin erfolgt zum einen durch Basenexzisionsreparatur, kürzlich wurde jedoch in dieser Arbeitsgruppe auch ein direkter Weg über einen C-C-Bindungsbruch entdeckt. Ein möglicher Mechanismus beinhaltet die Übertragung der Formylgruppe auf ein Hilfsnukleophil. Als Intermediat könnte dabei die Schiff'sche Base dienen, was letztlich zu der Bildung von Formyllysin führen würde. Um dies zu analysieren, wurden erneut Nukleosomen rekonstituiert und sowohl DNA als auch Histone auf Veränderungen der Formyl-Modifikation im Vergleich zu Nukleosomen mit unmodifizierter DNA untersucht, wobei keine Übertragung der Formylgruppe festgestellt werden konnte. Durch unspezifische Mono- bzw. Dimethylierung der Histon-Lysine ließ sich ein geringer Unterschied in der Deformylierung von fC-modifizierter und unmodifizierter DNA beobachten, was durch Anfertigung weiterer Replikate überprüft werden muss.

Ein weiterer Teil dieser Arbeit behandelte die Isoform 3 der TET-Enzyme (TET3). Diese ist die in Neuronen am stärksten exprimierte Isoform und daher hauptsächlich für die epigenetische Regulation der neuronalen Aktivität auf DNA-Ebene verantwortlich. Da TET-Enzyme abhängig von α -Ketoglutarat als Kofaktor sind, benötigen sie eine ständige Versorgung mit diesem Metaboliten, um dynamische Veränderungen der Methylierungs- sowie Hydroxymethylierungs-Muster und damit der Genexpression gewährleisten zu können. Vorangegangene Studien in dieser Arbeitsgruppe konnten mehrere metabolische Enzyme identifizieren, welche durch TET3-Coimmunpräzipitation angereichert wurden. Darunter befand sich auch die Glutamatdehydrogenase (*Glud1*, GDH), welche die Reaktion von Glutamat zu α -Ketoglutarat sowie deren Rückreaktion katalysiert. Aus diesen Experimenten ergab sich die Fragestellung, ob es sich hierbei um eine direkte Interaktion zwischen TET3 und GDH handelt.

Um dies zu untersuchen, wurde ein Quervernetzungs-Ansatz (Crosslinking) gewählt. Dabei werden Proteine, welche sich in räumlicher Nähe zueinander befinden, kovalent miteinander verknüpft. Somit kann man nach geeigneter Probenvorbereitung mittels MS^2 -Massenspektrometrie der zu Peptiden verdauten Proteine nachvollziehen, welche Proteine miteinander interagierten. Als Crosslinking-Reagenz (Crosslinker) werden dabei meist kleine Moleküle verwendet, welche zwei reaktive Gruppen besitzen. Diese reagieren mit je einem der interagierenden Proteine und sorgen so für die kovalente Verknüpfung zwischen zwei Peptiden (Crosslink) nach dem Protein-Verdau. Zusätzlich können diese Moleküle auch noch weitere Funktionen, wie etwa eine massenspaltbare Gruppe, besitzen.

Für die Experimente wurden beide Proteine als Fusionsproteine mit Tags zur Anreicherung in HEK-Zellen überexprimiert und anschließend mithilfe des literaturbekannten Crosslinkers Disuccinimidylsulfoxid (DSSO) quervernetzt. Die Probenkomplexität wurde entweder über eine Anreicherung mithilfe der fusionierten Tags oder mittels Größenausschluss-Chromatographie (SEC) verringert. Dabei wurden verschiedene Ansätze untersucht, bei denen das Crosslinking entweder an den intakten Zellen oder zu verschiedenen Zeitpunkten des Arbeitsablaufs durchgeführt wurde. Die Analyse der Proben erfolgte schließlich mithilfe von MS^2 -Massenspektrometrie. Mit diesen Experimenten konnte keine direkte Interaktion zwischen TET3 und GDH nachgewiesen werden, was nach neueren Erkenntnissen in dieser Arbeitsgruppe daran liegen könnte, dass die mitochondriale Transportsequenz von GDH, welche in HEK-Zellen größtenteils abgeschnitten wird, für die Interaktion mit TET3 von Bedeutung ist.

Die für diese Experimente angewendete Crosslinking-Massenspektrometrie befindet sich aufgrund ihrer Neuheit noch in einer Phase der Pionierarbeit. Es existieren viele verschiedene Reagenzien und nur für Standard-Verfahren finden einige wenige Moleküle breitere Anwendung. Daher wurde in dieser Arbeit „cliXlink“, ein neuer, in einer vorangegangenen Studie entwickelter und synthetisierter Crosslinker, etabliert.

Dieser kann zum einen für routinemäßige Anwendungen verwendet werden, bietet jedoch auch die Möglichkeit der modularen Anpassung an spezielle Anforderungen. Das Reagenz basiert auf der Struktur von DSSO: es besitzt zwei Succinimidylester, welche primär mit den Aminen von Lysin-Seitenketten reagieren, sowie ein Sulfoxid als massenspaltbare Gruppe. Letztere erleichtert die Identifikation von Crosslinks in MS²-Experimenten und ermöglicht zusätzlich die Analyse mittels MS³. Darüber hinaus enthält das Molekül ein Alkin, welches mittels Kupfer-katalysierter Alkin-Azid-Cycloaddition (CuAAC) funktionalisiert werden kann. Dadurch kann der Crosslinker um eine beliebige Funktionalität erweitert werden, insofern an diese ein terminales Azid geknüpft wurde.

Um cliXlink zu etablieren, wurde zunächst das Crosslinking-Protokoll optimiert. Dafür wurden sowohl Schritte in der Probenvorbereitung als auch die massenspektrometrische Analyse verbessert. Als Testsystem bei all diesen Versuchen diente aufgereinigtes Rinderserum-Albumin (BSA), welches als Standard für solche Experimente gilt. So konnte die Anzahl der mit cliXlink identifizierten Crosslinks deutlich gesteigert werden. Der Vorteil der Funktionalisierbarkeit des Reagenzes wurde beispielhaft anhand einer Anreicherung der Crosslinks gezeigt. Dies ist in komplexen Proben oftmals nötig, da die quervernetzten Peptide verglichen mit unmodifizierten Peptiden meist eine geringe Abundanz aufweisen und daher mit intensitätsabhängigen Analysemethoden oft nicht analysiert werden. Als Anreicherungs-methode wurde dabei die Wechselwirkung zwischen Biotin und Streptavidin ausgenutzt, indem ein Biotin-Azid und an Partikel immobilisiertes Streptavidin eingesetzt wurde. Das Biotin-Azid war zusätzlich reaktiv spaltbar, um eine Elution der angereicherten Spezies unter milden Bedingungen zu ermöglichen. Für die Anreicherung mussten zuerst Bedingungen gefunden werden, unter denen die CuAAC-Reaktion möglichst quantitativ abläuft, damit so viele Crosslinks wie möglich modifiziert werden und damit anreicherbar sind. Durch eine Reaktion auf Proteinebene in einem speziellen Puffer und anschließender Abtrennung von unreaktiertem Biotin-Azid war es möglich, eine erhebliche Anreicherung von BSA-Crosslinks aus einem komplexen Hintergrund von unmodifizierten Peptiden aus HEK-Proteinlysate zu erreichen.

II Summary

The first part of this thesis dealt with the question whether 5-formylcytosine can form covalent adducts with histones. In this case, the formyl group reacts with an amine of a lysine side chain of one of the histones whereby an imine, a so-called Schiff base, is generated reversibly. This linkage of DNA and histone could change the chromatin structure and thereby have an influence on the transcription of certain genes.

To investigate this, first, a short DNA single strand carrying a 5-formylcytosine or 5-formyluracil was synthesized by DNA solid phase synthesis. This strand was incubated with a short, lysine-containing peptide and the generated adducts were stabilised *via* reduction to secondary amines. Despite the use of a catalyst, it could be shown using gel electrophoresis that a reaction is only possible for 5-formyluracil. To be closer to the conditions *in vivo*, experiments with reconstituted nucleosomes were carried out in a next step. For this, a DNA strand which, besides the primer cytosines, contained only 5-formylcytosines was synthesized by PCR. This DNA together with a negative control consisting of an unmodified strand was used to generate mononucleosomes *in vitro*. Using denaturing polyacrylamide gel electrophoresis, it could be shown that DNA-protein adducts are indeed generated with 5-formylcytosine. For a direct identification of these adducts by mass spectrometry they were subjected to a combined DNA and protein digest. However, the resulting fdC nucleoside modification remaining on the histone peptides could not be identified. Through publication of the findings from two different research groups during the execution of these experiments, the results obtained by gel electrophoresis could be confirmed.

Since in a cell the quantity of adducts present at the same point in time is very low, next, an involvement as an intermediate in active demethylation was examined as a potential function of the covalent link between 5-formylcytosine and histones. The demethylation after oxidation to 5-formyl- or 5-carboxycytosine mainly takes place through base excision repair, however, recently also a direct path *via* a C-C bond cleavage was discovered in this research group. A possible mechanism contains the transfer of the formyl group to an auxiliary nucleophile. Herein, the Schiff base could serve as an intermediate leading to the formation of formyllysine. In order to investigate this, nucleosomes were reconstituted once again and DNA as well as histones were examined regarding changes in the formyl modification compared to nucleosomes with unmodified DNA. Here, no formyl transfer could be observed. Through unspecific mono- and dimethylation of the histone lysines, a small difference in deformylation between the fC-modified and unmodified strands could be observed. These findings need to be validated by preparing additional replicates.

Another part of this work was concerned with isoform 3 of the TET enzymes (TET3). In neurons, this is the most highly expressed isoform and thus is mainly responsible for epigenetic regulation of neuronal activity on DNA level.

Because TET enzymes are dependent on α -ketoglutarate as a cofactor, they need a constant supply of this metabolite to provide for dynamic changes in methylation and hydroxymethylation patterns and consequently in gene expression. Previous studies in this research group identified several metabolic enzymes which are enriched in TET3 co-immunoprecipitation. Among these was glutamate dehydrogenase (*Glud1*, GDH) which catalyses the reaction of glutamate to α -ketoglutarate as well as its back reaction. From these experiments, the question whether these findings were based on a direct interaction between TET3 and GDH arised.

To analyse this, a crosslinking approach was chosen. Here, proteins which are in close proximity to each other are covalently linked. That way, after sample preparation and performing a mass spectrometric analysis of the peptides derived from a digest of the proteins, protein interactions can be reconstructed. As crosslinkers, generally small molecules which carry two reactive groups are used.

These react with one of the interacting proteins each, thus making for the covalent linkage between two peptides (crosslink) after protein digest. In addition, these molecules can also contain secondary functions, for example a group that is cleavable within the mass spectrometer.

For this, both proteins were overexpressed as fusion proteins with enrichment tags in HEK cells following crosslinking using the literature-known crosslinker disuccinimidylsulfoxide (DSSO). The sample complexity was reduced either through enrichment with the fused tags or *via* size exclusion chromatography (SEC). Various approaches performing the crosslinking on intact cells or at different points of the workflow were performed. Finally, the sample analysis was carried out using MS² mass spectrometry. In these experiments, a direct interaction between TET3 and GDH could not be proven, which, according to recent findings in this group, could be due to the importance of the mitochondrial targeting sequence of GDH for the interaction with TET3 that is mostly cleaved off in HEK cells.

Crosslinking mass spectrometry, the technique used for these experiments, is still in a pioneering phase. There are many different reagents and only for standard applications a few of them have reached more widespread use. Therefore, crosslinking using “cliXlink”, a novel crosslinker designed and synthesized in a previous work, was established. This reagent can be used for routine practice but additionally offers the possibility of modular adaptation to special requirements. The crosslinker is based on the structure of DSSO: it features two succinimidyl esters, which primarily react with the amines of lysine side chains, and a sulfoxide as cleavable group in the mass spectrometer. The latter eases identification of crosslinks in MS² experiments and enables analysis *via* MS³.

Beyond that, the reagent contains an alkyne that can be functionalised with copper-catalysed alkyne-azide cycloaddition (CuAAC). Thereby, the crosslinker can be augmented by any user-defined functionality if an azide was attached to it beforehand.

To establish cliXlink, the crosslinking protocol was optimized first. For this, steps in sample preparation as well as the mass spectrometric analysis were improved. In all these tests, purified bovine serum albumin (BSA), which is regarded as a standard for such experiments, was used as test system. By doing this, the number of identified crosslinks using cliXlink could be drastically enhanced. The advantage of functionalization of the reagent was shown by carrying out an enrichment of crosslinks as an example. This is often necessary for complex samples because the crosslinked peptides usually possess a low abundance compared to unmodified peptides and are therefore frequently not analysed with intensity-dependent analytical methods. The interaction between biotin and streptavidin was employed as method of enrichment by using a biotin azide and immobilized streptavidin. In addition, the biotin azide was cleavable by reduction allowing for mild elution of the enriched species. For the enrichment, conditions under which the CuAAC is proceeding rather quantitatively had to be found first in order to achieve modification and thus enrichment of as many crosslinks as possible.

Through a reaction at the protein level in a special buffer following separation of unreacted biotin azide, a substantial enrichment of BSA crosslinks from a complex background of unmodified peptides derived from HEK protein lysate was achieved.

1 Einleitung

1.1 Epigenetik

Die Bedeutung des Begriffs „Epigenetik“ ist in der wissenschaftlichen Welt seit seiner Einführung eine umstrittene Angelegenheit.^[1] Im Jahre 1942 von *Conrad Waddington* erdacht, umfasst dessen ursprüngliche Definition die Veränderungen eines Phänotyps, welche nicht mit einem veränderten Genotyp einhergehen.^[2] Wie wir heute wissen, beruht diese Beobachtung auf Änderungen in der Expression von Genen, nicht jedoch in deren Sequenz. In jüngerer Vergangenheit kamen hauptsächlich zwei unterschiedliche Auslegungen auf: zum einen eine striktere Eingrenzung, welche die Vererbbarkeit dieser nicht mutationsbedingten Änderungen der Genexpression voraussetzt^[3-5] und zum anderen eine weiter gefasste Bezeichnung, die alle strukturellen Änderungen von Chromatin, welche mit der Regulation von Genaktivität zu tun haben, miteinbezieht.^[6] Allen Definitionen gemeinsam ist, dass die epigenetisch verstandenen Veränderungen der Genaktivität nicht durch Änderungen der Abfolge der DNA-Basen des Genoms, Adenin (A), Guanin (G), Thymin (T) sowie Cytosin (C), zustande kommen. Da diese Mechanismen in allen Bereichen des zentralen Dogmas der Molekularbiologie zu finden sind, kann man die Epigenetik in dessen drei verschiedene Ebenen von DNA, RNA und Proteinen unterteilen. Diese Teilbereiche werden im Folgenden genauer beschrieben.

1.1.1 Modifikation von DNA-Basen

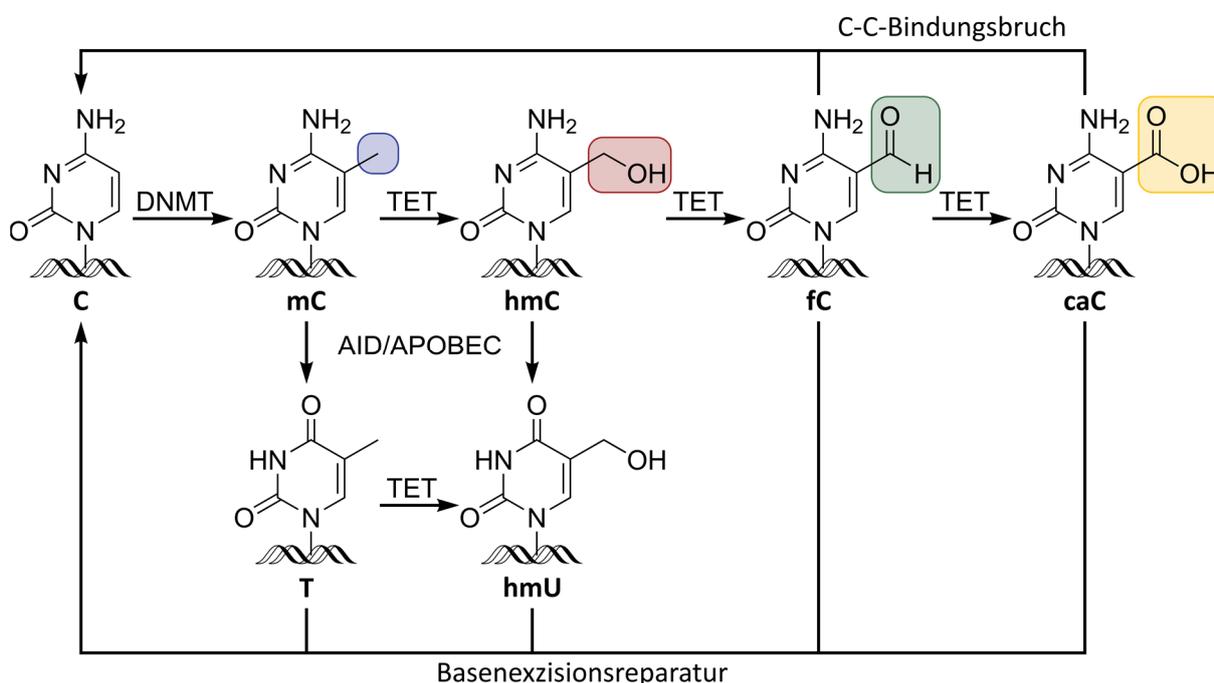
Zwar behandelt die Epigenetik keine Veränderungen der DNA-Sequenz an sich, dennoch kann durch kovalente Modifikationen der DNA-Basen auch auf Ebene der DNA ein epigenetischer Einfluss auf die Expression von Genen gewirkt werden. Im Gegensatz zu der hohen Anzahl bisher entdeckter Modifikationen der RNA (zum Zeitpunkt dieser Arbeit zählte die Datenbank *Modomics*^[7] 172^[8] RNA-Modifikationen), ist die Vielfalt der DNA-Modifikationen um einiges geringer. So beinhaltet die Datenbank *DNAMod*^[9] lediglich 49^[10] Modifikationen, von denen beinahe die Hälfte Methylierungen sind. Zudem entstehen viele DNA-Modifikationen als Folge von DNA-Schädigungen, wie z. B. der Oxidation durch reaktive Sauerstoffspezies.^[11] Somit ist nur für einen sehr geringen Teil aller DNA-Modifikationen eine epigenetische Funktion bekannt.

Eine dieser DNA-Modifikationen ist N^6 -Methyladenin. Diese Base tritt vor allem in Prokaryoten häufig auf,^[12, 13] wo sie eine Funktion bei der Abwehr von Bakteriophagen^[14] aber auch bei der Regulation der Transkription bestimmter Gene,^[15-19] der DNA-Replikation^[20] sowie der Reparatur von Replikationsfehlern^[21] besitzt. Mittlerweile gibt es Literatur, in der auch vom Vorhandensein von N^6 -Methyladenin in der DNA von einzelligen Eukaryoten berichtet wird,^[22] wo es, wie z. B. in *C. reinhardtii* als Markierung aktiver Gene,^[23] epigenetische Aufgaben haben soll. Studien, dass diese Base in den höheren Organismen *C. elegans* und *D. melanogaster* vorkommen soll, sind ebenfalls bekannt.^[24, 25]

In Wirbeltieren ist die Modifikation jedoch allenfalls in sehr geringen Mengen vorhanden.^[26-29] Nichtsdestotrotz ergaben Studien auch hier einen epigenetischen Einfluss dieser Base bei der Organisation von Chromosomen^[30] oder der Genexpression.^[31-33] Auch über eine Mitochondrien-spezifische Funktion^[29, 34] wurde berichtet. Zudem wurde die Oxidation zu *N*⁶-Hydroxymethyladenin^[35] nachgewiesen. Allerdings herrscht Unklarheit über die Aussagekraft dieser Ergebnisse, da eine kürzlich veröffentlichte Studie mehrere Möglichkeiten der Kontamination mit exogenem *N*⁶-Methyladenin und Limitationen der Analysemethoden als mögliche Gründe anführt. Dies führt zu Artefakten, die die ermittelte Menge der Modifikation höher erscheinen lassen, als sie eigentlich ist.^[36] Außerdem konnten *Mushev et al.* zeigen, dass der Großteil an genomischem *N*⁶-Methyladenin aus RNA stammt und fälschlicherweise in die DNA eingebaut wird.^[36] Es ist daher noch unklar, ob diese Base in Eukaryoten tatsächlich im Genom vorkommt bzw. dort eine Funktion besitzt.

Die in den meisten Eukaryoten hauptsächlich vorkommende und daher bereits ausgiebig untersuchte, epigenetisch relevante DNA-Modifikation ist 5-Methylcytosin (mC),^[37, 38] welche im Folgenden zusammen mit seinen oxidierten Formen 5-Hydroxymethylcytosin (hmC), 5-Formylcytosin (fC) sowie 5-Carboxylcytosin (caC) genauer beleuchtet wird (s. Schema 1.1). Die DNA-Base mC wird aufgrund ihrer hohen Abundanz von bis zu ca. 4 – 5 % im Genom einiger Säugetiere^[39] auch als fünfte Base bezeichnet. Die Methylierung ist dabei nicht gleichmäßig über das Genom verteilt, sondern tritt zum Großteil sequenzspezifisch auf. In den meisten differenzierten Säugetier-Zellen ist es größtenteils die Abfolge 5'-Cytosin-Guanin-3' (CpG), welche modifiziert wird,^[40] während in Pflanzen,^[41-46] embryonalen Stammzellen (ESCs)^[40, 47] sowie während der neuronalen Entwicklung^[48] auch andere Sequenzen methyliert sein können. CpGs sind in Organismen mit hohem mC-Gehalt im Vergleich zu anderen Dinukleotid-Kombinationen selten und man geht davon aus, dass dies an einer im Laufe der Evolution akkumulierten Desaminierung von mC zu T liegt.^[49, 50] Von den noch existierenden CpG-Stellen sind z. B. im humanen Genom ungefähr 70 – 80 % methyliert.^[40] Die restlichen CpGs bilden tausende sogenannte CpG-Inseln (CGIs): Bereiche von durchschnittlich 1000 Basenpaaren mit einem vergleichsweise hohen Gehalt an CpG-Dinukleotiden.^[51, 52] Diese Abschnitte überlappen beim Menschen mit den Promotor-Regionen von 60 – 70 % aller Gene, darunter nahezu sämtliche konstitutiv exprimierten Gene, wo sie größtenteils unmethyliert vorliegen.^[53, 54] Daher wird angenommen, dass sie beim Start der Transkription eine Rolle spielen.^[55, 56] Die meisten Promotoren bleiben im Laufe der Differenzierung einer Zelle unmethyliert und werden über andere Mechanismen, wie *Polycomb*-Komplexe, dynamisch reguliert.^[57, 58] Es konnten jedoch mittlerweile einige Gene mit CGI-haltigen Promotoren identifiziert werden, die mittels DNA-Methylierung inaktiviert werden, z. B. geprägte^[59, 60] oder nach der Differenzierung einer Stammzelle nicht mehr benötigte^[61-64] Gene. Es gilt als wahrscheinlich, dass die Methylierung nicht der erste Schritt in diesem Prozess ist, da Startpunkte der Transkription bei aktiven Genen meist arm an Nukleosomen sind.

Zudem besitzen in der Nähe des Startpunkts befindliche Nucleosomen bestimmte Histonmodifikationen, die eine Methylierung erschweren.^[65] Für die X-Chromosom-Inaktivierung konnte z. B. gezeigt werden, dass die DNA-Methylierung für den Beginn der Inaktivierung nicht notwendig ist.^[66, 67] Anschließend ist sie jedoch essenziell, da sonst eine Reaktivierung der Genexpression möglich ist.^[68] Weiterhin gibt es Studien in Krebszelllinien, die zeigen, dass CGI-haltige Gene, die bereits durch *Polycomb*-Komplexe transient inaktiviert wurden, leichter methyliert werden als andere Gene.^[69, 70] Die Verhinderung der Transkription nach der Methylierung erfolgt entweder durch direkte Inhibition des Bindens von Transkriptionsfaktoren^[71, 72] oder über Proteine mit einer Methyl-Bindedomäne, die weitere Chromatin-modifizierende Enzyme rekrutieren.^[73-77]



Schema 1.1: Übersicht von Methylierung und verschiedenen Demethylierungsmechanismen von Cytosin in DNA.

Wie bereits erwähnt, ist ein Großteil der nicht in CGIs auftretenden CpGs, welche sich größtenteils in den transkribierten Bereichen von Genen befinden, methyliert. Als Funktionen dieser intragenen Methylierung wurden die Stilllegung repetitiver Elemente wie Transposons^[78-80] sowie eine Rolle bei alternativem Spleißen^[81] postuliert.

Die Methylierung von Cytosin wird durch die sogenannten DNA-Methyltransferasen (in Säugetieren DNMTs genannt) unter Einsatz des Cofaktors S-Adenosylmethionin (SAM) katalysiert.^[82] Unterschieden werden die *maintenance*-Methylierung von hemimethylierter sowie die *de novo*-Methylierung von unmethylierter DNA. Erstere wird von DNMT1 katalysiert und sorgt für eine Methylierung der bei der DNA-Replikation neu gebildeten hemimethylierten Stränge, um die bestehenden Methylierungsmuster einer Zelle aufrechtzuerhalten.^[83]

Die Rekrutierung an Replikationszentren erfolgt nach neueren Erkenntnissen über eine Interaktion mit dem Protein „E3 Ubiquitin-Proteinligase“ (UHRF1), welches spezifisch an hemimethylierte DNA bindet.^[84, 85] Die *de novo*-Methylasen DNMT3A sowie DNMT3B sind im Gegensatz dazu für die Einführung neuer Methylierungen in unmethylierter DNA zuständig.^[86, 87] Dies ist zum einen für die bereits beschriebene Stilllegung von Genen bei der Differenzierung, zum anderen aber auch besonders während der frühen Embryogenese wichtig. Bereits kurz nach der Befruchtung^[88, 89] und auch während der Reifung der Urkeimzellen^[90] kommt es zu einer umfassenden Demethylierung großer Bereiche der DNA. Anschließend werden Teile davon durch DNMT3A und DNMT3B remethyliert, wobei das jeweilige Protein in einem tetrameren Komplex mit der katalytisch nicht aktiven Isoform DNMT3L agiert.^[91-93] Dieses erkennt unmethyliertes Lysin an Position vier des Histons H3 und dirigiert die Methylierung somit an bestimmte Stellen des Genoms.^[94]

Obwohl die Unterfamilien der DNMTs hauptsächlich entweder auf die *maintenance*- oder *de novo*-Methylierung spezialisiert sind, scheint es keine klare Abgrenzung zu geben, da mehrere Studien auch kooperative bzw. unterstützende Aktivität beobachten konnten.^[95-99]

Wie bereits erwähnt, findet an verschiedenen Zeitpunkten der Embryogenese eine weitreichende Demethylierung des Keimzell-Genoms, welches einen Methylierungsanteil von ca. 40 % in Oozyten^[100] und bis zu 85 % in Spermatozyten^[101] besitzt, statt. Während die Demethylierung der maternalen Erbinformation nach der Befruchtung durch DNA-Replikation^[102] und der Abwesenheit von DNMT1^[103, 104] passiv vonstattengeht, beginnt der Prozess bei der paternalen DNA bereits vor der ersten Zellteilung und verläuft so schnell, dass ein aktiver Mechanismus zugrunde liegen muss (Abb. 1.1).^[105, 106]

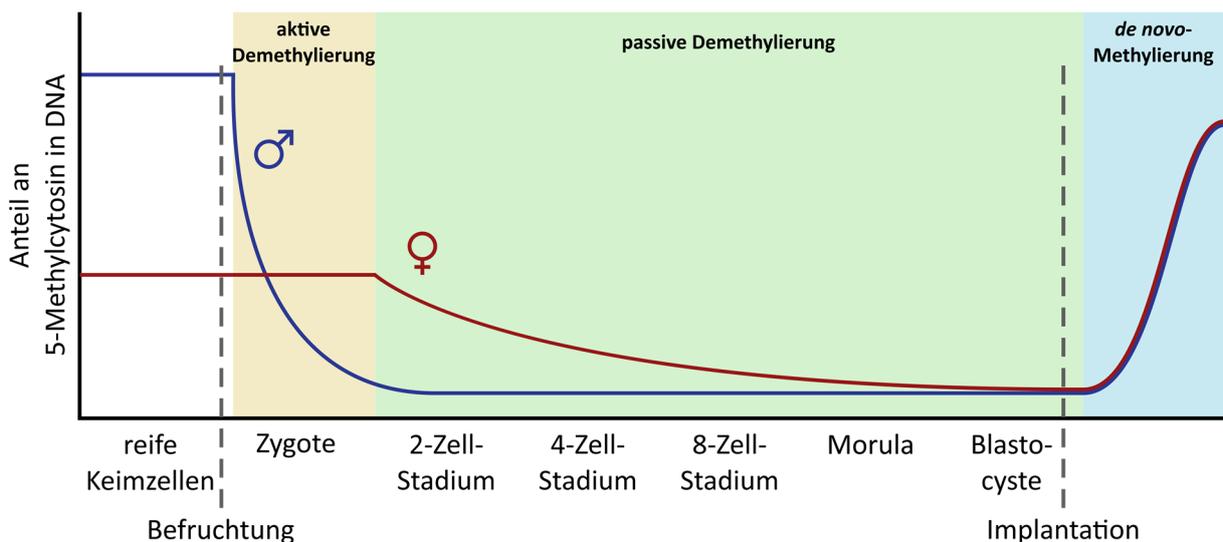


Abbildung 1.1: Unterschiede in der Demethylierung paternaler und maternaler DNA im Verlauf der Embryogenese. Die Demethylierung der maternalen Erbinformation läuft passiv durch Zellteilung ab, wohingegen die paternale DNA bereits vor der ersten Zellteilung aktiv demethyliert wird.

Das maternale Genom ist vor dieser aktiven Demethylierung durch verschiedene DNA-bindende Proteine geschützt.^[107] Auch bei der Reifung der Urkeimzellen scheint die aktive Demethylierung eine Rolle zu spielen.^[90] Die Mechanismen und beteiligten Enzyme dieser aktiven Demodifikation in Säugtieren erfahren großes Interesse in der aktuellen Forschung. Dabei herrscht noch Unklarheit über die Relevanz der verschiedenen Prozesse *in vivo*, es könnte aber sein, dass unterschiedliche Methoden an verschiedenen Zeitpunkten der Entwicklung von Bedeutung sind.

In Pflanzen können DNA-Glykosylasen die mC-Base direkt entfernen und somit für die weiteren Schritte der Basenexzisionsreparatur (BER), nämlich Ausschneiden der abasischen Stelle und Einbau eines unmodifizierten Cytosins, zugänglich machen.^[108-113] Im Gegensatz dazu wurden in Säugetieren noch keine stichhaltigen Belege für solch einen direkten Weg gefunden. Zwar existieren Berichte von schwacher Reaktivität zweier DNA-Glykosylasen („Thymin-DNA-Glykosylase“ (TDG) und „Methyl-CpG-Bindedomänen-Protein 4“ (MBD4)) gegenüber mC *in vitro*, allerdings ist diese Aktivität verglichen mit einer T:G-Fehlpaarungsreparatur weitaus niedriger.^[114-116]

Eine weitere Möglichkeit der aktiven Demethylierung besteht in der Nukleotidexzisionsreparatur (NER), bei der ein Abschnitt von ca. 25 – 30 Nukleotiden ausgeschnitten wird.^[117-119] Studien hierzu zeigen eine NER-abhängige und durch das „Wachstumsstillstands- und DNA-Schaden-induzierbare Protein 45 α “ (GADD45A) vermittelte Demethylierung,^[120, 121] was jedoch von einer anderen Forschungsgruppe angezweifelt wird.^[122] Auch eine Entfernung von lediglich der Methylgruppe, welche für das „Methyl-CpG-Bindedomänen-Protein 2“ (MBD2) berichtet wurde,^[123] konnte nicht bestätigt werden.^[124]

Ebenfalls untersucht wurde die Desaminierung von mC zu T und anschließende Reparatur der entstehenden T:G-Fehlpaarung (s. Schema 1.1). Es konnte gezeigt werden, dass die „Aktivierungs-induzierte Deaminase (AID) und das „Apolipoprotein B mRNA-editierende Enzym 1“ (APOBEC1), welche die Konvertierung von Cytosin zu Uracil (U) katalysieren,^[125, 126] auch mC zu T umwandeln können.^[127] Weitere Studien zeigen eine vergleichbare Funktion von DNMT3A und DNMT3B in humanen Brustkrebszellen, wo manche Promotoren einer zyklischen Methylierung und Demethylierung unterliegen.^[128, 129] Das durch diese Reaktion entstehende T:G-Paar kann anschließend durch BER unter Einfluss der bereits erwähnten DNA-Glykosylasen TDG^[130] und MBD4^[131] oder durch nicht-kanonische *Mismatch*-Reparatur (ncMMR)^[132] repariert werden. AID kann durch C-U-Konvertierung und anschließende *long-patch* BER, bei der 2 – 13 Nukleotide ausgeschnitten werden, auch indirekt den Grad der Methylierung verringern.^[133, 134]

Die heute am besten untersuchte Möglichkeit ist jedoch die schrittweise ablaufende Oxidation von mC zu den Basen hmC^[135-137] sowie fC und caC^[138-140], welche in Metazoen durch die Familie der Methylcytosin-Dioxygenasen (TETs) gebildet werden (s. Schema 1.1).

Dieser Prozess ist für die bereits angesprochene Demethylierung des paternalen Genoms nach der Befruchtung zuständig.^[141] Die oxidierten Basen werden anschließend zu einem großen Teil passiv durch Replikation verdünnt.^[142, 143]

Einige Stellen werden jedoch aktiv in unmethyliertes Cytosin überführt, wobei in der Literatur verschiedene Mechanismen beschrieben sind. Obwohl fC:G und caC:G keine Fehlpaarungen sind, besitzen sie aufgrund der Formyl- bzw. Carboxylgruppe eine verringerte Elektronendichte, was in einer Schwächung der glykosidischen Bindung resultiert.^[144, 145] Daher sind diese Basen Substrate der BER durch TDG^[140, 146] sowie „Nei-like“ (NEIL) 1 und 2.^[147-149] Außerdem ist die Stärke der Basenpaarung von fC und caC mit G geringer als für C,^[150] was dafür sorgen könnte, dass sie leichter in das aktive Zentrum des Enzyms gedreht werden können.^[144] Die Hydroxymethylgruppe von hmC übt diese Effekte nicht aus, weshalb hmC resistent gegenüber DNA-Glykosylasen ist.^[146] Einige Studien beschreiben allerdings die Möglichkeit der vorausgehenden Desaminierung von hmC zu 5-Hydroxymethyluracil (hmU) durch AID oder APOBEC1,^[130, 151] welches anschließend von TDG und MBD4^[130, 152] oder der „Einzelstrang-sensitiven monofunktionalen Uracil-DNA-Glykosylase 1“ (SMUG1)^[153] ausgeschnitten werden kann. Andere Forschungsarbeiten widersprechen dem jedoch, da hmC aus sterischen Gründen kein gutes Substrat für AID ist^[154, 155] und zudem gezeigt werden konnte, dass der Großteil an hmU in ESCs aus der TET-vermittelten Oxidation von Thymin stammt.^[156]

Die bisher beschriebenen Möglichkeiten der Demethylierung beruhen allesamt auf der Bildung von DNA-Strangbrüchen und Einbau neuer Nukleotide. Dies hat den Nachteil, dass bei fehlerhaftem Einbau Mutationen entstehen können. Unlängst konnten *Iwan et al.* jedoch auch einen TDG- sowie TET-unabhängigen, direkten Weg der Demodifikation von fC nachweisen, bei dem die Formylgruppe durch einen C-C-Bindungsbruch abgespalten wird, wodurch das Zucker-Phosphat-Rückgrat der DNA intakt bleibt.^[157] Eine daran anschließende Studie beschreibt die Notwendigkeit der nukleophilen Aktivierung der oxidierten Base, bevor die Bindung gespalten werden kann.^[158] Eine vergleichbare Reaktivität konnte von *Schiesser et al.* auch für caC gezeigt werden,^[159] was vor kurzem durch eine andere Forschungsgruppe bestätigt wurde.^[160]

Neben ihrer Rolle in der TET-abhängigen Methylierung, wurden mittlerweile auch diverse eigene Funktionen der oxidierten Cytosin-Derivate postuliert. Besonders hmC ist in dieser Hinsicht interessant, da es im Gehirn mit bis zu 1 % aller Cytosine sehr häufig vorkommt^[39, 161, 162] und sich im Laufe der Entwicklung von Neuronen dynamisch verändert.^[163, 164] *Bachman et al.* konnten zeigen, dass hmC in ESCs eine größtenteils stabile Modifikation ist.^[165] Mehrere Studien fanden unabhängig voneinander heraus, dass hmC hauptsächlich in *Enhancern* und intragenen Regionen aktiv transkribierter Gene sowie in bivalenten Promotoren, welche sowohl aktivierende als auch inhibierende Histonmodifikationen tragen und dadurch leicht aktiviert werden können, vorkommt.^[166-172]

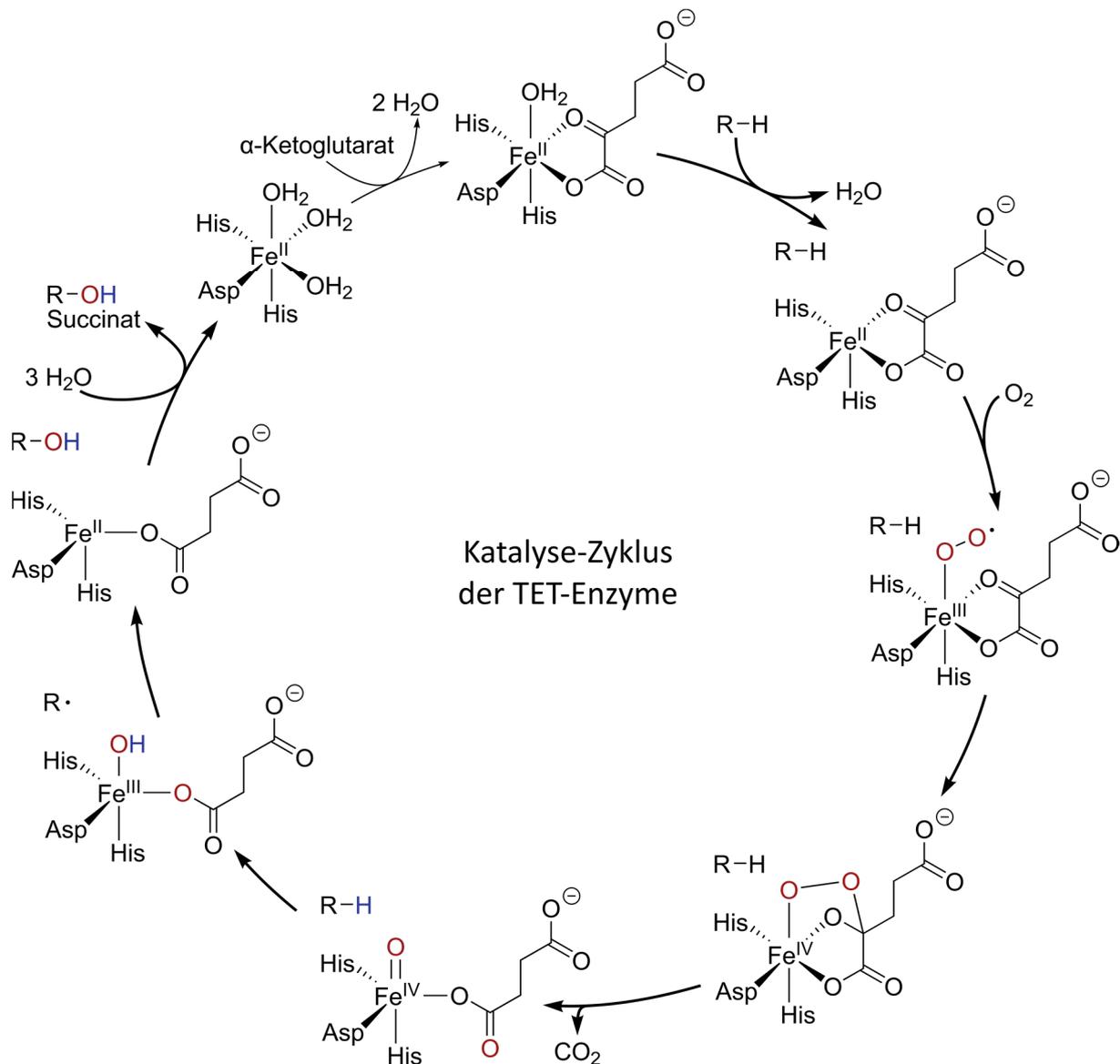
Da mC in solchen Promotoren nicht vermehrt zu finden ist,^[173] spricht dies für eine unterschiedliche Funktion von mC und hmC bei der Regulation von Promotor-abhängiger Genexpression. hmC besitzt insgesamt also wahrscheinlich eine aktivierende Funktion auf die Transkription und scheint besonders für die neuronale Entwicklung wichtig zu sein.

Auch im Fall von fC gibt es Hinweise darauf, dass diese Modifikation nicht nur Intermediat der aktiven Demethylierung ist. So konnte gezeigt werden, dass fC im Gehirn adulter Mäuse eine stabile Modifikation ist^[174] und genomweite Sequenzierungsexperimente zeigten ähnliche Lokalisierungsmuster wie für hmC auf.^[175-177] Die Tatsache, dass sowohl für hmC als auch fC Proteine gefunden wurden, welche präferenziell mit einer der Modifikationen interagieren,^[178, 179] könnte ebenfalls ein Hinweis darauf sein, dass diese oxidierten Cytosin-Derivate eigene epigenetische Funktionen besitzen. Weitere Studien *in vivo* müssen durchgeführt werden, um diese Zusammenhänge noch genauer zu beleuchten.

Wie bereits angesprochen, wird die Oxidation von mC durch die TET-Enzymfamilie katalysiert, welche aus den drei Proteinen TET1, TET2 und TET3 besteht und α -Ketoglutarat, ein Fe^{II}-Ion sowie molekularen Sauerstoff als Cofaktoren bzw. -substrate benötigt.^[136, 137] Die Reaktion folgt dabei dem in Schema 1.2 dargestellten Mechanismus.^[180] Nach der Bildung eines Fe(III)-Komplexes mit Sauerstoff und α -Ketoglutarat wird letzteres decarboxyliert, bevor die O-O-Bindung heterolytisch gespalten wird und eine Fe(IV)-Oxo-Spezies entsteht. Diese kann nun ein Wasserstoffatom der Substrat-Base (mC, hmC, fC) abstrahieren, worauf im Anschluss das entstandene Hydroxylradikal wieder auf die Base transferiert wird. Im Falle der Oxidation von hmC zu fC dehydratisiert das zunächst gebildete *gem*-Diol zum Aldehyd.

Die Expression der TET-Isoformen ist je nach Zelltyp und Entwicklungsstadium unterschiedlich. In Oozyten sowie der befruchteten Zygote ist hauptsächlich TET3 vorhanden und es konnte gezeigt werden, dass dieses nach der Fertilisation zum Großteil im paternalen Pronukleus zu finden ist und dort die bereits angesprochene schnelle Demethylierung des paternalen Genoms verursacht.^[141, 143, 181, 182] Bereits im Zwei-Zell-Stadium nimmt die TET3-Expression jedoch stark ab und in ESCs findet man ausschließlich TET1 und TET2.^[141] TET1 scheint dabei in erster Linie an Promotor-Regionen aktiv zu sein, während TET2 in intragenen Regionen aktiver Gene lokalisiert ist.^[183] Verschiedene Transkriptionsfaktoren, welche für Pluripotenz und Differenzierung wichtig sind, rekrutieren TET1 und TET2 außerdem zu *Enhancern* spezifischer Gene,^[184-186] was für eine Funktion dieser Dioxygenasen im Reifeprozess von Stammzellen spricht. Auch die schon erwähnte Demethylierung in Urkeimzellen ist abhängig von TET1 und TET2, wobei die Oxidation in diesem Fall bei hmC aufhört und anschließend durch Replikation passiv verdünnt wird.^[187, 188] Im Laufe der weiteren Differenzierung von murinen ESCs sinkt die Expression von TET1 und TET2 wieder, während die von TET3 hochreguliert wird.^[29, 136, 137]

All diese Beobachtungen weisen darauf hin, dass TET-Enzyme wichtige Funktionen während der Entwicklung komplexer Organismen ausüben, auch wenn das Ausmaß dieser Bedeutung sowie viele detaillierte Zusammenhänge noch nicht eindeutig geklärt sind. Eine Studie mit TET1/TET2-Doppel-*knock-out*-Mäusen zeigte beispielsweise, dass diese Mäuse zwar einige Defizite besitzen, aber prinzipiell überlebensfähig und fruchtbar sind.^[189]



Schema 1.2: Katalyse-Zyklus der TET-vermittelten Oxidation. Dargestellt ist das aktive Fe^{II}-Zentrum, welches im Enzym von zwei Histidin- (His) und einer Aspartat-Seitenkette (Asp) koordiniert wird.

Besonders bei der Differenzierung und Reifung von Neuronen scheinen TET-Enzyme eine große Rolle zu spielen, da mehrere Studien Defekte der neuronalen Entwicklung von Mäusen und Fröschen bei Verlust der TET-Aktivität zeigen konnten.^[151, 164, 190-192] Auch viele Arten von Krebs scheinen mit einer gestörten Aktivität durch TET-Enzyme zusammenzuhängen.^[193-198] Ursachen sind dabei zum einen Mutationen der TET-Proteine selbst, wie z. B. von TET2 bei myeloischer Leukämie.^[199, 200]

Oft ist jedoch nur der Metabolismus einer Zelle betroffen, sei es durch Mutation von metabolischen Enzymen^[194, 201, 202] oder Onkogenen.^[203-205] Da die TET-katalysierte Oxidation von mC durch ihre Abhängigkeit von α -Ketoglutarat als Cofaktor direkt an den Stoffwechsel der Zelle gekoppelt ist, wird sie von diesem beeinflusst. Wie sich dieses Zusammenspiel im Detail verhält und was dies für zukünftige therapeutische Ansätze bedeuten könnte, ist daher Gegenstand der aktuellen Forschung.^[206-211]

1.1.2 Nicht-codierende RNA als Inhibitor von Transkription und Translation

Die epigenetische Regulation auf RNA-Ebene steht nicht im Fokus dieser Forschungsarbeit. Da sich die verschiedenen epigenetischen Bereiche jedoch auch gegenseitig beeinflussen können, sollen im Folgenden der Übersicht halber kurz grundlegende Mechanismen dieser Art der Regulation von Genexpression erwähnt werden.

Heutzutage weiß man, dass nur ein sehr geringer Teil des Genoms von Eukaryoten (ca. 3 %) für Proteine codiert.^[212] Viele der restlichen, nicht-codierenden Bereiche werden jedoch trotzdem transkribiert und spielen eine Rolle bei der Regulation von zu ihnen komplementärer DNA oder RNA.^[213] Die wichtigsten Arten dieser regulatorischen RNA-Moleküle sind die *micro* (mi), *piwi-interacting* (pi) sowie *long non-coding* (lnc) RNA. Im Falle von miRNA und piRNA liegt dabei dasselbe Prinzip, RNA-Interferenz genannt, zugrunde. miRNA^[214-216] und piRNA^[217-220] wird zunächst aus endogener RNA gebildet, wobei miRNA durch Prozessierung von speziellen Hairpin-Transkripten^[221-223] und piRNA durch Spaltung langer ssRNA^[224-226] gebildet wird. Während miRNA in verschiedenen Zelltypen vorkommt,^[227] ist die Lokalisierung von piRNA hauptsächlich auf Keimzellen beschränkt, wo sie die Expression von Transposons inhibiert.^[228, 229] Anschließend bindet die RNA an Proteine der „Argonaut“ (AGO)-Familie von Endonucleasen.^[230-233] Bereits seit längerem ist bekannt, dass diese Proteine anschließend als Teil des sogenannten *RNA-induced silencing complex* (RISC)^[234] komplementäre mRNA binden und je nach Grad der Komplementarität entweder degradieren oder deren Translation unterdrücken können.^[235] Mittlerweile gibt es jedoch auch Studien, die zeigen, dass mit RNA beladene AGO-Proteine innerhalb des *RNA-induced initiation of transcriptional gene silencing complex* (RITS) bereits cotranskriptional an naszierende mRNA binden können,^[236] was neben der Degradation des Transkripts auch die lokale Bildung von inaktivem Chromatin durch Rekrutierung von Chromatin-Remodellierungsenzymen zur Folge hat.^[237-239]

LncRNAs werden als nicht-codierende RNA-Moleküle mit einer Länge von mehr als 200 Nukleotiden definiert.^[240, 241] Sie interagieren nicht mit AGO-Proteinen, sondern binden entweder an Transkriptionsfaktoren^[242] oder bilden einen DNA-RNA-Triplex.^[243, 244] Dadurch können sie Chromatin-Remodellierungs-Komplexe^[245-248] oder DNA-modifizierende Enzyme, z. B. DNMT3B,^[243] zu bestimmten Bereichen des Genoms dirigieren. Dies spielt unter anderem bei der X-Chromosom-Inaktivierung^[246] oder der Expression von geprägten Genen^[249, 250] eine Rolle.

1.1.3 Histone und die Grundlagen für plastisches Chromatin

Im Gegensatz zu Prokaryoten, bei denen sich die Erbinformation, kondensiert als sogenanntes Nucleoid, im Cytoplasma befindet,^[251, 252] ist das Genom von Eukaryoten durch seine Lokalisierung im Nucleus räumlich vom Rest der Zelle getrennt. Beim Menschen etwa beträgt die Länge der DNA einer diploiden Zelle ungefähr 2 m,^[253] weshalb sie durch Interaktion mit Proteinen als Chromatin um ein Vielfaches kompaktiert wird. Die wichtigsten dabei beteiligten Proteine sind Histone. Diese weisen einen verhältnismäßig hohen Anteil der positiv geladenen Aminosäuren Lysin und Arginin auf, was elektrostatische Wechselwirkungen mit der durch die Phosphat-Gruppen negativ geladenen DNA ermöglicht.^[254, 255]

Die grundlegende Struktur des Chromatins ist das Nucleosom (s. Abb. 1.2). Dieses bildet sich durch Interaktion eines ca. 147 bp langen Abschnitts der DNA mit den vier Kern-Histon-Proteinen H2A, H2B, H3 sowie H4.^[256] Durch Zusammenlagerung von je zwei Kopien aller Histone wird dabei ein Oktamer gebildet, um welches sich die DNA in 1.67 linksgängigen Umdrehungen windet, wodurch eine sogenannte Superhelix entsteht. Der dabei entstehende Komplex besitzt annähernd eine C_2 -Symmetrie,^[256] wobei das Basenpaar der DNA, welches die Symmetrieachse schneidet, als Dyade bezeichnet wird (s. Abb. 1.2). Die dafür notwendige Biegung und Steigung der DNA entsteht durch einen sogenannten *Roll-and-Slide*-Mechanismus: in Abständen von ca. 5 – 6 bp werden je zwei Basenpaare relativ zueinander verschoben und gedreht.^[257] Das führt zu alternierenden Knicken der großen und kleinen Furchen der Doppelhelix, wobei die hierbei sehr gestauchten kleinen Furchen durch Arginine der Histone stabilisiert werden.^[257-259] Die Deformierung ist sequenzspezifisch, da sich die benötigte Energie für die verschiedenen Dinukleotid-Kombinationen unterscheidet.^[259] Dies erklärt die besondere Stabilität von Nucleosomen mit bestimmten DNA-Sequenzen.^[260, 261]

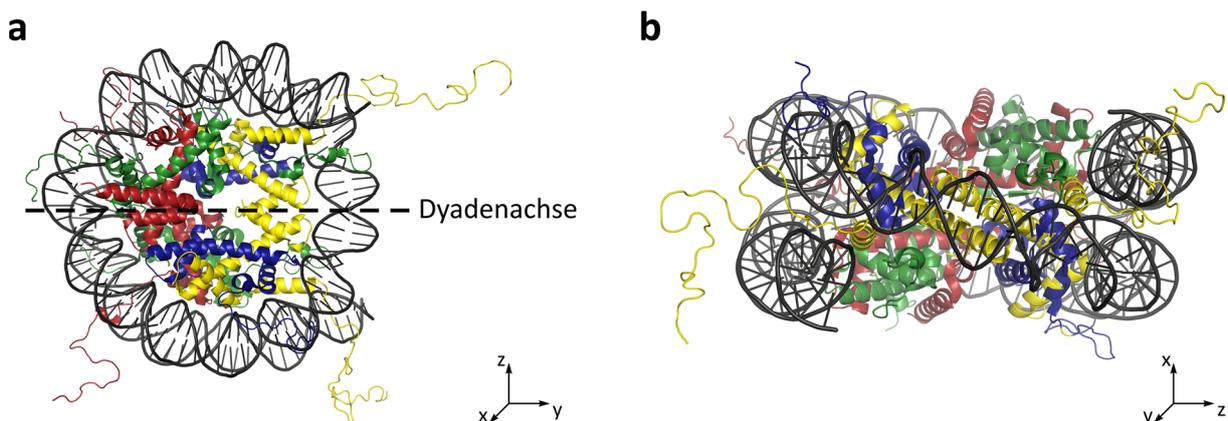


Abbildung 1.2: Kristallstruktur eines Nucleosom-Partikels bestehend aus Histon-Oktamer und einem 147 bp langen DNA-Abschnitt (pdb: 1KX5^[262]; H2A in rot, H2B in grün, H3 in gelb, H4 in blau). a) Draufsicht mit eingezeichneter Dyadenachse. b) Blick auf die Dyadenachse (Transformation ausgehend von a) zuerst -90° um die y-Achse und anschließend -90° um neue x-Achse).

Durch enzymatischen Verdau der DNA in Chromatin konnte gezeigt werden, dass es aus einer Kette von aneinandergereihten Nukleosomen mit Abständen von jeweils ca. 50 bp, auch 11 nm-Faser genannt, besteht (s. Abb. 1.3).^[263] Diese Kette kann durch die Bindung des Histons H1 weiter kondensiert werden, wodurch eine sekundäre Struktur, die sogenannte 30 nm-Faser, entsteht.^[264-266]

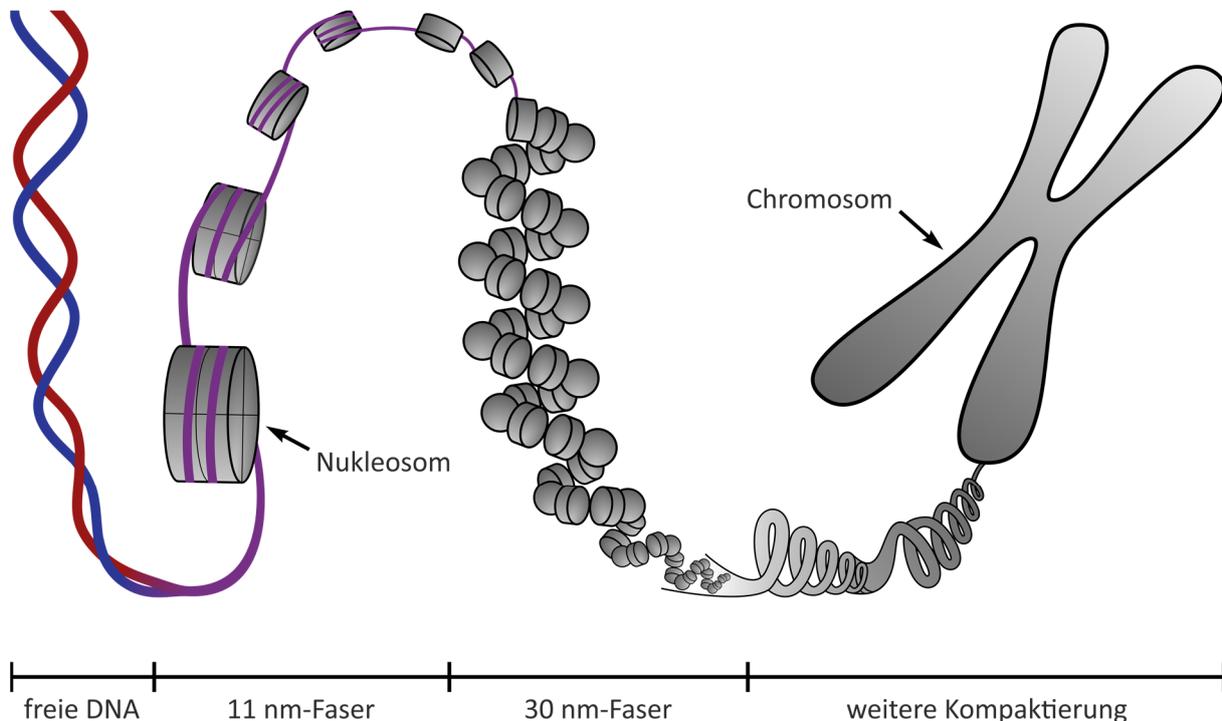


Abbildung 1.3: Kompaktierung eukaryotischer DNA durch Assoziation mit strukturgebenden Proteinen in Chromatin. Die freie DNA bildet zunächst mit Histonen eine Kette aus Nukleosomen, die sogenannte 11 nm-Faser. Durch Anlagerung des Histons H1 erfolgt eine weitere Kondensation in die 30 nm-Faser. Zusätzliche Strukturproteine ermöglichen eine weitere Kompaktierung dieser Faser zu Chromosomen während der Mitose.

H1 interagiert dabei wahrscheinlich asymmetrisch mit der Dyade und entweder dem eingehenden oder ausgehenden DNA-Strang.^[267, 268] Die Bildung der 30 nm-Faser scheint *in vivo* eine große Bedeutung für eine normale Entwicklung zu haben: laut zwei Studien besitzen Mäuse mit *H1-knockout* eine geringere Überlebenschance,^[269] da sich die Expression geprägter oder X-chromosomaler Gene verändert.^[270] Die genaue Struktur dieser Faser konnte bis jetzt jedoch nur *in vitro*^[266, 271-273] und kristallographisch^[274] direkt beobachtet werden. Es gibt zwar indirekte Hinweise auf ihre Existenz zumindest in Teilen des Genoms *in vivo*,^[275-278] allerdings konnten neuere Studien unter Verwendung von Kryo-Elektronenmikroskopie diese nicht bestätigen.^[279, 280] Die während der Mitose benötigte weitere Kondensierung des Chromatins zu Chromatiden erfolgt mithilfe von sogenannten Condensin-Proteinkomplexen,^[281-284] wobei auch hier die dabei entstehenden höhergeordneten Strukturen unklar sind.

Bereits früh fand man heraus, dass bis auf wenige Ausnahmen^[285-287] die Aminosäuren der N-terminalen Regionen der Histone wichtig für die Bildung der höheren Chromatinstrukturen^[288] und für die Zugänglichkeit der DNA bei der Bindung von Transkriptionsfaktoren^[289-292] sind.

Diese Prozesse werden durch kovalente Modifikationen der Aminosäuren-Seitenketten reguliert. Heutzutage sind eine Vielzahl an Histon-Modifikationen, sowie die dazugehörigen modifizierenden Enzyme bzw. Enzym-Komplexe bekannt. Es folgt daher eine Übersicht der wichtigsten Modifikationen und daraus resultierenden Effekte auf die Chromatinstruktur und Genexpression.

Eine der ältesten bekannten post-translationalen Histon-Modifikationen ist die Anfang der 1960er Jahre entdeckte Acetylierung.^[293, 294] Sie erfolgt an der ϵ -Aminogruppe von Lysin und wird durch Histon-Acetyltransferasen (HATs) unter Verwendung von Acetyl-CoA als Cofaktor katalysiert. In Eukaryoten gibt es verschiedene HATs,^[295-297] von denen einige sehr spezifisch nur bestimmte Positionen acetylieren.^[298-300] Durch die Acetylierung wird die Basizität von Lysin gesenkt, wodurch es unter physiologischen Bedingungen keine positive Ladung mehr trägt. Somit wird die Interaktion des Histons mit der DNA verringert.^[301] Mehrere Studien zeigen, dass dies die Bildung einer kompakten Chromatinstruktur verhindert und dadurch mit aktiver Transkription korreliert.^[302-305] Konsistent damit sind viele acetylierte Lysine vornehmlich in *Enhancer*- und Promotor-Regionen aktiver Gene zu finden.^[306] Um dynamische Veränderungen der Genexpression zu ermöglichen, können die Acetyl-Gruppen durch Histon-Deacetylasen (HDACs) wieder entfernt werden.^[307-309] Sowohl HATs als auch HDACs arbeiten dabei oft in Komplexen mit weiteren Proteinen, welche z. B. für die Substratspezifität wichtig sind.^[310-312]

Eine weitere Histon-Modifikation ist die Methylierung von Lysin und Arginin, wobei Lysin mono-, di- und tri-, Arginin hingegen mono- sowie dimethyliert werden kann. Sie wird durch Histon-Methyltransferasen unter Zuhilfenahme des Cofaktors SAM katalysiert. Dabei besitzen viele Histon-Lysin-Methyltransferasen eine sehr hohe Substrat-^[313, 314] aber auch Produktspezifität: das Auftreten von Tyrosin oder Phenylalanin an einer bestimmten Position im Enzym entscheidet, ob das Protein eine Monomethyltransferase ist oder auch di- und trimethylieren kann.^[315, 316] Im Gegensatz dazu sind Arginin-Methyltransferasen oft unspezifischer und können auch andere Proteine modifizieren.^[317-319] Sowohl Lysin- als auch Arginin-Methylierungen können abhängig von der Position der Aminosäure im Histon und der Methylierungs-Stufe sowohl einen aktivierenden als auch repressiven Einfluss auf die Genexpression ausüben.^[320-322] Die Methylierung von Histonen ist kein besonders dynamischer Prozess^[323] und lange Zeit war keine Histon-Demethylase bekannt, bis im Jahr 2004 die erste Lysin-Demethylase^[324] sowie eine Arginin-Deiminase, welche methyliertes Arginin durch Deiminierung zu Citrullin umwandelt,^[325, 326] identifiziert wurde. Kurz darauf wurden weitere Lysin-spezifische Enzyme^[327-333] und auch direkte Arginin-Demethylasen^[334, 335] entdeckt.

Andere weniger ausführlich studierte, epigenetisch relevante Histonmodifikationen beinhalten z. B. die Phosphorylierung von Serin,^[336, 337] Threonin^[338, 339] und Tyrosin,^[340, 341] die Ubiquitinylierung^[342, 343] oder Sumoylierung^[344] von Lysin sowie die ADP-Ribosylierung,^[345] welche nach bisherigem Kenntnisstand an Glutamat^[346, 347] und Lysin^[347, 348] stattfindet.

Das komplexe Zusammenspiel der verschiedenen Modifikationen bildet die Grundlage für unterschiedliche und sich dynamisch verändernde Chromatinstrukturen, welche für differenzielle Genaktivität nötig sind. Die dabei zugrunde liegenden Mechanismen sind die Nukleosomen-Belegung sowie die Nukleosomen-Positionierung. Erstere beschreibt, wie dicht ein bestimmter Abschnitt der DNA mit Nukleosomen besetzt ist, während letztere die genaue Position eines Nukleosoms im Genom behandelt. Damit ein Gen transkribiert werden kann, muss die DNA am Startpunkt der Transkription in dessen Promotorregion frei von Nukleosomen sein,^[349-351] was durch Einführung entsprechender Histon-Modifikationen ermöglicht wird.^[306, 352, 353] Dadurch können bestimmte Chromatin-Remodellierungskomplexe rekrutiert werden, die das Chromatin öffnen,^[354, 355] oder die Histone werden durch Transkriptionsfaktoren^[356-359] verdrängt. Somit entsteht ein Bereich freier DNA,^[360] welcher von zwei Nukleosomen flankiert wird.^[361] Besonders das erste Nukleosom der transkribierten Region ist dabei in vielen Genen aufgrund einer bestimmten DNA-Sequenz besonders stark positioniert,^[362] also fest an einer definierten Stelle gebunden. Dadurch bildet es eine Barriere, die dafür sorgt, dass sich ein definierter Nukleosom-freier Bereich bildet, an welchen der für die Transkription benötigten Präinitiations-Komplex binden kann.^[363] Durch das Zusammenspiel verschiedener Proteine kann die Elongation der Transkription nach der Initiation in vielen Fällen ohne eine vollständige Dissoziation der intragenen Nukleosomen erfolgen; lediglich ein H2A/H2B-Dimer wird temporär abgespalten.^[364-368]

Im Gegensatz zu der gerade beschriebenen Situation in aktiv transkribiertem Chromatin, auch Euchromatin genannt, ist die DNA in stillgelegten Bereichen des Genoms, dem Heterochromatin, sehr dicht gepackt.^[369] Dabei wird zwischen konstitutivem und fakultativem Heterochromatin unterschieden. Ersteres sind Genom-Abschnitte, deren Expression unabhängig vom Zelltyp unterdrückt wird, wie z. B. pericentromere^[370] und subtelomere^[371, 372] Regionen. Fakultatives Heterochromatin besteht hingegen aus codierenden Sequenzen, welche abhängig von Zelltyp und Stadium der Differenzierung abgeschaltet sind, aber zu einem späteren Zeitpunkt aktiviert werden können.^[373] Dies macht fakultatives Heterochromatin zum Hauptort der dynamischen epigenetischen Regulation.

Die Bildung von Heterochromatin erfolgt in mehreren Schritten. Zunächst werden durch Bindung bestimmter Protein-Komplexe oder RNA-Moleküle Histon-Deacetylasen und -Methyltransferasen rekrutiert,^[237, 374-378] wodurch ein Bereich mit charakteristischen Modifikationen für Heterochromatin entsteht.^[379] Besonders die Methylierung von H3K9 ist hierbei hervorzuheben, welche durch UHRF1-vermittelte Rekrutierung von DNMT1 ein Bindeglied zwischen Histon-Modifikationen und DNA-Methylierung darstellt.^[313, 380] Die neu entstandenen Modifikationen werden nun von weiteren Chromatin-Remodellierern, wie *Polycomb*-Proteinen^[381, 382] oder Heterochromatin-Protein 1,^[381, 383, 384] erkannt, wodurch sich das Heterochromatin weiter ausbreitet.^[385]

Wie bereits in Kapitel 1.1 erwähnt, wird darüber diskutiert, ob die Modifikation von Histonen tatsächlich als epigenetischer Mechanismus gewertet werden sollte, da es lange Zeit keine Belege gab, dass sie vererbt werden. Kürzlich fanden *Ragunathan et al.* und *Audergon et al.* jedoch unabhängig voneinander heraus, dass die Methylierungsmuster von H3K9 in *S. pombe* über mehrere Zellteilungen aufrechterhalten werden,^[386, 387] was den ersten Nachweis einer Histon-Modifikation als Träger epigenetischer Information darstellt. Weitere Studien müssen daher zeigen, ob dies auch für andere Modifikationen der Fall ist.

1.2 Massenspektrometrie-basierte Proteinanalyse

Die im vorigen Kapitel beschriebenen epigenetischen Mechanismen sorgen für eine Regulierung der Expression unterschiedlichster Gene, was abhängig von Stadium der Differenzierung und Typ einer Zelle in einer ganz bestimmten Protein-Zusammensetzung resultiert.^[388] Die Methoden zur Analyse der verschiedenen Aktivitätszustände im Bereich der Genomik, wie z. B. spezielle DNA-Sequenzierungen, die auch oxidierte Cytosine voneinander unterscheiden können,^[389-392] sowie der Transkriptomik, etwa die RNA-Sequenzierung,^[393-397] haben sich im Laufe der letzten Jahre stetig weiterentwickelt. Dennoch ist es oft unerlässlich, das Vorhandensein oder die Menge bestimmter Proteine direkt auf der Proteinebene selbst zu bestimmen, da Proteine verschiedene Halbwertszeiten besitzen,^[398] mRNA unterschiedlich stark translatiert werden kann^[399] und die Funktion bzw. Aktivität eines Proteins in großem Maße von den an ihm vorgenommenen post-translationalen Modifikationen abhängt.^[400-403] Das dazugehörige Fachgebiet wird als Proteomik bezeichnet, angelehnt an die Gesamtheit aller Proteine einer Zelle, dem Proteom. Sie beschäftigt sich mit Identifikation, Quantifizierung und Struktur von Proteinen sowie deren Interaktionen miteinander, um so unter anderem neue Biomarker zur Diagnose von Krankheiten zu identifizieren^[404-406] oder die Entwicklung neuer Wirkstoffe für deren Therapie zu ermöglichen.^[406-410]

Um dies zu erreichen, werden effiziente Analysemethoden benötigt, mit denen komplexe Proteingemische, wie beispielsweise Zelllysate, möglichst schnell und vollständig untersucht werden können. Bereits 1975 wurde hierfür von *O'Farrell* und *Klose* die hochauflösende 2D-Gelelektrophorese entwickelt, durch die mehrere hundert Proteine zunächst anhand ihres isoelektrischen Punktes und anschließend nach ihrer Masse aufgetrennt werden konnten.^[411, 412] Obwohl diese Technik einige Nachteile, wie einen niedrigen Durchsatz oder Probleme bei der Analyse von hydrophoben und wenig abundanten Proteinen^[413] besitzt, wurde sie im Laufe der Zeit weiterentwickelt und findet auch heutzutage noch Anwendung.^[414-418] Nichtsdestotrotz hat sich während der letzten Jahrzehnte die Massenspektrometrie (MS) als Analysemethode der Wahl durchgesetzt, weshalb sie im Folgenden eingehender besprochen wird.

1.2.1 Identifikation von Proteinen mittels Massenspektrometrie

Schon zu Beginn des 20. Jahrhunderts konnte *Sir J. J. Thomson* durch Gasentladungsexperimente Ionen erzeugen und diese anschließend anhand ihres Verhältnisses von Masse zu Ladung (m/z) auftrennen.^[419] Diese Technik wurde kurz darauf von *Dempster*^[420] und *Aston*^[421] beim Bau der ersten Massenspektrometer verwendet.

Die Massenspektrometrie kam vorerst jedoch nicht für die Analyse von Biomolekülen infrage, da diese aufgrund ihrer Größe nur schwer in die Gasphase überführt werden können und sich bei den damals etablierten harschen Ionisierungsbedingungen leicht zersetzen. Daher konzentrierte sich die Forschung zunächst auf chemische Methoden, um die Aminosäuresequenz eines Proteins zu bestimmen.^[422, 423] Als Pioniere der massenspektrometrischen Proteinanalyse gilt eine Gruppe um *K. Biemann*, die mit ihrem Massenspektrometer bereits 1959 Tripeptide identifizieren konnten. Dazu reduzierten sie diese vorher zu Polyaminoalkoholen, um mithilfe von Elektronenstoß-Ionisation charakteristische Fragmentierungen zu erzeugen.^[424] Dadurch war es ihnen möglich, 1976 die erste ausschließlich massenspektrometrisch identifizierte Sequenz eines Proteins zu präsentieren.^[425] Im Jahr 1981 wurde die vergleichsweise sanfte Ionisations-Methode *fast atom bombardment* (FAB) entwickelt, bei der eine Analyt-haltige Matrix mit einem Strahl aus teilweise ionisiertem Inertgas, z. B. Argon, beschossen wird.^[426] Diese Entdeckung führte in den darauffolgenden Jahren zu der Aufklärung von mehreren weiteren Proteinsequenzen.^[427-431] Doch erst Mitte der 1980er Jahre wurde die Massenspektrometrie zur Untersuchung von Proteinen durch die Entwicklung weiterer sanfter Ionisationstechniken populär: 1984 wurde von *Yamashita* und *Fenn*^[432] auf der Grundlage von Arbeiten von *Dole et al.*^[433] die Elektrospray-Ionisation (ESI) entwickelt. Im Jahr darauf folgte die von *Karas* und *Hillenkamp* publizierte Matrix-assistierte Laser-Desorption/Ionisation (MALDI).^[434] Beide Methoden wurden Ende der 1980er Jahre für die Analyse von großen Biomolekülen optimiert^[435, 436] und bilden seitdem die Grundlage der massenspektrometrischen Proteinanalyse.

1.2.1.1 Technische Grundlagen der massenspektrometrischen Analyse

Die heutzutage routinemäßig durchgeführte Untersuchung und Auswertung komplexer Protein-Gemische stellt hohe Anforderungen an die verwendeten Geräte. Moderne Massenspektrometer bestehen daher aus verschiedenen hochsensiblen Komponenten in einem Vakuum, welches dafür sorgt, dass die Ionen nicht mit Gasmolekülen kollidieren. Die essenziellen Bestandteile umfassen eine Ionenquelle, einen Analysator sowie einen Detektor. Erstere dient der Ionisierung der Analyten und deren Überführung in die Gasphase. Wie bereits erwähnt, werden für die Analyse von Proteinen heutzutage hauptsächlich MALDI und ESI verwendet, welche nun genauer erläutert werden.

Bei MALDI erfolgt die Ionisation ähnlich wie bei FAB durch Anregung einer Analyt-haltigen Matrix, wobei die Energie in diesem Fall von einem Laserpuls (meist im UV-Bereich) stammt (s. Abb. 1.4a). Der Analyt wird zunächst mit der Matrix auf einem Träger kokristallisiert. Geeignete Matrizen bestehen aus Substanzen, welche die Laserstrahlen möglichst effizient absorbieren können, wie z. B. konjugierte π -Systeme. Zudem sind sie oft Brønsted-Säuren oder -Basen und können daher bereits bei der Kokristallisation Protonentransfers mit Analyt-Molekülen eingehen. Die Ionisation erfolgt meist im Vakuum, allerdings besteht auch die Möglichkeit, bei atmosphärischem Druck zu ionisieren.^[437, 438] Der Mechanismus der Entstehung von Analyt-Ionen mittels MALDI ist dabei noch nicht vollständig geklärt. Nach dem sogenannten *lucky survivor*-Modell^[439] wird der betroffene Matrix-Bereich durch die Lasereinstrahlung erhitzt, wodurch große Cluster aus Matrix- und Analyt-Molekülen von der Festphase desorbiert werden^[440]. Obwohl die Cluster insgesamt nicht geladen sind, besitzen die Moleküle durch die beschriebenen Protonentransfers bereits teilweise Ladungen. Zerfallen die Cluster durch verbliebene Laserenergie zu kleineren Clustern, kann es zu einer Ladungstrennung und somit zur Entstehung von geladenen Clustern kommen, aus denen beim weiteren Zerfall Analyt-Ionen hervorgehen. Die Photoionisierung von Matrix-Molekülen und darauffolgende Protonen- sowie Elektronentransfer-Reaktionen tragen ebenfalls zur Generierung von Analyt-Ionen bei und könnten so erklären, weshalb mit MALDI hauptsächlich einfach geladene Spezies entstehen.^[439]

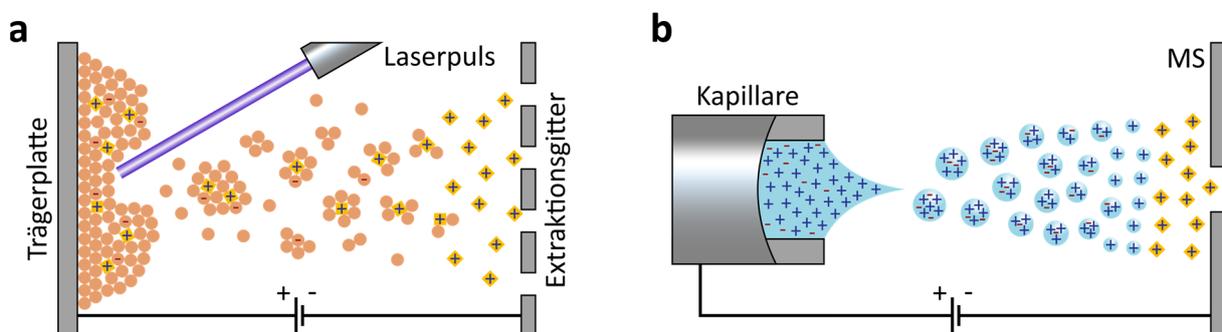


Abbildung 1.4: Ionisation von Biomolekülen mittels MALDI oder ESI. a) Funktionsprinzip von MALDI. Die Analyt-Ionen (gelb) befinden sich kokristallisiert mit Matrix-Molekülen (orange) auf einer metallischen Trägerplatte. Durch den Beschuss mit einem Laserpuls werden große Cluster desorbiert und bewegen sich entlang einer Spannung in Richtung des Massenspektrometers. Dabei fragmentieren die Cluster immer weiter, bis nur noch einzelne positiv geladene Analyt-Ionen übrigbleiben. Diese gelangen durch ein Extraktionsgitter ins Innere des Massenspektrometers. b) Ionisation mittels ESI. Die Analyt-Ionen befinden sich zunächst in einer flüssigen Phase, welche durch eine Kapillare geleitet wird. Durch das Anlegen einer Hochspannung bildet sich an der Spitze der Kapillare ein Taylor-Kegel, der daraufhin Tröpfchen emittiert. Durch Evaporation des Lösemittels steigt die Ladungsdichte der Tröpfchen, was bei Erreichen des Rayleigh-Limits zum Zerfall in mehrere kleinere Tröpfchen führt. Dieser Prozess wiederholt sich, bis schließlich nur noch einzelne Analyt-Ionen (gelb) vorliegen, welche in das Massenspektrometer gelangen.

Im Gegensatz zu MALDI findet die Ionisation bei ESI standardmäßig unter atmosphärischem Druck statt und der Analyt ist zudem in einer flüssigen Phase gelöst. Diese besteht meist aus einem Gemisch von Wasser und einem bei niedriger Temperatur siedenden Lösungsmittel, damit zum einen Protonen-Transfers zwischen Lösungsmittel und Analyt stattfinden können und zum anderen die Evaporation verbessert wird. Abhängig von der untersuchten Stoffklasse werden noch weitere Komponenten zugegeben, um die Bildung von Analyt-Ionen in der Lösung zu erhöhen. Zwischen dem Lösungsmittel-Emitter und der Ionenkapillare, durch welche die generierten Ionen in das Massenspektrometer gelangen, wird eine Hochspannung angelegt (s. Abb. 1.4). Dadurch entsteht an der Spitze des Emitters ein sogenannter Taylor-Kegel,^[441] der ab einer bestimmten Spannung ein Aerosol, bestehend aus kleinen Tröpfchen, emittiert. Diese Tröpfchen enthalten die geladenen Spezies, darunter auch die Analyt-Ionen. Durch Evaporation des Lösungsmittels und somit Abnahme des Tröpfchen-Volumens steigt die Ladungsdichte. Ist das sogenannte Rayleigh-Limit der Ladungsdichte erreicht, findet eine Coulomb-Explosion statt, durch die das Tröpfchen in mehrere kleinere Tröpfchen zerfällt.^[442] Dies wiederholt sich mehrmals, bis die in den Tröpfchen verbliebenen Ionen, abhängig von ihrer Größe, entweder durch Felddesorption in die Gasphase übergehen^[443] oder das Lösungsmittel von Tröpfchen, welche nur noch ein Ion enthalten, komplett evaporieren.^[444]

Da bei ESI ein Übergang von flüssiger zu Gasphase stattfindet, lässt sich dieser Ionisationsmethode leicht eine Hochleistungs-Flüssigkeitschromatographie (HPLC) vorschalten, wodurch eine Auftrennung der Analyten vor der massenspektrometrischen Analyse erreicht wird. Für MALDI ist dies auch möglich, jedoch mit höherem Aufwand verbunden.^[445]

Befinden sich die Ionen in der Gasphase, können sie durch Anlegen einer Spannung weiter zum Analysator geleitet werden, wo sie anhand ihres m/z -Verhältnisses aufgetrennt werden. Anschließend erfolgt die Aufzeichnung der Ionensignale durch den Detektor. Im Lauf der Zeit wurden viele verschiedene Systeme entwickelt, von denen die heutzutage relevanten nun genauer vorgestellt werden.

Bereits Mitte des 20. Jhdts. wurde die sogenannte *time-of-flight* (TOF)-Analyse vorgestellt,^[446] bei der die Ionen zunächst beschleunigt werden und dadurch abhängig von ihrem m/z -Verhältnis unterschiedliche Geschwindigkeiten erreichen. Anschließend durchqueren sie einen feldfreien Raum und es wird die Zeit gemessen, welche die Ionen benötigen, um auf einen Detektor zu treffen. Diese Zeit ist proportional zum m/z -Verhältnis, wodurch dieses bestimmt werden kann. Als Detektoren dienen hierbei meist Sekundärelektronenvervielfacher wie Mikrokanalplatten oder Ionen-Photonen-Wandler.^[447]

Zwei weitere Analysatoren, welche häufig mit Sekundärelektronenvervielfachern gekoppelt werden, sind der Quadrupol und die mit ihm verwandte Paul-Ionenfalle. Ein Quadrupol besteht aus vier stabförmigen, quadratisch angeordneten und parallelen Elektroden, wobei die gegenüberliegenden Elektroden jeweils elektrisch miteinander verbunden sind (Abb. 1.5a).

Nun wird eine Wechselspannung im Hochfrequenzbereich mit einem Gleichspannungsanteil angelegt, wodurch im Quadrupol-Inneren ein elektrisches Feld entsteht. Werden Ionen nun durch dieses Feld geführt, bewegen sie sich oszillierend durch den Quadrupol, wobei bei einer gegebenen Amplitude und Frequenz der angelegten Spannung nur Ionen aus einem bestimmten Teil des m/z -Bereichs eine stabile Bahn annehmen. Alle anderen Ionen oszillieren mit einer zu hohen Amplitude, sodass deren Bahnen instabil werden und sie mit einer der Elektroden kollidieren oder den Quadrupol radial verlassen. Dadurch gelangen nur spezifische Ionen zum nachgeschalteten Detektor. Indem nacheinander verschiedene Spannungen durchlaufen werden, kann der gesamte m/z -Bereich abgedeckt werden und man erhält ein Massenspektrum.

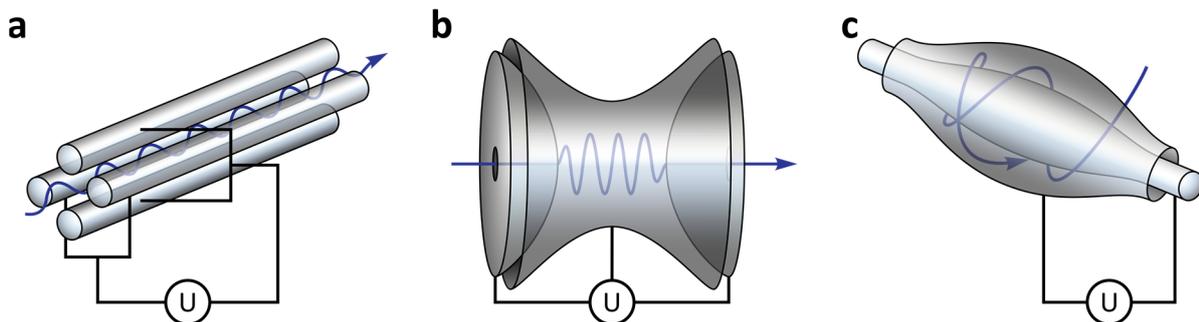


Abbildung 1.5: Schematische Darstellung einiger MS-Analysatoren. In blau ist jeweils der Ionenpfad angedeutet. a) Aufbau eines Quadrupols bestehend aus vier quadratisch angeordneten Elektroden. Diese sind an eine Wechselspannungsquelle angeschlossen, wobei die jeweils gegenüberliegenden Elektroden dieselbe Polarität besitzen. b) Paul Ionenfalle, bei der das elektrische Feld zwischen einer Ring- und zwei Endkappen-Elektroden aufgebaut wird. c) Darstellung einer Orbitrap bestehend aus einer zentralen Spindel-Elektrode, die von einer tonnenförmigen Elektrode umgeben ist. Die Ionen werden in die Tonnen-Elektrode injiziert und oszillieren abhängig von ihrem m/z -Verhältnis um die Spindel-Elektrode.

Die Paul-Ionenfalle, benannt nach ihrem Erfinder *Wolfgang Paul*,^[448] funktioniert nach demselben Prinzip, besteht jedoch aus einer ringförmigen Elektrode und zwei Elektroden-Kappen, welche die offenen Seiten des Rings begrenzen und zudem als Ein- bzw. Auslass für die Ionen dienen (Abb. 1.5b). Die Spannung wird zwischen den Elektroden-Kappen und der Ringelektrode angelegt. Diese besondere Geometrie erlaubt zusätzlich zur Selektion auch das Sammeln von Ionen in der Falle, wodurch z. B. intensivere Signale erreicht werden können. Mittlerweile gibt es weitere Arten von Paul-Ionenfallen, etwa lineare^[449] oder zylindrische^[450] Fallen.

Die bisher beschriebenen Systeme erfordern alle eine sequenzielle Analyse der Ionen mittels externer Detektoren. Es gibt jedoch auch Analysatoren, die mithilfe der Fourier-Transformation (FT) den gesamten m/z -Bereich direkt messen können. Unter diese Kategorie fallen die FT-Ionenzyklotronresonanz (FT-ICR)-Zelle^[451] sowie die Orbitrap^[452]. Ein FT-ICR-Analysator besteht aus einer Penning-Falle, in der die Ionen durch ein hohes Magnetfeld auf kreisförmige Bahnen mit für ihr m/z -Verhältnis spezifischen Zyklotron-Frequenzen gezwungen werden.

Durch Anlegen eines orthogonalen elektrischen Wechselfelds mit langsam steigender Frequenz werden die Ionen nacheinander auf größere Bahnen angeregt, wenn die angelegte und die Zyklotron-Frequenz übereinstimmen. Nach Abschalten dieser Anregungsfrequenz fallen die dadurch gebildeten Ionen-Pakete wieder auf ihre ursprüngliche Bahn zurück. Währenddessen induzieren sie Spiegelströme in gegenüberliegenden Detektor-Elektroden, die gemessen werden können. Dabei erhält man eine Frequenz, die aus der Überlagerung der Zyklotron-Frequenzen aller Ionen resultiert und *free induction decay* (FID) genannt wird. Mittels Fourier-Transformation lassen sich die zugrunde liegenden Frequenzen der einzelnen Ionen berechnen, welche umgekehrt proportional zum m/z -Verhältnis sind.

Ein Orbitrap-Analysator ist eine modifizierte Kingdon-Falle, die sich aus einer zentralen Spindel-Elektrode und einer umhüllenden Elektrode zusammensetzt (Abb. 1.5c). Die Ionen werden schräg injiziert, wodurch sie spiralförmig um die Spindel-Elektrode oszillieren. Es entstehen Ringe von Ionen mit gleichem m/z -Verhältnis, die mit spezifischer Frequenz entlang der Spindel oszillieren und dadurch ähnlich zu FT-ICR Spiegelströme in Detektor-Elektroden induzieren.

Dadurch entsteht eine Überlagerung der Frequenzen der einzelnen Ionen (Transient). Dieser Transient kann daraufhin, wie bereits für die Ionenzyklotronresonanz beschrieben, mithilfe einer Fourier-Transformation in ein Massenspektrum umgewandelt werden.

Der Vorteil dieser beiden FT-MS-Techniken gegenüber den anderen beschriebenen Methoden liegt in der sehr hohen Massengenauigkeit und Auflösung,^[453, 454] was für die Analyse komplexer Proteingemische unabdingbar ist. Moderne Massenspektrometer bestehen zudem oft aus Kombinationen von mehreren Analysatoren, um anspruchsvollere Experimente zu ermöglichen, wie sie heutzutage in der Proteomik üblich sind.

1.2.1.2 Methoden der Protein-Identifikation – zwei Wege, ein Ziel

Die massenspektrometrische Protein-Analyse findet meist mit positiv geladenen Analyten statt. Im sauren Milieu sind alle sauren und basischen Seitenketten protoniert, was für eine positive Gesamtladung sorgt. Die Untersuchung gestaltet sich besonders dann einfach, wenn ein einzelnes Protein analysiert wird, dessen Identität und Molekulargewicht bereits bekannt sind. Ist dies nicht der Fall, reicht es oftmals nicht aus, lediglich die Masse des intakten Proteins zu bestimmen, da verschiedene Proteine sehr ähnliche Massen besitzen können und viele von ihnen stark post-translational modifiziert werden. Daher wurden im späten 20. Jhdt. zwei Methoden entwickelt, die eine genauere Zuordnung der massenspektrometrischen Signale zu Proteinen ermöglichen. Beide Techniken machen sich hierfür die Fragmentierung von Ionen in der Gasphase zunutze, welche durch moderne Massenspektrometer ermöglicht wird. Die in der Proteomik am häufigsten verwendeten Fragmentierungs-Arten sind die kollisions-induzierte Dissoziation (CID)^[455] sowie die Elektronentransfer-Dissoziation (ETD).^[456]

Erstere basiert auf der Kollision von Ionen mit einem Inertgas, wobei man zwischen zwei Formen der Ionen-Beschleunigung unterscheidet. Bei klassischer CID werden Ionen in einer Ionenfalle anhand ihres m/z -Verhältnisses spezifisch angeregt und kollidieren durch ihre resultierende Schwingung mit dem Stoßgas. Im Gegensatz dazu werden die Ionen bei der höherenergetischen kollisions-induzierten Dissoziation (HCD) in eine Kollisionszelle beschleunigt, wo sie auf die Inertgas-Moleküle treffen.^[457] In beiden Fällen wird durch die Kollision kinetische Energie in innere Energie umgewandelt, was zur Fragmentierung bestimmter Bindungen führt. Während durch die spezifische Anregung von CID Ionen mit nur einem bestimmten m/z -Verhältnis fragmentiert werden können, werden durch HCD alle Ionen fragmentiert.

ETD beruht hingegen auf Elektronentransfer-Reaktionen zwischen den positiv geladenen Analyt-Ionen und negativ geladenen Radikalanionen, welche durch chemische Ionisation, z. B. von Fluoranthren,^[458] erhalten und anschließend zur Reaktion mit den Analyt-Ionen zusammengebracht werden. Durch die Übertragung eines Elektrons auf die Analyt-Ionen entstehen instabile Radikalkationen, die daraufhin fragmentieren.

Die Erste der angesprochenen Methoden zur massenspektrometrischen Proteinidentifikation ist der sogenannte *top-down*-Ansatz (Abb. 1.6).^[459] Hierbei wird ein intaktes Protein zunächst ionisiert und seine Masse analysiert. Aufgrund einer Verteilung von verschiedenen Ladungszuständen, die während der Ionisation entstehen, erhält man ein Massenspektrum (MS-Spektrum) mit je nach Ladung unterschiedlichen m/z -Verhältnissen. Diese sind durch natürlich vorkommende schwere Isotopen ebenfalls jeweils in mehrere Signale aufgeteilt. Anhand dieser Muster lässt sich das Molekulargewicht des untersuchten Proteins errechnen. Da dieses nicht immer ausreicht, um ein Protein eindeutig zu identifizieren, kann ein sogenanntes MS²-Experiment durchgeführt werden. Hierbei erfolgt nach Aufnahme des MS-Spektrums die Isolation der zu einem Signal gehörenden Protein-Ionen anhand ihres m/z -Verhältnisses, welche daraufhin fragmentiert werden. Je nach Ladungszustand und Sekundärstruktur des Proteins können dabei unterschiedliche Bindungen gespalten werden.^[460-463] Es folgt die Aufnahme eines weiteren Massenspektrums (MS²-Spektrum) der Fragment-Ionen. Dieses ist charakteristisch für das jeweilige Protein und ermöglicht dessen Identifikation mittels *sequence tag*-Strategie,^[464] *de novo*-Bestimmung der Aminosäuresequenz^[465] oder Abgleich mit bekannten Fragmentspektren einer Datenbank.^[466]

Die *top-down*-Strategie ist besonders für die Analyse von kleineren Proteinen und Identifikation von verschiedenen Isoformen,^[467, 468] sowie für die Untersuchung diverser Proteinmodifikationen^[469-471] geeignet. Die Effizienz der Ionisierung nimmt jedoch mit zunehmendem Molekulargewicht ab, was eine Einschränkung dieser Methode darstellt. Zudem können wenig abundante Proteine aufgrund der Aufteilung der Signalintensität auf viele Ladungszustände nur schlecht analysiert werden.

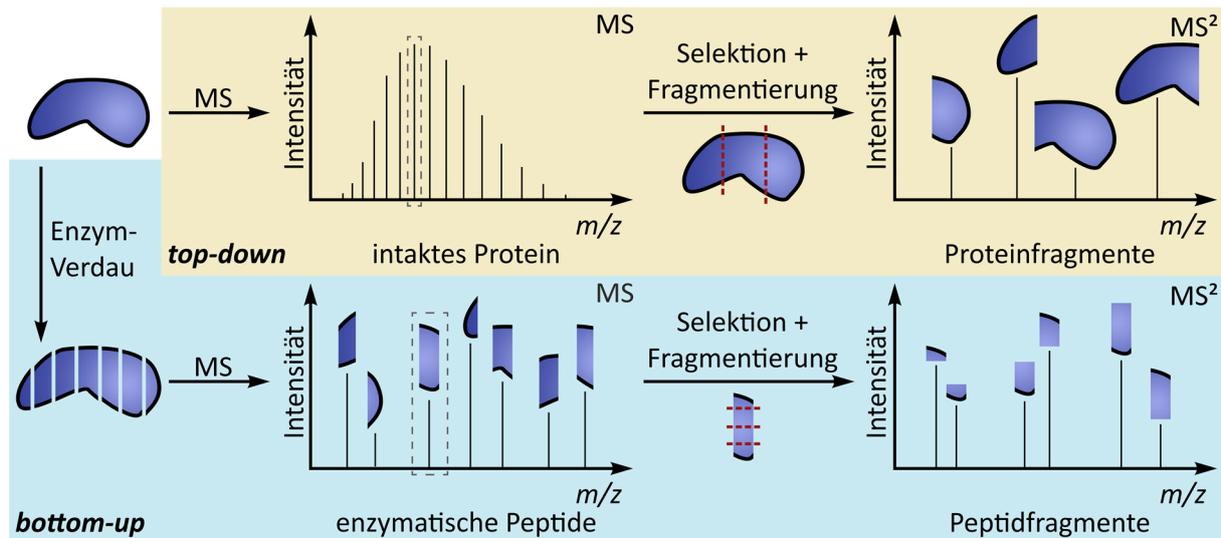


Abbildung 1.6: Vergleich der massenspektrometrischen Protein-Identifizierung mithilfe der *top-down*- bzw. *bottom up*-Strategie. Bei der *top-down*-Methode wird ein MS-Spektrum des intakten Proteins aufgenommen, was aufgrund vieler möglicher Ladungszustände zu mehreren Signalen führt. Anschließend wird ein einzelnes Signal selektiert und fragmentiert (mögliche Fragmentationsstellen in rot eingezeichnet). Anschließend wird eine MS-Analyse der Proteinfragmente durchgeführt. Im Falle der *bottom-up*-Strategie wird das Protein zuerst proteolytisch verdaut, bevor die resultierenden Peptide einer MS-Analyse unterzogen werden. Einzelne Peptide können daraufhin isoliert und fragmentiert werden. Die entstehenden Peptidfragmente werden schließlich erneut untersucht. Die Isotopenmuster der Signale wurden der Übersicht halber nicht dargestellt.

Aus diesem Grund wird für die massenspektrometrische Analyse von komplexen Proben mit vielen verschiedenen Proteinen in unterschiedlichen Konzentrationen, wie beispielsweise Zelllysaten, heutzutage meist die sogenannte *bottom-up*-Methode verwendet (Abb. 1.6). Bei dieser werden die Proteine zunächst proteolytisch zu Peptiden verdaut. Hierfür stehen mehrere Proteasen zur Auswahl; die am häufigsten verwendete ist jedoch Trypsin, welches die Peptidbindung C-terminal von den Aminosäuren Lysin und Arginin spaltet. Da bei einem tryptischen Verdau jedes Peptid durch den N-Terminus und Lysin oder Arginin als C-terminale Aminosäure im sauren Milieu mindestens zweifach positiv geladen ist, können die Peptidsignale im MS-Spektrum leicht von einfach geladenen Verunreinigungen unterschieden werden. Die Spaltung eines Proteins in mehrere Peptide, verbessert sowohl die Löslichkeit als auch die Ionisation und ermöglicht selbst die Analyse von großen oder wenig abundanten Proteinen.

Analog zur *top-down*-Strategie werden die Peptide zunächst ionisiert und analysiert. Bei der Untersuchung einzelner Proteine reicht das so erhaltene MS-Spektrum der proteolytischen Peptide oftmals bereits aus, um anhand eines Vergleichs mit einer Datenbank aus theoretisch verdauten Peptidmassen bekannter Proteinsequenzen die Identität eines Proteins festzustellen (*peptide mass fingerprinting*).^[472, 473] In komplexeren Proben tritt aufgrund der Vielzahl möglicher Peptide allerdings das Problem auf, dass die erhaltenen Peptidmassen zu verschiedenen Aminosäuresequenzen mit demselben Molekulargewicht passen, was eine Zuordnung verhindert.

Aus diesem Grund erfolgt meist eine Fragmentierung der Peptide und Aufnahme von MS²-Spektren. Ionen eines einzelnen Signals werden isoliert und anschließend fragmentiert, was in den meisten Fällen zur Spaltung einer Bindung im Peptidrückgrat pro Ion führt. Je nach Fragmentierungs-Methode brechen dabei unterschiedliche Bindungen (s. Abb. 1.7a). Da die Fragmentierungsenergien aller labilen Bindungen in einem Peptid ähnlich sind, erhält man aufgrund der hohen Anzahl an isolierten Ionen ein MS²-Spektrum mit Ionenserien, deren Masse sich jeweils nur um eine Aminosäure unterscheidet. Dabei bilden sich mit CID^[474] und HCD^[475] hauptsächlich b- und y-Ionen (Abb. 1.7b), während bei der Verwendung von ETD c- und z-Ionen entstehen.^[456] Anhand der Differenzen der Signale, die sich nur um eine Aminosäure unterscheiden lässt sich theoretisch die Aminosäuresequenz des Peptids bestimmen.

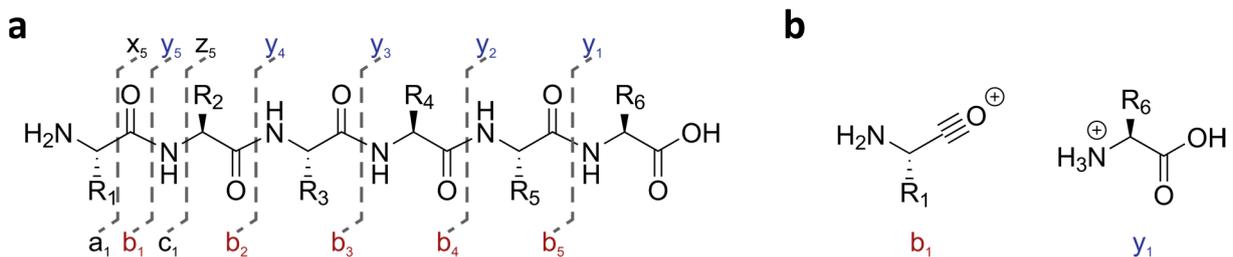


Abbildung 1.7: Mögliche Fragmentierungen von Peptiden in der Gasphase. a) Übersicht der durch Fragmentierung in der Gasphase gespaltenen Bindungen eines Peptids und der daraus resultierenden Fragmente. Der Übersicht halber ist jeweils nur eine der Fragmentierungsstellen, welche zu a- und x- bzw. c- und z-Ionen führen, dargestellt. Die Benennung der Fragmente folgt der von *Roepstorff* und *Fohlman* vorgeschlagenen^[476] und von *Johnson et al.* modifizierten^[477] Nomenklatur. b) Exemplarische Darstellung von b₁- sowie y₁-Ion des Peptids aus a).

Meist bilden sich jedoch nicht alle Fragment-Ionen, weshalb man sich heutzutage der automatisierten Datenbankanalyse unter Verwendung eines Algorithmus, wie z. B. *Sequest*,^[478] *Mascot*^[479] oder *Andromeda*^[480] bedient. Diese arbeiten alle nach einem ähnlichen Prinzip, indem sie Proteine aus einer vorher definierten Datenbank *in silico* zu Peptiden verdauen und die Massen der Fragment-Ionenserien dieser Peptide berechnen. Die so erhaltenen Fragmentmassen der theoretischen Peptide, deren Masse zu einem Signal im MS-Spektrum, dem sogenannten Vorläufer-Ion, passt, werden dann mit den Signalen im MS²-Spektrum abgeglichen. Die Peptidsequenz, bei der die theoretischen Massen am besten mit den beobachteten Fragment-Signalen übereinstimmen, wird dem entsprechenden MS²-Spektrum zugeordnet. Da die meisten Peptide eine einzigartige Sequenz besitzen, kann daraus das zugehörige Protein ermittelt werden. Auch Modifikationen können vorgegeben und so häufig genau lokalisiert werden.

Die *bottom-up*-Methode im Besonderen, aber auch *top-down*-Ansätze, profitieren von einer vorangehenden Separation der Analyt-Moleküle durch HPLC, wenn komplexere Proben analysiert werden sollen, da dies für eine Verringerung der Anzahl an Signalen in den einzelnen MS-Spektren sorgt und damit eine Fragmentierung von mehr Spezies ermöglicht. Als stationäre Phase dienen dabei meist Umkehrphasen-Materialien, welche Proteine oder Peptide anhand ihrer Hydrophobizität auftrennen.

Ein Lösungsmittelgradient aus Wasser und einem steigenden Anteil eines unpolaren Lösungsmittels, etwa Acetonitril, bewirkt eine sukzessive Elution von Analyt-Spezies. Wird ESI als Ionisationsmethode verwendet, kann die HPLC-Auftrennung, wie bereits erwähnt, direkt an die Ionisation gekoppelt werden. Nach der Aufnahme eines MS-Spektrums werden dann meist mehrere Signale nacheinander fragmentiert und analysiert, bevor das nächste MS-Spektrum aufgenommen wird und sich der Prozess wiederholt. Dies kann in Analysen mit mehreren zehntausend MS²-Spektren resultieren, was während der Auswertung aufgrund der vielen möglichen Kombinationen von Sequenzen und Spektren zu falsch-positiven Zuordnungen führen kann. Daher kommen bei der Auswertung häufig statistische Methoden zum Einsatz, die einen Teil der Peptid-Spektrums-Zuordnungen (PSMs) ausschließen, um den geschätzten Fehler unterhalb einer bestimmten Grenze zu halten. Dies kann z. B. durch Invertierung der verwendeten Datenbank geschehen, welche anschließend ebenfalls für die Auswertung der Spektren verwendet wird. Da die hierbei erhaltenen PSMs bis auf palindromische Sequenzen allesamt falsch-positiv sein müssen, kann durch Vergleich mit den aus der korrekten Datenbank gewonnenen PSMs die Falscherkennungsrate geschätzt werden.^[481] Da jedes einem Spektrum zugeordnete Peptid einen Wert erhält, der den Grad der Übereinstimmung angibt, können die Ergebnisse gefiltert werden, sodass eine geringe Falscherkennungsrate resultiert.

1.2.2 Crosslinking-Massenspektrometrie

Die im letzten Kapitel beschriebenen Methoden ermöglichen der modernen Wissenschaft eine umfassende Charakterisierung von kompletten Proteomen in kurzer Zeit. Informationen über Protein-Protein-Interaktionen und Struktur der Proteine gehen allerdings im Zuge der Probenvorbereitung verloren. Diese Auskünfte sind jedoch essenziell für das Verständnis von Signalwegen oder der Wirkungsweise von Proteinen und Proteinkomplexen. Hierfür bisher verwendete Techniken wie Röntgenkristallographie, NMR-Spektroskopie oder Elektronenmikroskopie zur Strukturanalyse sowie Coimmunpräzipitation für die Untersuchung von Protein-Protein-Interaktionen besitzen diverse Limitationen. So muss eine Probe für die Strukturanalyse eine vergleichsweise hohe Konzentration besitzen und darf kaum Verunreinigungen enthalten, während die Coimmunpräzipitation nur starke Interaktionen dokumentieren kann und nicht zwischen direkten und indirekten Interaktionen unterscheidet. In den letzten Jahren hat sich daher die sogenannte Crosslinking-Massenspektrometrie (XL-MS) entwickelt, welche diese Einschränkungen durch eine Kombination des *bottom-up*-Ansatzes und kovalenter Modifikation durch chemische Reagenzien umgeht. Üblicherweise gehen Proteinstruktur und Protein-Protein-Interaktionen beim proteolytischen Verdau zu Peptiden verloren. Bestimmte Moleküle, sogenannte Crosslinker, ermöglichen jedoch eine Fixierung dieser Informationen vor dem Verdau durch kovalente Verknüpfung zweier Aminosäuren, welche sich in unmittelbarer räumlicher Nähe zueinander befinden (s. Abb. 1.8a).

Nach der Proteolyse erhält man also neben unmodifizierten Peptiden auch Peptide, die durch den Crosslinker modifiziert wurden. Dabei gibt es verschiedene Reaktionsmöglichkeiten:^[482] Typ 2-Crosslinks bezeichnen die Vernetzung zweier verschiedener Peptide, welche entweder vom selben Protein (Intraprotein-Crosslink) oder von verschiedenen Proteinen (Interprotein-Crosslink) stammen können. Befinden sich die Reaktionsstellen nach dem Verdau im selben Peptid, spricht man von einem Typ 1-Crosslink. Typ 0-Crosslinks (Monolinks) schließlich entstehen, wenn nur eine der beiden reaktiven Gruppen des Crosslinkers mit einer Aminosäure reagiert hat. Besonders interessant sind folglich Typ 2-Crosslinks, weil Intraprotein-Crosslinks einen Aufschluss über die Proteinstruktur und Interprotein-Crosslinks Informationen über Protein-Protein-Interaktionen geben.

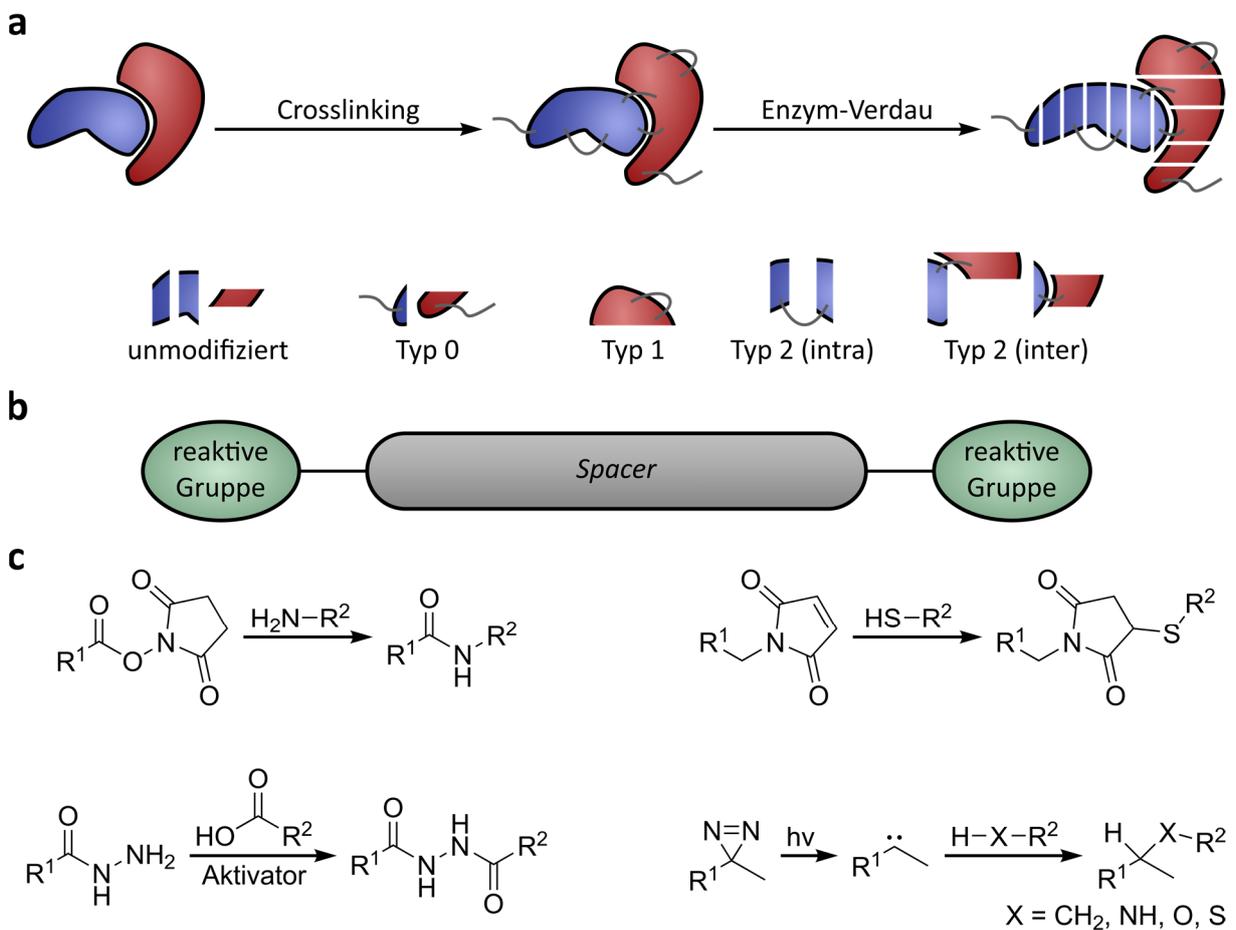


Abbildung 1.8: Grundlagen der Crosslinking-Massenspektrometrie. a) Schematische Darstellung der Probenvorbereitung für XL-MS-Analysen. Vor dem enzymatischen Verdau werden Peptid-Seitenketten, die sich in räumlicher Nähe zueinander befinden und mit den Reaktivgruppen des Crosslinkers reagieren können, kovalent miteinander verknüpft. Nach dem Verdau entstehen neben unmodifizierten Peptiden auch verschiedene Arten von Peptiden oder Peptid-Paaren, die eine Crosslinker-Modifikation enthalten. b) Grundlegender Aufbau eines Crosslinkers. Zwei reaktive Gruppen sind über einen sogenannten *Spacer* miteinander verbunden. Letzterer bestimmt die Länge des Crosslinkers und kann weitere Funktionalitäten wie eine spaltbare Gruppe oder eine Möglichkeit zur Anreicherung der Crosslinks enthalten. c) Übersicht einiger reaktiver Gruppen für verschiedene Aminosäuren-Seitenketten.

Da eine erfolgreiche XL-MS-Analyse mit hohen Anforderungen hinsichtlich einer möglichst quantitativen Reaktion, Bioverträglichkeit und MS-Kompatibilität verbunden ist, gilt ein besonderes Augenmerk dem Design der zur Anwendung kommenden Reagenzien. Grundsätzlich besitzt ein Crosslinker mindestens zwei reaktive Gruppen, welche mit Seitenketten der Protein-Aminosäuren reagieren können (Abb. 1.8b und c). So existieren z. B. N-Hydroxysuccinimidyl-Ester (NHS-Ester) zur Reaktion mit Aminen wie dem Protein-N-Terminus oder dem ϵ -Amin von Lysin-Seitenketten, wobei in geringerem Ausmaß auch die Hydroxyl-Gruppen von Serin, Threonin und Tyrosin reagieren können.^[483, 484] Diese Funktionalisierung hat einen vergleichsweise hohen Grad der Vernetzung zum Vorteil, da Proteine meist mehrere Lysine enthalten, die sich zudem durch ihre Hydrophilie oft in Lösungsmittel-zugänglichen Positionen befinden. Ein Nachteil ist die Hydrolyse-Empfindlichkeit der NHS-Ester, was die Bildung von Monolinks als Nebenreaktion hervorruft. Andere reaktive Gruppen sind z. B. Maleimide zur Reaktion mit Thiolen,^[485] Hydrazide, welche an zuvor aktivierte Carbonsäure-Funktionen von Asparagin- und Glutaminsäure angreifen können,^[486, 487] sowie unspezifisch in X-H-Bindungen insertierende photoreaktive Gruppen wie Diazirine^[488, 489] oder Benzophenone.^[489, 490]

Die beiden zum Einsatz kommenden reaktiven Gruppen sind im Crosslinker durch den sogenannten *Spacer* miteinander verbunden. Die Länge dieser Region legt den Abstand der Reaktivgruppen zueinander fest^[491] und bestimmt dadurch, wie weit die vom Crosslinker verknüpften Funktionen voneinander entfernt sein können. Eine Reaktion ist allerdings durch die dynamischen Bewegungen der Proteine in Lösung in einem größeren Bereich möglich.^[492] Zudem enthält der *Spacer* oft eine weitere Funktionalität: eine in der Gasphase spaltbare Gruppe, welche die Auswertung der massenspektrometrischen Analyse erleichtert. Da sich das m/z -Verhältnis eines Vorläufer-Ions im Falle eines Typ 2-Crosslinks aus der Masse der beiden Peptide sowie des Crosslinkers zusammensetzt, müssen alle Kombinationen von zwei Peptiden aus der eingesetzten Datenbank berücksichtigt werden, was eine Quadrierung der Datenbankgröße zur Folge hat. Vor allem bei komplexeren Proben mit Proteom-weitem Datenbank-Vergleich führt dies zu einem bis jetzt noch nicht handhabbaren Rechenaufwand. Daher werden Crosslinker ohne spaltbare Gruppe, beispielsweise DSS (Abb. 1.9a), hauptsächlich für die Strukturanalyse einzelner Proteine oder Proteinkomplexe verwendet. Ein chemisch spaltbarer Crosslinker, wie z. B. der Disulfid-haltige Crosslinker DSP^[493] (Abb. 1.9b), beseitigt dieses Problem zwar, allerdings geht durch die Spaltung vor der massenspektrometrischen Analyse auch die Information verloren, welche Peptide miteinander interagiert hatten. Daher besitzen solche Crosslinker ebenfalls limitierte Anwendungsgebiete.

Im Gegensatz dazu ermöglichen in der Gasphase spaltbare Gruppen auf mehrere Arten die Analyse von komplexen Proben. Dabei konnten sich bereits einige dieser Gruppen etablieren, etwa die Harnstoff-Gruppe in DSBU^[494] (Abb. 1.9c) oder das Sulfoxid von DSSO^[495] (Abb. 1.9d). Ihnen ist gemeinsam, dass sie bei geringerer Energie als das Peptidrückgrat spalten.

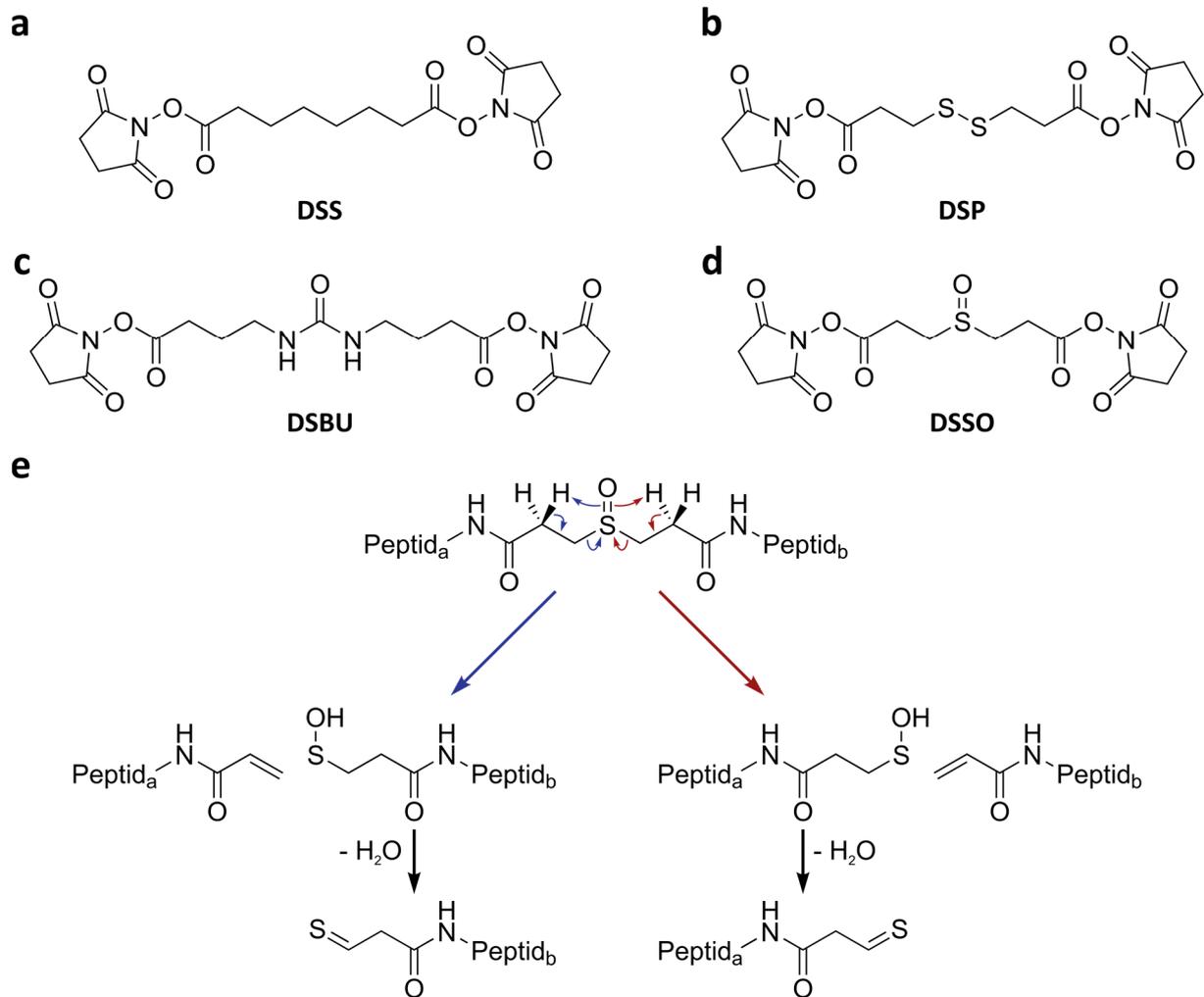


Abbildung 1.9: Übersicht einiger Crosslinker und Prinzip der Fragmentierung eines spaltbaren Crosslinkers. a) Struktur des nicht-spaltbaren Crosslinkers DSS. b) Struktur des chemisch spaltbaren Crosslinkers DSP. c) Struktur des in der Gasphase spaltbaren Crosslinkers DSBU. d) Struktur des in der Gasphase spaltbaren Crosslinkers DSSO. e) Schema der Fragmentierung eines DSSO-Crosslinks in der Gasphase.

Beim Sulfoxid erfolgt die Spaltung durch eine pericyclische Reaktion, was im Bruch von einer der beiden C-S-Bindungen resultiert (s. Abb. 1.9e). Durch diese Spaltung entsteht ein Alken- sowie ein Sulfensäure-Fragment, wobei letzteres durch den Verlust von H₂O weiter zu einem Thial zerfallen kann.

Abbildung 1.10 zeigt eine typische MS²-Analyse eines solchen Typ 2-Crosslinks. Wird im MS-Spektrum ein Crosslink isoliert und fragmentiert, kommt es im MS²-Spektrum zur Ausbildung von zwei Arten von Signalen: zum einen entstehen die Ionenserien der Peptide. Zusätzlich bilden sich auch Ionen, welche aus der Fragmentierung des Crosslinkers entstehen und folglich aus einem einzelnen, intakten Peptid und dem Crosslinker-Fragment bestehen. Da die beiden C-S-Bindungen des Crosslinkers stochastisch mit ähnlicher Wahrscheinlichkeit gespalten werden, entstehen zwei Signalaare, welche jeweils eine definierte Massendifferenz aufweisen. Diese ergibt sich aus dem Unterschied der Molekulargewichte der beiden Crosslinker-Fragmente am selben Peptid.

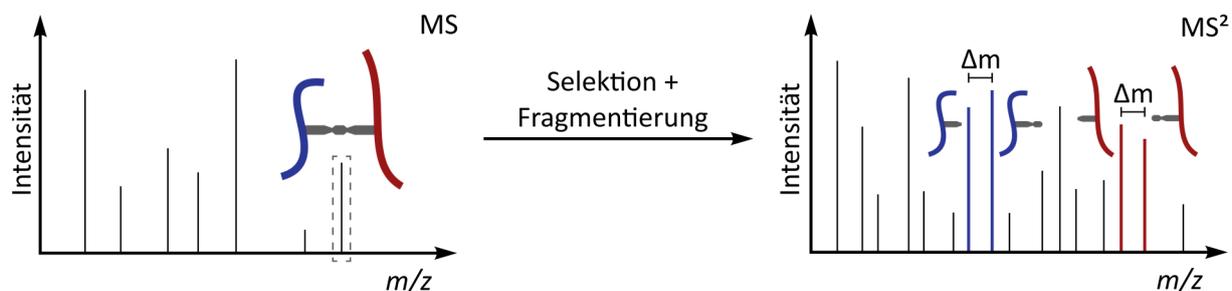


Abbildung 1.10: MS²-Analyse eines Typ 2-Crosslinks bei Verwendung eines symmetrisch fragmentierenden Crosslinker. Wird aus dem MS-Spektrum ein Crosslink für die Fragmentierung ausgewählt, können neben dem Peptidrückgrat auch die labilen Bindungen des Crosslinkers gespalten werden. Dadurch entstehen im MS²-Spektrum neben den Ionenserien auch zwei Signale pro Peptid, welche sich jeweils nur um eine bestimmte, ausschließlich vom Crosslinker abhängige Masse unterscheiden.

Moderne Analyse-Programme wie *MeroX*^[496] oder *XlinkX*^[497, 498] können diese Reporter-Signale aufgrund der charakteristischen Differenz identifizieren und für eine robustere Crosslink-Identifizierung nutzen.

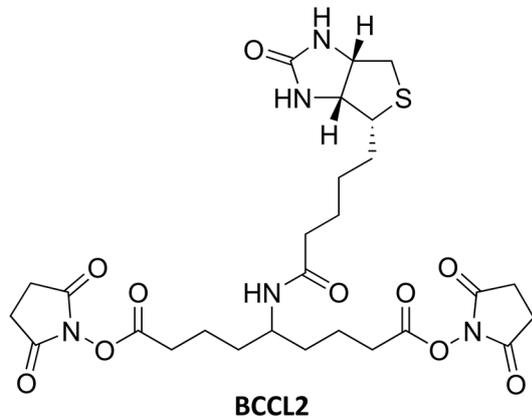
Liegt die Fragmentierungsenergie des Crosslinkers unterhalb der von Peptiden, ist es zusätzlich möglich, MS³-Analysen durchzuführen, bei denen zunächst mit niedriger Energie nur der Crosslinker fragmentiert wird, bevor die Reporter-Signale aus dem MS²-Spektrum einzeln isoliert und erneut fragmentiert werden, wodurch man MS³-Spektren mit den Fragment-Ionenserien einzelner Peptide erhält.

Trotz dieser Möglichkeiten, die Identifikation von Crosslinks zu verbessern, gibt es eine weitere Herausforderung bei deren Analyse in komplexen Proben, nämlich ihre geringe Abundanz verglichen mit den unmodifizierten Peptiden. Daher ist es oftmals notwendig, Typ 2-Crosslinks, welche aufgrund ihrer Information meist am interessantesten sind, vor der Analyse anzureichern. Dies kann zum einen durch chromatographische Methoden wie Größenausschluss- (SEC)^[499] oder Kationenaustausch-Chromatographie^[500] erfolgen, da Crosslinks in der Regel größer sind und eine höhere positive Ladung besitzen als lineare Peptide.

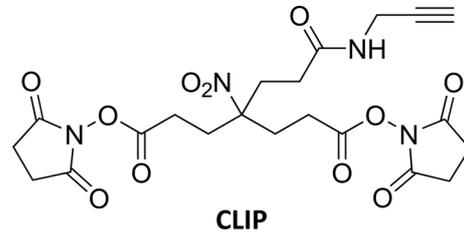
Eine selektivere Art der Anreicherung wird durch den Einbau bestimmter funktioneller Gruppen, welche eine Affinitätschromatographie ermöglichen, in die *Spacer*-Region des Crosslinkers erreicht. Oft wird hierfür Biotin verwendet, welches eine starke Bindung mit Streptavidin eingehen kann.^[501] Das Biotin kann zum einen direkt in den Crosslinker eingebaut werden, wie es z. B. in den PIR-Reagenzien^[502] oder den nicht-spaltbaren Crosslinkern BCCL2^[503] (Abb. 1.11a) und *One-Piece Leiker*^[504] der Fall ist. Eine weitere Möglichkeit ist der Einbau von Gruppen wie Aziden oder Alkinen, welche durch kupferkatalysierte^[505, 506] oder spannungsinduzierte^[507] Alkin-Azid-Cycloaddition bioorthogonal mit entsprechend modifizierten Affinitätsgruppen verknüpft werden können. Auch eine direkte Konjugation an eine Festphase wurde bereits beschrieben.^[508] Beispiele für diese Strategie sind die nicht-spaltbaren Crosslinker *Two-Piece Leiker*^[504] und CLIP^[509] (Abb. 1.11b) sowie das auf dem Sulfoxid als spaltbare Gruppe basierende DSBSO^[510] (Abb. 1.11c).

Da alle der vorgestellten Moleküle jedoch Defizite wie eine fehlende Spaltbarkeit (BCCL2, Leiker, CLIP), hohe Komplexität (PIR) oder nachteilige Fragmentierungseigenschaften (DSBSO) besitzen, existieren bislang größtenteils konzeptionelle Beweisstudien.

a



b



c

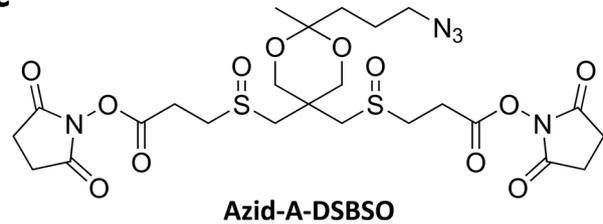


Abbildung 1.11: Übersicht einiger in der Literatur bekannter, Amin-reaktiver und anreicherbarer Crosslinker. a) BCCL2-Crosslinker von *Kang et al.*,^[503] welcher eine Biotin-Einheit für die Aufreinigung mittels Streptavidin enthält. b) CLIP-Crosslinker von *Chowdhury et al.*,^[509] besitzt ein Alkin, welches über CuAAC funktionalisiert werden kann. c) Azid-A-DSBSO-Crosslinker von *Burke et al.*,^[510] Der Crosslinker lässt sich durch das Azid mittels CuAAC funktionalisieren. Durch säurekatalysierte Spaltung des Ketals kann die verknüpfte Gruppe wieder abgespalten werden. Die beiden Sulfoxide sorgen für eine Spaltbarkeit des Crosslinkers in der Gasphase.

2 Zielsetzung

Der Fokus der Doktorarbeit sollte auf der Beantwortung epigenetisch relevanter Fragestellungen mithilfe von massenspektrometrischen Methoden liegen. Hierfür sollten einige bei Beginn dieser Arbeit noch unbekannte Prozesse aus dem Gebiet der mit DNA interagierenden Proteine genauer untersucht werden. Da die Erforschung dieser Zusammenhänge teilweise auf der sich noch in der Entwicklung befindlichen Crosslinking-Massenspektrometrie beruht, sollte hierbei ebenfalls zur Methodenentwicklung beigetragen werden.

Zum einen sollte eine mögliche Schnittstelle der epigenetischen Regulation zwischen DNA und Proteinen genauer beleuchtet werden. So besteht chemisch betrachtet die Möglichkeit, dass sich zwischen DNA, welche die epigenetische Base 5-Formylcytosin enthält, sowie Histonen im Kontext von Chromatin kovalente Addukte zwischen der Formylgruppe und Aminen der Lysin-Seitenketten ausbilden. Diese sogenannten Schiff'schen Basen bilden sich reversibel und könnten einen Einfluss auf die Nukleosom-Positionierung, Chromatinstabilität oder auch aktive Cytosin-Demethylierung besitzen und über die Modifikation der Lysine durch Methyl- oder Acetyltransferasen reguliert werden. Für die Untersuchung sollten *in vitro* Nukleosomen ausgebildet werden, um an diesen die Adduktbildung zu untersuchen. Als DNA sollten dabei mittels PCR hergestellte Stränge mit vielen 5-Formylcytosin-Basen verwendet werden, um einen möglichst großen Effekt zu erzielen. Als Analysemethode sollte hierbei neben gelelektrophoretischen Verfahren, welche erste Anhaltspunkte für die weitere Untersuchung liefern sollten, vor allem Massenspektrometrie verwendet werden. Diese ermöglicht heutzutage bei der Wahl einer geeigneten Probenvorbereitung den detaillierten Aufschluss über die Position von Modifikationen bekannter Masse in Proteinen. Durch eine Stabilisierung der eigentlich labilen Addukte und dem anschließenden Verdau von DNA und Proteinen sollte der am Histon verbleibende DNA-Rest identifiziert und lokalisiert werden, um die kovalente Wechselwirkung direkt nachzuweisen.

Auch eine mögliche Funktion dieser Addukte in Form einer Beteiligung bei der aktiven Demethylierung sollte untersucht werden. Dafür sollte sowohl eine Übertragung der Formylgruppe von der DNA auf die Histone als auch in Kooperation quantitativ eine Auswirkung auf die Menge an Formyl-Modifikation in der DNA analysiert werden. Dies sollte erneut mithilfe von massenspektrometrischen Methoden erfolgen, wobei für Ersteres nach formyl-modifizierten Lysinen in den Histonen gesucht und für Letzteres eine quantitative Analyse von 5-Formylcytosin und Cytosin durchgeführt werden sollte.

Es gibt mittlerweile viele Hinweise darauf, dass die epigenetische Modifikation der DNA und der Metabolismus einer Zelle eng miteinander verknüpft sind. Im Rahmen einer größeren Studie über das Zusammenspiel von TET3 und dem Citratzyklus sollte in dieser Arbeit eine weitere epigenetisch relevante Fragestellung untersucht werden.

Es konnten in der Arbeitsgruppe Hinweise auf eine mögliche direkte Interaktion zwischen dem für die Oxidation von 5-Methylcytosin unter anderem zu 5-Formylcytosin verantwortlichen Enzym TET3 sowie der Glutamatdehydrogenase GDH gefunden werden. Letztere könnte für eine unmittelbare Bereitstellung des für die Oxidation benötigten Cofaktors α -Ketoglutarat verantwortlich sein. Mithilfe von Crosslinking-Massenspektrometrie sollte daher analysiert werden, ob eine solche räumliche Nähe zwischen den beiden Proteinen existiert. Als Crosslinker sollte dabei DSSO eingesetzt und die Probenkomplexität durch Affinitätstag- oder Chromatographie-basierte Methoden reduziert werden.

In Bezug auf die herausfordernde Komplexität des im oberen Abschnitt erwähnten Ansatzes und allgemein zur Verbesserung der Struktur- oder Interaktomanalyse von Proteinen mittels Crosslinking-Massenspektrometrie, sollte zusätzlich ein in dieser Arbeitsgruppe entwickelter Crosslinker etabliert werden. Dieser wurde so konzipiert, dass er eine modulare Funktionalisierung erlaubt. In dieser Doktorarbeit sollte daher zum einen die routinemäßige Verwendung des Moleküls in Crosslinking-Experimenten von geringer Komplexität erreicht werden. Zum anderen sollte der Nutzen der Modularität des Crosslinkers anhand eines Beispiels gezeigt werden. Hierfür sollte eine Anreicherung von Crosslinks aus einer komplexen Probe erfolgen, da diese oft nur einen geringen Teil der gesamten Proteinmenge ausmachen, was deren Identifikation mit gängigen massenspektrometrischen Methoden erschwert. Die Anreicherung sollte auf Grundlage der sehr hohen Affinität von Biotin zu Streptavidin erfolgen. Mithilfe von Kupfer-katalysierter Cycloadditions-Chemie sollten die Crosslinks in einer bioorthogonalen Reaktion mit Biotin funktionalisiert und anschließend an immobilisiertem Streptavidin angereichert werden.

3 Ergebnisse und Diskussion

3.1 Untersuchung von Addukten zwischen DNA-Aldehyden und Histonen

Dieser Teil der Doktorarbeit beschäftigte sich mit der Fragestellung, ob sich zwischen primären Aminen von Histonen und der Formylgruppe von fC in DNA kovalente Interaktionen in Form von Iminen, sogenannte Schiff'sche Basen, bilden können. Diese könnten *in vivo* einen großen Einfluss auf die Genregulation oder Genomstabilität ausüben.

3.1.1 Addukte zwischen einem Peptid und einem einzelsträngigen ODN

Um zunächst die generelle Reaktivität von fC-haltiger DNA mit einem primären Amin zu untersuchen, wurde in einem ersten Schritt die Reaktion eines fC-haltigen DNA-Einzelstrangs mit einem Lysin-haltigen Peptid untersucht. Außerdem wurde auch ein entsprechender fU-haltiger Strang eingesetzt, um eventuelle Unterschiede dieser Formyl-Pyrimidine zu beleuchten. Als Peptid wurde dabei ein 7-mer mit der Sequenz Ac-NH-IEAKGER-COOH (Ac-IR7) verwendet. Dieses Peptid enthält durch die Lysin-Seitenkette ein primäres Amin, welches mit der Formylgruppe unter der Bildung eines Imins reagieren kann. Damit die Reaktion ausschließlich am Lysin-Amin stattfinden kann, wurde der N-Terminus acetyliert, wodurch das freie Elektronenpaar des Stickstoffs nicht mehr für einen nukleophilen Angriff zur Verfügung steht. Als ODN wurde ein 27 bp langer Teil der 601 Sequenz verwendet, welcher an einer Position eine formyl-modifizierte Base enthielt (5'-AGCTGTCTACGACCAGTTXAGCGGCCT-3'; X = fC, fU). Die 601 Sequenz ist eine künstliche Sequenz, die stark mit Histonen interagiert und leicht Nukleosomen ausbildet. Die beiden DNA-Einzelstränge wurden mittels DNA-Festphasensynthese synthetisiert. Dabei wurde die Synthese des fC-Strangs (fC-ODN) von *Dr. Iacovos Michaelides* und *Dr. Franziska Traube* durchgeführt, während der fU-Strang (fU-ODN) im Rahmen dieser Arbeit in Kooperation mit *Dr. Iacovos Michaelides* und *Dr. Franziska Traube* hergestellt wurde.

Abbildung 3.1 zeigt die Kontrolle der erfolgreichen Synthese und Aufreinigung von fC- sowie fU-ODN. Während die HPLC-Chromatogramme der Stränge vor der Aufreinigung durch semi-präparative HPLC (Abb. 3.1a und 3.1d) noch mehrere Verunreinigungen aufweisen, zeigen die Chromatogramme des jeweiligen aufgereinigten Stranges (Abb. 3.1b und 3.1e), dass die Stränge in hoher Reinheit erhalten wurden. Die MALDI-TOF-Analysen (Abb. 3.1c und 3.1f) bestätigen zudem die theoretisch berechneten molekularen Massen der gewünschten ODNs.

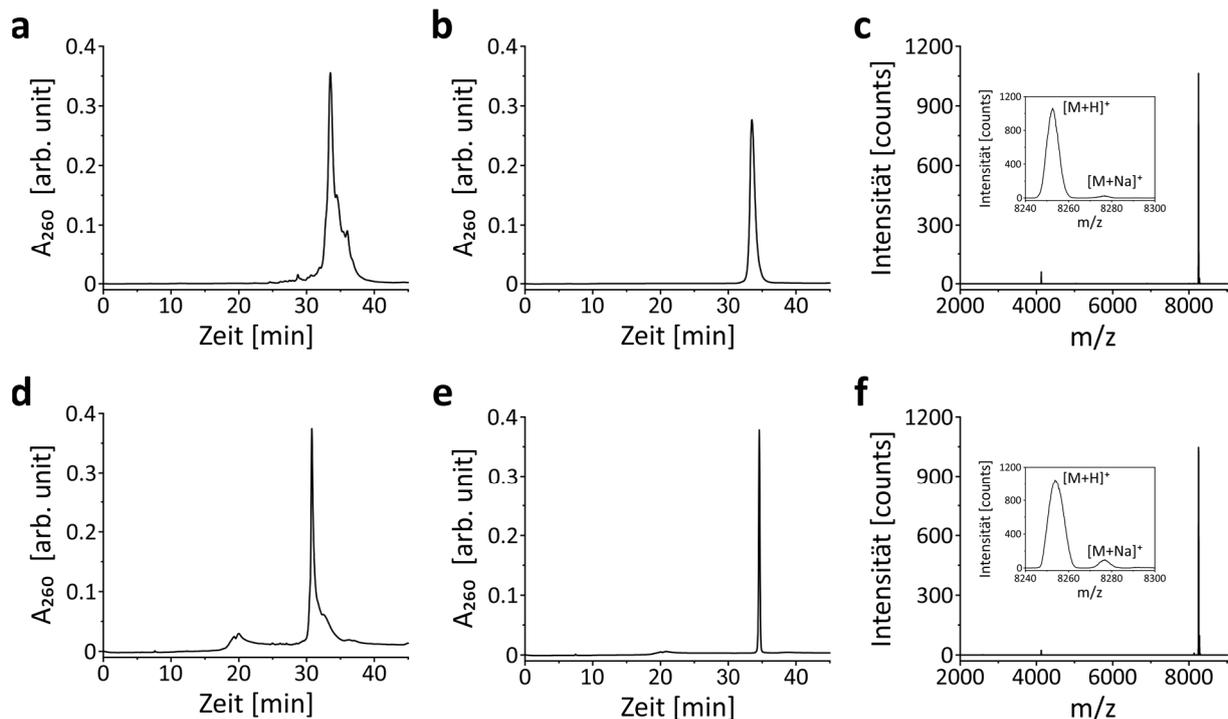
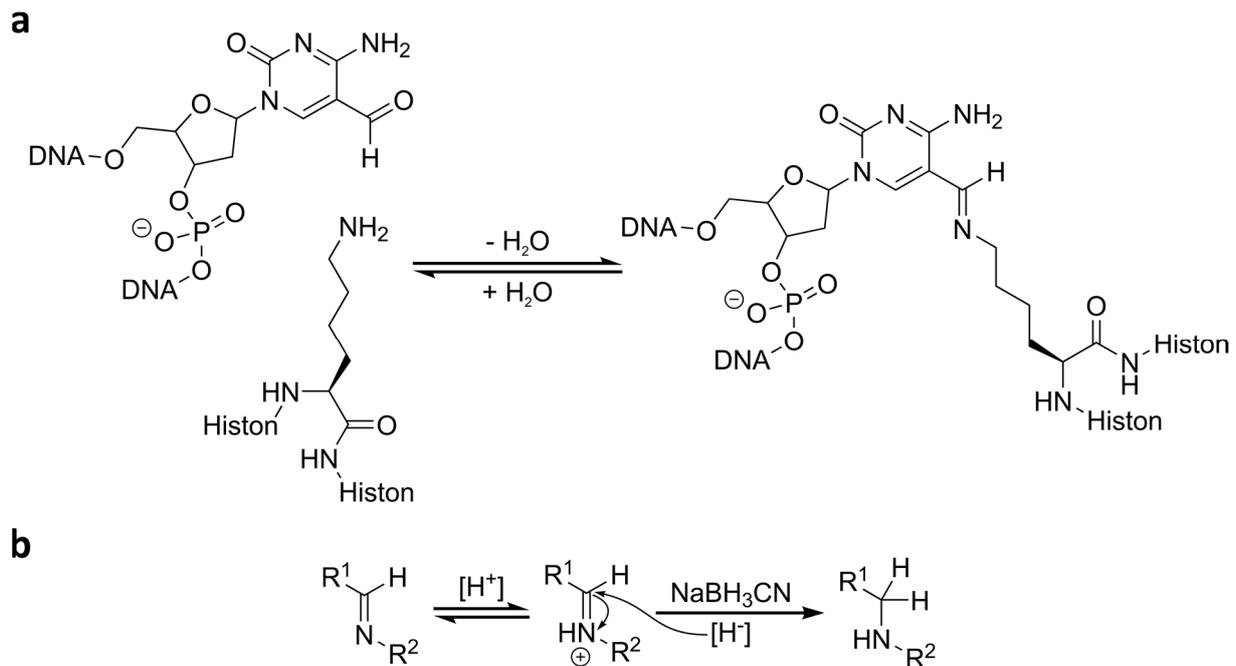


Abbildung 3.1: Ergebnisse der Synthesen des fC sowie fU-haltigen DNA-Einzelstranges. a) HPLC-Chromatogramm des fC-Stranges vor der Aufreinigung. b) HPLC-Chromatogramm des fC-Stranges nach Aufreinigung mittels semi-präparativer HPLC. c) MALDI-TOF-MS-Spektrum des aufgereinigten fC-Stranges. d) – f) wie a) – c), jedoch mit fU-Strang.

Während der Inkubation von ODN und Peptid können sich Schiff'sche Basen zwischen dem primären Amin und der Formylgruppe ausbilden (Schema 3.1a). Da diese jedoch gegenüber Hydrolyse labil sind, wurden sie in einem nächsten Schritt reduziert, wie es bei einer reduktiven Aminierung der Fall ist. Als Reduktionsmittel kam dabei NaBH_3CN zum Einsatz, welches als Hydrid-Donor das Imin zu einem in Wasser stabilen sekundären Amin reduzieren kann (Schema 3.1b). Dabei hat es den Vorteil gegenüber z. B. NaBH_4 , dass es weniger stark reduzierend ist. Bei physiologischem pH-Wert reduziert es freie Formylgruppen daher nicht zu Hydroxymethyl-Gruppen.^[511] Somit kann sich das Gleichgewicht der Imin-Bildung während der Reduktion weiter in Richtung Produktseite verschieben und man erhält potentiell mehr Addukte. Zum Beenden der Reduktion wurde schließlich NaBH_4 zugegeben, wodurch alle noch nicht reagierten Formylgruppen reduziert werden und damit für eine Reaktion nicht mehr zur Verfügung stehen.

Um die Addukt-Bildung zu beschleunigen, wurden in manchen Proben Katalysatoren verwendet, welche den nukleophilen Angriff des Peptid-Amins über die Bildung eines intermediären Imins zwischen Katalysator und Formylgruppe erleichtern. Zum einen kam dabei *p*-Anisidin (Kat. 1) zum Einsatz, da bereits bekannt ist, dass dieses Reagenz in Kombination mit einem pH-Wert von 5 – 6 die Reaktion von fC bzw. fU mit Oxyaminen, Hydraziden und aromatischen Aminen katalysiert.^[138, 175, 512, 513]



Schema 3.1: Ausbildung und Stabilisierung der Schiff'schen Basen zwischen DNA und Histonen. a) Darstellung der reversiblen Reaktion von fC-haltiger DNA und dem Lysin eines Histons zu einem Imin. b) Stabilisierung des Imins durch Reduktion mittels NaBH_3CN .

Der zweite verwendete Katalysator (Kat. 2) war *p*-Diaminobenzol, welcher ebenfalls bei einem leicht sauren pH-Wert eingesetzt wurde. *Su et al.* konnten bereits zeigen, dass die Oxim-Bildung von fC mit einem Oxyamin durch diesen Katalysator noch stärker beschleunigt wird als bei der Verwendung von *p*-Anisidin.^[513]

Die Analyse der Addukt-Bildung erfolgte mittels denaturierender Harnstoff-PAGE. Hierbei werden die Moleküle anhand ihrer Größe aufgetrennt und nicht-kovalente Interaktionen durch die hohe Konzentration an Harnstoff aufgelöst. Kleinere Moleküle bewegen sich schneller durch das Gel, weshalb sich die Bildung von ODN-Peptid-Addukten anhand einer Verschiebung der DNA zu höheren Molekulargewichten nachverfolgen lassen sollte. Die Visualisierung der DNA erfolgte dabei über ein DNA-interkalierendes Fluorophor.

In Abbildung 3.2a sind die Ergebnisse dieses Versuchs dargestellt. Wurde die Inkubation von ODN und Peptid in Gegenwart von Puffer mit einem pH-Wert von 7.4 und ohne Katalysator durchgeführt, konnte keine Reaktion zwischen Peptid und ODN erreicht werden. Dasselbe gilt für die Reaktion von fC-ODN und Peptid unter leicht sauren Bedingungen und Katalyse durch *p*-Anisidin. Bei fU-ODN unter den gleichen Bedingungen erkennt man jedoch zusätzlich zur unreaktierten Bande oberhalb von dieser eine diffuse weitere Bande. Diese stellt das Addukt zwischen ODN und Peptid dar, welches aufgrund seiner größeren räumlichen Struktur und geringeren Ladungsdichte langsamer durch das Gel läuft. Der größte Teil der Intensität ist jedoch weiterhin bei der Bande der unreaktierten DNA zu finden, was bedeutet, dass nur ein kleiner Teil der DNA Addukte mit dem Peptid eingegangen ist.

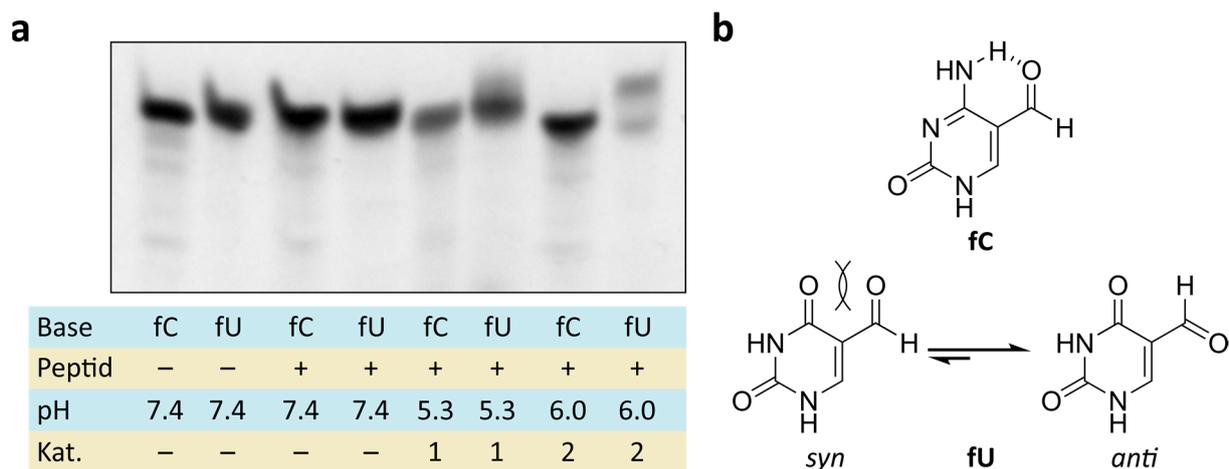


Abbildung 3.2: Reaktion von fC- und fU-ODN mit Ac-IR7. a) Harnstoff-PAGE der verschiedenen Reaktionsbedingungen (Kat. 1: *p*-Anisidin, Kat. 2: *p*-Diaminobenzol). b) Unterschiedliche Konformationen der Basen fC und fU.

Ein deutlicherer Umsatz lässt sich bei der Verwendung von *p*-Diaminobenzol als Katalysator erkennen. Während der fC-Strang weiterhin keine Reaktion zeigt, sieht man für das fU-ODN eine deutliche Aufteilung der Intensität in zwei Banden. Unter Verwendung dieses Katalysatorsystems bilden sich also mehr Schiff'sche Basen als mit *p*-Anisidin. Dieser Unterschied der Effektivität der Katalysatoren steht wie bereits erwähnt im Einklang zur Literatur.^[513] Der fC-haltige Strang zeigt in keiner Probe eine Reaktion, was an der geringeren Reaktivität gegenüber Nucleophilen der Formylgruppe von fC im Vergleich zu fU liegt. *Hardisty et al.* kommen anhand von Experimenten mit fC/fU und Oxyaminen sowie Hydraziden zu demselben Ergebnis.^[512] Der Grund hierfür könnte die höhere Elektronendichte des Cytosin-Ringsystems im Vergleich zu Uracil sein, wodurch der nukleophile Angriff des Amins erschwert wird. Zudem weisen Studien darauf hin, dass für fU im Gegensatz zu fC die *anti*-Konformation bevorzugt vorliegt (s. Abb. 3.2b).^[514, 515] Während Amin und Formylgruppe bei fC in der *syn*-Konformation eine Wasserstoffbrückenbindung ausbilden können, kommt es für fU zu sterischer Hinderung zwischen den beiden Oxo-Gruppen. Dies könnte ebenfalls einen Einfluss auf die Reaktivität haben. Die Literatur zeigt jedoch, dass fC, wenn auch in geringerem Ausmaß als fU, mit Nucleophilen reagiert.^[512, 513] Als Endprodukte entstehen bei den verwendeten Oxyaminen und Hydraziden jedoch gegenüber Hydrolyse viel stabilere Oxime bzw. Hydrazone als das bei der Reaktion mit einem primären Amin gebildete Imin, dessen Bildung nur eine Gleichgewichtsreaktion ist.^[516, 517] Daher ist eine geringere Reaktivität zwischen Formylgruppe und Lysin-Seitenketten-Amin plausibel.

Zusammenfassend konnte gezeigt werden, dass ein fU-haltiges ODN in Anwesenheit von verschiedenen Katalysatoren unter leicht sauren Bedingungen Addukte mit einem kurzen, Lysin-haltigen Peptid bildet. Unter denselben Bedingungen reicht die Reaktivität der Formylgruppe von fC gegenüber dem nukleophilen Angriff des primären aliphatischen Amins von Lysin allerdings nicht aus, um eine Reaktion einzugehen.

3.1.2 Addukte zwischen dsDNA und Histonen in rekonstituierten Nukleosomen

Obwohl das unter 3.1.1 beschriebene Experiment mit einem kurzen fC-haltigen ODN und einem Lysin-haltigen Peptid eine geringe Reaktivität der Formylgruppe gegenüber dem nukleophilen Angriff des Amins nahelegt, wurde diese Reaktion genauer untersucht. Im Zellkern bilden DNA und Histone Chromatin, was mit einem geringen räumlichen Abstand der Makromoleküle einhergeht. Dadurch entsteht eine sehr hohe effektive Molarität an den Stellen, an denen eine Formylgruppe und ein Lysin-Amin in räumlicher Nähe vorliegen. Ähnlich wie bei intramolekularen oder enzymatischen Reaktionen^[518] könnte dadurch die Reaktionsgeschwindigkeit stark erhöht werden.

Zur Nachbildung der Chromatinstruktur *in vitro* wurden rekombinante Histone (human, H2A, H2B, H3.1, H4) sowie eine kurze DNA-Sequenz verwendet. Als DNA wurde die stark positionierende 601 Sequenz^[261] verwendet, was bedeutet, dass sie mit dem Histon-Oktamer definierte Nukleosomen ausbildet. Es wurde bereits von *Vasudevan et al.* gezeigt, dass ein 145 bp langes Fragment dieser Sequenz für die Nukleosom-Bildung ausreicht,^[519] weshalb in dieser Arbeit mit dem besagten Teil der 601 DNA gearbeitet wurde. Dieses Modell ist eine gute Repräsentation der Grundeinheit von Chromatin, der Nukleosomen, welche aus der um die Histone gewundenen DNA bestehen und wird standardmäßig für Untersuchungen an Chromatin *in vitro* verwendet.^[520-522]

Um fC bzw. fU in den Strang einzubauen, wurde das unmodifizierte Templat mittels PCR vervielfältigt, wobei im Falle der fC-haltigen DNA (601_{fC}) 5-Formyl-dCTP statt dCTP und für die fU-haltige DNA (601_{fU}) 5-Formyl-dUTP statt dTTP verwendet wurde. Dadurch erhält man DNA-Stränge, welche bis auf die in den Primern enthaltenen Basen nur noch die modifizierten Basen enthalten. Wie bereits von *Steigenberger et al.* beschrieben,^[523] musste eine spezielle Polymerase verwendet sowie die Extensionszeit erhöht werden, um einen effizienten Einbau der modifizierten Triphosphate zu erreichen. Die erfolgreiche Synthese der modifizierten Stränge ist in Abbildung 3.3a dargestellt. Man erkennt auf dem Agarose-Gel sowohl für 601_{fC} als auch 601_{fU} eine Bande auf derselben Höhe, welcher durch die Verwendung eines DNA-Markers eine Größe zwischen 50 und 200 bp zugeordnet werden kann. Dies stimmt mit der theoretischen Größe der 601 DNA überein. Es sind keine weiteren DNA-Banden zu erkennen, was bedeutet, dass nur die gewünschte Sequenz als definiertes Produkt synthetisiert wurde.

Interessanterweise sieht man einen großen Unterschied in der Intensität zwischen 601_{fC} und 601_{fU}, obwohl dieselbe Menge an DNA auf das Gel aufgetragen wurde. Dies könnte daran liegen, dass die hohe Anzahl an fU-Substitutionen die DNA-Konformation durch die bereits in Kapitel 3.1.1 und Abbildung 3.2b beschriebene *anti*-Konformation so verändert, dass die Basenpaarung mit Adenin beeinträchtigt wird und der Fluoreszenzfarbstoff daher nicht mehr interkalieren kann.^[514, 524]

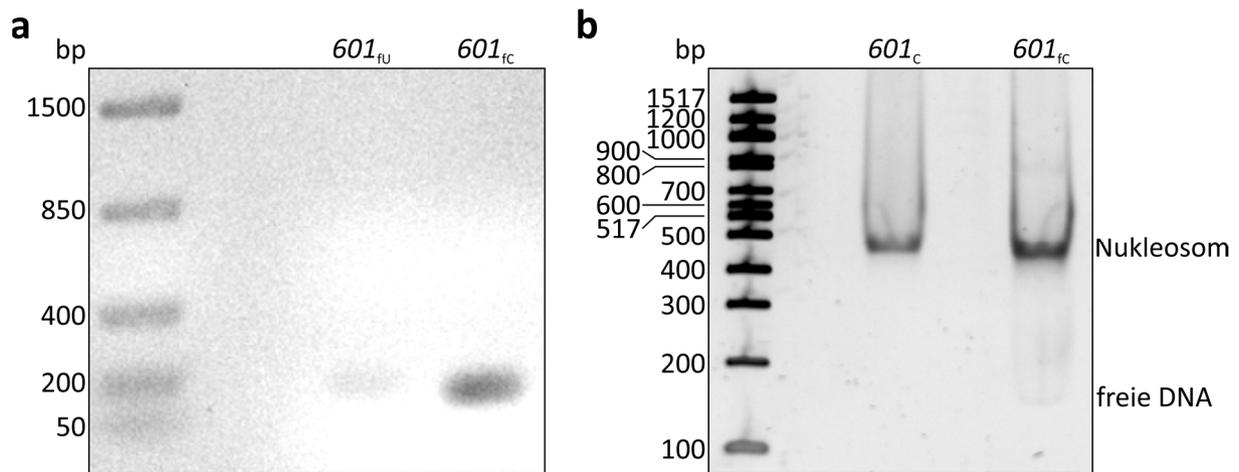


Abbildung 3.3: PCR von fC- bzw. fU-haltiger dsDNA mittels PCR und anschließende Nukleosom-Rekonstitution: a) Agarose-Gel der PCR-Synthesen von 601_{fU} sowie 601_{fC} . Der Kontrast und die Helligkeit dieser Aufnahme wurden nachträglich angepasst. b) Native PAGE der Nukleosom-Rekonstitution mit 601_c und 601_{fC} . Durch die Ausbildung des Nukleosoms wird die Bande der DNA zu höheren Molekulargewichten verschoben.

Die mittels PCR erhaltenen DNA-Stränge wurden im Anschluss für die Bildung der Nukleosomen *in vitro* verwendet. Diese erfolgte durch eine schrittweise Verringerung der NaCl-Konzentration einer Lösung des H2A/H2B-Dimers, H3.1/H4-Tetramers sowie der DNA von 2 M auf 0.25 M. Die Rekonstitution beginnt bei dieser Methode bei 1.2 – 1.0 M NaCl, indem zunächst das H3.1/H4-Tetramer mit der DNA interagiert. Ab einer NaCl-Konzentration von 0.8 M lagern sich dann auch die beiden Exemplare des H2A/H2B-Dimers an und bei einer Konzentration von 0.6 M NaCl ist die Nukleosom-Bildung abgeschlossen.^[525] Bei einer weiteren Verringerung der NaCl-Konzentration auf 0.1 M und besonders bei einem zu hohen Überschuss an Histonen relativ zur DNA, beginnt das Nukleosom wieder zu dissoziieren oder auch Aggregate mit anderer Zusammensetzung zu bilden,^[525, 526] weshalb für die Rekonstitution ein Verhältnis von Tetramer zu DNA von 1:1 und Dimer zu DNA von 2:1 gewählt wurde. Anstelle des häufig verwendeten TRIS-Puffers wurde für die Rekonstitution HEPES-Puffer verwendet, da TRIS ein primäres Amin besitzt, welches ebenfalls mit den Formylgruppen von fC und fU reagieren und somit die Bildung der Schiff'schen Basen zwischen DNA und Histonen verhindern könnte.

Die Analyse der rekonstituierten Nukleosomen geschah mittels nativer Polyacrylamid-Gelelektrophorese (PAGE), vergleichbar mit einem EMSA-Experiment. Hierfür wurde ein Teil der Nukleosomen auf ein nicht-denaturierendes Gel aufgetragen und anschließend eine Spannung entlang des Gels angelegt. Dabei werden die erfolgreich rekonstituierten Nukleosomen von Molekülen anderer Größe wie freier DNA, einzelnen Histonen oder unvollständig bzw. übermäßig assoziierten DNA-Protein-Aggregaten getrennt. Dies liegt an der veränderten Größe und Ladung eines DNA-Protein-Komplexes, der sich langsamer durch das Gel bewegt als ungebundene DNA oder Proteine.

Die Dissoziation des Komplexes wird dabei durch einen „Käfig-Effekt“ der Gelmatrix verhindert,^[527] zudem wurde die Elektrophorese bei 4 °C durchgeführt, um eine Denaturierung aufgrund der Aufheizung des Gels während der Elektrophorese zu vermeiden. Die Visualisierung erfolgte mithilfe eines in die DNA interkalierenden Fluorophors, welches die Detektion der verschiedenen DNA-haltigen Fraktionen (ungebunden und in Nukleosomen) erlaubt. In Abbildung 3.3b ist eine solche Gelelektrophorese für die Nukleosomen mit 601_c sowie 601_{fc} dargestellt. Hierbei ist zu beachten, dass der verwendete und hauptsächlich für Agarose-Gelelektrophorese konzipierte DNA-Marker (100 bp *ladder*, *New England Biolabs*) durch die Eigenschaften des Polyacrylamid-Gels eine zusätzliche Bande zeigt.^[528] Da eine genaue Größenbestimmung jedoch nicht das Ziel der Experimente war, wurde der Marker nur als Anhaltspunkt für die Unterscheidung zwischen freier und Nukleosom-gebundener DNA verwendet. Man sieht für beide Proben eine Bande zwischen 400 und 500 bp, welche laut Kit-Hersteller der DNA entspricht, die nukleosomal vorliegt. Für 601_{fc} lässt sich noch schwach eine Bande zwischen 100 und 200 bp erkennen, die einen kleinen Rest freier DNA darstellt. Da der verwendete Fluorophor allerdings in die kleine Furche der DNA interkaliert^[529] und ein Teil der kleinen Furchen in Richtung des Histon-Oktamers zeigt und durch die Windung der DNA zudem teilweise stark gekrümmt ist,^[519] ist es wahrscheinlich, dass die ungebundene DNA für die Interkalation des Fluorophors zugänglicher ist als die nukleosomale. Somit ist der Unterschied zwischen diesen beiden Populationen vermutlich noch größer und man kann davon ausgehen, dass die Rekonstitution in beiden Fällen mit einer sehr hohen Ausbeute erfolgreich war. Auch Aggregate höherer Ordnung, also mit mehr als einem Oktamer assoziierte DNA, welche oberhalb der Nukleosom-Bande sichtbar wären, treten nicht auf. Es wurde also nur eine definierte Spezies gebildet.

Auch mit 601_{fu} wurde eine Nukleosom-Rekonstitution durchgeführt und analysiert, allerdings konnte weder für freie noch für nukleosomale DNA eine Bande beobachtet werden, was wahrscheinlich an der bereits nach der PCR festgestellten und diskutierten niedrigeren Intensität des fU-Strangs bei Anfärbung mit einem interkalierenden Fluorophor liegt. Die weiteren gelelektrophoretischen Experimente wurden daher nur mit 601_c sowie 601_{fc} durchgeführt, da der Fokus dieser Arbeit auf Addukten zwischen fC und den Histonen lag. Für entsprechende Experimente mit 601_{fu} könnten in zusätzlichen Experimenten z. B. fluoreszierende PCR-Primer verwendet werden.

Mit den rekonstituierten Nukleosomen, welche entweder unmodifizierte oder fC-haltige DNA enthielten, wurde nun zunächst elektrophoretisch untersucht, ob kovalente Addukte ausgebildet werden. Dafür wurden die Proben über Nacht inkubiert, damit sich die Schiff'schen Basen ausbilden konnten. Um eine Dissoziation der Nukleosomen zu verhindern, wurde diese Inkubation bei 18 °C durchgeführt. Ein Teil der Proben wurde dabei mit NaBH_3CN reduziert, um die sich bildenden Imine zu sekundären Aminen umzusetzen und so gegenüber Hydrolyse zu stabilisieren.

Von den so erhaltenen Proben wurde wieder ein Teil für kurze Zeit bei 95 °C erhitzt, um die Tertiärstruktur des Komplexes aufzuheben und die nicht-reduzierten Addukte möglichst zu hydrolysieren. Die Analyse erfolgte erneut mittels Polyacrylamid-Gelelektrophorese; diesmal wurden jedoch durch die Verwendung von SDS denaturierende Bedingungen gewählt, um nicht-kovalente Wechselwirkungen zu unterbinden. Dadurch sollte eine Verschiebung der DNA zu höheren Molekulargewichten auf dem Gel nur noch möglich sein, wenn sie kovalent mit einem oder mehreren Histonen verknüpft ist. Da im Probenpuffer normalerweise vorhandene Thiole, wie DTT oder β -Mercaptoethanol, mit der Formylgruppe von fC zu Hemi- oder Dithioacetalen reagieren könnten^[530] und somit ebenfalls zu einer Verschiebung der DNA-Bande führen würden, wurde ein Thiol-freier Probenpuffer verwendet. Abbildung 3.4a zeigt die Gele der unterschiedlich behandelten Nukleosomen. Die Proben mit 601_C zeigen sowohl mit als auch ohne Reduktion nur eine Bande zwischen 100 und 200 bp, welche der ungebundenen DNA entspricht. Dies stellt die Negativkontrolle dar, da die unmodifizierte DNA keine kovalenten Interaktionen mit den Histonen ausbilden kann und bestätigt außerdem, dass die denaturierenden Bedingungen der SDS-PAGE ausreichen, um den Komplex zwischen DNA und Histonen zu dissoziieren. Anders sieht es bei den mit 601_{fC} rekonstituierten Nukleosomen aus. In allen Proben sieht man, dass ein Teil der DNA ebenfalls in ungebundener Form vorliegt, also dissoziiert ist. Allerdings gibt es auch Banden bei höheren Molekulargewichten, was bedeutet, dass die DNA noch mit einem oder mehreren Histonen assoziiert sein könnte. Diese Banden sind teilweise sehr diffus, was darauf hindeutet, dass die DNA-Protein-Komplexe nicht immer definiert sind, sondern eventuell verschiedene Stöchiometrien von DNA und den einzelnen Histonen oder auch unterschiedliche Bindungsstellen aufweisen.

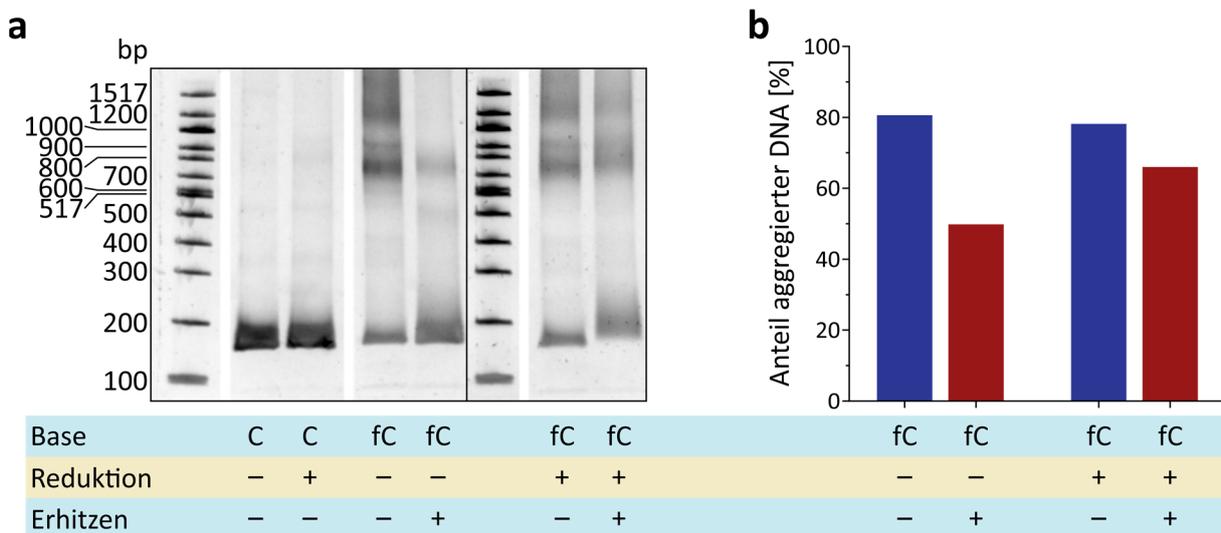


Abbildung 3.4: Denaturierende gelelektrophoretische Analyse der rekonstituierten und teilweise reduzierten Nukleosomen. a) SDS-PAGE der unterschiedlich behandelten Nukleosomen. b) Quantitative Auswertung des Anteils der aggregierten DNA in den Proben mit 601_{fC}.

Um Unterschiede im Anteil an Protein-gebundener DNA zwischen den einzelnen Proben besser einschätzen zu können, wurde der Anteil der DNA-Protein-Aggregate quantitativ ausgewertet (Abb. 3.4b). Dafür wurde die Intensität des Bereichs über der 400 bp-Markerbande, der alle teilweise diffusen Aggregat-Banden enthält, durch die Gesamtintensität der jeweiligen Spur geteilt, um zu bestimmen, wie viel Prozent der DNA an Histone gebunden sind. Interessanterweise sieht man, dass bereits in der nicht-reduzierten Probe ca. 80 % der DNA aggregiert vorliegen. Die Stabilität der reversibel kovalenten Interaktionen scheint also bereits so hoch zu sein, dass die denaturierenden Bedingungen nicht ausreichen, um den DNA-Protein-Komplex zu dissoziieren. Das könnte daran liegen, dass einige Imine nicht hydrolysiert werden können, da sie aufgrund der Struktur des Nukleosom-Komplexes nicht Lösungsmittel-zugänglich sind. Aufgrund der Verwendung von DNA mit vielen modifizierten Cytosinen ist es aber auch möglich, dass an verschiedenen Stellen innerhalb eines Nukleosoms gleichzeitig Addukte gebildet werden, sodass die Hydrolyse einzelner Interaktionspunkte keine besonderen Auswirkungen auf den gesamten Komplex hat. Betrachtet man die nicht-reduzierte Probe nach Inkubation bei 95 °C, sieht man, dass der Anteil an assoziierter DNA auf unter 50 % sinkt. Durch das Erhitzen werden Sekundär- und Tertiärstruktur der Histone aufgehoben und auch der DNA-Duplex schmilzt, sodass die Hydrolyse der Schiff'schen Basen vermutlich besser ablaufen kann. Anscheinend bildet jedoch bereits das nicht-reduzierte Nukleosom mit fC-haltiger DNA so stabile Addukte, dass diese kurze thermische Denaturierung nicht ausreicht, um die Addukte komplett zu zerstören. Eventuell könnte man eine komplette Dissoziation durch eine Verlängerung der Inkubationszeit erreichen.

Auch beim Vergleich der reduzierten Proben zeigt sich diese Tendenz, allerdings ist die Abnahme der aggregierten DNA bei Erhitzen hier nicht so stark wie für das nicht-reduzierte Nukleosom. Durch die Reduktion werden stabile sekundäre Amine gebildet, welche weder bei erhöhter Temperatur noch Verlust der Tertiärstruktur des Komplexes dissoziieren können. Offensichtlich werden jedoch nicht alle Schiff'schen Basen reduziert, da man sonst keinen Unterschied gegenüber der Probe, welche nicht erhitzt wurde, erwarten würde. Eventuell ist ein Teil der Addukte also nicht nur, wie oben bereits ausgeführt, gegenüber der Hydrolyse, sondern auch gegenüber der Reduktion geschützt.

Damit ausgeschlossen werden kann, dass die Banden bei höheren Molekulargewichten auf einen unspezifischen Effekt zurückzuführen sind, wurde eine SDS-PAGE mit zwei Kontrollproben durchgeführt (Abb. 3.5a). Man sieht, dass 601_{fC} allein (Spur 1) ebenso wie eine kurz vor der Elektrophorese hergestellte Lösung aus 601_{fC} und Histonen ohne Rekonstitutions-Protokoll (Spur 2) keine Addukt-Banden erzeugt. Die Rekonstitution von DNA und Histonen zu Nukleosomen ist also notwendig, was darauf hinweist, dass die Banden durch spezifische Interaktionen zwischen der fC-haltigen DNA und dem Histone-Oktamer entstehen, welche von der Nukleosom-Struktur vermittelt werden.

Um zu zeigen, dass die auf dem Gel sichtbaren Addukt-Banden zusätzlich zu DNA auch Protein enthalten, wurde das Experiment für 601_{fc} ohne (Spur 3) und mit (Spur 4) Reduktion wiederholt (Abb. 3.5b). Daraufhin wurden einige Bereiche der reduzierten Probe ausgeschnitten, ein In-Gel-Protein-Verdau durchgeführt und massenspektrometrisch analysiert. Abbildungen 3.5c und d zeigen die Anzahl der identifizierten Peptid-Spektrums-Zuordnungen (PSMs) der jeweiligen Histone in den ausgeschnittenen Banden. Diese Analyse ist lediglich semi-quantitativer Natur und erlaubt daher nur einen Vergleich zwischen den Histonen in einer Probe, da die unterschiedlichen Größen der ausgeschnittenen Gelstücke einen Einfluss auf die Gesamt-Proteinmenge haben.

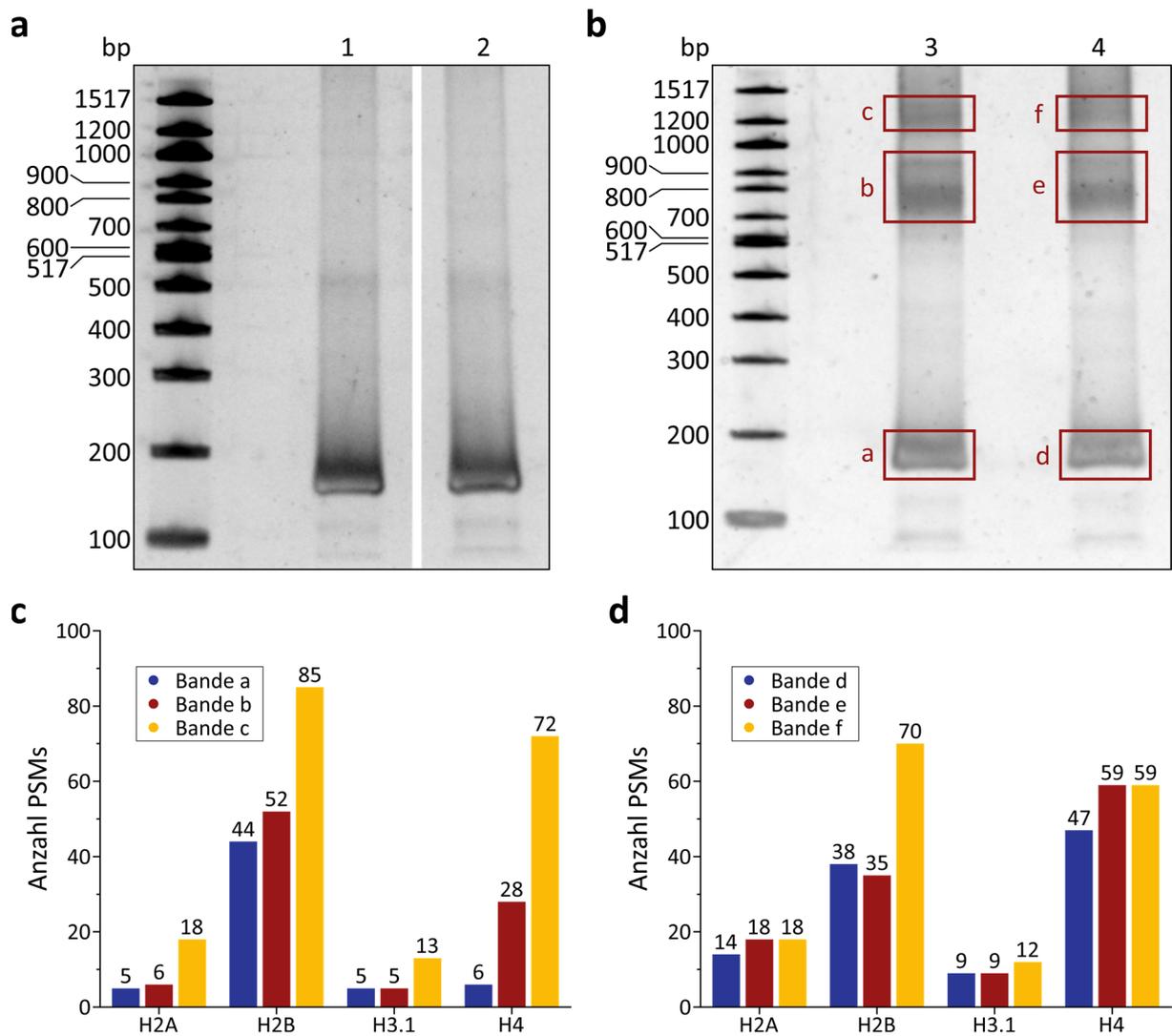


Abbildung 3.5: Kontroll-Experimente zur Konsolidierung der erhaltenen Ergebnisse. a) SDS-PAGE von 601_{fc} ohne Histone (Spur 1) und einer Lösung von 601_{fc} und Histonen welche nicht rekonstituiert wurde (Spur 2). b) SDS-PAGE von nicht-reduzierten (Spur 3) sowie reduzierten (Spur 4) 601_{fc} -Nukleosomen. Die roten Bereiche stellen die für die massenspektrometrische Analyse ausgeschnittenen Gelregionen dar. c) Massenspektrometrische Analyse der ausgeschnittenen Gelbanden der nicht-reduzierten Nukleosomen. d) Massenspektrometrische Analyse der ausgeschnittenen Gelbanden der reduzierten Nukleosomen.

Wie man sieht, konnten in jeder Bande Peptide von allen vier Kern-Histonen identifiziert werden. Es wurden allerdings auch in den Banden mit unmodifizierter DNA (Banden a und d) einige PSMs gefunden. Man würde erwarten, dass nur in den verschobenen Banden Histone identifiziert werden können, da die freie DNA nicht mehr mit Proteinen assoziiert ist. Möglicherweise stammen die Identifikationen in den Banden a und d jedoch daher, dass sich ein Teil der nicht kovalent an die DNA gebundenen Histone mit vergleichbarer Geschwindigkeit wie die freie DNA durch das Gel bewegt, wodurch die Banden von DNA und Protein später unabhängig voneinander auf einer ähnlichen Höhe lägen. Für die Banden bei höheren Molekulargewichten ist dies unwahrscheinlicher, da die Molekulargewichte der einzelnen Histone relativ gering sind. Zudem ist die Zahl der PSMs in den Banden b, c, e und f in fast allen Fällen höher als die der jeweiligen freien DNA-Banden, was ein Hinweis auf eine höhere Proteinmenge sein kann. Besonders auffällig ist dieser Trend für die Histone H2B sowie H4, was eine Interaktion dieser Proteine mit der DNA nahelegt. Zusammenfassend weisen diese Ergebnisse also darauf hin, dass die beobachtete Verschiebung der DNA zu höheren Molekulargewichten wahrscheinlich von der Interaktion mit den Histonen stammt.

Um den Effekt der Formylgruppe auf die Addukt-Bildung nachzuweisen, wurde ein weiteres Experiment durchgeführt, bei dem rekonstituierte Nukleosomen mit Hydroxylamin behandelt wurden. Dieses reagiert mit der Formylgruppe zu einem Oxim (Abb. 3.6a), welches daraufhin nicht mehr für die Reaktion mit einem Lysin-Amin zur Verfügung steht. Die Derivatisierung erfolgte dabei entweder vor oder nach der Inkubation der Nukleosomen mit bzw. ohne NaBH_3CN . In Abbildung 3.6b ist das SDS-Polyacrylamidgel dieses Versuchs dargestellt. Erneut wurde der Anteil der aggregierten DNA ausgewertet, um den Vergleich zwischen den Proben zu erleichtern (Abb. 3.6c). Wie man sieht, ist der Anteil der DNA in Aggregaten bei den 601_{rc} -Nukleosomen ohne Hydroxylamin-Zugabe wie zuvor sehr hoch. Die Probe, bei der die Hydroxylamin-Zugabe vor der Inkubation ohne Reduktion erfolgte, weist hingegen einen weitaus geringeren Anteil von ca. 32 % auf, was zeigt, dass die Blockierung der Formylgruppen die Ausbildung der Addukte beeinträchtigt. Diese Abnahme ist größer als bei der in Abbildung 3.4 gezeigten Erhitzung, was ein weiterer Hinweis darauf ist, dass die Stabilität der lediglich Imin-verknüpften Nukleosomen auf einem kooperativen Effekt mehrerer gleichzeitig vorliegender Addukte beruht. Ist ein Großteil der Formylgruppen derivatisiert, reichen eventuell verbleibende Addukte nicht aus, um den denaturierenden Bedingungen der SDS-PAGE zu widerstehen. Bei der Probe, bei der Hydroxylamin erst nach der Inkubation ohne NaBH_3CN zugegeben wurde, lässt sich der gleiche Effekt beobachten. Allerdings ist die Abnahme der assoziierten DNA hierbei nicht so stark ausgeprägt. Dies liegt vermutlich daran, dass sich durch die vorangegangene Inkubationszeit mehr Addukte ausbilden konnten und die anschließende 15-minütige Inkubation mit NH_2OH nicht ausreichte, um mit allen Formylgruppen zu reagieren.

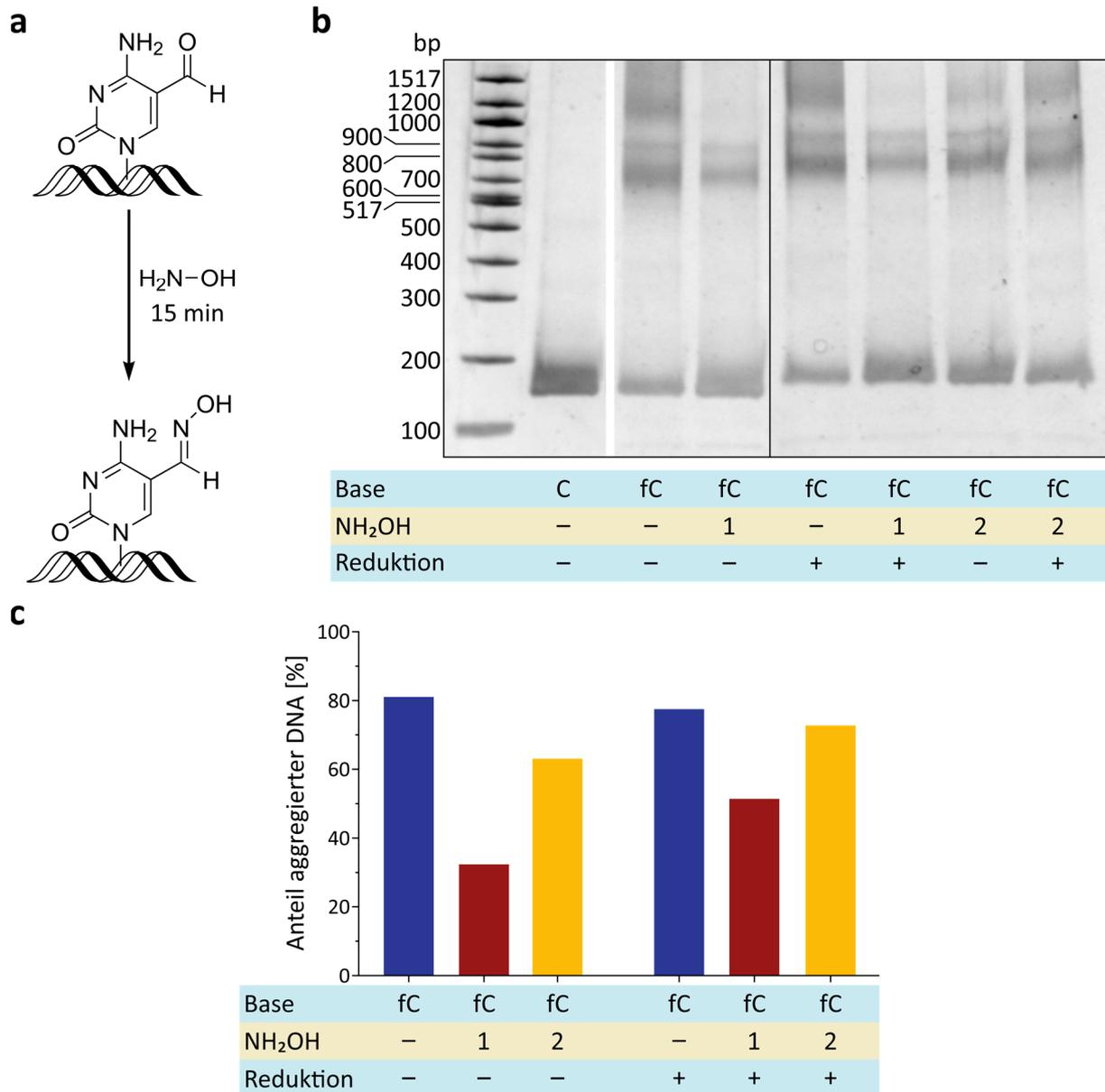


Abbildung 3.6: Auswirkungen der Behandlung mit Hydroxylamin auf die Stabilität der 601_{fC}-Nukleosomen. a) Reaktion von Hydroxylamin mit der Formylgruppe von fC. b) SDS-PAGE der unterschiedlich behandelten Nukleosomen. Die Hydroxylamin-Zugabe erfolgte zum einen vor der Inkubation mit bzw. ohne NaBH₃CN (1) oder danach (2). c) Quantitative Auswertung des Anteils der aggregierten DNA in den Proben mit 601_{fC}.

Im Prinzip zeigt sich bei den Hydroxylamin-behandelten Proben mit Reduktion ein ähnliches Bild. Wie bereits im Experiment, bei dem die Nukleosomen erhitzt wurden, beobachtet werden konnte, sind die Verluste der DNA-Protein-Konjugate jedoch nicht so ausgeprägt wie bei den Proben ohne Reduktion: so sinkt der Anteil der aggregierten DNA bei Hydroxylamin-Zugabe vor der Reduktion nicht so stark, da die wenigen nach der Derivatisierung eventuell noch verbleibenden Addukte durch die anschließende Reduktion stabilisiert werden. Analog dazu hat die Derivatisierung mit Hydroxylamin nach der Reduktion beinahe keinen Effekt, da die Imine bereits zu sekundären Aminen reduziert wurden.

Insgesamt lassen die durchgeführten Gelelektrophorese-Experimente darauf schließen, dass sich zwischen fC-haltiger DNA und Histonen im Kontext von Nukleosomen Addukte ausbilden, an denen die Formylgruppe von fC beteiligt ist. Diese Addukte sind recht stabil gegenüber denaturierenden Bedingungen, was darauf zurückzuführen sein könnte, dass sie entweder schlecht Lösungsmittel-zugänglich sind oder an verschiedenen Stellen in demselben Molekül gleichzeitig auftreten.

Da die gelelektrophoretischen Versuche jedoch nur indirekte Hinweise darauf geben können, ob formyl-haltige DNA und Histon-Amine tatsächlich Addukte bilden, sollten diese mittels Protein-Massenspektrometrie direkt nachgewiesen werden. Dafür wurde ein *bottom-up* Ansatz gewählt, bei dem man die Proteine zunächst zu Peptiden verdaut, welche dann analysiert werden. Dies hat den Vorteil, dass die genaue Position von Modifikationen innerhalb eines Proteins ermittelt werden kann. Da diese Modifikationen allerdings recht klein sein sowie eine definierte Masse haben müssen, war es notwendig, nach der Nukleosom-Rekonstitution und -Reduktion auch die DNA zu verdauen.

Somit verbleibt an den Stellen des Proteins, an denen sich Addukte ausgebildet hatten, nur noch das fdC- bzw. fdU-Nukleosid (Abb. 3.7), das dann bei der Analyse als Markierung für die Lokalisierung der kovalenten Interaktionsstellen genutzt werden kann.

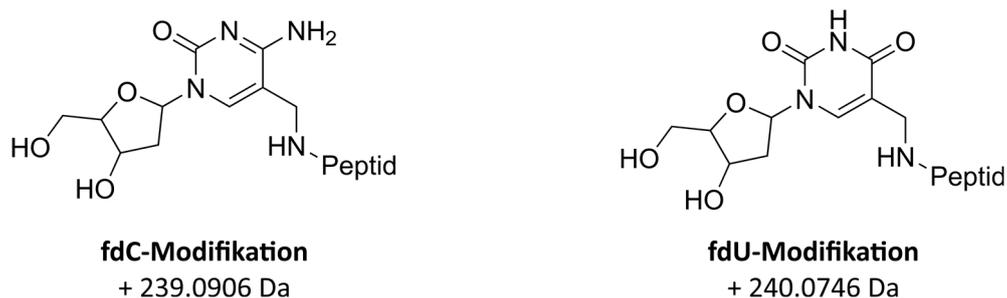


Abbildung 3.7: Struktur der Modifikationen, welche nach dem Verdau der DNA zu Nucleosiden am Peptid verbleiben und deren Addukt-Masse in der massenspektrometrischen Analyse.

Die Nukleosomen mit fC-haltiger DNA wurden hierfür als Erstes, wie bereits beschrieben, rekonstituiert und reduziert. Im Rahmen dieser Arbeit wurden mehrere Versuche unternommen, Addukte zwischen der formyl-haltigen DNA und den Lysinen der Histone mittels Massenspektrometrie nachzuweisen. Dabei wurden verschiedene Parameter zusätzlich zu dem für die Positivkontrolle beschriebenen Protokoll untersucht. Nachfolgend soll daher ein Überblick über die angewendeten Methoden gegeben werden.

Ein erster wichtiger Aspekt des Protokolls war die Wahl einer geeigneten Protease. Histone besitzen einen hohen Anteil der basischen und damit in physiologischer Lösung positiv geladenen Aminosäuren Lysin und Arginin, um eine starke Assoziation an die negativ geladene DNA zu erreichen. Diese Aminosäuren sind jedoch auch die Stellen, an denen Trypsin, die am häufigsten verwendete Protease in der *bottom-up*-Massenspektrometrie, die Polypeptidkette schneidet.

Dadurch entstehen beim Verdau der Histone viele sehr kurze Peptide mit weniger als fünf Aminosäuren, welche in einer komplexen Peptid-Mischung nur sehr schlecht identifiziert werden können, da sie oft zu polar für Umkehrphasen-Chromatographie sind^[531] und in der bioinformatischen Auswertung Schwierigkeiten bereiten.^[532] Daher wurde untersucht, ob durch die Verwendung einer anderen Protease der identifizierbare Anteil der Protein-Sequenzen und vor allem der Lysine gesteigert werden kann. Trypsin wurde dabei mit den Proteasen Chymotrypsin sowie Thermolysin verglichen. Chymotrypsin spaltet nach aromatischen Aminosäuren, während Thermolysin nach hydrophoben Aminosäuren schneidet. In Abbildung 3.8a ist der identifizierte Anteil der Sequenzen durch Verdau mit den verschiedenen Proteasen in Prozent dargestellt.

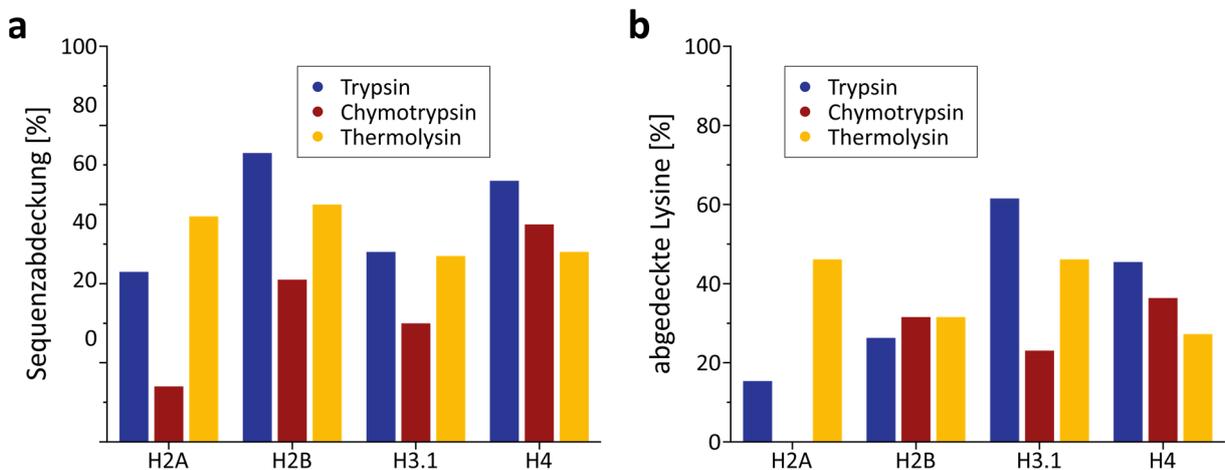


Abbildung 3.8: Vergleich der Eignung verschiedener Proteasen für den Histon-Verdau. a) Identifizierter Anteil der Histon-Sequenzen. b) Identifizierter Anteil der Histon-Lysine.

Wie man sieht, ist die Sequenzabdeckung von Trypsin für die Histone H2B, H3.1 sowie H4 am höchsten. Lediglich für H2A konnte mit Thermolysin eine höhere Abdeckung erreicht werden. Die identifizierten Sequenzen bei Verwendung von Chymotrypsin sind vergleichsweise gering. Ein ähnliches Bild zeigt sich bei Betrachtung der identifizierten Lysine relativ zur gesamten Anzahl im Histon (Abb. 3.8b). Hier ist Thermolysin bei zwei Histonen, H2A und H2B, die beste Wahl, während Trypsin für H3.1 und H4 die besten Werte liefert. Auch mit Chymotrypsin konnten einige Lysine identifiziert werden, allerdings ist wie bei der Sequenzabdeckung mindestens eine andere Protease besser oder vergleichbar gut.

Da Trypsin nur nach unmodifizierten Lysinen schneidet, könnten die mit Lysin verknüpften Nukleoside im eigentlichen Experiment dafür sorgen, dass Peptide entstehen, die zu lang für eine Identifikation sind. Aus diesem Grund und da Thermolysin eine über alle Histone gleichmäßigere Abdeckung zeigt, wurde es für den Nukleosomen-Verdau ausgewählt.

Eine weitere Herausforderung des Protokolls liegt in der komplexen Nukleosom-Struktur, welche den Verdau der Proteine erschwert. Mehrere Studien konnten zeigen, dass bei der Behandlung von Chromatin, aber auch einzelnen Nukleosomen, mit Proteasen nur einige wenige Peptide abgespalten werden.^[526, 533-535] Diese Peptide entstehen größtenteils durch den proteolytischen Abbau der N- und C-Termini der Histone, da diese leicht zugänglich sind, während die meisten anderen Schnittstellen durch die Nukleosom-Struktur abgeschirmt werden.^[536] Da zuerst ein Protein- und dann ein DNA-Verdau durchgeführt wurde, könnte ersterer unvollständig ablaufen, wodurch keine Addukte in den vor Proteolyse geschützten Bereichen identifizierbar sind. Daher wurde vor dem Protein-Verdau ein Vorverdau der DNA mithilfe von Benzonase durchgeführt, welche die DNA zu Di-, Tri- und Tetranukleotiden degradiert.^[537] Dadurch sollten die Histone für den Protein-Verdau zugänglicher sein. Zudem wurde nach Protein- und DNA-Verdau ein zweiter Protein-Verdau durchgeführt, um eine vollständige Degradation der Histone zu Peptiden zu gewährleisten.

Es konnten jedoch in keinem der durchgeführten Experimente zum direkten Nachweis der Addukte mittels Massenspektrometrie Peptide mit der gesuchten Modifikation gefunden werden, was verschiedene Gründe haben kann. So ist es möglich, dass das Protokoll nicht geeignet war, um einen kompletten Verdau der Makromoleküle zu erreichen. Nur wenn die DNA tatsächlich bis zu den einzelnen Nukleosiden degradiert wird, entsteht die Modifikation, welche während der Analyse definiert wurde. Zwar wurden auch testweise einige andere mögliche Reste untersucht, wie z. B. das entsprechende Monophosphat oder das durch einen Angriff des Amins an der C₆-Position von fC bzw. fU entstehende Addukt, allerdings kann es sein, dass durch den Verdau ein anderer, unbekannter Rest entsteht. Auch ein unvollständiger Protein-Verdau könnte dazu geführt haben, dass Peptide entstanden sind, welche nicht mittels Massenspektrometrie analysiert oder durch die eingestellten Parameter bei der Analyse *in silico* nicht identifiziert werden können. Andere Proteasen oder eine Derivatisierung der Lysine, zum Beispiel mit Propionsäureanhydrid,^[538] könnten die Sequenzabdeckung eventuell erhöhen. Eine Optimierung der Parameter der massenspektrometrischen Analyse könnte ebenfalls für eine höhere Anzahl unterschiedlicher Peptide sorgen, was nötig wäre, falls die Intensität der modifizierten Peptid-Ionen nicht ausreicht, um für die Fragmentierung ausgewählt zu werden.

Um zu untersuchen, ob das verwendete Protokoll mit dem kombinierten DNA-Protein-Verdau prinzipiell für die Identifikation von DNA-Protein-Addukten geeignet ist, wurde eine Positivkontrolle verwendet. Diese wurde von *Dr. Ralf Strasser* und *Irene Ponzo, M.Sc. von Dynamic Biosensors* zur Verfügung gestellt und bestand, wie in Abbildung 3.9a dargestellt, aus einem über einen Crosslinker kovalent verknüpften und anschließend aufgereinigten Komplex einer einzelsträngigen DNA und der katalytischen α -Untereinheit der cAMP-abhängigen Proteinkinase (Prkaca) aus *C. griseus*.

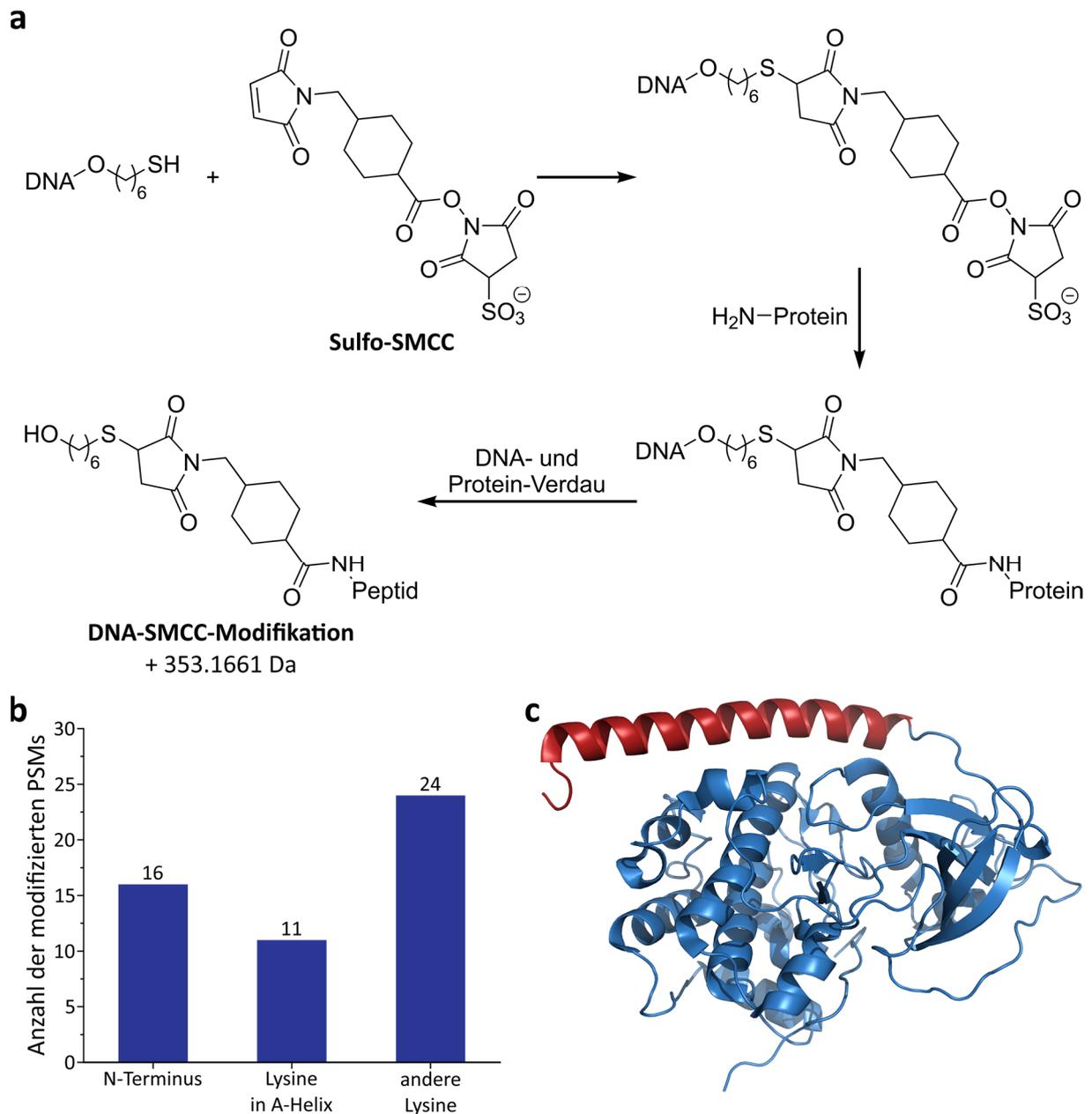


Abbildung 3.9: Überprüfung des kombinierten Protein- sowie DNA-Verdaus anhand einer Positivkontrolle. a) Funktionalisierung eines Thiolhaltigen DNA-Strangs mit einem DNA-Protein-Crosslinker und Reaktion mit einem Amin. Nach DNA- und Protein-Verdau bleibt ein Fragment mit definierter Masse am Peptid zurück. b) Ergebnis der massenspektrometrischen Analyse der verdauten Probe. c) Kristallstruktur von Prkaca aus *C. griseus* (pdb: 4WIH^[539]). Die N-terminale A-Helix ist rot eingefärbt.

Die Modifikation erfolgt durch die Verwendung eines Aktivesters ebenfalls an Aminen der Lysin-Seitenketten sowie am Protein-N-Terminus, weshalb dieses System ein gutes Modell sein sollte, um zu untersuchen, ob der nach dem Verdau am Protein verbleibende Rest der DNA massenspektrometrisch nachgewiesen werden kann. Da hierbei jedoch nur eine einzelsträngige DNA verwendet wurde und das DNA-Protein-Addukt keine komplexe Quartärstruktur wie die der Nukleosomen besitzt, kann dieses Experiment nur einen groben Anhaltspunkt der grundsätzlichen Funktionstüchtigkeit des kombinierten DNA-Protein-Verdaus geben.

Auch die Aminosäuresequenz des Proteins unterscheidet sich, indem sie nicht so viele Lysine und Arginine enthält, wie es bei den Histonen der Fall ist.

Die erhaltenen MS-Daten wurden *in silico* ausgewertet, wobei nach der in Abbildung 3.9a gezeigten Modifikation an Lysinen sowie dem Protein-N-Terminus gesucht wurde, welche sich aus dem Crosslinker und dem Thiol-modifizierten Rest der DNA ergibt. Die Validierung der den einzelnen MS²-Spektren zugeordneten Peptide erfolgte unter Berücksichtigung einer FDR von 1 %. In Abbildung 3.9b ist die Anzahl an Spektren angegeben, denen ein Peptid mit der Modifikation an einem Lysin bzw. dem Protein-N-Terminus zugeordnet werden konnte. Auffallend ist, dass die Modifikation überwiegend in der N-terminalen Region, welche A-Helix genannt wird,^[540] und dort besonders häufig am N-Terminus des Proteins auftritt. Betrachtet man die Kristallstruktur des Proteins (Abb. 3.9c),^[539] sieht man, dass diese A-Helix relativ isoliert vom restlichen, eher globulär aufgebauten Teil des Proteins vorliegt. Es ist also plausibel, dass die Verknüpfung von DNA und Protein zum großen Teil an den Aminen in diesem Teil des Proteins erfolgt, da diese Region sehr gut zugänglich ist.

Auch *in vivo* kann die A-Helix auf verschiedene Art und Weise modifiziert werden oder mit Proteinen interagieren.^[540-543] Dieses Experiment zeigt also, dass die Kombination aus einem DNA- sowie einem Protein-Verdau eines kovalent verknüpften DNA-Protein-Komplexes durchaus zur Identifikation der an den interagierenden Peptiden verbleibenden Modifikation führen kann. Mit hoher Wahrscheinlichkeit liegt das Ausbleiben der Identifikation im Falle der Nukleosomen also an deren komplexer Struktur sowie an der begrenzten Sequenzabdeckung, die für diese ungewöhnlichen Aminosäuresequenzen erreicht werden kann.

Während die beschriebenen MS-Analysen der Nukleosomen durchgeführt wurden, wurde von *Li et al.* sowie *Ji et al.* unabhängig voneinander publiziert, dass Histone tatsächlich Addukte mit fC-haltiger DNA bilden.^[544, 545] Die verwendeten Methoden sind zum Teil ähnlich zu denen in dieser Doktorarbeit, allerdings wurden auch andere Ansätze verfolgt, um einen Nachweis der Addukte zu erbringen. Daher sollen die von den Arbeitsgruppen publizierten Ergebnisse im Folgenden diskutiert und mit den hier dargestellten Versuchen verglichen werden.

So zeigen *Ji et al.* ebenfalls Experimente mit kurzen Peptiden und einem DNA-Einzelstrang, welcher ein fC enthält. Diese wurden ebenfalls mittels denaturierender PAGE ausgewertet. Für einige der Peptide konnten die Autor*innen Addukte nachweisen, allerdings zeigte ein 10-mer Peptid, welches drei Glutaminsäuren enthält, keine Reaktion mit der DNA. Als Erklärung dafür nennen die Autor*innen die negativen Ladungen der Carboxyl-Seitenketten, welche eine Assoziation an die ebenfalls negativ geladene DNA erschweren könnten. Dies könnte eine mögliche Erklärung sein, warum das in Kapitel 3.1.1 verwendete Peptid nicht mit der fC-haltigen DNA reagierte, da es bei einer Länge von sieben Aminosäuren ebenfalls zwei Glutaminsäuren besitzt.

Interessanterweise haben die Autor*innen in diesem Experiment auf eine Schützung des Peptid-N-Terminus, wie es in dieser Arbeit erfolgte, verzichtet. Somit könnte die Addukt-Bildung auch an diesem primären Amin stattfinden. Für ein Peptid zeigen die Autor*innen beispielhaft ein MS²-Spektrum der Peptidfragmente des Addukts mit fC-Nukleosid, allerdings wurde lediglich ein einziges Fragment identifiziert, welches eine Unterscheidung, ob fC am N-Terminus oder am Lysin gebunden ist, ermöglicht. Dieses Fragment hat zusätzlich eine sehr geringe Intensität, sodass eine Aussage über die Lokalisierung der Modifikation am Lysin eher unsicher erscheint.

Auch Addukte zwischen fC-haltiger DNA und Histonen wurden von den Autor*innen gelelektrophoretisch sowie massenspektrometrisch untersucht. Allerdings wurden hierbei nur einzelne Histone sowie ein kurzer DNA-Einzelstrang anstelle von Nukleosomen und DNA zusammen inkubiert. Die Ergebnisse dieser gelelektrophoretischen Experimente unterstützen die in dieser Doktorarbeit durchgeführten Versuche, da in der Publikation ebenfalls eine Bande bei höherem Molekulargewicht sichtbar ist, wenn Protein und DNA inkubiert wurden. Die Intensität dieser Bande sinkt durch Erhitzen, außer wenn vorher mit NaBH₃CN reduziert wurde, so wie es auch in dem in Abbildung 3.4 dargestellten Versuch beobachtet werden konnte.

Die in der Publikation gezeigte massenspektrometrische Analyse ähnelt zum Teil den in dieser Arbeit durchgeführten Experimenten. Dabei wurde einerseits intaktes Histon mit kovalent verknüpfter ssDNA nachgewiesen. Ebenso konnten in einer *bottom up*-Analyse einige Lysine und interessanterweise auch Arginine identifiziert werden, die mit fdC modifiziert waren. Die Modifikation von Arginin steht im Gegensatz zu den Versuchen von *Li et al.*, die keine Reaktion zwischen fdC-Nukleosid und N-terminal geschütztem Arginin beobachten konnten.^[544] Eine ergleichenbare Identifikation von modifizierten Lysinen konnte jedoch mit den Methoden in dieser Doktorarbeit nicht erreicht werden, weshalb nachfolgend mögliche Ursachen für die unterschiedlichen Ergebnisse diskutiert werden sollen.

Zum einen wurden in der besagten Publikation nur einzelne Histone (H2A bzw. H4) und ein kurzer DNA-Einzelstrang anstelle von rekonstituierten Nukleosomen verwendet. Dadurch sinkt die Komplexität der Addukte und es gibt keine schwer zugänglichen Stellen, die den DNA- oder Protein-Verdau behindern könnten.

Ein weiterer Unterschied besteht in der MS-Methode. Hier verwenden die Autor*innen eine *in-source* CID von 20 V, also eine Fragmentierung in der Ionenquelle vor der Aufnahme des MS-Spektrums. Dadurch soll die Ribose des mit den modifizierten Peptiden verknüpften fdC-Nukleosids abgespalten werden, weshalb die Autor*innen bei der Analyse *in silico* die Modifikationsmasse der fC-Base verwenden. Allerdings scheint es, als hätten die Autor*innen die Falscherkennungsrate (FDR) nicht beachtet, sondern alle identifizierten Peptide berücksichtigt, da große Bereiche der Histon-Sequenzen gelb oder rot gefärbt sind (Daten aus der *Supporting Information*).

Dies bedeutet, dass in diesen Abschnitten der Sequenz nur schlecht zugeordnete PSMs gefunden wurden, die nicht mit dem in der Analyse verwendeten *FDR-cut off* vereinbar sind. Da jedoch generell nur 10–50 % der PSMs, welche in MS²-Experimenten identifiziert werden, tatsächlich korrekt sind,^[546] ist es unabdingbar, für die Interpretation der Ergebnisse lediglich die Peptide heranzuziehen, welche eine niedrige FDR, meist 1 %, ergeben.^[481, 547, 548] Da für die aufgelisteten modifizierten Peptide nicht angegeben wurde, ob sie dieses Kriterium erfüllen und auch keine anderen Qualitäts-Parameter wie *q value*, *posterior error probability* oder *XCorr* der zugehörigen PSMs spezifiziert wurden, ist also nicht eindeutig, ob diese Peptide eventuell falsch-positive Zuordnungen sind. Aus diesen Gründen wurde bei den in dieser Doktorarbeit generierten Daten ein *FDR-cut off* von 1 % angewendet, was ebenfalls erklären könnte, weshalb im Gegensatz zu *Ji et al.* keine modifizierten Peptide identifiziert werden konnten.

In einem weiteren Experiment weisen die Autor*innen das Vorkommen der Schiff'schen Basen in Zellen nach. Hierfür wurde DNA aus HEK-Zellen extrahiert und reduziert und DNA und Proteine bis hin zu einzelnen Nucleosiden bzw. Aminosäuren verdaut. Das übrigbleibende fC-Nucleosid-Lysin-Addukt wurde zur Signalsteigerung derivatisiert und quantifiziert. Somit konnte berechnet werden, dass ca. 1.2 % aller Formylcytosine Addukte gebildet hatten. Bei dieser Analyse geht jedoch die Information verloren, welche Proteine kovalent an die DNA gebunden waren und es lässt sich somit keine Aussage treffen, ob die Addukte von Nucleosomen stammten.

Auch *Li et al.* zeigen Experimente, die Ähnlichkeit mit denen aus dieser Doktorarbeit besitzen.^[544] So werden in dieser Publikation ebenfalls Nucleosomen mit einer fC-haltigen 601 DNA rekonstituiert. Die DNA enthielt hierbei jedoch nur ein einzelnes fC und es wurden vier verschiedene Positionen dieser Base getestet. Mittels denaturierender Gelelektrophorese wurde die Bildung der Addukte untersucht. Dabei nutzten die Autor*innen eine photo-reaktive Schutzgruppe für die Formylgruppe und konnten zeigen, dass erst nach Entschützung durch Lichteinwirkung Addukte ausgebildet werden, was eine Rolle der Formylgruppe bei der Addukt-Bildung bestätigt. Eine Kinetik zeigt, dass nach 70 h etwa 20 % der Nucleosomen Addukte gebildet haben und der Gleichgewichtszustand nach dieser Zeit bereits fast erreicht ist. Durch Zugabe von Hydroxylamin zu nicht-reduzierten Nucleosomen sinkt der Anteil an Addukten wieder, wie es auch in Abbildung 3.6 dieser Arbeit zu sehen ist. Mittels In-Gel-Verdau und massenspektrometrischer Analyse konnten die Autor*innen außerdem das Histon H4 in der Addukt-Gelbande nachweisen. Dieses Histon konnte neben H2B auch bei dem in dieser Arbeit durchgeführten In-Gel-Verdau der Addukt-Banden mit hoher Wahrscheinlichkeit identifiziert werden (s. Abb. 3.5). Die von *Li et al.* durchgeführten Experimente bestätigen somit allesamt die Erkenntnisse, welche durch die gelelektrophoretischen Versuche in dieser Doktorarbeit gewonnen werden konnten.

In späteren Publikationen wurde die Rolle der kovalenten DNA-Histon-Addukte weiter untersucht. So zeigten *Raiber et al.* einen Einfluss auf die Genexpression *in vivo* durch Organisation von Nukleosomen an *Enhancern* und konnten eine Adduktbildung am Histon H3 nachweisen,^[549] während die Gruppe um *Prof. Natalia Tretyakova* mit Modellsystemen von Effekten auf Geschwindigkeit und Genauigkeit der Replikation und Transkription berichtet.^[550-553] Ein direkter Nachweis der kovalenten Interaktion zwischen DNA und Histonen im Kontext von Nukleosomen mittels Protein-MS ist in der Literatur jedoch nach wie vor nicht zu finden.

3.1.3 Untersuchung einer möglichen Deformylierung von fC in Nukleosomen

Wie bereits erwähnt, wurde die Menge an fC-Lysin-Addukten in HEK-Zellen von *Ji et al.* als ca. 1.2 % aller fC-Basen quantifiziert.^[545] Geht man davon aus, dass der Anteil von fC in Zellen 0.002–0.02 % aller Cytosine entspricht,^[138, 139] ergibt sich daraus bei einer Länge von 6.32×10^9 bp und einem GC-Gehalt von 40.9 %^[253] eine absolute Anzahl von ca. 260–2600 Addukten im Genom einer diploiden humanen Zelle. Da die Autor*innen die Addukte für die Quantifizierung allerdings für 3 h mit NaBH_3CN reduzieren mussten, wodurch das Gleichgewicht auf die Seite des Imin-Addukts verschoben wurde, ist davon auszugehen, dass die reale Anzahl noch weit darunter liegt. Es stellte sich daher die Frage, ob die Schiff'schen Basen lediglich ein kurzlebiges Intermediat einer anderen Reaktion darstellen.

Bereits vor einiger Zeit wurde spekuliert, dass fC neben der durch TDG initiierten BER^[140] entweder direkt oder nach Oxidation zu caC durch Abspaltung der Formyl- bzw. Carboxylgruppe in Cytosin umgewandelt werden kann (Abb. 3.10a).^[159, 554] Vor kurzem konnten *Iwan et al.* zeigen, dass diese Reaktion in Zellen tatsächlich abläuft.^[157] Die dabei involvierten Enzyme sind allerdings noch nicht bekannt, weshalb die Überlegung aufkam, ob die Adduktbildung von fC mit den Histonen eine Rolle in dieser aktiven Demethylierung spielen könnte.

Der Reaktionsmechanismus (Abb. 3.10b) könnte dabei ähnlich zu dem von *Schiesser et al.* vorgeschlagenen Mechanismus für die ausschließlich Thiol-katalysierte Deformylierung ablaufen.^[554] Zunächst bildet sich das Imin zwischen Formylgruppe und Histon-Amin, welches den Angriff eines Hilfs-Nukleophils (z. B. eines Thiols) an der C6-Position der Base erleichtert. Im nächsten Schritt folgt die Anlagerung eines Protons an der C5-Position. Nach Anlagerung eines Wassermoleküls an das C-Atom der ehemaligen Formylgruppe, kann das Formyl-Amin sowie das Hilfs-Nukleophil abgespalten werden, wobei die Aromatizität wiederhergestellt wird. Als Resultat wurde die Formylgruppe auf das Amin des Proteins übertragen. Das für die Reaktion notwendige Thiol könnte zum einen direkt von einem Histon stammen, oder aber von einem weiteren Protein, welches für die Deformylierung rekrutiert wird.

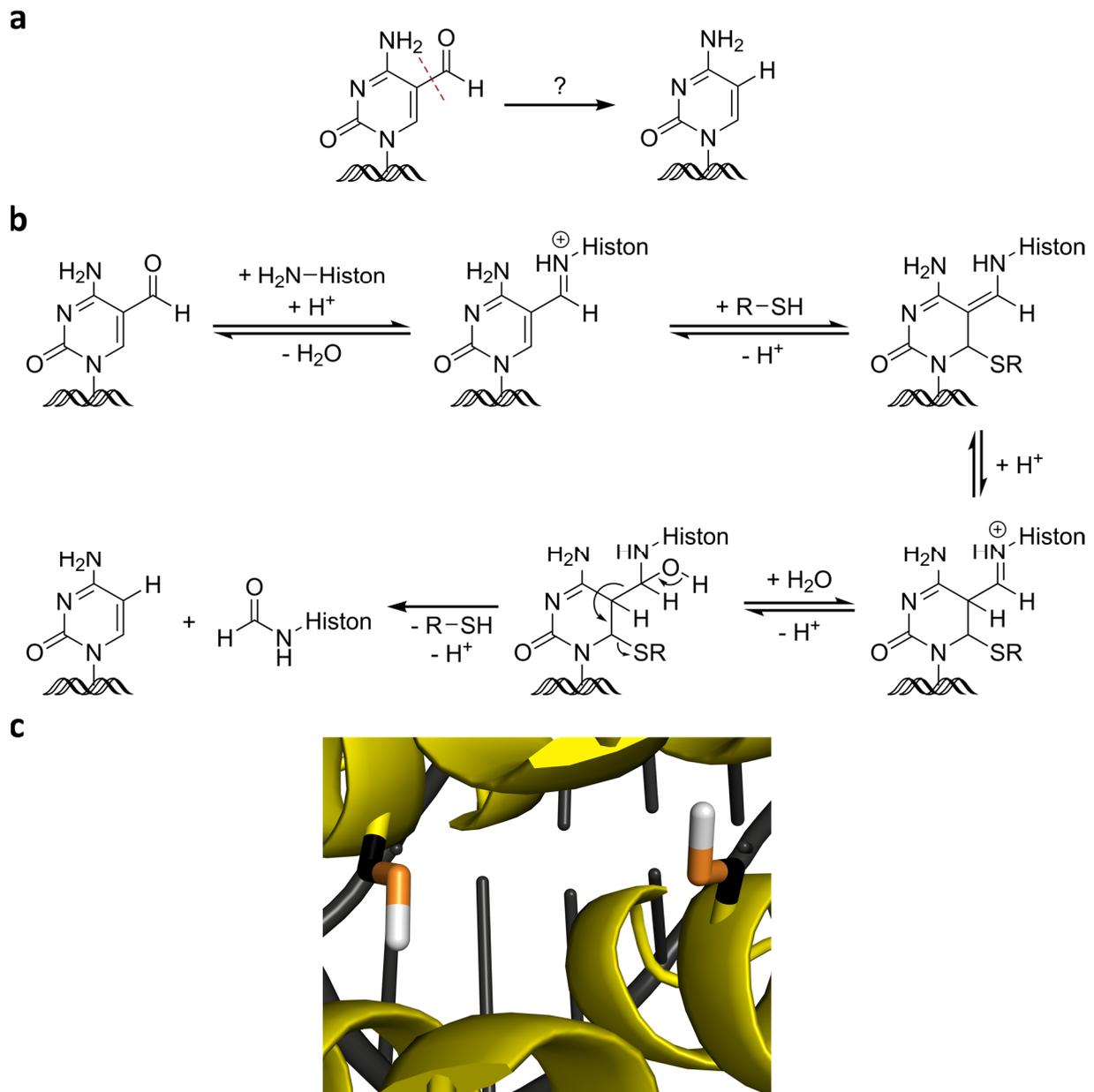


Abbildung 3.10: Möglicher Mechanismus einer Histon-katalysierten Deformylierung. a) Der C-C-Bindungsbruch wird durch bislang unbekannte Enzyme katalysiert. b) Möglicher Mechanismus der Übertragung der Formylgruppe von fC auf ein Histon-Lysin. c) Räumliche Anordnung der Cysteine beider H3-Monomere im Nucleosom (pdb: 1KX5^[262], H3-Sequenz dargestellt in gelb, S-Atome in orange und hinzugefügte H-Atome in weiß).

Histon H3 besitzt ein Cystein und ist damit das einzige Cystein-haltige Histon des Oktamers. Die Cysteine der beiden H3-Moleküle im Nucleosom (s. Abb. 3.10c) bilden *in vivo* größtenteils eine Disulfidbrücke,^[555] allerdings konnte ebenfalls *in vivo* gezeigt werden, dass sie teilweise auch reduziert vorliegen und z. B. mit Glutathion reagieren können.^[556] Das Gleichgewicht zwischen freien und oxidierten Cysteinen könnte eine Möglichkeit für die Zelle sein, die Deformylierung abhängig vom Redox-Potential in der Zelle zu kontrollieren. Wie bereits weiter oben erwähnt, konnten Raiber *et al.* zudem bereits zeigen, dass H3 *in vivo* Addukte mit fC-haltiger DNA ausbildet.^[549]

Um diese Fragestellung zu untersuchen, wurden erneut Nukleosomen mit fC-haltiger 601 DNA rekonstituiert. Hierzu wurden in Zusammenarbeit mit *Dr. Angie Kirchner* mittels PCR, wie schon in Kapitel 3.1.2 beschrieben, Konstrukte der 601 Sequenz hergestellt. Für die fC-haltige DNA wurde jedoch $^{15}\text{N}_2$ -isotopenmarkiertes fdCTP verwendet (Abb. 3.11a), welches von *Dr. Alexander Schön* zur Verfügung gestellt wurde, da die DNA nach dem Verdau zu Nucleosiden quantifiziert werden sollte. Normales fdC würde nach der Deformylierung nicht von dem in den Primern enthaltenen dC unterscheidbar sein, was die Quantifizierung erschweren könnte. Demgegenüber würde $^{15}\text{N}_2$ -fdC zu $^{15}\text{N}_2$ -dC umgewandelt werden, was sich getrennt von unmarkiertem dC analysieren ließe. Als Kontrolle wurde ein Strang mit *superheavy*-dCTP (Abb. 3.11a) statt normalem dCTP synthetisiert, da erste Tests zeigten, dass das natürliche Isotopenmuster von dC in 601_C einen sehr großen Hintergrund verursacht. Bei shdC sind alle Kohlenstoffatome komplett ^{13}C - und die drei Stickstoffatome der Base ^{15}N -markiert. Somit besitzt der Kontrollstrang den gleichen, durch die Primer-dCs verursachten Hintergrund wie der formyl-haltige Strang.

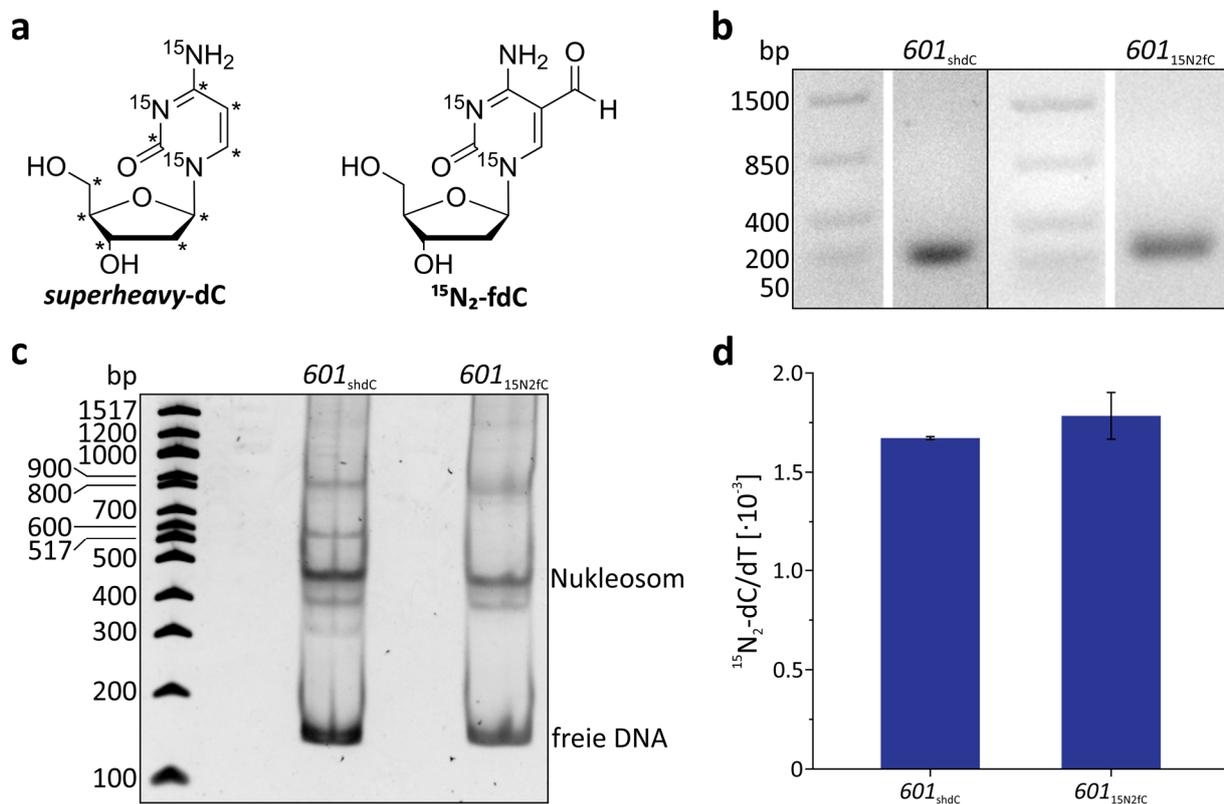


Abbildung 3.11: Analyse einer möglichen Deformylierung fC-haltiger DNA in Nucleosomen *in vitro*.

a) Übersicht der als Triphosphate eingesetzten und später quantifizierten isotopenmarkierten Desoxynucleoside. Im Falle von shdC sind die ^{13}C -Isotopenmarkierungen als * dargestellt. b) Agarose-Gel der PCR-Synthese von 601_{shdC} und 601_{15N2fC}. c) Native PAGE der Nucleosomrekonstitution. d) Ergebnis der von *Dr. Katharina Iwan* durchgeführten DNA-Quantifikation. Dargestellt ist die gefundene Menge des Deformylierungsprodukts $^{15}\text{N}_2$ -dC normiert auf die Gesamtmenge an DNA (ausgedrückt durch die Menge an dT).

Abbildung 3.11b zeigt, dass durch PCR mit beiden isotonenmarkierten Triphosphaten ein einzelnes Produkt im Bereich der Länge der 601 Sequenz synthetisiert werden konnte. Wie bereits in Kapitel 3.1.2 beschrieben, wurden diese DNA-Doppelstränge verwendet, um Nukleosomen zu rekonstituieren (Abb. 3.11c), welche im Anschluss bei 37 °C inkubiert wurden, um eine mögliche Deformylierung unter physiologischen Bedingungen nachzustellen. Daraufhin wurde die DNA von *Dr. Katharina Iwan* verdaut und mittels Massenspektrometrie quantitativ in Mess-Triplikaten untersucht (Abb. 3.11d). Dargestellt ist die Menge an gefundenem Deformylierungsprodukt $^{15}\text{N}_2\text{-dC}$ pro dT aus jeweils zwei Replikaten, da die Intensität für dT, die als Normierung für die eingesetzte DNA-Menge diente, bei jeweils einem Replikat außerhalb der Eichgerade lag. Man sieht, dass auch in der Probe mit 601_{shdC} $^{15}\text{N}_2\text{-dC}$ gefunden wurde. Dies ist mit dem bereits angesprochenen Hintergrund während der massenspektrometrischen Analyse zu erklären, der durch das Isotopenmuster der Primer-dCs zustande kommt. In der Probe mit 601_{15N2fC} ist ebenfalls Deformylierungsprodukt vorhanden, allerdings unter Berücksichtigung der Standardabweichung nicht mehr als in der Kontrollprobe, wodurch es wahrscheinlich ist, dass dieser Wert auch dem Hintergrund der Messung geschuldet ist.

Da bei diesem Experiment nicht festgestellt werden konnte, ob die fC-DNA in den Nukleosomen deformyliert wurde, wurde untersucht, ob die Deformylierungsreaktion eventuell durch eine *N*-Methylierung der Lysine beschleunigt werden kann. Diese sollte die Nukleophilie des Amins erhöhen und könnte daher zu einer Steigerung der Reaktivität führen. Eine Abschätzung der Unterschiede zwischen Lysin und *N*-Methyllysin hinsichtlich ihrer Nukleophilie kann durch den Vergleich von MeNH_2 und Me_2NH erfolgen. *Brotzel et al.* zeigten anhand der Reaktion mit Benzhydrylium-Ionen, dass obwohl sich die Basizität von MeNH_2 ($\text{pK}_a = 10.66$) und Me_2NH ($\text{pK}_a = 10.73$) kaum unterscheidet, die Nukleophilie von Me_2NH um mehrere Größenordnungen höher ist als die von MeNH_2 .^[557] Es könnte also sein, dass der Angriff des Amins an den Carbonyl-Kohlenstoff der Formylgruppe durch die zusätzliche Methylgruppe erleichtert wird, wodurch der beobachtete Effekt in der quantitativen Analyse eventuell eindeutiger ausfallen könnte.

Um dies zu untersuchen, wurden die Histone einem bereits etablierten Protokoll zur Methylierung von Proteinen^[558] unterzogen. Der zugrunde liegende Mechanismus entspricht einer reduktiven Alkylierung, bei der zunächst aus dem unmodifizierten Lysin-Amin und Formaldehyd ein Imin gebildet wird, welches anschließend durch einen Hydrid-Donor zu einem sekundären Amin reduziert wird (Abb. 3.12a). Eine massenspektrometrische Analyse der so behandelten Histone zeigte, dass dieses Protokoll zwar für eine beinahe vollständige Methylierung sorgte, jedoch fast ausschließlich dimethylierte Lysine erzeugte (Abb. 3.12b). Das ist allerdings ein Nachteil, da Dimethyl-Lysin als tertiäres Amin keine Schiff'schen Basen bilden kann. Daher wurde das Protokoll angepasst, indem eine geringere Menge der Reagenzien verwendet und auf eine zweistufige Reaktion verzichtet wurde.

Das Verhältnis von mono- und dimethylierten Peptiden ist in diesen Proben sehr viel besser als mit dem originalen Methylierungsprotokoll; eine vollständige oder zumindest überwiegende Monomethylierung konnte allerdings nicht erreicht werden (Abb. 3.12b). Außerdem ist auch der Anteil an nicht-methylierten Peptiden in dieser Probe viel höher. Für einen höheren Anteil an monomethylierten Peptiden wäre eine weitere Optimierung nötig. Da die Nukleophilie von Methyllysin allerdings, wie oben ausgeführt, höher ist als die von Lysin und die Reaktion zu Dimethyllysin daher schneller abläuft, sollte eine selektive und gleichzeitig vollständige Monomethylierung, die zudem die Histone nicht degradieren darf, kaum zu erreichen sein. Mit den Histonen, welche mit dem modifizierten Protokoll methyliert wurden, wurden wie zuvor Nucleosomen mit 601_{shdC} und $601_{15\text{N}2\text{fc}}$ rekonstituiert (Abb. 3.12c) und die Nucleoside nach einem DNA-Verdau erneut von *Dr. Katharina Iwan* quantitativ untersucht.

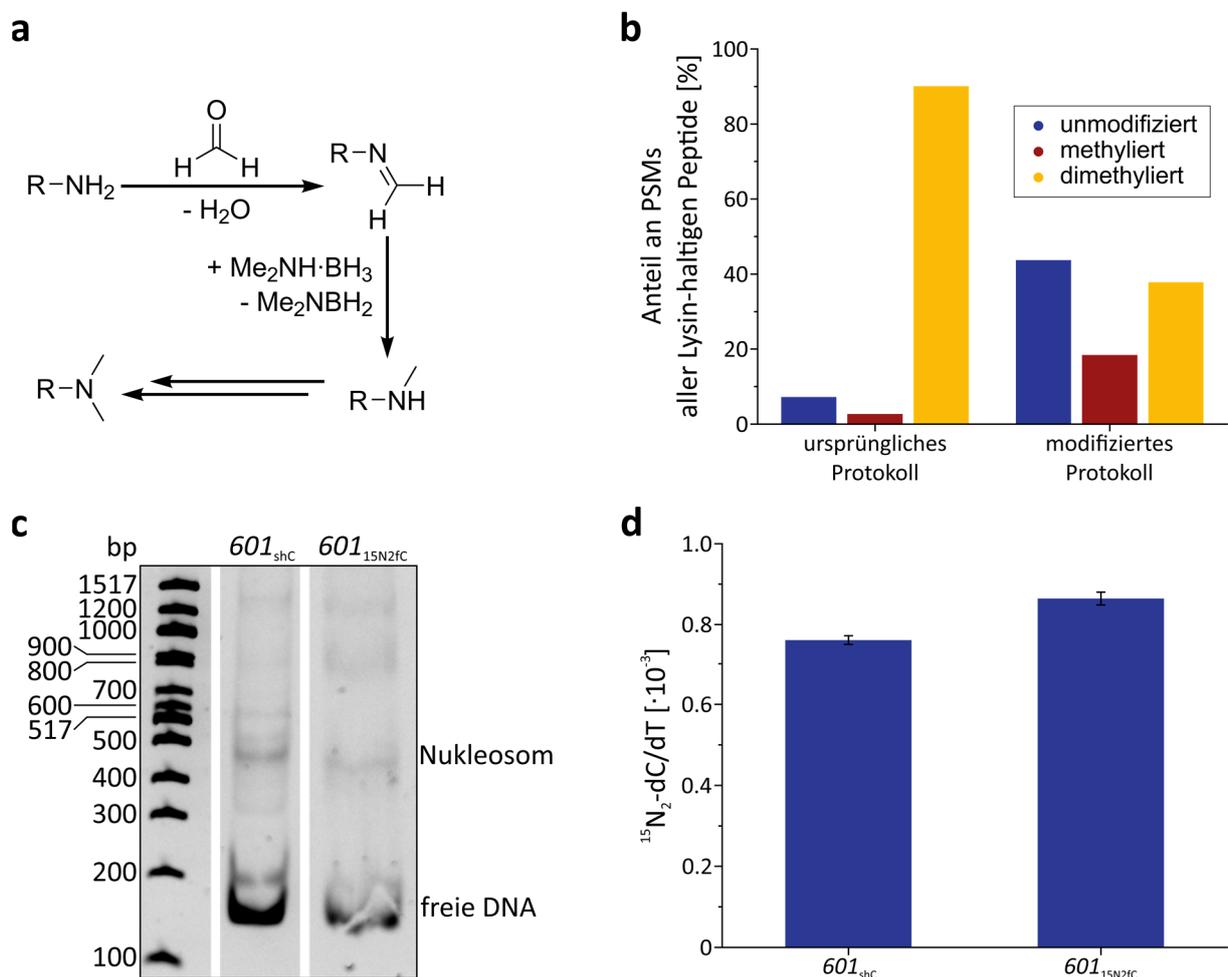


Abbildung 3.12: Analyse der Deformylierung mit vorheriger Methylierung der Histone zur Steigerung der Reaktivität. a) Mono- und Dimethylierung eines primären Amins mit Formaldehyd und Dimethylaminoboran. b) Produktverteilung der Histon-Methylierung bei Verwendung zweier unterschiedlicher Protokolle. c) Nucleosom-Rekonstitution der mit dem modifizierten Protokoll methylierten Histone und 601_{shdC} bzw. $601_{15\text{N}2\text{fc}}$. d) Ergebnis der von *Dr. Katharina Iwan* durchgeführten DNA-Quantifikation. Dargestellt ist die Menge des Deformylierungsprodukts $^{15}\text{N}_2\text{dC}$ normiert auf die Gesamtmenge an DNA (ausgedrückt durch die Menge an dT).

Abbildung 3.12d zeigt die Menge an detektiertem Deformylierungsprodukt $^{15}\text{N}_2\text{-dC}$ pro dT aus drei Mess-Replikaten. Wie zuvor ist eine geringe Differenz zwischen Nukleosomen mit 601_{shdC} und $601_{15\text{N}2\text{fC}}$ erkennbar, allerdings sind die Fehlerbalken in diesem Fall kleiner. Weitere biologische Replikate müssen angefertigt werden, um eine statistische Analyse der Signifikanz dieser Beobachtung zu ermöglichen. Da trotz der Verwendung eines größtenteils shdC-modifizierten Kontrollstranges ein großer Hintergrund zu sehen ist, könnte in weiteren Versuchen ein $2'\text{-F-fdC-Triphosphat}$ verwendet werden, was die Überlappung der Masse des Deformylierungsprodukts mit dem Isotopenmuster der Primer-dCs verhindern würde. Auch die Synthese eines caC-haltigen Stranges ist möglich, da die Decarboxylierung möglicherweise eine schneller ablaufende Reaktion darstellt. Die Analyse der modifizierten DNA nach Inkubation ohne Histone wäre zudem als weitere Kontrolle wichtig, um eine spontane Demodifikation ausschließen zu können.

Als letztes wurde überprüft, ob sich die Formylgruppe, welche bei einer Deformylierung auf ein Histon-Lysin übertragen werden würde, mittels Proteomik nachweisen lässt. Dafür wurden die Nukleosomen sowohl mit als auch ohne Methylierung einem Protein-Verdau unterzogen und mittels Massenspektrometrie untersucht. Die Analyse erfolgte unter Einbeziehung von Formyllysin sowie im Falle der methylierten Histone zusätzlich von Methyl- und Dimethyllysin sowie Formylmethyllysin als mögliche Modifikationen. Die Auswertung ergab jedoch keine eindeutig identifizierten formyl-modifizierten Peptide, was ein Indiz dafür ist, dass die Reaktion mit diesem System nur in geringem Umfang oder über einen anderen Mechanismus erfolgt.

Es kann zudem nicht ausgeschlossen werden, dass die Deformylierung andere Histon-Modifikationen benötigt oder neben den Histonen die Beteiligung weiterer Enzyme erfordert. Besonderer Fokus liegt hierbei auf dem Hilfs-Nukleophil, welches an der C6-Position angreift. Erst kürzlich konnte nachgewiesen werden, dass dieser Schritt essenziell für die Deformylierung *in vivo* ist^[158] und es wurde vorgeschlagen, dass dieses Nukleophil ein Thiol in Form eines Cysteins sein könnte.^[554] Weiter oben wurden bereits die speziellen Eigenschaften von H3 erwähnt, welches als einziges Histon des Oktamers Cysteine besitzt. Diese müssen jedoch in reduzierter Form vorliegen, um nukleophil angreifen zu können. Es ist möglich, dass die Cysteine des rekombinanten Proteins, welches für die Rekonstitutionen verwendet wurde, unter den *in vitro* vorliegenden Bedingungen oxidiert als Disulfide vorlagen und somit nicht für die Reaktion zur Verfügung standen. Ebenso könnte es sein, dass *in vivo* ein weiteres Protein das Hilfs-Nukleophil bereitstellt und die Addukte von DNA und Histon nur der Unterstützung der von diesem Protein katalysierten Deformylierung dienen. Neuere Experimente, welche in dieser Arbeitsgruppe durchgeführt wurden, deuten z. B. auf eine Beteiligung von Dnmt3b und Sirtuinen hin.^[559] Schließlich besteht auch die Möglichkeit, dass das Formylcytosin *in vivo* erst zu Carboxylcytosin oxidiert wird, um eine effizientere Demodifikation zu erreichen.

3.2 Untersuchung einer direkten Interaktion von TET3 und GDH

Die Oxidation von mC zu hmC, fC sowie caC wird durch die TET-Enzyme katalysiert, welche als Cosubstrat α -Ketoglutarat, ein Intermediat des Citrat-Zyklus, benötigen.^[136-139] Es existieren drei verschiedene TET-Isoenzyme, von denen TET3 besonders in Neuronen exprimiert wird.^[164, 192] Studien, die in dieser Arbeitsgruppe durchgeführt wurden, lieferten Hinweise darauf, dass das benötigte α -Ketoglutarat im Nukleus durch eine räumliche Nähe von TET3 und GDH bereitgestellt wird.^[560] Um zu überprüfen, ob es sich dabei um eine direkte Interaktion handelt, wurden in Zusammenarbeit mit *Dr. Michael Stadlmeier* XL-MS-Experimente durchgeführt. Dafür wurden in mehreren Versuchen nicht-kovalente Protein-Interaktionen durch Crosslinking mit DSSO chemisch stabilisiert und nach einem tryptischen Verdau massenspektrometrisch analysiert.

3.2.1 Etablierung eines Crosslinking-Protokolls mit DSSO

In einem ersten Schritt sollte das Crosslinking unter Verwendung von DSSO an einem Modellsystem getestet werden, um es in der Arbeitsgruppe zu etablieren. Hierfür wurde BSA verwendet, welches ein weitverbreitetes Modell für XL-MS-Experimente darstellt.^[561] Dabei entstehen zwar vorrangig Intraprotein-Crosslinks, dennoch ist die Anzahl der gefundenen Crosslinks ein gutes Maß für die Sensitivität der verwendeten Methode.

Eine Lösung von BSA wurde mit DSSO behandelt und nach Reduktion und Alkylierung der Cystein-Reste zu Peptiden verdaut. Die Analyse erfolgte mit einer MS²-Methode. Hierbei macht man sich die asymmetrische Spaltbarkeit von DSSO zunutze, da bei der Fragmentierung mit geeigneter HCD-Energie neben den Peptid-Fragment-Ionen auch Reporter-Ionen gebildet werden, bei denen nur der Crosslinker gespalten wird. Durch die asymmetrische Spaltung entstehen zwei Fragment-Paare, deren Massen sich jeweils um ungefähr 32 Da unterscheiden. Die verwendeten Analyse-Softwares *XlinkX*^[497, 498] (implementiert in der *Proteome Discoverer* Software) oder *MeroX*^[496] können diese Signale dazu nutzen, die Zuordnung eines Crosslinks zu einem Spektrum besser zu verifizieren, was essentiell bei der Verwendung größerer Protein-Datenbanken ist. Da die HCD-Energie einen großen Einfluss auf die Bildung von Reporter- und Fragment-Ionen und damit auf die Identifikation von Crosslinks hat, wurden zunächst zwei verschiedene Ansätze verglichen. Das Ziel hierbei war es, eine Methode zu entwickeln, bei der sich zum einen die Reporter-Fragmente ausbilden und zum anderen genug Peptid-Fragmente entstehen, um den Crosslink als solchen identifizieren zu können aber gleichzeitig auch die Identifikation der Peptide zu ermöglichen. Zum einen wurde eine normalisierte HCD-Energie von 30 % angewendet und zum anderen eine gestufte Fragmentierung, in welcher drei Fragmentierungen mit Energien von 25, 30 sowie 35 % durchgeführt und die entstehenden Fragmente anschließend in einem gemeinsamen MS²-Scan analysiert wurden. Die Daten wurden mit beiden genannten Programmen analysiert, um ihre Eignung für den MS²-Ansatz zu vergleichen.

Da es für *MeroX* häufig Software-Aktualisierungen gibt, die die Datenanalyse weiter optimieren, wurden zwei verschiedene Versionen, nämlich 2.0.0.6 und 2.0.1.4 verwendet. Die Auswertung erfolgte für *XlinkX* im MS²- und für *MeroX* im RISEUP-Modus und wurde unter Anwendung eines FDR-cutoffs von 1 % durchgeführt. Zusätzlich wurden bei *XlinkX* nur CSMs berücksichtigt, welche zusätzlich einen Score von mindestens 20 besaßen, während für *MeroX* ein Score von mindestens 50 vorausgesetzt wurde. Diese Werte werden von den Programmen unterschiedlich berechnet, weshalb sie nicht miteinander vergleichbar sind. Als Datenbank diente dabei eine Kombination aus der Sequenz von BSA und den *dust/contact proteins* des *common Repository of Adventitious Proteins (cRAP)* von *The Global Proteome Machine*.^[562] Diese Kombination aus Zielprotein und kontaminierenden Proteinen wurde gewählt, um die FDR-Kontrolle durch eine Erhöhung der Komplexität zu verbessern. In Abbildung 3.13 ist jeweils die Anzahl der Spektren, denen ein Crosslink zugeordnet werden konnte (Crosslink-Spektrums-Zuordnungen, CSMs) sowie die Anzahl einzigartigen Crosslinks dargestellt.

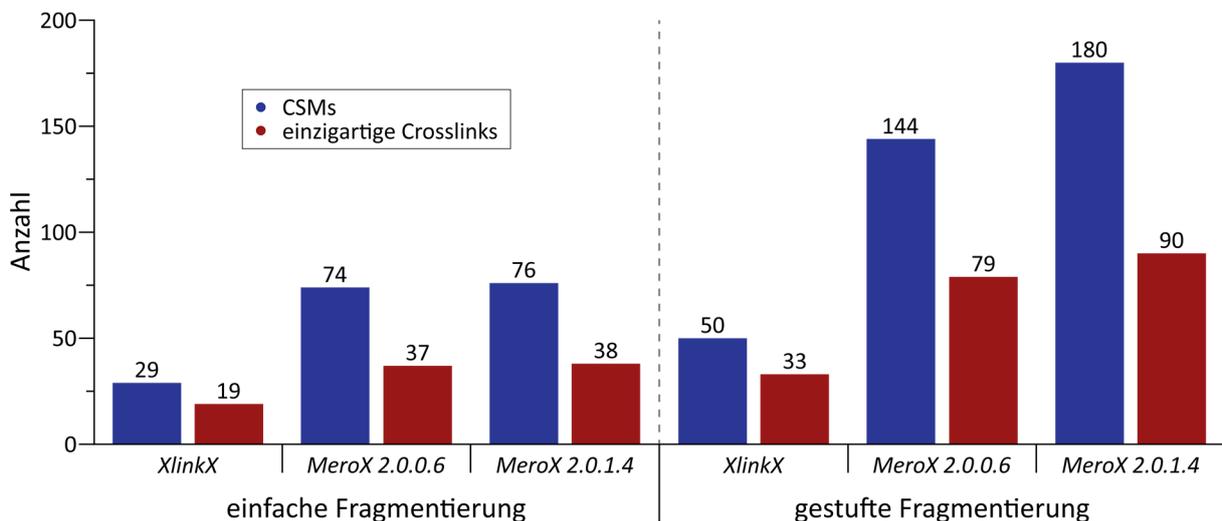


Abbildung 3.13: Vergleich des Crosslinkings von BSA mit DSSO unter Verwendung verschiedener Auswerteprogramme bei zwei unterschiedlichen Fragmentierungen. Abgebildet ist die Anzahl der CSMs sowie der einzigartigen Crosslinks, welche die Programme *XlinkX* sowie zwei Versionen von *MeroX* identifizierten. Die Versuche wurden mit einfacher (normalisierte HCD-Energie: 30 %) sowie gestufter (normalisierte HCD-Energien: 25, 30, 35) Fragmentierung durchgeführt.

Wie man sieht, bewirkt die gestufte normalisierte HCD-Energie bei allen Programmen eine drastische Erhöhung sowohl der CSMs als auch der einzigartigen Crosslinks. Dies liegt daran, dass bei einer Energie von 25 % hauptsächlich der Crosslinker gespalten wird, während bei den höheren Energien vermehrt Peptid-Fragment-Ionen entstehen. Dadurch gibt es zum einen genug Ionen für die Peptid-Identifikation, zum anderen bilden sich auch intensive Reporter-Signale, welche für die Verifikation der Crosslinks von Vorteil sind. Die Entwickler von *MeroX* empfehlen daher sogar die Verwendung einer gestuften HCD-Energie.^[563]

Beim Vergleich der Analyse-Programme zeigen sich ebenfalls große Unterschiede. So ist zu erkennen, dass mit *MeroX* generell mehr Crosslinks identifiziert werden konnten als mit *XlinkX*. Dies liegt wahrscheinlich an den unterschiedlichen zugrunde liegenden Algorithmen. Während *MeroX* alle theoretisch möglichen Peptid-Kombinationen mit der Precursor-Masse des Spektrums vergleicht und danach die beste Zuordnung auswählt,^[496] ist bei *XlinkX* in der aktuellsten Version 2.0 das Vorhandensein aller Reporter-Ionen oder zumindest des Precursors und eines sehr intensiven Reporter-Ions nötig, damit eine Zuordnung überhaupt durchgeführt wird.^[498] Besonders interessant ist jedoch auch der Unterschied zwischen den beiden getesteten Versionen von *MeroX*: in der neueren Version steigt der Wert für CSMs um 25 % im Vergleich zur älteren Version und auch bei den einzigartigen Crosslinks lässt sich ein Anstieg von immerhin 14 % beobachten. Der Grund dafür kann nicht eindeutig bestimmt werden, da vom Entwickler zwischen den Versionen viele verschiedene Änderungen vorgenommen wurden. Diese Ergebnisse zeigen jedoch anschaulich den großen Einfluss der verwendeten Analyse-Software auf die Ergebnisse, was ein Zeichen dafür ist, dass sich das Feld der XL-MS noch in der Entwicklung befindet. Für die weiteren Experimente in Kapitel 3.2 und 3.3 wurde *MeroX 2.0.1.4* verwendet, da eine kürzlich erschienene Vergleichs-Studie zeigte, dass die reale FDR bei *MeroX* im Vergleich zu *XlinkX* eher erreicht wird.^[564] Lediglich die Daten der in Kapitel 3.3.3 vorgestellten Publikation wurden mit *MeroX 2.0.0.6* ausgewertet, da dies zum Zeitpunkt der Veröffentlichung die aktuellste Version war.

Meist besitzen Crosslinks eine geringe Abundanz, weshalb mit MS-Methoden, welche die intensivsten Signale eines MS-Spektrums für die Fragmentierung auswählen, oft die intensiveren unmodifizierten Peptide fragmentiert werden. Zur Anreicherung der wenig abundanten Crosslinks wurden 40 µg von mit DSSO quervernetztem und mit Trypsin verdautem BSA mittels Größenausschluss-Chromatographie (SEC) fraktioniert, wodurch die Crosslinks von regulären tryptischen Peptiden abgetrennt werden können. Da erstere aus zwei kovalent verknüpften Peptiden bestehen, sind sie im Mittel größer als unmodifizierte Peptide und eluieren daher früher von einer porösen Gelmatrix. Dies führt bei Auftrennung der Elution in mehrere Fraktionen zu einer Verringerung der Proben-Komplexität. Um den Einfluss des Crosslinkings auf die SEC nachverfolgen zu können, wurde auch eine nicht quervernetzte BSA-Probe fraktioniert. Abbildung 3.14a zeigt die Chromatogramme der beiden Fraktionierungen, wobei die Elution der Peptide durch deren Absorption bei 215 nm nachverfolgt wurde. Man sieht, dass die Intensität der Absorption in den Fraktionen 12 bis 15 durch das Crosslinking erhöht wird. Dies stellt die durch DSSO kovalent verknüpften Peptid-Paare dar, welche nach dem Protein-Verdau des quervernetzten Proteins entstehen und vor den unmodifizierten Peptiden eluieren. Der Rest der beiden Chromatogramme unterscheidet sich kaum, da hier die unmodifizierten linearen Peptide eluieren, welche für beide Proben identisch sind.

Für die massenspektrometrische Analyse wurden die Elutionsfraktionen 12 bis 18 der Crosslink-Probe getrennt voneinander analysiert (Abb. 3.14b). Es ist zu sehen, dass sich der Großteil der identifizierten CSMs und einzigartigen Crosslinks auf nur wenige Fraktionen verteilt. Die Anzahl der identifizierten Spezies liegt dabei um ein Vielfaches über den in Abbildung 3.13 ohne Fraktionierung gezeigten Werten. Durch die Fraktionierung lässt sich also ein enormer Zugewinn an Information erreichen, solange man eine längere Gesamtmesszeit in Kauf nehmen kann und über genügend Ausgangsmaterial verfügt. Letzteres ist besonders in der Strukturanalyse von überexprimierten und aufgereinigten Proteinen der Fall, welche daher ein gutes Anwendungsbeispiel darstellt.

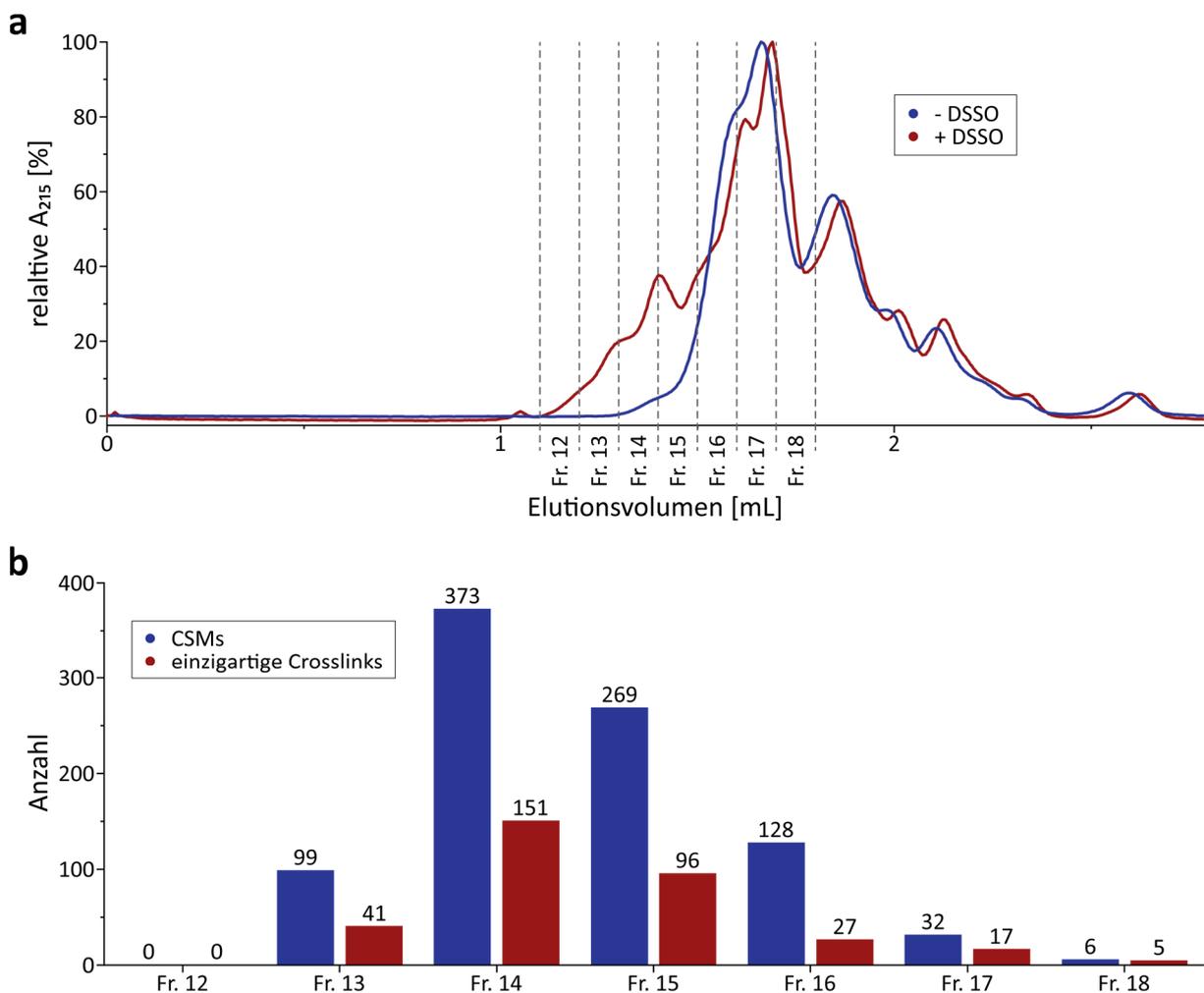


Abbildung 3.14: Fraktionierung von mit DSSO quervernetztem und verdautem BSA mittels SEC. a) Chromatogramme der Fraktionierungen eines BSA-Verdaus ohne bzw. mit vorherigem DSSO-Crosslinking. Die leichte Verschiebung der Chromatogramme entsteht durch eine geräteabhängige geringfügige Variabilität zwischen den Auftrennungen. b) Ergebnis der XL-MS-Analyse einiger SEC-Fractionen des mit DSSO quervernetzten und verdauten BSAs.

3.2.2 Crosslinking-Experimente mit TET3 und GDH

Da das Crosslinking von BSA mit DSSO als Testsystem gute Ergebnisse lieferte, wurde das Protokoll im nächsten Schritt zur Untersuchung einer möglichen direkten Interaktion von TET3 und GDH verwendet. Dabei wurde das Crosslinking zu verschiedenen Zeitpunkten des Protokolls durchgeführt, die im Folgenden besprochen werden.

Die beiden Enzyme wurden in allen Versuchen durch Transfektion entsprechender Plasmide und einer anschließenden Überexpression in HEK-Zellen gewonnen. Das Plasmid für TET3 exprimierte dabei ein Fusionsprotein aus eGFP und TET3, was dessen Aufreinigung über GFP-Antikörper-Partikel ermöglichte. Die Bindung des Fusionsproteins an die Partikel konnte dabei über die eGFP-Fluoreszenz kontrolliert werden. Dabei wurde in einigen Versuchen mit der kompletten TET3-Sequenz (TET3_{fl}) und in anderen Experimenten mit der katalytischen Domäne von TET3 (TET3_{cd}) gearbeitet.

GDH wiederum wurde mit einem C-terminalen FLAG-Tag^[565] exprimiert, wodurch es in einigen Versuchen mithilfe von FLAG-Antikörper-Partikeln aufgereinigt werden konnte. Die erhaltenen massenspektrometrischen Daten wurden mit *MeroX 2.0.1.4* hinsichtlich der Crosslinks zwischen TET3 und GDH, aber auch der Intralinks innerhalb der beiden Proteine, analysiert. Zusätzlich wurde mithilfe des *Proteome Discoverer* untersucht, ob und wie abundant die überexprimierten Proteine vorhanden waren, indem die Anzahl der PSMs der Ziel-Proteine relativ zur Gesamtzahl der PSMs in der Probe betrachtet wurde. Dieser Wert ist ein Indiz für den Anteil des entsprechenden Proteins am Gesamtprotein in der Probe. Auch wenn dies keine quantitative Analyse darstellt, lassen sich Unterschiede zwischen verschieden behandelten Proben so sehr gut nachverfolgen.

In einem ersten Versuch wurden HEK-Zellen nach Plasmid-Cotransfektion und Expression von eGFP-TET3_{fl} sowie GDH-FLAG mit DSSO behandelt und anschließend mit RIPA-Puffer^[566] lysiert. Anschließend wurde eGFP-TET3_{fl} mithilfe der GFP-Antikörper-Partikel aufgereinigt und nach Elution von der Festphase verdaut und massenspektrometrisch analysiert. Mit diesem Protokoll konnten weder Crosslinks zwischen TET3_{fl} und GDH noch Intralinks innerhalb der beiden Proteine identifiziert werden. Die Analyse der Überexpression (Abb. 3.15a, Ansatz 1) ergab jedoch, dass ungefähr 10 % aller identifizierten PSMs von TET3_{fl} stammen. Dies ist zu erwarten, da eGFP-TET3_{fl} durch den Einsatz der GFP-Antikörper-Partikel stark angereichert wurde. Es wurden allerdings keine PSMs von GDH-FLAG gefunden, was darauf hinweisen könnte, dass dieses Plasmid nicht gut exprimiert wurde. Daher wurde im nächsten Ansatz nach der Cotransfektion der HEK-Zellen Natriumbutyrat zum Kulturmedium zugegeben, da dieses die transiente Expression von Plasmiden in humanen Zelllinien erhöht.^[567] Das Crosslinking erfolgte zusätzlich zu der Reaktion in intakten Zellen auch in Proben mit zuvor isolierten Zellkernen. Mit dieser Methode konnten jedoch ebenfalls keine Crosslinks nachgewiesen werden. Die Expressionsanalyse (Abb. 3.15a, Ansatz 2) ergab erneut hohe Werte für eGFP-TET3_{fl} von 4.7 % bzw. 9.0 % aller identifizierten PSMs in der Probe mit kompletten Zellen bzw. in der Zellkernprobe.

In diesen Versuchen konnte auch GDH-FLAG identifiziert werden, jedoch in geringerem Maße (1.0 bzw. 1.6 % aller PSMs). Dies könnte für eine immer noch schlechtere Expression des GDH-Plasmids sprechen oder aber daran liegen, dass GDH mit anderen natürlich vorkommenden TET3-Bindern konkurrieren muss. Prinzipiell bedeutet dies aber, dass das verwendete Protokoll für die Coimmunpräzipitation von TET3 und GDH geeignet ist. Die möglicherweise niedrigere Expression von GDH-FLAG könnte allerdings dafür sorgen, dass eventuell auftretende Crosslinks nicht abundant genug sind, um identifiziert zu werden. Es konnte außerdem festgestellt werden, dass nach der Zelllyse und anschließender Zentrifugation ein großer Teil der eGFP-Fluoreszenz im unlöslichen Rückstand verblieb, während das Proteinlysate nur schwach fluoreszierte.

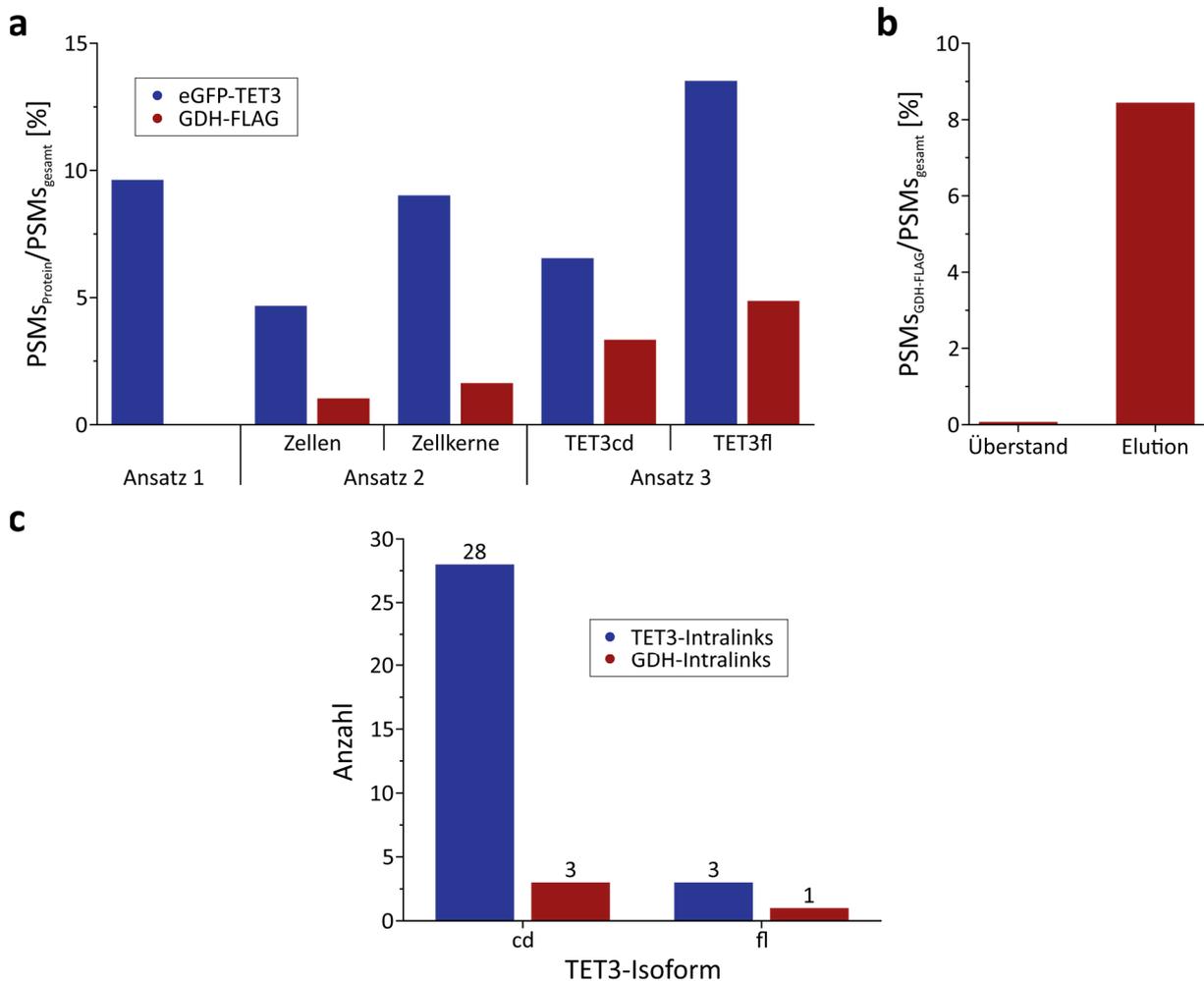


Abbildung 3.15: Ergebnisse der durchgeführten Experimente zum Nachweis einer direkten Interaktion zwischen TET3 und GDH. a) Expressionsanalyse von TET3 und GDH. Dargestellt ist der Anteil der identifizierten PSMs des jeweiligen Proteins relativ zu den insgesamt in der Probe identifizierten PSMs, was ein Hinweis auf den Anteil des entsprechenden Proteins an gesamtem Protein ist. b) Validierung der FLAG-Anreicherung von GDH-FLAG. c) Identifizierte TET3- bzw. GDH-Intralinks in den Proben von Ansatz 3.

Daher wurde vermutet, dass eGFP-TET3_{fl} sich nicht gut in RIPA-Puffer löst und der Großteil des überexprimierten Proteins sowie eventuell damit quervernetztes GDH-FLAG nicht analysiert werden kann. Dies könnte durch das hohe Molekulargewicht von TET3 erklärt werden, da von *Ngoka* gezeigt werden konnte, dass große Proteine teilweise nur schlecht mit RIPA-Puffer solubilisiert werden können.^[568] Aus diesen Gründen wurde das Experiment erneut überarbeitet, indem die Plasmide für eGFP-TET3 und GDH-FLAG in zwei getrennten Kulturen transfiziert und exprimiert wurden. Dies sollte die Expression beider Plasmide verbessern, da die Plasmide nicht mehr um die Ressourcen einer Zelle konkurrieren müssen. Die Lyse erfolgte anschließend in einem anderen Puffer mit 1 % Triton X-100 als Detergens. Dadurch konnte der Anteil an eGFP-TET3_{fl} im Lysat erhöht werden, was sich durch eine stärkere eGFP-Fluoreszenz äußerte. Zusätzlich wurden auch Zellen mit eGFP-TET3_{cd} transfiziert und lysiert. Dieses Protein stellt nur den katalytisch aktiven Teil des Enzyms dar, wodurch das Molekulargewicht des Fusionsproteins verringert wird. Zudem konnte *Dr. Franziska Traube* in dieser Arbeitsgruppe bereits zeigen, dass GDH auch mit TET3_{cd} interagiert. Tatsächlich wurde im Lysat von eGFP-TET3_{cd} eine höhere Fluoreszenz als bei eGFP-TET3_{fl} festgestellt, was bedeutet, dass die katalytische Domäne von TET3 entweder besser exprimiert wird oder besser löslich ist als das Volllängen-Protein.

Das GDH-FLAG-Fusionsprotein wurde nach der Lyse mithilfe von FLAG-Antikörper-Partikeln aufgereinigt und durch die Inkubation mit freiem FLAG-Peptid eluiert, um kompetitive TET3-Binder zu verringern. Der Erfolg dieser Anreicherung wurde massenspektrometrisch untersucht (Abb. 3.15b) und es konnte festgestellt werden, dass 8.4 % aller in der Elutionsfraktion identifizierten PSMs zu GDH gehörten, während im Überstand nach Inkubation mit den Partikeln kaum noch Peptide von GDH-FLAG gefunden wurden. Die TET3-Fusionsproteine wurden wie zuvor an die GFP-Antikörper-Partikel gebunden und aufgereinigt und die so beladene Festphase mit dem zuvor aufgereinigten GDH-FLAG inkubiert. Unspezifische Wechselwirkungen wurden durch mehrmaliges Waschen minimiert, bevor das Crosslinking mit DSSO erfolgte. Reduktion, Alkylierung und Verdau der Proteine erfolgten im Anschluss auf der Festphase. Mit diesem Protokoll (s. Abb. 3.15a, Ansatz 3) konnte der Anteil von eGFP-TET3_{fl} erneut erhöht werden (13.5 % aller PSMs). Der Anteil von eGFP-TET3_{cd} betrug lediglich 6.5 % aller PSMs, was jedoch daran liegen könnte, dass durch die geringere Länge des Proteins weniger einzigartige Peptide entstehen. Auch der Anteil von GDH-FLAG stieg mit dieser Methode auf 4.9 % (Probe mit TET3_{fl}) bzw. 3.3 % (Probe mit TET3_{cd}). Zudem konnten sowohl für TET3_{cd} als auch TET3_{fl} Intralinks zwischen TET3-Peptiden und ebenso GDH-Intralinks identifiziert werden (Abb. 3.15c). Für TET3_{cd} wurden dabei um einiges mehr TET3-Intralinks identifiziert als mit TET3_{fl}. Dies könnte daran liegen, dass durch die geringere Größe des verkürzten TET3 mehr Moleküle an die GFP-Antikörper-Partikel binden können, wodurch die Abundanz der Intralinks erhöht werden würde. Interlinks zwischen TET3 und GDH wurden jedoch erneut nicht identifiziert.

Diese Ergebnisse sprechen prinzipiell dafür, dass das verwendete Protokoll geeignet ist, um TET3 und GDH quervernetzen, deuten allerdings ebenfalls darauf hin, dass eventuell gebildete Interlinks nach wie vor nicht abundant genug sind, um identifiziert zu werden. Ebenfalls wäre es möglich, dass durch die FLAG-Aufreinigung die homo-hexamere Struktur von GDH zerstört wird und die Identifikation von GDH in den Proben nur auf einer unspezifischen Bindung des Proteins an die Partikel beruht. Daher wurde das Crosslinking auch vor der Anreicherung durchgeführt, indem beide Zelllysate zusammen inkubiert und anschließend quervernetzt wurden. Die Komplexität dieser Probe war folglich sehr hoch, weshalb sie nach dem Verdau mittels SEC aufgetrennt wurde, da bereits in Kapitel 3.2.1 der positive Effekt einer Separation der quervernetzten Peptide durch SEC auf die Identifikation von Crosslinks gezeigt werden konnte. In diesem Fall führte die Fraktionierung allerdings nicht zur Identifikation von TET3-GDH-Interlinks, was möglicherweise an deren geringer Abundanz liegen könnte. Weitere von *Dr. Michael Stadlmeier* durchgeführte Experimente mit MS³-Analyse konnten neben TET3-Intralinks einige Crosslinks zwischen TET3 und verschiedenen Proteinen aus dem HEK-Zelllysate identifizieren, eine Interaktion mit GDH konnte allerdings auch hier nicht beobachtet werden.^[569]

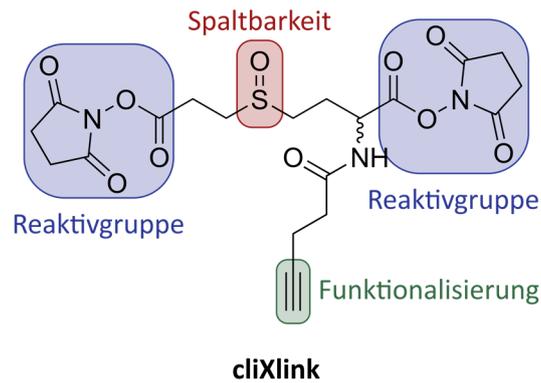
Lange Zeit konnte das Ausbleiben des Nachweises der Interaktion mit diesen Methoden nicht in Einklang mit den in der Gruppe bereits durchgeführten Experimenten gebracht werden. Daher lag der Fokus der Protokoll-Optimierung hauptsächlich auf der Erhöhung der Crosslink-Intensität. Kürzlich konnte durch andere Arbeiten in dieser Gruppe jedoch gezeigt werden, dass die mitochondriale Transportsequenz von GDH eventuell eine große Rolle bei der Interaktion mit TET3 im Nukleus spielt. Diese Sequenz dient dem Transfer des Großteils des von einer eukaryotischen Zelle exprimierten GDH in die Mitochondrien, wo sie im Anschluss abgespalten wird.^[570] Es ist also auch möglich, dass das meiste in den Versuchen durch die humanen Zellen überexprimierte GDH diese Sequenz eventuell nicht mehr besaß und somit nicht mit TET3 interagieren konnte. Weitere Experimente mit gezieltem Einsatz von GDH mit mitochondrialer Transportsequenz könnten also einen erfolgreichen Nachweis der Interlinks zwischen TET3 und GDH ermöglichen.

3.3 Etablierung eines neuen, anreicherbaren Protein-Protein-Crosslinkers

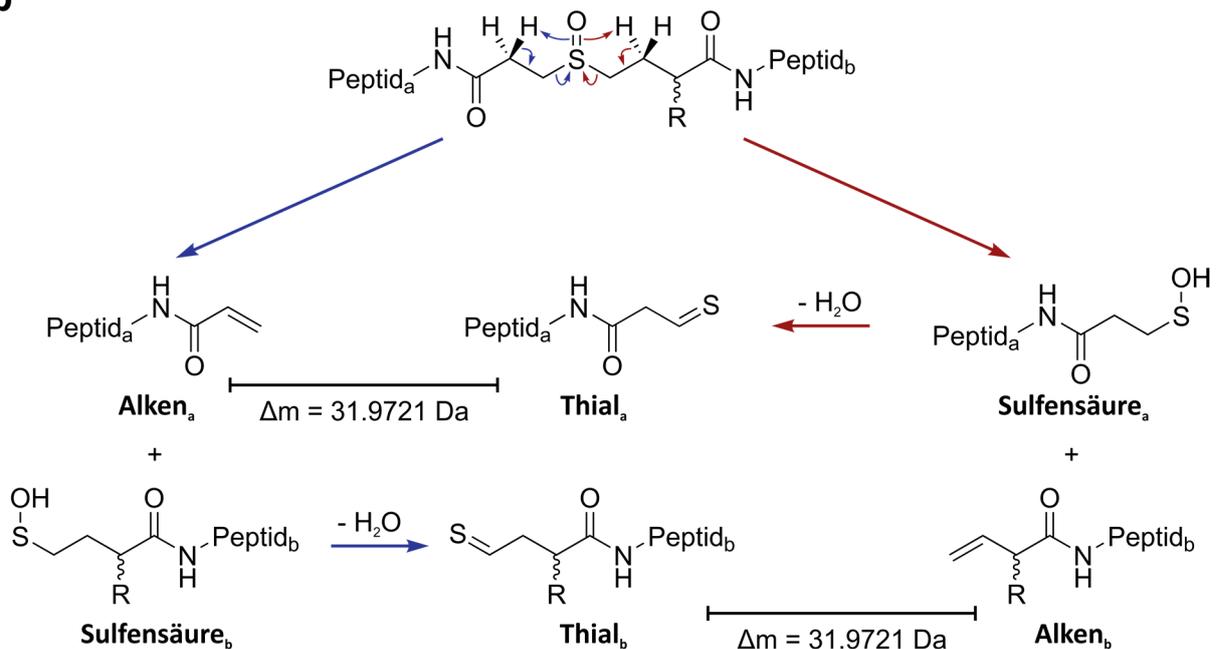
In Kapitel 3.2.1 konnte bereits gezeigt werden, dass die Fraktionierung von Crosslink-Proben, z. B. mit SEC, die Anzahl der detektierten Crosslinks in einer Probe stark erhöhen kann. Ein Nachteil einer solchen Fraktionierung ist jedoch die Notwendigkeit einer großen Menge an Probe. So wurden für die gezeigten Ergebnisse 40 µg von quervernetztem und zu Peptiden verdautem BSA verwendet. Gerade für die Detektion von wenig abundanten Crosslinks in komplexen Proben ist also ein hoher Materialeinsatz vonnöten, welcher abhängig vom untersuchten System eventuell nur schwer bereitgestellt werden kann. Außerdem steigt die Messzeit, da für jede Probe mehrere Fraktionen analysiert werden müssen. In diesem Teil der Doktorarbeit wurde daher in Zusammenarbeit mit *Dr. Michael Stadlmeier* ein neues Crosslinking-Reagenz etabliert, welches die Anreicherung von quervernetzten Peptiden ohne Fraktionierung ermöglicht. Hierfür wurden Methoden entwickelt und optimiert, welche die Grundlage für Protein-Strukturaufklärung und -Interaktomanalyse in komplexen Proben mittels XL-MS bilden. Die Struktur des von *Dr. Michael Stadlmeier* entwickelten und synthetisierten Reagenzes mit dem Namen cliXlink ist in Schema 3.2a dargestellt. Es besitzt zwei NHS-Ester, welche mit räumlich benachbarten Lysinen und in geringerem Umfang auch mit Serinen, Threoninen und Tyrosinen reagieren können. Die Sulfoxid-Gruppe ermöglicht analog zu DSSO die Fragmentierung in der Gasphase bei niedrigen Kollisionsenergien (Schema 3.2b). Dabei entsteht an einem Peptid ein Alken- und am anderen Peptid ein Sulfensäurefragment, wobei letzteres durch den Neutralverlust von H₂O weiter zu einem Thial reagieren kann. Durch die zwei asymmetrisch spaltenden Fragmentierungsstellen bilden sich je Peptid durch Alken- bzw. Thial-Modifikation zwei Reporter-Fragmente mit einer Massendifferenz von ca. 32 Da, welche, wie bereits in Kapitel 3.2.1 beschrieben, die Identifikation eines Crosslinks erleichtern und somit die Verwendung von umfangreichen Protein-Datenbanken ermöglichen.

Zusätzlich besitzt das Reagenz durch die Alkin-Gruppe die Möglichkeit der Funktionalisierung mittels CuAAC-Chemie.^[505, 506] Diese Reaktion ermöglicht eine 1,3-dipolare Cycloaddition, auch Huisgen-Reaktion genannt,^[571, 572] eines primären Alkins mit einem Azid unter milden Bedingungen. So können verschiedene Azid-haltige Moleküle vor oder nach der Crosslink-Bildung an das Reagenz gekoppelt werden, ohne dass die Biomoleküle degradiert werden. Zum Beispiel können Gruppen eingeführt werden, welche die Anreicherung der Crosslinks ermöglichen und damit die Komplexität der Probe reduzieren, was die Identifikation wenig abundanter Crosslinks vereinfacht. Die Vorteile dieses Ansatzes wurden im Rahmen dieser Arbeit anhand der Anreicherung über die Biotin-Streptavidin-Wechselwirkung an einem Testsystem gezeigt.

a



b



Schema 3.2: Übersicht des neuartigen Crosslinkers cliXlink. a) Struktur von cliXlink mit hervorgehobenen Eigenschaften der einzelnen Teile des Moleküls. b) Übersicht der Fragmentierung von cliXlink in der Gasphase. Die Fragmentierung erfolgt durch eine pericyclische Reaktion. Dadurch, dass die Fragmentierung auf beiden Seiten des Sulfoxids erfolgen kann, entstehen Peptidpaare mit charakteristischer Massendifferenz.

3.3.1 Optimierung der Crosslinking-Reaktion

Zunächst wurde die Crosslinking-Reaktion unter Verwendung von cliXlink untersucht. Dabei wurde erneut das in Kapitel 3.2.1 beschriebene Testsystem BSA verwendet, um die erhaltenen Ergebnisse mit denen von DSSO vergleichen zu können. Da erste Tests in nur sehr wenigen Crosslinks resultierten, wurde das Protokoll jedoch an verschiedenen Stellen optimiert (Abb. 3.17). So wurde in einem ersten Schritt die Aufreinigung der Peptide statt mit *StageTips* mithilfe von *ZipTips* durchgeführt. Erstere bestehen aus einem Styrol-Divinylbenzol-Copolymer, welches mit Sulfonsäure-Gruppen funktionalisiert ist. Das Copolymer ermöglicht π - π -Wechselwirkungen mit den aromatischen Aminosäuren der Peptide, während die Sulfonate mit den kationischen Aminosäuren Lysin und Arginin sowie dem Peptid-N-Terminus interagieren.

Im Gegensatz dazu bestehen *ZipTips* aus C₁₈-modifiziertem SiO₂, was die Interaktion mit hydrophoben Aminosäuren ermöglicht. Durch die Verwendung von *ZipTips* konnten die identifizierten CSMs sowie einzigartigen Crosslinks mehr als verdreifacht werden. Dies könnte daran liegen, dass obwohl zwei quervernetzte tryptische Peptide in Lösung mindestens vier positive Ladungen besitzen, die Ladungsdichte von Crosslinks niedriger ist als bei einzelnen Peptiden, da Trypsin nach den mit Crosslinker modifizierten Lysinen nicht schneidet. Somit könnte die Interaktion zwischen Crosslinks und den Sulfonaten des *StageTip*-Materials schwächer sein als bei einzelnen Peptiden, wodurch letztere präferenziell an diese Festphase binden würden. Die Bindung an das C₁₈-Material der *ZipTips* beruht dagegen nur auf hydrophoben Wechselwirkungen, wobei eine geringere Ladungsdichte sogar von Vorteil sein könnte.

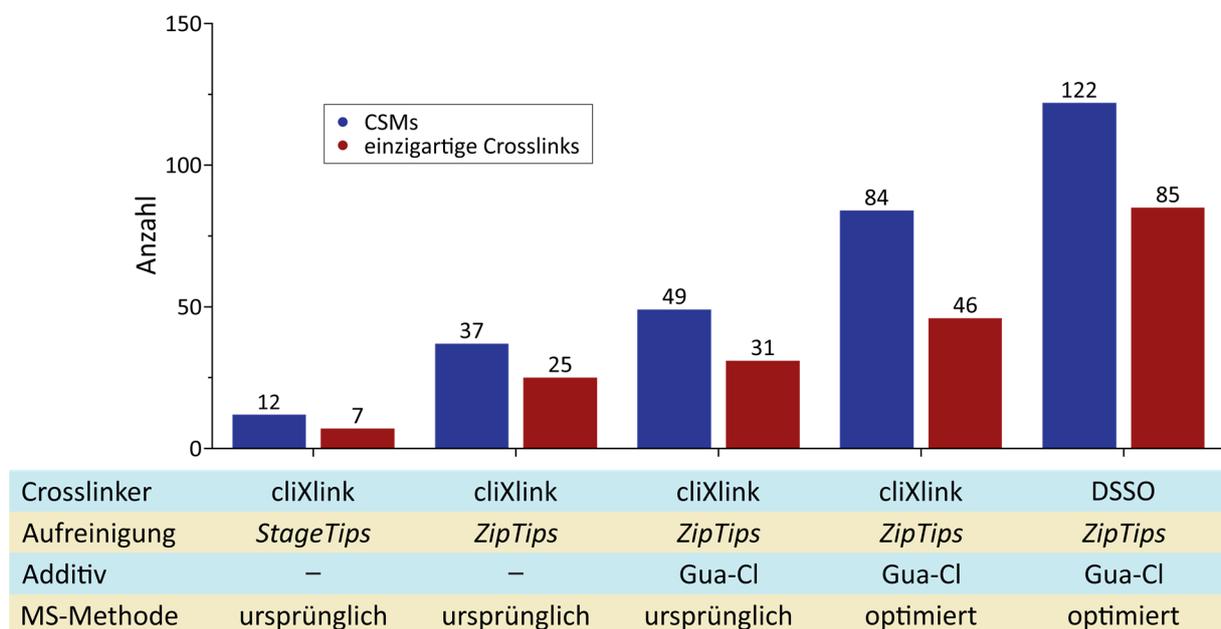


Abbildung 3.16: Optimierung des Crosslinkings von BSA mit cliXlink. Die optimierten Parameter umfassten die Aufreinigung der Peptide, die Resuspendierung der Proteine (mit oder ohne Additiv) nach Acetonfällung sowie die MS-Methode.

Weiterhin wurde die Resuspendierung der Proteine nach Acetonfällung optimiert, indem ein Puffer mit einer hohen Konzentration an Guanidiniumchlorid (6 M) verwendet wurde. Dieses chaotrope Salz hat einen „einsalzenden“ Effekt, was bedeutet, dass es eine bessere Löslichkeit von Proteinen bewirkt, wobei diese denaturiert werden.^[573] Neben der Vorteile für die Resuspendierung sollte diese Methode aufgrund der besseren Zugänglichkeit der Primärstrukturen durch die Denaturierung auch für den tryptischen Verdau vorteilhaft sein. Um allerdings eine Denaturierung von Trypsin zu vermeiden, wurde die Guanidiniumchlorid-Konzentration vor dem Verdau auf 0.45 M gesenkt. Diese Änderung im Protokoll sorgte für eine weitere Erhöhung der identifizierten CSMs und einzigartigen Crosslinks.

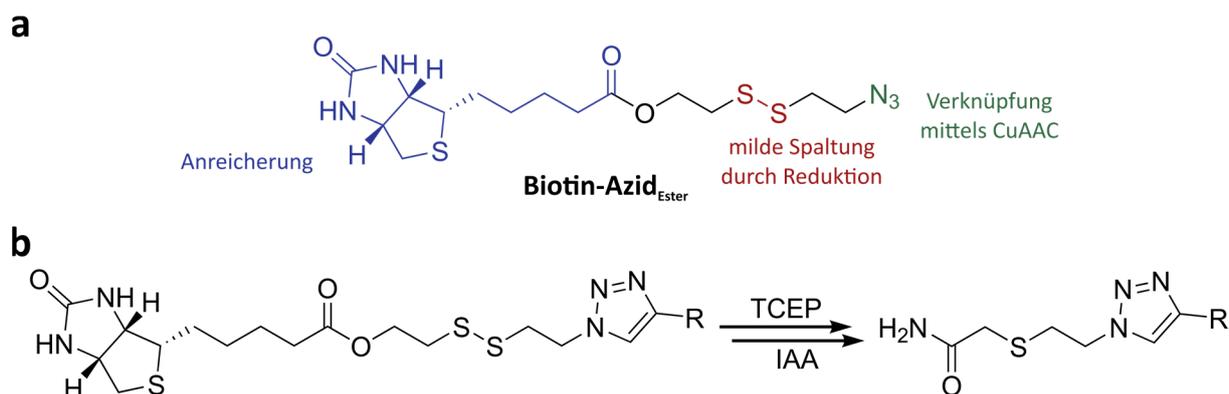
Der Mechanismus hinter der „einsalzenden“ Eigenschaft von chaotropen Salzen ist noch nicht eindeutig geklärt. Ursprünglich ging man davon aus, dass die Wasserstoffbrücken-Bildung zwischen Wassermolekülen reduziert wird, wodurch wiederum die hydrophoben Wechselwirkungen der Proteine geschwächt werden.^[574-576] Mittlerweile gibt es jedoch einige Studien, die Belege gegen diese Theorie anführen^[577-579] und es konnte gezeigt werden, dass direkte Interaktionen zwischen Salz-Ionen mit den Makromolekülen bzw. ihrer Hydrathülle einen größeren Einfluss auf diesen Löslichkeits-Effekt ausüben.^[580]

Die letzte Optimierung des Crosslinkings fand auf der Ebene der massenspektrometrischen Analyse statt. Zum einen wurde die Anzahl der gesammelten Ladungen für das MS²-Spektrum verdoppelt und die maximale Injektionszeit erhöht, was dafür sorgt, dass eine höhere Molekülanzahl pro Peptid fragmentiert wird und die Signale der Fragment-Ionen im MS²-Spektrum dadurch intensiver werden. Da Crosslinks aus zwei vernetzten Peptiden bestehen und somit im Mittel mehr Fragment-Ionen gebildet werden als bei einzelnen Peptiden, ist eine ausreichend hohe Intensität des Vorläufer-Moleküls besonders wichtig. Dies begründet auch die zweite Änderung, nämlich eine Erhöhung der zum Auslösen eines Fragmentierungs-Ereignisses minimal notwendigen Intensität des Vorläufer-Ions. Wenn die Fragmentierung eines Crosslinks bei zu geringer Vorläufer-Intensität ausgelöst wird, können im Rahmen der maximal eingestellten Injektionszeit nicht genug Ionen gesammelt werden, was je nach Vorläufer-Ion in einem MS²-Spektrum von geringer Qualität resultieren könnte. Die Wichtigkeit dieser Tatsache wird durch die mit diesen Anpassungen analysierte Probe deutlich, da die CSMs hierdurch noch einmal fast verdoppelt und die einzigartigen Crosslinks auf 46 angehoben werden konnten.

Zum Vergleich wurde auch ein Experiment mit DSSO unter den optimierten Bedingungen durchgeführt. Man sieht, dass DSSO beim Crosslinking von wenig komplexen Proben überlegen ist, da deutlich mehr CSMs sowie einzigartige Crosslinks identifiziert werden konnten als mit cliXlink. Dies liegt vermutlich an der komplexeren Struktur des neuen Reagenzes. Nichtsdestotrotz konnte gezeigt werden, dass das Crosslinking mit cliXlink funktioniert, weshalb die nächste Aufgabe darin bestand, die Vorteile der CuAAC-Funktionalisierung zu demonstrieren.

3.3.2 Optimierung der Anreicherung von cliXlink-Crosslinks mithilfe von Biotin

Die Möglichkeit der Funktionalisierung von cliXlink wurde am Beispiel der Anreicherung von quervernetzten Peptiden untersucht. Dafür musste zunächst ein entsprechendes Protokoll für die Cycloaddition sowie die eigentliche Anreicherung etabliert werden. Dies geschah erneut am bereits untersuchten Testsystem BSA, indem ein Biotin-Azid (Schema 3.3a) auf Protein- bzw. Peptidebene an das Alkin der Crosslinks gekoppelt wurde. Dadurch konnten die Crosslinks mit immobilisiertem Streptavidin angereichert werden.



Schema 3.3: Biotin-Azid für die Funktionalisierung von cliXlink zur Anreicherung von Crosslinks. a) Struktur des verwendeten Biotin-Azids. b) Die Spaltung des Disulfids mithilfe von milden Reduktionsmitteln wie TCEP erlaubt die Elution der angereicherten Crosslinks von der Streptavidin-Festphase. Im Anschluss werden die freien Thiole mit IAA alkyliert, um eine Reoxidation zu verhindern.

Streptavidin ist ein homotetrameres Protein und hat eine sehr hohe Affinität zu Biotin,^[501] weshalb die Elution von Biotin-haltigen Molekülen nach der Anreicherung eine große Herausforderung ist. Daher besitzt das hier verwendete und von *Dr. Michael Stadlmeier* synthetisierte Reagenz eine Disulfidbrücke, welche reaktiv gespalten werden kann. Durch diese spaltbare Gruppe können die Crosslinks nach der Anreicherung daher leicht durch ein mildes Reduktionsmittel, z. B. TCEP, von der Streptavidin-Festphase abgetrennt werden (s. Schema 3.3b).^[581] Um eine Reoxidation des Thiols zu verhindern, wurden die dadurch entstandenen freien Thiole mit Iodacetamid derivatisiert. Ein weiterer Vorteil dieser Herangehensweise ist der nach Spaltung lediglich sehr kleine Rest, der am Crosslinker verbleibt und daher keinen negativen Einfluss auf die Eigenschaften des Crosslinks bei der MS-Analyse ausübt.

3.3.2.1 CuAAC-Reaktion auf Peptidebene

In einem ersten Ansatz wurde die Funktionalisierung von mit cliXlink quervernetztem BSA nach dem tryptischen Verdau an den Peptiden untersucht. Die dabei durchgeführte CuAAC-Reaktion wurde mithilfe einer Fluoreszenz-Kinetik getestet und optimiert. Hierfür wurde ein Coumarin-Azid verwendet (s. Abb. 3.17a), welches von *Simon Veth, M.Sc.* zur Verfügung gestellt wurde. Während die Fluoreszenz des Coumarins durch das Azid, welches Elektronendichte in das Ringsystem bringt, in den nicht sichtbaren Bereich verschoben wird, sorgt die Reaktion mit einem Alkin und die resultierende Bildung eines Triazols für eine Verschiebung der Coumarin-Fluoreszenz auf ca. 480 nm.^[582] Somit lässt sich das Fortschreiten der Reaktion anhand der Zunahme der Fluoreszenz über die Zeit beobachten. In Abbildung 3.17b ist eine solche Kinetik mit den optimierten Reaktionsbedingungen dargestellt. Als Katalysator wurde hierbei ein System aus CuSO_4 und Natriumascorbat verwendet, wodurch *in situ* Cu(I) gebildet wird. Als Ligand für den aktiven Kupferkomplex diente THPTA.

Man sieht, dass die Fluoreszenz bereits zu Beginn recht hoch ist, was mit der Zeit, die nach dem Mischen der Reagenzien benötigt wurde, um die Messung zu starten, begründet werden kann. Die Fluoreszenz steigt daraufhin zunächst schnell an, bis nach ca. 50 min ein Plateau und damit das Ende der Reaktion erreicht ist.

Es bildet sich also wie erwartet das Triazol aus Crosslinker-Alkin und Coumarin-Azid, wodurch die Fluoreszenz des Coumarins verschoben wird. Die darauffolgende leichte Abnahme der Fluoreszenz lässt sich durch eine mit der Zeit eintretende Photobleichung des Coumarins erklären.

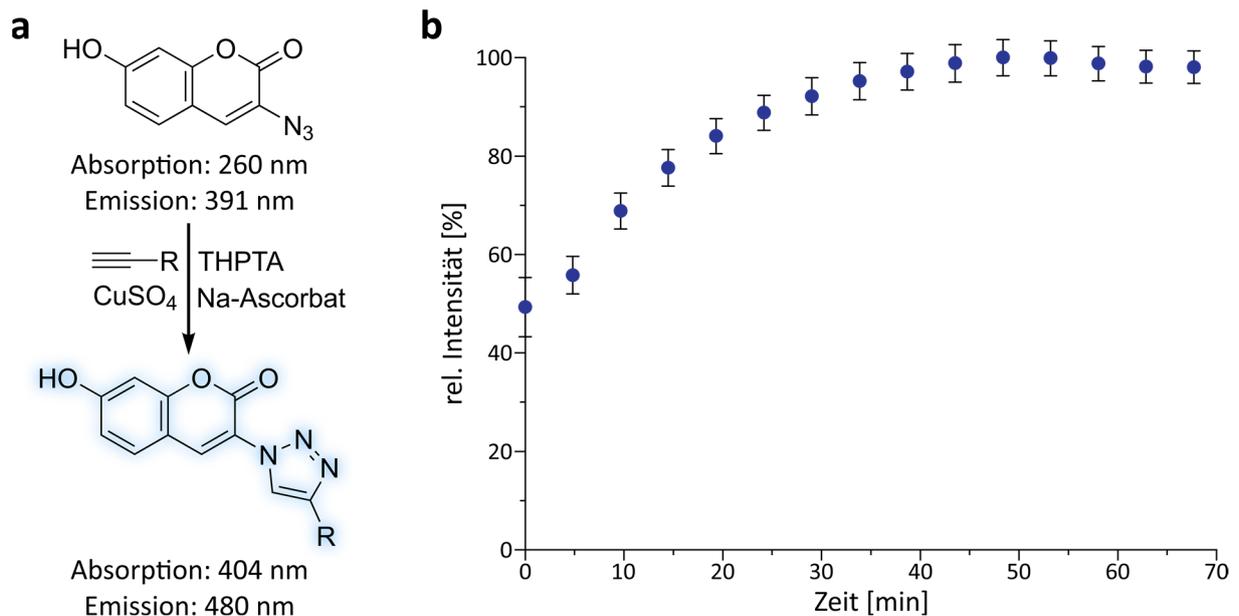


Abbildung 3.17: Fluoreszenzkinetik zur Nachverfolgung der CuAAC-Reaktion. a) Funktionsweise des verwendeten 3-Azido-7-Hydroxycoumarins. Durch die CuAAC-Reaktion wird das Emissionsmaximum der Coumarin-Fluoreszenz in den sichtbaren Bereich verschoben. b) Kinetik der CuAAC-Reaktion mit Coumarin-Azid und verdautem, mit cliXlink quervernetztem BSA in Triplikaten.

Die somit erfolgreich getesteten Reaktionsbedingungen wurden daraufhin auf die Reaktion mit dem Biotin-Azid angewendet, wonach die Proben mit magnetischen Streptavidin-Partikeln angereichert wurden. Die Analyse erfolgte mittels Massenspektrometrie, es konnten jedoch nach Anreicherung keine Crosslinks mehr detektiert werden. Zwei mögliche Ursachen hierfür wurden daher genauer untersucht. Beispielsweise könnten während der CuAAC-Reaktion auftretende Nebenreaktionen für die schlechte Identifikation nach Cycloaddition und Anreicherung verantwortlich sein, z. B. mit den Peptiden an sich. Diese könnten so modifiziert worden sein, dass sie nicht mehr als tryptische Peptide identifiziert werden konnten. So berichten *Allen et al.* beispielsweise von der Spaltung von Histidin-haltigen Peptiden durch Cu^{II}.^[583] Um dies zu testen, wurde mit cliXlink quervernetztes BSA nach dem Verdau mit den Reagenzien der CuAAC-Reaktion, jedoch ohne Azid, inkubiert. Anschließend wurde die Anzahl der identifizierten Crosslinks, aber auch der unmodifizierten Peptide vor und nach der Inkubation verglichen (Abb. 3.18).

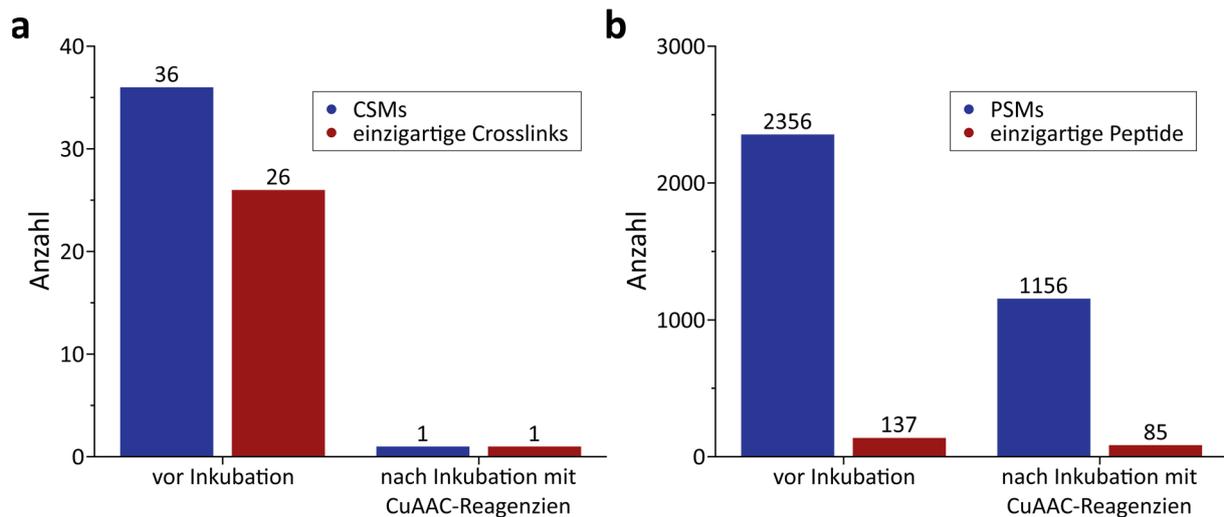


Abbildung 3.18: Analyse der Peptid-Degradation durch die CuAAC-Reagenzien. a) Effekt der CuAAC-Reagenzien auf die Crosslink-Identifikationen. b) Effekt der CuAAC-Reagenzien auf die Peptid-Identifikationen.

Man sieht in Abbildung 3.18a, dass allein durch die Zugabe der CuAAC-Reagenzien beinahe alle CSMs verloren werden. Die Analyse der Peptide (Abb. 3.18b) zeigt ein ähnliches Bild, da nach der Inkubation nur noch ca. die Hälfte der identifizierten PSMs übrig sind. Tatsächlich sorgen die Bedingungen der Cycloaddition bei der Reaktion auf Peptidebene also für eine Degradation oder unbekannte Modifizierung der Peptide im Allgemeinen und nicht nur spezifisch der Crosslinks. Die Crosslinks zeigen jedoch einen größeren Effekt, was damit erklärt werden kann, dass sie durch das Crosslinker-Azid eher in die Nähe des Kupfer-Liganden-Komplexes gelangen.

Da das Biotin-Azid für die Cycloaddition im Überschuss zugegeben wurde, ist es zudem möglich, dass die Streptavidin-Partikel zu einem großen Teil durch unreaktiertes Biotin-Azid gesättigt wurden, weshalb die Biotin-funktionalisierten Crosslinks nicht binden konnten. Zwar erfolgte nach der Reaktion ein Aufreinigungsschritt mittels *StageTips*, allerdings kann es sein, dass auch das unreaktierte Biotin-Azid an das Säulenmaterial binden konnte. Um dies zu untersuchen, wurde ein vereinfachtes System gewählt, indem ein Pentinsäure-Aktivester (s. Abb. 3.19a) eingesetzt wurde. Dieser modifiziert ebenso wie der Crosslinker Lysine, durch die einzelne reaktive Gruppe werden jedoch statt Crosslinks endständige Modifikationen, ähnlich wie Monolinks, gebildet. Dadurch nimmt die Anzahl der modifizierten Spezies zu, weil nicht zwei Lysine in räumlicher Nähe vorliegen müssen und die Auswertung wird vereinfacht, da statt einer Crosslink-Analyse eine normale Peptididentifikation durchgeführt werden kann. Durch das terminale Alkin ist die Modifikation und anschließende Anreicherung mit Biotin-Azid genau wie bei cliXlink möglich. Dies sollte dazu führen, dass nach der Anreicherung ausschließlich Azid-funktionalisierte Peptide identifiziert werden, da nur diese an die Streptavidin-Partikel binden können.

Ein solches Experiment wurde mit Pentinsäure-funktionalisiertem und verdautem BSA durchgeführt. Hierbei wurde die Anzahl der PSMs und der einzigartigen Peptide mit Modifikation vor und nach Anreicherung verglichen (Abb. 3.19b). Dabei wurde zusätzlich zur erwarteten Modifikation nach erfolgreicher Cycloaddition auch die ursprüngliche Pentinsäure-Modifikation berücksichtigt, um Rückschlüsse auf die Effizienz der Reaktion zu treffen zu können. Wie man sieht, wurde durch die Anreicherung eine Steigerung der PSMs mit Azid-Funktionalisierung von ca. 29 % erreicht. Beim Vergleich der einzigartigen Peptide zeigt sich allerdings, dass nach der Anreicherung sogar drei Peptide weniger identifiziert wurden als davor, was an einer Sättigung der Partikel durch unreaktiertes Biotin-Azid liegen könnte.

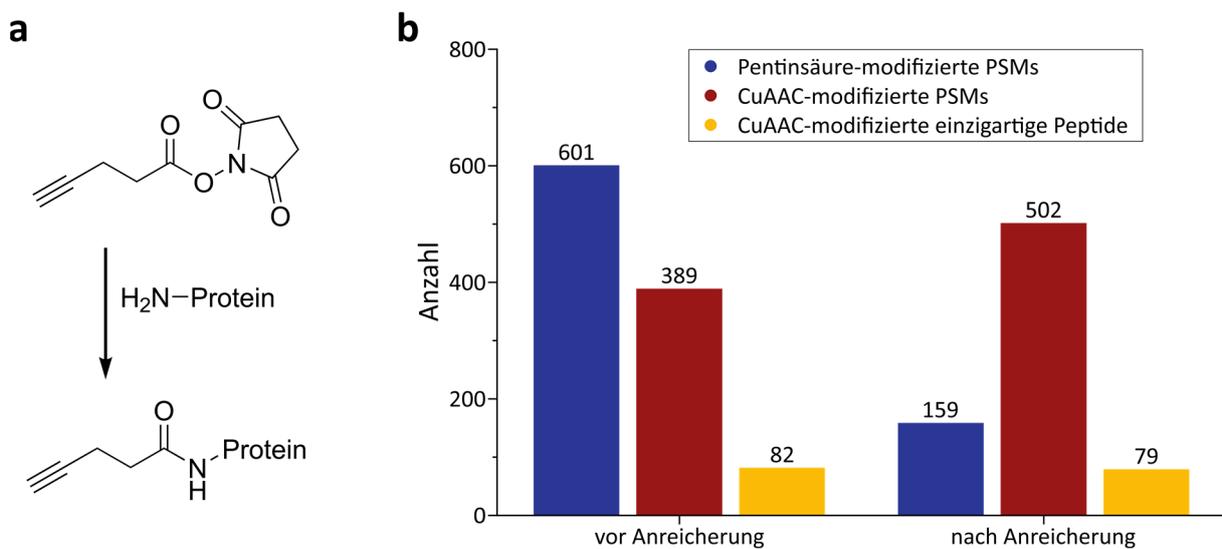


Abbildung 3.19: Kontrolle der Anreicherung mit einem vereinfachten System. a) Funktionalisierung von primären Aminen mit einem Pentinsäure-Aktivester. b) Anzahl der Pentinsäure-modifizierten Peptide ohne (Pentinsäure-modifiziert) sowie mit (CuAAC-modifiziert) Biotin-Azid-Funktionalisierung vor und nach der Anreicherung mithilfe von Streptavidin-Partikeln.

Beim Betrachten der lediglich Pentinsäure-modifizierten PSMs fallen zudem zwei weitere Herausforderungen auf. Erstens ist der Anteil dieser Peptide nach der Cycloaddition ohne Anreicherung recht hoch, was darauf schließen lässt, dass die Reaktion nicht effizient genug ist. Idealerweise sollten alle Alkine reagieren, um eine möglichst hohe Intensität der Crosslinks nach der Anreicherung zu erreichen. Zweitens wurden auch nach der Anreicherung immer noch einige dieser Peptide identifiziert, was bedeutet, dass die Waschschritte nach der Inkubation von Probe mit Streptavidin-Partikeln noch nicht ausreichen, um unspezifische Wechselwirkungen zu unterbinden.

Insgesamt konnte also festgestellt werden, dass die CuAAC-Reaktion auf Peptidebene die Peptide zum Teil degradiert und außerdem nicht quantitativ verläuft. Letzteres könnte zwar ein für Albumin spezifischer Effekt sein, da dieses Protein Cu^{II} -Ionen bindet und damit womöglich aus der Reaktion entfernt,^[584] nichtsdestotrotz war ein Ziel der Alkin-Funktionalisierung ihre universelle Einsetzbarkeit, weshalb weitere Optimierungsschritte durchgeführt wurden.

Diese erfolgten allerdings auf der Proteinebene, also vor dem tryptischen Verdau, um die negativen Effekte der Reagenzien bei der Cycloaddition zu verringern. Die Anreicherung sollte dadurch ebenfalls profitieren, da die Abtrennung von überschüssigem Biotin-Azid auf Proteinebene durch einen Protein-Aufreinigungsschritt leicht möglich sein sollte.

3.3.2.2 CuAAC-Reaktion auf Proteinebene

Durch die Durchführung der CuAAC-Reaktion vor dem Protein-Verdau sollte es möglich sein, die oben genannten Nachteile der Reaktion auf Peptidebene zu vermeiden. Zwar sind auch für Proteine einige selektive Spaltungen durch Cu^{II} in Verbindung mit Ascorbat bekannt,^[585, 586] allerdings sind diese stark abhängig von der Aminosäuresequenz und sollten daher nicht so sehr ins Gewicht fallen.

Dafür musste das Protokoll jedoch an einigen Stellen angepasst werden. Besonders wichtig war hierbei die Abtrennung von überschüssigem Crosslinker nach der Inkubation mit Protein, da dieser sonst ebenfalls mit dem Azid reagieren würde. Dasselbe gilt für überschüssiges Biotin-Azid nach der Cycloaddition aus den in Kapitel 3.3.2.1 beschriebenen Gründen. Daher wurden die Proteine zunächst nach Crosslinking sowie nach CuAAC über eine 10 kDa MWCO-Säule umgepuffert. Somit wurden die überschüssigen Reagenzien als kleine Moleküle unterhalb des Filterlimits entfernt. In Abbildung 3.20a ist das verwendete Protokoll dargestellt. Nach Durchführung des Crosslinkings und der CuAAC-Reaktion erfolgte der enzymatische Protein-Verdau, wonach die Biotin-funktionalisierten Crosslinks mittels Streptavidin-Partikeln aus der komplexen Peptidmischung angereichert wurden. Durch Reduktion des Disulfids konnten die Crosslinks im nächsten Schritt eluiert werden.

Abbildung 3.20b zeigt die nach Crosslinking, Cycloaddition bzw. Anreicherung identifizierten CSMs und einzigartigen Crosslinks. Man erkennt, dass bereits nach dem Crosslinking kaum und nach der Reaktion mit Biotin-Azid gar keine Crosslinks identifiziert werden konnten. Durch die Anreicherung war es möglich, wieder einige Crosslinks zu detektieren, allerdings liegt die Zahl der Identifikationen weit unter den bereits in Kapitel 3.3.1 ohne Anreicherung erhaltenen Werten.

Da bei den im genannten Abschnitt durchgeführten Experimenten bereits der positive Effekt der Protein-Denaturierung mithilfe von Guanidiniumchlorid bei der Resuspendierung gezeigt werden konnte, wurde in einem nächsten Ansatz auch die Cycloaddition in einem Guanidiniumchlorid-haltigen Puffer durchgeführt. Durch die Denaturierung könnte die Zugänglichkeit des Crosslinker-Alkyls und damit die Effizienz der Reaktion erhöht werden. Dafür wurde bereits bei der MWCO-Aufreinigung nach dem Crosslinking ein Puffer mit Guanidiniumchlorid verwendet. Dies sorgte für einen sehr hohen Zuwachs an identifizierten Crosslinks vor der CuAAC-Reaktion (s. Abb. 3.20b), was dafürsprechen könnte, dass die Proteine im Versuch ohne Guanidiniumchlorid auf dem Säulenmaterial ausgefällt oder anderweitig adsorbiert wurden, sodass sie nicht wieder eluiert werden konnten.

Nach der CuAAC-Reaktion konnten jedoch erneut keine CSMs mehr identifiziert werden. Im Vergleich zu der Probe ohne Gua-Cl konnten nach der Anreicherung jedoch zumindest mehr CSMs und einzigartige Crosslinks identifiziert werden, was dafürspricht, dass die Cycloaddition effizienter abläuft. Es wird jedoch auch deutlich, dass durch die Verwendung von MWCO-Säulen noch viele Crosslinks verloren gehen, da nach der Anreicherung nicht mehr vergleichbare Zahlen wie nach dem reinen Crosslinking erreicht werden. Daher wurden weitere Proben angefertigt, bei denen die Aufreinigungsschritte statt mit MWCO-Säulen mithilfe von Acetonfällungen durchgeführt wurden.

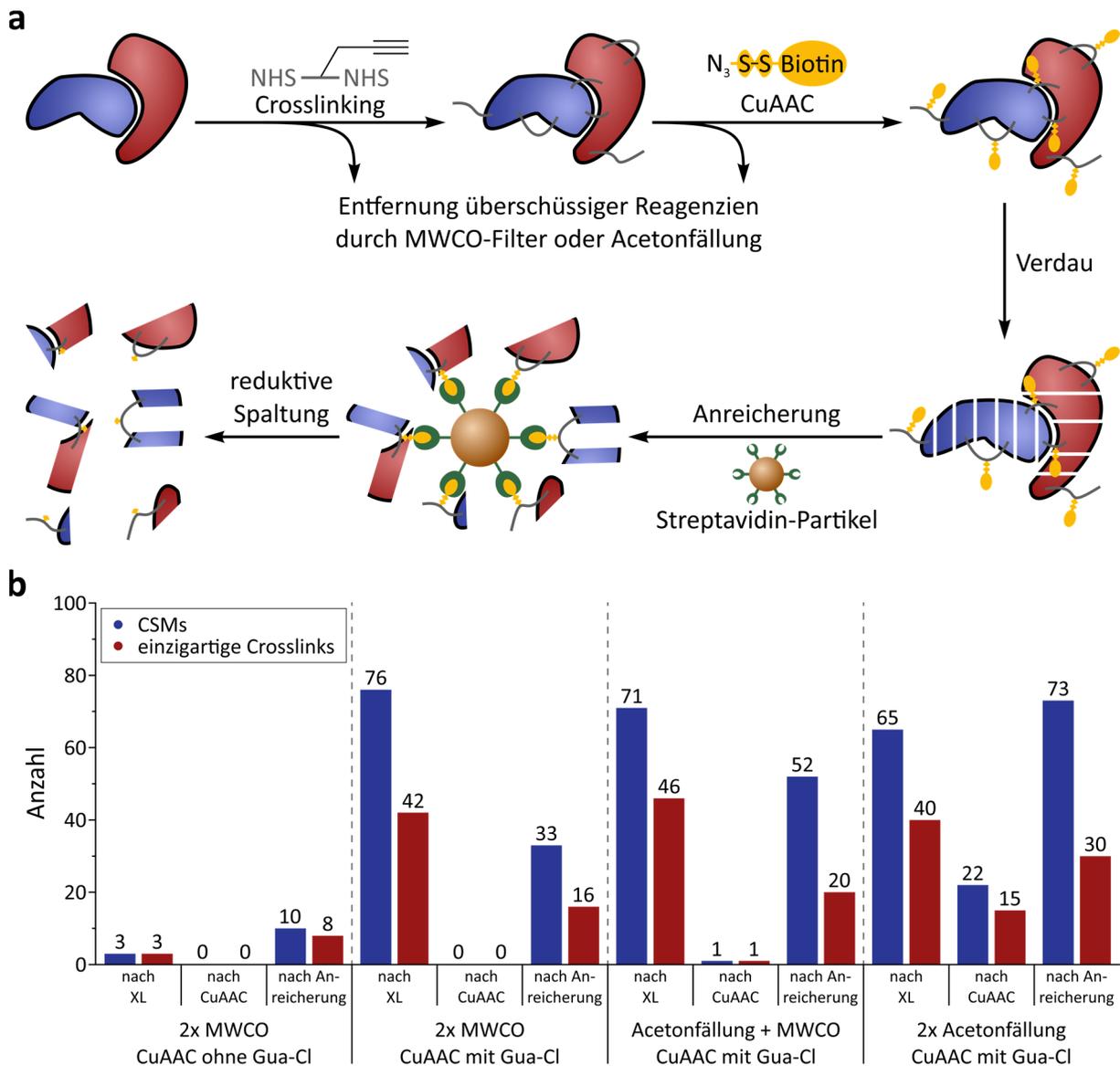


Abbildung 3.20: Optimierung des Protokolls mit CuAAC-Reaktion auf Proteinebene. a) Übersicht über die verschiedenen Schritte des verwendeten Protokolls. In diesem Beispiel ist die Vernetzung zweier verschiedener Proteine dargestellt, dasselbe Prinzip liegt jedoch auch bei der Analyse einzelner Proteine zugrunde. b) Ergebnisse der Protokoll-Optimierung. Die optimierten Parameter umfassen die Aufreinigung nach Crosslinking bzw. CuAAC sowie den Puffer der CuAAC-Reaktion.

Wurde lediglich die Aufreinigung nach dem Crosslinking ersetzt, konnte die Zahl der identifizierten Spezies nach Anreicherung bereits deutlich gesteigert werden. Nach der CuAAC-Reaktion wurde jedoch nur ein einzelner Crosslink identifiziert, da hierbei wieder ein MWCO-Filter verwendet wurde. Bei zusätzlichem Umstellen dieses Aufreinigungsschritts auf eine Acetonfällung konnten auch nach der CuAAC-Reaktion noch 22 CSMs und 15 einzigartige Crosslinks identifiziert werden. Die Identifikationen nach Anreicherung waren ebenfalls höher als zuvor und reichten an die Zahlen nach dem Crosslinking heran.

Diese Ergebnisse zeigen deutlich, dass Funktionalisierung und Anreicherung auf Proteinebene verbessert werden konnten. Dennoch ist zu erkennen, dass besonders die Cycloaddition auch nach der Optimierung für den Verlust eines großen Teils der Identifikationen sorgt, weshalb ein weiteres Screening dieser Reaktion durchgeführt wurde, um die Reaktionsbedingungen noch weiter zu verbessern. Dafür wurde 2-Azidoethanol anstelle des Biotin-Azids verwendet, da eine Aufreinigung hierbei nicht notwendig war. Es wurde sowohl die Anzahl der unmodifizierten als auch Azid-funktionalisierten CSMs nach der Cycloaddition untersucht, um Informationen zum einen über die Effizienz der CuAAC-Reaktion und zum anderen über mögliche negative Nebeneffekte der Reaktionsbedingungen zu erhalten.

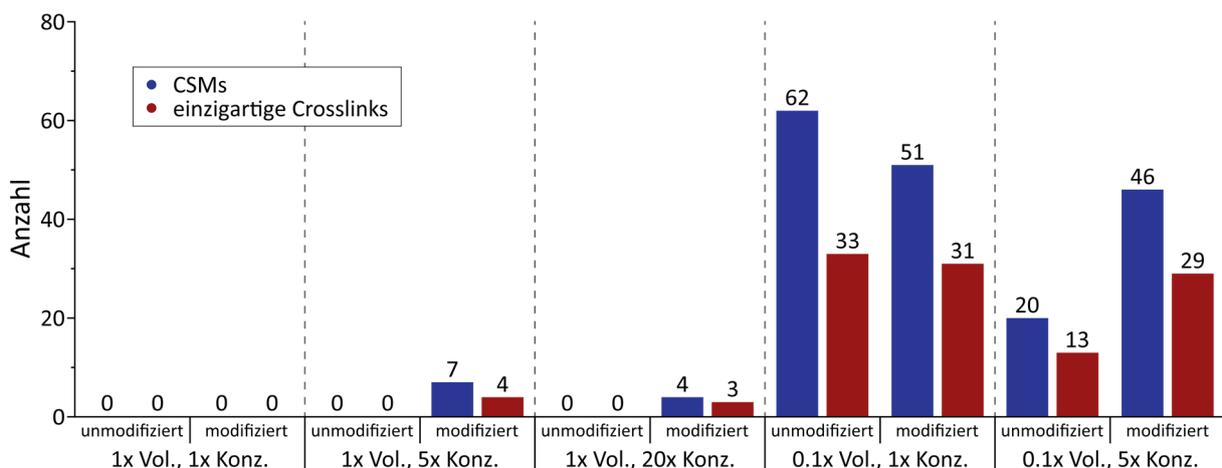


Abbildung 3.21: Optimierung der CuAAC-Reaktion unter Verwendung von 2-Azidoethanol. Analysiert wurden BSA-clipXlink-Crosslinks ohne (unmodifiziert) sowie mit (modifiziert) 2-Azidoethanol-Funktionalisierung. Die Angaben der Volumina und Konzentrationen sind relativ zu den ersten Bedingungen zu verstehen.

Mit den bisher verwendeten Reaktionsbedingungen wurden auch mit diesem vereinfachten System keine Azid-modifizierten CSMs identifiziert (s. Abb. 3.21). Auch unmodifizierte CSMs wurden nicht gefunden, was dafürspricht, dass ein großer Teil der CSMs degradiert wird. Wurde die Konzentration der Reagenzien CuSO_4 , THPTA sowie 2-Azidoethanol jedoch um das Fünffache erhöht, konnten 7 CSMs und 4 einzigartige Crosslinks detektiert werden. Eine Erhöhung der Konzentration von CuSO_4 und THPTA um das Zwanzigfache führte allerdings wieder zu einer Abnahme der Identifikationen, möglicherweise bedingt durch die bereits bei den Peptiden beobachtete Degradation.

In einer weiteren Probe wurde das Volumen für die Resuspendierung nach der ersten Acetonfällung und auch der anschließenden CuAAC-Reaktion auf ein Zehntel verringert, wodurch nur die Konzentration des Crosslinker-Alkins steigt. Dies sorgte für einen starken Anstieg an Azid-modifizierten CSMs auf 51, allerdings konnten auch 62 unreaktierte CSMs identifiziert werden, was vermutlich an den veränderten relativen Konzentrationen der Reaktanden liegt. In der Tat zeigte sich bei einer Kombination aus verringertem Volumen und einer fünffach höheren Konzentration der CuAAC-Reagenzien, dass die Zahl der unmodifizierten CSMs deutlich abnahm, während die Azid-funktionalisierten CSMs mit einem Wert von 46 auf einem vergleichbaren Niveau blieben, weshalb diese Bedingungen für die weiteren Experimente verwendet wurden.

Mit den optimierten Bedingungen wurde jedoch festgestellt, dass sich bei Verwendung des Biotin-Azids anstelle von 2-Azidoethanol im Laufe der Reaktion ein Niederschlag bildete. Dies könnte an der Verringerung der Löslichkeit durch eine erhöhte Hydrophobizität des an vielen Stellen mit Biotin-Azid reaktierten Proteins liegen. Daher wurden einige verschiedene Puffer getestet, um das Ausfallen von BSA zu verhindern und die Identifikationsrate noch weiter zu steigern. Um die massenspektrometrischen Eigenschaften der Crosslinks zu verbessern, wurde das Disulfid reduziert und die Thiole alkyliert, obwohl nicht angereichert wurde. Zunächst wurde untersucht, bei welchen Bedingungen optisch kein Niederschlag mehr zu erkennen war. Dies konnte durch eine Erhöhung der Konzentration von Guanidiniumchlorid auf 1 M, 3 M oder 6 M erreicht werden und auch bei Verwendung von 1 % SDS anstelle von Guanidiniumchlorid oder der Durchführung der Reaktion in 30 % MeCN. Anschließend wurde die Effizienz der Cycloaddition unter den verschiedenen Reaktionsbedingungen verglichen (Abb. 3.22). Unter den bereits zuvor etablierten Bedingungen mit 0.45 M Guanidiniumchlorid wurden 59 modifizierte und nur 5 unreaktierte CSMs identifiziert, was bedeutet, dass die Funktionalisierung des Alkins bereits mit hoher Effizienz abläuft. Durch die Verwendung von SDS sowie eines Puffers mit 30 % MeCN konnte das Ausfallen des Proteins zwar verhindert werden, jedoch lief die CuAAC-Reaktion unter diesen Bedingungen anscheinend nicht effizient ab, da deutlich mehr unreaktierte als Azid-funktionalisierte CSMs identifiziert wurden. Dies könnte an einer Bindung des Dodecylsulfats bzw. Cyanids an die aktive Spezies des Cu^I-THPTA-Komplexes und dessen dadurch bedingte Inaktivierung liegen. Bei einer Erhöhung der Guanidiniumchlorid-Konzentration auf 1 M wurden dagegen 75 CSMs detektiert, was einer Steigerung von 27 % im Vergleich zu den bisherigen Standardbedingungen entspricht. Die Zahl der unreaktierten Spezies blieb dabei gering, es handelt sich hierbei also tatsächlich um einen Effekt der besseren Löslichkeit, der die Funktionalisierung von mehr Crosslinks pro Protein-Molekül ermöglicht.

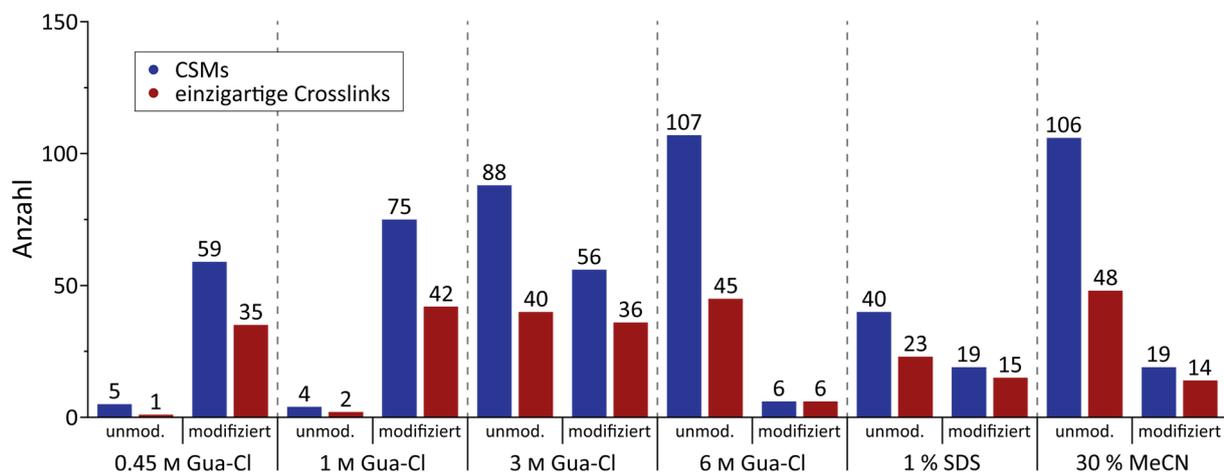


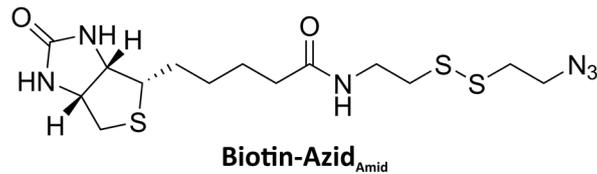
Abbildung 3.22: Verbesserung der Löslichkeit von mit Biotin-Azid_{Ester} modifizierten Crosslinks. Analysiert wurden BSA-cliXlink-Crosslinks ohne (unmod.) sowie mit (modifiziert) Biotin-Azid-Funktionalisierung.

Bei einer weiteren Erhöhung der Konzentration auf 3 M oder sogar 6 M Guanidiniumchlorid sank die Zahl der modifizierten CSMs allerdings wieder und es wurden sehr viele unmodifizierte Spezies identifiziert. Auch dies könnte ein Effekt der Inaktivierung des katalytisch aktiven Komplexes durch die Bindung von Guanidin bei höheren Konzentrationen sein.

In einem letzten Ansatz zur Verbesserung der Anreicherung wurde das verwendete Biotin-Azid_{Ester} leicht modifiziert. Da es einen Ester enthielt, wurde die Möglichkeit in Betracht gezogen, dass dieser im Verlauf der Probenbehandlung teilweise hydrolysierte und dadurch ebenfalls einen negativen Einfluss auf die Anzahl der Crosslink-Identifikationen hatte. Von *Filipp Streshnev, M.Sc.* wurde daher ein weiteres Azid synthetisiert, welches statt des Esters ein gegenüber Hydrolyse stabileres Amid besitzt (Abb. 3.23a). Abbildung 3.23b zeigt einen Vergleich der identifizierten Crosslinks nach Cycloaddition mit den beiden Biotin-Reagenzien. In diesem Fall wurde auf eine Spaltung der Disulfidbrücke des Azids verzichtet, um den Effekt der Ester-Hydrolyse untersuchen zu können. Wie man sieht, wurden mit dem Amid-haltigen Reagenz eine vergleichbare Anzahl von modifizierten CSMs sowie einzigartigen Crosslinks wie mit dem bisher verwendeten Biotin-Azid_{Ester} gefunden. Auch die Anzahl der nicht durch die CuAAC-Reaktion modifizierten Spezies unterscheidet sich nur geringfügig. Dies ist wahrscheinlich auf die Variabilität zwischen den MS-Analysen zurückzuführen und deutet darauf hin, dass das Amid keinen großen Einfluss auf die Effizienz der Cycloaddition ausübt. Im Vergleich der hydrolysierten CSMs fällt jedoch auf, dass der Effekt der unerwünschten Hydrolyse beim Ester-haltigen Biotin-Azid deutlich größer ist, da für dieses 19 CSMs und 13 einzigartige Crosslinks identifiziert werden konnten. Bei der Reaktion mit Amid-haltigem Biotin-Azid wurden hingegen nur zwei Spezies detektiert, weshalb es in den weiteren Experimenten verwendet wurde.

Insgesamt war es also möglich, die CuAAC-Reaktion von mit cliXlink quervernetztem BSA und einem Biotin-Azid und die anschließende Anreicherung an verschiedenen Stellen zu verbessern. Dies ermöglichte somit den konzeptionellen Nachweis der Funktionsweise des neuen Crosslinkers cliXlink. Dadurch konnte eine Anreicherung von BSA-Crosslinks aus einem komplexen Hintergrund von unmodifizierten Peptiden erreicht werden, was zu einer Veröffentlichung dieser Daten in der Zeitschrift *ChemBioChem* führte. Im Folgenden wird daher die Publikation vorgestellt und auf den Beitrag, der im Rahmen dieser Doktorarbeit geleistet wurde, eingegangen.

a



b

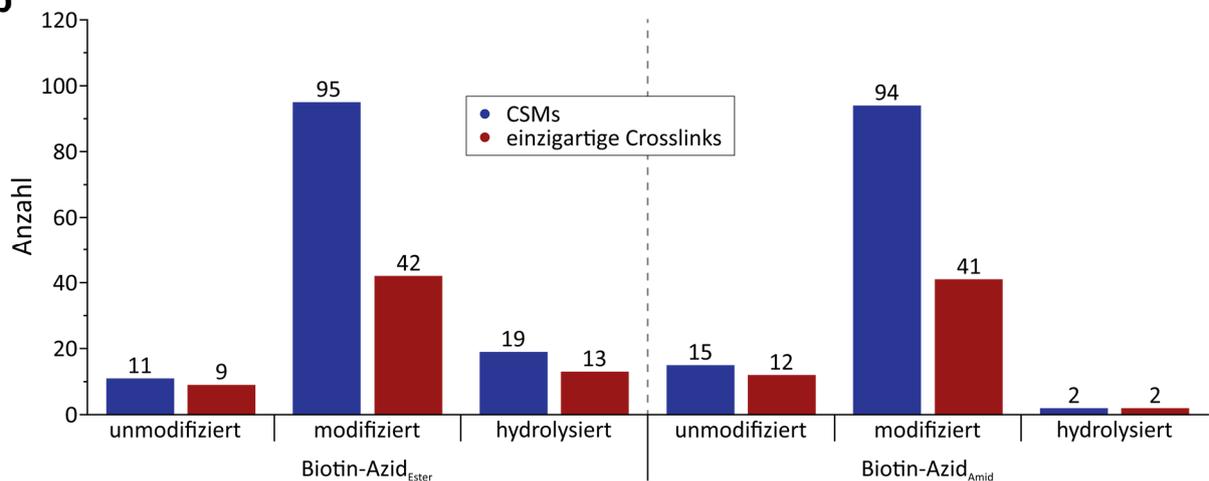


Abbildung 3.23: Untersuchung der Hydrolyse von Biotin-Azid_{Ester} im Vergleich zu einem neuen, Amid-haltigen Reagenz. a) Struktur des Amid-haltigen Biotin-Azids. b) Analyse der CuAAC-Reaktion mit den beiden Biotin-Azid-Reagenzien. Untersucht wurden BSA-cliXlink-Crosslinks ohne (unmodifiziert) bzw. mit (modifiziert) Biotin-Azid-Funktionalisierung und zusätzlich Crosslinks, die zwar mit dem Azid reagiert hatten, bei denen der Ester bzw. das Amid jedoch im Laufe der Probenvorbereitung hydrolysierte.

3.3.3 Publikation

A Click-Chemistry-Based Enrichable Crosslinker for Structural and Protein Interaction Analysis by Mass Spectrometry

Michael Stadlmeier*, [Leander S. Runtsch](#)*, Filipp Streshnev, Martin Wühr, Thomas Carell
ChemBioChem, **2020**, *21*, 103–107.

* Beitrag der Autoren zu gleichen Teilen.

3.3.3.1 Zusammenfassung

Ausgehend von den in dieser Arbeit durchgeführten Optimierungen konnte in der Publikation anhand des bereits erwähnten Modellsystems BSA gezeigt werden, dass das neu entwickelte Reagenz cliXlink für die XL-MS²-Analyse von Proteinen geeignet ist und es konnten zusätzlich Hinweise gefunden werden, dass auch ein MS³-Ansatz möglich sein könnte. Für wenig komplexe Proben wie aufgereinigte Proteine kann das neue Molekül analog zu bereits kommerziell erhältlichen Reagenzien verwendet werden und erzielt ähnliche Ergebnisse. Zusätzlich wurden Experimente durchgeführt, die den Nutzen der mittels CuAAC-Chemie funktionalisierbaren Azid-Gruppe des neuen Crosslinkers aufzeigen. Dafür wurde mit cliXlink quervernetztes BSA mit einem spaltbaren Biotin-Azid gekoppelt, welches die Anreicherung der Crosslinks aus einem komplexen, nicht quervernetzten Peptidhintergrund mithilfe von Streptavidin-beschichteten Partikeln und deren anschließende Elution unter milden Bedingungen ermöglichte. Dies führte im Vergleich zu einer nicht angereicherten Probe zu einem deutlichen Anstieg der identifizierten Crosslinks. Somit konnte gezeigt werden, dass cliXlink sowohl für die Strukturanalyse von aufgereinigten Proteinen als auch für die Untersuchung von Protein-Interaktomen in komplexen Proben geeignet ist.

3.3.3.2 Autorenbeitrag

Das in dieser Studie gezeigte Crosslinking von BSA sowie die Anreicherungs-Experimente mit dem neuartigen Crosslinking-Reagenz wurden von mir in Zusammenarbeit mit *Dr. Michael Stadlmeier* konzeptualisiert, durchgeführt und ausgewertet. Auch die Analyse der Zellgängigkeit des Reagenzes sowie die Berechnung der euklidischen C_α-C_α-Distanzen der Crosslinks wurde von mir durchgeführt. Des Weiteren war ich an der Anfertigung des Manuskripts beteiligt.



A Click-Chemistry-Based Enrichable Crosslinker for Structural and Protein Interaction Analysis by Mass Spectrometry

Michael Stadlmeier⁺,^[a, b] Leander Simon Runtsch⁺,^[a] Filipp Streshnev,^[a] Martin Wühr,^[b] and Thomas Carell^{*[a]}

Mass spectrometry is the method of choice for the characterisation of proteomes. Most proteins operate in protein complexes, in which their close association modulates their function. However, with standard MS analysis, information on protein–protein interactions is lost and no structural information is retained. To gain structural and interactome data, new crosslinking reagents are needed that freeze inter- and intramolecular interactions. Herein, the development of a new reagent, which has several features that enable highly sensitive crosslinking MS, is reported. The reagent enables enrichment of crosslinked peptides from the majority of background peptides to facilitate efficient detection of low-abundant crosslinked peptides. Due to the special cleavable properties, the reagent can be used for MS² and potentially for MS³ experiments. Thus, the new crosslinking reagent, in combination with high-end MS, should enable sensitive analysis of interactomes, which will help researchers to obtain important insights into cellular states in health and diseases.

Proteins need to interact with other proteins to form functional complexes. In many cases, protein function inside cells cannot be understood without information about the protein structure and knowledge of in which complex the protein is situated.^[1,2] This is, for example, of paramount importance for proteins that modulate epigenetic information on DNA. Almost all of these chromatin-modifying proteins require intensive interaction with metabolic enzymes, which provide the cofactors needed for histone acetylation, deacetylation or methylation

and demethylation.^[3–5] This showcases the need to study the complex environment of a given protein to analyse its function and activity state. Protein crosslinking, in combination with analysis by mass spectrometry (XL-MS), is ideally suited as a method to gain information about protein structure and the composition of protein complexes.^[6–9] For XL-MS, specialised chemical reagents, crosslinkers, are required, which are able to covalently connect protein residues that are in close proximity, for example, interacting in a complex.^[10,11] Characterisation of the crosslinked peptides by means of mass spectrometry, however, poses a formidable challenge for two reasons. Firstly, crosslink identification has to be achieved by analysing fragment ions of not only one, but two, connected peptides, thereby massively complicating MS² spectra. Secondly, the crosslinked peptide species have only a very low abundance and are part of a peptide mixture that is overwhelmingly dominated by non-crosslinked peptides. In many cases, this leads to dramatic information loss that hampers accurate crosslink identification, and therefore, interactome analysis. A few crosslinking reagents were developed, which introduced MS-cleavable groups, and the possibility for XL enrichment to tackle these problems.^[12–18]

Herein, we report the development of a new crosslinking reagent that is enrichable and has cleavage properties which allow accurate crosslink identification. The designed cliXlink (**1**) is depicted in Figure 1. The cell-permeable reagent (Figure 1A and Figure S1 in the Supporting Information) features two succinimidyl ester units that can react with nucleophiles in proteins and are spaced roughly 9 Å apart; thus enabling the fixation of short distances to gain structural information on proteins and to capture proteins in close proximity to each other. Efficient MS cleavability is ensured by the well-established sulfide group, which can be cleaved under low-energy collision-induced dissociation (CID) conditions prior to peptide fragmentation. In addition, an alkyne unit allows an enrichment moiety, for example, biotin, to be attached to the crosslinked sites with the help of CuAAC reaction.^[19]

Application of the crosslinker can follow the workflow depicted in Figure 1B. It involves the addition of reagent **1** to a complex proteome, fixing the native state distance information of peptides in close proximity and preserving them throughout the further workflow for MS analysis. Attachment of the affinity group for enrichment is performed after crosslinking to avoid interference from bulky enrichment groups. By modification of the crosslinked peptides at the protein level, excess small-molecule reagents can readily be removed by acetone precipita-

[a] Dr. M. Stadlmeier,⁺ L. S. Runtsch,⁺ F. Streshnev, Prof. Dr. T. Carell
Department of Chemistry, Ludwig-Maximilians-Universität München
Butenandtstrasse 5–13, 81377 Munich (Germany)
E-mail: thomas.carell@lmu.de

[b] Dr. M. Stadlmeier,⁺ Prof. Dr. M. Wühr
Department of Molecular Biology and
the Lewis-Sigler Institute for Integrative Genomics, Princeton University
Washington Road, Princeton, NJ 08544 (USA)

[*] These authors contributed equally to this work.

 Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201900611>.

 © 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

 This article is part of a Special Issue commemorating the 20th Anniversary of ChemBioChem. To view the complete issue, visit our homepage

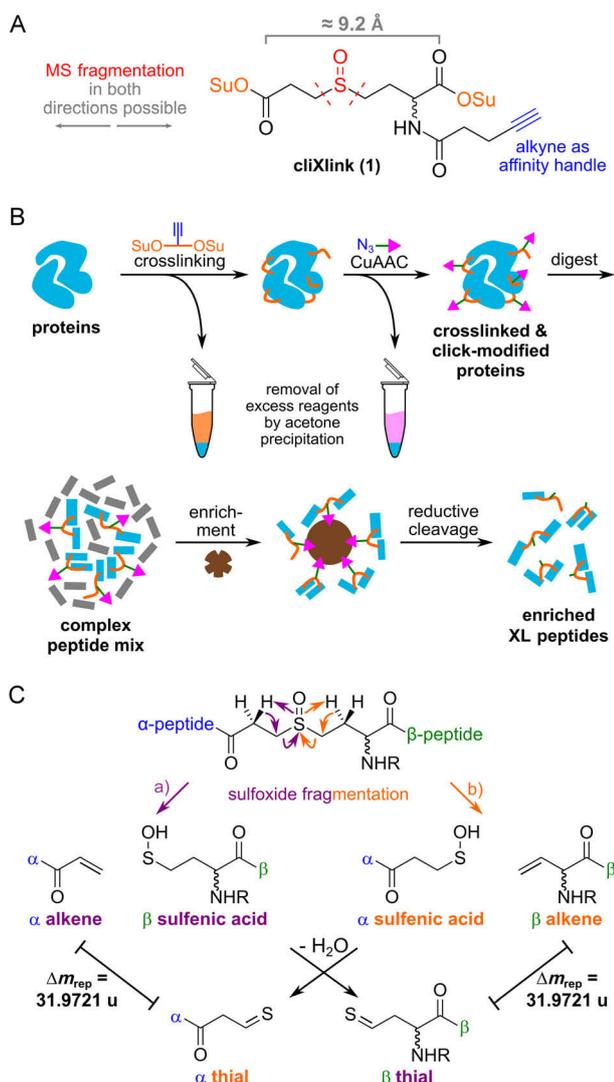
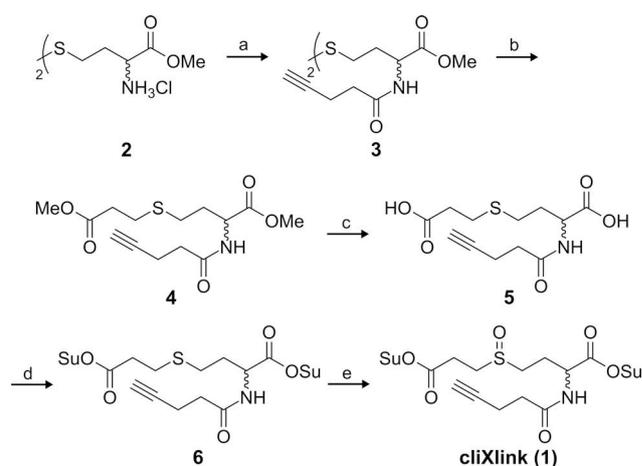


Figure 1. A) Structure of **1**. Su: succinimidyl. B) Workflow for XL-MS experiments. After the addition of **1**, protein sites in close proximity are covalently linked. The generated crosslinks can be functionalised with a copper-catalysed Huisgen reaction (CuAAC). Excess reagents are removed by acetone precipitations. After enzymatic digestion and enrichment on magnetic streptavidin beads, crosslinked peptides were cleaved off under mild conditions and subsequently analysed by means of LC-MS². C) Fragmentation pathways of the crosslinked peptides, forming two peptide pairs with a characteristic Δm_{rep} of 31.9721 u.

tion. This is followed by enzymatic digestion of the proteins. The crosslinked peptides are in this way labelled, for example, with biotin, which allows for their enrichment. Afterwards, they are analysed by means of mass spectrometry. Because the individual peptide masses in a crosslink are not immediately accessible, crosslink assignment is difficult, especially in complex samples. This requires the use of MS-cleavable reagents, which facilitate MS² identification by forming specific fragment ions.^[20,21] At best, the reagent should also enable MS³ experiments, which requires that the crosslinker is cleaved before the peptides. This allows peptides to be separated for individual MS³-based identification. Our reagent **1** contains β -hydrogen atoms on both sides of the sulfoxide, so that fragmentation occurs in two directions, as depicted in Figure 1C. This leads

to a clean separation of the peptides (α , β) and generates two fragments for each peptide: an alkene and a sulfenic acid fragment; the latter forms a thial upon water loss. This provides two mass pairs with a characteristic $\Delta m/z \approx 32$. Importantly, we observed that fragmentation pathway a (Figure 1C) predominated. Therefore, for the α -peptide the alkene fragment, and for the β -peptide the thial fragment, are the main cleavage products. Due to the asymmetric cleavage properties, most of the signal intensity is retained on one fragment per peptide, which should provide excellent sensitivity in MS³ experiments.

The synthesis of reagent **1** is straightforward (Scheme 1). The starting point is the oxidised disulfide dimer of homocysteine **2**, which is treated with 4-pentynoic acid to give disulfide



Scheme 1. Synthesis of cliXlink (1). a) 4-Pentynoic acid, *N*-(3-dimethylamino-propyl)-*N'*-ethylcarbodiimide hydrochloride (EDC-HCl), 1-hydroxybenzotriazole hydrate (HOBT·H₂O), NEt₃, DMF, RT, 15 h, 71 %; b) over two steps: 1) tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl), NaHCO₃, DMF, H₂O, RT, 3 h; 2) methyl 3-bromopropionate, 45 °C, 44 h, 55 %; c) LiOH, THF, H₂O, 0 °C RT, overnight (o/n), 91 %; d) *N*-hydroxysuccinimide (NHS), pyridine, trifluoroacetic anhydride, MeCN, 0 °C RT, o/n, 78 %; e) *meta*-chloroperoxybenzoic acid (mCPBA), AcOEt, RT, 30 min, 57 %.

3. Reductive cleavage of the disulfide and alkylation of the generated thiols with methyl 3-bromopropionate generates sulfide compound **4**. Saponification of both methyl esters to **5** and conversion of **5** into activated bis-succinimidyl ester **6**, followed by oxidation of the thioether to the sulfoxide, gives reagent **1** in a total yield of 16%. Of particular importance is that the reagent is very pure and the reactive ester units are in place at both sides. Partial hydrolysis needs to be avoided. This was ensured by a final precipitation purification. Reagent **1** was dissolved in a mixture of ethyl acetate and dichloromethane and precipitated upon addition of hexanes.

We next investigated the MS properties of **1**. To this end, we added **1** to the commonly used model protein, bovine serum albumin (BSA). We followed the workflow depicted in Figure 1B without performing the enrichment step. In short, after the addition of **1** to the protein solution and reaction for 1 h at room temperature and physiological pH, we precipitated the protein with acetone, resuspended it in buffer (see the

Supporting Information) and digested the protein subsequently with a mixture of trypsin and Lys-C.

The obtained peptide mixture was desalted with C18 Tip columns and analysed by means of HPLC-MS² (Figure 2). Figure 2A shows two crosslinked BSA peptides ($\alpha + \beta$) as an example. After determining the exact mass of the intact crosslink (m/z 677.1; Figure 2B), we performed both CID and HCD fragmentation on the precursor to evaluate the fragmentation properties (Figure 2C). More selective CID fragmentation (resonance excitation of the crosslink ion; Figure 2C, top spectrum) at a low normalised collision energy of 25% provides, as expected, only a small set of intense signals. Two prominent signal pairs, with $\Delta m/z$ of 32 (Δm_{rep}), are obtained due to crosslinker fragmentation, which results in a thial (α/β -thial) and alkene (α/β -alkene) fragment for each peptide. As mentioned before, fragmentation pathway a (Figure 1C) is preferred, leading to an asymmetric intensity distribution. The dominant formation of the expected intact, separated peptides consequently makes the reagent amenable for more sophisticated MS³ experiments.

For our study, we used HCD fragmentation. This method provides simultaneous crosslinker cleavage and formation of the peptide fragments required for identification (Figure 2C, bottom spectrum). To help with crosslink/peptide recognition, it is important that HCD fragmentation still provides the alkene and thial fragments. The HCD data, therefore, allow the identification of peptides by MS².

Using this method, we analysed the data with the freely available MeroX software, strictly filtering all crosslink identifications (score > 50, false discovery rate (FDR) 1%) and requiring the presence of at least three out of four ions from the two characteristic mass pairs.^[23,24] Without exploiting the possibility for CuAAC-based enrichment, we were able to identify 61 unique BSA crosslinks in a single measurement (Figure 2D).^[25] This result is representative for measurements conducted with purified BSA in our laboratory. Additionally, we depicted the crosslinks found in the BSA crystal structure and noted that the majority of measured distances were reasonable. Importantly, the crosslinks show a mean C_{α} - C_{α} distance between the crosslinked amino acids of 22.1 Å and a distance distribution that is in perfect agreement of what is expected for crosslinking with a reagent that spaces the active esters by about 9 Å (Figures 2E and S2)^[26] and taking into account side-chain lengths and protein molecular dynamics.^[27] Crosslinking occurred mostly between two Lys residues (49%), but we also detected Lys-Thr (30%), Lys-Ser (13%) and Lys-Tyr (8%) connections, in agreement with the reactivity of the succinimidyl esters (Figure 2F).^[28]

This success allowed us to validate the new crosslinker in a more complex environment. We particularly wanted to show the additional value of the CuAAC-based enrichment possibility. For this experiment, we again crosslinked the BSA protein (10 µg), precipitated it with acetone, redissolved the crosslinked protein and subsequently performed the click reaction (Figure S3A) by adding **7** (Figure 3A) and CuSO₄/tris(3-hydroxypropyl)triazolylmethylamine (THPTA) followed by reduction of Cu^{II} to Cu^I upon addition of sodium ascorbate. After 1 h at

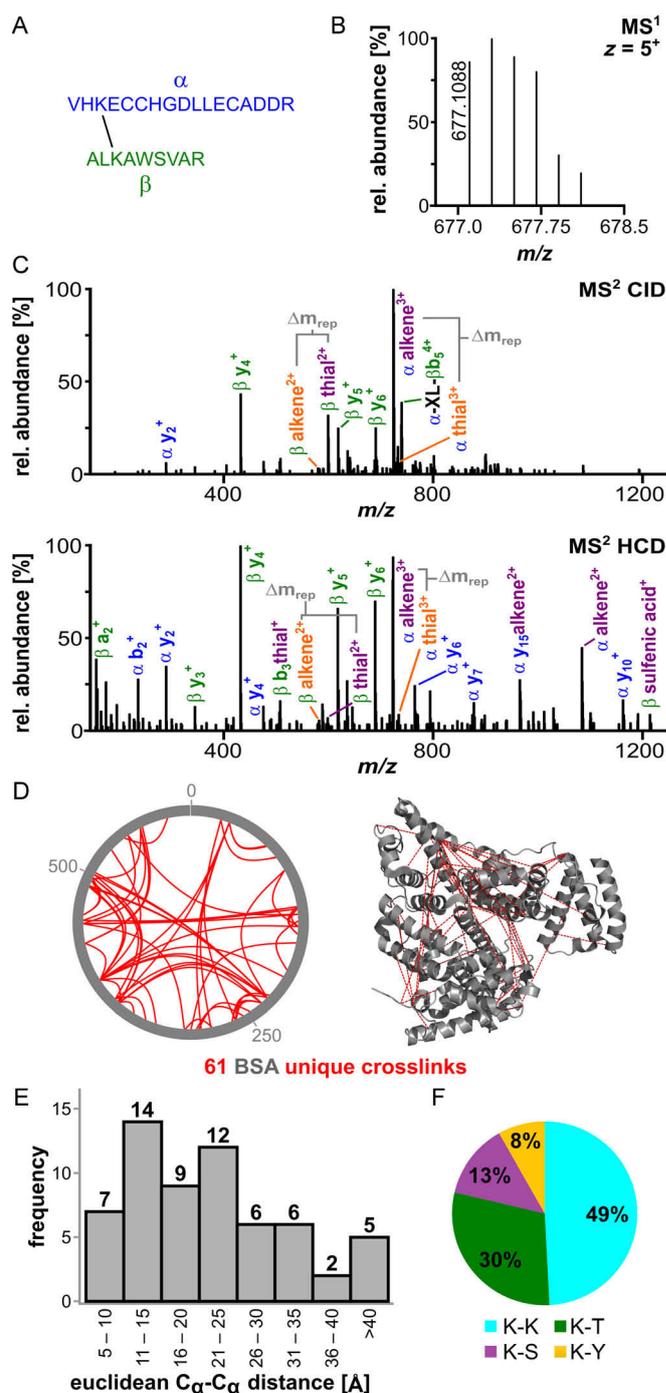


Figure 2. Analysis of crosslinked BSA prior to enrichment. A) Sequences of a representative crosslink within BSA: peptide α : VHKECCHGDLLECADDR (IAA-modified on Cys) and β : ALKAWSVAR. B) The isotope envelope of the 5+ precursor is shown. C) Fragmentation of the crosslinked peptides under CID (25%, top spectrum) and stepped higher energy collisional dissociation (HCD) conditions (25/30/32%, bottom spectrum) showing the product ions of sulfoxide cleavage. The different fragmentation pathways are shown in orange and purple (Figure 1). D) MS² identification of BSA crosslinking sites prior to enrichment. The xVis representation shows the positions of crosslinked amino acids as red lines (left). The corresponding C_{α} - C_{α} distances are depicted in the BSA crystal structure (PDB ID: 4F5S^[22]). E) Histogram of the measured Euclidean C_{α} - C_{α} distances in the crystal structure; the majority are below 40 Å, with a mean measured distance of 22.1 Å. F) Distribution of the identified crosslinking sites between Lys-Lys, Lys-Thr, Lys-Ser and Lys-Tyr.

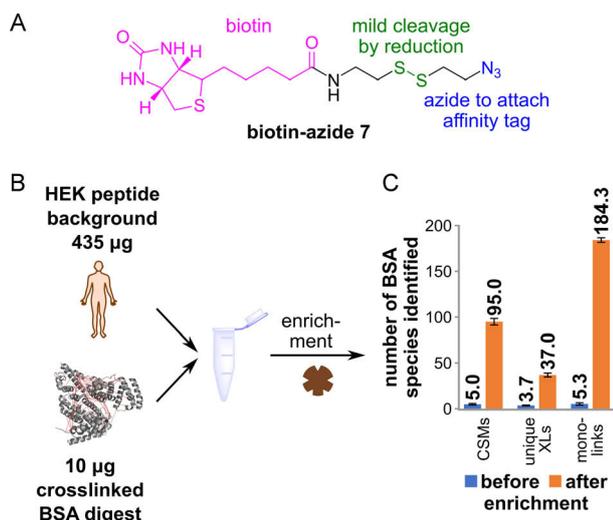


Figure 3. Enrichment of crosslinked BSA from a complex sample. A) Structure of **7**, with the biotin group (pink), the disulfide bond (green) and the azide moiety (blue). Upon reduction of the disulfide, the biotin moiety is removed from the crosslink. B) Depiction of the spike-in experiment. 10 µg of crosslinked and CuAAC-modified BSA digest were added to 435 µg of HEK digest as a complex background. Magnetic streptavidin beads were used to enrich the crosslinked peptides. C) Results of the HPLC-MS² analysis, showing the effect of enrichment. The experiment was performed in technical triplicates; numbers show the mean value and the standard deviation is indicated by the error bar.

room temperature, we performed another acetone precipitation to remove excess biotin azide. We then added trypsin and Lys-C for digestion.

These peptides we next combined with a protein digest obtained from a HEK cell extract (435 µg of protein; Figure 3B). This creates a massive background of non-crosslinked peptides. If we now analysed the crosslinks within this large background (Figure 3C), we identified only a very small number of crosslink spectral matches (mean CSMs=5.0), unique crosslinked sites (mean unique XLS=3.7) as well as monolinks (mean=5.3), in which one succinimidyl ester hydrolysed before reaction with the protein. However, if we performed the enrichment with streptavidin-coated magnetic beads, the number of crosslink identifications improved dramatically. Following incubation of the peptide mixture with the magnetic particles; extensive washing (6 ×) of the beads with buffer (see the Supporting Information) and mild, reductive cleavage of the disulfide to liberate the captured crosslinks and thereby removing the biotin moiety (Figure S3B), we now, on average, detected 95.0 CSMs and 37.0 unique XLS together with 184.3 monolinks. The number of identified unique XLS was lower than that for the purified BSA (Figure 2D), which could be explained by inefficiencies in the CuAAC reaction or interference of the non-crosslinked complex background. However, this result shows that our new crosslinker, **1**, is able to enrich crosslinked peptides in a vast excess of unmodified peptides; thus facilitating the analysis of complex samples with low-abundant crosslinks. Originally, we were sceptical about the asymmetric structure of **1**. The data, however, show that, despite this, a large number of crosslinks are identified.

In summary, our new crosslinker **1** combines the following advantages: firstly, the size of the reagent is ideally suited to gain valuable distance information for structural proteomics. In addition, the molecule is cell permeable and the alkyne unit does not cause major interference with the crosslinking reaction. Furthermore, the robust, efficient sulfoxide fragmentation simplifies crosslink identification. The possibility of functionalising peptides crosslinked with **1** by click chemistry allows diverse modifications and enrichment strategies that enable analysis of low-abundant crosslinks to be employed. To conclude, our reagent **1** now paves the way for complex interactome analysis by using MS² and MS³ experiments.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for financial support through SFB1309 (TP-A4), SFB1032 (TP-A5) and SPP1784 and the Einzelverfahren CA275-11/1. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 765266 (LightDyNAmics). Research was furthermore supported by a European Research Council (ERC) advanced grant under the European Union's Horizon 2020 research and innovation programme (grant agreement no. EPIR 741912) and the Volkswagen foundation (Initiative •Life : EcoRib). M.W. thanks the U.S. Department of Energy (DOE grant DE-SC0018260) and the National Institutes of Health (NIH grant R35GM128813) for support. M.S. and L.S.R. thank the Fonds der Chemischen Industrie for predoctoral fellowships. The mass spectrometry proteomics data have been deposited with the ProteomeXchange Consortium through the PRIDE²⁹ partner repository with the accession no. PXD015080.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: click chemistry • crosslinking mass spectrometry • interactome analysis • proteomics • structure elucidation

- [1] I. N. Berezovsky, E. Guarnera, Z. Zheng, B. Eisenhaber, F. Eisenhaber, *Curr. Opin. Struct. Biol.* **2017**, *42*, 67–74.
- [2] H. Currinn, T. Wassmer, *Biochem. Soc. Trans.* **2016**, *44*, 185–190.
- [3] S. Allard, R. T. Utley, J. Savard, A. Clarke, P. Grant, C. J. Brandl, L. Pillus, J. L. Workman, J. C. t., *EMBO J.* **1999**, *18*, 5108–5119.
- [4] Y. Xue, J. Wong, G. T. Moreno, M. K. Young, J. C. t., W. Wang, *Mol. Cell* **1998**, *2*, 851–861.
- [5] M. Sauvageau, G. Sauvageau, *PLoS Biol.* **2008**, *6*, e113.
- [6] A. Leitner, T. Walzthoeni, A. Kahraman, F. Herzog, O. Rinner, M. Beck, R. Aebersold, *Mol. Cell. Proteomics* **2010**, *9*, 1634–1649.
- [7] A. Sinz, *Angew. Chem. Int. Ed.* **2018**, *57*, 6390–6396; *Angew. Chem.* **2018**, *130*, 6498–6504.
- [8] J. Rappsilber, *J. Struct. Biol.* **2011**, *173*, 530–540.
- [9] J. D. Chavez, J. E. Bruce, *Curr. Opin. Chem. Biol.* **2019**, *48*, 8–18.
- [10] F. J. O'Reilly, J. Rappsilber, *Nat. Struct. Mol. Biol.* **2018**, *25*, 1000–1008.
- [11] A. Leitner, M. Faini, F. Stengel, R. Aebersold, *Trends Biochem. Sci.* **2016**, *41*, 20–32.

- [12] A. Kao, C.-I. Chiu, D. Vellucci, Y. Yang, V. R. Patel, S. Guan, A. Randall, P. Baldi, S. D. Rychnovsky, L. Huang, *Mol. Cell. Proteomics* **2011**, *10*, M110.002212.
- [13] M. Q. Müller, F. Dreiocker, C. H. Ihling, M. Sch fer, A. Sinz, *Anal. Chem.* **2010**, *82*, 6958–6968.
- [14] A. M. Burke, W. Kandur, E. J. Novitsky, R. M. Kaake, C. Yu, A. Kao, D. Vellucci, L. Huang, S. D. Rychnovsky, *Org. Biomol. Chem.* **2015**, *13*, 5030–5037.
- [15] X. Tang, J. E. Bruce, *Mol. Biosyst.* **2010**, *6*, 939–947.
- [16] R. Huang, W. Zhu, Y. Wu, J. Chen, J. Yu, B. Jiang, H. Chen, W. Chen, *Chem. Sci.* **2019**, *10*, 6443–6447.
- [17] M. Rey, M. Dupr , I. Lopez-Neira, M. Duchateau, J. Chamot-Rooke, *Anal. Chem.* **2018**, *90*, 10707–10714.
- [18] B. Steigenberger, R. J. Pieters, A. J. R. Heck, R. A. Scheltema, *ACS Cent. Sci.* **2019**, *5*, 1514–1522.
- [19] M. Ehrlich, M. J. Gattner, B. Viverge, J. Bretzler, D. Eisen, M. Stadlmeier, M. Vrabel, T. Carell, *Chem. Eur. J.* **2015**, *21*, 7701–7704.
- [20] F. Liu, P. Lössl, R. Scheltema, R. Viner, A. J. R. Heck, *Nat. Commun.* **2017**, *8*, 15473.
- [21] O. Klykov, B. Steigenberger, S. Pektaş, D. Fasci, A. J. R. Heck, R. A. Scheltema, *Nat. Protoc.* **2018**, *13*, 2964–2990.
- [22] A. Bujacz, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 1278–1289.
- [23] M. Götze, J. Pettelkau, R. Fritzsche, C. H. Ihling, M. Sch fer, A. Sinz, *J. Am. Soc. Mass. Spectrom.* **2015**, *26*, 83–97.
- [24] M. Götze, C. Iacobucci, C. H. Ihling, A. Sinz, *Anal. Chem.* **2019**, *91*, 10236–10244.
- [25] M. Grimm, T. Zimniak, A. Kahraman, F. Herzog, *Nucleic Acids Res.* **2015**, *43*, W362–W369.
- [26] E. D. Merkley, S. Rysavy, A. Kahraman, R. P. Hafen, V. Daggett, J. N. Adkins, *Protein Sci.* **2014**, *23*, 747–759.
- [27] Z. Orb n-N meth, R. Beveridge, D. M. Hollenstein, E. Rampler, T. Stranzl, O. Hudecz, J. Doblmann, P. Schlögelhofer, K. Mechtler, *Nat. Protoc.* **2018**, *13*, 478–494.
- [28] S. Kalkhof, A. Sinz, *Anal. Bioanal. Chem.* **2008**, *392*, 305–312.
- [29] Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. P rez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, Ş. Y Imaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A. F. Jarnuczak, T. Ternent, A. Brazma, J. A. Vizca no, *Nucleic Acids Res.* **2019**, *47*, D442–D450.

Manuscript received: October 8, 2019

Accepted manuscript online: October 8, 2019

Version of record online: November 28, 2019

CHEM**BIO**CHEM

Supporting Information

A Click-Chemistry-Based Enrichable Crosslinker for Structural and Protein Interaction Analysis by Mass Spectrometry

Michael Stadlmeier^{+, [a, b]} Leander Simon Runtsch^{+, [a]} Philipp Streshnev,^[a] Martin Wühr,^[b] and Thomas Carell^{*[a]}

cbic_201900611_sm_miscellaneous_information.pdf

Author Contributions

M.S. Conceptualization: Equal; Data curation: Lead; Formal analysis: Lead; Investigation: Equal; Methodology: Equal; Software: Equal; Writing - Review & Editing: Equal

L.R. Conceptualization: Supporting; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing - Review & Editing: Supporting

F.S. Investigation: Supporting; Methodology: Supporting

M.W. Conceptualization: Supporting; Formal analysis: Supporting; Investigation: Supporting; Writing - Review & Editing: Supporting

T.C. Formal analysis: Equal; Funding acquisition: Lead; Project administration: Lead; Resources: Lead; Supervision: Lead; Writing - Original Draft: Lead.

Supporting Information

Contents

| | |
|---|----|
| 1. General information..... | 3 |
| 2. Chemical Synthesis..... | 4 |
| 2.1. Synthesis of cliXlink..... | 4 |
| 2.1.1. Methyl S-(((R)-4-methoxy-4-oxo-3-(pent-4-ynamido)butyl)thio)-N-(pent-4-ynoyl)-L-homocysteinate (3)..... | 4 |
| 2.1.2. Methyl S-(3-methoxy-3-oxopropyl)-N-(pent-4-ynoyl)homocysteinate (4)..... | 5 |
| 2.1.3. S-(2-carboxyethyl)-N-(pent-4-ynoyl)homocysteine (5)..... | 6 |
| 2.1.4. N-Succinimidyl S-(3-((N-succinimidyl)oxy)-3-oxopropyl)-N-(pent-4-ynoyl)homocysteinate (6)..... | 6 |
| 2.1.5. Bis-N-hydroxydisuccinimidyl-S-oxo-S-(carboxyethyl)-N'-(pent-4-ynoyl)homocysteinate (cliXlink, 1)..... | 8 |
| 2.2. Synthesis of cleavable biotin-azide 7..... | 9 |
| 2.2.1. 2-((2-azidoethyl)disulfaneyl)ethan-1-amine (9)..... | 9 |
| 2.2.2. N-(2-((2-azidoethyl)disulfaneyl)ethyl)-biotinamide (7)..... | 10 |
| 3. Sample preparation & mass spectrometric analysis..... | 11 |
| 3.1. Cross-linking for BSA analysis..... | 11 |
| 3.2. CuAAC-modification & enrichment..... | 11 |
| 3.3. MS-Analysis..... | 12 |
| 3.4. Data analysis..... | 13 |
| 3.5. Cell permeability..... | 15 |
| 4. Additional Material..... | 16 |
| 5. NMR-Spectra..... | 18 |
| 5.1. Methyl S-(((R)-4-methoxy-4-oxo-3-(pent-4-ynamido)butyl)thio)-N-(pent-4-ynoyl)-L-homocysteinate (3)..... | 18 |
| 5.2. Methyl S-(3-methoxy-3-oxopropyl)-N-(pent-4-ynoyl)homocysteinate (4)..... | 19 |
| 5.3. S-(2-carboxyethyl)-N-(pent-4-ynoyl)homocysteine (5)..... | 20 |

| | | |
|--|---|----|
| 5.4. <i>N</i> -Succinimidyl | <i>S</i> -(3-((<i>N</i> -succinimidyl)oxy)-3-oxopropyl)- <i>N</i> -(pent-4-ynoyl)homocysteinate (6) | 21 |
| 5.5. <i>Bis-N-hydroxydisuccinimidyl-S-oxo-S</i> -(carboxyethyl)- <i>N'</i> -(pent-4-inoyl)homocysteinate (cliXlink, 1) | | 22 |
| 5.6. 2-((2-azidoethyl)disulfaneyl)ethan-1-amine (9) | | 23 |
| 5.7. <i>N</i> -(2-((2-azidoethyl)disulfaneyl)ethyl)-biotinamide (7) | | 24 |
| 6. References | | 25 |

1. General information

Unless noted otherwise, all reactions were performed using oven-dried glassware under an atmosphere of argon. Molsieve-dried solvents were used from *Sigma Aldrich* and chemicals were bought from *Sigma Aldrich*, *TCl*, *Carbolution*, *Roth* and *Carbosynth*. For extraction and chromatography purposes, technical grade solvents were distilled prior to their usage. Reaction controls were performed using TLC-Plates from *Merck* (Merck 60 F²⁵⁴), flash column chromatography purifications were performed on *Merck* Geduran Si 60 (40–63 μM). Visualization of the TLC plates was achieved through UV-absorption or through staining with *Ceric ammonium molybdate stain*. NMR spectra were recorded in deuterated solvents on *Varian VXR400S*, *Varian Inova 400*, *Bruker AMX 600*, *Bruker Ascend 400* or *Bruker Avance III HD*. HR-ESI-MS spectra were obtained from a *Thermo Finnigan LTQ FT-ICR*. IR-measurements were performed on a *Perkin Elmer Spectrum BX FT-IR* spectrometer with a diamond-ATR (*Attenuated Total Reflection*) unit.

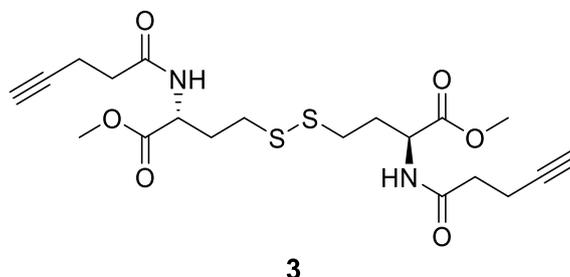
For MS-experiments, only highest grade solvents and reagents were used.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium^[1] via the PRIDE partner repository^[2] with the dataset identifier PXD015080.

2. Chemical Synthesis

2.1. Synthesis of cliXlink

2.1.1. Methyl S-(((R)-4-methoxy-4-oxo-3-(pent-4-ynamido)butyl)thio)-N-(pent-4-ynoyl)-L-homocysteinate (**3**)



To a solution of 4-pentynoic acid (174 mg, 1.77 mmol, 2.6 eq.) in dry DMF (5 mL), triethylamine (416 mg, 570 μ L, 4.08 mmol, 6.0 eq.), 1-HOBt \cdot H₂O (220 mg, 1.63 mmol, 2.4 eq.), EDC \cdot HCl (312 mg, 1.63 mmol, 2.4 eq.) and DL-homocystin bis(methylester) dihydrochloride (**2**) (250 mg, 677 μ mol, 1.0 eq.) were added and the mixture was stirred at room temperature for 15 h. Water (20 mL) was added and the suspension was acidified with HCl_(aq) (2 M) to a pH of \approx 2. The resulting suspension was extracted with AcOEt (2x30 mL). Combined organics were dried over MgSO₄, filtered and the solvent was removed under reduced pressure. The crude product was purified by flash column chromatography (silica gel, hexanes:AcOEt 1:2) to yield **3** (221 mg, 480 μ mol, 71%) as colorless oil.

¹H-NMR (599 MHz, CDCl₃): δ /ppm = 2.03 (t, ⁴J = 2.6 Hz, 2H, -C \equiv CH), 2.03 – 2.12 (m, 2H, -S-CH₂-CH₂-), 2.21 – 2.29 (m, 2H, -S-CH₂-CH₂-), 2.43 – 2.60 (m, 8H, HC \equiv C-CH₂CH₂-), 2.73 (m, 4H, -S-CH₂-CH₂-), 3.76 (s, 3H, -O-CH₃), 3.77 (s, 3H, -O-CH₃), 4.74 (m, 2H, -CH-), 6.49 (d, ³J = 7.8 Hz, 1H, -NH-), 6.49 (d, ³J = 7.9 Hz, 1H, -NH-).

Because of the mixture of diastereomeres, twice the number of signals was observed in the ¹³C-NMR-spectrum. The star (*) indicates the signals of the second stereoisomer.

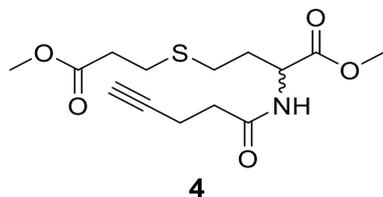
¹³C-NMR (151 MHz, CDCl₃): δ /ppm = 15.0 (HC \equiv C-CH₂-CH₂-), 15.0* (HC \equiv C-CH₂-CH₂-), 32.4 (-S-CH₂-CH₂-), 32.5* (-S-CH₂-CH₂-), 34.7 (-S-CH₂-CH₂-), 35.1* (-S-CH₂-CH₂-), 35.3 (HC \equiv C-CH₂-CH₂-), 35.3* (HC \equiv C-CH₂-CH₂-), 51.4 (-CH-), 51.4* (-CH-), 52.8 (-O-CH₃), 52.8* (-O-CH₃), 69.7 (HC \equiv C-), 69.7* (HC \equiv C-), 82.9 (HC \equiv C-), 82.9* (HC \equiv C-), 171.1 (-CONH-), 171.2* (-CONH-), 172.4 (-COMe), 172.4* (-COMe).

IR (ATR): $\tilde{\nu}$ /cm⁻¹ = 3286 (w), 2854 (m), 2002 (w), 1742 (s), 1647 (m), 1514 (m), 1437 (m) 1375 (m), 1286 (m), 1220 (s), 1178 (s), 1070 (m), 1025 (w), 982 (m), 942 (w), 905 (w), 797 (w), 750 (m), 634 (m).

HRMS (ESI): m/z : calculated for $C_{20}H_{29}N_2O_6S_2^+$ ($[M+H]^+$): 457.1462; found: 457.1458 ($\Delta m = 0.87$ ppm).

Retardation factor: R_F (hexanes:AcOEt 1:2) = 0.27.

2.1.2. Methyl *S*-(3-methoxy-3-oxopropyl)-*N*-(pent-4-ynoyl)homocysteinate (**4**)



To a solution of disulfide **3** (1.50 g, 3.29 mmol, 1.0 eq.) in a mixture of DMF and Wasser (4:1, 25 mL), TCEP·HCl (940 mg, 3.29 mmol, 1.0 eq.) and $NaHCO_3$ (1.16 g, 13.8 mmol, 4.2 eq.) were added. The reaction mixture was stirred for 3 h at room temperature and additional TCEP·HCl (284 mg, 0.991 mmol, 0.3 eq.) as well as methyl 3-bromopropionate (1.37 g, 895 μ L, 8.21 mmol, 2.5 eq.) were added. After stirring at 45 °C for 44 h, water (100 mL) was added and the mixture was extracted with AcOEt (2x100 mL). Combined organics were dried over $MgSO_4$, filtered and the solvent was removed under reduced pressure. Purification by flash column chromatography (silica gel, hexanes:AcOEt 3:1→1:1) yielded **4** (1.13 g, 61 mmol, 55%) as colorless oil.

1H -NMR (600 MHz, $CDCl_3$): $\delta/ppm = 1.93 - 2.01$ (m, 1H, -CH- CH_2 -), 2.02 (t, $^4J = 2.6$ Hz, 1H, $HC\equiv C$ -), 2.17 (m, 1H, -CH- CH_2 -), 2.40 – 2.62 (m, 8H, $HC\equiv C$ - CH_2CH_2 -, -S- CH_2 - CH_2 -COOMe, -CH- CH_2 - CH_2 -S-), 2.76 (t, $^3J = 7.1$ Hz, 2H, -S- CH_2 - CH_2 -COOMe), 3.69 (s, 3H, - OCH_3), 3.75 (s, 3H, - OCH_3), 4.72 (td, $^3J = 7.5$ Hz, 4.9 Hz, 1H, -CH-), 6.36 (d, $^3J = 7.8$ Hz, 1H, -NH-).

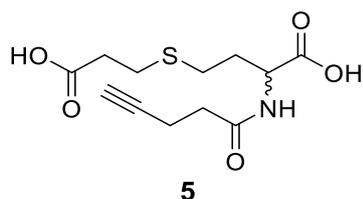
^{13}C -NMR (151 MHz, $CDCl_3$): $\delta/ppm = 14.9$ ($HC\equiv C$ - CH_2 - CH_2 -), 27.0 (-S- CH_2 - CH_2 -COOMe), 28.0 ($HC\equiv C$ - CH_2 - CH_2 -), 32.3 (-CH- CH_2 -), 34.7 (-S- CH_2 - CH_2 -COOMe), 35.3 (-CH- CH_2 - CH_2 -S-), 51.7 (-CH-), 52.0, (- OCH_3), 52.7 (- OCH_3), 69.6 ($HC\equiv C$ -), 82.9 ($HC\equiv C$ -), 170.9 (-CH-COOMe), 172.4 (-COX), 172.5 (-COX).

IR (ATR): $\tilde{\nu}/cm^{-1} = 3288$ (w), 2953 (w), 1737 (s), 1655 (m), 1537 (s), 1437 (m), 1361 (m), 1215 (m).

HRMS (ESI): m/z : calculated for $C_{14}H_{22}NO_5S^+$ ($[M+H]^+$): 316.1213; found: 316.1216 ($\Delta m = 0.95$ ppm).

Retardation factor: R_F (hexanes:AcOEt 1:1) = 0.36.

2.1.3. *S*-(2-carboxyethyl)-*N*-(pent-4-ynoyl)homocysteine (**5**)



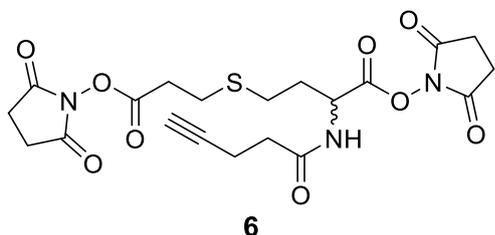
To a solution of diester **4** (156 mg, 0.495 mmol, 1.0 eq.) in a mixture of THF (3 mL) and water (3 mL) at 0 °C, LiOH (60 mg, 2.5 mmol, 5.0 eq.) was added and the reaction mixture was stirred for 30 min at 0 °C. Afterwards, the reaction mixture was allowed to warm to room temperature and stirring continued overnight. Water (15 mL) was added, the solution was acidified with HCl_(aq) (2M) to a pH of 1 and extracted with AcOEt (5x15 mL). Combined organics were dried over MgSO₄, filtered and the solvent was removed under reduced pressure. Diacid **5** was obtained as colorless oil (128 mg, 0.445 mmol, 91%) and used without further purification.

¹H-NMR (400 MHz, CD₃OD): δ/ppm = 1.89 – 2.03 (m, 1H, -CH-CH₂-), 2.13 (m, 1H, -CH-CH₂-), 2.25 – 2.28 (m, 1H, HC≡C-), 2.43 – 2.51 (m, 4H, HC≡C-CH₂CH₂-), 2.54 – 2.70 (m, 4H, -CH-CH₂-CH₂-S-, -S-CH₂-CH₂-COOH), 2.76 (m, 2H, -S-CH₂-CH₂-COOH), 4.53 (m, 1H, -CH-).

¹³C-NMR (101 MHz, CD₃OD): δ/ppm = 14.2 (HC≡C-CH₂-CH₂-), 26.4 (HC≡C-CH₂-CH₂-), 27.7 (-S-CH₂-CH₂-COOH), 31.6 (-CH-CH₂-), 34.4 (-S-CH₂-CH₂-COOH), 34.5 (HC≡C-CH₂-CH₂-), 51.7(-CH-), 68.9 (HC≡C-), 82.2 (HC≡C-), 172.7 (-COX), 174.2 (-COX), 174.5 (-COX).

HRMS (ESI): *m/z*: calculated for: C₁₂H₁₆NO₅S⁻ ([M-H]⁻): 286.0755; found: 286.0756 (Δ*m* = 0.36 ppm).

2.1.4. *N*-Succinimidyl *S*-(3-((*N*-succinimidyl)oxy)-3-oxopropyl)-*N*-(pent-4-ynoyl)homocysteinate (**6**)



To a solution of diacid **5** (120 mg, 0.418 mmol, 1.0 eq.), *N*-hydroxysuccinimide (192 mg, 1.67 mmol, 4.0 eq.) and pyridine (264 mg, 270 μL, 3.34 mmol, 8.0 eq.) in dry MeCN (15 mL) at 0 °C, trifluoroacetic anhydride (351 mg, 232 μL, 1.67 mmol, 4.0 eq.) was added dropwise. The reaction mixture was allowed to warm to room temperature and stirred overnight. Afterwards, the solvent was removed under reduced pressure and the residue was taken up

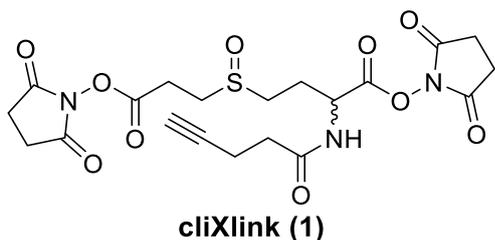
in CHCl_3 (20 mL), washed with basic water (pH = 9 by NaOH, 1x30 mL), acidic water (pH = 3 by $\text{HCl}_{(\text{aq})}$, 1x30 mL) and brine (1x30 mL). The organic layer was dried over MgSO_4 , filtered and the solvent was reduced under reduced pressure. Without further purification, di-NHS-ester **6** (158 mg, 0.328 mmol, 78%) was obtained in form of a colorless foam.

$^1\text{H-NMR}$ (400 MHz, CDCl_3): δ/ppm = 2.04 (t, 4J = 2.4 Hz, 1H, $-\text{C}\equiv\text{CH}$), 2.16 – 2.36 (m, 2H, $-\text{S}-\text{CH}_2-\text{CH}_2-\text{CH}-$), 2.45 – 2.57 (m, 4H, $\text{HC}\equiv\text{C}-\text{CH}_2\text{CH}_2-$), 2.71 – 2.79 (m, 2H, $-\text{S}-\text{CH}_2-\text{CH}_2-\text{CH}-$), 2.82 – 2.99 (m, 12H, NHS-Hs, $-\text{S}-\text{CH}_2-\text{CH}_2-\text{CONHS}$), 5.17 (m, 1H, $-\text{CH}-$), 6.51 (d, 3J = 8.0 Hz, 1H, $-\text{NH}-$).

$^{13}\text{C-NMR}$ (101 MHz, CD_3OD): δ/ppm = 14.8 ($-\text{CH}_2-$), 25.7 ($-\text{CH}_2-$), 26.5 ($-\text{CH}_2-$), 27.4 ($-\text{CH}_2-$), 32.2 ($-\text{CH}_2-$), 32.3 ($-\text{CH}_2-$), 35.1 ($-\text{CH}_2-$), 50.1 ($-\text{CH}-$), 69.8 ($\text{HC}\equiv\text{C}-$), 82.9 ($\text{HC}\equiv\text{C}-$), 167.4 ($-\text{COX}$), 167.7 ($-\text{COX}$), 169.2 ($-\text{COX}$), 170.9 ($-\text{COX}$).

HRMS (ESI): m/z : calculated for $\text{C}_{20}\text{H}_{24}\text{N}_3\text{O}_9\text{S}^+$ ($[\text{M}+\text{H}]^+$): 482.1228; found: 482.1233 (Δm = 1.04 ppm); calculated for $\text{C}_{20}\text{H}_{27}\text{N}_4\text{O}_9\text{S}^+$ ($[\text{M}+\text{NH}_4]^+$): 499.1493; found: 499.1496 (Δm = 0.60 ppm).

2.1.5. *Bis-N-hydroxydisuccinimidyl-S-oxo-S-(carboxyethyl)-N'-(pent-4-inoyl)homocysteinate* (cliXlink, **1**)



To a solution of di-NHS-Ester **6** (169 mg, 0.351 mmol, 1.0 eq.) in AcOEt (2 mL), *m*CPBA (77%, 79 mg, 0.35 mmol, 1.0 eq.) was added. The resulting suspension was stirred for 30 min at room temperature. DCM (4 mL) was added and the solution was mixed with hexanes (8 mL), leading to the formation of a colorless precipitate. After filtration, the filtrate was again mixed with hexanes (5 mL) and the precipitate was filtered off. The combined precipitates were washed with DCM (5 mL), ice-cold AcOEt (10 mL) and ice-cold acetone (5 mL) and dried under reduced pressure. By this, crosslinker **cliXlink (1)** (99 mg, 0.20 mmol, 57%) was obtained as a colorless foam.

The isolated mixture of diastereomeres was used without further purification for the MS-experiments.

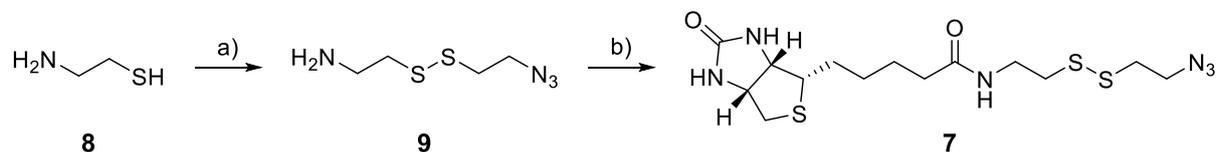
¹H-NMR (599 MHz, CDCl₃): δ /ppm = 2.04 (m, 1H, HC≡C-), 2.44 – 2.59 (m, 6H, -CH-CH₂-CH₂-SO-, HC≡C-CH₂-CH₂-), 2.86 (s, 8H, NHS-Hs), 2.93 – 3.27 (m, 6H, -SO-CH₂-CH₂-CONHS, -CH-CH₂-CH₂-SO-), 5.10 – 5.22 (m, 1H, -CH-), 7.02 (d, ³*J* = 7.6 Hz, 0.5H, -NH-), 7.30 (d, ³*J* = 7.9 Hz, 0.5H, -NH-).

Because of the mixture of diastereomeres, multiple, partly overlapping signals were observed in the ¹³C-NMR-spectrum. The star (*) indicates the signals of the second stereoisomer.

¹³C-NMR (151 MHz, CDCl₃): δ /ppm = 14.6 (HC≡C-CH₂-CH₂-), 14.7* (HC≡C-CH₂-CH₂-), 23.5 (-SO-CH₂-CH₂-CONHS), 24.1* (-SO-CH₂-CH₂-CONHS), 25.6 (NHS-CH₂-), 25.8 (HC≡C-CH₂-CH₂-), 26.4* (HC≡C-CH₂-CH₂-), 34.8 (-CH-CH₂-CH₂-SO-), 34.9* (-CH-CH₂-CH₂-SO-), 45.5 (-SO-CH₂-), 45.6* (-SO-CH₂-), 47.2 (-SO-CH₂-), 49.3 (-CH-), 49.4* (-CH-), 69.5 (HC≡C-), 69.7* (HC≡C-), 82.7 (HC≡C-), 82.8* (HC≡C-), 167.0 (-CONHS), 167.0* (-CONHS), 167.1 (-CONHS), 168.8 (-CO-N-CO-), 171.1 (-CONH-), 171.2* (-CONH-).

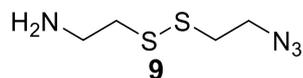
HRMS (ESI): m/z : calculated for $C_{20}H_{24}N_3O_{10}S^+$ ($[M+H]^+$): 498.1177; found: 498.1188 ($\Delta m = 2.21$ ppm); calculated for $C_{20}H_{27}N_4O_{10}S^+$ ($[M+NH_4]^+$): 515.1442; found: 515.1453 ($\Delta m = 2.14$ ppm).

2.2. Synthesis of cleavable biotin-azide **7**



Supplementary Scheme 1: Synthesis of biotin-azide **7**. a) $(N_3CH_2CH_2S)_2$, MeOH, rt, 3 h, 76%. b) Biotin-pentafluorophenyl ester, NEt_3 , DMF, rt, 3 h, 77%.

2.2.1. 2-((2-azidoethyl)disulfaneyl)ethan-1-amine (**9**)



To a solution of cysteamine (**8**, 0.38 g, 4.9 mmol, 1.0 eq) in dry MeOH (30 mL), a solution of 1,2-bis(2-azidoethyl)disulfane (1.0 g, 4.9 mmol, 1.0 eq.) in dry MeOH (20 mL) was added dropwise at room temperature. The reaction mixture was stirred for 3 h at room temperature. Then, the solvent was removed under reduced pressure and the crude mixture was purified by flash column chromatography (silica gel, DCM:MeOH 20:1). By this, azide **9** (0.66 g, 3.7 mmol, 76%) was obtained as colorless oil.

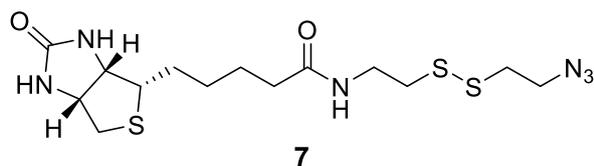
1H -NMR (400 MHz, $DMSO-d_6$): δ /ppm = 1.75 (s, 2H, $-NH_2$), 2.71 – 2.82 (m, 4H), 2.91 (t, $^3J = 6.5$ Hz, 2H), 3.61 (t, $^3J = 6.5$ Hz, 2H).

Because of cis-trans-isomerism of the disulfide bond, twice the number of signals was observed in the ^{13}C -NMR-spectrum. The star (*) indicates the signals of the second stereoisomer.

^{13}C -NMR (101 MHz, $DMSO-d_6$): δ /ppm = 36.8, 36.9*, 40.9, 41.0*, 41.9, 42.0*, 49.1, 49.2*.

IR (ATR): $\tilde{\nu}/cm^{-1}$ = 3360 (w), 2919 (w), 2870 (w), 2097 (s), 1989 (w), 1679 (w), 1589 (w), 1448 (w), 1418 (w), 1391 (w), 1347 (w), 1255 (m), 1227 (m), 1196 (m), 1118 (w), 1071 (w), 1030 (m), 893 (m), 858 (w), 809 (w), 733 (m), 700 (w).

2.2.2. N-(2-((2-azidoethyl)disulfaneyl)ethyl)-biotinamide (**7**)



To a solution of biotin-pentafluorophenyl ester (500 mg, 1.22 mmol, 1.0 eq.) in dry DMF (10 mL), azide **9** (326 mg, 1.83 mmol, 1.5 eq.) was added, followed by triethylamine (0.44 g, 0.60 mL, 4.4 mmol, 3.6 eq.). The reaction mixture was stirred at room temperature for 3 h. Afterwards, the solvent was removed under reduced pressure and the crude product was triturated with Et₂O (20 mL). The resulting solid was filtered, washed with Et₂O (10 mL) and then purified by flash column chromatography (silica gel, DCM:MeOH 20:1), yielding biotin-azide **7** (380 mg, 0.939 mmol, 77%) as colorless solid.

¹H-NMR (400 MHz, CD₃OD): δ /ppm = 1.39 – 1.51 (m, 2H), 1.54 – 1.80 (m, 4H), 2.22 (t, J = 7.6 Hz, 2H), 2.71 (d, J = 12.7 Hz, 1H), 2.83 – 2.96 (m, 5H), 3.21 (ddd, J = 8.9, 5.9, 4.4 Hz, 1H), 3.49 (td, J = 6.6, 1.2 Hz, 2H), 3.60 (t, J = 6.5 Hz, 2H), 4.31 (dd, J = 7.9, 4.4 Hz, 1H), 4.49 (dd, J = 8.0, 5.0 Hz, 1H).

¹³C-NMR (101 MHz, CD₃OD): δ /ppm = 26.8, 29.5, 29.8, 36.7, 38.5, 38.6, 39.4, 41.1, 51.0, 57.0, 61.6, 63.4, 166.1, 176.3.

IR (ATR): $\tilde{\nu}$ /cm⁻¹ = 3304 (w), 3224 (w), 3087 (w), 2928 (w), 2870 (w), 2239 (w), 2174 (w), 2097 (m), 2043 (w), 1698 (s), 1636 (m), 1540 (m), 1478 (w), 1459 (m), 1435 (w), 1364 (w), 1337 (w), 1315 (w), 1293 (w), 1272 (m), 1241 (m), 1228 (m), 1207 (m), 1174 (w), 1152 (w), 1120 (w), 1076 (w), 1043 (w), 978 (w), 890 (w), 856 (w), 847 (w), 803 (w).

HRMS (ESI): m/z : calculated for: C₁₄H₂₅N₆O₂S₃⁺ ([M+H]⁺): 405.1196; found: 405.1218 (Δm = 5.43 ppm).

3. Sample preparation & mass spectrometric analysis

3.1. Cross-linking for BSA analysis

BSA was diluted to a concentration of 1 $\mu\text{g}/\mu\text{L}$ in 10 mM HEPES (pH 7.4). A 20 mM solution of cross-linker (cliXlink, **1**) in dry DMSO was prepared and added to the samples for a final concentration of 1 mM following incubation at room temperature for 45 min. To quench the reaction, an aqueous solution of 1 M NH_4HCO_3 was added (final concentration 20 mM, incubation at room temperature for 15 min). In the next step, an aqueous solution of 500 mM DTT was added (final concentration 5 mM, incubation at 60 °C for 30 min) to reduce disulfide bonds. The resulting free thiols were alkylated with an aqueous solution of 1 M iodoacetamide (final concentration 20 mM, incubation at room temperature for 30 min in the dark). Cold acetone (HPLC grade) was added to a concentration of 80% (v/v) and proteins were precipitated at -20 °C overnight. The samples were centrifuged at 15000 \times g and 4 °C for 10 min, the supernatant was discarded and the pellet was air-dried to remove leftover acetone. Proteins were resuspended to a concentration of 13.3 $\mu\text{g}/\mu\text{L}$ with resuspension buffer (6 M guanidinium chloride, 10 mM EPPS, pH 8.5) and incubated at 37 °C for 15 min following addition of dilution buffer (10 mM EPPS, pH 8.5) to achieve a final guanidinium chloride concentration of 0.45 M and a protein concentration of 1 $\mu\text{g}/\mu\text{L}$. Samples were digested by adding Trypsin/Lys-C (*Promega*, dissolved in 50 mM aqueous AcOH) 1:40 (w/w) relative to protein amount and incubating at 37 °C overnight. The resulting peptide mixture was acidified with 5% aqueous TFA, purified with C_{18} -ZipTips (0.6 μL resin, *Merck*) and the solvent was evaporated. Before MS analysis, the samples were resuspended in 2% MeCN, 0.1% TFA in water to a concentration of \approx 1 $\mu\text{g}/\mu\text{L}$. 2 μL were injected for each measurement.

3.2. CuAAC-modification & enrichment

Samples that were modified and enriched were for the first part prepared the same way as the cross-linking-only samples. After acetone precipitation however, these samples were resuspended to a concentration of 6 $\mu\text{g}/\mu\text{L}$ with resuspension buffer (6 M guanidinium chloride, 10 mM EPPS, pH 8.5) and incubated at 37 °C for 15 min following addition of dilution buffer (10 mM EPPS, pH 8.5) to achieve a final guanidinium chloride concentration of 1 M and a protein concentration of 1 $\mu\text{g}/\mu\text{L}$ after addition of the CuAAC reagents. For the CuAAC-reaction, first, 10 mM biotin-azide **7** (dissolved in DMSO) was added to a final concentration of 500 μM . Secondly, 50 mM CuSO_4 and 250 mM THPTA (both dissolved in H_2O) were pre-mixed in a ratio of 1:1 (v/v) and added to the reaction mixture to a final concentration of 250 μM and 1.25 mM, respectively. Finally, 500 mM sodium ascorbate was added to a final concentration of 5 mM and the samples were incubated at room temperature for 1 h. Subsequently, a second acetone precipitation was carried out analogous to the first one and the samples were digested and purified as described for the cross-linking-only samples.

10 µg of digested and purified BSA (cross-linked and clicked) were resuspended in 10 µL PBS before addition of 435 µg HEK peptide background (dissolved in PBS) resulting in a concentration of 0.33 fmol BSA per 1 µg HEK peptides. 100 µL magnetic streptavidin-coated particles (Dynabeads™ MyOne™ Streptavidin T1, *Thermo Fisher Scientific*) were washed three times with 500 µL PBS, resuspended in 50 µL PBS and added to the samples. The total volume was adjusted to 500 µL and the samples were incubated at room temperature in an orbital shaker for 1 h. The supernatant was discarded and the beads were washed three times with PBS and three times with 50 mM potassium phosphate buffer (pH 7.4) to remove non-specifically binding peptides. After resuspending the beads in 100 µL 50 mM potassium phosphate buffer (pH 7.4), samples were treated with an aqueous solution of TCEP (final concentration 10 mM, incubation at room temperature for 30 min) to cleave off the cross-linked peptides from the beads. The supernatant was transferred to a new vial while the beads were washed once with 50 mM potassium phosphate buffer (pH 7.4) and the supernatant was again transferred to the same vial. To ensure complete alkylation in the next step, the sample pH was checked and adjusted to 7-8 if necessary by adding more buffer. Alkylation of the reduced thiols was carried out using an aqueous solution of iodoacetamide (final concentration 40 mM, incubation at room temperature for 45 min). The reaction was quenched with an aqueous solution of DTT (final concentration 20 mM, incubation at room temperature for 15 min) and the samples were concentrated almost to dryness before being desalted and concentrated using *C₁₈-ZipTips*. For MS analysis, the samples were resuspended in 6 µL 2% MeCN, 0.1% TFA in water and 5.8 µL were injected per measurement. For the control without enrichment, 0.23 µg BSA was spiked into 10 µg HEK peptide background and the volume was adjusted to 15 µL. The samples were reduced, alkylated, quenched and desalted as described above. 2 µL were injected per measurement.

3.3. MS-Analysis

BSA cross-link-samples before and after enrichment were analysed on a *Q Exactive HF* mass spectrometer (*Thermo Fisher Scientific*), coupled to an *Ultimate 3000* nano-HPLC-system (*Thermo Fisher Scientific*). Peptides were first loaded with 0.1% formic acid_(aq) onto a *Acclaim PepMap 100* µ-precolumm cartridge (5 µm, 100 Å, 300 µm ID × 5 mm, *Thermo Fisher Scientific*). For separation of the peptides, a *PicoTip* emitter (noncoated, 15 cm, 75 µm ID, 8 µm tip, *New Objective*), packed in house with *Reprosil-Pur 120 C18-AQ* material (1.9 µm, 120 Å, *Dr. A. Maisch GmbH*), was utilised. The flow-rate was set to 300 nL/min and the 60 min gradient was programmed with percentages B as follows:

1% for 5 min; 1% – 6% in 1 min; 6% – 28% in 38 min; 28% – 85% in 3 min; 85% for 3 min; 85% – 1% in 3 min; 1% for 7 min.

As solvent A, MS-grade water (*Sigma-Aldrich*) with 0.1% (v/v) formic acid (*Thermo-Fisher Scientific*) was used, for solvent B, MS-grade MeCN with 0.1% (v/v) (*Carl Roth*) was used.

The Pico-Tip emitter was positioned in front of the MS-inlet with a modified *Nanospray Flex ion source* (*Thermo Fisher Scientific*), a voltage of 2.2 kV in respect to the mass spectrometer was applied through a T-junction. A column oven (*Sonation GmbH*) was used to heat the column to a constant temperature of 30 °C.

The Q Exactive HF was operated in dd-MS² mode with the following settings:

Polarity: positive; MS¹ resolution: 120k; MS¹ AGC target: 3e6 charges; MS¹ maximum IT: 20 ms; MS¹ scan range: *m/z* 300 – 1750; MS² resolution: 15k; MS² AGC target: 2e5 charges; MS² maximum IT: 125 ms; Top N: 15; isolation window: *m/z* 1.6; isolation offset: *m/z* 0.2; HCD stepped normalised collision energy: 25%, 30%; intensity threshold: 8e5 counts; charge exclusion: unassigned, 1, 2, 7, 8, >8; peptide match: off; exclude isotopes: on; dynamic exclusion: 10 s.

For comparison of CID and HCD fragmentation, an *Orbitrap Fusion Lumos* mass spectrometer (*Thermo Fisher Scientific*) was utilized with the following MS²-scan parameters:

CID-fragmentation:

isolation mode: quadrupole; isolation window: *m/z* 1.6; CID collision energy: 25%; activation Q: 0.25; activation time: 10 ms; MS² detector: orbitrap; MS² resolution: 15k; MS² AGC target: 5e4 charges; MS² maximum IT: 50 ms.

HCD-fragmentation:

isolation mode: quadrupole; isolation window: *m/z* 1.4; HCD stepped normalised collision energy: 25%, 30%, 32%; MS² detector: orbitrap; MS² resolution: 15k; MS² AGC target: 2e5 charges; MS² maximum IT: 125 ms.

3.4. Data analysis

The acquired .raw data files of the BSA-cross-linked samples were converted to .mgf-files with MSConvert by the ProteoWizard software suite (version 3.0.9576).^[3]

For analysis of the .mgf-files, MeroX 2.0.0.6 was used.^[4-5] A .FASTA-file was supplied which consisted of the sequence of bovine serum albumin (BSA; Uniprot: P02769) in addition to 38 possible interferences from the dust/contact protein section of the cRAP-database (<https://www.thegpm.org/crap/>) to increase the size of the utilised .FASTA-file to aid with FDR verification.

The following cross-linkers were defined:

cliXlink:

Supplementary Table 1: MeroX 2.0.0.6 settings for unmodified cliXlink-cross-linked peptides.

| Setting | Value |
|---|--|
| Composition | C ₁₂ H ₁₃ NO ₄ S |
| Composition Mass | 267.05652849 |
| Specificity Site 1 | K{ |
| Specificity Site 2 | KSTY{ |
| Ignore consecutive peptides as cross-links | True |
| Specific Precursor Loss | False |
| Dead-end cross-links (type 0) | True; reacted with "H ₂ O" |
| Intrapeptidal cross-links (type 1) | False |
| Modification of Fragments at Site 1 | SO1, C ₃ H ₄ O ₂ S, 103.9932; thial1, C ₃ H ₂ OS, 85.9826, essential; alk1, C ₃ H ₂ O, 54.0106, essential |
| Modification of Fragments at Site 2 | SO2, C ₉ H ₁₁ NO ₃ S, 213.0460; thial2, C ₉ H ₉ NO ₂ S, 195.0354, essential; alk2, C ₉ H ₉ NO ₂ , 163.0633, essential |

cliXlink after CuAAC, reduction & alkylation:

Supplementary Table 2: MeroX 2.0.0.6 settings for modified cliXlink-cross-linked peptides.

| Setting | Value |
|---|---|
| Composition | C ₁₆ H ₂₁ N ₅ O ₅ S ₂ |
| Composition Mass | 427.09841003 |
| Specificity Site 1 | K{ |
| Specificity Site 2 | KSTY{ |
| Ignore consecutive peptides as cross-links | True |
| Specific Precursor Loss | False |
| Dead-end cross-links (type 0) | True; reacted with "H ₂ O" |
| Intrapeptidal cross-links (type 1) | False |
| Modification of Fragments at Site 1 | SO1, C ₃ H ₄ O ₂ S, 103.9932; thial1, C ₃ H ₂ OS, 85.9826, essential; alk1, C ₃ H ₂ O, 54.0106, essential |
| Modification of Fragments at Site 2 | SO2, C ₁₃ H ₁₉ N ₅ O ₄ S ₂ , 373.0878; thial2, C ₁₃ H ₁₇ N ₅ O ₃ S ₂ , 355.0773 essential; alk2, C ₁₃ H ₁₇ N ₅ O ₃ S, 323.1052, essential |

For all MS²-data analyses, the following settings were used:

K blocked as XL-site; peptide length 5 – 30; missed cleavages: 3; static modifications: carbamidomethyl (C); dynamic modifications: oxidation (M); precursor precision (MS¹): 6 ppm; fragment ion precision (MS²): 15 ppm; mass recalibration (MS¹ and MS²): 0 ppm; mass limits: 1000 – 6000 Da; S/N ratio: 1.5; ion types: a, b, y; minimum number of fragments/peptide: 3; deisotoping: true; correct precursor masses: true, max. 3 isotope shifts; RISEUP-mode; max. missing ions: 1; max. 1 neutral loss of identified fragments; apply prescore: 5% intensity; slow, precise scoring; decoy database generation: shuffle sequences, but keep protease sites.

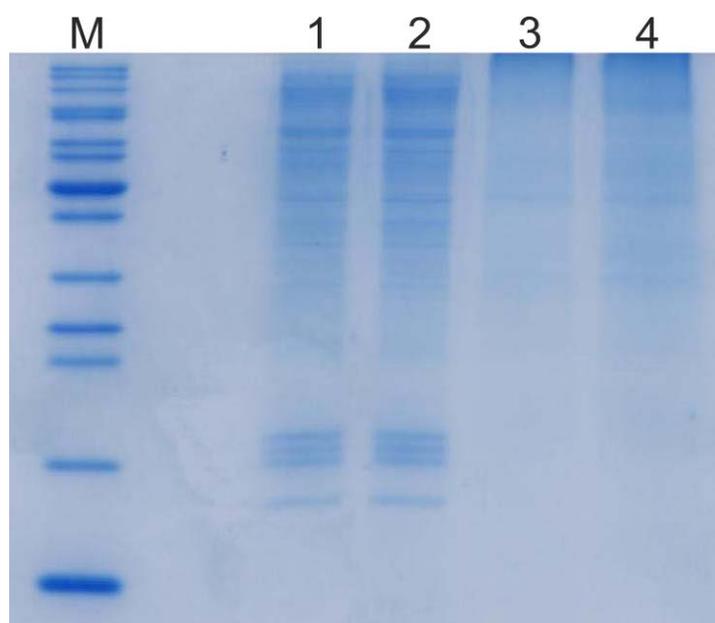
We only considered cross-link spectral matches (CSMs) and unique XLs which were FDR-controlled to 1% and had at least a MeroX-Score of 50.

For further analysis, we used the xVis server of the Herzog lab (LMU Munich) for generating the xVis-representation.^[6] Euclidean C_α-C_α-distances of the identified cross-links in the crystal structure of BSA (pdb: 4f5s)^[7] were measured with xWalk.^[8]

For the data analysis in Fig. 3C, for the sample before enrichment, cliXlink was used unmodified. Only after enrichment, modified cliXlink was utilised. The datasets were analysed with the respective settings. The measurements were carried out from technical triplicates of both the samples prior to and after enrichment. Standard deviation is indicated by the error bar.

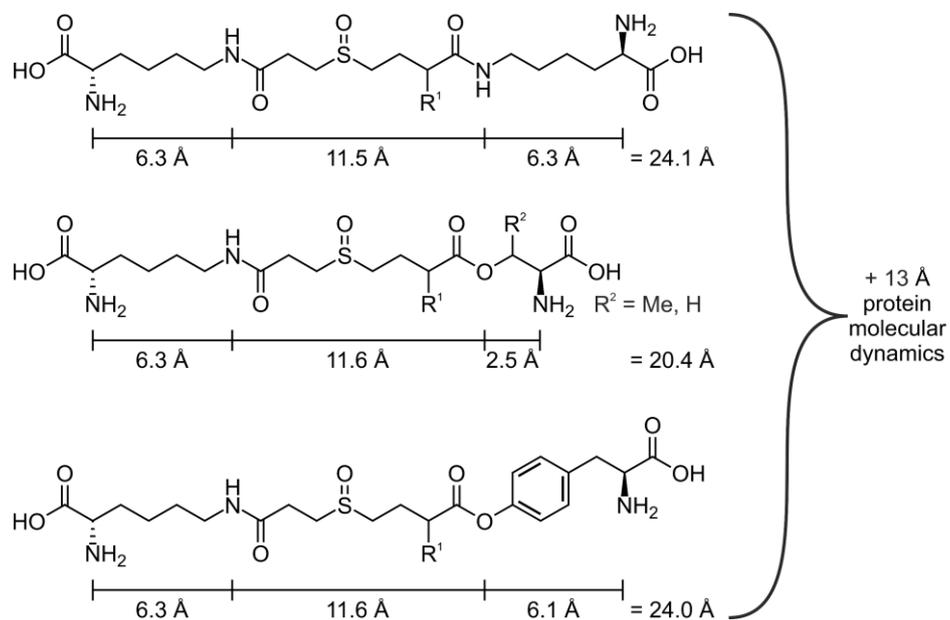
3.5. Cell permeability

For each sample, 2 million cells (HEK293T) were detached from the incubation vessel and washed twice with PBS. After resuspension in 100 µL PBS, 50 mM DSSO or cliXlink (**1**) were added for a final concentration of 5 mM. Two controls with the same volume of water or DMSO were also prepared. All samples were incubated on ice for 1 h following quenching with NH₄HCO₃ (20 mM final concentration, incubation on ice for 15 min). The cells were pelleted by centrifugation at 400 × g and 4 °C for 3 min, resuspended in 500 µL RIPA buffer with 1x protease inhibitor cocktail (cOmplete™ EDTA-free, Roche) and 90 U/mL Benzonase and incubated at 4 °C for 1 h in an orbital shaker. The samples were centrifuged at 15000 × g and 4 °C for 10 min and the supernatant was transferred to a new Eppendorf vial. To determine the protein concentration, a Bradford assay was conducted and 5 µg of each sample were mixed with 5x SDS loading buffer and loaded on a 15% SDS gel to analyze the degree of cross-linking. The gel was run at 50 V for the first 30 min and then at 150 V until the Bromophenol Blue dye reached the bottom of the gel. Staining was carried out using Coomassie Brilliant Blue G-250 (Sigma-Aldrich).

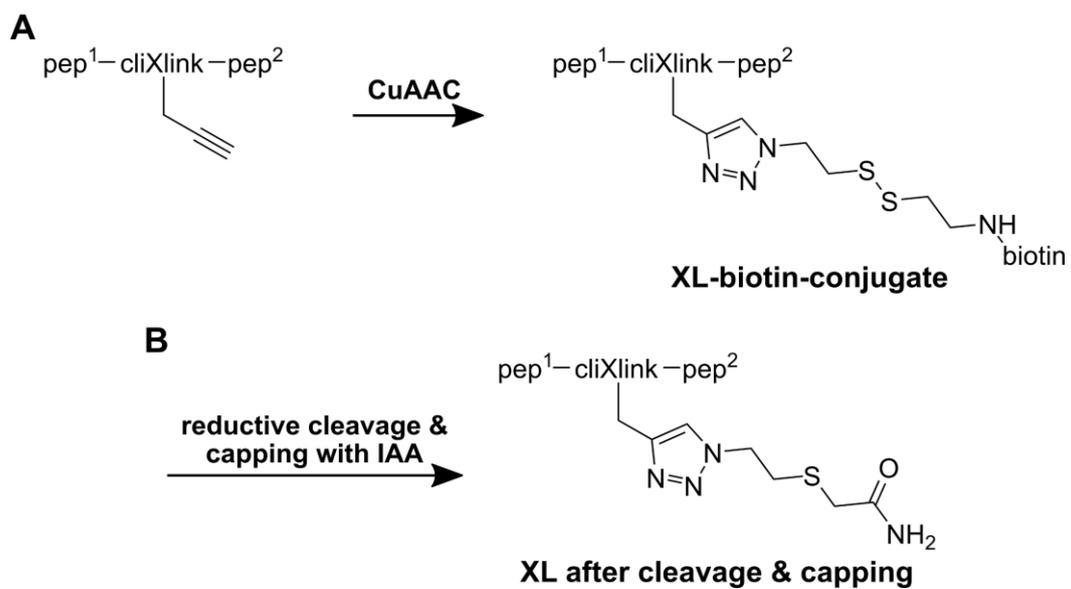


Supplementary Figure 1: 15%-SDS-PAGE of cell lysates (5 μ g each) from crosslinking experiments on HEK293T cells. Samples applied were: lane M: PageRuler™ Unstained Protein Ladder (5 μ L, Thermo Fisher Scientific), lane 1: negative control, lane 2: incubation with DMSO, lane 3: incubation with DSSO, lane 4: incubation with cliXlink (**1**). For the known cell-permeable cross-linker DSSO, the observed molecular weights are shifted upwards. The same trend is visible for cliXlink (**1**), indicating its cell permeability.

4. Additional Material



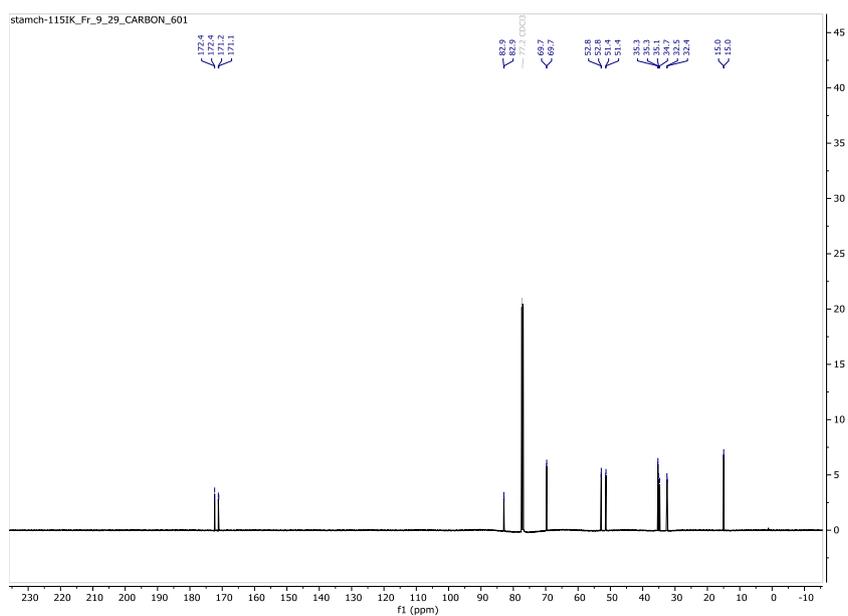
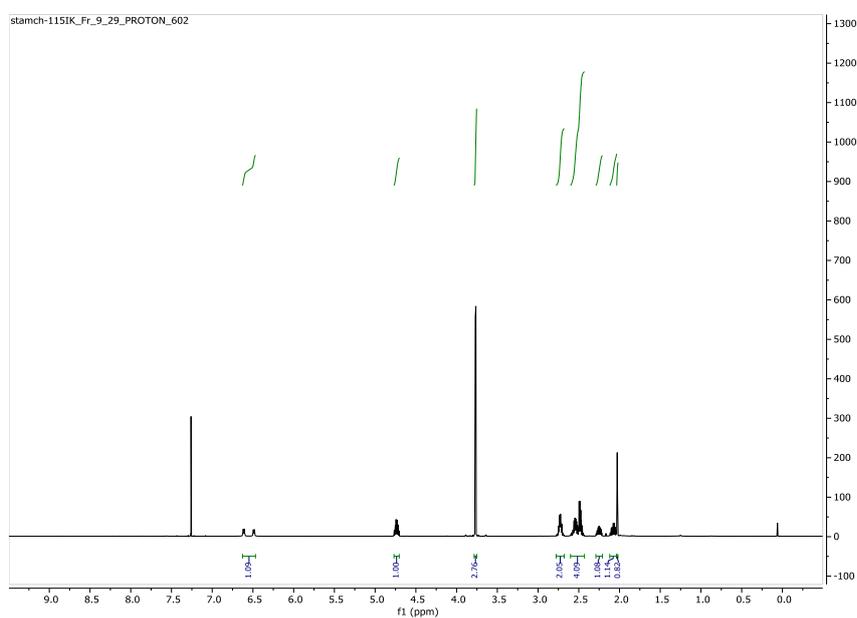
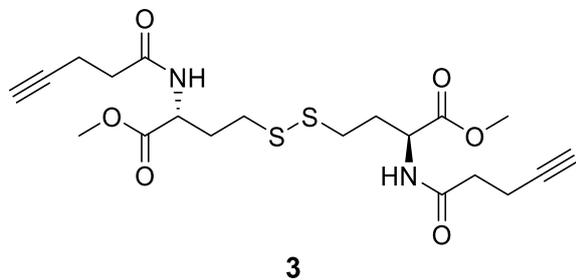
Supplementary Figure 2: Euclidean C_{α} - C_{α} distances for Lys-Lys, Lys-Ser/Thr and Lys-Tyr cross-links with cliXlink (**1**), calculated according to their bond lengths and angles. In addition to the calculated distances, $\approx 13 \text{ \AA}$ have to be considered for molecular dynamics.^[9]



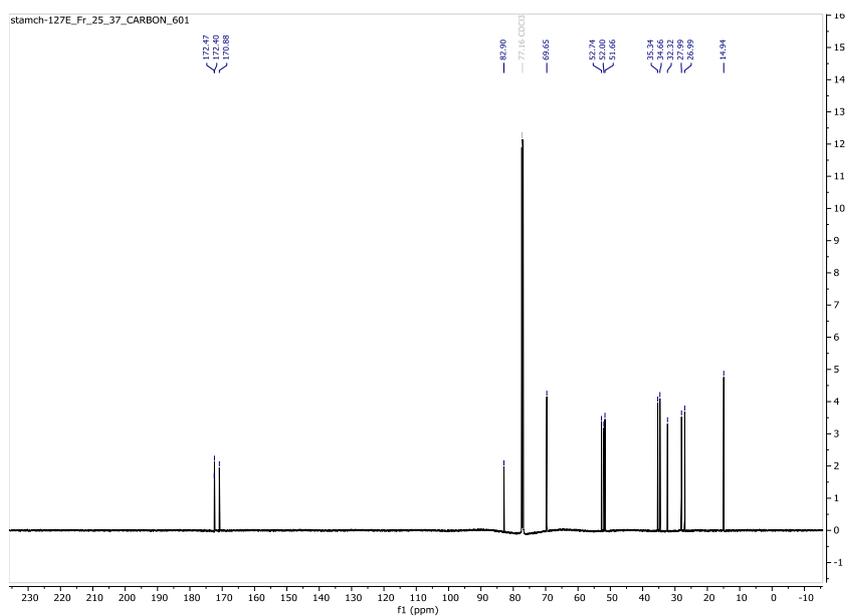
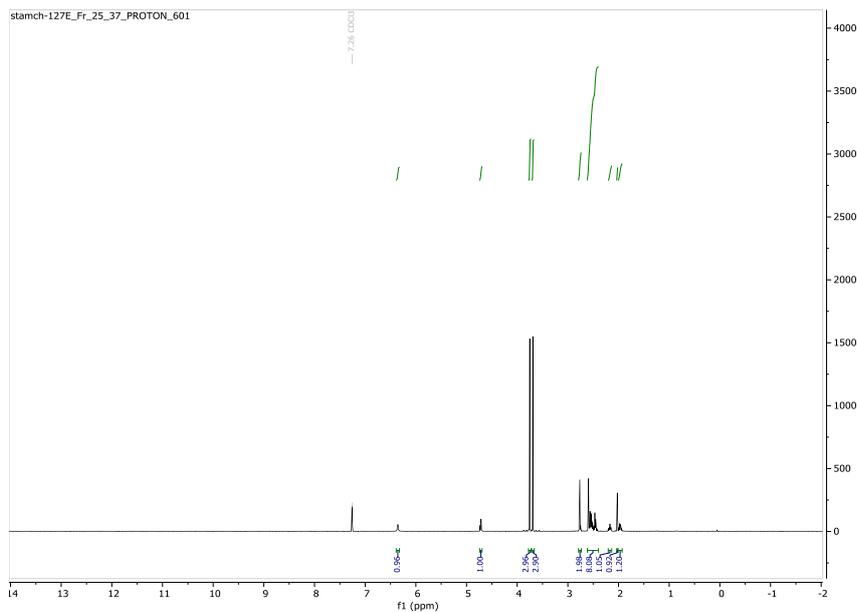
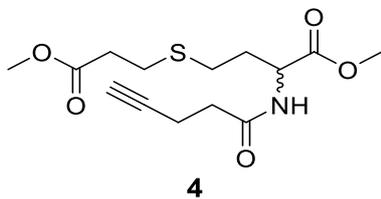
Supplementary Figure 3: Structures involved in the enrichment process. A) Resulting triazole-compound after CuAAC-reaction with biotin **7**. B) After reductive cleavage, the formed free thiol is capped with iodacetamide to avoid reoxidation, resulting in the shown structure.

5. NMR-Spectra

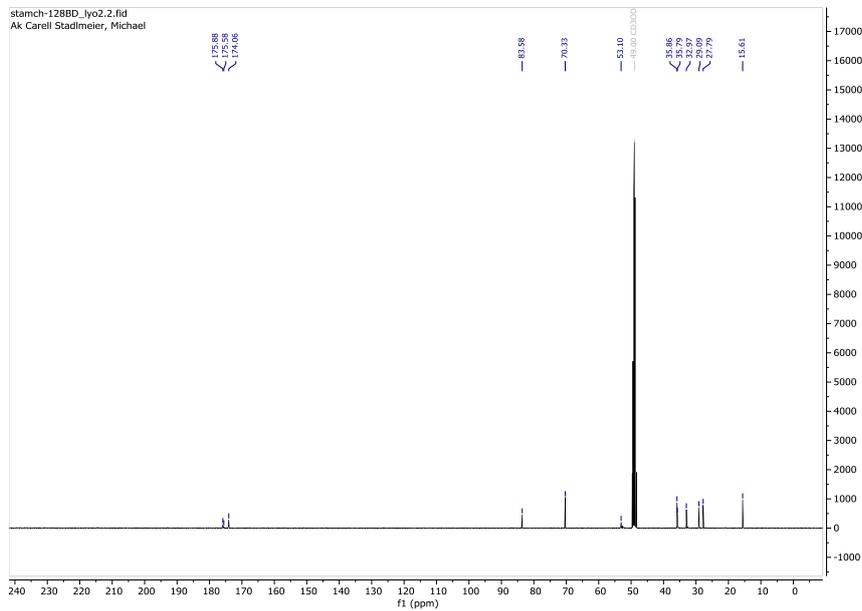
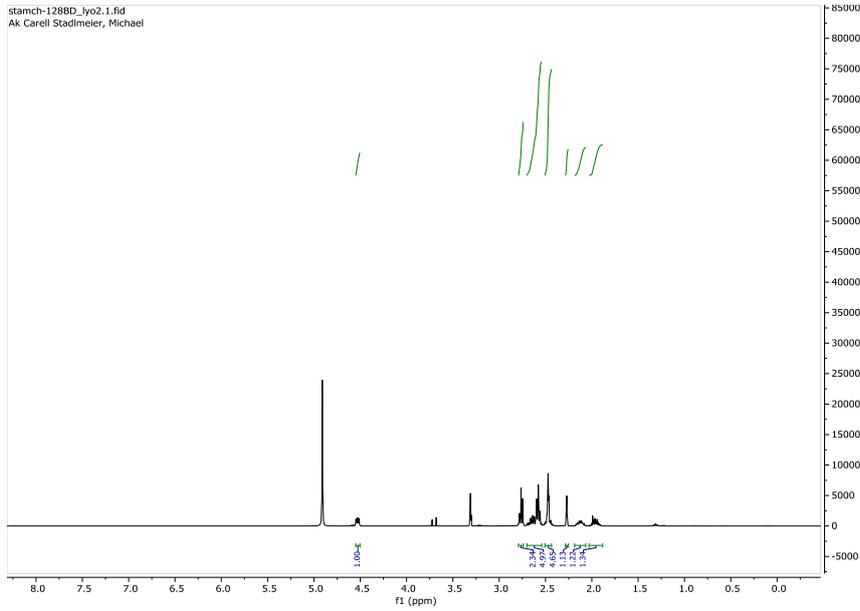
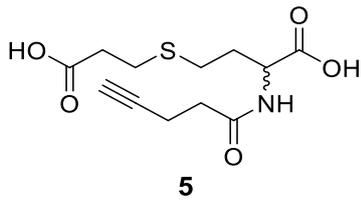
5.1. Methyl S-(((R)-4-methoxy-4-oxo-3-(pent-4-ynamido)butyl)thio)-N-(pent-4-ynoyl)-L-homocysteinate (3)



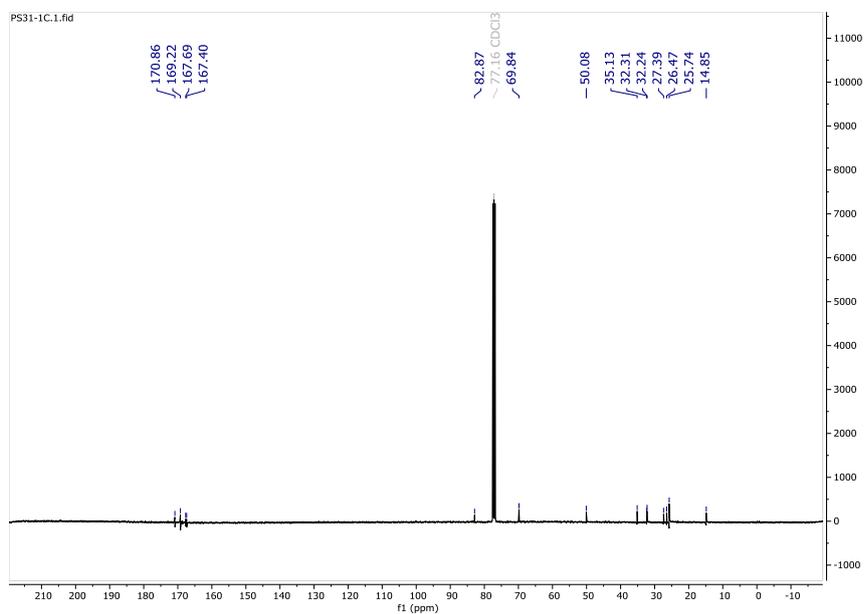
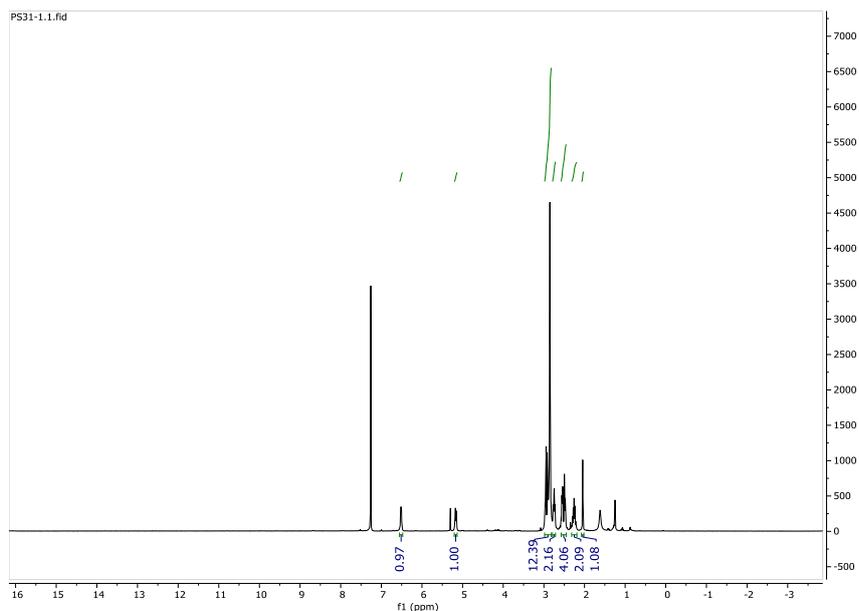
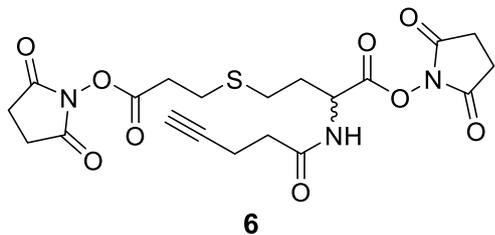
5.2. Methyl S-(3-methoxy-3-oxopropyl)-N-(pent-4-ynoyl)homocysteinate (4)



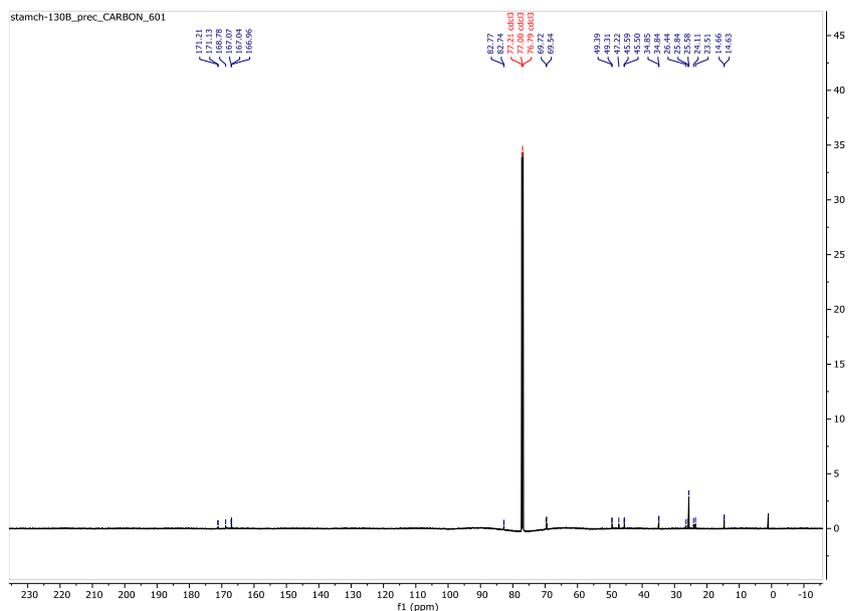
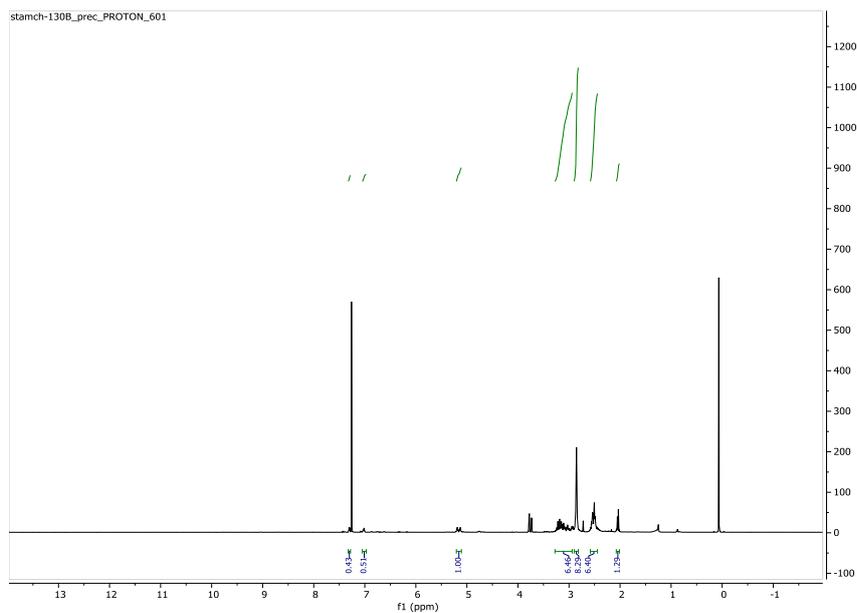
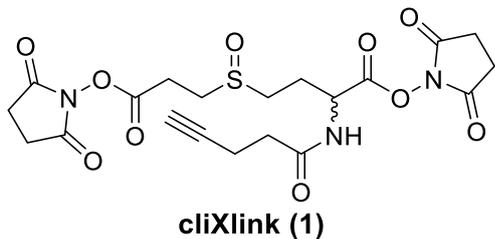
5.3. S-(2-carboxyethyl)-N-(pent-4-ynoyl)homocysteine (5)



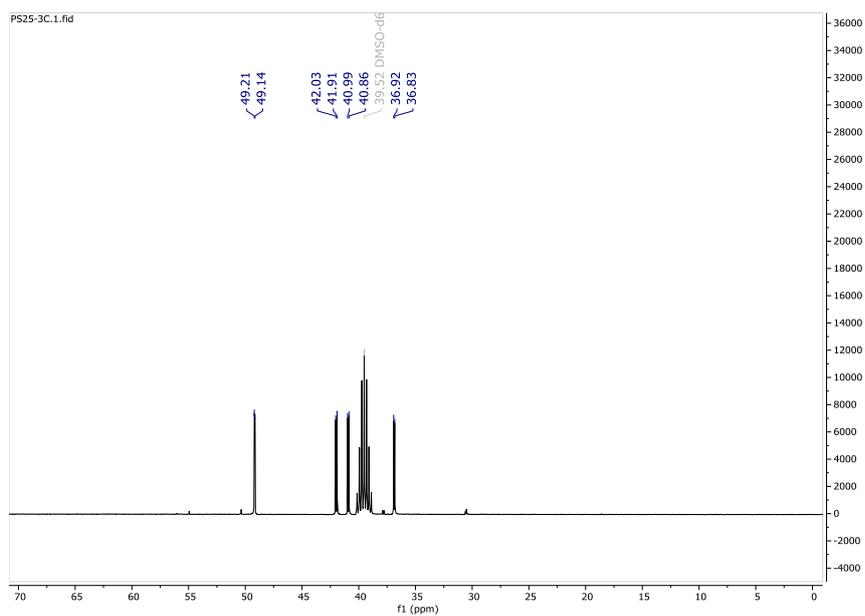
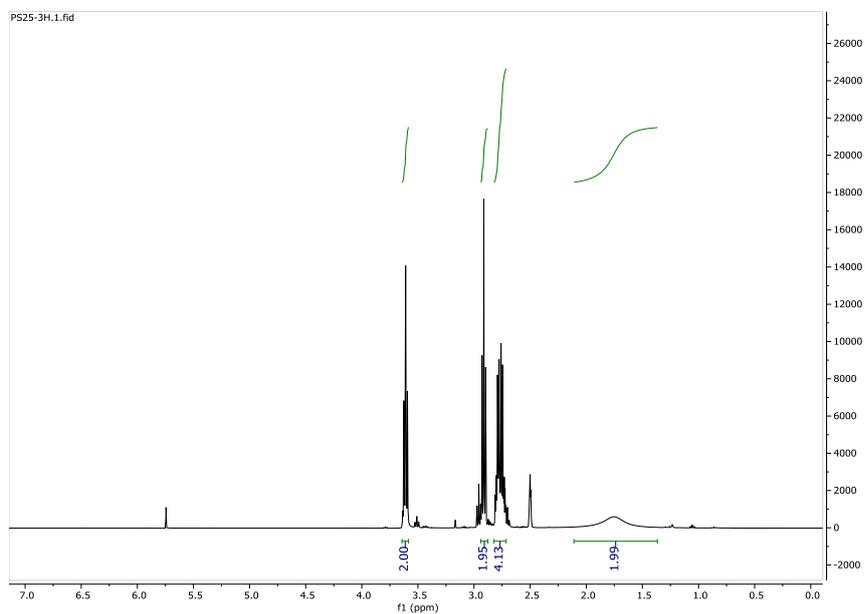
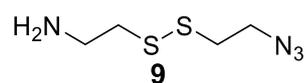
5.4. *N*-Succinimidyl *S*-(3-((*N*-succinimidyl)oxy)-3-oxopropyl)-*N*-(pent-4-ynoyl)homocysteinate (6)



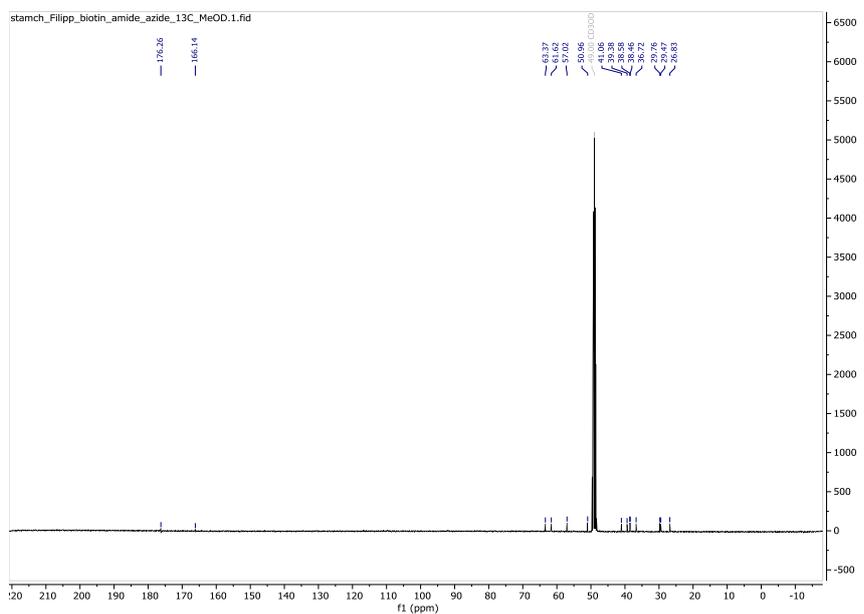
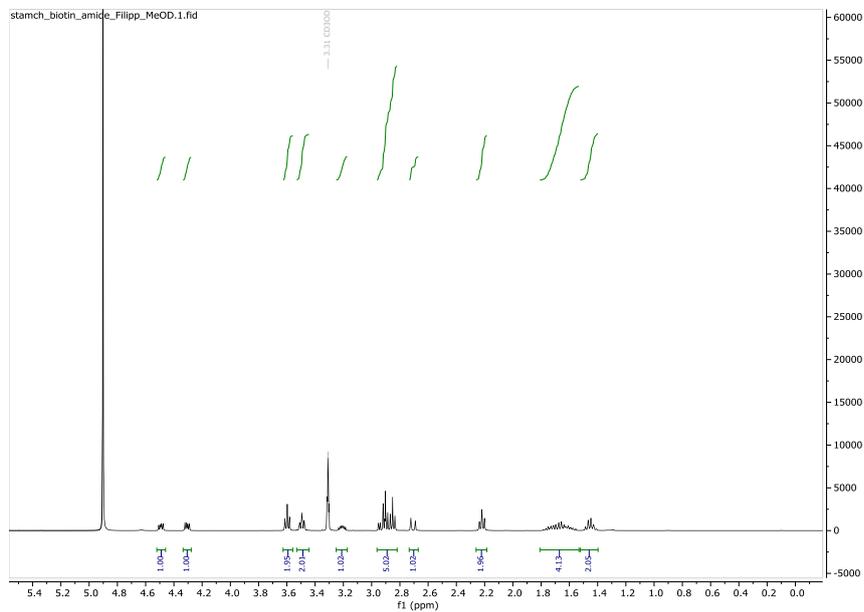
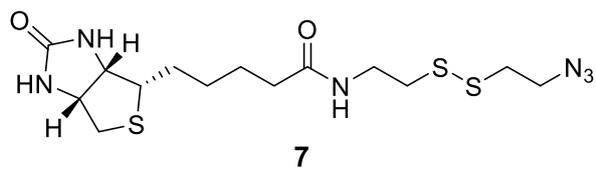
5.5. Bis-N-hydroxydisuccinimidyl-S-oxo-S-(carboxyethyl)-N'-(pent-4-inoyl)homocysteinate (cliXlink, 1)



5.6. 2-((2-azidoethyl)disulfaneyl)ethan-1-amine (9)



5.7. N-(2-((2-azidoethyl)disulfaneyl)ethyl)biotinamide (7)



6. References

- [1] E. W. Deutsch, A. Csordas, Z. Sun, *et al.*, *Nucleic Acids Res.* **2017**, *45*, D1100-1106.
- [2] Y. Perez-Riverol, A. Csordas, J. Bai, *et al.*, *Nucleic Acids Res.* **2019**, *47*, D442-450.
- [3] M. C. Chambers, B. Maclean, R. Burke, *et al.*, *Nat. Biotechnol.* **2012**, *30*, 918-920.
- [4] M. Götze, J. Pettelkau, R. Fritzsche, *et al.*, *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 83-97.
- [5] M. Götze, C. Iacobucci, C. H. Ihling, *et al.*, *Anal. Chem.* **2019**, *91*, 10236-10244.
- [6] M. Grimm, T. Zimniak, A. Kahraman, *et al.*, *Nucleic Acids Res.* **2015**, *43*, W362-369.
- [7] A. Bujacz, *Acta Crystallogr. D* **2012**, *68*, 1278-1289.
- [8] A. Kahraman, L. Malmstrom, R. Aebersold, *Bioinformatics* **2011**, *27*, 2163-2164.
- [9] Z. Orbán-Németh, R. Beveridge, D. M. Hollenstein, *et al.*, *Nat. Protoc.* **2018**, *13*, 478-494.

4 Ausblick

Im Rahmen dieser Arbeit wurden verschiedene Fragestellungen mithilfe von Protein-Massenspektrometrie als vorwiegende Analysemethode untersucht. Dabei konnten interessante Zusammenhänge aufgedeckt und verschiedene Methoden etabliert werden, die für weiterführende Studien von großer Bedeutung sind.

Zum einen konnte mithilfe von gelelektrophoretischen Methoden die Ausbildung kovalenter Addukte zwischen 5-Formylcytosin-haltiger DNA und Lysinen des Histon-Oktamers im Kontext von Nukleosomen *in vitro* gezeigt werden. Dies steht in Übereinstimmung mit während der Durchführung der Experimente publizierten Ergebnissen anderer Forschungsgruppen.^[544, 545] Ein direkter Nachweis dieser Addukte in Nukleosomen mittels Massenspektrometrie durch die Identifikation von DNA-modifizierten Peptiden konnte allerdings nicht erbracht werden und ist auch in der Literatur bislang nicht zu finden. Eine weitere Optimierung des mehrstufigen DNA-Protein-Verdaus könnte die Verdau-Effizienz erhöhen und für eine bessere Identifikation der modifizierten Peptide bei der massenspektrometrischen Analyse sorgen.

Mittlerweile sind einige Publikationen erschienen, die verschiedene Funktionen und Eigenschaften der DNA-Protein-Addukte nahelegen.^[549-553] Auch in dieser Arbeit wurde mit der Beteiligung in der aktiven Deformylierung eine mögliche Funktion dieser kovalenten Verknüpfung untersucht. Dabei konnte *in vitro* nahegelegt werden, dass monomethylierte Lysine des Histon-Oktamers eventuell für eine Deformylierung von fC-haltiger DNA sorgen. Um genauere Aussagen treffen zu können sowie eine statistische Signifikanz-Analyse zu ermöglichen, müssen weitere biologische Replikate angefertigt werden. Die Verwendung von speziell modifizierten Triphosphaten könnte hierbei für eine Reduzierung des Hintergrunds in der MS-Analyse sorgen. Auch eine mögliche Decarboxylierung könnte mit diesem System untersucht werden.

Da andere in der Arbeitsgruppe durchgeführte Experimente auf eine Beteiligung von weiteren Proteinen hindeuten, könnte auch eine Inkubation der formyl-haltigen Nukleosomen mit den entsprechenden Proteinen zu einer effizienteren Deformylierung der DNA führen. Ebenso sind DNA-Quantifizierungs-Experimente *in vivo* in Kombination mit einer Quantifizierung der möglicherweise übertragenen Formylgruppe auf Histon-Peptide denkbar. Hierfür könnte man sowohl Wildtyp-Zellen als auch unterschiedliche *knockout*-Zelllinien verwenden, welche die in Frage kommenden Proteine nicht exprimieren können.

Ausgehend von dem in dieser Arbeit durchgeführten Vergleich der Reaktivität zwischen fC und fU hinsichtlich der Bildung von Addukten mit einem Peptid könnten zudem Unterschiede dieser beiden Basen im Kontext von Nukleosomen untersucht werden. Dafür kann die hier optimierte PCR-Reaktion unter Verwendung von fluoreszierenden Primern angewendet werden, um auch 601_{fU} zu analysieren.

Mit den durchgeführten Experimenten zur Untersuchung einer Interaktion zwischen TET3 und GDH mithilfe von Crosslinking-Massenspektrometrie konnte eine direkte Wechselwirkung nicht nachgewiesen werden. Gründe hierfür sind die hohe Komplexität der verwendeten Proben, die ungleiche Überexpression der beiden Proteine sowie eine mögliche Zerstörung der hexameren Struktur von GDH während des Arbeitsablaufs. Da jedoch im Rahmen dieser Arbeit mehrfach TET3- bzw. GDH-Intralinks identifiziert werden konnten, stellt die Dissertation die grundlegenden Methoden für weitere Versuche in diese Richtung zur Verfügung. Denkbar ist die Expression der beiden Proteine in einem anderen System, wie z. B. Hefe oder Insektenzellen, in größerem Maßstab, um eine chromatographische Aufreinigung mittels Affinitätsgruppen oder Größenausschluss-Chromatographie zu erlauben. Dies könnte für eine erhebliche Reduktion der Komplexität sorgen und das Crosslinking in einer kontrollierteren Umgebung ermöglichen. Auch Experimente mit Fokus auf der für die TET3-GDH-Interaktion eventuell wichtigen mitochondrialen Importsequenz sind denkbar.

Schließlich ist auch die Verwendung eines anderen Crosslinkers eine Möglichkeit, wobei sich hierfür unter anderem das in dieser Arbeit etablierte Molekül cliXlink anbieten würde. Dieses könnte für eine Anreicherung der möglicherweise wenig abundanten Crosslinks zwischen TET3 und GDH sorgen und die Identifikation somit vereinfachen.

Der neuartige Sulfoxid-basierte Crosslinker besitzt durch die Alkin-Gruppe eine modulare Struktur und es konnte an einem Beispiel gezeigt werden, dass diese Funktionalität eine Anreicherung von Crosslinks aus einem komplexen, nicht-modifizierten Peptidhintergrund ermöglicht. Die dabei etablierten Methoden können ohne Weiteres für die Strukturanalyse von aufgereinigten Proteinen oder Proteinkomplexen sowie für Interaktomanalysen verwendet werden. Gleichzeitig dienen sie aber auch als Grundlage für die Erforschung zahlreicher weiterer Fragestellungen, welche von der Modularität des Crosslinkers profitieren. Dabei können allein durch die Einführung einer Azid-Gruppe bereits bekannte Systeme zugänglich gemacht werden. So ist die Entwicklung neuer Reagenzien für die Anreicherung, z. B. auf Basis der Wechselwirkung zwischen Adamantan und β -Cyclodextrin,^[587] denkbar.

Die in dieser Arbeit durchgeführten Crosslinking-Experimente beruhten allesamt auf MS²-Methoden, jedoch wurden auch vielversprechende Messungen durchgeführt, welche eine interessante Fragmentierung von cliXlink unter CID-Bedingungen in einer Ionenfalle zeigen. Durch die Asymmetrie des Crosslinkers bilden sich die Reporter-Fragmentpaare mit unterschiedlichen Intensitäten. Dadurch könnte in MS³-Experimenten eventuell auf eine Analyse aller vier Fragmente verzichtet werden, was die Dauer eines Scanzzyklus reduzieren könnte. Dies deutet auf die vorteilhaften Eigenschaften des neuen Crosslinkers für den Einsatz in MS³-Messungen hin und würde damit die Analyse von noch komplexeren Proben erlauben.

Hierfür müssten weitere Experimente an MS³-fähigen Massenspektrometern durchgeführt und die mit verschiedenen CID-Energien erhaltenen Fragmente genauer untersucht werden, um optimale Bedingungen für die asymmetrische Fragmentierung zu erhalten.

Die Massenspektrometrie und besonders Crosslinking-Massenspektrometrie zur Untersuchung von Proteinen und deren Wechselwirkungen ist immer noch ein sehr dynamisches Forschungsfeld und wird sich in den nächsten Jahren stetig weiterentwickeln. Dabei spielen sowohl Neuerungen in Hinsicht auf Geräte sowie Software als auch die Entwicklung und Verbesserung der verwendeten Crosslinking-Reagenzien und -Methoden eine wichtige Rolle. So wird nach und nach eine Art „Werkzeugkasten“ an Molekülen entstehen, von denen jedes für ein bestimmtes Anwendungsgebiet besonders geeignet ist und von Wissenschaftler*innen gezielt ausgesucht werden kann. Hierbei stellen modulare Reagenzien, wie das hier vorgestellte cliXlink, aufgrund ihrer vielseitigen Einsetzbarkeit eine besonders interessante Gruppe dar.

5 Material und Methoden

5.1 Reagenzien

Die verwendeten Chemikalien (außer anders gekennzeichnete) wurden von *Abnova*, *Applichem*, *Carl Roth*, *Linktech*, *New England Biolabs*, *Sigma-Aldrich* und *Thermo Fisher Scientific* bezogen. Soweit nicht anders spezifiziert, wurden für alle massenspektrometrischen Experimente Chemikalien und Lösungsmittel in HPLC- bzw. LC-MS-Qualität verwendet. Trockene Lösungsmittel wurden unter Argon-Atmosphäre über Molekularsieb gelagert und mittels Septums unter Argon-Überdruck entnommen.

Puffer und Lösungen von Reagenzien sind, falls nicht anderweitig gekennzeichnet, wässrige Lösungen der jeweils aufgeführten Komponenten. Als PBS wird in dieser Arbeit eine Lösung von 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, 137 mM NaCl sowie 2.7 mM KCl mit einem pH-Wert von 7.4 bezeichnet.

Primer und andere Oligodesoxynukleotide wurden lyophilisiert von *Sigma-Aldrich* bezogen, wobei als Aufreinigungsmethode *Desalt* ausgewählt wurde. Die Resuspension auf eine Konzentration von 100 µM erfolgte in H₂O.

Trypsin (*Thermo Fisher Scientific*) sowie Trypsin/Lys-C (*Promega*) für den Verdau von Proteinen zur massenspektrometrischen Analyse wurden in MS-Qualität bezogen und mit 50 mM AcOH auf die gewünschte Stamm-Konzentration resuspendiert.

5.2 Geräte

Zentrifugationsschritte wurden mithilfe der Zentrifugen *5810R* und *5424R* von *Eppendorf* durchgeführt.

Für Inkubationsschritte, bei denen die Proben durchmischt werden mussten, wurden ein *Thermomixer comfort* sowie ein *ThermoMixer C* von *Eppendorf* bzw. ein Rotationsmischer von *VWR* verwendet.

Die Messung von Bradford-Assays zur Bestimmung der Proteinkonzentration von Zelllysaten erfolgte an einem *BioPhotometer Plus* von *Eppendorf*.

DNA-Konzentrationen wurden an einem *NanoDrop ND-1000*-Spektrophotometer (*PEQLAB Biotechnology*) gemessen.

PCR-Ansätze sowie DNA-Hybridisierungen wurden mithilfe eines *Mastercycler personal* von *Eppendorf* ausgeführt. Die vorhergehende Optimierung der Primer-Hybridisierungstemperatur einer PCR erfolgte an einem *T100 Thermal Cycler* von *Bio-Rad*.

Gelelektrophoretische Analysen erfolgten im Falle von SDS- und nativer PAGE in einer *Mini-PROTEAN Tetra* Elektrophoresekammer von *Bio-Rad* sowie bei Harnstoff-PAGE mit einer *PROTEAN II xi Cell* von *Bio-Rad*, welche mit einem *DC30-K20* Wasserbad von *Thermo Haake* temperiert wurde. PCR-Verifizierungen wurden mit einem *PerfectBlue Gelsystem Mini S* von *PEQLAB Biotechnologie* durchgeführt. Für die Belichtung der EMSA- sowie Harnstoff-Gele sowie die Digitalisierung von SDS-Gele wurde ein *LAS-3000 Image Reader* von *Fujifilm* oder ein *Amersham Imager 680* von *GE Healthcare* und für PCR-Gele eine *E-BOX VX5* von *Vilber Lourmat* verwendet.

Fluoreszenz-Kinetiken zur Verfolgung der CuAAC-Reaktion wurden an einem *GENios Pro* Mikrotiterplatten-Lesegerät von *Tecan* durchgeführt.

5.3 Methoden zum Addukt-Nachweis zwischen fC-haltiger DNA und Histonen

In Vorversuchen wurde zunächst die Interaktion von fC- bzw. fU-haltigen, einzelsträngigen Oligodesoxyribonukleotiden mit einem kurzen Lysin-haltigen Peptid mittels Gelelektrophorese getestet. Zur Untersuchung von möglichen kovalenten DNA-Protein-Addukten zwischen der Formylgruppe von fC bzw. fU und den primären Aminen der Lysin-Seitenketten von Histonen wurden Nukleosomen mit formyl-haltiger DNA rekonstituiert. Die potenziellen Addukte in Form von gegenüber Hydrolyse labilen Schiff'schen Basen wurden durch Reduktion zu sekundären Aminen stabilisiert. Anschließend wurden DNA sowie Proteine verdaut und die Proben mittels LC-ESI-MS² analysiert.

5.3.1 Oligonukleotid-Peptid-Addukte

Zu Beginn wurden Addukte zwischen fC/fU und Lysinen an einem einfachen Testsystem untersucht, indem die Reaktion eines einzelsträngigen Oligonukleotids mit einer Formylgruppe und einem Lysin-haltigen Peptid mittels Gelelektrophorese analysiert wurde.

5.3.1.1 Synthese von fC/fU-haltigen Oligodesoxynukleotiden

| Puffer und Reagenzien | |
|-----------------------|---|
| Aktivierungs-Lösung | 0.25 M <i>Activator 42</i> (<i>Sigma-Aldrich</i>) |
| Cap A-Lösung | 20 % (v/v) Ac ₂ O, 30 % (v/v) 2,6-Lutidin, 50 % (v/v) MeCN |
| Cap B-Lösung | 20 % (v/v) <i>N</i> -Methylimidazol, 80 % (v/v) MeCN |
| Entschützungs-Lösung | 3% (v/v) DCA, 97 % (v/v) DCM |
| Oxidations-Lösung | 25 mM I ₂ in 2,6-Lutidin/MeCN/H ₂ O (6:65:30 (v/v)) |
| HPLC-Puffer A | 0.1 M NEt ₃ /AcOH |
| HPLC-Puffer B | 0.1 M NEt ₃ /AcOH in 80 % MeCN |
| MALDI-Matrix | 0.7 M 3-HPA, 0.07 M Diammoniumcitrat in 50 % MeCN |

Die Oligonukleotide mit einem fC bzw. fU wurden mittels DNA-Festphasensynthese an einem *ABI 394 DNA/RNA Synthesizer* von *Applied Biosystems* im 200 nmol-Maßstab hergestellt. Als Sequenz wurde dabei ein kurzer Abschnitt der 1998 beschriebenen *601* DNA gewählt.^[261]

Die Synthese des fC-Strangs (fC-ODN: 5'-AGCTGTCTACGACCAGTT**f**CAGCGGCCT-3') wurde dabei von *Dr. Iacovos Michaelides* und *Dr. Franziska Traube* durchgeführt, während der entsprechende fU-Strang (fU-ODN: 5'-AGCTGTCTACGACCAGTT**f**UAGCGGCCT-3') in Zusammenarbeit mit *Dr. Iacovos Michaelides* und *Dr. Franziska Traube* hergestellt wurde. Die verwendeten Phosphoramidite Bz-dA-CE, dmf-dG-CE, Ac-dC-CE, dT-CE (*Link Technologies*) und 5'-dox-dU-CE (fU mit Acetal-geschützter Formylgruppe; in der Arbeitsgruppe synthetisiert) wurden als 0.1 M Lösungen in trockenem Acetonitril eingesetzt. Als Festphase diente ein 0.2 μmol dT LV-PS von *Glen Research*, da die 3'-Base des Stranges ein Thymin war. Die Kupplungszeit für die kanonischen Basen betrug 1 min, während das 5'-dox-dU-Phosphoramidit 4 min lang gekuppelt wurde.

Im Anschluss an die Synthese wurde der Festphasenträger lyophilisiert und das Säulenmaterial daraufhin vom Träger in ein Reaktionsgefäß überführt. Für die Abspaltung des Stranges von der Festphase sowie Entschützung wurde das Material zunächst in 800 μL 30 % NH_4OH resuspendiert und 17 h bei 25 °C inkubiert. Die Suspension wurde durch einen 0.2 μm Filter filtriert und lyophilisiert. Anschließend wurde die DNA gefällt, indem der Rückstand in 100 μL H_2O gelöst und 60 μL 3 M NaOAc sowie 640 μL kaltes EtOH zugegeben wurde, bevor die Lösung 15 min auf Trockeneis inkubiert und 15 min bei $21130 \times g$ und 4 °C zentrifugiert wurde. Der Überstand wurde verworfen und der Rückstand lyophilisiert. In einem zweiten Entschützungsschritt wurde die DNA in 100 μL 80 % Essigsäure resuspendiert und 5 h bei 20 °C inkubiert. Schließlich erfolgte eine weitere Ethanol-fällung wie oben beschrieben, wobei 1.04 mL Ethanol verwendet wurden.

Die Aufreinigung des Strangs wurde mittels präparativer Umkehrphasen-HPLC an einer 1525 *ef*-Pumpe mit 2487 UV-Detektor (*Waters*) unter Verwendung einer CC 250/4 *Nucleosil 120-3 C18*-Säule (*Macherey-Nagel*) durchgeführt. Hierfür wurde die Probe anhand eines Gradienten von 0 % HPLC-Puffer B (entspricht 100 % HPLC-Puffer A) bis 25 % HPLC-Puffer B über einen Zeitraum von 45 min bei einem Fluss von 5 mL/min aufgetrennt und in Fraktionen von ca. 1 – 2 mL aufgefangen. Die Fraktionen wurden mithilfe von analytischer HPLC sowie MALDI-TOF-MS analysiert. Zum Vergleich wurde auch die Probe vor Aufreinigung mittels analytischer HPLC untersucht. Analytische Chromatogramme wurden mit einer *alliance 2695*-Pumpe mit 2996 Photodiodenarray-Detektor (*Waters*) angefertigt. Säule und Puffer waren dieselben wie für die präparative HPLC; der Fluss betrug jedoch 0.5 mL/min. MALDI-TOF-Spektren der einzelnen HPLC-Fraktionen wurden mit einem *autoflex II*-Spektrometer von *Bruker Daltonics* aufgenommen. Die Proben wurden dafür zunächst mittels einer Zellulosemischester-Membran mit einer Porengröße von 0.025 μm (*Merck Millipore*) gegen H_2O entsalzt. 1 μL der MALDI-Matrix wurde auf die Trägerplatte aufgetragen und durch Evaporation des Lösemittels kristallisiert. Anschließend wurde 1 μL der entsalzten Probe auf die Matrix gegeben und ebenfalls getrocknet. Die Messung erfolgte im Negativ-Modus und Ionen unterhalb eines *m/z*-Wertes von 1900 wurden supprimiert.

Die Fraktionen der semi-präparativen HPLC, welche ausschließlich Produkt enthielten, wurden vereint und lyophilisiert. Die Resuspension erfolgte in H₂O, bevor die DNA-Konzentration UV/Vis-spektroskopisch bestimmt wurde.

5.3.1.2 Reaktion von Peptid und ODN

| Puffer | |
|-----------------------|---|
| Katalysator-Puffer I | 50 mM HEPES pH 5.3, 50 mM <i>p</i> -Anisidin |
| Katalysator-Puffer II | 50 mM K-Phosphat pH 6.0, 1 mM <i>p</i> -Diaminobenzol |

Eine Lösung von 1.5 µM ODN (fC-ODN bzw. fU-ODN) und 1.5 mM Peptid (Ac-NH-IEAKGER-COOH; Ac-IR7, kommerziell erworben von *ProteoGenix*) in 50 mM K-Phosphat pH 7.4, Katalysator-Puffer I oder II wurde über Nacht bei RT inkubiert. Anschließend wurde 1 M NaBH₃CN (Endkonzentration 100 mM) zugegeben und über Nacht bei 25 °C inkubiert. Nun wurde 1 M NaBH₄ (Endkonzentration 100 mM) zugegeben und 1 h bei RT unter Agitation inkubiert.

5.3.1.3 Harnstoff-PAGE

| Puffer und Reagenzien | |
|--|--|
| 10x TBE | 0.89 M TRIS, 0.89 M Borsäure, 20 mM Na ₂ H ₂ EDTA, pH 8.3 |
| 15 % Harnstoff-Polyacrylamid-Gel (Mengenangaben für ein Gel) | 24 mL <i>Rotiphorese</i> Sequenziergel Konzentrat, 12 mL 8 M Harnstoff, 4 mL 10x TBE, 200 µL 0.1 g/mL APS, 20 µL TEMED |
| 2x Harnstoff-Probenpuffer | 7 M Harnstoff, 12 % <i>Ficoll</i> 400, 1x TBE, 0.3 mg/mL Bromphenolblau |

Die Peptid-ODN-Addukte wurden mittels 15 %-Harnstoff-PAGE analysiert. Hierfür wurden alle Gelkomponenten bis auf APS und TEMED gemischt. Nun wurde zunächst APS zugegeben, die Lösung durch Invertieren gut durchmischt und direkt im Anschluss TEMED zugegeben und erneut vermischt. Die Gelösung wurde sofort zwischen zwei Glasplatten (22.3 cm × 20 cm und 20 cm × 20 cm; *biostep*) mit 1 mm dicken Abstandshaltern (*biostep*) gegossen und ein Kamm mit 20 Taschen eingesetzt (resultierendes Gelvolumen ohne Taschen ca. 17.5 cm × 16 cm × 0.1 cm).

15 pmol DNA wurden von den Proben abgenommen und mit dem gleichen Volumen an 2x Harnstoff-Probenpuffer verdünnt, zusammen mit einer Kontroll-DNA ohne Peptid auf das Gel aufgetragen und für 3.5 h in 1x TBE bei 40 °C und einer konstanten Leistung von 20 W aufgetrennt. Anschließend wurde das Gel mit H₂O gewaschen, 30 min mit *SYBR Green I* (*Sigma-Aldrich*; 1:10000 in 1x TBE verdünnt) gefärbt und mittels UV-*trans*-Illumination analysiert.

5.3.2 DNA-Nukleosom-Addukte

Um die Interaktion von DNA und Histonen *in vitro* nachzustellen, wurden Nukleosomen mit formylhaltiger DNA rekonstituiert und auf Addukte zwischen der Formylgruppe und primären Aminen der Histone untersucht.

5.3.2.1 Synthese einer überwiegend fC/fU-modifizierten DNA mittels PCR

| Reagenzien | |
|------------|---------------------------------|
| FWD-Primer | 5'-ATCGATGTATATATCTGACACGTGC-3' |
| REV-Primer | 5'-ATCAGAATCCCGGTGCCGAG-3' |

Für die Nukleosom-Experimente *in vitro* wurde mittels PCR ein 145 bp langer Abschnitt der 601 DNA-Sequenz^[261] amplifiziert, welcher für die Nukleosom-Positionierung verantwortlich ist.^[519] Hierbei wurden alle Cytosine (außer die in den Primern enthaltenen) durch 5-Formylcytosin bzw. alle Thyminen durch 5-Formyluracil ersetzt. Als DNA-Polymerase wurde *KOD XL (Merck Millipore)* verwendet, da diese sich besonders gut für den Einbau modifizierter Nukleotide eignet. Als Templat wurde die unmodifizierte 601 Sequenz verwendet (s. Tab. 5.1), welche von Dr. Franziska Traube zur Verfügung gestellt wurde. Um Kontrollexperimente mit der unmodifizierten Sequenz durchführen zu können, wurde die PCR auch mit dCTP durchgeführt.

Tabelle 5.1: Sequenzen der beiden unmodifizierten 601 DNA-Einzelstränge.

| Sequenz | |
|--------------------|--|
| 601 _{FWD} | 5'-ATCGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCCTTGGCGGTTAAAACG CGGGGACAGCCGCTACGTGCGTTTAAAGCGTGCTAGAGCTGTCTACGACCAATTGAGCGGCCT CGGCACCGGGATTCTGAT-3' |
| 601 _{REV} | 5'-ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGC ACGTACGCGCTGTCCCCGCGTTTTAACCGCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTC AGATATATACATCGAT-3' |

Die Bedingungen der PCR sowie die Konzentrationen der Reagenzien wurden zunächst mit einzelnen 50 µL-Ansätzen optimiert, bevor zehn 50 µL-Ansätze bei den optimalen Bedingungen (Tab. 5.2) angesetzt wurden, um eine ausreichende Menge des jeweiligen DNA-Strangs zu erhalten. Die optimierte Zusammensetzung der Proben war dabei wie folgt: 0.2 pg/µL Templat, 0.2 mM dNTPs (N = A, G, C, T für Kontrollstrang (601_C); N = A, G, fC, T für fC-Strang (601_{fC}); N = A, G, C, fU für fU-Strang (601_{fU})), je 0.4 µM Primer (FWD und REV) sowie 0.025 U/µL *KOD XL* DNA Polymerase in 1x *KOD XL*-Puffer (bei Polymerase mitgeliefert).

Tabelle 5.2: PCR-Programm für die Synthese der verschiedenen 601 DNA-Stränge.

| Schritt | Bedingungen | |
|--|------------------|----------------------|
| | unmodifiziert | modifiziert (fC, fU) |
| Anfangs-Denaturierung | 1 min bei 94 °C | |
| Denaturierung | 30 s bei 94 °C | |
| Hybridisierung | 15 s bei 52 °C | |
| Elongation | 8 s bei 72 °C | 30 s bei 72 °C |
| Zyklen (Denaturierung – Hybridisierung – Elongation) | 30 | 45 |
| finale Elongation | 10 min bei 72 °C | |

Im Anschluss an die PCR wurde die DNA mithilfe des *GeneJET PCR Purification Kit* (*Thermo Fisher Scientific*) aufgereinigt, wobei aufgrund der geringen Größe des PCR-Produkts gemäß dem Herstellerprotokoll *iso*-Propanol zugegeben wurde. Zur Erhöhung der Ausbeute wurde der Durchfluss der Aufreinigung aufgehoben, nach der ersten Elution erneut auf die Säule aufgetragen und ein weiteres Mal aufgereinigt und separat eluiert. Die Konzentration des PCR-Produkts wurde UV/Vis-spektroskopisch bestimmt.

5.3.2.2 Agarose-Gelelektrophorese

| Puffer und Reagenzien | |
|---|---|
| 1x TAE | 40 mM TRIS, 20 mM AcOH, 1 mM Na ₂ H ₂ EDTA, pH 8.5 |
| 1 % Agarose-Gel (Mengenangabe für ein <i>Mini S</i> -Gel) | 0.5 g Agarose in 50 mL 1x TAE |
| 6x Agarose-Probenpuffer | 20 mM TRIS, 0.15 g/mL <i>Ficoll</i> 400, 60 mM Na ₂ H ₂ EDTA, 4.8 mg/mL SDS, 0.3 mg/mL Xylencyanol FF, 0.3 mg/mL Bromphenolblau, 1.2 mg/mL Orange G |

Zur Kontrolle der PCR wurde eine 1 %-Agarose-Gelelektrophorese durchgeführt. Hierfür wurde die Suspension von Agarose in 1x TAE in einem Mikrowellengerät erhitzt, bis die Agarose komplett gelöst war. Anschließend wurde die Lösung bei RT etwas abgekühlt, mit 5 µL *peqGREEN* (*Peqlab*) versetzt und in die Gelkammer gefüllt. Ein Kamm mit zehn Taschen wurde verwendet.

Je 200 ng der DNA-Proben wurden mit 6x Agarose-Probenpuffer (Endkonzentration 1x) gemischt, in jeweils eine Geltasche gefüllt und in 0.5x TAE-Puffer für 90 min bei einer konstanten Spannung von 70 V aufgetrennt. Zur groben Bestimmung der Länge der DNA-Fragmente wurden zusätzlich 10 µL eines DNA-Standards (*FastRuler Low Range DNA Ladder* von *Thermo Fisher Scientific*) auf das Gel aufgetragen. Die Analyse der Gelelektrophorese erfolgte anschließend durch *trans*-Illumination mit UV-Licht.

5.3.2.3 Nukleosom-Rekonstitution

Die Rekonstitution der Nukleosomen mit unmodifizierter sowie formyl-haltiger DNA folgte dem *Dilution Assembly Protocol* des *EpiMark Nucleosome Assembly Kit* (New England Biolabs), mit einem Verhältnis zwischen DNA und Proteinen von 1:1. Als Verdünnungspuffer wurde allerdings 10 mM HEPES (pH 7.9) statt TRIS verwendet, um zu vermeiden, dass die primären Amine von TRIS mit den Formylgruppen interagieren. Es wurden Nukleosomen mit 601_C , 601_{fC} sowie 601_{fU} rekonstituiert.

5.3.2.4 Native PAGE

| Reagenzien | |
|---|---|
| 4.5 % natives Polyacrylamid-Gel (Mengenangaben für vier Gele) | 4 mL <i>Rotiphorese Gel 30 (37.5:1)</i> (Carl Roth), 23.25 mL H ₂ O, 1.5 mL 10x TBE, 750 µL Glycerin, 300 µL 0.1 g/mL APS, 30 µL TEMED |

Zur Untersuchung der Nukleosom-Bildung wurde eine EMSA-ähnliche Methode verwendet, bei der die Bindung von DNA und Proteinen anhand einer Verschiebung der DNA-Bande zu höheren Molekulargewichten visualisiert werden kann.

Alle Gelkomponenten außer APS und TEMED wurden gemischt, bevor APS und anschließend TEMED zugegeben wurde. Die Lösung wurde sofort zwischen zwei Glasplatten (10.1 cm × 8.3 cm mit integrierten, 0.75 mm dicken Abstandhaltern und 10.1 cm × 7.1 cm; *Bio-Rad*) gegossen und ein Kamm mit zehn Taschen eingesetzt (resultierendes Gelvolumen ohne Taschen ca. 6.1 cm × 8.3 cm × 0.75 mm).

Je 3.125 pmol rekonstituiertes Nukleosom wurden mit Glycerin (1/5 des Probenvolumens) verdünnt, auf das Gel aufgetragen und in 0.25x TBE bei 4 °C und einer konstanten Spannung von 100 V für 30 min und bei 150 V für weitere 30 min aufgetrennt. Zur Größenabschätzung der DNA wurden 5 µL eines DNA-Standards (*100 bp DNA Ladder* von New England Biolabs) aufgetragen. Anschließend wurde das Gel mit H₂O gewaschen, 30 min mit *SYBR Green I* (*Sigma-Aldrich*; 1:10000 in 0.25x TBE verdünnt) gefärbt und mittels UV-trans-Illumination analysiert.

5.3.2.5 Reduktion der DNA-Protein-Addukte und Hydroxylamin-Behandlung

Zur Stabilisierung der Schiff'schen Basen-Addukte zwischen fC und den Histonen des Nukleosom-Komplexes wurden die Nukleosomen mit 601_C sowie 601_{fC} nach der Rekonstitution (s. 5.3.2.4) reduziert. Dafür wurden je 6.25 pmol Nukleosom durch Zugabe von 1 M NaBH₃CN (Endkonzentration 100 mM) und Inkubation über Nacht bei 18 °C zusammen mit Kontrollproben ohne Reduktionsmittel inkubiert. Anschließend wurde die Hälfte der Proben in ein neues Reaktionsgefäß überführt und 5 min bei 95 °C inkubiert.

Je 3.125 pmol der rekonstituierten Nukleosomen mit 601_{fC} wurden zusätzlich mit Hydroxylamin behandelt. Dafür wurde vor oder nach der oben beschriebenen Inkubation (mit bzw. ohne Reduktion) 5 % Hydroxylamin zugegeben (Endkonzentration 0.25 %) und die Proben 15 min bei RT inkubiert.

5.3.2.6 SDS-PAGE von DNA-Protein-Addukten

| Puffer und Reagenzien | |
|---|---|
| 10 % Trenngel (Mengenangaben für vier Gele) | 6.7 mL <i>Rotiphorese Gel 30 (37.5:1) (Carl Roth)</i> , 5 mL 1.5 M TRIS pH 8.8, 200 µL 0.1 g/mL SDS, 7.9 mL H ₂ O, 200 µL 0.1 g/mL APS, 20 µL TEMED |
| 5 % Sammelgel (Mengenangaben für vier Gele) | 1.7 mL <i>Rotiphorese Gel 30 (37.5:1) (Carl Roth)</i> , 1.25 mL 1 M TRIS pH 6.8, 100 µL 0.1 g/mL SDS, 6.8 mL H ₂ O, 100 µL 0.1 g/mL APS, 10 µL TEMED |
| 5x SDS-Probenpuffer | 225 mM TRIS, 50 % (v/v) Glycerin, 50 mg/mL SDS, pH 6.8 |
| 1x SDS-Elektrophoresepuffer | 25 mM TRIS, 192 mM Glycin, 3.5 mM SDS |

Für die Analyse der Addukt-Stabilität wurden analytische SDS-Polyacrylamid-Gele nach *Laemmli*^[588] verwendet. Alle Gelkomponenten des Trenngels außer APS und TEMED wurden gemischt, bevor APS und anschließend TEMED zugegeben wurde. Die Lösung wurde sofort zwischen zwei Glasplatten (10.1 cm × 8.3 cm mit integrierten, 0.75 mm dicken Abstandshaltern und 10.1 cm × 7.1 cm; *Bio-Rad*) bis 2 cm unter den Rand gegossen, mit 1 mL *iso*-Propanol bedeckt und aushärten gelassen (resultierendes Trenngelvolumen ca. 5.1 cm × 10.1 cm × 0.75 mm). Anschließend wurde das *iso*-Propanol wieder entfernt, das Sammelgel auf dieselbe Weise wie das Trenngel vorbereitet, auf letzteres gegossen und ein Kamm mit zehn Taschen eingesetzt (resultierendes Sammelgelvolumen ohne Taschen ca. 1 cm × 10.1 cm × 0.75 mm).

Die Proben wurden mit 5x SDS-Probenpuffer (Endkonzentration 1x) versetzt, auf das Gel aufgetragen und in 1x SDS-Elektrophoresepuffer bei einer konstanten Spannung von 50 V für 30 min und bei 100 V für weitere 60 min aufgetrennt. Zur Größenabschätzung der DNA wurden 5 µL eines DNA-Standards (*100 bp DNA Ladder von New England Biolabs*) aufgetragen. Nach der Elektrophorese wurde das Gel mit H₂O gewaschen, 30 min mit *SYBR Green I (Sigma-Aldrich; 1:10000 in 0.25x TBE verdünnt)* gefärbt und mittels UV-*trans*-Illumination analysiert.

Die quantitative Auswertung der Intensitäten verschiedener Gelbereiche erfolgte mit *Image-Quant TL v8.2.0* von *GE Healthcare* im Modus *Analysis Toolbox*. Die Intensitäten der gewählten Bereiche wurden quantifiziert und der Hintergrund der Aufnahme subtrahiert, wobei die mittlere Pixel-Intensität eines unbenutzten Teils des Gels als Hintergrund diente.

5.3.2.7 In-Gel-Verdau

| Puffer und Reagenzien | |
|-----------------------|--|
| Gel-Verdaulösung | 0.1 mg/mL Trypsin, 100 mM NH ₄ HCO ₃ |
| Extraktionslösung | 50 % (v/v) MeCN, 0.1 % (v/v) TFA |

Um zu überprüfen, welche der sichtbaren DNA-Banden auch Histone enthalten, wurde ein In-Gel-Verdau von ausgewählten Banden durchgeführt. Hierzu wurden die Banden unter UV-Bestrahlung mithilfe eines Skalpellens ausgeschnitten, in kleine Stücke (ca. 1 mm × 1 mm) zerteilt und in ein Reaktionsgefäß pro Bande gegeben. Nun wurde 100 µL trockenes MeCN zugegeben, um dem Gel die Flüssigkeit zu entziehen, bis die Gelstücke schrumpften und opak wurden. Der Überstand wurde verworfen und die Gelstücke in 50 µL Gel-Verdaulösung resuspendiert. Nach einer Inkubation von 30 min bei RT wurden erneut 50 µL der Gel-Verdaulösung zugegeben und die Proben über Nacht bei 37 °C unter Agitation inkubiert. Anschließend wurde der Überstand in ein neues Reaktionsgefäß überführt und die Gelstücke mit 50 µL Extraktionslösung 10 min bei 37 °C inkubiert. Diese Extraktion wurde zwei Mal wiederholt, wobei der Überstand jedes Mal in das neue Reaktionsgefäß gegeben wurde. Das Lösungsmittel wurde *in vacuo* entfernt und die Proben wurden über *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 1 und Tab. 8.2). Die Auswertung erfolgte mithilfe des *Proteome Discoverers* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6) unter Verwendung einer Datenbank mit den humanen Kernhistonen HIST3H2A (Uniprot: Q7L7L0), HIST2H2BE (Uniprot: Q16778), H3C1 (Uniprot: P68431) und H4C1 (Uniprot: P62805), jeweils ohne das N-terminale Methionin. Als dynamische Modifikation wurde die Oxidation von Methionin (+15.9949 Da) eingestellt.

5.3.2.8 Verdau der fC-Histon-Addukte

| Puffer und Reagenzien | |
|-----------------------|---|
| Verdau-Puffer I | 50 mM HEPES pH 7.9, 1 mM MgCl ₂ |
| 10x Verdau-Puffer II | 500 mM TRIS pH 8.0, 10 mM CaCl ₂ |
| DNA-Verdaumix I | 0.53 mM ZnSO ₄ , 0.67 Units/µL Antarktische Phosphatase, 8.2 Units/µL Nuclease S ₁ aus <i>A. oryzae</i> |
| DNA-Verdaumix II | 0.53 mM Na ₂ H ₂ EDTA, 26.7 Units/mL Phosphodiesterase I aus <i>C. adamanteus</i> |

Um eine geeignete Protease für den Verdau der Histone zu finden, wurden 20 pmol des humanen H2A/H2B Dimers und 10 pmol des H3.1/H4-Tetramers mit je 0.1 µg Trypsin (in Verdau-Puffer I), Chymotrypsin (in 1x Verdau-Puffer II) oder Thermolysin (in 1x Verdau-Puffer II) über Nacht bei 37 °C unter Agitation inkubiert, mithilfe von *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 1 und Tab. 8.2).

Die Auswertung erfolgte mithilfe des *Proteome Discoverers* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6) unter Verwendung einer Datenbank mit den humanen Kernhistonen HIST3H2A (Uniprot: Q7L7L0), HIST2H2BE (Uniprot: Q16778), H3C1 (Uniprot: P68431) und H4C1 (Uniprot: P62805), jeweils ohne das N-terminale Methionin. Als dynamische Modifikation wurde die Oxidation von Methionin (+15.9949 Da) eingestellt.

Für den Verdau der rekonstituierten (s. 5.3.2.3) und reduzierten (s. 5.3.2.5) Nucleosomen mit formylhaltiger DNA wurden zunächst 10 pmol der Nucleosomen mit 1 M TEAB (pH 8.5) auf 70 µL aufgefüllt und ein DNA-Vorverdau durch Inkubation mit Benzonase (Endkonzentration 50 Units/mL) für 3 h bei 37 °C durchgeführt. Daraufhin wurde 10x Verdau-Puffer II (Endkonzentration 1x) zugegeben und ein erster Protein-Verdau mit 0.1 µg Thermolysin über Nacht bei 37 °C unter Agitation durchgeführt.

Der DNA-Verdau folgte einem in dieser Arbeitsgruppe entwickelten Protokoll.^[589] Die Proben wurden mit 18.75 µL DNA-Verdaumix I 3 h bei 37 °C unter Agitation inkubiert. Anschließend wurden 18.75 µL DNA-Verdaumix II zugegeben und wieder 3 h bei 37 °C inkubiert. Erneut wurde 10x Verdau-Puffer II zugegeben (Endkonzentration 1x) und ein zweiter Protein-Verdau mit 0.1 µg Thermolysin über Nacht bei 37 °C durchgeführt. Die Proben wurden mittels *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 1 und Tab. 8.2). Die Auswertung erfolgte wie oben beschrieben, wobei als zusätzliche dynamische Modifikation der nach dem DNA-Verdau übrigbleibende Rest der verdauten Addukte am Protein-N-Terminus oder den Lysin (fC: +239.0906 Da; fU: +240.0746 Da) eingestellt wurde.

5.3.2.9 DNA-Protein-Verdauprotokoll mit einer Positivkontrolle

| Puffer und Reagenzien | |
|-----------------------|--|
| DNA-Verdaumix I | 0.53 mM ZnSO ₄ , 0.67 Units/µL Antarktische Phosphatase, 8.2 Units/µL Nuclease S ₁ aus <i>Aspergillus oryzae</i> |
| DNA-Verdaumix II | 0.53 mM Na ₂ H ₂ EDTA, 26.7 Units/mL Phosphodiesterase I aus <i>Crotalus adamanteus</i> |
| Verdau-Puffer I | 50 mM HEPES pH 7.9, 1 mM MgCl ₂ |

Der in Kapitel 5.3.2.8 angewendete kombinierte DNA- und Protein-Verdau wurde an einer Positivkontrolle getestet, welche von *Dr. Ralf Strasser* von *Dynamic Biosensors* zur Verfügung gestellt wurde. Diese bestand aus einem Addukt zwischen der katalytischen α-Untereinheit der cAMP-abhängigen Proteinkinase aus *C. griseus* (Prkaca, N-terminal Glycin-Histidin vor eigentlicher Sequenz) und einer mit einem Amin-reaktiven Succinimidyl-Ester modifizierten DNA. Die Synthese der Addukte erfolgte durch *Irene Ponzo, M.Sc.* unter Verwendung des *Amine Coupling Kit 1* (*Dynamic Biosensors*). Die Aufreinigung erfolgte ebenfalls durch *Dynamic Biosensors* an einem *proFIRE*-Gerät.

Zunächst wurden 66.7 pmol der Addukt-Probe mit H₂O auf 35 µL aufgefüllt und mit 7.5 µL DNA-Verdaumix I 3 h bei 37 °C unter Agitation inkubiert. Anschließend wurden 7.5 µL DNA-Verdaumix II zugegeben und erneut 3 h bei 37 °C inkubiert. Für den Protein-Verdau wurden die Proben zunächst mit 500 mM DTT (Endkonzentration 5 mM) versetzt und 20 min bei 60 °C unter Agitation inkubiert. Danach wurden die Proben auf RT abgekühlt und nach Zugabe von 1 M IAA (Endkonzentration 20 mM) 20 min bei RT und unter Lichtausschluss sowie Agitation inkubiert. Darauf folgte eine erneute Zugabe von 500 mM DTT (Endkonzentration 10 mM, zuvor zugegebenes DTT nicht mit eingerechnet) sowie Inkubation für 20 min bei RT unter Agitation. Die Proben wurden daraufhin mit Verdau-Puffer I auf 100 µL aufgefüllt und es wurde 1 µg/µL Trypsin (Endkonzentration 10 ng/µL) zugegeben, bevor die Proben über Nacht bei 37 °C unter Agitation inkubiert wurden.

Die Probe wurden mittels *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 2 und Tab. 8.3).

Die Auswertung erfolgte mithilfe des *Proteome Discoverers* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6) unter Verwendung der Sequenz der modifizierten Prkaca (Uniprot: P25321, N-terminal Glycin-Histidin vor eigentlicher Sequenz). Als dynamische Modifikationen wurden die Oxidation von Methionin (+15.9949 Da) sowie der nach Verdau übrig bleibende Rest des DNA-Protein-Crosslinkers am Protein-N-Terminus sowie an Lysinen (+353.1661 Da) eingestellt. Als statische Modifikation wurde die durch Reduktion und Alkylierung hervorgerufene Carbamidomethylierung der Cysteine (+57.0125 Da) ausgewählt.

5.3.2.10 Deformylierungs-Experimente

| Puffer und Reagenzien | |
|-----------------------|--|
| Histon-Dialysepuffer | 20 mM HEPES pH 7.9, 2 M NaCl |
| Verdau-Puffer I | 50 mM HEPES pH 7.9, 1 mM MgCl ₂ |

Um die Möglichkeit einer Deformylierung von fC zu C über die Bildung von Addukten mit Histonen zu überprüfen, wurden DNA-Quantifizierungs-Experimente mit Nukleosomen durchgeführt.

Dafür wurde zunächst in Zusammenarbeit mit *Dr. Angie Kirchner* mittels PCR, wie in Kapitel 5.3.2.1 beschrieben, 601 DNA hergestellt, welche entweder ¹⁵N₂-isotopenmarkiertes fC (601_{15N2fc}) oder *super-heavy*-dC (¹³C₉¹⁵N₃-markiertes dC, 601_{shdC}) enthält. Mit diesen Strängen wurden Nukleosomen rekonstituiert, gelelektrophoretisch analysiert (s. 5.3.2.3 sowie 5.3.2.4) und danach 2 h bei 37 °C inkubiert. Anschließend wurde die DNA von *Dr. Katharina Iwan* verdaut und an einem *TripleQuad*-Massenspektrometer quantitativ analysiert.

Außerdem wurde der Effekt einer Methylierung der Histon-Lysine untersucht. Dafür wurden das H2A/H2B-Dimer und das H3.1/H4-Tetramer aus dem *EpiMark Nucleosome Assembly Kit* (*New England Biolabs*) nach einem in der Literatur beschriebenen Protokoll reduktiv alkylert.^[558]

Hierfür wurden Dimer und Tetramer zunächst in *Slide-A-Lyzer MINI 2K MWCO*-Dialysevorrichtungen (*Thermo Fisher Scientific*) überführt und bei 4 °C gegen 100 mL Histon-Dialysepuffer dialysiert. Nach 2 h wurde die Dialyse-Lösung erneuert und über Nacht weiter dialysiert. Im Anschluss wurde zu je 7.5 µL 20 µM Dimer bzw. 10 µM Tetramer 0.5 µL 200 mM Dimethylaminoboran und 0.5 µL 400 mM Formaldehyd gegeben und die Lösung 2 h bei 4 °C inkubiert. Anschließend wurden wieder 0.5 µL 200 mM Dimethylaminoboran und 0.5 µL 400 mM Formaldehyd zugegeben und erneut 2 h bei 4 °C inkubiert, bevor 0.25 µL 200 mM Dimethylaminoboran zugegeben und über Nacht bei 4 °C inkubiert wurde. Um das Formaldehyd zu quenchen, wurde 1 µL 1 M TRIS (pH 8.0) zugegeben und 2 h bei 4 °C inkubiert. Daraufhin wurde ein weiterer Dialyse-Schritt, wie zuvor beschrieben, durchgeführt. Je 1.5 pmol des Dimers sowie Tetramers wurden mit Verdau-Puffer I auf 50 µL aufgefüllt und nach Zugabe von 0.1 µg/µL Trypsin (Endkonzentration 1 ng/µL) über Nacht bei 37 °C unter Agitation verdaut. Die Proben wurden mithilfe von *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 2 und Tab. 8.3). Die Auswertung erfolgte mithilfe des *Proteome Discoverers* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6) unter Verwendung einer Datenbank mit den humanen Kernhistonen HIST3H2A (Uniprot: Q7L7L0), HIST2H2BE (Uniprot: Q16778), H3C1 (Uniprot: P68431) und H4C1 (Uniprot: P62805), jeweils ohne das N-terminale Methionin. Als dynamische Modifikation wurde die Oxidation von Methionin (+15.9949 Da) sowie mono- (+14.0157 Da) und dimethyliertes (+28.0313 Da) Lysin eingestellt.

Da bei der beschriebenen Methylierung sehr viele dimethylierte Lysine generiert wurden, wurde das Methylierungs-Protokoll angepasst. Zum einen wurden lediglich 50 mM Dimethylaminoboran und 100 mM Formaldehyd (ca. 5 Äquivalente relativ zu Lysinen) verwendet und anschließend 4 h bei 4 °C inkubiert. Zudem wurden die Reagenzien nicht ein zweites Mal zugegeben, sondern die Proben sofort mit 0.25 µL 50 mM Dimethylaminoboran versetzt, 2 h bei 4 °C inkubiert und anschließend mit 0.5 µL 1 M TRIS für 30 min bei 4 °C gequencht. Nach der finalen Dialyse wurden Nukleosomen mit 601_{shc} sowie 601_{15N2FC}, wie in Kapitel 5.3.2.3 beschrieben, rekonstituiert, wobei auch Proben mit nicht-methylierten Histonen angesetzt wurden. Ein Teil dieser Proben wurde von *Dr. Katharina Iwan* zur DNA-Quantifizierung verwendet.

Die gebildeten Nukleosomen aus Histonen mit und ohne Methylierung wurden zudem mit Trypsin verdaut. Dies erfolgte durch Auffüllen der Proben auf 50 µL mit Verdau-Puffer I, Zugabe von 0.1 µg/µL Trypsin (Endkonzentration 2 ng/µL) und Inkubation über Nacht bei 37 °C. Die verdauten Proben wurden mittels *StageTips* aufgereinigt (s. Kap. 5.6.1) und wie oben beschrieben mittels MS² analysiert, wobei als zusätzliche dynamische Modifikation Formyllysin (+27.9949 Da), sowie für die methylierten Proben Methyllysin (+14.0157 Da), Dimethyllysin (+28.0313 Da) und Formylmethyllysin (+42.0106 Da) eingestellt wurde.

5.4 Crosslinking von TET3 und GDH mithilfe von DSSO

In Zusammenarbeit mit *Dr. Michael Stadlmeier* wurde mithilfe von XL-MS eine mögliche direkte Interaktion zwischen TET3 und GDH untersucht.

5.4.1 Crosslinking-Tests mit BSA

| Puffer und Reagenzien | |
|-----------------------|--|
| Verdau-Puffer I | 50 mM HEPES pH 7.9, 1 mM MgCl ₂ |
| SEC-Puffer | 50 mM K-Phosphat pH 7.4, 150 mM NaCl |

Zunächst wurde das Crosslinking mit DSSO in der Arbeitsgruppe etabliert. Hierfür wurde BSA quervernetzt und die so entstehenden Intralinks mittels XL-MS analysiert.

Dafür wurde 1 µg/µL BSA in 20 mM HEPES pH 7.9 mit 50 mM DSSO (gelöst in trockenem DMSO, Endkonzentration 2.5 mM) versetzt und 1 h bei Raumtemperatur inkubiert. Anschließend wurde unreaktiver Crosslinker durch Zugabe von 100 mM NH₄HCO₃ (Endkonzentration 20 mM) und Inkubation für 5 min bei RT gequenchet. Daraufhin erfolgte die Reduktion von Disulfidbrücken, indem 1 M DTT (Endkonzentration 20 mM) zugegeben und die Probe 1 h bei 60 °C unter Agitation inkubiert wurde, sowie die Alkylierung der freien Cysteine durch Zugabe von 1 M IAA (Endkonzentration 40 mM) und Inkubation für 30 min bei RT unter Agitation und Lichtausschluss. Um das Protein zu fällen, wurden vier Probenvolumen kaltes Aceton zugegeben und die Probe über Nacht bei -20 °C inkubiert. Im Anschluss wurde 10 min bei 15000 × g und 4 °C zentrifugiert, der Überstand verworfen und der Rückstand an Luft getrocknet, bevor letzterer in 150 µL Verdau-Puffer I resuspendiert und das Protein durch Zugabe von 1 µg Trypsin pro 20 µg BSA und Inkubation über Nacht bei 37 °C verdaut wurde. Die Peptide wurden mittels *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 3 und Tab. 8.4). Die Auswertung erfolgte mithilfe von *MeroX* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.7 und 8.8) sowie *XlinkX* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6) unter Verwendung einer Datenbank mit der Sequenz von BSA (Uniprot: P02769) sowie den Sequenzen der 38 sogenannten *dust/contact proteins* des *common Repository of Adventitious Proteins (cRAP)* von *The Global Proteome Machine*^[562] (Uniprot: P04745, P13645, P35527, P04264, P35908, Q15323, Q14532, O76011, Q92764, O76013, O76014, O76015, O76009, Q14525, Q14533, Q9NSB4, P78385, Q9NSB2, P78386, O43790, O77727, P02534, P25690, P02539, P15241, P25691, P02444, P02445, P02443, P02441, Q02958, P02438, P02439, P02440, P08131, P26372, O82803, P15252).

Im Falle der Fraktionierung mithilfe von SEC wurden 40 µg von quervernetztem und verdautem BSA nach der *StageTip*-Aufreinigung in 10 µL SEC-Puffer resuspendiert und 10 min bei 15000 × g zentrifugiert, um eventuelle unlösliche Komponenten abzutrennen. Die Fraktionierung erfolgte an einem *ÄKTApurifier*-System unter Verwendung einer *Superdex Peptide 3.2/300*-Säule (beide bezogen von *GE Healthcare*).

Der Überstand der zentrifugierten Peptid-Probe wurde über einen 20 µL fassenden Loop auf die Säule geladen und mit einer Flussrate von 50 µL/min aufgetrennt, wobei als mobile Phase SEC-Puffer verwendet wurde, welcher zuvor mithilfe einer Zellulosemischester-Membran mit einer Porengröße von 0.2 µm (*GE Healthcare*) filtriert und anschließend *in vacuo* entgast wurde. Die Absorption bei 215 nm wurde gemessen und das Eluat wurde in Fraktionen von 100 µL aufgefangen. Ausgewählte Fraktionen wurden *in vacuo* eingeeengt, mittels *StageTips* aufgereinigt (s. Kap. 5.6.1) und wie oben beschrieben massenspektrometrisch analysiert.

5.4.2 Methoden für den Nachweis einer Interaktion zwischen TET3 und GDH

5.4.2.1 Expression der Proteine in humanen Zellen und Crosslinking in Zellen

| Puffer | |
|-------------|--|
| RIPA-Puffer | 10 mM TRIS pH 7.75, 150 mM NaCl, 0.5 mM Na ₂ H ₂ EDTA, 1 % Triton X-100, 1 % Natriumdesoxycholat |
| Lyse-Puffer | 50 mM HEPES pH 7.9, 150 mM NaCl, 1 mM Na ₂ H ₂ EDTA, 1 % Triton X-100 |
| Puffer A | 10 mM HEPES pH 7.9, 10 mM NaCl, 1.5 mM MgCl ₂ , 1 mM DTT |
| Puffer C | 20 mM HEPES pH 7.9, 420 mM NaCl, 0.2 mM Na ₂ H ₂ EDTA, 1 mM DTT |

Um eine direkte Interaktion zwischen TET3 und GDH nachzuweisen, wurden die beiden Proteine zunächst in HEK293T-Zellen überexprimiert. Die Zellen wurden in Petrischalen mit einem Durchmesser von 15 cm und einer für adhärenzte Zellkulturen speziell behandelten Oberfläche (*Sarstedt*) kultiviert. Als Kulturmedium wurde dabei RPMI-Medium^[590] (*Sigma-Aldrich*) mit 10 % FBS (*Sigma-Aldrich*), 1 % 200 mM Alaninyl-Glutamin (*Sigma-Aldrich*, in 0.85 % NaCl) sowie 1 % *Pen-Strep* (*Sigma-Aldrich*, 10 U/µL Penicillin und 10 µg/µL Streptomycin in 0.9 % NaCl) verwendet. Die Transfektion erfolgte mit *jetPrime* (*Polyplus transfection*) nach dem Herstellerprotokoll, wobei 10 µg Plasmid pro Petrischale verwendet wurden. Es wurden Plasmide verwendet, die für eines der Fusionsproteine eGFP-TET3_{cd}, eGFP-TET3_{fl} oder GDH-FLAG codierten (s. Kap. 8.1). Die eGFP-TET3-Plasmide wurden von der Forschungsgruppe um *Prof. Dr. Heinrich Leonhardt* zur Verfügung gestellt, wohingegen das GDH-FLAG-Plasmid kommerziell bei *GENEART* erworben wurde. Bei Bedarf wurden die Plasmide in chemisch kompetenten *E. coli*-Zellen (*5-α*, *New England Biolabs*) amplifiziert und mithilfe des *ZymoPURE II Plasmid Midiprep Kit* (*Zymo Research*) nach Herstellerprotokoll extrahiert und aufgereinigt.

Die HEK293T-Zellen wurden für 5 h transfiziert, wobei bei manchen Versuchen eine Cotransfektion eines TET3- und des GDH-Plasmids durchgeführt wurde, während bei anderen Experimenten getrennte Kulturen angelegt wurden. Nach der Transfektion wurde das Medium gewechselt und bei einigen Versuchen 1 M Natriumbutyrat (Endkonzentration 5 mM) zum neuen Medium gegeben, um eine bessere Expression der Plasmide zu erreichen.

Daraufhin wurden die Zellen zwei Tage lang kultiviert und anschließend geerntet sowie zwei Mal mit PBS gewaschen, wobei nach jedem Waschschrift 3 min bei $400 \times g$ zentrifugiert und der Überstand abgenommen wurde. Bei manchen Versuchen mit Cotransfektion erfolgte nun das Crosslinking durch Zugabe von 50 mM DSSO (gelöst in trockenem DMSO, Endkonzentration 10 mM) und Inkubation für 15 min bei RT unter Agitation. In diesen Fällen wurde die Reaktion durch Zugabe von 500 mM NH_4HCO_3 (Endkonzentration 20 mM) und Inkubation für 5 min bei RT unter Agitation gestoppt, bevor die Zellen zwei Mal mit 500 μL PBS gewaschen wurden, um überschüssiges DSSO zu entfernen.

Die Lyse erfolgte auf unterschiedliche Art und Weise: bei einigen Versuchen wurde komplettes Zelllysat hergestellt, indem zum Zellpellet 1 mL RIPA-Puffer^[566] bzw. Lyse-Puffer zugegeben wurde, wobei dem jeweiligen Puffer kurz vorher 25x *cOmplete EDTA-free* Protease-Inhibitor-Cocktail (*Roche*, Endkonzentration 1x) und Benzonase (*Merck Millipore*, Endkonzentration 90 U/mL) beigefügt wurden. Die Suspension wurde in einen *Dounce*-Handhomogenisator (*Wheaton*) gegeben und durch zehnmaliges auf und ab bewegen eines Typ B-Pistills auf Eis homogenisiert, bevor die Probe 1 h auf Eis inkubiert wurde. Das Lysat wurde 10 min bei $15000 \times g$ und 4°C zentrifugiert und der Überstand in ein neues Reaktionsgefäß überführt. Für andere Experimente wurde Zellkernlysat verwendet, welches nach einem modifizierten Protokoll von *Dignam et al.* hergestellt wurde.^[591] Dabei wurden zunächst die Zellkerne isoliert, indem das Zellpellet in 5 Zellvolumen Puffer A (+ 25x *cOmplete EDTA-free* Protease-Inhibitor-Cocktail, Endkonzentration 1x) resuspendiert und 10 min auf Eis inkubiert wurde. Die Suspension wurde anschließend 5 min bei $400 \times g$ und 4°C zentrifugiert, der Überstand abgenommen und das Pellet in 2 Zellvolumen Puffer A resuspendiert. Die Probe wurde, wie oben beschrieben, mit einem *Dounce*-Homogenisator homogenisiert und 15 min bei $3000 \times g$ zentrifugiert. Der Überstand, der das Cytoplasma darstellte, wurde verworfen und das Zellkern-Pellet entweder, wie oben beschrieben, mit DSSO behandelt und mit RIPA-Puffer lysiert, oder in zwei Zellvolumen Puffer C (+ 25x *cOmplete EDTA-free* Protease-Inhibitor-Cocktail, Endkonzentration 1x; + Benzonase, Endkonzentration 90 U/mL) resuspendiert, durch *Dounce*-Homogenisation (s. o.) lysiert und 1 h bei 4°C in einem Rotationsmischer inkubiert. Nach Zentrifugation für 10 min bei $15000 \times g$ und 4°C wurde das Zellkernlysat abgenommen und in ein neues Reaktionsgefäß überführt.

Die Proteinkonzentration der Lysate wurde mithilfe eines Bradford-Assays bestimmt. Dazu wurde 1 μL der Proben mit 49 μL H_2O verdünnt und in einer Küvette mit 950 μL einer Lösung von 20 % *Protein Assay Dye Reagent-Konzentrat* (*Bio-Rad*) gemischt. Nach 5 min wurde mithilfe eines Spektrophotometers die Extinktion bei 595 nm gemessen und anhand einer im Gerät gespeicherten Eichgerade die Proteinkonzentration berechnet.

5.4.2.2 Anreicherung von eGFP-TET3 bzw. GDH-FLAG und Crosslinking in Lysat

| Puffer | |
|---------------------|---|
| FLAG-Waschpuffer I | 50 mM HEPES pH 7.4, 150 mM NaCl |
| FLAG-Waschpuffer II | 50 mM HEPES pH 7.4, 500 mM NaCl |
| GFP-Waschpuffer I | 10 mM HEPES pH 7.4, 150 mM NaCl, 0.5 mM Na ₂ H ₂ EDTA |
| GFP-Waschpuffer II | 10 mM HEPES pH 7.4, 500 mM NaCl, 0.5 mM Na ₂ H ₂ EDTA |
| Verdau-Puffer I | 50 mM HEPES pH 7.9, 1 mM MgCl ₂ |

Um die Probenkomplexität zu verringern, wurden verschiedene Ansätze der Aufreinigung mittels Affinitäts-Tags verfolgt. Nachfolgend ist das Protokoll mit den durch getrennte Transfektion erhaltenen eGFP-TET3- und GDH-FLAG-Lysaten beschrieben. Bei den cotransfizierten und bereits quervernetzten Proben wurde lediglich eine Aufreinigung mittels des eGFP-Tags durchgeführt.

Das GDH-FLAG-Lysat wurde zunächst über magnetische *Anti-FLAG M2*-Partikel (Sigma-Aldrich) aufgereinigt. Dafür wurden 100 µL der Partikel-Suspension drei Mal mit 250 µL FLAG-Waschpuffer I gewaschen, das GDH-FLAG-Zelllysate zu den Partikeln gegeben und 1 h bei 4 °C im Rotationsmischer inkubiert. Der Überstand wurde abgenommen und die Festphase einmal mit je 250 µL FLAG-Waschpuffer I, drei Mal mit FLAG-Waschpuffer II und ein weiteres Mal mit FLAG-Waschpuffer I gewaschen. Die Elution erfolgte durch Zugabe von 500 µL FLAG-Peptid (*Sigma-Aldrich*) in FLAG-Waschpuffer I und Inkubation für 15 min bei 4 °C im Rotationsmischer. Der Überstand wurde in ein neues Reaktionsgefäß überführt.

Das Lysat mit überexprimiertem eGFP-TET3 (cd bzw. fl) wurde seinerseits über magnetische *GFP-Trap*-Agarose-Partikel (*Chromotek*) aufgereinigt. Hierfür wurden 50 µL der Partikel-Suspension, wie bereits für die FLAG-Aufreinigung beschrieben, vorbereitet, mit eGFP-TET3-Zelllysate inkubiert und gewaschen, wobei als Puffer GFP-Waschpuffer I und II verwendet wurden. Anschließend wurden die Partikel in 100 µL GFP-Waschpuffer I resuspendiert und 150 µL des aufgereinigten GDH-FLAG zugegeben. Diese Suspension wurde 30 min auf Eis inkubiert, wobei darauf geachtet wurde, dass die Partikel sich nicht absetzen. Der Überstand wurde abgenommen und die Festphase drei Mal mit GFP-Waschpuffer I gewaschen. Anschließend wurden die Partikel in 237.5 µL GFP-Waschpuffer I resuspendiert und 50 mM DSSO (Endkonzentration 2.5 mM) zugegeben, bevor 1 h auf Eis inkubiert wurde. Die Reaktion wurde durch Zugabe von 500 mM NH₄HCO₃ (Endkonzentration 20 mM) und Inkubation für 5 min auf Eis gestoppt. Der Überstand wurde abgenommen und die Partikel in 100 µL Verdau-Puffer I resuspendiert. Nun erfolgte die Reduktion der Disulfidbrücken durch Zugabe von 100 mM DTT (Endkonzentration 5 mM) und Inkubation für 30 min bei RT unter Agitation, sowie die Alkylierung der Cysteine mit 100 mM IAA (Endkonzentration 10 mM) und Inkubation für 10 min bei RT unter Agitation und Lichtausschluss. Die Suspension wurde mit Verdau-Puffer I auf 150 µL aufgefüllt und 1 µg/µL Trypsin (Endkonzentration 3.33 ng/µL) zugegeben.

Die Proben wurden über Nacht bei 37 °C unter Agitation inkubiert und der Überstand daraufhin in ein neues Reaktionsgefäß gegeben. Die Partikel wurden in 100 µL Verdau-Puffer I resuspendiert und dieser Überstand ebenfalls in das neue Reaktionsgefäß gegeben. Nun erfolgte ein zweiter Verdau, indem 2.5 µL 1 µg/µL Trypsin zugegeben und 3 h bei 37 °C unter Agitation inkubiert wurde.

Alle Ansätze wurden mittels *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 4 und Tab. 8.4). Die Analyse der Crosslinks erfolgte mithilfe von *MeroX* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.7 und 8.8) unter Verwendung einer Datenbank mit den in Kapitel 8.1 aufgeführten Sequenzen, wobei je nach Plasmid die Sequenz für eGFP-Tet3_{cd} bzw. eGFP-TET3_{fi} eingesetzt wurde. Zudem wurde *Proteome Discoverer* verwendet (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6), um die in den Proben vorliegenden Proteine zu identifizieren. Als dynamische Modifikation wurde hierbei die Oxidation von Methionin (+15.9949 Da) und als statische Modifikation die Carbamidomethylierung von Cystein (+57.0215 Da) eingestellt.

5.4.2.3 Anreicherung mittels SEC

| Puffer | |
|-----------------|--|
| Verdau-Puffer I | 50 mM HEPES pH 7.9, 1 mM MgCl ₂ |

Mit einem Teil der mit den Plasmiden für eGFP-TET3_{fi} und GDH-FLAG cotransfizierten und mit Lyse-Puffer lysierten Zellen (s. 5.4.2.1) wurde zusätzlich eine Anreicherung der Crosslinks durch SEC versucht. Hierfür wurden 120 µg des Lysats mit Verdau-Puffer I auf 1.2 µg/µL verdünnt und mit 50 mM DSSO (gelöst in trockenem DMSO, Endkonzentration 2.5 mM) für 30 min bei 4 °C quervernetzt und durch Zugabe von 500 mM NH₄HCO₃ (Endkonzentration 20 mM) gequenchet. Anschließend wurde 1 M DTT (Endkonzentration 20 mM) zugegeben und 1 h bei 60 °C unter Agitation inkubiert, bevor die Probe mit 1 M IAA (Endkonzentration 40 mM) versetzt und 30 min bei RT unter Agitation und Lichtausschluss inkubiert wurde. Die Proteine wurden durch Zugabe von vier Probenvolumen kaltem Aceton und Inkubation über Nacht bei -20 °C gefällt, durch Zentrifugation für 10 min bei 15000 × *g* abgetrennt, der Überstand abgenommen und der Rückstand an Luft getrocknet. Anschließend wurde der Rückstand in 150 µL Verdau-Puffer I resuspendiert und 1 µg/µL Trypsin (Endkonzentration 20 ng/µL) zugegeben und über Nacht bei 37 °C inkubiert. Die Aufreinigung erfolgte mittels *StageTips* (s. Kap. 5.6.1), bevor die SEC, wie in Kapitel 5.4.1 beschrieben, durchgeführt wurde. Ausgewählte Fraktionen wurden *in vacuo* eingeeengt, mittels *StageTips* (s. Kap. 5.6.1) aufgereinigt und mittels MS² analysiert (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 4 und Tab. 8.4). Die Auswertung erfolgte mithilfe von *MeroX* (s. Kap. 5.7.2, sowie 8.3, Tab. 8.7 und 8.8) unter Verwendung einer Datenbank der Sequenzen von GDH-FLAG (s. Kap. 8.1.1) und eGFP-TET3_{fi} (s. Kap. 8.1.2).

5.5 Methoden zur Etablierung eines neuen Protein-Protein-Crosslinkers

Die Anwendung eines von *Dr. Michael Stadlmeier* entwickelten und synthetisierten Reagenzes zur Identifikation und Anreicherung von Protein-Protein-Wechselwirkungen mittels Massenspektrometrie wurde gemeinsam mit *Dr. Michael Stadlmeier* mithilfe von verschiedenen Methoden optimiert.

5.5.1 Optimierung der Identifikation von Protein-Protein-Crosslinks

| Puffer | |
|-------------------------|--------------------------------|
| Crosslink-Puffer | 10 mM HEPES pH 7.4 |
| Resuspendierungs-Puffer | 10 mM HEPPS pH 8.5, 6 M Gua-Cl |
| Verdünnungs-Puffer | 10 mM HEPPS pH 8.5 |

Das Protokoll der Crosslinking-Reaktion wurde im Vergleich zu dem in Kapitel 5.4.1 beschriebenen Ablauf an verschiedenen Stellen optimiert. Im Folgenden ist das resultierende optimierte Protokoll aufgeführt, die optimierten Parameter sind in Kapitel 3.3.1 beschrieben.

BSA wurde auf eine Konzentration von 1 µg/µL in Crosslink-Puffer angelöst. Eine frisch hergestellte Lösung von DSSO bzw. cliXlink (20 mM in trockenem DMSO) wurde zu den Proben gegeben (Crosslinker-Endkonzentration 1 mM) und 45 min bei RT unter Agitation inkubiert. Um die Reaktion zu stoppen, wurde 1 M NH₄HCO₃ zugegeben (Endkonzentration 20 mM) und 15 min bei Raumtemperatur inkubiert. Anschließend wurde 500 mM DTT zugegeben (Endkonzentration 5 mM) und die Proben 30 min bei 60 °C unter Agitation inkubiert, bevor 1 M IAA zugegeben (Endkonzentration 20 mM) und 30 min bei RT unter Lichtausschluss und Agitation inkubiert wurde. Das Protein wurde durch Zugabe von vier Probenvolumen Aceton und Inkubation bei -20 °C über Nacht gefällt. Darauf folgte eine Zentrifugation für 10 min bei 15000 × *g* und 4 °C, wonach der Überstand verworfen und der Rückstand bei RT an Luft getrocknet wurde. Durch Zugabe von Resuspendierungs-Puffer wurde das Protein auf eine Konzentration von 13.3 µg/µL angelöst und 15 min bei 37 °C unter Agitation inkubiert, bevor durch Zugabe von Verdünnungs-Puffer die Proteinkonzentration auf 1 µg/µL und die Gua-Cl-Konzentration auf 0.45 M eingestellt wurde. Die Proben wurden durch Zugabe von 1 µg/µL Trypsin/Lys-C (Endkonzentration 25 ng/µL) und Inkubation über Nacht bei 37 °C unter Agitation verdaut. Die Aufreinigung erfolgte mittels *ZipTips* (s. Kap. 5.6.2), bevor die Proben mittels MS² (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 5 und Tab. 8.5) analysiert wurden, wobei verschiedene Geräteparameter variiert wurden. Die Auswertung erfolgte mithilfe von *MeroX* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.7 und 8.8) unter Verwendung der in Kapitel 5.4.1 beschriebenen Datenbank.

5.5.2 Untersuchung der CuAAC-Reaktion auf Peptidebene

Zunächst wurde die CuAAC-Reaktion an quervernetzten und tryptisch verdauten Peptiden untersucht. Dies geschah erst mithilfe einer auf Fluoreszenz basierenden Untersuchung als Testsystem und anschließend mittels Massenspektrometrie.

5.5.2.1 Fluoreszenz-basierte Untersuchung mithilfe von Coumarin-Azid

| Puffer | |
|--------------|-------------------------|
| CuAAC-Puffer | 10 mM K-Phosphat pH 7.4 |

Zunächst wurde die CuAAC-Reaktion an einem Testsystem untersucht. Dazu wurden quervernetzte BSA-Peptide, wie unter 5.5.1 beschrieben, hergestellt, wobei die Aufreinigung über *StageTips* erfolgte (s. Kap. 5.6.1). Das Peptidpellet wurde auf eine Konzentration von 0.5 µg/µL in CuAAC-Puffer resuspendiert. Die CuAAC-Reaktion wurde in 100 µL in Triplikaten durchgeführt; die Endkonzentrationen der Reagenzien beziehen sich daher auf dieses Volumen. Als Reaktionsgefäß wurde eine Mikrotiterplatte (*Nunc F96 MicroWell, black* von *Thermo Fisher Scientific*) verwendet. 10 µg der Peptide wurden mit 10 mM 3-Azido-7-hydroxycoumarin (hergestellt von *Simon Veth, M.Sc.*, gelöst in DMSO) versetzt (Endkonzentration 100 µM) und mit CuAAC-Puffer auf 98 µL aufgefüllt. 10 mM CuSO₄ und 50 mM THPTA wurden zu gleichen Teilen gemischt und zu den Proben gegeben (Endkonzentrationen 50 µM (CuSO₄), 250 µM (THPTA)). Schließlich wurde die Reaktion durch Zugabe von 500 mM Natriumascorbat (Endkonzentration 5 mM) gestartet. Um später die Hintergrund-Fluoreszenz abziehen zu können, wurden zusätzlich Proben ohne BSA-Peptide vorbereitet. Das Fortschreiten der CuAAC-Reaktion wurde über eine kinetische Fluoreszenz-Messung an einem Mikrotiterplatten-Lesegerät verfolgt. Dabei wurde die Fluoreszenz-Intensität ca. alle 5 min mit den in Tabelle 5.3 aufgeführten Einstellungen gemessen, bis die Intensität nicht mehr weiter anstieg.

Tabelle 5.3: Einstellungen für die Verfolgung der CuAAC-Reaktion mittels Fluoreszenz-Kinetik.

| Parameter | Einstellung |
|-----------------------------------|----------------------------|
| Filter für Anregung | 360 nm (Bandbreite: 35 nm) |
| Filter für Emission | 465 nm (Bandbreite: 35 nm) |
| <i>Gain</i> | <i>manual: 40</i> |
| <i>Lag time</i> | 0 µs |
| <i>Integration time</i> | 40 µs |
| <i>Number of reads</i> | 10 |
| <i>Time between move and read</i> | 2 ms |
| <i>Read mode</i> | <i>Top</i> |
| <i>Filter switch mode</i> | <i>per Plate</i> |
| <i>Mirror selection</i> | <i>50% mirror</i> |

Zwischen den Messungen wurde die Mikrotiterplatte mit einer Amplitude von 5 mm horizontal geschüttelt. Für die grafische Auswertung wurde der Mittelwert der Triplikate der Leerproben bestimmt und von den Werten der Proben subtrahiert. Anschließend wurden der Mittelwert sowie die Standardabweichung der Hintergrund-korrigierten Proben-Triplikate berechnet und die Fluoreszenz-Intensität gegen die Zeit aufgetragen.

5.5.2.2 Anreicherung und massenspektrometrische Analyse

Um einen Effekt der Anreicherung auf die identifizierte Zahl von Crosslinks zu untersuchen, wurde eine analog zu Kapitel 5.5.1 hergestellte, mit cliXlink funktionalisierte BSA-Peptidmischung, wie in Kapitel 5.5.2.1 beschrieben, mittels CuAAC funktionalisiert und anschließend erneut mithilfe von *StageTips* aufgereinigt (s. Kap. 5.6.1). Als Azid wurde in diesem Fall jedoch Biotin-Azid_{Ester} (s. Schema 3.3) verwendet. Für die Anreicherung wurden je Probe 100 μ L magnetische und mit Streptavidin funktionalisierte Partikel (*Dynabeads MyOne Streptavidin T1* von *Thermo Fisher Scientific*) drei Mal mit 500 μ L PBS gewaschen und in 100 μ L PBS resuspendiert. Die entsalzten Proben wurden in 400 μ L PBS resuspendiert, zu den gewaschenen Partikeln gegeben und 1 h bei RT in einem Rotationsmischer inkubiert. Anschließend wurden die Partikel drei Mal mit 500 μ L PBS gewaschen, bevor sie in 500 μ L PBS resuspendiert wurden. Durch Zugabe von 1 M TCEP (Endkonzentration 5 mM) und Inkubation für 30 min bei RT unter Agitation wurden die Crosslinks von der Festphase abgespalten. Der Überstand wurde in ein neues Reaktionsgefäß überführt, die Partikel mit 100 μ L PBS gespült und dieser Überstand ebenfalls in das neue Reaktionsgefäß gegeben. Der pH-Wert dieser Lösung wurde überprüft und gegebenenfalls mit mehr Puffer auf einen neutralen bis leicht basischen Wert eingestellt. Das durch die Reduktion entstandene Thiol wurde alkyliert, indem 1 M IAA (Endkonzentration 20 mM) zugegeben und 45 min bei RT unter Lichtausschluss und Agitation inkubiert wurde. Die Peptide wurden mithilfe von *StageTips* aufgereinigt (s. Kap. 5.6.1) und mittels MS² (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 5 und Tab. 8.5: optimierte Methode) analysiert. Die Auswertung erfolgte mit *MeroX* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.7 und 8.8) unter Verwendung der in Kapitel 5.4.1 beschriebenen Datenbank.

Da mit dieser Analyse keine Crosslinks identifiziert werden konnten, wurde in einem weiteren Experiment ein möglicherweise negativer Effekt der CuAAC-Reagenzien auf die Crosslinks bzw. generell Peptide untersucht. Hierfür wurde die CuAAC-Reaktion ohne das Azid durchgeführt und die Peptide anschließend wie oben beschrieben aufgereinigt und massenspektrometrisch untersucht. Die Auswertung erfolgte zusätzlich mit *Proteome Discoverer* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6), wobei als dynamische Modifikation die Oxidation von Methionin (+15.9949 Da) sowie als statische Modifikation die Carbamidomethylierung von Cystein (+57.0125 Da) eingestellt wurde.

Um das Protokoll der Anreicherung zu verifizieren, wurde ein vereinfachtes System verwendet, indem BSA statt mit cliXlink mit 4-Pentinsäure-*N*-Hydroxysuccinimid umgesetzt wurde. Dadurch entstehen viele modifizierte Peptide, die Typ 0-Crosslinks ähneln. Die Funktionalisierung mit Biotin-Azid_{Ester} sowie die Anreicherung erfolgten wie oben beschrieben. Ein Teil der Probe wurde dabei zum Vergleich nicht angereichert. Die Proben wurden mittels *StageTips* aufgereinigt und massenspektrometrisch analysiert. Die Auswertung erfolgte mit *Proteome Discoverer* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.6), wobei als dynamische Modifikationen die Oxidation von Methionin (+15.9949 Da), die freie Pentinsäure-Modifikation (+80.0262 Da) und die mit Biotin-Azid_{Ester} funktionalisierte sowie reduzierte und alkylierte Pentinsäure-Modifikation (+240.0681 Da) eingestellt wurde. Als statische Modifikation wurde die Carbamidomethylierung von Cystein (+57.0125 Da) gewählt.

5.5.3 CuAAC-Reaktion auf Proteinebene und Anreicherung

| Puffer | |
|-------------------------|--------------------------------|
| Resuspendierungs-Puffer | 10 mM HEPPS pH 8.5, 6 M Gua-Cl |
| Verdünnungs-Puffer | 10 mM HEPPS pH 8.5 |
| Wasch-Puffer | 50 mM K-Phosphat pH 7.4 |

Das Protokoll mit CuAAC-Reaktion auf Proteinebene wurde an mehreren Stellen optimiert. So wurde der Effekt einer Acetonfällung und einer Umpufferung mittels MWCO-Säule als Aufreinigungsmethoden nach den verschiedenen Schritten verglichen. Besonders die CuAAC-Reaktion wurde verbessert, unter anderem mithilfe eines vereinfachten Testsystems, indem quervernetztes BSA unter verschiedenen Bedingungen mit 2-Azidoethanol funktionalisiert und anschließend ohne Anreicherung verdaut und analysiert wurde. Auch eine Analyse der Löslichkeit des mit Biotin-Azid_{Ester} reagierte BSAs wurde durchgeführt. Schließlich wurde die Eignung von Biotin-Azid_{Ester} und Biotin-Azid_{Amid} (s. Schema 3.3 sowie Abb. 3.23) verglichen. Die folgenden Bedingungen stellen das bereits optimierte Protokoll dar, die genauen veränderten Parameter werden in Kapitel 3.3.2.2 diskutiert.

Die Proben wurden zunächst, wie in Kapitel 5.5.1 beschrieben, quervernetzt. Nach der ersten Acetonfällung wurde der Proteinrückstand jedoch mit Resuspendierungs-Puffer auf eine Proteinkonzentration von 6 µg/µL angelöst und 15 min bei 37 °C unter Agitation inkubiert. Anschließend wurde Verdünnungs-Puffer zugefügt, so dass nach der darauffolgenden Zugabe der CuAAC-Reagenzien die Proteinkonzentration 1 µg/µL und die Gua-Cl-Konzentration 1 M betrug. Für die CuAAC-Reaktion wurde zunächst 10 mM Biotin-Azid_{Amid} (gelöst in DMSO, Endkonzentration 500 µM) zugegeben. Als Nächstes wurden 50 mM CuSO₄ und 250 mM THPTA zu gleichen Teilen miteinander vermischt und anschließend zu den Proben gegeben (Endkonzentrationen 250 µM (CuSO₄), 1.25 mM (THPTA)). Zuletzt wurde 500 mM Natriumascorbat zugegeben (Endkonzentration 5 mM) und die Proben 1 h bei RT unter Agitation inkubiert.

Nun erfolgte eine zweite Acetonfällung und schließlich, wie in Kapitel 5.5.1 beschrieben, ein Protein-Verdau. Die Proben wurden mittels *ZipTips* aufgereinigt (s. Kap. 5.6.2), wobei maximal 10 µg Protein pro *ZipTip* aufgetragen wurden. Proben mit einer höheren Proteinmenge wurden auf mehrere *ZipTips* aufgeteilt und die Eluate anschließend vereinigt. Für die Anreicherung wurden je Probe 100 µL magnetische und mit Streptavidin funktionalisierte Partikel (*Dynabeads MyOne Streptavidin T1* von *Thermo Fisher Scientific*) drei Mal mit 500 µL PBS gewaschen und in 100 µL PBS resuspendiert. Die entsalzten Proben wurden in 400 µL PBS resuspendiert, zu den gewaschenen Partikeln gegeben und 1 h bei RT in einem Rotationsmischer inkubiert. Anschließend wurden die Partikel drei Mal mit 500 µL PBS und drei Mal mit 500 µL Wasch-Puffer gewaschen, bevor sie in 100 µL Wasch-Puffer resuspendiert wurden. Durch Zugabe von 1 M TCEP (Endkonzentration 10 mM) und Inkubation für 30 min bei RT unter Agitation wurden die Crosslinks von der Festphase abgespalten. Der Überstand wurde in ein neues Reaktionsgefäß überführt, die Partikel mit 50 µL Wasch-Puffer gespült und dieser Überstand ebenfalls in das neue Reaktionsgefäß gegeben. Der pH-Wert dieser Lösung wurde überprüft und gegebenenfalls mit mehr Puffer auf einen neutralen bis leicht basischen Wert eingestellt. Das durch die Reduktion entstandene Thiol wurde alkyliert, indem 1 M IAA (Endkonzentration 40 mM) zugegeben und 45 min bei RT unter Lichtausschluss und Agitation inkubiert wurde. Überschüssiges IAA wurde durch Zugabe von 500 mM DTT (Endkonzentration 20 mM) gequench, das Probenvolumen durch Evaporation des Lösungsmittels *in vacuo* auf ca. 20 µL reduziert, die Peptide mithilfe von *ZipTips* aufgereinigt (s. Kap. 5.6.2) und mittels MS² (s. Kap. 5.7.1 sowie 8.2, Tab. 8.1: Gradient 5 und Tab. 8.5: optimierte Methode) analysiert. Die Auswertung erfolgte mit *MeroX* (s. Kap. 5.7.2 sowie 8.3, Tab. 8.7 und 8.8) unter Verwendung der in Kapitel 5.4.1 beschriebenen Datenbank.

5.6 Aufreinigung von enzymatisch verdauten Protein-Proben

Immer wenn durch Protein-Verdau erhaltene Peptide mittels Massenspektrometrie analysiert wurden, wurden diese zuvor aufgereinigt, um Salze sowie potenzielle Verunreinigungen wie Detergenzien oder Reaktionsnebenprodukte zu entfernen, welche die nachfolgende massenspektrometrische Analyse negativ beeinflussen könnten.

5.6.1 Aufreinigung mithilfe von *StageTips*

| Reagenzien | |
|---------------------------------|-----------------------------------|
| <i>StageTip</i> -Elutionslösung | 80 % MeCN, 5 % NH ₄ OH |
| MS-Lösung | 2 % MeCN, 0.1 % FA |

Diese Methode wurde von *Kulak et al.* entwickelt^[592] und hier leicht modifiziert angewendet. Die *StageTips* (selbst hergestellt, 200 µL Pipettenspitzen (*Brand*) mit drei Schichten *Empore SDB-RPS*-Material gepackt) wurden zunächst mit 100 µL 1 % TFA 1.5 min bei 6000 × *g* zentrifugiert.

Anschließend wurden die zuvor mit 5 % TFA auf pH < 2 angesäuerten Proben auf die *StageTips* aufgetragen und bei 1500 × *g* zentrifugiert, bis das Volumen komplett durch die Säule gelaufen war. Nun wurden die *StageTips* mit je 100 µL 0.2 % TFA, *iso*-Propanol und schließlich erneut 0.2 % TFA gewaschen, wobei nach jedem Waschschrift bei 6000 × *g* zentrifugiert wurde, bis das Volumen durch die Säule gelaufen war. Die Elution erfolgte durch Zugabe von 150 µL *StageTip*-Elutionslösung und Zentrifugation bei 1500 × *g* in ein neues Reaktionsgefäß. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand auf eine Peptid-Konzentration von 1 µg/µL in MS-Lösung resuspendiert.

5.6.2 Aufreinigung mithilfe von C₁₈-*ZipTips*

| Reagenzien | |
|-------------------------------|----------------------|
| <i>ZipTip</i> -Waschlösung | 5 % MeOH, 0.1 % TFA |
| <i>ZipTip</i> -Elutionslösung | 80 % MeCN, 0.1 % TFA |
| MS-Lösung | 2 % MeCN, 0.1 % FA |

Für diese Aufreinigungs-Methode wurden *ZipTips* mit 0.6 µL C₁₈-Material (*Merck Millipore*) verwendet. Zunächst wurden die Proben mit 5 % TFA auf pH < 2 angesäuert, bevor die *ZipTips* durch je zweimaliges Aufziehen und Ausstoßen von 10 µL MeCN und danach 10 µL 0.1 % TFA gewaschen wurden. Nun wurden die Peptide durch zehnmaliges Auf- und Abpipettieren der *ZipTip* in der Probenlösung an die Säule gebunden. Das Säulenmaterial wurde durch dreimaliges Aufziehen und Ausstoßen von 10 µL *ZipTip*-Waschlösung gewaschen, bevor die Peptide durch zehnmaliges Auf- und Abpipettieren von 15 µL *ZipTip*-Elutionslösung in einem neuen Reaktionsgefäß eluiert wurden. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand auf eine Peptid-Konzentration von 1 µg/µL in MS-Lösung resuspendiert.

5.7 Massenspektrometrische Methoden

Die Analyse der verdauten Proteine erfolgte mittels einer HPLC-Trennung der Peptide und anschließender ESI-MS²-Analyse. Die so erhaltenen Peptid-Fragmentspektren wurden anschließend mithilfe von Datenbank-gestützten Proteinidentifikations-Programmen ausgewertet.

5.7.1 nanoLC-Tandem-Massenspektrometrie

Die Peptid-Analyse erfolgte durch chromatographische Auftrennung der Peptide mithilfe einer *Ultimate 3000 RSLCnano*-UHPLC von *Dionex/Thermo Fisher Scientific*, wonach *on-line* eine MS²-Analyse an einem *Q Exactive HF*-Massenspektrometers von *Thermo Fisher Scientific* durchgeführt wurde.

Um die Komplexität der Proben zu reduzieren, wurden die Peptide zunächst mittels Umkehrphasen-UHPLC entsprechend ihrer physikochemischen Eigenschaften aufgetrennt. 5 µL der Proben wurden dabei in einem ersten Schritt an eine Vorsäule (*Acclaim C18 PepMap100*, 300 µm Innendurchmesser, 5 mm Länge; *Thermo Fisher Scientific*) gebunden, um etwaige Verunreinigungen zu entfernen.

Anschließend wurde der Lösungsmittelfluss durch Umschalten eines Ventils umgekehrt und die Peptide für die Auftrennung auf eine selbst gepackte analytische Quarzglas-Säule (*SilicaTip Emitter*, 75 µm Innendurchmesser, 8 µm Spitze, 15 cm Länge; *New Objective*; gepackt mit *ReproSil-Pur 120 C18-AQ*, 1.9 µm, 120 Å; *Dr. Maisch GmbH*) gespült, welche über einen Säulenofen von *Sonation* auf 30 °C temperiert wurde.

Die Ionisation erfolgte durch Anlegen einer Spannung von 2.2 kV an die Säule über ein T-Stück. Die Peptide wurden mit einer Flussrate von 0.3 µL/min aufgetrennt, wobei als Lösungsmittel 0.1 % FA in H₂O (A) und 0.1 % FA in MeCN (B) verwendet wurden. Um die Peptide entsprechend ihrer Hydrophobizität aufzutrennen, wurde ein Gradient mit steigender Konzentration an Lösungsmittel B angewendet. Die für die LC-ESI-MS²-Analysen jeweils verwendeten Parameter sind in Kapitel 8.2 aufgeführt.

5.7.2 Auswertung der massenspektrometrischen Analysen

Die aus den MS-Analysen erhaltenen Spektren wurden mit den Programmen *Proteome Discoverer 2.2* (*Thermo Fisher Scientific*, im Falle von XL-Analysen mit Plugin *XlinkX 2.0*^[498]) oder mit *MeroX 2.0.0.6* bzw. *2.0.1.4*^[496] (*Michael Götze*, Martin-Luther-Universität Halle-Wittenberg) ausgewertet. Die dabei vorgenommenen Einstellungen sind in Kapitel 8.3 aufgeführt. Für die Auswertung mit *MeroX* wurden die bei der MS-Analyse generierten .RAW-Dateien zunächst mithilfe von *MSConvert* aus dem *ProteoWizard* Anwendungspaket (Version 3.0.9576)^[593] zu .*mgf*-Dateien konvertiert. Für Proben, bei denen die MS²-Spektren im Modus *Profile* aufgenommen wurden, wurde bei der Konvertierung das *Peak Picking* auf dem MS²-Level aktiviert.

6 Abkürzungsverzeichnis

| Abkürzung | Bedeutung |
|-----------|---|
| A | Adenin |
| Ac | Acetyl- |
| APS | Ammoniumperoxodisulfat |
| ATP | Adenosintriphosphat |
| BER | Basenexzisionsreparatur |
| BCCL | Abkürzung des Crosslinkers von Kang <i>et al.</i> ^[503] |
| bp | Basenpaar |
| BSA | bovines Serumalbumin |
| Bz | Benzoyl- |
| C | Cytosin |
| caC | 5-Carboxylcytosin |
| cd | engl. <i>catalytic domain</i> , katalytische Domäne |
| CE | Cyanoethyl- |
| CID | engl. <i>collision-induced dissociation</i> , kollisionsinduzierte Dissoziation |
| CLIP | engl. <i>click-enabled linker for interacting proteins</i> , Click-aktivierter Linker für interagierende Proteine |
| CoA | Coenzym A |
| CSM | engl. <i>crosslink-spectrum match</i> , Crosslink-Spektrum-Zuordnung |
| CuAAC | Cu(I)-katalysierte Alkin-Azid-Cycloaddition |
| dA | 2'-Desoxyadenosin |
| dC | 2'-Desoxycytidin |
| DCA | engl. <i>dichloroacetic acid</i> , Dichloressigsäure |
| DCM | Dichlormethan |
| dG | 2'-Desoxyguanosin |
| dmf | Dimethylformamidinyl- |
| DMSO | Dimethylsulfoxid |
| DNA | engl. <i>desoxyribonucleic acid</i> , Desoxyribonukleinsäure |
| dNTP | Desoxynukleotid-Triphosphat |
| dox | 1,3-Dioxan-2-yl- |
| ds | engl. <i>double-stranded</i> , doppelsträngig |
| DSBSO | Disuccinimidylbissulfoxid |
| DSBU | engl. <i>disuccinimidyldibutyric urea</i> , Disuccinimidyl-dibutyrat-Harnstoff |
| DSS | Disuccinimidylsuberat |
| DSSO | Disuccinimidylsulfoxid |
| dT | 2'-Desoxythymidin |
| DTT | Dithiothreitol |
| dU | 2'-Desoxyuridin |
| EDTA | Ethylendiamintetraacetat |
| EMSA | engl. <i>electrophoretic mobility shift assay</i> , elektrophoretischer Mobilitätsverschiebungs-Assay |

| Abkürzung | Bedeutung |
|----------------------------|--|
| engl. | englisch |
| ESCs | engl. <i>embryonic stem cells</i> , embryonale Stammzellen |
| ESI | Elektrospray-Ionisation |
| Et | Ethyl- |
| ETD | Elektronentransfer-Dissoziation |
| FA | engl. <i>formic acid</i> , Ameisensäure |
| FBS | engl. <i>fetal bovine serum</i> , fetales Kälberserum |
| fC | 5-Formylcytosin |
| fdC | 2'-Desoxy-5-formylcytidin |
| FDR | engl. <i>false discovery rate</i> , Falscherkennungsrate |
| fl | engl. <i>full length</i> , Volllängenprotein |
| Fr. | Fraktion |
| fU | 5-Formyluracil |
| fdU | 5-Formyldesoxyuridin |
| G | Guanin |
| Gua | Guanidinium |
| HCD | engl. <i>higher energy C-trap dissociation</i> , höherenergetische C-Trap-Dissoziation |
| HEK[-Zellen] | engl. <i>human embryonic kidney</i> , humane embryonale Niere[n-Zellen] |
| HEPES | 2-(4-(2-Hydroxyethyl)-piperazin-1-yl)ethansulfonsäure |
| HEPPS | 3-(4-(2-Hydroxyethyl)-piperazin-1-yl)propansulfonsäure |
| hmC | 5-Hydroxymethylcytosin |
| hmU | 5-Hydroxymethyluracil |
| HPA | engl. <i>hydroxypicolinic acid</i> , Hydroxypicolinsäure |
| HPLC | engl. <i>high performance liquid chromatography</i> , Hochleistungs-Flüssigkeitschromatographie |
| IAA | Iodacetamid |
| Kat. | Katalysator |
| Konz. | Konzentration |
| LC | engl. <i>liquid chromatography</i> , Flüssigkeitschromatographie |
| MALDI | engl. <i>matrix-assisted laser desorption/ionization</i> , Matrix-unterstützte Laser-Desorption/Ionisation |
| mC | 5-Methylcytosin |
| Me | Methyl- |
| MS | Massenspektrometrie |
| MS ⁿ (n = 2, 3) | Mehrfach-Massenspektrometrie |
| MWCO | engl. <i>molecular weight cut-off</i> , Größenausschluss |
| NER | Nukleotidexzisionsreparatur |
| NHS | <i>N</i> -Hydroxysuccinimidyl- |
| ODN | Oligodesoxyribonukleotid |
| <i>p</i> - | <i>para</i> - |
| PAGE | Polyacrylamid-Gelelektrophorese |
| PBS | engl. <i>phosphate buffered saline</i> , phosphat-gepufferte NaCl-Lösung |

| Abkürzung | Bedeutung |
|-----------|--|
| PCR | engl. <i>polymerase chain reaction</i> , Polymerase-Kettenreaktion |
| PSM | engl. <i>peptide-spectrum match</i> , Peptid-Spektrum-Zuordnung |
| RIPA | engl. <i>radioimmunoprecipitation assay</i> , Radioimmunpräzipitations-Assay |
| RNA | engl. <i>ribonucleic acid</i> , Ribonukleinsäure |
| RPMI | Roswell Park Memorial-Institut |
| RT | Raumtemperatur |
| S/N | engl. <i>signal to noise ratio</i> , Signal-Rausch-Verhältnis |
| SAM | S-Adenosylmethionin |
| SDS | engl. <i>sodium dodecyl sulfate</i> , Natriumdodecylsulfat |
| shdC | <i>superheavy</i> -dC, ¹³ C ₉ ¹⁵ N ₃ -markiertes Desoxycytidin |
| siRNA | engl. <i>small interfering ribonucleic acid</i> , kleine eingreifende Ribonukleinsäure |
| SMCC | Succinimidyl-4-(<i>N</i> -maleimidomethyl)-cyclohexan-1-carboxylat |
| ss | engl. <i>single-stranded</i> , einzelsträngig |
| T | Thymin |
| TAE | engl. <i>TRIS-acetic acid-EDTA</i> , TRIS-Essigsäure-EDTA |
| TBE | TRIS-Borsäure-EDTA |
| TCEP | Tris(2-carboxyethyl)phosphin |
| TEMED | <i>N,N,N',N'</i> -Tetramethylethan-1,2-diamin |
| TFA | engl. <i>trifluoroacetic acid</i> , Trifluoressigsäure |
| THPTA | Tris(3-hydroxypropyltriazolylmethyl)amin |
| TOF | engl. <i>time of flight</i> , Flugzeit-Massenanalysator |
| TRIS | Tris(hydroxymethyl)aminomethan |
| unmod. | unmodifiziert |
| UV | Ultraviolett |
| Vis | sichtbares Licht |
| Vol. | Volumen |
| XL | Crosslinking |

7 Literaturverzeichnis

- [1] H. Pearson, „Epigenetic“ in „Language: disputed definitions“, *Nature* **2008**, *455*, 1023–1028.
- [2] C. H. Waddington, „The epigenotype“, *Endeavour* **1942**, *1*, 18–20.
- [3] V. E. A. Russo, R. A. Martienssen, A. D. Riggs, *Epigenetic Mechanisms of Gene Regulation*, Cold Spring Harbor Laboratory Press, Woodbury, **1996**.
- [4] M. Ptashne, „On the use of the word ‘epigenetic’“, *Curr. Biol.* **2007**, *17*, R233–R236.
- [5] S. L. Berger, T. Kouzarides, R. Shiekhattar, A. Shilatifard, „An operational definition of epigenetics“, *Genes Dev.* **2009**, *23*, 781–783.
- [6] A. Bird, „Perceptions of epigenetics“, *Nature* **2007**, *447*, 396–398.
- [7] P. Boccaletto, M. A. Machnicka, E. Purta, P. Piątkowski, B. Bagiński, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm, J. M. Bujnicki, „MODOMICS: a database of RNA modification pathways. 2017 update“, *Nucleic Acids Res.* **2018**, *46*, D303–D307.
- [8] Genesilico, *Modomics*, <https://iimcb.genesilico.pl/modomics> (abgerufen am 16. April 2020).
- [9] A. J. Sood, C. Viner, M. M. Hoffman, „DNAMod: the DNA modification database“, *J. Cheminformatics* **2019**, *11*, 30.
- [10] *DNAMod: the DNA modification database*, <https://dnamod.hoffmanlab.org> (abgerufen am 16. April 2020).
- [11] C. von Sonntag, *Free-Radical-Induced DNA Damage and Its Repair*, 1st ed., Springer, Berlin/Heidelberg, **2006**, 357–482.
- [12] D. B. Dunn, J. D. Smith, „The occurrence of 6-methylaminopurine in deoxyribonucleic acids“, *Biochem. J.* **1958**, *68*, 627–636.
- [13] M. Ehrlich, M. A. Gama-Sosa, L. H. Carreira, L. G. Ljungdahl, K. C. Kuo, C. W. Gehrke, „DNA methylation in thermophilic bacteria: N⁴-methylcytosine, 5-methylcytosine, and N⁶-methyladenine“, *Nucleic Acids Res.* **1985**, *13*, 1399–1412.
- [14] J. E. Brooks, R. J. Roberts, „Modification profiles of bacterial genomes“, *Nucleic Acids Res.* **1982**, *10*, 913–934.
- [15] D. M. Heithoff, R. L. Sinsheimer, D. A. Low, M. J. Mahan, „An essential role for DNA adenine methylation in bacterial virulence“, *Science* **1999**, *284*, 967–970.
- [16] A. Reisenauer, L. Shapiro, „DNA methylation affects the cell cycle transcription of the CtrA global regulator in *Caulobacter*“, *EMBO J.* **2002**, *21*, 4969–4977.
- [17] T. Oshima, C. Wada, Y. Kawagoe, T. Ara, M. Maeda, Y. Masuda, S. Hiraga, H. Mori, „Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*“, *Mol. Microbiol.* **2002**, *45*, 673–695.
- [18] A. D. Hernday, B. A. Braaten, D. A. Low, „The mechanism by which DNA adenine methylase and PapI activate the pap epigenetic switch“, *Mol. Cell* **2003**, *12*, 947–957.

- [19] A. Løbner-Olesen, M. G. Marinus, F. G. Hansen, „Role of SeqA and Dam in *Escherichia coli* gene expression: a global/microarray analysis“, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4672–4677.
- [20] E. Boye, A. Løbner-Olesen, „The role of *dam* methyltransferase in the control of DNA replication in *E. coli*“, *Cell* **1990**, *62*, 981–989.
- [21] M. G. Marinus, A. Poteete, J. A. Arraj, „Correlation of DNA adenine methylase activity with spontaneous mutability in *Escherichia coli* K-12“, *Gene* **1984**, *28*, 123–125.
- [22] S. Hattman, C. Kenny, L. Berger, K. Pratt, „Comparative study of DNA methylation in three unicellular eucaryotes“, *J. Bacteriol.* **1978**, *135*, 1156–1157.
- [23] Y. Fu, G. Z. Luo, K. Chen, X. Deng, M. Yu, D. Han, Z. Hao, J. Liu, X. Lu, L. C. Doré, X. Weng, Q. Ji, L. Mets, C. He, „N⁶-Methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*“, *Cell* **2015**, *161*, 879–892.
- [24] G. Zhang, H. Huang, D. Liu, Y. Cheng, X. Liu, W. Zhang, R. Yin, D. Zhang, P. Zhang, J. Liu, C. Li, B. Liu, Y. Luo, Y. Zhu, N. Zhang, S. He, C. He, H. Wang, D. Chen, „N⁶-methyladenine DNA modification in *Drosophila*“, *Cell* **2015**, *161*, 893–906.
- [25] E. L. Greer, M. A. Blanco, L. Gu, E. Sendinc, J. Liu, D. Aristizábal-Corrales, C. H. Hsu, L. Aravind, C. He, Y. Shi, „DNA methylation on N⁶-adenine in *C. elegans*“, *Cell* **2015**, *161*, 868–878.
- [26] M. J. Koziol, C. R. Bradshaw, G. E. Allen, A. S. H. Costa, C. Frezza, J. B. Gurdon, „Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications“, *Nat. Struct. Mol. Biol.* **2016**, *23*, 24–30.
- [27] T. P. Wu, T. Wang, M. G. Seetin, Y. Lai, S. Zhu, K. Lin, Y. Liu, S. D. Byrum, S. G. Mackintosh, M. Zhong, A. Tackett, G. Wang, L. S. Hon, G. Fang, J. A. Swenberg, A. Z. Xiao, „DNA methylation on N⁶-adenine in mammalian embryonic stem cells“, *Nature* **2016**, *532*, 329–333.
- [28] S. Schiffers, C. Ebert, R. Rahimoff, O. Kosmatchev, J. Steinbacher, A.-V. Bohne, F. Spada, S. Michalakis, J. Nickelsen, M. Müller, T. Carell, „Quantitative LC–MS provides no evidence for m⁶dA or m⁴dC in the genome of mouse embryonic stem cells and tissues“, *Angew. Chem. Int. Ed.* **2017**, *56*, 11268–11271; „Quantitative LC-MS liefert keinen Hinweis auf m⁶dA oder m⁴dC im Genom von Mausstammzellen und -geweben“, *Angew. Chem.* **2017**, *129*, 11422–11425.
- [29] C. W. Q. Koh, Y. T. Goh, J. D. W. Toh, S. P. Neo, S. B. Ng, J. Gunaratne, Y.-G. Gao, S. R. Quake, W. F. Burkholder, W. S. S. Goh, „Single-nucleotide-resolution sequencing of human N⁶-methyldeoxyadenosine reveals strand-asymmetric clusters associated with SSBP1 on the mitochondrial genome“, *Nucleic Acids Res.* **2018**, *46*, 11659–11670.
- [30] L. Y. Beh, G. T. Debelouchina, D. M. Clay, R. E. Thompson, K. A. Lindblad, E. R. Hutton, J. R. Bracht, R. P. Sebra, T. W. Muir, L. F. Landweber, „Identification of a DNA N⁶-adenine methyltransferase complex and its impact on chromatin organization“, *Cell* **2019**, *177*, 1781–1796.e1725.
- [31] B. Yao, Y. Cheng, Z. Wang, Y. Li, L. Chen, L. Huang, W. Zhang, D. Chen, H. Wu, B. Tang, P. Jin, „DNA N⁶-methyladenine is dynamically regulated in the mouse brain following environmental stress“, *Nat. Commun.* **2017**, *8*, 1122.

- [32] X. Li, Q. Zhao, W. Wei, Q. Lin, C. Magnan, M. R. Emami, L. E. Wearick-Silva, T. W. Viola, P. R. Marshall, J. Yin, S. U. Madugalle, Z. Wang, S. Nainar, C. B. Vågbo, L. J. Leighton, E. L. Zajackowski, K. Ke, R. Grassi-Oliveira, M. Bjørås, P. F. Baldi, R. C. Spitale, T. W. Bredy, „The DNA modification N⁶-methyl-2'-deoxyadenosine (m₆dA) drives activity-induced gene expression and is required for fear extinction“, *Nat. Neurosci.* **2019**, *22*, 534–544.
- [33] S.-M. Kweon, Y. Chen, E. Moon, K. Kvederaviciutė, S. Klimasauskas, D. E. Feldman, „An adversarial DNA N⁶-methyladenine-sensor network preserves Polycomb silencing“, *Mol. Cell* **2019**, *74*, 1138–1147.e1136.
- [34] Z. Hao, T. Wu, X. Cui, P. Zhu, C. Tan, X. Dou, K.-W. Hsu, Y.-T. Lin, P.-H. Peng, L.-S. Zhang, Y. Gao, L. Hu, H.-L. Sun, A. Zhu, J. Liu, K.-J. Wu, C. He, „N⁶-deoxyadenosine methylation in mammalian mitochondrial DNA“, *Mol. Cell* **2020**, *78*, 382–395.e388.
- [35] J. Xiong, T.-T. Ye, C.-J. Ma, Q.-Y. Cheng, B.-F. Yuan, Y.-Q. Feng, „N⁶-hydroxymethyladenine: a hydroxylation derivative of N⁶-methyladenine in genomic DNA of mammals“, *Nucleic Acids Res.* **2018**, *47*, 1268–1277.
- [36] Z. K. O’Brown, K. Boulias, J. Wang, S. Y. Wang, N. M. O’Brown, Z. Hao, H. Shibuya, P.-E. Fady, Y. Shi, C. He, S. G. Megason, T. Liu, E. L. Greer, „Sources of artifact in measurements of 6mA and 4mC abundance in eukaryotic genomic DNA“, *BMC Genomics* **2019**, *20*, 445.
- [37] S. Feng, S. J. Cokus, X. Zhang, P. Y. Chen, M. Bostick, M. G. Goll, J. Hetzel, J. Jain, S. H. Strauss, M. E. Halpern, C. Ukomadu, K. C. Sadler, S. Pradhan, M. Pellegrini, S. E. Jacobsen, „Conservation and divergence of methylation patterning in plants and animals“, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 8689–8694.
- [38] A. Zemach, I. E. McDaniel, P. Silva, D. Zilberman, „Genome-wide evolutionary analysis of eukaryotic DNA methylation“, *Science* **2010**, *328*, 916–919.
- [39] D. Globisch, M. Münzel, M. Müller, S. Michalakis, M. Wagner, S. Koch, T. Brückl, M. Biel, T. Carell, „Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates“, *PLoS One* **2010**, *5*, e15367.
- [40] R. Lister, M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q.-M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren, J. R. Ecker, „Human DNA methylomes at base resolution show widespread epigenomic differences“, *Nature* **2009**, *462*, 315–322.
- [41] L. Bonen, T. Y. Huh', M. W. Gray, „Can partial methylation explain the complex fragment patterns observed when plant mitochondrial DNA is cleaved with restriction endonucleases?“, *FEBS Lett.* **1980**, *111*, 340–346.
- [42] B. Deumling, „Sequence arrangement of a highly methylated satellite DNA of a plant, *Scilla*: a tandemly repeated inverted repeat“, *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 338–342.
- [43] Y. Gruenbaum, T. Naveh-Many, H. Cedar, A. Razin, „Sequence specificity of methylation in higher plant DNA“, *Nature* **1981**, *292*, 860–862.
- [44] P. Meyer, I. Niedenhof, M. ten Lohuis, „Evidence for cytosine methylation of non-symmetrical sequences in transgenic *Petunia hybrida*“, *EMBO J.* **1994**, *13*, 2084–2088.

- [45] E. J. Oakeley, J.-P. Jost, „Non-symmetrical cytosine methylation in tobacco pollen DNA“, *Plant Mol. Biol.* **1996**, *31*, 927–930.
- [46] C. Goubely, P. Arnaud, C. Tatout, J. S. Heslop-Harrison, J. M. Deragon, „S1 SINE retroposons are methylated at symmetrical and non-symmetrical positions in *Brassica napus*: identification of a preferred target site for asymmetrical methylation“, *Plant Mol. Biol.* **1999**, *39*, 243–255.
- [47] B. H. Ramsahoye, D. Biniszkiwicz, F. Lyko, V. Clark, A. P. Bird, R. Jaenisch, „Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a“, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5237–5242.
- [48] A. J. Price, L. Collado-Torres, N. A. Ivanov, W. Xia, E. E. Burke, J. H. Shin, R. Tao, L. Ma, Y. Jia, T. M. Hyde, J. E. Kleinman, D. R. Weinberger, A. E. Jaffe, „Divergent neuronal DNA methylation patterns across human cortical development reveal critical periods and a unique role of CpH methylation“, *Genome Biol.* **2019**, *20*, 196.
- [49] W. Salser, „Globin mRNA sequences: analysis of base pairing and evolutionary implications“, *Cold Spring Harb. Symp. Quant. Biol.* **1978**, *42*, 985–1002.
- [50] A. P. Bird, „DNA methylation and the frequency of CpG in animal DNA“, *Nucleic Acids Res.* **1980**, *8*, 1499–1504.
- [51] M. Gardiner-Garden, M. Frommer, „CpG islands in vertebrate genomes“, *J. Mol. Biol.* **1987**, *196*, 261–282.
- [52] F. Antequera, A. Bird, „Number of CpG islands and genes in human and mouse“, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11995–11999.
- [53] D. N. Cooper, M. H. Taggart, A. P. Bird, „Unmethylated domains in vertebrate DNA“, *Nucleic Acids Res.* **1983**, *11*, 647–658.
- [54] A. Bird, M. Taggart, M. Frommer, O. J. Miller, D. Macleod, „A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA“, *Cell* **1985**, *40*, 91–99.
- [55] F. Larsen, G. Gundersen, R. Lopez, H. Prydz, „CpG islands as gene markers in the human genome“, *Genomics* **1992**, *13*, 1095–1107.
- [56] S. Saxonov, P. Berg, D. L. Brutlag, „A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters“, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1412–1417.
- [57] B. E. Bernstein, T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber, E. S. Lander, „A bivalent chromatin structure marks key developmental genes in embryonic stem cells“, *Cell* **2006**, *125*, 315–326.
- [58] M. Ku, R. P. Koche, E. Rheinbay, E. M. Mendenhall, M. Endoh, T. S. Mikkelsen, A. Presser, C. Nusbaum, X. Xie, A. S. Chi, M. Adli, S. Kasif, L. M. Ptaszek, C. A. Cowan, E. S. Lander, H. Koseki, B. E. Bernstein, „Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains“, *PLoS Genet.* **2008**, *4*, e1000242.

-
- [59] N. Kikyo, C. M. Williamson, R. M. John, S. C. Barton, C. V. Beechey, S. T. Ball, B. M. Cattanach, M. A. Surani, J. Peters, „Genetic and functional analysis of *neuronatin* in mice with maternal or paternal duplication of distal Chr 2“, *Dev. Biol.* **1997**, *190*, 66–77.
- [60] J. D. Choi, L. A. Underkoffler, A. J. Wood, J. N. Collins, P. T. Williams, J. A. Golden, E. F. Schuster Jr., K. M. Loomes, R. J. Oakey, „A novel variant of *Inpp5f* is imprinted in brain, and its expression is correlated with differential methylation of an internal CpG island“, *Mol. Cell. Biol.* **2005**, *25*, 5514–5522.
- [61] M. Weber, I. Hellmann, M. B. Stadler, L. Ramos, S. Pääbo, M. Rebhan, D. Schübeler, „Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome“, *Nat. Genet.* **2007**, *39*, 457–466.
- [62] E. Schilling, M. Rehli, „Global, comparative analysis of tissue-specific promoter CpG methylation“, *Genomics* **2007**, *90*, 314–323.
- [63] L. Shen, Y. Kondo, Y. Guo, J. Zhang, L. Zhang, S. Ahmed, J. Shu, X. Chen, R. A. Waterland, J. P. Issa, „Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters“, *PLoS Genet.* **2007**, *3*, 2023–2036.
- [64] F. Mohn, M. Weber, M. Rebhan, T. C. Roloff, J. Richter, M. B. Stadler, M. Bibel, D. Schübeler, „Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors“, *Mol. Cell* **2008**, *30*, 755–766.
- [65] T. K. Kelly, T. B. Miranda, G. Liang, B. P. Berman, J. C. Lin, A. Tanay, P. A. Jones, „H2A.Z maintenance during mitosis reveals nucleosome shifting on mitotically silenced genes“, *Mol. Cell* **2010**, *39*, 901–911.
- [66] T. Sado, M. Okano, E. Li, H. Sasaki, „De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation“, *Development* **2004**, *131*, 975–982.
- [67] L. F. Lock, N. Takagi, G. R. Martin, „Methylation of the *Hprt* gene on the inactive X occurs after chromosome inactivation“, *Cell* **1987**, *48*, 39–46.
- [68] T. Sado, M. H. Fenner, S. S. Tan, P. Tam, T. Shioda, E. Li, „X inactivation in the mouse embryo deficient for *Dnmt1*: distinct effect of hypomethylation on imprinted and random X inactivation“, *Dev. Biol.* **2000**, *225*, 294–303.
- [69] J. E. Ohm, K. M. McGarvey, X. Yu, L. Cheng, K. E. Schuebel, L. Cope, H. P. Mohammad, W. Chen, V. C. Daniel, W. Yu, D. M. Berman, T. Jenuwein, K. Pruitt, S. J. Sharkis, D. N. Watkins, J. G. Herman, S. B. Baylin, „A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing“, *Nat. Genet.* **2007**, *39*, 237–242.
- [70] Y. Schlesinger, R. Straussman, I. Keshet, S. Farkash, M. Hecht, J. Zimmerman, E. Eden, Z. Yakhini, E. Ben-Shushan, B. E. Reubinoff, Y. Bergman, I. Simon, H. Cedar, „Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer“, *Nat. Genet.* **2007**, *39*, 232–236.
- [71] S. M. Iguchi-Ariga, W. Schaffner, „CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation“, *Genes Dev.* **1989**, *3*, 612–619.
-

- [72] M. R. Campanero, M. I. Armstrong, E. K. Flemington, „CpG methylation as a mechanism for the regulation of E2F activity“, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6481–6486.
- [73] X. Nan, H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman, A. Bird, „Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex“, *Nature* **1998**, *393*, 386–389.
- [74] P. L. Jones, G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis, A. P. Wolffe, „Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription“, *Nat. Genet.* **1998**, *19*, 187–191.
- [75] Y. Zhang, H. H. Ng, H. Erdjument-Bromage, P. Tempst, A. Bird, D. Reinberg, „Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation“, *Genes Dev.* **1999**, *13*, 1924–1935.
- [76] P. A. Wade, A. Geggion, P. L. Jones, E. Ballestar, F. Aubry, A. P. Wolffe, „Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation“, *Nat. Genet.* **1999**, *23*, 62–66.
- [77] F. Fuks, P. J. Hurd, D. Wolf, X. Nan, A. P. Bird, T. Kouzarides, „The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation“, *J. Biol. Chem.* **2003**, *278*, 4035–4040.
- [78] C. P. Walsh, J. R. Chaillet, T. H. Bestor, „Transcription of IAP endogenous retroviruses is constrained by cytosine methylation“, *Nat. Genet.* **1998**, *20*, 116–117.
- [79] D. Bourc'his, T. H. Bestor, „Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L“, *Nature* **2004**, *431*, 96–99.
- [80] J. Barau, A. Teissandier, N. Zamudio, S. Roy, V. Nalesso, Y. Héroult, F. Guillou, D. Bourc'his, „The DNA methyltransferase DNMT3C protects male germ cells from transposon activity“, *Science* **2016**, *354*, 909–912.
- [81] S. Shukla, E. Kavak, M. Gregory, M. Imashimizu, B. Shutinoski, M. Kashlev, P. Oberdoerffer, R. Sandberg, S. Oberdoerffer, „CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing“, *Nature* **2011**, *479*, 74–79.
- [82] J. C. Wu, D. V. Santi, „Kinetic and catalytic mechanism of HhaI methyltransferase“, *J. Biol. Chem.* **1987**, *262*, 4778–4786.
- [83] S. Pradhan, A. Bacolla, R. D. Wells, R. J. Roberts, „Recombinant human DNA (cytosine-5) methyltransferase: I. Expression, purification, and comparison of *de novo* and maintenance methylation“, *J. Biol. Chem.* **1999**, *274*, 33002–33010.
- [84] M. Bostick, J. K. Kim, P. O. Estève, A. Clark, S. Pradhan, S. E. Jacobsen, „UHRF1 plays a role in maintaining DNA methylation in mammalian cells“, *Science* **2007**, *317*, 1760–1764.
- [85] J. Sharif, M. Muto, S. Takebayashi, I. Suetake, A. Iwamatsu, T. A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, K. Okamura, S. Tajima, K. Mitsuya, M. Okano, H. Koseki, „The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA“, *Nature* **2007**, *450*, 908–912.
- [86] M. Okano, S. Xie, E. Li, „Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases“, *Nat. Genet.* **1998**, *19*, 219–220.

-
- [87] M. Okano, D. W. Bell, D. A. Haber, E. Li, „DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development“, *Cell* **1999**, *99*, 247–257.
- [88] M. Monk, M. Boubelik, S. Lehnert, „Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development“, *Development* **1987**, *99*, 371–382.
- [89] F. Santos, B. Hendrich, W. Reik, W. Dean, „Dynamic reprogramming of DNA methylation in the early mouse embryo“, *Dev. Biol.* **2002**, *241*, 172–182.
- [90] P. Hajkova, S. Erhardt, N. Lane, T. Haaf, O. El-Maarri, W. Reik, J. Walter, M. A. Surani, „Epigenetic reprogramming in mouse primordial germ cells“, *Mech. Dev.* **2002**, *117*, 15–23.
- [91] D. Jia, R. Z. Jurkowska, X. Zhang, A. Jeltsch, X. Cheng, „Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation“, *Nature* **2007**, *449*, 248–251.
- [92] I. Suetake, F. Shinozaki, J. Miyagawa, H. Takeshima, S. Tajima, „DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction“, *J. Biol. Chem.* **2004**, *279*, 27816–27823.
- [93] K. Hata, M. Okano, H. Lei, E. Li, „Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice“, *Development* **2002**, *129*, 1983–1993.
- [94] S. K. T. Ooi, C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S.-P. Lin, C. D. Allis, X. Cheng, T. H. Bestor, „DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA“, *Nature* **2007**, *448*, 714–717.
- [95] G. Liang, M. F. Chan, Y. Tomigahara, Y. C. Tsai, F. A. Gonzales, E. Li, P. W. Laird, P. A. Jones, „Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements“, *Mol. Cell. Biol.* **2002**, *22*, 480–491.
- [96] G. D. Kim, J. Ni, N. Kelesoglu, R. J. Roberts, S. Pradhan, „Co-operation and communication between the human maintenance and *de novo* DNA (cytosine-5) methyltransferases“, *EMBO J.* **2002**, *21*, 4183–4195.
- [97] M. Fatemi, A. Hermann, H. Gowher, A. Jeltsch, „Dnmt3a and Dnmt1 functionally cooperate during *de novo* methylation of DNA“, *Eur. J. Biochem.* **2002**, *269*, 4981–4984.
- [98] T. Chen, Y. Ueda, J. E. Dodge, Z. Wang, E. Li, „Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b“, *Mol. Cell. Biol.* **2003**, *23*, 5594–5605.
- [99] S. Jeong, G. Liang, S. Sharma, J. C. Lin, S. H. Choi, H. Han, C. B. Yoo, G. Egger, A. S. Yang, P. A. Jones, „Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA“, *Mol. Cell. Biol.* **2009**, *29*, 5366–5376.
- [100] M. D. Saenz-de-Juano, E. Ivanova, K. Billooye, A.-C. Herta, J. Smitz, G. Kelsey, E. Anckaert, „Genome-wide assessment of DNA methylation in mouse oocytes reveals effects associated with in vitro growth, superovulation, and sexual maturity“, *Clin. Epigenetics* **2019**, *11*, 197.
-

- [101] C. Popp, W. Dean, S. Feng, S. J. Cokus, S. Andrews, M. Pellegrini, S. E. Jacobsen, W. Reik, „Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency“, *Nature* **2010**, *463*, 1101–1105.
- [102] N. Rougier, D. Bourc'his, D. M. Gomes, A. Niveleau, M. Plachot, A. Paldi, E. Viegas-Péquignot, „Chromosome methylation patterns during mammalian preimplantation development“, *Genes Dev.* **1998**, *12*, 2108–2113.
- [103] L. L. Carlson, A. W. Page, T. H. Bestor, „Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting“, *Genes Dev.* **1992**, *6*, 2536–2541.
- [104] S. Ratnam, C. Mertineit, F. Ding, C. Y. Howell, H. J. Clarke, T. H. Bestor, J. R. Chaillet, J. M. Trasler, „Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development“, *Dev. Biol.* **2002**, *245*, 304–314.
- [105] W. Mayer, A. Niveleau, J. Walter, R. Fundele, T. Haaf, „Demethylation of the zygotic paternal genome“, *Nature* **2000**, *403*, 501–502.
- [106] J. Oswald, S. Engemann, N. Lane, W. Mayer, A. Olek, R. Fundele, W. Dean, W. Reik, J. Walter, „Active demethylation of the paternal genome in the mouse zygote“, *Curr. Biol.* **2000**, *10*, 475–478.
- [107] T. Nakamura, Y. Arai, H. Umehara, M. Masuhara, T. Kimura, H. Taniguchi, T. Sekimoto, M. Ikawa, Y. Yoneda, M. Okabe, S. Tanaka, K. Shiota, T. Nakano, „PGC7/Stella protects against DNA demethylation in early embryogenesis“, *Nat. Cell Biol.* **2007**, *9*, 64–71.
- [108] Y. Choi, M. Gehring, L. Johnson, M. Hannon, J. J. Harada, R. B. Goldberg, S. E. Jacobsen, R. L. Fischer, „DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*“, *Cell* **2002**, *110*, 33–42.
- [109] Z. Gong, T. Morales-Ruiz, R. R. Ariza, T. Roldán-Arjona, L. David, J. K. Zhu, „*ROS1*, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase“, *Cell* **2002**, *111*, 803–814.
- [110] J. Penterman, D. Zilberman, J. H. Huh, T. Ballinger, S. Henikoff, R. L. Fischer, „DNA demethylation in the *Arabidopsis* genome“, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6752–6757.
- [111] A. P. Ortega-Galisteo, T. Morales-Ruiz, R. R. Ariza, T. Roldán-Arjona, „*Arabidopsis* DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks“, *Plant Mol. Biol.* **2008**, *67*, 671–681.
- [112] M. I. Ponferrada-Marín, T. Roldán-Arjona, R. R. Ariza, „*ROS1* 5-methylcytosine DNA glycosylase is a slow-turnover catalyst that initiates DNA demethylation in a distributive fashion“, *Nucleic Acids Res.* **2009**, *37*, 4264–4274.
- [113] H. La, B. Ding, G. P. Mishra, B. Zhou, H. Yang, R. Bellizzi Mdel, S. Chen, B. C. Meyers, Z. Peng, J. K. Zhu, G. L. Wang, „A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon *Tos17* and promotes its transposition in rice“, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 15498–15503.

- [114] J. P. Jost, M. Siegmann, L. Sun, R. Leung, „Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase“, *J. Biol. Chem.* **1995**, *270*, 9734–9739.
- [115] B. Zhu, Y. Zheng, D. Hess, H. Angliker, S. Schwarz, M. Siegmann, S. Thiry, J. P. Jost, „5-Methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex“, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5135–5139.
- [116] B. Zhu, Y. Zheng, H. Angliker, S. Schwarz, S. Thiry, M. Siegmann, J. P. Jost, „5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence“, *Nucleic Acids Res.* **2000**, *28*, 4157–4165.
- [117] J. C. Huang, D. L. Svoboda, J. T. Reardon, A. Sancar, „Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer“, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3664–3668.
- [118] J. G. Moggs, K. J. Yarema, J. M. Essigmann, R. D. Wood, „Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct“, *J. Biol. Chem.* **1996**, *271*, 7177–7186.
- [119] A. D. D'Andrea, „Chapter 4 - DNA repair pathways and human cancer“ in *The Molecular Basis of Cancer*, 3rd Ed. (Eds.: J. Mendelsohn, P. M. Howley, M. A. Israel, J. W. Gray, C. B. Thompson), W.B. Saunders, Philadelphia, **2008**, 42.
- [120] G. Barreto, A. Schäfer, J. Marhold, D. Stach, S. K. Swaminathan, V. Handa, G. Döderlein, N. Maltry, W. Wu, F. Lyko, C. Niehrs, „Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation“, *Nature* **2007**, *445*, 671–675.
- [121] K.-M. Schmitz, N. Schmitt, U. Hoffmann-Rohrer, A. Schäfer, I. Grummt, C. Mayer, „TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation“, *Mol. Cell* **2009**, *33*, 344–353.
- [122] S. G. Jin, C. Guo, G. P. Pfeifer, „GADD45A does not promote DNA demethylation“, *PLoS Genet.* **2008**, *4*, e1000013.
- [123] S. K. Bhattacharya, S. Ramchandani, N. Cervoni, M. Szyf, „A mammalian protein with specific demethylase activity for mCpG DNA“, *Nature* **1999**, *397*, 579–583.
- [124] B. Hendrich, J. Guy, B. Ramsahoye, V. A. Wilson, A. Bird, „Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development“, *Genes Dev.* **2001**, *15*, 710–723.
- [125] N. Navaratnam, J. R. Morrison, S. Bhattacharya, D. Patel, T. Funahashi, F. Giannoni, B. B. Teng, N. O. Davidson, J. Scott, „The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase“, *J. Biol. Chem.* **1993**, *268*, 20709–20712.
- [126] M. Muramatsu, V. S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N. O. Davidson, T. Honjo, „Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells“, *J. Biol. Chem.* **1999**, *274*, 18470–18476.

- [127] H. D. Morgan, W. Dean, H. A. Coker, W. Reik, S. K. Petersen-Mahrt, „Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming“, *J. Biol. Chem.* **2004**, *279*, 52353–52360.
- [128] R. Métivier, R. Gallais, C. Tiffoche, C. Le Péron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, V. Benes, A. Jeltsch, F. Gannon, G. Salbert, „Cyclical DNA methylation of a transcriptionally active promoter“, *Nature* **2008**, *452*, 45–50.
- [129] S. Kangaspeska, B. Stride, R. Métivier, M. Polycarpou-Schwarz, D. Ibberson, R. P. Carmouche, V. Benes, F. Gannon, G. Reid, „Transient cyclical methylation of promoter DNA“, *Nature* **2008**, *452*, 112–115.
- [130] S. Cortellino, J. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, L. K. Abramowitz, M. S. Bartolomei, F. Rambow, M. R. Bassi, T. Bruno, M. Fanciulli, C. Renner, A. J. Klein-Szanto, Y. Matsumoto, D. Kobi, I. Davidson, C. Alberti, L. Larue, A. Bellacosa, „Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair“, *Cell* **2011**, *146*, 67–79.
- [131] C. B. Millar, J. Guy, O. J. Sansom, J. Selfridge, E. MacDougall, B. Hendrich, P. D. Keightley, S. M. Bishop, A. R. Clarke, A. Bird, „Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice“, *Science* **2002**, *297*, 403–405.
- [132] I. Grin, A. A. Ishchenko, „An interplay of the base excision repair and mismatch repair pathways in active DNA demethylation“, *Nucleic Acids Res.* **2016**, *44*, 3713–3727.
- [133] F. Santos, J. Peat, H. Burgess, C. Rada, W. Reik, W. Dean, „Active demethylation in mouse zygotes involves cytosine deamination and base excision repair“, *Epigenet. Chromatin* **2013**, *6*, 39.
- [134] D.-M. Franchini, C.-F. Chan, H. Morgan, E. Incorvaia, G. Rangam, W. Dean, F. Santos, W. Reik, S. K. Petersen-Mahrt, „Processive DNA demethylation via DNA deaminase-induced lesion resolution“, *PLoS One* **2014**, *9*, e97754.
- [135] N. W. Penn, R. Suwalski, C. O'Riley, K. Bojanowski, R. Yura, „The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid“, *Biochem. J.* **1972**, *126*, 781–790.
- [136] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, „Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1“, *Science* **2009**, *324*, 930–935.
- [137] S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, „Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification“, *Nature* **2010**, *466*, 1129–1133.
- [138] T. Pfaffeneder, B. Hackner, M. Truss, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeyer, T. Carell, „The discovery of 5-formylcytosine in embryonic stem cell DNA“, *Angew. Chem. Int. Ed.* **2011**, *50*, 7008–7012; „The discovery of 5-formylcytosine in embryonic stem cell DNA“, *Angew. Chem.* **2011**, *123*, 7146–7150.
- [139] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, „Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine“, *Science* **2011**, *333*, 1300–1303.

- [140] Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, G. L. Xu, „Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA“, *Science* **2011**, *333*, 1303–1307.
- [141] M. Wossidlo, T. Nakamura, K. Lepikhov, C. J. Marques, V. Zakhartchenko, M. Boiani, J. Arand, T. Nakano, W. Reik, J. Walter, „5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming“, *Nat. Commun.* **2011**, *2*, 241.
- [142] A. Inoue, Y. Zhang, „Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos“, *Science* **2011**, *334*, 194.
- [143] A. Inoue, L. Shen, Q. Dai, C. He, Y. Zhang, „Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development“, *Cell Res.* **2011**, *21*, 1670–1676.
- [144] Q. Dai, P. J. Sanstead, C. S. Peng, D. Han, C. He, A. Tokmakoff, „Weakened N3 hydrogen bonding by 5-formylcytosine and 5-carboxylcytosine reduces their base-pairing stability“, *ACS Chem. Biol.* **2016**, *11*, 470–477.
- [145] A. Maiti, A. Z. Michelson, C. J. Armwood, J. K. Lee, A. C. Drohat, „Divergent mechanisms for enzymatic excision of 5-formylcytosine and 5-carboxylcytosine from DNA“, *J. Am. Chem. Soc.* **2013**, *135*, 15813–15822.
- [146] A. Maiti, A. C. Drohat, „Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites“, *J. Biol. Chem.* **2011**, *286*, 35334–35338.
- [147] U. Müller, C. Bauer, M. Siegl, A. Rottach, H. Leonhardt, „TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation“, *Nucleic Acids Res.* **2014**, *42*, 8592–8604.
- [148] L. Schomacher, D. Han, M. U. Musheev, K. Arab, S. Kienhöfer, A. von Seggern, C. Niehrs, „Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation“, *Nat. Struct. Mol. Biol.* **2016**, *23*, 116–124.
- [149] A. Slyvka, K. Mierzejewska, M. Bochtler, „Nei-like 1 (NEIL1) excises 5-carboxylcytosine directly and stimulates TDG-mediated 5-formyl and 5-carboxylcytosine excision“, *Sci. Rep.* **2017**, *7*, 9001.
- [150] R. C. A. Dubini, A. Schön, M. Müller, T. Carell, P. Rovó, „Impact of 5-formylcytosine on the melting kinetics of DNA by ¹H NMR chemical exchange“, *Nucleic Acids Res.* **2020**, *48*, 8796–8807.
- [151] J. U. Guo, Y. Su, C. Zhong, G. L. Ming, H. Song, „Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain“, *Cell* **2011**, *145*, 423–434.
- [152] H. Hashimoto, Y. Liu, A. K. Upadhyay, Y. Chang, S. B. Howerton, P. M. Vertino, X. Zhang, X. Cheng, „Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation“, *Nucleic Acids Res.* **2012**, *40*, 4841–4849.
- [153] R. J. Boorstein, A. Cummings Jr., D. R. Marenstein, M. K. Chan, Y. Ma, T. A. Neubert, S. M. Brown, G. W. Teebor, „Definitive identification of mammalian 5-hydroxymethyluracil DNAN-glycosylase activity as SMUG1“, *J. Biol. Chem.* **2001**, *276*, 41991–41997.

- [154] C. S. Nabel, H. Jia, Y. Ye, L. Shen, H. L. Goldschmidt, J. T. Stivers, Y. Zhang, R. M. Kohli, „AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation“, *Nat. Chem. Biol.* **2012**, *8*, 751–758.
- [155] G. Rangam, K. M. Schmitz, A. J. Cobb, S. K. Petersen-Mahrt, „AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud“, *PLoS One* **2012**, *7*, e43279.
- [156] T. Pfaffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljarova, D. Schuermann, S. Michalakis, O. Kosmatchev, S. Schiesser, B. Steigenberger, N. Raddaoui, G. Kashiwazaki, U. Müller, C. G. Spruijt, M. Vermeulen, H. Leonhardt, P. Schär, M. Müller, T. Carell, „Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA“, *Nat. Chem. Biol.* **2014**, *10*, 574–581.
- [157] K. Iwan, R. Rahimoff, A. Kirchner, F. Spada, A. S. Schröder, O. Kosmatchev, S. Ferizaj, J. Steinbacher, E. Parsa, M. Müller, T. Carell, „5-Formylcytosine to cytosine conversion by C-C bond cleavage *in vivo*“, *Nat. Chem. Biol.* **2018**, *14*, 72–78.
- [158] A. Schön, E. Kaminska, F. Schelter, E. Ponkkonen, E. Korytiaková, S. Schiffers, T. Carell, „Analysis of an active deformylation mechanism of 5-formyl-deoxycytidine (fdC) in stem cells“, *Angew. Chem. Int. Ed.* **2020**, *59*, 5591–5594; „Analyse des aktiven Deformylierungsmechanismus von 5-Formyl-2'-Desoxycytidin in Stammzellen“, *Angew. Chem.* **2020**, *132*, 5639–5643.
- [159] S. Schiesser, B. Hackner, T. Pfaffeneder, M. Müller, C. Hagemeyer, M. Truss, T. Carell, „Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing“, *Angew. Chem. Int. Ed.* **2012**, *51*, 6516–6520; „Mechanismus und Stammzellaktivität der Decarboxylierung von 5-Carboxycytosin, bestimmt mittels Isotopenverfolgung“, *Angew. Chem.* **2012**, *124*, 6622–6626.
- [160] Y. Feng, N.-B. Xie, W.-B. Tao, J.-H. Ding, X.-J. You, C.-J. Ma, X. Zhang, C. Yi, X. Zhou, B.-F. Yuan, Y.-Q. Feng, „Transformation of 5-carboxycytosine to cytosine through C–C bond cleavage in human cells constitutes a novel pathway for DNA demethylation“, *CCS Chem.* **2020**, *2*, 994–1008.
- [161] S. Kriaucionis, N. Heintz, „The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain“, *Science* **2009**, *324*, 929–930.
- [162] M. Münzel, D. Globisch, T. Brückl, M. Wagner, V. Welzmler, S. Michalakis, M. Müller, M. Biel, T. Carell, „Quantification of the sixth DNA base hydroxymethylcytosine in the brain“, *Angew. Chem. Int. Ed.* **2010**, *49*, 5375–5377; „Quantitative Bestimmung der sechsten DNA-Base Hydroxymethylcytosin im Gehirn“, *Angew. Chem.* **2010**, *122*, 5503–5505.
- [163] K. E. Szulwach, X. Li, Y. Li, C. X. Song, H. Wu, Q. Dai, H. Irier, A. K. Upadhyay, M. Gearing, A. I. Levey, A. Vasanthakumar, L. A. Godley, Q. Chang, X. Cheng, C. He, P. Jin, „5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging“, *Nat. Neurosci.* **2011**, *14*, 1607–1616.
- [164] M. A. Hahn, R. Qiu, X. Wu, A. X. Li, H. Zhang, J. Wang, J. Jui, S. G. Jin, Y. Jiang, G. P. Pfeifer, Q. Lu, „Dynamics of 5-hydroxymethylcytosine and chromatin marks in mammalian neurogenesis“, *Cell Rep.* **2013**, *3*, 291–300.

-
- [165] M. Bachman, S. Uribe-Lewis, X. Yang, M. Williams, A. Murrell, S. Balasubramanian, „5-Hydroxymethylcytosine is a predominantly stable DNA modification“, *Nat. Chem.* **2014**, *6*, 1049–1055.
- [166] H. Wu, A. C. D'Alessio, S. Ito, Z. Wang, K. Cui, K. Zhao, Y. E. Sun, Y. Zhang, „Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells“, *Genes Dev.* **2011**, *25*, 679–684.
- [167] K. Williams, J. Christensen, M. T. Pedersen, J. V. Johansen, P. A. Cloos, J. Rappsilber, K. Helin, „TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity“, *Nature* **2011**, *473*, 343–348.
- [168] W. A. Pastor, U. J. Pape, Y. Huang, H. R. Henderson, R. Lister, M. Ko, E. M. McLoughlin, Y. Brudno, S. Mahapatra, P. Kapranov, M. Tahiliani, G. Q. Daley, X. S. Liu, J. R. Ecker, P. M. Milos, S. Agarwal, A. Rao, „Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells“, *Nature* **2011**, *473*, 394–397.
- [169] G. Ficz, M. R. Branco, S. Seisenberger, F. Santos, F. Krueger, T. A. Hore, C. J. Marques, S. Andrews, W. Reik, „Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation“, *Nature* **2011**, *473*, 398–402.
- [170] Y. Xu, F. Wu, L. Tan, L. Kong, L. Xiong, J. Deng, A. J. Barbera, L. Zheng, H. Zhang, S. Huang, J. Min, T. Nicholson, T. Chen, G. Xu, Y. Shi, K. Zhang, Y. G. Shi, „Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells“, *Mol. Cell* **2011**, *42*, 451–464.
- [171] K. E. Szulwach, X. Li, Y. Li, C. X. Song, J. W. Han, S. Kim, S. Namburi, K. Hermetz, J. J. Kim, M. K. Rudd, Y. S. Yoon, B. Ren, C. He, P. Jin, „Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells“, *PLoS Genet.* **2011**, *7*, e1002154.
- [172] H. Stroud, S. Feng, S. Morey Kinney, S. Pradhan, S. E. Jacobsen, „5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells“, *Genome Biol.* **2011**, *12*, R54.
- [173] S. D. Fouse, Y. Shen, M. Pellegrini, S. Cole, A. Meissner, L. Van Neste, R. Jaenisch, G. Fan, „Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation“, *Cell Stem Cell* **2008**, *2*, 160–169.
- [174] M. Bachman, S. Uribe-Lewis, X. Yang, H. E. Burgess, M. Iurlaro, W. Reik, A. Murrell, S. Balasubramanian, „5-Formylcytosine can be a stable DNA modification in mammals“, *Nat. Chem. Biol.* **2015**, *11*, 555–557.
- [175] E. A. Raiber, D. Beraldi, G. Ficz, H. E. Burgess, M. R. Branco, P. Murat, D. Oxley, M. J. Booth, W. Reik, S. Balasubramanian, „Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase“, *Genome Biol.* **2012**, *13*, R69.
- [176] C. X. Song, K. E. Szulwach, Q. Dai, Y. Fu, S. Q. Mao, L. Lin, C. Street, Y. Li, M. Poidevin, H. Wu, J. Gao, P. Liu, L. Li, G. L. Xu, P. Jin, C. He, „Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming“, *Cell* **2013**, *153*, 678–691.
-

- [177] L. Shen, H. Wu, D. Diep, S. Yamaguchi, A. C. D'Alessio, H. L. Fung, K. Zhang, Y. Zhang, „Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics“, *Cell* **2013**, *153*, 692–706.
- [178] C. G. Spruijt, F. Gnerlich, A. H. Smits, T. Pfaffeneder, P. W. Jansen, C. Bauer, M. Münzel, M. Wagner, M. Müller, F. Khan, H. C. Eberl, A. Mensinga, A. B. Brinkman, K. Lephikov, U. Müller, J. Walter, R. Boelens, H. van Ingen, H. Leonhardt, T. Carell, M. Vermeulen, „Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives“, *Cell* **2013**, *152*, 1146–1159.
- [179] M. Iurlaro, G. Ficz, D. Oxley, E. A. Raiber, M. Bachman, M. J. Booth, S. Andrews, S. Balasubramanian, W. Reik, „A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation“, *Genome Biol.* **2013**, *14*, R119.
- [180] T. Borowski, A. Bassan, P. E. M. Siegbahn, „Mechanism of dioxygen activation in 2-oxoglutarate-dependent enzymes: a hybrid DFT study“, *Chem.-Eur. J.* **2004**, *10*, 1031–1041.
- [181] K. Iqbal, S. G. Jin, G. P. Pfeifer, P. E. Szabó, „Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine“, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3642–3647.
- [182] T.-P. Gu, F. Guo, H. Yang, H.-P. Wu, G.-F. Xu, W. Liu, Z.-G. Xie, L. Shi, X. He, S.-g. Jin, K. Iqbal, Y. G. Shi, Z. Deng, P. E. Szabó, G. P. Pfeifer, J. Li, G.-L. Xu, „The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes“, *Nature* **2011**, *477*, 606–610.
- [183] Y. Huang, L. Chavez, X. Chang, X. Wang, W. A. Pastor, J. Kang, J. A. Zepeda-Martínez, U. J. Pape, S. E. Jacobsen, B. Peters, A. Rao, „Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells“, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 1361–1366.
- [184] Y. Costa, J. Ding, T. W. Theunissen, F. Faiola, T. A. Hore, P. V. Shliaha, M. Fidalgo, A. Saunders, M. Lawrence, S. Dietmann, S. Das, D. N. Lévassieur, Z. Li, M. Xu, W. Reik, J. C. Silva, J. Wang, „NANOG-dependent function of TET1 and TET2 in establishment of pluripotency“, *Nature* **2013**, *495*, 370–374.
- [185] L. de la Rica, J. Rodríguez-Ubrea, M. García, A. B. Islam, J. M. Urquiza, H. Hernando, J. Christensen, K. Helin, C. Gómez-Vaquero, E. Ballestar, „PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation“, *Genome Biol.* **2013**, *14*, R99.
- [186] J. Xiong, Z. Zhang, J. Chen, H. Huang, Y. Xu, X. Ding, Y. Zheng, R. Nishinakamura, G. L. Xu, H. Wang, S. Chen, S. Gao, B. Zhu, „Cooperative action between SALL4A and TET proteins in stepwise oxidation of 5-methylcytosine“, *Mol. Cell* **2016**, *64*, 913–925.
- [187] J. A. Hackett, R. Sengupta, J. J. Zylitz, K. Murakami, C. Lee, T. A. Down, M. A. Surani, „Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine“, *Science* **2013**, *339*, 448–452.
- [188] J. J. Vincent, Y. Huang, P. Y. Chen, S. Feng, J. H. Calvopiña, K. Nee, S. A. Lee, T. Le, A. J. Yoon, K. Faull, G. Fan, A. Rao, S. E. Jacobsen, M. Pellegrini, A. T. Clark, „Stage-specific roles for Tet1 and Tet2 in DNA demethylation in primordial germ cells“, *Cell Stem Cell* **2013**, *12*, 470–478.

- [189] M. M. Dawlaty, A. Breiling, T. Le, G. Raddatz, M. I. Barrasa, A. W. Cheng, Q. Gao, B. E. Powell, Z. Li, M. Xu, K. F. Faull, F. Lyko, R. Jaenisch, „Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development“, *Dev. Cell* **2013**, *24*, 310–323.
- [190] D. K. Ma, M. H. Jang, J. U. Guo, Y. Kitabatake, M. L. Chang, N. Pow-Anpongkul, R. A. Flavell, B. Lu, G. L. Ming, H. Song, „Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis“, *Science* **2009**, *323*, 1074–1077.
- [191] Y. Xu, C. Xu, A. Kato, W. Tempel, J. G. Abreu, C. Bian, Y. Hu, D. Hu, B. Zhao, T. Cerovina, J. Diao, F. Wu, H. H. He, Q. Cui, E. Clark, C. Ma, A. Barbara, G. J. Veenstra, G. Xu, U. B. Kaiser, X. S. Liu, S. P. Sugrue, X. He, J. Min, Y. Kato, Y. G. Shi, „Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development“, *Cell* **2012**, *151*, 1200–1213.
- [192] T. Li, D. Yang, J. Li, Y. Tang, J. Yang, W. Le, „Critical role of Tet3 in neural progenitor cell maintenance and terminal differentiation“, *Mol. Neurobiol.* **2015**, *51*, 142–154.
- [193] M. Ko, Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski, A. Rao, „Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2“, *Nature* **2010**, *468*, 839–843.
- [194] S. Turcan, D. Rohle, A. Goenka, L. A. Walsh, F. Fang, E. Yilmaz, C. Campos, A. W. Fabius, C. Lu, P. S. Ward, C. B. Thompson, A. Kaufman, O. Guryanova, R. Levine, A. Heguy, A. Viale, L. G. Morris, J. T. Huse, I. K. Mellinghoff, T. A. Chan, „IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype“, *Nature* **2012**, *483*, 479–483.
- [195] Y. Kudo, K. Tateishi, K. Yamamoto, S. Yamamoto, Y. Asaoka, H. Ijichi, G. Nagae, H. Yoshida, H. Aburatani, K. Koike, „Loss of 5-hydroxymethylcytosine is accompanied with malignant cellular transformation“, *Cancer Sci.* **2012**, *103*, 670–676.
- [196] C. G. Lian, Y. Xu, C. Ceol, F. Wu, A. Larson, K. Dresser, W. Xu, L. Tan, Y. Hu, Q. Zhan, C. W. Lee, D. Hu, B. Q. Lian, S. Kleffel, Y. Yang, J. Neiswender, A. J. Khorasani, R. Fang, C. Lezcano, L. M. Duncan, R. A. Scolyer, J. F. Thompson, H. Kakavand, Y. Houvras, L. I. Zon, M. C. Mihm Jr., U. B. Kaiser, T. Schatton, B. A. Woda, G. F. Murphy, Y. G. Shi, „Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma“, *Cell* **2012**, *150*, 1135–1146.
- [197] H. Yang, Y. Liu, F. Bai, J. Y. Zhang, S. H. Ma, J. Liu, Z. D. Xu, H. G. Zhu, Z. Q. Ling, D. Ye, K. L. Guan, Y. Xiong, „Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation“, *Oncogene* **2013**, *32*, 663–669.
- [198] C. Liu, L. Liu, X. Chen, J. Shen, J. Shan, Y. Xu, Z. Yang, L. Wu, F. Xia, P. Bie, Y. Cui, X. W. Bian, C. Qian, „Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of TET1“, *PLoS One* **2013**, *8*, e62828.
- [199] M. Jan, T. M. Snyder, M. R. Corces-Zimmerman, P. Vyas, I. L. Weissman, S. R. Quake, R. Majeti, „Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia“, *Sci. Transl. Med.* **2012**, *4*, 149ra118.

- [200] R. Itzykson, O. Kosmider, A. Renneville, M. Morabito, C. Preudhomme, C. Berthon, L. Adès, P. Fenaux, U. Platzbecker, O. Gagey, P. Rameau, G. Meurice, C. Oréar, F. Delhommeau, O. A. Bernard, M. Fontenay, W. Vainchenker, N. Droin, E. Solary, „Clonal architecture of chronic myelomonocytic leukemias“, *Blood* **2013**, *121*, 2186–2198.
- [201] B. E. Baysal, R. E. Ferrell, J. E. Willett-Brozick, E. C. Lawrence, D. Myssiorek, A. Bosch, A. van der Mey, P. E. Taschner, W. S. Rubinstein, E. N. Myers, C. W. Richard III, C. J. Cornelisse, P. Devilee, B. Devlin, „Mutations in *SDHD*, a mitochondrial complex II gene, in hereditary paraganglioma“, *Science* **2000**, *287*, 848–851.
- [202] I. P. Tomlinson, N. A. Alam, A. J. Rowan, E. Barclay, E. E. Jaeger, D. Kelsell, I. Leigh, P. Gorman, H. Lamlum, S. Rahman, R. R. Roylance, S. Olpin, S. Bevan, K. Barker, N. Hearle, R. S. Houlston, M. Kiuru, R. Lehtonen, A. Karhu, S. Vilkki, P. Laiho, C. Eklund, O. Vierimaa, K. Aittomäki, M. Hietala, P. Sistonen, A. Paetau, R. Salovaara, R. Herva, V. Launonen, L. A. Aaltonen, „Germline mutations in *FH* predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer“, *Nat. Genet.* **2002**, *30*, 406–410.
- [203] H. Shim, C. Dolde, B. C. Lewis, C. S. Wu, G. Dang, R. A. Jungmann, R. Dalla-Favera, C. V. Dang, „c-Myc transactivation of *LDH-A*: implications for tumor metabolism and growth“, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6658–6663.
- [204] S. P. Mathupala, C. Heese, P. L. Pedersen, „Glucose catabolism in cancer cells. The type II hexokinase promoter contains functionally active response elements for the tumor suppressor p53“, *J. Biol. Chem.* **1997**, *272*, 22776–22780.
- [205] J. Yun, C. Rago, I. Cheong, R. Pagliarini, P. Angenendt, H. Rajagopalan, K. Schmidt, J. K. Willson, S. Markowitz, S. Zhou, L. A. Diaz Jr., V. E. Velculescu, C. Lengauer, K. W. Kinzler, B. Vogelstein, N. Papadopoulos, „Glucose deprivation contributes to the development of *KRAS* pathway mutations in tumor cells“, *Science* **2009**, *325*, 1555–1559.
- [206] W. Xu, H. Yang, Y. Liu, Y. Yang, P. Wang, S. H. Kim, S. Ito, C. Yang, P. Wang, M. T. Xiao, L. X. Liu, W. Q. Jiang, J. Liu, J. Y. Zhang, B. Wang, S. Frye, Y. Zhang, Y. H. Xu, Q. Y. Lei, K. L. Guan, S. M. Zhao, Y. Xiong, „Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α -ketoglutarate-dependent dioxygenases“, *Cancer Cell* **2011**, *19*, 17–30.
- [207] M. Xiao, H. Yang, W. Xu, S. Ma, H. Lin, H. Zhu, L. Liu, Y. Liu, C. Yang, Y. Xu, S. Zhao, D. Ye, Y. Xiong, K. L. Guan, „Inhibition of α -KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of *FH* and *SDH* tumor suppressors“, *Genes Dev.* **2012**, *26*, 1326–1338.
- [208] M. Agathocleous, C. E. Meacham, R. J. Burgess, E. Piskounova, Z. Zhao, G. M. Crane, B. L. Cowin, E. Bruner, M. M. Murphy, W. Chen, G. J. Spangrude, Z. Hu, R. J. DeBerardinis, S. J. Morrison, „Ascorbate regulates haematopoietic stem cell function and leukaemogenesis“, *Nature* **2017**, *549*, 476–481.
- [209] D. Wu, D. Hu, H. Chen, G. Shi, I. S. Fetahu, F. Wu, K. Rabidou, R. Fang, L. Tan, S. Xu, H. Liu, C. Argueta, L. Zhang, F. Mao, G. Yan, J. Chen, Z. Dong, R. Lv, Y. Xu, M. Wang, Y. Ye, S. Zhang, D. Duquette, S. Geng, C. Yin, C. G. Lian, G. F. Murphy, G. K. Adler, R. Garg, L. Lynch, P. Yang, Y. Li, F. Lan, J. Fan, Y. Shi, Y. G. Shi, „Glucose-regulated phosphorylation of *TET2* by *AMPK* reveals a pathway linking diabetes to cancer“, *Nature* **2018**, *559*, 637–641.

- [210] J. Alves, G. M. Dexheimer, L. Reckzigel, M. Goettert, V. Biolchi, A. L. Abujamra, „Changes in IDH2, TET2 and KDM2B gene expression after treatment with classic chemotherapeutic agents and decitabine in myelogenous leukemia cell lines“, *J. Hematol.* **2019**, *8*, 89–101.
- [211] A. Da Cruz Paula, E. M. da Silva, S. E. Segura, F. Pareja, R. Bi, P. Selenica, S. H. Kim, L. Ferrando, M. Vahdatinia, R. A. Soslow, A. Vidal, S. Gatus, C. G. Przybycin, N. R. Abu-Rustum, X. Matias-Guiu, B. P. Rubin, J. S. Reis-Filho, D. F. DeLair, B. Weigelt, „Genomic profiling of primary and recurrent adult granulosa cell tumors of the ovary“, *Mod. Pathol.* **2020**, *33*, 1606–1617.
- [212] The ENCODE Project Consortium, „An integrated encyclopedia of DNA elements in the human genome“, *Nature* **2012**, *489*, 57–74.
- [213] S. Djebali, C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J. Lagarde, W. Lin, F. Schlesinger, C. Xue, G. K. Marinov, J. Khatun, B. A. Williams, C. Zaleski, J. Rozowsky, M. Röder, F. Kokocinski, R. F. Abdelhamid, T. Alioto, I. Antoshechkin, M. T. Baer, N. S. Bar, P. Batut, K. Bell, I. Bell, S. Chakraborty, X. Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R. Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M. J. Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R. Johnson, P. Kapranov, B. King, C. Kingswood, O. J. Luo, E. Park, K. Persaud, J. B. Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L.-H. See, A. Shahab, J. Skancke, A. M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J. Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow, M. Gerstein, T. Hubbard, A. Reymond, S. E. Antonarakis, G. Hannon, M. C. Giddings, Y. Ruan, B. Wold, P. Carninci, R. Guigó, T. R. Gingeras, „Landscape of transcription in human cells“, *Nature* **2012**, *489*, 101–108.
- [214] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, „Identification of novel genes coding for small expressed RNAs“, *Science* **2001**, *294*, 853–858.
- [215] N. C. Lau, L. P. Lim, E. G. Weinstein, D. P. Bartel, „An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*“, *Science* **2001**, *294*, 858–862.
- [216] R. C. Lee, V. Ambros, „An extensive class of small RNAs in *Caenorhabditis elegans*“, *Science* **2001**, *294*, 862–864.
- [217] S. T. Grivna, E. Beyret, Z. Wang, H. Lin, „A novel class of small RNAs in mouse spermatogenic cells“, *Genes Dev.* **2006**, *20*, 1709–1714.
- [218] T. Watanabe, A. Takeda, T. Tsukiyama, K. Mise, T. Okuno, H. Sasaki, N. Minami, H. Imai, „Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes“, *Genes Dev.* **2006**, *20*, 1732–1743.
- [219] A. Girard, R. Sachidanandam, G. J. Hannon, M. A. Carmell, „A germline-specific class of small RNAs binds mammalian Piwi proteins“, *Nature* **2006**, *442*, 199–202.
- [220] A. Aravin, D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, N. Iovino, P. Morris, M. J. Brownstein, S. Kuramochi-Miyagawa, T. Nakano, M. Chien, J. J. Russo, J. Ju, R. Sheridan, C. Sander, M. Zavolan, T. Tuschl, „A novel class of small RNAs bind to MILI protein in mouse testes“, *Nature* **2006**, *442*, 203–207.
- [221] Y. Lee, C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Rådmark, S. Kim, V. N. Kim, „The nuclear RNase III Drosha initiates microRNA processing“, *Nature* **2003**, *425*, 415–419.

- [222] J. Han, Y. Lee, K.-H. Yeom, Y.-K. Kim, H. Jin, V. N. Kim, „The Drosha-DGCR8 complex in primary microRNA processing“, *Genes Dev.* **2004**, *18*, 3016–3027.
- [223] E. Bernstein, A. A. Caudy, S. M. Hammond, G. J. Hannon, „Role for a bidentate ribonuclease in the initiation step of RNA interference“, *Nature* **2001**, *409*, 363–366.
- [224] J. J. Ipsaro, A. D. Haase, S. R. Knott, L. Joshua-Tor, G. J. Hannon, „The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis“, *Nature* **2012**, *491*, 279–283.
- [225] H. Nishimasu, H. Ishizu, K. Saito, S. Fukuhara, M. K. Kamatani, L. Bonnefond, N. Matsumoto, T. Nishizawa, K. Nakanaga, J. Aoki, R. Ishitani, H. Siomi, M. C. Siomi, O. Nureki, „Structure and function of Zucchini endoribonuclease in piRNA biogenesis“, *Nature* **2012**, *491*, 284–287.
- [226] B. W. Han, W. Wang, C. Li, Z. Weng, P. D. Zamore, „piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production“, *Science* **2015**, *348*, 817–821.
- [227] N. Pomper, Y. Liu, M. L. Hoye, J. D. Dougherty, T. M. Miller, „CNS microRNA profiles: a database for cell type enriched microRNA expression across the mouse central nervous system“, *Sci. Rep.* **2020**, *10*, 4921.
- [228] J. Brennecke, A. A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, G. J. Hannon, „Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*“, *Cell* **2007**, *128*, 1089–1103.
- [229] J. Brennecke, C. D. Malone, A. A. Aravin, R. Sachidanandam, A. Stark, G. J. Hannon, „An epigenetic role for maternally inherited piRNAs in transposon silencing“, *Science* **2008**, *322*, 1387–1392.
- [230] K. M. Nishida, K. Saito, T. Mori, Y. Kawamura, T. Nagami-Okada, S. Inagaki, H. Siomi, M. C. Siomi, „Gene silencing mechanisms mediated by Aubergine–piRNA complexes in *Drosophila* male gonad“, *Rna* **2007**, *13*, 1911–1922.
- [231] S. Diederichs, D. A. Haber, „Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression“, *Cell* **2007**, *131*, 1097–1108.
- [232] A. Turchinovich, B. Burwinkel, „Distinct AGO1 and AGO2 associated miRNA profiles in human cells and blood plasma“, *RNA Biol.* **2012**, *9*, 1066–1075.
- [233] A. Webster, S. Li, J. K. Hur, M. Wachsmuth, J. S. Bois, E. M. Perkins, D. J. Patel, A. A. Aravin, „Aub and Ago3 are recruited to nuage through two mechanisms to form a ping-pong complex assembled by Krimper“, *Mol. Cell* **2015**, *59*, 564–575.
- [234] S. M. Hammond, E. Bernstein, D. Beach, G. J. Hannon, „An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells“, *Nature* **2000**, *404*, 293–296.
- [235] Y. Zeng, R. Yi, B. R. Cullen, „MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms“, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9779–9784.
- [236] A. Verdel, S. Jia, S. Gerber, T. Sugiyama, S. Gygi, S. I. S. Grewal, D. Moazed, „RNAi-mediated targeting of heterochromatin by the RITS complex“, *Science* **2004**, *303*, 672–676.

- [237] T. A. Volpe, C. Kidner, I. M. Hall, G. Teng, S. I. Grewal, R. A. Martienssen, „Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi“, *Science* **2002**, *297*, 1833–1837.
- [238] D. Castanotto, S. Tommasi, M. Li, H. Li, S. Yanow, G. P. Pfeifer, J. J. Rossi, „Short hairpin RNA-directed cytosine (CpG) methylation of the *RASSF1A* gene promoter in HeLa cells“, *Mol. Ther.* **2005**, *12*, 179–183.
- [239] D. H. Kim, P. Sætrom, O. Snøve Jr., J. J. Rossi, „MicroRNA-directed transcriptional gene silencing in mammalian cells“, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 16230–16235.
- [240] B. Bánfai, H. Jia, J. Khatun, E. Wood, B. Risk, W. E. Gundling Jr., A. Kundaje, H. P. Gunawardena, Y. Yu, L. Xie, K. Krajewski, B. D. Strahl, X. Chen, P. Bickel, M. C. Giddings, J. B. Brown, L. Lipovich, „Long noncoding RNAs are rarely translated in two human cell lines“, *Genome Res.* **2012**, *22*, 1646–1657.
- [241] P. Kapranov, J. Cheng, S. Dike, D. A. Nix, R. Dutttagupta, A. T. Willingham, P. F. Stadler, J. Hertel, J. Hackermüller, I. L. Hofacker, I. Bell, E. Cheung, J. Drenkow, E. Dumais, S. Patel, G. Helt, M. Ganesh, S. Ghosh, A. Piccolboni, V. Sementchenko, H. Tamma, T. R. Gingeras, „RNA maps reveal new RNA classes and a possible function for pervasive transcription“, *Science* **2007**, *316*, 1484–1488.
- [242] Y. Jeon, J. T. Lee, „YY1 tethers Xist RNA to the inactive X nucleation center“, *Cell* **2011**, *146*, 119–133.
- [243] K. M. Schmitz, C. Mayer, A. Postepska, I. Grummt, „Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes“, *Genes Dev.* **2010**, *24*, 2264–2269.
- [244] I. Martianov, A. Ramadass, A. Serra Barros, N. Chow, A. Akoulitchev, „Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript“, *Nature* **2007**, *445*, 666–670.
- [245] J. L. Rinn, M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H. Goodnough, J. A. Helms, P. J. Farnham, E. Segal, H. Y. Chang, „Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs“, *Cell* **2007**, *129*, 1311–1323.
- [246] J. Zhao, B. K. Sun, J. A. Erwin, J. J. Song, J. T. Lee, „Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome“, *Science* **2008**, *322*, 750–756.
- [247] T. Nagano, J. A. Mitchell, L. A. Sanz, F. M. Pauler, A. C. Ferguson-Smith, R. Feil, P. Fraser, „The *Air* noncoding RNA epigenetically silences transcription by targeting G9a to chromatin“, *Science* **2008**, *322*, 1717–1720.
- [248] A. M. Khalil, M. Guttman, M. Huarte, M. Garber, A. Raj, D. Rivea Morales, K. Thomas, A. Presser, B. E. Bernstein, A. van Oudenaarden, A. Regev, E. S. Lander, J. L. Rinn, „Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression“, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 11667–11672.
- [249] N. Thakur, V. K. Tiwari, H. Thomassin, R. R. Pandey, M. Kanduri, A. Göndör, T. Grange, R. Ohlsson, C. Kanduri, „An antisense RNA regulates the bidirectional silencing property of the *Kcnq1* imprinting control region“, *Mol. Cell. Biol.* **2004**, *24*, 7855–7862.

- [250] S. Schoenfelder, G. Smits, P. Fraser, W. Reik, R. Paro, „Non-coding transcripts in the *H19* imprinting control region mediate gene silencing in transgenic *Drosophila*“, *EMBO Rep.* **2007**, *8*, 1068–1073.
- [251] G. Piekarski, „Cytologische Untersuchungen an Paratyphus-und Colibakterien“, *Arch. Mikrobiol.* **1937**, *8*, 428–439.
- [252] M. Y. Tolstorukov, K. Virnik, V. B. Zhurkin, S. Adhya, „Organization of DNA in a bacterial nucleoid“, *BMC Microbiol.* **2016**, *16*, 22.
- [253] A. Piovesan, M. C. Pelleri, F. Antonaros, P. Strippoli, M. Caracausi, L. Vitale, „On the length, weight and GC content of the human genome“, *BMC Res. Notes* **2019**, *12*, 106.
- [254] S. M. West, R. Rohs, R. S. Mann, B. Honig, „Electrostatic interactions between arginines and the minor groove in the nucleosome“, *J. Biomol. Struct. Dyn.* **2010**, *27*, 861–866.
- [255] A. T. Fenley, D. A. Adams, A. V. Onufriev, „Charge state of the globular histone core controls stability of the nucleosome“, *Biophys. J.* **2010**, *99*, 1577–1585.
- [256] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, T. J. Richmond, „Crystal structure of the nucleosome core particle at 2.8 Å resolution“, *Nature* **1997**, *389*, 251–260.
- [257] M. Y. Tolstorukov, A. V. Colasanti, D. M. McCandlish, W. K. Olson, V. B. Zhurkin, „A novel roll-and-slide mechanism of DNA folding in chromatin: implications for nucleosome positioning“, *J. Mol. Biol.* **2007**, *371*, 725–738.
- [258] M. García-Pérez, M. Pinto, J. A. Subirana, „Nonsequence-specific arginine interactions in the nucleosome core particle“, *Biopolymers* **2003**, *69*, 432–439.
- [259] D. Wang, N. B. Ulyanov, V. B. Zhurkin, „Sequence-dependent kink-and-slide deformations of nucleosomal DNA facilitated by histone arginines bound in the minor groove“, *J. Biomol. Struct. Dyn.* **2010**, *27*, 843–859.
- [260] S. Pennings, S. Muyldermans, G. Meersseman, L. Wyns, „Formation, stability and core histone positioning of nucleosomes reassembled on bent and other nucleosome-derived DNA“, *J. Mol. Biol.* **1989**, *207*, 183–192.
- [261] P. T. Lowary, J. Widom, „New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning“, *J. Mol. Biol.* **1998**, *276*, 19–42.
- [262] C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, T. J. Richmond, „Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution“, *J. Mol. Biol.* **2002**, *319*, 1097–1113.
- [263] J. P. Whitlock Jr., R. T. Simpson, „Removal of histone H1 exposes a fifty base pair DNA segment between nucleosomes“, *Biochemistry* **1976**, *15*, 3307–3314.
- [264] P. Oudet, M. Gross-Bellard, P. Chambon, „Electron microscopic and biochemical evidence that chromatin structure is a repeating unit“, *Cell* **1975**, *4*, 281–300.
- [265] J. Allan, T. Mitchell, N. Harborne, L. Bohm, C. Crane-Robinson, „Roles of H1 domains in determining higher order chromatin structure and H1 location“, *J. Mol. Biol.* **1986**, *187*, 591–601.

-
- [266] B. Dorigo, T. Schalch, A. Kulangara, S. Duda, R. R. Schroeder, T. J. Richmond, „Nucleosome arrays reveal the two-start organization of the chromatin fiber“, *Science* **2004**, *306*, 1571–1573.
- [267] Y. B. Zhou, S. E. Gerchman, V. Ramakrishnan, A. Travers, S. Muyldermans, „Position and orientation of the globular domain of linker histone H5 on the nucleosome“, *Nature* **1998**, *395*, 402–405.
- [268] B.-R. Zhou, H. Feng, H. Kato, L. Dai, Y. Yang, Y. Zhou, Y. Bai, „Structural insights into the histone H1-nucleosome complex“, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 19390–19395.
- [269] Y. Fan, T. Nikitina, E. M. Morin-Kensicki, J. Zhao, T. R. Magnuson, C. L. Woodcock, A. I. Skoultchi, „H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo“, *Mol. Cell. Biol.* **2003**, *23*, 4559–4572.
- [270] Y. Fan, T. Nikitina, J. Zhao, T. J. Fleury, R. Bhattacharyya, E. E. Bouhassira, A. Stein, C. L. Woodcock, A. I. Skoultchi, „Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation“, *Cell* **2005**, *123*, 1199–1212.
- [271] Y. Cui, C. Bustamante, „Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure“, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 127–132.
- [272] J. Bednar, R. A. Horowitz, S. A. Grigoryev, L. M. Carruthers, J. C. Hansen, A. J. Koster, C. L. Woodcock, „Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin“, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14173–14178.
- [273] P. J. Robinson, L. Fairall, V. A. Huynh, D. Rhodes, „EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure“, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6506–6511.
- [274] T. Schalch, S. Duda, D. F. Sargent, T. J. Richmond, „X-ray structure of a tetranucleosome and its implications for the chromatin fibre“, *Nature* **2005**, *436*, 138–141.
- [275] A. S. Belmont, M. B. Braunfeld, J. W. Sedat, D. A. Agard, „Large-scale chromatin structural domains within mitotic and interphase chromosomes in vivo and in vitro“, *Chromosoma* **1989**, *98*, 129–143.
- [276] A. S. Belmont, K. Bruce, „Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure“, *J. Cell Biol.* **1994**, *127*, 287–302.
- [277] B. Rydberg, W. R. Holley, I. S. Mian, A. Chatterjee, „Chromatin conformation in living cells: support for a zig-zag model of the 30 nm chromatin fiber“, *J. Mol. Biol.* **1998**, *284*, 71–84.
- [278] K. Bystricky, P. Heun, L. Gehlen, J. Langowski, S. M. Gasser, „Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques“, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16495–16500.
- [279] M. Eltsov, K. M. Maclellan, K. Maeshima, A. S. Frangakis, J. Dubochet, „Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ“, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 19732–19737.
-

- [280] Y. Nishino, M. Eltsov, Y. Joti, K. Ito, H. Takata, Y. Takahashi, S. Hihara, A. S. Frangakis, N. Imamoto, T. Ishikawa, K. Maeshima, „Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure“, *EMBO J.* **2012**, *31*, 1644–1653.
- [281] T. Hirano, T. J. Mitchison, „A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro“, *Cell* **1994**, *79*, 449–458.
- [282] T. Hirano, R. Kobayashi, M. Hirano, „Condensins, chromosome condensation protein complexes containing XCAP-C, XCAP-E and a *Xenopus* homolog of the *Drosophila* Barren protein“, *Cell* **1997**, *89*, 511–521.
- [283] F. Cubizolles, V. Legagneux, R. Le Guellec, I. Chartrain, R. Uzbekov, C. Ford, K. Le Guellec, „pEg7, a new *Xenopus* protein required for mitotic chromosome condensation in egg extracts“, *J. Cell Biol.* **1998**, *143*, 1437–1446.
- [284] T. Sutani, T. Yuasa, T. Tomonaga, N. Dohmae, K. Takio, M. Yanagida, „Fission yeast condensin complex: essential roles of non-SMC subunits for condensation and Cdc2 phosphorylation of Cut3/SMC4“, *Genes Dev.* **1999**, *13*, 2271–2283.
- [285] Q. Feng, H. Wang, H. H. Ng, H. Erdjument-Bromage, P. Tempst, K. Struhl, Y. Zhang, „Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain“, *Curr. Biol.* **2002**, *12*, 1052–1058.
- [286] F. Xu, K. Zhang, M. Grunstein, „Acetylation in histone H3 globular domain regulates gene expression in yeast“, *Cell* **2005**, *121*, 375–385.
- [287] M. A. Dawson, A. J. Bannister, B. Göttgens, S. D. Foster, T. Bartke, A. R. Green, T. Kouzarides, „JAK2 phosphorylates histone H3Y41 and excludes HP1 α from chromatin“, *Nature* **2009**, *461*, 819–822.
- [288] J. Allan, N. Harborne, D. C. Rau, H. Gould, „Participation of core histone "tails" in the stabilization of the chromatin solenoid“, *J. Cell Biol.* **1982**, *93*, 285–297.
- [289] D. Y. Lee, J. J. Hayes, D. Pruss, A. P. Wolffe, „A positive role for histone acetylation in transcription factor access to nucleosomal DNA“, *Cell* **1993**, *72*, 73–84.
- [290] M. Vettese-Dadey, P. Walter, H. Chen, L. J. Juan, J. L. Workman, „Role of the histone amino termini in facilitated binding of a transcription factor, GAL4-AH, to nucleosome cores“, *Mol. Cell. Biol.* **1994**, *14*, 970–981.
- [291] K. J. Polach, P. T. Lowary, J. Widom, „Effects of core histone tail domains on the equilibrium constants for dynamic DNA site accessibility in nucleosomes“, *J. Mol. Biol.* **2000**, *298*, 211–223.
- [292] Z. Yang, C. Zheng, C. Thiriet, J. J. Hayes, „The core histone N-terminal tail domains negatively regulate binding of transcription factor IIIA to a nucleosome containing a 5S RNA gene via a novel mechanism“, *Mol. Cell. Biol.* **2005**, *25*, 241–249.
- [293] D. M. Phillips, „The presence of acetyl groups of histones“, *Biochem. J.* **1963**, *87*, 258–263.

- [294] V. G. Allfrey, R. Faulkner, A. E. Mirsky, „Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis“, *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *51*, 786–794.
- [295] J. Borrow, V. P. Stanton Jr., J. M. Andresen, R. Becher, F. G. Behm, R. S. Chaganti, C. I. Civin, C. Disteche, I. Dube, A. M. Frischauf, D. Horsman, F. Mitelman, S. Volinia, A. E. Watmore, D. E. Housman, „The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein“, *Nat. Genet.* **1996**, *14*, 33–41.
- [296] A. J. Bannister, T. Kouzarides, „The CBP co-activator is a histone acetyltransferase“, *Nature* **1996**, *384*, 641–643.
- [297] A. F. Neuwald, D. Landsman, „GCN5-related histone *N*-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein“, *Trends Biochem. Sci.* **1997**, *22*, 154–155.
- [298] M. H. Kuo, J. E. Brownell, R. E. Sobel, T. A. Ranalli, R. G. Cook, D. G. Edmondson, S. Y. Roth, C. D. Allis, „Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines“, *Nature* **1996**, *383*, 269–272.
- [299] A. Hilfiker, D. Hilfiker-Kleiner, A. Pannuti, J. C. Lucchesi, „*mof*, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*“, *EMBO J.* **1997**, *16*, 2054–2060.
- [300] S. Takechi, T. Nakayama, „Sas3 is a histone acetyltransferase and requires a zinc finger motif“, *Biochem. Biophys. Res. Commun.* **1999**, *266*, 405–410.
- [301] L. Hong, G. P. Schroth, H. R. Matthews, P. Yau, E. M. Bradbury, „Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA“, *J. Biol. Chem.* **1993**, *268*, 305–314.
- [302] T. R. Hebbes, A. W. Thorne, C. Crane-Robinson, „A direct link between core histone acetylation and transcriptionally active chromatin“, *EMBO J.* **1988**, *7*, 1395–1402.
- [303] M. Garcia-Ramirez, C. Rocchini, J. Ausio, „Modulation of chromatin folding by histone acetylation“, *J. Biol. Chem.* **1995**, *270*, 17923–17928.
- [304] C. Tse, T. Sera, A. P. Wolffe, J. C. Hansen, „Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III“, *Mol. Cell. Biol.* **1998**, *18*, 4629–4638.
- [305] M. Shogren-Knaak, H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie, C. L. Peterson, „Histone H4-K16 acetylation controls chromatin structure and protein interactions“, *Science* **2006**, *311*, 844–847.
- [306] Z. Wang, C. Zang, J. A. Rosenfeld, D. E. Schones, A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, W. Peng, M. Q. Zhang, K. Zhao, „Combinatorial patterns of histone acetylations and methylations in the human genome“, *Nat. Genet.* **2008**, *40*, 897–903.
- [307] M. Braunstein, A. B. Rose, S. G. Holmes, C. D. Allis, J. R. Broach, „Transcriptional silencing in yeast is associated with reduced nucleosome acetylation“, *Genes Dev.* **1993**, *7*, 592–604.

- [308] J. Taunton, C. A. Hassig, S. L. Schreiber, „A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p“, *Science* **1996**, 272, 408–411.
- [309] C. A. Hassig, T. C. Fleischer, A. N. Billin, S. L. Schreiber, D. E. Ayer, „Histone deacetylase activity is required for full transcriptional repression by mSin3A“, *Cell* **1997**, 89, 341–347.
- [310] P. A. Grant, L. Duggan, J. Côté, S. M. Roberts, J. E. Brownell, R. Candau, R. Ohba, T. Owen-Hughes, C. D. Allis, F. Winston, S. L. Berger, J. L. Workman, „Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex“, *Genes Dev.* **1997**, 11, 1640–1650.
- [311] Y. Xue, J. Wong, G. T. Moreno, M. K. Young, J. Côté, W. Wang, „NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities“, *Mol. Cell* **1998**, 2, 851–861.
- [312] S. Allard, R. T. Utley, J. Savard, A. Clarke, P. Grant, C. J. Brandl, L. Pillus, J. L. Workman, J. Côté, „NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p“, *EMBO J.* **1999**, 18, 5108–5119.
- [313] H. Tamaru, X. Zhang, D. McMillen, P. B. Singh, J. Nakayama, S. I. Grewal, C. D. Allis, X. Cheng, E. U. Selker, „Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*“, *Nat. Genet.* **2003**, 34, 75–79.
- [314] B. Xiao, C. Jing, J. R. Wilson, P. A. Walker, N. Vasisht, G. Kelly, S. Howell, I. A. Taylor, G. M. Blackburn, S. J. Gamblin, „Structure and catalytic mechanism of the human histone methyltransferase SET7/9“, *Nature* **2003**, 421, 652–656.
- [315] X. Zhang, Z. Yang, S. I. Khan, J. R. Horton, H. Tamaru, E. U. Selker, X. Cheng, „Structural basis for the product specificity of histone lysine methyltransferases“, *Mol. Cell* **2003**, 12, 177–185.
- [316] R. E. Collins, M. Tachibana, H. Tamaru, K. M. Smith, D. Jia, X. Zhang, E. U. Selker, Y. Shinkai, X. Cheng, „*In vitro* and *in vivo* analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases“, *J. Biol. Chem.* **2005**, 280, 5563–5570.
- [317] J. Tang, A. Frankel, R. J. Cook, S. Kim, W. K. Paik, K. R. Williams, S. Clarke, H. R. Herschman, „PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells“, *J. Biol. Chem.* **2000**, 275, 7723–7730.
- [318] T. L. Branscombe, A. Frankel, J. H. Lee, J. R. Cook, Z. Yang, S. Pestka, S. Clarke, „PRMT5 (Janus kinase-binding protein 1) catalyzes the formation of symmetric dimethylarginine residues in proteins“, *J. Biol. Chem.* **2001**, 276, 32971–32976.
- [319] D. Cheng, J. Cote, S. Shaaban, M. T. Bedford, „The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing“, *Mol. Cell* **2007**, 25, 71–83.
- [320] H. Wang, Z. Q. Huang, L. Xia, Q. Feng, H. Erdjument-Bromage, B. D. Strahl, S. D. Briggs, C. D. Allis, J. Wong, P. Tempst, Y. Zhang, „Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor“, *Science* **2001**, 293, 853–857.
- [321] E. Fabbrizio, S. El Messaoudi, J. Polanowska, C. Paul, J. R. Cook, J. H. Lee, V. Negre, M. Rousset, S. Pestka, A. Le Cam, C. Sardet, „Negative regulation of transcription by the type II arginine methyltransferase PRMT5“, *EMBO Rep.* **2002**, 3, 641–645.

-
- [322] A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, K. Zhao, „High-resolution profiling of histone methylations in the human genome“, *Cell* **2007**, *129*, 823–837.
- [323] P. Byvoet, G. R. Shepherd, J. M. Hardin, B. J. Noland, „The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells“, *Arch. Biochem. Biophys.* **1972**, *148*, 558–567.
- [324] Y. Shi, F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero, Y. Shi, „Histone demethylation mediated by the nuclear amine oxidase homolog LSD1“, *Cell* **2004**, *119*, 941–953.
- [325] G. L. Cuthbert, S. Daujat, A. W. Snowden, H. Erdjument-Bromage, T. Hagiwara, M. Yamada, R. Schneider, P. D. Gregory, P. Tempst, A. J. Bannister, T. Kouzarides, „Histone deimination antagonizes arginine methylation“, *Cell* **2004**, *118*, 545–553.
- [326] Y. Wang, J. Wysocka, J. Sayegh, Y. H. Lee, J. R. Perlin, L. Leonelli, L. S. Sonbuchner, C. H. McDonald, R. G. Cook, Y. Dou, R. G. Roeder, S. Clarke, M. R. Stallcup, C. D. Allis, S. A. Coonrod, „Human PAD4 regulates histone arginine methylation levels via demethylination“, *Science* **2004**, *306*, 279–283.
- [327] Y. Tsukada, J. Fang, H. Erdjument-Bromage, M. E. Warren, C. H. Borchers, P. Tempst, Y. Zhang, „Histone demethylation by a family of JmjC domain-containing proteins“, *Nature* **2006**, *439*, 811–816.
- [328] J. R. Whetstine, A. Nottke, F. Lan, M. Huarte, S. Smolikov, Z. Chen, E. Spooner, E. Li, G. Zhang, M. Colaiacovo, Y. Shi, „Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases“, *Cell* **2006**, *125*, 467–481.
- [329] K. Yamane, C. Toumazou, Y. Tsukada, H. Erdjument-Bromage, P. Tempst, J. Wong, Y. Zhang, „JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor“, *Cell* **2006**, *125*, 483–495.
- [330] B. D. Fodor, S. Kubicek, M. Yonezawa, R. J. O'Sullivan, R. Sengupta, L. Perez-Burgos, S. Opravil, K. Mechtler, G. Schotta, T. Jenuwein, „Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells“, *Genes Dev.* **2006**, *20*, 1557–1562.
- [331] P. A. Cloos, J. Christensen, K. Agger, A. Maiolica, J. Rappsilber, T. Antal, K. H. Hansen, K. Helin, „The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3“, *Nature* **2006**, *442*, 307–311.
- [332] R. J. Klose, K. Yamane, Y. Bae, D. Zhang, H. Erdjument-Bromage, P. Tempst, J. Wong, Y. Zhang, „The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36“, *Nature* **2006**, *442*, 312–316.
- [333] S. Shin, R. Janknecht, „Diversity within the JMJD2 histone demethylase family“, *Biochem. Biophys. Res. Commun.* **2007**, *353*, 973–977.
- [334] B. Chang, Y. Chen, Y. Zhao, R. K. Bruick, „JMJD6 is a histone arginine demethylase“, *Science* **2007**, *318*, 444–447.
-

- [335] L. J. Walport, R. J. Hopkinson, R. Chowdhury, R. Schiller, W. Ge, A. Kawamura, C. J. Schofield, „Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases“, *Nat. Commun.* **2016**, *7*, 11974.
- [336] H. S. Choi, B. Y. Choi, Y. Y. Cho, H. Mizuno, B. S. Kang, A. M. Bode, Z. Dong, „Phosphorylation of histone H3 at serine 10 is indispensable for neoplastic cell transformation“, *Cancer Res.* **2005**, *65*, 5818–5827.
- [337] P. N. Lau, P. Cheung, „Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing“, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2801–2806.
- [338] E. Metzger, N. Yin, M. Wissmann, N. Kunowska, K. Fischer, N. Friedrichs, D. Patnaik, J. M. Higgins, N. Potier, K. H. Scheidtmann, R. Buettner, R. Schüle, „Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation“, *Nat. Cell Biol.* **2008**, *10*, 53–60.
- [339] E. Metzger, A. Imhof, D. Patel, P. Kahl, K. Hoffmeyer, N. Friedrichs, J. M. Müller, H. Greschik, J. Kirfel, S. Ji, N. Kunowska, C. Beisenherz-Huss, T. Günther, R. Buettner, R. Schüle, „Phosphorylation of histone H3T6 by PKC β controls demethylation at histone H3K4“, *Nature* **2010**, *464*, 792–796.
- [340] A. Xiao, H. Li, D. Shechter, S. H. Ahn, L. A. Fabrizio, H. Erdjument-Bromage, S. Ishibe-Murakami, B. Wang, P. Tempst, K. Hofmann, D. J. Patel, S. J. Elledge, C. D. Allis, „WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity“, *Nature* **2009**, *457*, 57–62.
- [341] K. Mahajan, B. Fang, J. M. Koomen, N. P. Mahajan, „H2B Tyr37 phosphorylation suppresses expression of replication-dependent core histone genes“, *Nat. Struct. Mol. Biol.* **2012**, *19*, 930–937.
- [342] B. Zhu, Y. Zheng, A. D. Pham, S. S. Mandal, H. Erdjument-Bromage, P. Tempst, D. Reinberg, „Monoubiquitination of human histone H2B: the factors involved and their roles in *HOX* gene regulation“, *Mol. Cell* **2005**, *20*, 601–611.
- [343] H. Wang, L. Zhai, J. Xu, H. Y. Joo, S. Jackson, H. Erdjument-Bromage, P. Tempst, Y. Xiong, Y. Zhang, „Histone H3 and H4 ubiquitylation by the CUL4-DDB-ROC1 ubiquitin ligase facilitates cellular response to DNA damage“, *Mol. Cell* **2006**, *22*, 383–394.
- [344] D. Nathan, K. Ingvarsdottir, D. E. Sterner, G. R. Bylebyl, M. Dokmanovic, J. A. Dorsey, K. A. Whelan, M. Krsmanovic, W. S. Lane, P. B. Meluh, E. S. Johnson, S. L. Berger, „Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications“, *Genes Dev.* **2006**, *20*, 966–976.
- [345] R. Martinez-Zamudio, H. C. Ha, „Histone ADP-ribosylation facilitates gene transcription by directly remodeling nucleosomes“, *Mol. Cell. Biol.* **2012**, *32*, 2490–2502.
- [346] N. Ogata, K. Ueda, O. Hayaishi, „ADP-ribosylation of histone H2B. Identification of glutamic acid residue 2 as the modification site“, *J. Biol. Chem.* **1980**, *255*, 7610–7615.
- [347] N. Ogata, K. Ueda, H. Kagamiyama, O. Hayaishi, „ADP-ribosylation of histone H1. Identification of glutamic acid residues 2, 14, and the COOH-terminal lysine residue as modification sites“, *J. Biol. Chem.* **1980**, *255*, 7616–7620.

- [348] S. Messner, M. Altmeyer, H. Zhao, A. Pozivil, B. Roschitzki, P. Gehrig, D. Rutishauser, D. Huang, A. Caflisch, M. O. Hottiger, „PARP1 ADP-ribosylates lysine residues of the core histone tails“, *Nucleic Acids Res.* **2010**, *38*, 6350–6362.
- [349] Y. Lorch, J. W. LaPointe, R. D. Kornberg, „Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones“, *Cell* **1987**, *49*, 203–210.
- [350] H. Boeger, J. Griesenbeck, J. S. Strattan, R. D. Kornberg, „Nucleosomes unfold completely at a transcriptionally active promoter“, *Mol. Cell* **2003**, *11*, 1587–1598.
- [351] C. K. Lee, Y. Shibata, B. Rao, B. D. Strahl, J. D. Lieb, „Evidence for nucleosome depletion at active regulatory regions genome-wide“, *Nat. Genet.* **2004**, *36*, 900–905.
- [352] M. Kininis, B. S. Chen, A. G. Diehl, G. D. Isaacs, T. Zhang, A. C. Siepel, A. G. Clark, W. L. Kraus, „Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters“, *Mol. Cell. Biol.* **2007**, *27*, 5090–5104.
- [353] L. Tserel, R. Kolde, A. Rebane, K. Kisand, T. Org, H. Peterson, J. Vilo, P. Peterson, „Genome-wide promoter analysis of histone modifications in human monocyte-derived antigen presenting cells“, *BMC Genomics* **2010**, *11*, 642–642.
- [354] A. H. Hassan, P. Prochasson, K. E. Neely, S. C. Galasinski, M. Chandy, M. J. Carrozza, J. L. Workman, „Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes“, *Cell* **2002**, *111*, 369–379.
- [355] X. Bao, A. J. Rubin, K. Qu, J. Zhang, P. G. Giresi, H. Y. Chang, P. A. Khavari, „A novel ATAC-seq approach reveals lineage-specific reinforcement of the open chromatin landscape via cooperation between BAF and p63“, *Genome Biol.* **2015**, *16*, 284.
- [356] L. A. Cirillo, F. R. Lin, I. Cuesta, D. Friedman, M. Jarnik, K. S. Zaret, „Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4“, *Mol. Cell* **2002**, *9*, 279–289.
- [357] X. Guo, K. Tatsuoka, R. Liu, „Histone acetylation and transcriptional regulation in the genome of *Saccharomyces cerevisiae*“, *Bioinformatics* **2005**, *22*, 392–399.
- [358] S. John, P. J. Sabo, R. E. Thurman, M. H. Sung, S. C. Biddie, T. A. Johnson, G. L. Hager, J. A. Stamatoyannopoulos, „Chromatin accessibility pre-determines glucocorticoid receptor binding patterns“, *Nat. Genet.* **2011**, *43*, 264–268.
- [359] R. E. Thurman, E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano, E. Haugen, N. C. Sheffield, A. B. Stergachis, H. Wang, B. Vernot, K. Garg, S. John, R. Sandstrom, D. Bates, L. Boatman, T. K. Canfield, M. Diegel, D. Dunn, A. K. Ebersol, T. Frum, E. Giste, A. K. Johnson, E. M. Johnson, T. Kutuyavin, B. Lajoie, B. K. Lee, K. Lee, D. London, D. Lotakis, S. Neph, F. Neri, E. D. Nguyen, H. Qu, A. P. Reynolds, V. Roach, A. Safi, M. E. Sanchez, A. Sanyal, A. Shafer, J. M. Simon, L. Song, S. Vong, M. Weaver, Y. Yan, Z. Zhang, Z. Zhang, B. Lenhard, M. Tewari, M. O. Dorschner, R. S. Hansen, P. A. Navas, G. Stamatoyannopoulos, V. R. Iyer, J. D. Lieb, S. R. Sunyaev, J. M. Akey, P. J. Sabo, R. Kaul, T. S. Furey, J. Dekker, G. E. Crawford, J. A. Stamatoyannopoulos, „The accessible chromatin landscape of the human genome“, *Nature* **2012**, *489*, 75–82.

- [360] G. C. Yuan, Y. J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, O. J. Rando, „Genome-scale identification of nucleosome positions in *S. cerevisiae*“, *Science* **2005**, *309*, 626–630.
- [361] A. Almer, W. Hörz, „Nuclease hypersensitive regions with adjacent positioned nucleosomes mark the gene boundaries of the PHO5/PHO3 locus in yeast“, *EMBO J.* **1986**, *5*, 2681–2687.
- [362] T. N. Mavrich, I. P. Ioshikhes, B. J. Venters, C. Jiang, L. P. Tomsho, J. Qi, S. C. Schuster, I. Albert, B. F. Pugh, „A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome“, *Genome Res.* **2008**, *18*, 1073–1083.
- [363] H. S. Rhee, B. F. Pugh, „Genome-wide structure and organization of eukaryotic pre-initiation complexes“, *Nature* **2012**, *483*, 295–301.
- [364] H. Kimura, P. R. Cook, „Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B“, *J. Cell. Biol.* **2001**, *153*, 1341–1353.
- [365] M. L. Kireeva, W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, V. M. Studitsky, „Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription“, *Mol. Cell* **2002**, *9*, 541–552.
- [366] O. I. Kulaeva, D. A. Gaykalova, N. A. Pestov, V. V. Golovastov, D. G. Vassilyev, I. Artsimovitch, V. M. Studitsky, „Mechanism of chromatin remodeling and recovery during passage of RNA polymerase II“, *Nat. Struct. Mol. Biol.* **2009**, *16*, 1272–1278.
- [367] D. S. Luse, L. C. Spangler, A. Újvári, „Efficient and rapid nucleosome traversal by RNA polymerase II depends on a combination of transcript elongation factors“, *J. Biol. Chem.* **2011**, *286*, 6040–6048.
- [368] A. Nock, J. M. Ascano, M. J. Barrero, S. Malik, „Mediator-regulated transcription through the +1 nucleosome“, *Mol. Cell* **2012**, *48*, 837–848.
- [369] H. D. Ou, S. Phan, T. J. Deerinck, A. Thor, M. H. Ellisman, C. C. O'Shea, „ChromEMT: visualizing 3D chromatin structure and compaction in interphase and mitotic cells“, *Science* **2017**, *357*, eaag0025.
- [370] Y. C. G. Lee, Y. Ogiyama, N. M. C. Martins, B. J. Beliveau, D. Acevedo, C.-t. Wu, G. Cavalli, G. H. Karpen, „Pericentromeric heterochromatin is hierarchically organized and spatially contacts H3K9me2 islands in euchromatin“, *PLoS Genet.* **2020**, *16*, e1008673.
- [371] J. Kanoh, M. Sadaie, T. Urano, F. Ishikawa, „Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres“, *Curr. Biol.* **2005**, *15*, 1808–1819.
- [372] J. Wang, A. L. Cohen, A. Letian, X. Tadeo, J. J. Moresco, J. Liu, J. R. Yates III, F. Qiao, S. Jia, „The proper connection between shelterin components is required for telomeric heterochromatin assembly“, *Genes Dev.* **2016**, *30*, 827–839.
- [373] J. Ueda, K. Maehara, D. Mashiko, T. Ichinose, T. Yao, M. Hori, Y. Sato, H. Kimura, Y. Ohkawa, K. Yamagata, „Heterochromatin dynamics during the differentiation process revealed by the DNA methylation reporter mouse, *MethylRO*“, *Stem Cell Rep.* **2014**, *2*, 910–924.
- [374] S. Jia, K. Noma, S. I. Grewal, „RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins“, *Science* **2004**, *304*, 1971–1976.

- [375] H. S. Kim, E. S. Choi, J. A. Shin, Y. K. Jang, S. D. Park, „Regulation of Swi6/HP1-dependent heterochromatin assembly by cooperation of components of the mitogen-activated protein kinase pathway and a histone deacetylase Clr6“, *J. Biol. Chem.* **2004**, *279*, 42850–42859.
- [376] E. J. Hong, J. Villén, E. L. Gerace, S. P. Gygi, D. Moazed, „A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation“, *RNA Biol.* **2005**, *2*, 106–111.
- [377] E. H. Bayne, S. A. White, A. Kagansky, D. A. Bijos, L. Sanchez-Pulido, K. L. Hoe, D. U. Kim, H. O. Park, C. P. Ponting, J. Rappsilber, R. C. Allshire, „Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity“, *Cell* **2010**, *140*, 666–677.
- [378] A. Bulut-Karslioglu, V. Perrera, M. Scaranaro, I. A. de la Rosa-Velazquez, S. van de Nobelen, N. Shukeir, J. Popow, B. Gerle, S. Opravil, M. Pagani, S. Meidhof, T. Brabletz, T. Manke, M. Lachner, T. Jenuwein, „A transcription factor-based mechanism for mouse heterochromatin formation“, *Nat. Struct. Mol. Biol.* **2012**, *19*, 1023–1030.
- [379] J. A. Rosenfeld, Z. Wang, D. E. Schones, K. Zhao, R. DeSalle, M. Q. Zhang, „Determination of enriched histone modifications in non-genic portions of the human genome“, *BMC Genomics* **2009**, *10*, 143.
- [380] S. B. Rothbart, K. Krajewski, N. Nady, W. Tempel, S. Xue, A. I. Badeaux, D. Barsyte-Lovejoy, J. Y. Martinez, M. T. Bedford, S. M. Fuchs, C. H. Arrowsmith, B. D. Strahl, „Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation“, *Nat. Struct. Mol. Biol.* **2012**, *19*, 1155–1160.
- [381] W. Fischle, Y. Wang, S. A. Jacobs, Y. Kim, C. D. Allis, S. Khorasanizadeh, „Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains“, *Genes Dev.* **2003**, *17*, 1870–1881.
- [382] E. Bernstein, E. M. Duncan, O. Masui, J. Gil, E. Heard, C. D. Allis, „Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin“, *Mol. Cell. Biol.* **2006**, *26*, 2560–2569.
- [383] M. Lachner, D. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, „Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins“, *Nature* **2001**, *410*, 116–120.
- [384] A. J. Bannister, P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire, T. Kouzarides, „Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain“, *Nature* **2001**, *410*, 120–124.
- [385] B. Al-Sady, H. D. Madhani, G. J. Narlikar, „Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread“, *Mol. Cell* **2013**, *51*, 80–91.
- [386] P. N. C. B. Audergon, S. Catania, A. Kagansky, P. Tong, M. Shukla, A. L. Pidoux, R. C. Allshire, „Restricted epigenetic inheritance of H3K9 methylation“, *Science* **2015**, *348*, 132–135.
- [387] K. Ragunathan, G. Jih, D. Moazed, „Epigenetic inheritance uncoupled from sequence-specific recruitment“, *Science* **2015**, *348*, 1258699.

- [388] M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A.-K. Szigartyo, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, F. Pontén, „Tissue-based map of the human proteome“, *Science* **2015**, *347*, 1260419.
- [389] M. Frommer, L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy, C. L. Paul, „A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands“, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1827–1831.
- [390] M. Yu, G. C. Hon, K. E. Szulwach, C.-X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J.-H. Min, P. Jin, B. Ren, C. He, „Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome“, *Cell* **2012**, *149*, 1368–1380.
- [391] M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik, S. Balasubramanian, „Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution“, *Science* **2012**, *336*, 934–937.
- [392] F. Gao, Y. Xia, J. Wang, H. Luo, Z. Gao, X. Han, J. Zhang, X. Huang, Y. Yao, H. Lu, N. Yi, B. Zhou, Z. Lin, B. Wen, X. Zhang, H. Yang, J. Wang, „Integrated detection of both 5-mC and 5-hmC by high-throughput tag sequencing technology highlights methylation reprogramming of bivalent genes during cellular differentiation“, *Epigenetics* **2013**, *8*, 421–430.
- [393] M. N. Bainbridge, R. L. Warren, M. Hirst, T. Romanuik, T. Zeng, A. Go, A. Delaney, M. Griffith, M. Hickenbotham, V. Magrini, E. R. Mardis, M. D. Sadar, A. S. Siddiqui, M. A. Marra, S. J. M. Jones, „Analysis of the prostate cancer cell line LNCaP transcriptome using a sequencing-by-synthesis approach“, *BMC Genomics* **2006**, *7*, 246.
- [394] U. Nagalakshmi, Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein, M. Snyder, „The transcriptional landscape of the yeast genome defined by RNA sequencing“, *Science* **2008**, *320*, 1344–1349.
- [395] B. T. Wilhelm, S. Marguerat, S. Watt, F. Schubert, V. Wood, I. Goodhead, C. J. Penkett, J. Rogers, J. Bähler, „Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution“, *Nature* **2008**, *453*, 1239-1243.
- [396] N. Cloonan, A. R. Forrest, G. Kolle, B. B. Gardiner, G. J. Faulkner, M. K. Brown, D. F. Taylor, A. L. Steptoe, S. Wani, G. Bethel, A. J. Robertson, A. C. Perkins, S. J. Bruce, C. C. Lee, S. S. Ranade, H. E. Peckham, J. M. Manning, K. J. McKernan, S. M. Grimmond, „Stem cell transcriptome profiling via massive-scale mRNA sequencing“, *Nat. Methods* **2008**, *5*, 613–619.
- [397] A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, B. Wold, „Mapping and quantifying mammalian transcriptomes by RNA-Seq“, *Nat. Methods* **2008**, *5*, 621–628.
- [398] A. Bachmair, D. Finley, A. Varshavsky, „In vivo half-life of a protein is a function of its amino-terminal residue“, *Science* **1986**, *234*, 179–186.
- [399] S. Rogers, M. Girolami, W. Kolch, K. M. Waters, T. Liu, B. Thrall, H. S. Wiley, „Investigating the correspondence between transcriptomic and proteomic expression profiles using coupled cluster models“, *Bioinformatics* **2008**, *24*, 2894–2900.

-
- [400] P. C. van Weeren, K. M. T. de Bruyn, A. M. M. de Vries-Smits, J. van Lint, B. M. T. Burgering, „Essential role for protein kinase B (PKB) in insulin-induced glycogen synthase kinase 3 inactivation: characterization of dominant-negative mutant of PKB“, *J. Biol. Chem.* **1998**, *273*, 13150–13156.
- [401] M. Ashcroft, M. H. G. Kubbutat, K. H. Vousden, „Regulation of p53 function and stability by phosphorylation“, *Mol. Cell. Biol.* **1999**, *19*, 1751–1758.
- [402] O. A. Malakhova, M. Yan, M. P. Malakhov, Y. Yuan, K. J. Ritchie, K. I. Kim, L. F. Peterson, K. Shuai, D.-E. Zhang, „Protein ISGylation modulates the JAK-STAT signaling pathway“, *Genes Dev.* **2003**, *17*, 455–460.
- [403] J. Luo, M. Li, Y. Tang, M. Laszkowska, R. G. Roeder, W. Gu, „Acetylation of p53 augments its site-specific DNA binding both *in vitro* and *in vivo*“, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2259–2264.
- [404] Z. Zhang, R. C. Bast Jr., Y. Yu, J. Li, L. J. Sokoll, A. J. Rai, J. M. Rosenzweig, B. Cameron, Y. Y. Wang, X. Y. Meng, A. Berchuck, C. Van Haaften-Day, N. F. Hacker, H. W. de Bruijn, A. G. van der Zee, I. J. Jacobs, E. T. Fung, D. W. Chan, „Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer“, *Cancer Res.* **2004**, *64*, 5882–5890.
- [405] M. A. Meuwis, M. Fillet, P. Geurts, D. de Seny, L. Lutteri, J. P. Chapelle, V. Bours, L. Wehenkel, J. Belaiche, M. Malaise, E. Louis, M. P. Merville, „Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling“, *Biochem. Pharmacol.* **2007**, *73*, 1422–1433.
- [406] H. F. Hu, W. W. Xu, Y. Wang, C. C. Zheng, W. X. Zhang, B. Li, Q. Y. He, „Comparative proteomics analysis identifies Cdc42-Cdc42BPA signaling as prognostic biomarker and therapeutic target for colon cancer invasion“, *J. Proteome Res.* **2018**, *17*, 265–275.
- [407] K. Bassiri, S. Ferluga, V. Sharma, N. Syed, C. L. Adams, E. Lasonder, C. O. Hanemann, „Global proteome and phospho-proteome analysis of Merlin-deficient meningioma and schwannoma identifies PDLIM2 as a novel therapeutic target“, *EBioMedicine* **2017**, *16*, 76–86.
- [408] Y. Zhou, W. Sun, N. Chen, C. Xu, X. Wang, K. Dong, B. Zhang, J. Zhang, N. Hao, A. Sun, H. Wei, F. He, Y. Jiang, „Discovery of NKCC1 as a potential therapeutic target to inhibit hepatocellular carcinoma cell growth and metastasis“, *Oncotarget* **2017**, *8*, 66328–66342.
- [409] M. Peled, A. S. Tocheva, S. Sandigursky, S. Nayak, E. A. Philips, K. E. Nichols, M. Strazza, I. Azoulay-Alfaguter, M. Askenazi, B. G. Neel, A. J. Pelzek, B. Ueberheide, A. Mor, „Affinity purification mass spectrometry analysis of PD-1 uncovers SAP as a new checkpoint inhibitor“, *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E468–E477.
- [410] J. K. Mulligan, K. Patel, T. Williamson, N. Reaves, W. Carroll, S. E. Stephenson, P. Gao, R. R. Drake, B. A. Neely, S. Tomlinson, R. J. Schlosser, C. Atkinson, „C3a receptor antagonism as a novel therapeutic target for chronic rhinosinusitis“, *Mucosal Immunol.* **2018**, *11*, 1375–1385.
- [411] P. H. O'Farrell, „High resolution two-dimensional electrophoresis of proteins“, *J. Biol. Chem.* **1975**, *250*, 4007–4021.
- [412] J. Klose, „Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues“, *Hum. Genet.* **1975**, *26*, 231–243.
-

- [413] M. R. Wilkins, E. Gasteiger, J. C. Sanchez, A. Bairoch, D. F. Hochstrasser, „Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicity and copy number“, *Electrophoresis* **1998**, *19*, 1501–1505.
- [414] T. Rabilloud, C. Adessi, A. Giraudel, J. Lunardi, „Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients“, *Electrophoresis* **1997**, *18*, 307–316.
- [415] R. P. Zahedi, C. Meisinger, A. Sickmann, „Two-dimensional benzyldimethyl-n-hexadecylammonium chloride/SDS-PAGE for membrane proteomics“, *Proteomics* **2005**, *5*, 3581–3588.
- [416] S. Falvo, M. Di Carli, A. Desiderio, E. Benvenuto, A. Moglia, T. America, S. Lanteri, A. Acquadro, „2-D DIGE analysis of UV-C radiation-responsive proteins in globe artichoke leaves“, *Proteomics* **2012**, *12*, 448–460.
- [417] E. P. Wright, M. A. Partridge, M. P. Padula, V. J. Gauci, C. S. Malladi, J. R. Coorsen, „Top-down proteomics: enhancing 2D gel electrophoresis from tissue processing to high-sensitivity protein detection“, *Proteomics* **2014**, *14*, 872–889.
- [418] S. Y. Wong, O. H. Hashim, N. Hayashi, „Development of high-performance two-dimensional gel electrophoresis for human hair shaft proteome“, *PLoS One* **2019**, *14*, e0213947.
- [419] J. J. Thomson, „Bakerian lecture: rays of positive electricity“, *Proc. R. Soc. A* **1913**, *89*, 1–20.
- [420] A. J. Dempster, „A new method of positive ray analysis“, *Phys. Rev.* **1918**, *11*, 316–325.
- [421] F. W. Aston, „LXXIV. A positive ray spectrograph“, *Philos. Mag.* **1919**, *38*, 707–714.
- [422] F. Sanger, E. O. P. Thompson, „The amino-acid sequence in the glycol chain of insulin. I. The identification of lower peptides from partial hydrolysates“, *Biochem. J.* **1953**, *53*, 353–366.
- [423] P. Edman, „Mechanism of the phenyl isothiocyanate degradation of peptides“, *Nature* **1956**, *177*, 667–668.
- [424] K. Biemann, G. Gapp, J. Seibl, „Application of mass spectrometry to structure problems. I. Amino acid sequence in peptides“, *J. Am. Chem. Soc.* **1959**, *81*, 2274–2275.
- [425] G. Hudson, K. Biemann, „Mass spectrometric sequencing of proteins. The structure of subunit I of monellin“, *Biochem. Biophys. Res. Commun.* **1976**, *71*, 212–220.
- [426] M. Barber, R. S. Bordoli, R. D. Sedgwick, A. N. Tyler, „Fast atom bombardment of solids (F.A.B.): a new ion source for mass spectrometry“, *J. Chem. Soc., Chem. Commun.* **1981**, 325–327.
- [427] P. Hoben, N. Royal, A. Cheung, F. Yamao, K. Biemann, D. Söll, „*Escherichia coli* glutaminyl-tRNA synthetase. II. Characterization of the *glnS* gene product“, *J. Biol. Chem.* **1982**, *257*, 11644–11650.
- [428] T. A. Webster, B. W. Gibson, T. Keng, K. Biemann, P. Schimmel, „Primary structures of both subunits of *Escherichia coli* glycol-tRNA synthetase“, *J. Biol. Chem.* **1983**, *258*, 10637–10641.

- [429] R. Freedman, B. Gibson, D. Donovan, K. Biemann, S. Eisenbeis, J. Parker, P. Schimmel, „Primary structure of histidine-tRNA synthetase and characterization of *hisS* transcripts“, *J. Biol. Chem.* **1985**, *260*, 10063–10068.
- [430] F. Fasiolo, B. W. Gibson, P. Walter, B. Chatton, K. Biemann, Y. Boulanger, „Cytoplasmic methionyl-tRNA synthetase from Bakers' yeast. A monomer with a post-translationally modified N terminus“, *J. Biol. Chem.* **1985**, *260*, 15571–15576.
- [431] R. Breton, H. Sanfaçon, I. Papayannopoulos, K. Biemann, J. Lapointe, „Glutamyl-tRNA synthetase of *Escherichia coli*. Isolation and primary structure of the *gltX* gene and homology with other aminoacyl-tRNA synthetases“, *J. Biol. Chem.* **1986**, *261*, 10610–10617.
- [432] M. Yamashita, J. B. Fenn, „Electrospray ion source. Another variation on the free-jet theme“, *J. Phys. Chem.* **1984**, *88*, 4451–4459.
- [433] M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson, M. B. Alice, „Molecular beams of macroions“, *J. Chem. Phys.* **1968**, *49*, 2240–2249.
- [434] M. Karas, D. Bachmann, F. Hillenkamp, „Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules“, *Anal. Chem.* **1985**, *57*, 2935–2939.
- [435] M. Karas, F. Hillenkamp, „Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons“, *Anal. Chem.* **1988**, *60*, 2299–2301.
- [436] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, „Electrospray ionization for mass spectrometry of large biomolecules“, *Science* **1989**, *246*, 64–71.
- [437] V. V. Laiko, M. A. Baldwin, A. L. Burlingame, „Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry“, *Anal. Chem.* **2000**, *72*, 652–657.
- [438] V. V. Laiko, S. C. Moyer, R. J. Cotter, „Atmospheric pressure MALDI/ion trap mass spectrometry“, *Anal. Chem.* **2000**, *72*, 5239–5243.
- [439] M. Karas, M. Glückmann, J. Schäfer, „Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors“, *J. Mass Spectrom.* **2000**, *35*, 1–12.
- [440] L. V. Zhigilei, P. B. S. Kodali, B. J. Garrison, „On the threshold behavior in laser ablation of organic solids“, *Chem. Phys. Lett.* **1997**, *276*, 269–273.
- [441] G. I. Taylor, „Disintegration of water drops in an electric field“, *Proc. R. Soc. A* **1964**, *280*, 383–397.
- [442] L. Rayleigh, „XX. On the equilibrium of liquid conducting masses charged with electricity“, *Philos. Mag.* **1882**, *14*, 184–186.
- [443] S. Nguyen, J. B. Fenn, „Gas-phase ions of solute species from charged droplets of solutions“, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1111–1117.
- [444] J. Fernandez de la Mora, „Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism“, *Anal. Chim. Acta* **2000**, *406*, 93–104.

- [445] D. R. Mueller, H. Voshol, A. Waldt, B. Wiedmann, J. van Oostrum, „LC-MALDI MS and MS/MS — an efficient tool in proteome analysis“ in *Subcellular Proteomics: From Cell Deconstruction to System Reconstruction* (Eds.: E. Bertrand, M. Faupel), Springer Netherlands, Dordrecht, **2007**, 355–380.
- [446] M. M. Wolff, W. E. Stephens, „A pulsed mass spectrometer with time dispersion“, *Rev. Sci. Instrum.* **1953**, *24*, 616–617.
- [447] F. Dubois, R. Knochenmuss, R. Zenobi, A. Brunelle, C. Deprun, Y. Le Beyec, „A comparison between ion-to-photon and microchannel plate detectors“, *Rapid Commun. Mass Spectrom.* **1999**, *13*, 786–791.
- [448] W. Paul, H. Steinwedel, „Ein neues Massenspektrometer ohne Magnetfeld“, *Z. Naturforsch. A* **1953**, *8*, 448–450.
- [449] G. G. Dolnikowski, M. J. Kristo, C. G. Enke, J. T. Watson, „Ion-trapping technique for ion/molecule reaction studies in the center quadrupole of a triple quadrupole mass spectrometer“, *Int. J. Mass Spectrom. Ion Proc.* **1988**, *82*, 1–15.
- [450] O. Kornienko, P. T. A. Reilly, W. B. Whitten, J. M. Ramsey, „Micro ion trap mass spectrometry“, *Rapid Commun. Mass Spectrom.* **1999**, *13*, 50–53.
- [451] M. B. Comisarow, A. G. Marshall, „Fourier transform ion cyclotron resonance spectroscopy“, *Chem. Phys. Lett.* **1974**, *25*, 282–283.
- [452] A. Makarov, „Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis“, *Anal. Chem.* **2000**, *72*, 1156–1162.
- [453] F. He, C. L. Hendrickson, A. G. Marshall, „Baseline mass resolution of peptide isobars: a record for molecular mass resolution“, *Anal. Chem.* **2001**, *73*, 647–650.
- [454] A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, S. Horning, „Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer“, *Anal. Chem.* **2006**, *78*, 2113–2120.
- [455] J. M. Wells, S. A. McLuckey, „Collision-induced dissociation (CID) of peptides and proteins“ in *Methods in Enzymology*, Vol. 402 (Ed.: A. L. Burlingame), Academic Press, Cambridge, **2005**, 148–185.
- [456] J. E. P. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, D. F. Hunt, „Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry“, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528–9533.
- [457] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, M. Mann, „Higher-energy C-trap dissociation for peptide modification analysis“, *Nat. Methods* **2007**, *4*, 709–712.
- [458] A. Chi, C. Huttenhower, L. Y. Geer, J. J. Coon, J. E. P. Syka, D. L. Bai, J. Shabanowitz, D. J. Burke, O. G. Troyanskaya, D. F. Hunt, „Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry“, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2193–2198.

- [459] N. L. Kelleher, H. Y. Lin, G. A. Valaskovic, D. J. Aaserud, E. K. Fridriksson, F. W. McLafferty, „Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry“, *J. Am. Chem. Soc.* **1999**, *121*, 806–812.
- [460] T. G. Schaaff, B. J. Cargile, J. L. Stephenson Jr., S. A. McLuckey, „Ion trap collisional activation of the $(M + 2H)^{2+}$ - $(M + 17H)^{17+}$ ions of human hemoglobin β -chain“, *Anal. Chem.* **2000**, *72*, 899–907.
- [461] J. M. Wells, J. L. Stephenson, S. A. McLuckey, „Charge dependence of protonated insulin decompositions“, *Int. J. Mass Spectrom.* **2000**, *203*, A1–A9.
- [462] K. A. Newton, P. A. Chrisman, G. E. Reid, J. M. Wells, S. A. McLuckey, „Gaseous apomyoglobin ion dissociation in a quadrupole ion trap: $[M + 2H]^{2+}$ - $[M + 21H]^{21+}$ “, *Int. J. Mass Spectrom.* **2001**, *212*, 359–376.
- [463] J. S. Cobb, M. L. Easterling, J. N. Agar, „Structural characterization of intact proteins is enhanced by prevalent fragmentation pathways rarely observed for peptides“, *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 949–959.
- [464] E. Mørtz, P. B. Connor, P. Roepstorff, N. L. Kelleher, T. D. Wood, F. W. McLafferty, M. Mann, „Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases“, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8264–8267.
- [465] D. M. Horn, R. A. Zubarev, F. W. McLafferty, „Automated *de novo* sequencing of proteins by tandem high-resolution mass spectrometry“, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10313–10317.
- [466] F. Meng, B. J. Cargile, L. M. Miller, A. J. Forbes, J. R. Johnson, N. L. Kelleher, „Informatics and multiplexing of intact protein identification in bacteria and the archaea“, *Nat. Biotechnol.* **2001**, *19*, 952–957.
- [467] K. R. Lorenzatto, K. Kim, I. Ntai, G. P. Paludo, J. Camargo de Lima, P. M. Thomas, N. L. Kelleher, H. B. Ferreira, „Top down proteomics reveals mature proteoforms expressed in subcellular fractions of the *Echinococcus granulosus* preadult stage“, *J. Proteome Res.* **2015**, *14*, 4805–4814.
- [468] K. R. Durbin, L. Fornelli, R. T. Fellers, P. F. Doubleday, M. Narita, N. L. Kelleher, „Quantitation and identification of thousands of human proteoforms below 30 kDa“, *J. Proteome Res.* **2016**, *15*, 976–982.
- [469] S. K. Sze, Y. Ge, H. Oh, F. W. McLafferty, „Top-down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue“, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1774–1779.
- [470] J. J. Pesavento, Y. B. Kim, G. K. Taylor, N. L. Kelleher, „Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry“, *J. Am. Chem. Soc.* **2004**, *126*, 3386–3387.
- [471] M. Schirm, I. C. Schoenhofen, S. M. Logan, K. C. Waldron, P. Thibault, „Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins“, *Anal. Chem.* **2005**, *77*, 7774–7782.

- [472] M. Mann, P. Højrup, P. Roepstorff, „Use of mass spectrometric molecular weight information to identify proteins in sequence databases“, *Biol. Mass Spectrom.* **1993**, *22*, 338–345.
- [473] J. R. Yates III, S. Speicher, P. R. Griffin, T. Hunkapiller, „Peptide mass maps: a highly informative approach to protein identification“, *Anal. Biochem.* **1993**, *214*, 397–408.
- [474] D. L. Tabb, L. L. Smith, L. A. Breci, V. H. Wysocki, D. Lin, J. R. Yates III, „Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides“, *Anal. Chem.* **2003**, *75*, 1155–1163.
- [475] A. Michalski, N. Neuhauser, J. Cox, M. Mann, „A systematic investigation into the nature of tryptic HCD spectra“, *J. Proteome Res.* **2012**, *11*, 5479–5491.
- [476] P. Roepstorff, J. Fohlman, „Proposal for a common nomenclature for sequence ions in mass spectra of peptides“, *Biomed. Mass Spectrom.* **1984**, *11*, 601.
- [477] R. S. Johnson, S. A. Martin, K. Biemann, J. T. Stults, J. T. Watson, „Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine“, *Anal. Chem.* **1987**, *59*, 2621–2625.
- [478] J. K. Eng, A. L. McCormack, J. R. Yates III, „An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database“, *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989.
- [479] D. N. Perkins, D. J. Pappin, D. M. Creasy, J. S. Cottrell, „Probability-based protein identification by searching sequence databases using mass spectrometry data“, *Electrophoresis* **1999**, *20*, 3551–3567.
- [480] J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, M. Mann, „Andromeda: a peptide search engine integrated into the MaxQuant environment“, *J. Proteome Res.* **2011**, *10*, 1794–1805.
- [481] J. E. Elias, S. P. Gygi, „Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry“, *Nat. Methods* **2007**, *4*, 207–214.
- [482] B. Schilling, R. H. Row, B. W. Gibson, X. Guo, M. M. Young, „MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides“, *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 834–850.
- [483] S. Kalkhof, A. Sinz, „Chances and pitfalls of chemical cross-linking with amine-reactive *N*-hydroxysuccinimide esters“, *Anal. Bioanal. Chem.* **2008**, *392*, 305–312.
- [484] S. Mädler, C. Bich, D. Touboul, R. Zenobi, „Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities“, *J. Mass Spectrom.* **2009**, *44*, 694–706.
- [485] J. Prahlad, D. N. Hauser, N. M. Milkovic, M. R. Cookson, M. A. Wilson, „Use of cysteine-reactive cross-linkers to probe conformational flexibility of human DJ-1 demonstrates that Glu18 mutations are dimers“, *J. Neurochem.* **2014**, *130*, 839–853.
- [486] A. Leitner, L. A. Joachimiak, P. Unverdorben, T. Walzthoeni, J. Frydman, F. Förster, R. Aebersold, „Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes“, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 9455–9460.

- [487] C. B. Gutierrez, C. Yu, E. J. Novitsky, A. S. Huszagh, S. D. Rychnovsky, L. Huang, „Developing an acidic residue reactive and sulfoxide-containing MS-cleavable homobifunctional cross-linker for probing protein–protein interactions“, *Anal. Chem.* **2016**, *88*, 8315–8322.
- [488] A. F. Gomes, F. C. Gozzo, „Chemical cross-linking with a diazirine photoactivatable cross-linker investigated by MALDI- and ESI-MS/MS“, *J. Mass Spectrom.* **2010**, *45*, 892–899.
- [489] N. I. Brodie, K. A. Makepeace, E. V. Petrotchenko, C. H. Borchers, „Isotopically-coded short-range hetero-bifunctional photo-reactive crosslinkers for studying protein structure“, *J. Proteomics* **2015**, *118*, 12–20.
- [490] A. Belsom, G. Mudd, S. Giese, M. Auer, J. Rappsilber, „Complementary benzophenone cross-linking/mass spectrometry photochemistry“, *Anal. Chem.* **2017**, *89*, 5319–5324.
- [491] N. S. Green, E. Reisler, K. N. Houk, „Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers“, *Protein Sci.* **2001**, *10*, 1293–1304.
- [492] Z. Orbán-Németh, R. Beveridge, D. M. Hollenstein, E. Rampler, T. Stranzl, O. Hudecz, J. Doblmann, P. Schlögelhofer, K. Mechtler, „Structural prediction of protein models using distance restraints derived from cross-linking mass spectrometry data“, *Nat. Protoc.* **2018**, *13*, 478–494.
- [493] A. J. Lomant, G. Fairbanks, „Chemical probes of extended biological structures: synthesis and properties of the cleavable protein cross-linking reagent [³⁵S]dithiobis(succinimidyl propionate)“, *J. Mol. Biol.* **1976**, *104*, 243–261.
- [494] M. Q. Müller, F. Dreiocker, C. H. Ihling, M. Schäfer, A. Sinz, „Cleavable cross-linker for protein structure analysis: reliable identification of cross-linking products by tandem MS“, *Anal. Chem.* **2010**, *82*, 6958–6968.
- [495] A. Kao, C. L. Chiu, D. Vellucci, Y. Yang, V. R. Patel, S. Guan, A. Randall, P. Baldi, S. D. Rychnovsky, L. Huang, „Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes“, *Mol. Cell. Proteomics* **2011**, *10*, M110.002212.
- [496] M. Götze, J. Pettelkau, R. Fritzsche, C. H. Ihling, M. Schäfer, A. Sinz, „Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis“, *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 83–97.
- [497] F. Liu, D. T. Rijkers, H. Post, A. J. Heck, „Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry“, *Nat. Methods* **2015**, *12*, 1179–1184.
- [498] F. Liu, P. Lössl, R. Scheltema, R. Viner, A. J. R. Heck, „Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification“, *Nat. Commun.* **2017**, *8*, 15473.
- [499] A. Leitner, R. Reischl, T. Walzthoeni, F. Herzog, S. Bohn, F. Förster, R. Aebersold, „Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography“, *Mol. Cell. Proteomics* **2012**, *11*, M111.014126.

- [500] R. Fritzsche, C. H. Ihling, M. Götze, A. Sinz, „Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis“, *Rapid Commun. Mass Spectrom.* **2012**, *26*, 653–658.
- [501] L. Chaiet, F. J. Wolf, „The properties of streptavidin, a biotin-binding protein produced by *Streptomyces*“, *Arch. Biochem. Biophys.* **1964**, *106*, 1–5.
- [502] H. Zhang, X. Tang, G. R. Munske, N. Tolic, G. A. Anderson, J. E. Bruce, „Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry“, *Mol. Cell. Proteomics* **2009**, *8*, 409–420.
- [503] S. Kang, L. Mou, J. Lanman, S. Velu, W. J. Brouillette, P. E. Prevelige Jr., „Synthesis of biotin-tagged chemical cross-linkers and their applications for mass spectrometry“, *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1719–1726.
- [504] D. Tan, Q. Li, M.-J. Zhang, C. Liu, C. Ma, P. Zhang, Y.-H. Ding, S.-B. Fan, L. Tao, B. Yang, X. Li, S. Ma, J. Liu, B. Feng, X. Liu, H.-W. Wang, S.-M. He, N. Gao, K. Ye, M.-Q. Dong, X. Lei, „Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states“, *eLife* **2016**, *5*, e12509.
- [505] C. W. Tornøe, C. Christensen, M. Meldal, „Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides“, *J. Org. Chem.* **2002**, *67*, 3057–3064.
- [506] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, „A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes“, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599; „A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes“, *Angew. Chem.* **2002**, *114*, 2708–2711.
- [507] N. J. Agard, J. A. Prescher, C. R. Bertozzi, „A strain-promoted [3 + 2] azide–alkyne cycloaddition for covalent modification of biomolecules in living systems“, *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
- [508] M. A. Nessen, G. Kramer, J. Back, J. M. Baskin, L. E. Smeenk, L. J. de Koning, J. H. van Maarseveen, L. de Jong, C. R. Bertozzi, H. Hiemstra, C. G. de Koster, „Selective enrichment of azide-containing peptides from complex mixtures“, *J. Proteome Res.* **2009**, *8*, 3702–3711.
- [509] S. M. Chowdhury, X. Du, N. Tolić, S. Wu, R. J. Moore, M. U. Mayer, R. D. Smith, J. N. Adkins, „Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry“, *Anal. Chem.* **2009**, *81*, 5524–5532.
- [510] A. M. Burke, W. Kandur, E. J. Novitsky, R. M. Kaake, C. Yu, A. Kao, D. Vellucci, L. Huang, S. D. Rychnovsky, „Synthesis of two new enrichable and MS-cleavable cross-linkers to define protein-protein interactions by mass spectrometry“, *Org. Biomol. Chem.* **2015**, *13*, 5030–5037.
- [511] R. F. Borch, M. D. Bernstein, H. D. Durst, „Cyanohydridoborate anion as a selective reducing agent“, *J. Am. Chem. Soc.* **1971**, *93*, 2897–2904.

- [512] R. E. Hardisty, F. Kawasaki, A. B. Sahakyan, S. Balasubramanian, „Selective chemical labeling of natural T modifications in DNA“, *J. Am. Chem. Soc.* **2015**, *137*, 9270–9272.
- [513] M. Su, A. Kirchner, S. Stazzoni, M. Müller, M. Wagner, A. Schröder, T. Carell, „5-Formylcytosine could be a semipermanent base in specific genome sites“, *Angew. Chem. Int. Ed.* **2016**, *55*, 11797–11800; „5-Formylcytosin ist vermutlich eine semipermanente Base an definierten Genompositionen“, *Angew. Chem.* **2016**, *128*, 11974–11978.
- [514] S. Bjelland, L. Eide, R. W. Time, R. Stote, I. Eftedal, G. Volden, E. Seeberg, „Oxidation of thymine to 5-formyluracil in DNA: mechanisms of formation, structural implications, and base excision by human cell free extracts“, *Biochemistry* **1995**, *34*, 14758–14764.
- [515] D. K. Rogstad, J. Heo, N. Vaidehi, W. A. Goddard, A. Burdzy, L. C. Sowers, „5-Formyluracil-induced perturbations of DNA function“, *Biochemistry* **2004**, *43*, 5688–5697.
- [516] F. A. Carey, R. J. Sundberg, *Advanced Organic Chemistry: Part A: Structure and Mechanisms*, 5th ed., Springer US, New York City, **2007**, 648–651.
- [517] K. B. Wiberg, R. Glaser, „Resonance interactions in acyclic systems. 4. Stereochemistry, energetics, and electron distributions in 3-center-four- π -electron systems A:BC“, *J. Am. Chem. Soc.* **1992**, *114*, 841–850.
- [518] F. M. Menger, „On the source of intramolecular and enzymatic reactivity“, *Accounts Chem. Res.* **1985**, *18*, 128–134.
- [519] D. Vasudevan, E. Y. Chua, C. A. Davey, „Crystal structures of nucleosome core particles containing the '601' strong positioning sequence“, *J. Mol. Biol.* **2010**, *403*, 1–10.
- [520] K. Luger, T. J. Rechsteiner, T. J. Richmond, „Preparation of nucleosome core particle from recombinant histones“ in *Methods in Enzymology*, Vol. 304 (Eds.: P. M. Wassarman, A. P. Wolffe), Academic Press, Cambridge, **1999**, 3–19.
- [521] C. Zhou, J. T. Szczepanski, M. M. Greenberg, „Histone modification via rapid cleavage of C4'-oxidized abasic sites in nucleosome core particles“, *J. Am. Chem. Soc.* **2013**, *135*, 5274–5277.
- [522] J. Winger, G. D. Bowman, „The sequence of nucleosomal DNA modulates sliding by the Chd1 chromatin remodeler“, *J. Mol. Biol.* **2017**, *429*, 808–822.
- [523] B. Steigenberger, S. Schiesser, B. Hackner, C. Brandmayr, S. K. Laube, J. Steinbacher, T. Pfaffeneder, T. Carell, „Synthesis of 5-hydroxymethyl-, 5-formyl-, and 5-carboxycytidine-triphosphates and their incorporation into oligonucleotides by polymerase chain reaction“, *Org. Lett.* **2013**, *15*, 366–369.
- [524] F. Kawasaki, P. Murat, Z. Li, T. Santner, S. Balasubramanian, „Synthesis and biophysical analysis of modified thymine-containing DNA oligonucleotides“, *Chem. Commun.* **2017**, *53*, 1389–1392.
- [525] J. C. Hansen, K. E. van Holde, D. Lohr, „The mechanism of nucleosome assembly onto oligomers of the sea urchin 5 S DNA positioning sequence“, *J. Biol. Chem.* **1991**, *266*, 4276–4282.

- [526] K. Tatchell, K. E. Van Holde, „Reconstitution of chromatin core particles“, *Biochemistry* **1977**, *16*, 5295–5303.
- [527] M. G. Fried, G. Liu, „Molecular sequestration stabilizes CAP-DNA complexes during polyacrylamide gel electrophoresis“, *Nucleic Acids Res.* **1994**, *22*, 5054–5059.
- [528] New England Biolabs, *FAQ: why are there extra bands visible on polyacrylamide gels?*, <https://international.neb.com/faqs/0001/01/01/why-are-there-extra-bands-visible-on-polyacrylamide-gels> (abgerufen am 13. Januar 2020).
- [529] A. I. Dragan, R. Pavlovic, J. B. McGivney, J. R. Casas-Finet, E. S. Bishop, R. J. Strouse, M. A. Schenerman, C. D. Geddes, „SYBR Green I: fluorescence properties and interaction with DNA“, *J. Fluoresc.* **2012**, *22*, 1189–1199.
- [530] R. J. Ouellette, J. D. Rawn, „19 - Aldehydes and ketones: nucleophilic addition reactions“ in *Organic Chemistry* (Eds.: R. J. Ouellette, J. D. Rawn), Elsevier, Boston, **2014**, 629–657.
- [531] S. Le Maux, A. B. Nongonierma, R. J. FitzGerald, „Improved short peptide identification using HILIC–MS/MS: retention time prediction model based on the impact of amino acid position in the peptide sequence“, *Food Chem.* **2015**, *173*, 847–854.
- [532] V. R. Koskinen, P. A. Emery, D. M. Creasy, J. S. Cottrell, „Hierarchical clustering of shotgun proteomics data“, *Mol. Cell. Proteomics* **2011**, *10*, M110.003822.
- [533] H. Weintraub, F. V. Lente, „Dissection of chromosome structure with trypsin and nucleases“, *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4249–4253.
- [534] R. L. Rill, D. K. Oosterhof, „The accessibilities of histones in nucleosome cores to an arginine-specific protease“, *J. Biol. Chem.* **1982**, *257*, 14875–14880.
- [535] N. L. Rosenberg, R. M. Smith, R. L. Rill, „The action of chymotrypsin on nucleosome cores. Histone products and conformational effects of limited digestion“, *J. Biol. Chem.* **1986**, *261*, 12375–12383.
- [536] L. Böhm, G. Briand, P. Sautière, C. Crane-Robinson, „Proteolytic digestion studies of chromatin core-histone structure. Identification of the limit peptides of histones H3 and H4“, *Eur. J. Biochem.* **1981**, *119*, 67–74.
- [537] M. Nestle, W. K. Roberts, „An extracellular nuclease from *Serratia marcescens*. II. Specificity of the enzyme“, *J. Biol. Chem.* **1969**, *244*, 5219–5225.
- [538] B. A. Garcia, S. Mollah, B. M. Ueberheide, S. A. Busby, T. L. Muratore, J. Shabanowitz, D. F. Hunt, „Chemical derivatization of histones for facilitated analysis by mass spectrometry“, *Nat. Protoc.* **2007**, *2*, 933–938.
- [539] D. Kudlinzki, V. L. Linhard, K. Saxena, S. Sreeramulu, S. Gande, U. Schieborr, M. Dreyer, H. Schwalbe, „High-resolution crystal structure of cAMP-dependent protein kinase from *Cricetulus griseus*“, *Acta Crystallogr. F* **2015**, *71*, 1088–1093.
- [540] A. C. Bastidas, M. S. Deal, J. M. Steichen, M. M. Keshwani, Y. Guo, S. S. Taylor, „Role of N-terminal myristylation in the structure and regulation of cAMP-dependent protein kinase“, *J. Mol. Biol.* **2012**, *422*, 215–229.

- [541] A. Tholey, R. Pipkorn, D. Bossemeyer, V. Kinzel, J. Reed, „Influence of myristoylation, phosphorylation, and deamidation on the structural behavior of the N-Terminus of the catalytic subunit of cAMP-dependent protein kinase“, *Biochemistry* **2001**, *40*, 225–231.
- [542] M. Sastri, D. M. Barraclough, P. T. Carmichael, S. S. Taylor, „A-kinase-interacting protein localizes protein kinase A in the nucleus“, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 349–354.
- [543] E. C. Gaffarogullari, L. R. Masterson, E. E. Metcalfe, N. J. Traaseth, E. Balatri, M. M. Musa, D. Mullen, M. D. Distefano, G. Veglia, „A myristoyl/phosphoserine switch controls cAMP-dependent protein kinase association to membranes“, *J. Mol. Biol.* **2011**, *411*, 823–836.
- [544] F. Li, Y. Zhang, J. Bai, M. M. Greenberg, Z. Xi, C. Zhou, „5-Formylcytosine yields DNA–protein cross-links in nucleosome core particles“, *J. Am. Chem. Soc.* **2017**, *139*, 10617–10620.
- [545] S. Ji, H. Shao, Q. Han, C. L. Seiler, N. Y. Tretyakova, „Reversible DNA–protein cross-linking at epigenetic DNA marks“, *Angew. Chem. Int. Ed.* **2017**, *56*, 14130–14134; „Reversible DNA–protein cross-linking at epigenetic DNA marks“, *Angew. Chem.* **2017**, *129*, 14318–14322.
- [546] J. E. Elias, W. Haas, B. K. Faherty, S. P. Gygi, „Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations“, *Nat. Methods* **2005**, *2*, 667–675.
- [547] A. Keller, A. I. Nesvizhskii, E. Kolker, R. Aebersold, „Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search“, *Anal. Chem.* **2002**, *74*, 5383–5392.
- [548] J. Peng, J. E. Elias, C. C. Thoreen, L. J. Licklider, S. P. Gygi, „Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC–MS/MS) for large-scale protein analysis: the yeast proteome“, *J. Proteome Res.* **2003**, *2*, 43–50.
- [549] E. A. Raiber, G. Portella, S. Martinez Cuesta, R. Hardisty, P. Murat, Z. Li, M. Iurlaro, W. Dean, J. Spindel, D. Beraldi, Z. Liu, M. A. Dawson, W. Reik, S. Balasubramanian, „5-Formylcytosine organizes nucleosomes and forms Schiff base interactions with histones in mouse embryonic stem cells“, *Nat. Chem.* **2018**, *10*, 1258–1266.
- [550] S. Ji, I. Fu, S. Naldiga, H. Shao, A. K. Basu, S. Broyde, N. Y. Tretyakova, „5-Formylcytosine mediated DNA–protein cross-links block DNA replication and induce mutations in human cells“, *Nucleic Acids Res.* **2018**, *46*, 6455–6469.
- [551] S. Naldiga, S. Ji, J. Thomforde, C. M. Nicolae, M. Lee, Z. Zhang, G.-L. Moldovan, N. Y. Tretyakova, A. K. Basu, „Error-prone replication of a 5-formylcytosine-mediated DNA-peptide cross-link in human cells“, *J. Biol. Chem.* **2019**, *294*, 10619–10627.
- [552] S. Ji, J. Thomforde, C. Rogers, I. Fu, S. Broyde, N. Y. Tretyakova, „Transcriptional bypass of DNA–protein and DNA–peptide conjugates by T7 RNA polymerase“, *ACS Chem. Biol.* **2019**, *14*, 2564–2575.
- [553] S. Ji, D. Park, K. Kropachev, M. Kolbanovskiy, I. Fu, S. Broyde, M. Essawy, N. E. Geacintov, N. Y. Tretyakova, „5-Formylcytosine-induced DNA–peptide cross-links reduce transcription efficiency, but do not cause transcription errors in human cells“, *J. Biol. Chem.* **2019**, *294*, 18387–18397.

- [554] S. Schiesser, T. Pfaffeneder, K. Sadeghian, B. Hackner, B. Steigenberger, A. S. Schröder, J. Steinbacher, G. Kashiwazaki, G. Höfner, K. T. Wanner, C. Ochsenfeld, T. Carell, „Deamination, oxidation, and C-C bond cleavage reactivity of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine“, *J. Am. Chem. Soc.* **2013**, *135*, 14593–14599.
- [555] R. D. Camerini-Otero, G. Felsenfeld, „Histone H3 disulfide dimers and nucleosome structure“, *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5519–5523.
- [556] J. L. García-Giménez, G. Olaso, S. B. Hake, C. Bönisch, S. M. Wiedemann, J. Markovic, F. Dasí, A. Gimeno, C. Pérez-Quilis, O. Palacios, M. Capdevila, J. Viña, F. V. Pallardó, „Histone H3 glutathionylation in proliferating mammalian cells destabilizes nucleosomal structure“, *Antioxid. Redox. Sign.* **2013**, *19*, 1305–1320.
- [557] F. Brotzel, Y. C. Chu, H. Mayr, „Nucleophilicities of primary and secondary amines in water“, *J. Org. Chem.* **2007**, *72*, 3679–3688.
- [558] I. Rayment, „Reductive alkylation of lysine residues to alter crystallization properties of proteins“, *Methods Enzymol.* **1997**, *276*, 171–179.
- [559] A. Kirchner, *Untersuchung des Metabolismus nicht-kanonischer DNA-Basen in embryonalen Stammzellen*, Ludwig-Maximilians-Universität (München), **2020**.
- [560] F. R. Traube, *Regulation of α -ketoglutarate dependent dioxygenase TET3 – novel insights into epigenetic mechanisms in neurons*, Ludwig-Maximilians-Universität (München), **2020**.
- [561] C. Iacobucci, C. Piotrowski, R. Aebbersold, B. C. Amaral, P. Andrews, K. Bernfur, C. Borchers, N. I. Brodie, J. E. Bruce, Y. Cao, S. Chaignepain, J. D. Chavez, S. Claverol, J. Cox, T. Davis, G. Degliesposti, M.-Q. Dong, N. Edinger, C. Emanuelsson, M. Gay, M. Götze, F. Gomes-Neto, F. C. Gozzo, C. Gutierrez, C. Haupt, A. J. R. Heck, F. Herzog, L. Huang, M. R. Hoopmann, N. Kalisman, O. Klykov, Z. Kukačka, F. Liu, M. J. MacCoss, K. Mechtler, R. Mesika, R. L. Moritz, N. Nagaraj, V. Nesati, A. G. C. Neves-Ferreira, R. Ninnis, P. Novák, F. J. O'Reilly, M. Pelzing, E. Petrotchenko, L. Piersimoni, M. Plasencia, T. Pukala, K. D. Rand, J. Rappsilber, D. Reichmann, C. Sailer, C. P. Sarnowski, R. A. Scheltema, C. Schmidt, D. C. Schriemer, Y. Shi, J. M. Skehel, M. Slavin, F. Sobott, V. Solis-Mezarino, H. Stephanowitz, F. Stengel, C. E. Stieger, E. Trabjerg, M. Trnka, M. Vilaseca, R. Viner, Y. Xiang, S. Yilmaz, A. Zelter, D. Ziemianowicz, A. Leitner, A. Sinz, „First community-wide, comparative cross-linking mass spectrometry study“, *Anal. Chem.* **2019**, *91*, 6953–6961.
- [562] The Global Proteome Machine, *cRAP protein sequences*, <https://thegpm.org/crap/index.html> (abgerufen am 12. Februar 2020).
- [563] C. Iacobucci, M. Götze, C. H. Ihling, C. Piotrowski, C. Arlt, M. Schäfer, C. Hage, R. Schmidt, A. Sinz, „A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein–protein interactions“, *Nat. Protoc.* **2018**, *13*, 2864–2889.
- [564] R. Beveridge, J. Stadlmann, J. M. Penninger, K. Mechtler, „A synthetic peptide library for benchmarking crosslinking-mass spectrometry search engines for proteins and protein complexes“, *Nat. Commun.* **2020**, *11*, 742.
- [565] T. P. Hopp, K. S. Prickett, V. L. Price, R. T. Libby, C. J. March, D. Pat Cerretti, D. L. Urdal, P. J. Conlon, „A short polypeptide marker sequence useful for recombinant protein identification and purification“, *Bio/Technology* **1988**, *6*, 1204–1210.

-
- [566] C. Alcaraz, M. De Diego, M. J. Pastor, J. M. Escribano, „Comparison of a radioimmunoprecipitation assay to immunoblotting and ELISA for detection of antibody to African swine fever virus“, *J. Vet. Diagn. Invest.* **1990**, *2*, 191–196.
- [567] C. M. Gorman, B. H. Howard, R. Reeves, „Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate“, *Nucleic Acids Res.* **1983**, *11*, 7631–7648.
- [568] L. C. M. Ngoka, „Sample prep for proteomics of breast cancer: proteomics and gene ontology reveal dramatic differences in protein solubilization preferences of radioimmunoprecipitation assay and urea lysis buffers“, *Proteome Sci.* **2008**, *6*, 30.
- [569] M. Stadlmeier, *Synthese Sulfoxid-basierter Reagenzien für die Proteomik und deren Anwendung*, Ludwig-Maximilians-Universität (München), **2019**.
- [570] V. Mastorodemos, D. Kotzamani, I. Zaganas, G. Arianoglou, H. Latsoudis, A. Plaitakis, „Human *GLUD1* and *GLUD2* glutamate dehydrogenase localize to mitochondria and endoplasmic reticulum“, *Biochem. Cell Biol.* **2009**, *87*, 505–516.
- [571] R. Huisgen, „1,3-Dipolar cycloadditions. Past and future“, *Angew. Chem. Int. Ed. Engl.* **1963**, *2*, 565–598; „1.3-Dipolare Cycloadditionen. Rückschau und Ausblick“, *Angew. Chem.* **1963**, *75*, 604–637.
- [572] R. Huisgen, „Kinetics and mechanism of 1,3-dipolar cycloadditions“, *Angew. Chem. Int. Ed. Engl.* **1963**, *2*, 633–645; „Kinetik und Mechanismus 1.3-Dipolarer Cycloadditionen“, *Angew. Chem.* **1963**, *75*, 742–754.
- [573] C. Tanford, „Protein denaturation“ in *Advances in Protein Chemistry, Vol. 23* (Eds.: C. B. Anfinsen, M. L. Anson, J. T. Edsall, F. M. Richards), Academic Press, Cambridge, **1968**, 121–282.
- [574] P. H. von Hippel, K.-Y. Wong, „On the conformational stability of globular proteins: the effects of various electrolytes and nonelectrolytes on the thermal ribonuclease transition“, *J. Biol. Chem.* **1965**, *240*, 3909–3923.
- [575] M. W. Washabaugh, K. D. Collins, „The systematic characterization by aqueous column chromatography of solutes which affect protein stability“, *J. Biol. Chem.* **1986**, *261*, 12477–12485.
- [576] F. Vanzi, B. Madan, K. Sharp, „Effect of the protein denaturants urea and guanidinium on water structure: a structural and thermodynamic study“, *J. Am. Chem. Soc.* **1998**, *120*, 10748–10753.
- [577] A. W. Omta, M. F. Kropman, S. Woutersen, H. J. Bakker, „Negligible effect of ions on the hydrogen-bond structure in liquid water“, *Science* **2003**, *301*, 347–349.
- [578] J. D. Batchelor, A. Olteanu, A. Tripathy, G. J. Pielak, „Impact of protein denaturants and stabilizers on water structure“, *J. Am. Chem. Soc.* **2004**, *126*, 1958–1961.
- [579] M. C. Gurau, S.-M. Lim, E. T. Castellana, F. Albertorio, S. Kataoka, P. S. Cremer, „On the mechanism of the Hofmeister effect“, *J. Am. Chem. Soc.* **2004**, *126*, 10522–10523.

- [580] Y. Zhang, S. Furyk, D. E. Bergbreiter, P. S. Cremer, „Specific ion effects on the water solubility of macromolecules: PNIPAM and the Hofmeister series“, *J. Am. Chem. Soc.* **2005**, *127*, 14505–14510.
- [581] J. Szychowski, A. Mahdavi, J. J. L. Hodas, J. D. Bagert, J. T. Ngo, P. Landgraf, D. C. Dieterich, E. M. Schuman, D. A. Tirrell, „Cleavable biotin probes for labeling of biomolecules via azide–alkyne cycloaddition“, *J. Am. Chem. Soc.* **2010**, *132*, 18351–18360.
- [582] K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill, Q. Wang, „A fluorogenic 1,3-dipolar cycloaddition reaction of 3-azidocoumarins and acetylenes“, *Org. Lett.* **2004**, *6*, 4603–4606.
- [583] G. Allen, R. O. Campbell, „Specific cleavage of histidine-containing peptides by copper(II)“, *Int. J. Pept. Protein Res.* **1996**, *48*, 265–273.
- [584] I. M. Klotz, H. G. Curme, „The thermodynamics of metallo-protein combinations. Copper with bovine serum albumin“, *J. Am. Chem. Soc.* **1948**, *70*, 939–943.
- [585] C. W. M. Orr, „Studies on ascorbic acid. II. Physical changes in catalase following incubation with ascorbate or ascorbate and copper(II)“, *Biochemistry* **1967**, *6*, 3000–3006.
- [586] G. Marx, M. Chevion, „Site-specific modification of albumin by free radicals. Reaction with copper(II) and ascorbate“, *Biochem. J.* **1986**, *236*, 397–400.
- [587] K. Sadrerafi, E. E. Moore, M. W. Lee, „Association constant of β -cyclodextrin with carboranes, adamantane, and their derivatives using displacement binding technique“, *J. Incl. Phenom. Macro.* **2015**, *83*, 159–166.
- [588] U. K. Laemmli, „Cleavage of structural proteins during the assembly of the head of bacteriophage T4“, *Nature* **1970**, *227*, 680–685.
- [589] F. R. Traube, S. Schiffers, K. Iwan, S. Kellner, F. Spada, M. Müller, T. Carell, „Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides“, *Nat. Protoc.* **2019**, *14*, 283–312.
- [590] G. E. Moore, R. E. Gerner, H. A. Franklin, „Culture of normal human leukocytes“, *J. Amer. Med. Assoc.* **1967**, *199*, 519–524.
- [591] J. D. Dignam, R. M. Lebovitz, R. G. Roeder, „Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei“, *Nucleic Acids Res.* **1983**, *11*, 1475–1489.
- [592] N. A. Kulak, G. Pichler, I. Paron, N. Nagaraj, M. Mann, „Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells“, *Nat. Methods* **2014**, *11*, 319–324.
- [593] M. C. Chambers, B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T. A. Baker, M.-Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S. L. Seymour, L. M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E. W. Deutsch, R. L. Moritz, J. E. Katz, D. B. Agus, M. MacCoss, D. L. Tabb, P. Mallick, „A cross-platform toolkit for mass spectrometry and proteomics“, *Nat. Biotechnol.* **2012**, *30*, 918–920.

8 Anhang

8.1 Sequenzen der verwendeten Fusionsproteine

8.1.1 GDH-FLAG

MYRRLGEALLSRAGPAALGSAAADSAALLGWARGQPSAAPQPLTPVARRHYSEAAADREDDPNFFKMVEGFF
 DRGASIVEDKLVEDLKTRESEEQKRNRVRGILRIIKPCNHVLSLSPFIRRDDGSWEVIEGYRAQHSQHRTPCGGIRYS
 TDVSVDEVKALASLMTYKCAVVDVPPFGGAKAGVKINPKNYTDNELEKITRRFTMELAKKGFIPGIDVPAPDMSTG
 EREMSWIADTYASTIGHYDINAHACVTGKPISQGGIHRISATGRGVFHGIENFINEASYMSILGMTPGFGDKTFVV
 QGFGNVGLHSMRYLHRFGAKCVGVGESDGSIWNPDGIDPKELEDFKLQHGSLGFPKAKVYEGSILEADCILIPAA
 SEKQLTKSNAPRVKAKIIAEGANGPTTPEADKIFLERNIMVIPDLYLNAGGVTVSYFEWLKNLNHVSYGRLLTFKYERD
 SNYHLLMSVQESLERKFGKHGGTIPVVPTAEFQDRISGASEKDIVHSLAYTMERSARQIMRTAMKYNLGLDLRTA
 AYNVAIEKVFVYNEAGVTFTDYKDDDDK

8.1.2 eGFP-TET3_{fl}

MVSKGEELFTGVVPIVELDGDVNGHKFSVSGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLYGVQCFSRYPD
 HMKQHDFFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKIDFKEDGNILGHKLEYNYNNSHNVI
 MADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAA
 GITLGMDELYKSGENLYFQGRSRGDRMSQFQVPLAVQPDLGLYDFPQQQVMVGGFQGPGLPMAGSETQLRGG
 GDGRKKRRCGTCDPCRRLENGSCTSCTNRRTHQICKLRKCEVLKKGAGLLKEVEINAREGTGPWAQGATVKTGS
 ELSVDGPPVPGQMDSGPVYHGDSRQLSTSGAPVNGAREPAGPGLLGAAGPWRVDQKPDWEAASGPTHAAARLE
 DAHDLVAFSAVAEAVSSYGALSTRLYETFNREMSREAGSNGRGRPESCSSEGEDLDTLQTALALARHGMKPPNCT
 CDGPECPDFLEWLEGKIKSMAMEGGQGRPRLPGALPPSEAGLPAPSTRPPLSSEVPQVPPLEGLPLSQSALSIAKEK
 NISLQTAIAIEALTQLSSALPQPSHSTSQASCLPEALSPSAPFRSQSYLRAPSWPVVPEEHPFAPDSPAFFPATPR
 PEFSEAWGTDTPPATPRNSWPVPRPSPDMAELEQLLGSASDIQSVFKRPEALPTKPKVKVEAPSSSPAPVPSPI
 QREAPLLSEPDTHQKAQTALQQHLHHRNLFLEQAQDASFTSTEPQAPGWWAPPGSPAPRPPDKPPKEKPKKPP
 PTPAGGPVGAEKTPGIKTSVRKPIQIKKSRSDMQFLPLVVRQIVLEGLKQASEGQAPLPAQLSVPPPASQGAAS
 QSCATPLTPEPSLALFAPSPSGDLSLLPPTQEMRSPSPMVALQSGSTGGPLPPADDKLEELIRQFEAFGDSFGLPGPP
 SVPIQEPENQSTCLPAPESPFATRSPKKIKIESSGAVTVLSTTCFHSEEGQEATPTKAENLPTPLSGFLESPLKYLDTP
 TKSLLDTPAKKAQSEFPTCDCVEQIVEKDEGPYYTHLGSQPTVASIRELMEDRYGEKGAIRIEKVIYTGKEGKSSRGC
 PIAKWVIRRHTLEEKLLCLRHRAGHHCQNAVIVILILAWEGIPRSLGDTLYQELDTLRKYGNPTSRRRCGLNDDRTC
 ACQGKDPNTCGASFSFGCSWSMYFNGCKYARSKTPRKFRLTGDNPKEEEVLRNSFQDLATEVAPLYKRLAPQAYQ
 NQVTNEDVAIDCRLGLKEGRPFSGVTACMDFCAHAKDQHNLNGCTVVCTLTKEDNRCVGGQIPEDEQLHVLP
 KMASTDEFGEENQNAKVSSGAIQVLTAFPREVRRLEPEAKSCRQRQLEARKAAAEKLLQKEKLSTPEKIQEALEL
 AGVTTDPGLSLKGLSQSLKPSLKVPEQNHFSFKYSGNAVVESYVLGSCRPSDPYSMSSVYSYHSRYAQPGLAS
 VNGFHSKYTLPSFGYYGFPSSNPVFPQFLGPSAWGHGGSGGSFEKKPDLHALHNSLNPAYGGAFAELPGQAVAT
 DNHHPIPHHQQPAYPGPKEYLLPKVPLHPASRDPSFPAQSSSCYNRSIKQEPIDPLTQAESIPRDSAKMSRTPLEA
 SQNGGPSHLWGQYSGGSPMSPKRTNSVGGNWGVFPGESPTIVPDKLNSFGASCLTPSHFPESQWGLFTGEGQ
 QSAPHAGARLRGKPWSPCKFGNGTSALTGPSLTEKPWGMGTGDFNPALKGGPGFQDKLWNPVKVEEGRIPTPG
 ANPLDKAWQAFGMPLSSNEKLFALKSEEKLWDPFSLEEGTAEPPSKGVVKEEKSGPTVEEDEEELWSDSEHNFL
 DENIGGVAVAPAHCSILIECARRELHATTPLKKNRCHPTRISLVFYQHKNLNQPNHGLALWEAKMKQLAERARQR
 QEEAARLGLGQQEAKLYGKRRKWGGAMVAEPQHKEKKAIPTRQALAMPTDSAVTVSSYAYTKVTGPYSRWI

8.1.3 eGFP-TET3_{cd}

MVSKGEELFTGVVPIVELDGDVNGHKFSVSGEGEGDATYGKLTLLKFICTTGKLPVPWPTLVTTLTLYGVQCFSRYPD
 HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYI
 MADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSQDPNEKRDHMLLEFVTAA
 GITLGMDELYKSGENLYFQGRSRGDRMVALQSGSTGGPLPPADDKLEELIRQFEAEFGDSFGLPGPPSVPIQEPENQ
 STCLPAPESPFATRSPKIKIESSGAVTVLSTTCFHSEEGQEATPTKAENPLTPTLSGFLESPLKYLDTPTKSLLDTPAK
 KAQSEFPTCDCVEQIVEKDEGPYYTHLGSQPTVASIRELMEDRYGEKGAIRIEKVIYTGKEGKSSRGCPIAKWVIRRH
 TLEEKLLCLRHRAGHHCQNAVIVILILAWEGIPRSLGDTLYQELDTLRKYGNPTSRRCGLNDDRTCACQGKDPNT
 CGASFSGCWSMYFNGCKYARSKTPRKFRLTGDNPKEEVLRNSFQDLATEVAPLYKRLAPQAYQNQVTNEDVAI
 DCRLGLKEGRPFSGVTACMDFCAHAHKDQHNLYNGCTVVCTLTKEDNRCVQIPEDEQLHVLPLYKMASTDEFGS
 EENQNAKVSSGAIQVLTAFPVRRRLPEPAKSCRQRQLEARKAAAEEKKQKEKLTPEKIQEAELELAGVTTDPGLS
 LKGGLSQQSLKPSLKVPEQNHFFSKYSGNAVVESYVLGSCRPSDPYSMSSVSYHSRYAQPLASVNGFHSKYTLF
 SFGYYGFPSNPVFPVFLGSAWGHGGSGSFEKPKDLHALHNSLNPAWGGAFAELPGQAVATDNHHPIPHHQ
 QPAYPGKEYLLPKVPQLHPASRDSPFAQSSSCYNRSIKQEPIDPLTQAESIPRDSAKMSRTPLEASQNGGPSHL
 WGQYSGGSPMSPKRTNSVGGNWGVFPPGESPTIVPKLNSFGASCLTPSHFPESQWGLFTGEGQQSAPHAGARL
 RGKPWSPCKFGNGTSALTGPSLTKPWGMGTGDFNPALKGGPGFQDKLWNPVKVEEGRIPTPGANPLDKAWQA
 FGMPSSNEKLFGALKSEEKLWDPFSLEEGTAEPPSKGVVKEEKSGPTVEEDEEELWSDSEHNFLDENIGGVAVAP
 AHCSILIECARRELHATTPLKKNRCHPTRISLVFYQHKNLNQPNHGLALWEAKMKQLAERARQRQEAAARLGLGQ
 QEAKLYGKKRWGGAMVAEPQHKEKGAIPTRQALAMPTDSAVTVSSYAYTKVTGPYSRWI

8.2 Parameter der LC-MS/MS-Analysen

Tabelle 8.1: verwendete Gradienten der vorgeschalteten HPLC-Trennung der Peptidproben

| HPLC-Gradient | | |
|---------------|--------------------------|---------------------------|
| | 1 | 2 |
| Gradient | 6 – 34 % Lösungsmittel B | 14 – 34 % Lösungsmittel B |
| Dauer | 85 min | 46 min |
| | 3 | 4 |
| Gradient | 8 – 32 % Lösungsmittel B | 8 – 32 % Lösungsmittel B |
| Dauer | 35 min | 95 min |
| | 5 | |
| Gradient | 6 – 28 % Lösungsmittel B | |
| Dauer | 34 min | |

Tabelle 8.2: MS/MS-Parameter der in Kapitel 5.3.2.7 sowie 5.3.2.8 durchgeführten Analysen

| Overall Method Settings | |
|---------------------------------|-----------------------------------|
| <i>Use lock masses</i> | <i>best</i> |
| <i>Lock mass injection</i> | – |
| <i>Chrom. peak width (FWHM)</i> | 15 s |
| General | |
| <i>Polarity</i> | <i>positive</i> |
| <i>In-source CID</i> | 0.0 eV |
| <i>Default charge state</i> | 2 |
| Full MS | |
| <i>Microscans</i> | 1 |
| <i>Resolution</i> | 120000 |
| <i>AGC target</i> | 3×10^6 |
| <i>Maximum IT</i> | 20 ms |
| <i>Number of scan ranges</i> | 1 |
| <i>Scan range</i> | 300 – 1750 <i>m/z</i> |
| <i>Spectrum data type</i> | <i>Profile</i> |
| MS2 | |
| <i>Microscans</i> | 1 |
| <i>Resolution</i> | 15000 |
| <i>AGC target</i> | 2×10^5 |
| <i>Maximum IT</i> | 110 ms |
| <i>Loop count</i> | 15 |
| <i>MSX count</i> | 1 |
| <i>TopN</i> | 15 |
| <i>Isolation window</i> | 2.0 <i>m/z</i> |
| <i>Isolation offset</i> | 0.3 <i>m/z</i> |
| <i>Scan range</i> | 200 – 2000 <i>m/z</i> |
| <i>Fixed first mass</i> | – |
| <i>NCE</i> | 28 |
| <i>Spectrum data type</i> | <i>Profile</i> |
| <i>Minimum AGC target</i> | 2.00×10^3 |
| <i>Intensity threshold</i> | 1.8×10^4 |
| <i>Apex trigger</i> | – |
| <i>Charge exclusion</i> | <i>unassigned, 1, 7, 8, >8</i> |
| <i>Peptide match</i> | <i>off</i> |
| <i>Exclude isotopes</i> | <i>on</i> |
| <i>Dynamic exclusion</i> | 20.0 s |
| <i>If idle...</i> | <i>do not pick others</i> |

Tabelle 8.3: LC-MS/MS-Parameter der in Kapitel 5.3.2.9 und 5.3.2.10 durchgeführten Analysen

| Overall Method Settings | |
|---------------------------------|-----------------------------------|
| <i>Use lock masses</i> | <i>best</i> |
| <i>Lock mass injection</i> | – |
| <i>Chrom. peak width (FWHM)</i> | 15 s |
| General | |
| <i>Polarity</i> | <i>positive</i> |
| <i>In-source CID</i> | 0.0 eV |
| <i>Default charge state</i> | 2 |
| Full MS | |
| <i>Microscans</i> | 1 |
| <i>Resolution</i> | 120000 |
| <i>AGC target</i> | 1×10^6 |
| <i>Maximum IT</i> | 175 ms |
| <i>Number of scan ranges</i> | 1 |
| <i>Scan range</i> | 300 – 1750 <i>m/z</i> |
| <i>Spectrum data type</i> | <i>Profile</i> |
| MS2 | |
| <i>Microscans</i> | 1 |
| <i>Resolution</i> | 30000 |
| <i>AGC target</i> | 1×10^5 |
| <i>Maximum IT</i> | 70 ms |
| <i>Loop count</i> | 20 |
| <i>MSX count</i> | 1 |
| <i>TopN</i> | 20 |
| <i>Isolation window</i> | 2.0 <i>m/z</i> |
| <i>Isolation offset</i> | 0.5 <i>m/z</i> |
| <i>Scan range</i> | 200 – 2000 <i>m/z</i> |
| <i>Fixed first mass</i> | – |
| <i>NCE</i> | 28 |
| <i>Spectrum data type</i> | <i>Centroid</i> |
| <i>Minimum AGC target</i> | 1.00×10^3 |
| <i>Intensity threshold</i> | 1.4×10^4 |
| <i>Apex trigger</i> | 10 – 20 s |
| <i>Charge exclusion</i> | <i>unassigned, 1, 7, 8, >8</i> |
| <i>Peptide match</i> | <i>Preferred</i> |
| <i>Exclude isotopes</i> | <i>on</i> |
| <i>Dynamic exclusion</i> | 25.0 s |
| <i>If idle...</i> | <i>do not pick others</i> |

Tabelle 8.4: LC-MS/MS-Parameter der in Kapitel 5.4 durchgeführten Analysen

| Overall Method Settings | |
|---------------------------------|--------------------------------|
| <i>Use lock masses</i> | <i>best</i> |
| <i>Lock mass injection</i> | – |
| <i>Chrom. peak width (FWHM)</i> | 15 s |
| General | |
| <i>Polarity</i> | <i>positive</i> |
| <i>In-source CID</i> | 0.0 eV |
| <i>Default charge state</i> | 2 |
| Full MS | |
| <i>Microscans</i> | 1 |
| <i>Resolution</i> | 120000 |
| <i>AGC target</i> | 3×10^6 |
| <i>Maximum IT</i> | 20 ms |
| <i>Number of scan ranges</i> | 1 |
| <i>Scan range</i> | 300 – 1750 <i>m/z</i> |
| <i>Spectrum data type</i> | <i>Profile</i> |
| MS2 | |
| <i>Microscans</i> | 1 |
| <i>Resolution</i> | 30000 |
| <i>AGC target</i> | 2×10^5 |
| <i>Maximum IT</i> | 100 ms |
| <i>Loop count</i> | 15 |
| <i>MSX count</i> | 1 |
| <i>TopN</i> | 15 |
| <i>Isolation window</i> | 1.6 <i>m/z</i> |
| <i>Isolation offset</i> | 0.2 <i>m/z</i> |
| <i>Scan range</i> | 200 – 2000 <i>m/z</i> |
| <i>Fixed first mass</i> | – |
| <i>NCE</i> | 25, 30, 35 |
| <i>Spectrum data type</i> | <i>Profile</i> |
| <i>Minimum AGC target</i> | 5.00×10^3 |
| <i>Intensity threshold</i> | 5.0×10^4 |
| <i>Apex trigger</i> | – |
| <i>Charge exclusion</i> | <i>unassigned, 1, 2, >8</i> |
| <i>Peptide match</i> | <i>off</i> |
| <i>Exclude isotopes</i> | <i>on</i> |
| <i>Dynamic exclusion</i> | 20.0 s |
| <i>If idle...</i> | <i>do not pick others</i> |

Tabelle 8.5: LC-MS/MS-Parameter der in Kapitel 5.5.1 sowie 5.5.3 durchgeführten Analysen

| Parameter | ursprüngliche Methode | optimierte Methode |
|---------------------------------|-----------------------|--------------------------------------|
| Overall Method Settings | | |
| <i>Use lock masses</i> | | <i>best</i> |
| <i>Lock mass injection</i> | | – |
| <i>Chrom. peak width (FWHM)</i> | | 15 s |
| General | | |
| <i>Polarity</i> | | <i>positive</i> |
| <i>In-source CID</i> | | 0.0 eV |
| <i>Default charge state</i> | | 2 |
| Full MS | | |
| <i>Microscans</i> | | 1 |
| <i>Resolution</i> | | 120000 |
| <i>AGC target</i> | | 3×10^6 |
| <i>Maximum IT</i> | | 20 ms |
| <i>Number of scan ranges</i> | | 1 |
| <i>Scan range</i> | | 300 – 1750 <i>m/z</i> |
| <i>Spectrum data type</i> | | <i>Profile</i> |
| MS2 | | |
| <i>Microscans</i> | | 1 |
| <i>Resolution</i> | | 15000 |
| <i>AGC target</i> | 1×10^5 | 2×10^5 |
| <i>Maximum IT</i> | 100 ms | 125 ms |
| <i>Loop count</i> | | 15 |
| <i>MSX count</i> | | 1 |
| <i>TopN</i> | | 15 |
| <i>Isolation window</i> | | 1.6 <i>m/z</i> |
| <i>Isolation offset</i> | | 0.2 <i>m/z</i> |
| <i>Scan range</i> | | 200 – 2000 <i>m/z</i> |
| <i>Fixed first mass</i> | | – |
| <i>stepped NCE</i> | | 25, 30 |
| <i>Spectrum data type</i> | | <i>Centroid</i> |
| <i>Minimum AGC target</i> | 5.00×10^3 | 1.00×10^5 |
| <i>Intensity threshold</i> | 5.0×10^4 | 8.0×10^5 |
| <i>Apex trigger</i> | | – |
| <i>Charge exclusion</i> | | <i>unassigned, 1, 2, 7, 8, >8</i> |
| <i>Peptide match</i> | | <i>off</i> |
| <i>Exclude isotopes</i> | | <i>on</i> |
| <i>Dynamic exclusion</i> | | 10.0 s |
| <i>If idle...</i> | | <i>do not pick others</i> |

8.3 Einstellungen der Programme zur Auswertung der MS-Analysen

Tabelle 8.6: Einstellungen in *Proteome Discoverer 2.2*

| Spectrum Selector | | | |
|------------------------------------|----------------|-----------------------------------|-----------------------------|
| <i>Min. precursor mass</i> | 350 Da | | |
| <i>Max. precursor mass</i> | 5000 Da | | |
| <i>S/N Threshold</i> | 1.5 | | |
| Sequest HT | | XlinkX 2.0 | |
| <i>Max. missed cleavage sites</i> | 3 | <i>Max. missed cleavage sites</i> | 3 |
| <i>Min. peptide length</i> | 5 | <i>Min. peptide length</i> | 5 |
| <i>Max. peptide length</i> | 150 | <i>Precursor mass tolerance</i> | 10 ppm |
| <i>Precursor mass tolerance</i> | 10 ppm | <i>Fragment mass tolerance</i> | 20 ppm |
| <i>Fragment mass tolerance</i> | 0.02 Da | <i>Target FDR (strict)</i> | 0.01 (FDR strategy: simple) |
| <i>Fragment ions</i> | a, b, y | <i>Minimum score</i> | 40 |
| <i>Max. equal modifications</i> | 3 | <i>Minimum score difference</i> | 4 |
| Target Decoy PSM Validator | | | |
| <i>Maximum delta Cn</i> | 0.05 | | |
| <i>Target FDR (strict)</i> | 0.01 | | |
| <i>Target FDR (relaxed)</i> | 0.05 | | |
| Peptide Validator | | | |
| <i>Validation based on</i> | <i>q-Value</i> | | |
| Peptide Filter | | | |
| <i>Peptide confidence at least</i> | <i>High</i> | | |
| <i>Minimum peptide length</i> | 5 | | |

Tabelle 8.7: allgemeine Einstellungen in *MeroX 2.0.1.4*

| Protease sites | |
|--|---|
| <i>Protease site 1</i> | R?, missed cleavages: 3, blocked by: P, blocked as XL-site: no |
| <i>Protease site 2</i> | K?, missed cleavages: 3, blocked by: P, blocked as XL-site: yes |
| <i>Maximum total missed cleavages</i> | 3 |
| <i>Minimum peptide length</i> | 5 |
| <i>Maximum peptide length</i> | 30 |
| Modifications | |
| <i>Static modification</i> | from C to B (Carbamidomethylierung von Cystein) |
| <i>Variable modification</i> | from M to m (Oxidation von Methionin, max mod.: 1) |
| Crosslinker (s. Tab. 8.8 für Crosslinker-spezifische Einstellungen) | |
| <i>Specificity site 1</i> | K, Protein-N-Terminus |
| <i>Specificity site 2</i> | K, S, T, Y, Protein-N-Terminus |
| <i>Ignore consecutive peptides as cross-links</i> | enabled |
| <i>Dead-end cross-links (type 0)</i> | enabled |
| <i>Dead-ends reacted with</i> | H ₂ O |
| Mass comparison | |
| <i>Precursor precision (MS1)</i> | 6 ppm |
| <i>Mass recalibration (MS1)</i> | 0 ppm |
| <i>Fragment ion precision (MS2)</i> | 15 ppm |
| <i>Mass recalibration (MS2)</i> | 0 ppm |
| <i>Lower mass limit</i> | 1000 Da |
| <i>Upper mass limit</i> | 6000 Da |
| <i>S/N ratio</i> | 1.5 |
| <i>Ion types</i> | a, b, y |
| <i>Minimum nr. of fragments/peptide</i> | 3 |
| <i>Deisotoping</i> | enabled |
| <i>Correct precursor masses</i> | enabled (max. number of isotope shifts: 3) |
| <i>Analysis mode</i> | RISEUP (max. missing ions: 1) |
| <i>Neutral losses of fragment ions</i> | neutral loss of identified fragments (max. losses: 1) |
| Score & FDR | |
| <i>Apply prescore (%intensity)</i> | 5.0 |
| <i>FDR cut off (%)</i> | 1.0 |
| <i>Score cut off</i> | 50 |
| <i>Score Settings</i> | slow, precise scoring |
| <i>Decoy database generation</i> | shuffle sequences but keep protease sites |
| <i>Decoy database level</i> | Decoys based on peptide sequence shuffling/reversing |

Tabelle 8.8: Crosslinker-spezifische Einstellungen in *MeroX 2.0.1.4*

| Crosslinker | | |
|---|--|--|
| | DSSO | cliXlink |
| <i>Composition/Mass</i> | C ₆ H ₆ O ₃ S | C ₁₂ H ₁₃ NO ₄ S |
| <i>Maximum C_α-C_α-distance (Å)</i> | 23.5 | 30.0 |
| <i>Modification of fragments at site 1</i> | C ₃ H ₂ O* (* = essential) C ₃ H ₄ O ₂ S C ₃ H ₂ OS* | |
| <i>Modification of fragments at site 2</i> | C ₃ H ₂ O* C ₃ H ₄ O ₂ S C ₃ H ₂ OS* | C ₉ H ₉ NO ₂ * C ₉ H ₁₁ NO ₃ S C ₉ H ₉ NO ₂ S* |
| | cliXlink (+Biotin-Azid, reduziert und alkyliert) | cliXlink (+Azidoethanol) |
| <i>Composition/Mass</i> | C ₁₆ H ₂₁ N ₅ O ₅ S ₂ | C ₁₄ H ₁₃ N ₄ O ₅ S |
| <i>Maximum C_α-C_α-distance (Å)</i> | 30.0 | |
| <i>Modification of fragments at site 1</i> | C ₃ H ₂ O* C ₃ H ₄ O ₂ S C ₃ H ₂ OS* | |
| <i>Modification of fragments at site 2</i> | C ₁₃ H ₁₇ N ₅ O ₃ S* C ₁₃ H ₁₉ N ₅ O ₄ S ₂ C ₁₃ H ₁₇ N ₅ O ₃ S ₂ * | C ₁₁ H ₁₄ N ₄ O ₃ * C ₁₁ H ₁₆ N ₄ O ₄ S C ₁₁ H ₁₄ N ₄ O ₃ S* |
| | cliXlink (+Biotin-Azid_{Ester}) | cliXlink (+Biotin-Azid_{Amid}) |
| <i>Composition/Mass</i> | C ₂₆ H ₃₆ N ₆ O ₇ S ₄ | C ₂₆ H ₃₇ N ₇ O ₆ S ₄ |
| <i>Maximum C_α-C_α-distance (Å)</i> | 30.0 | |
| <i>Modification of fragments at site 1</i> | C ₃ H ₂ O* C ₃ H ₄ O ₂ S C ₃ H ₂ OS* | |
| <i>Modification of fragments at site 2</i> | C ₂₃ H ₃₂ N ₆ O ₅ S ₃ * C ₂₃ H ₃₄ N ₆ O ₆ S ₄ C ₂₃ H ₃₂ N ₆ O ₅ S ₄ * | C ₂₃ H ₃₃ N ₇ O ₄ S ₃ * C ₂₃ H ₃₅ N ₇ O ₅ S ₄ C ₂₃ H ₃₃ N ₇ O ₄ S ₄ * |
| | cliXlink (+Biotin-Azid_{Ester}, hydrolysierter Ester) | cliXlink (+Biotin-Azid_{Amid}, hydrolysiertes Amid) |
| <i>Composition/Mass</i> | C ₁₆ H ₂₂ N ₄ O ₅ S ₃ | C ₁₆ H ₂₃ N ₅ O ₄ S ₃ |
| <i>Maximum C_α-C_α-distance (Å)</i> | 30.0 | |
| <i>Modification of fragments at site 1</i> | C ₃ H ₂ O* C ₃ H ₄ O ₂ S C ₃ H ₂ OS* | |
| <i>Modification of fragments at site 2</i> | C ₁₃ H ₁₈ N ₄ O ₃ S ₂ * C ₁₃ H ₂₀ N ₄ O ₄ S ₃ C ₁₃ H ₁₈ N ₄ O ₃ S ₃ * | C ₁₃ H ₁₉ N ₅ O ₂ S ₂ * C ₁₃ H ₂₁ N ₅ O ₃ S ₃ C ₁₃ H ₁₉ N ₅ O ₂ S ₃ * |