

Untersuchung der
nanos-mRNA-Regulation in *Drosophila*
und
Softwareentwicklung für die Auswertung
von *Cross-Linking*/MS-Analysen

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt der

Naturwissenschaftlichen Fakultät I - Biowissenschaften

Martin-Luther-Universität Halle-Wittenberg

von Michael Götze

geboren am 03. Mai 1987 in Zwenkau

Gutachter:

Prof. Dr. Elmar Wahle

Prof. Dr. Andrea Sinz

Prof. Dr. Bernhard Spengler

verteidigt am 5.4.2017

Inhaltsverzeichnis

Inhaltsverzeichnis	iii
Zusammenfassung	ix
Summary	xi
I Regulation der <i>nanos</i>-mRNA in der frühen Embryogenese von <i>Drosophila melanogaster</i>	1
1 Einleitung	3
1.1 <i>Drosophila melanogaster</i> als Modellorganismus	3
1.1.1 Oogenese und frühe Embryonalentwicklung	3
1.1.2 Übergang von maternaler zu zygotischer Regulation der Genexpression	5
1.1.3 Deadenylierung von Transkripten durch den Ccr4-Not-Komplex	6
1.1.4 Achsenbildung im frühen Embryo	7
1.2 <i>nanos</i> -mRNA-Regulation in der frühen Entwicklung von <i>Drosophila</i>	8
1.3 Komponenten des reprimierten RNPs	11
1.3.1 Smaug	12
1.3.2 Cup	13
1.3.3 Me31B	14
1.3.4 Trailer Hitch	15
1.4 Zielstellung	17
2 Ergebnisse	19
2.1 Reinigung des Repressor-Komplexes - Vorüberlegungen	19
2.2 Reinigung des Repressor-Komplexes	21
2.2.1 Stabilisierung der <i>Bait</i> -RNAs	21
2.2.2 Reinigung des Repressor-Komplexes durch Gradientenzentrifugation	22
2.3 MS-Analyse	26
2.3.1 Identifizierte Proteine	31
Repressor-Komplex	31

	Ccr4-Not-Komplex	31
	CTLH-Komplex	31
	Die DEAD-Box-ATPase Belle	34
	Andere Proteine	35
2.4	Quantifizierung der identifizierten Repressor-Komplexkomponenten . .	36
2.4.1	Konzentration der identifizierten Proteine im Embryoextrakt	36
2.4.2	Stöchiometrie der Komponenten im Repressor-Komplex	37
2.5	Sedimentationsverhalten reprimierter RNA	41
2.5.1	Assoziation mit Ribosomen	41
2.5.2	Oligomerisierung der RNAs	41
2.5.3	Identifizierung regulatorischer RNAs	43
2.6	Interaktionspartner von Smaug	44
2.7	<i>Cross-Linking</i> der angereicherten Proteine	46
3	Diskussion	49
3.1	Subkomplexe im SRE-abhängigen RNP	49
3.1.1	Repressor-Komplex	50
3.1.2	Ccr4-Not-Komplex	51
3.1.3	CTLH-Komplex	53
	Ubiquitinierung von Me31B	53
3.1.4	Weitere identifizierte Proteine	54
3.2	Die Rolle von Belle im Repressor-Komplex	57
3.3	Me31B und Trailer Hitch bedecken die reprimierte RNA	59
3.3.1	Sedimentationsverhalten als Folge der Me31B•Tral-Oligomerisierung .	61
3.4	Modell der <i>nanos</i> -mRNA-Regulation während der frühen Embryogenese	62
3.4.1	<i>Cis</i> -Interaktion des 5'- und 3'-Endes einer mRNA	63
3.4.2	<i>Cross-Linking</i> /MS-Analyse des Smaug-RNP	64
3.5	Ausblick	65
II	Bioinformatische Analyse von <i>Cross-Linking</i>/MS-Analysen	69
1	Einleitung	71
1.1	Massenspektrometrische Methoden der strukturellen Proteomik	71
1.2	<i>Cross-Linking</i> in Kombination mit Massenspektrometrie	72
1.2.1	<i>Cross-Linking</i> -Reagenzien	75
	Reaktive Gruppen	75
1.2.2	Massenspektrometrische Analyse	76
	Fragmentierung von <i>Cross-Links</i>	76

	Fragmentierungsarten	79
1.2.3	Vereinfachung der Auswertung von <i>Cross-Linking</i> /MS-Datensätzen . .	80
	Anreicherung quervernetzter Peptide	80
	Auflösung von MS/MS-Spektren	81
	Isotopen-markierte <i>Cross-Linker</i>	81
	MS/MS-spaltbare <i>Cross-Linker</i>	81
1.3	Bioinformatische Auswertung von XL-MS-Experimenten	84
1.3.1	Verfügbare <i>Cross-Linking</i> /MS-Software	85
	Visualisierung von identifizierten <i>Cross-Links</i>	87
1.3.2	Angewendete Algorithmen zur Identifizierung von <i>Cross-Links</i>	87
	Definitionen zum Algorithmus	87
	Verketteten von Peptiden zu einer <i>Cross-Link</i> -Datenbank	88
	Suche von Peptiden mit unbekannter Modifikation	89
	Durchsuchen aller möglichen Kombinationen	89
2	Ergebnisse	91
2.1	Weiterentwicklung von StavroX und MeroX	91
2.2	Suchalgorithmen in StavroX und MeroX	92
2.2.1	Standard-Suchalgorithmus von StavroX und MeroX	92
2.2.2	MeroX-spezifischer Suchalgorithmus: RISE-Modus	95
2.3	Validierung von <i>Cross-Links</i>	96
2.3.1	<i>Score</i> -Berechnung	96
	Anzahl der identifizierten Ionen	98
	Ionenserien	98
	Reporterionen der <i>Cross-Linker</i> -Fragmentierung	99
	Vereinigung der Faktoren zum finalen <i>score</i>	99
2.3.2	Abschätzung der <i>False Discovery Rate</i> (FDR)	102
2.4	Beschreibung der Programmfunktionen	107
2.4.1	Programmeinstellungen	107
	Aminosäuren und Modifizierungen	107
	Proteaseschnittstellen	108
	<i>Cross-Linker</i> -Einstellungen	108
	Eingabe der Einstellungen in StavroX und MeroX	108
2.4.2	Kommandozeilenoptionen	109
2.4.3	Grafische Benutzeroberfläche (GUI)	109
2.4.4	Integration von xVis	113
2.4.5	Dateiformate	113

3	Diskussion	115
3.1	Anwendungsbeispiele der Software-Tools	116
3.2	Limitationen der <i>Cross-Link</i> -Analyse	118
3.3	Mögliche Optimierungen von StavroX und MeroX	119
3.3.1	Proteomweite Analysen	119
3.3.2	Optimierung der Suchalgorithmen	120
3.3.3	Standard-Datenformate	121
3.4	Zusammenfassung	122
M	- Material & Methoden	123
M.1	Material	123
M.1.1	Chemikalien	123
M.1.2	Geräte	126
M.1.3	Software	126
M.1.4	Plasmide	127
M.1.5	Antikörper	127
M.2	Methoden	128
M.2.1	Erzeugung eines Me31B-spezifischen Antikörpers	128
M.2.2	Präparation von <i>short</i> RNAs	128
M.2.3	<i>In-vitro</i> -Synthese von RNAs	129
	Synthese 5'-Cap-markierter RNAs	130
M.2.4	Präparation von <i>Drosophila</i> -Embryoextrakt (DEE)	130
M.2.5	<i>In-vitro</i> -Translations- und Repressionsmessungen	131
	Translation in <i>Drosophila</i> -Embryoextrakt (DEE)	131
	Translation in Kaninchen-Retikulozytenlysat (RRL)	132
M.2.6	Reinigung des reprimierten RNPs	132
	Saccharosegradientenzentrifugation	132
	Streptavidin-Biotin- <i>pull-down</i>	133
	<i>Pull-down</i> zur RNA-Analyse	134
M.2.7	<i>Deep-sequencing</i> -Analyse	134
M.2.8	MS-Analyse	135
	Auswertung der massenspektrometrischen Daten	135
M.2.9	<i>Western blotting</i>	136
	Quantitative Auswertung von <i>Western-blot</i> -Analysen	136
M.2.10	Immunpräzipitation von Smaug	137
M.2.11	Identifizierung von <i>Cross-Links</i> im Smaug-RNP	137
M.2.12	Suche nach SRE- <i>stem-loops</i> in 3'-UTRs von mRNAs	138
M.2.13	Verwendete Bibliotheken in StavroX und MeroX	138
M.2.14	Verwendete Datensätze zur Analyse mit MeroX	138

Literaturverzeichnis	I
Abbildungsverzeichnis	XXXI
Tabellenverzeichnis	XXXIII
Abkürzungsverzeichnis	XXXV
Anhang	XLI
Publikationsliste	LVII
Lebenslauf	LIX
Selbstständigkeitserklärung	LXI
Danksagung	LXIII

Zusammenfassung

In der vorliegenden Promotionsarbeit wurden zwei unterschiedliche Themen bearbeitet: die Regulation der *nanos*-mRNA in der frühen Embryogenese von *Drosophila melanogaster* und die bioinformatische Analyse von *Cross-Linking*/MS-Experimenten.

Die Regulation der Nanos-Expression während der frühen Embryonalentwicklung der Taufliege ist wichtig für die Ausbildung der anterior-posterioren Körperachse des Embryos sowie für die Erhaltung der Keimbahnzellen. Die Regulation beruht auf posttranskriptionalen Mechanismen, die unter anderem vom Protein Smaug abhängen. Smaug vermittelt die Bildung eines regulatorischen Komplexes an der *nanos*-mRNA, der sowohl den Abbau der mRNA als auch die Hemmung der Translation bewirkt. In dieser Arbeit wurden die Zusammensetzung des gebildeten Komplexes massenspektrometrisch im Detail untersucht und die Stöchiometrie der Komponenten gemessen. Die bisher bekannten Komponenten dieses Repressor-Komplexes konnten zuverlässig bestätigt werden. Zusätzlich wurde eine bisher unbekannte Komponente identifiziert: Belle, ein Protein der DEAD-Box-ATPase/Helikase-Familie. Belle konnte zum einen im Repressor-Komplex *in vitro* identifiziert werden. Zum anderen legen genetische Experimente eine Rolle von Belle in der posttranskriptionalen Regulation der *nanos*-mRNA *in vivo* nahe. Welche Rolle Belle dabei spielt, ist allerdings noch unklar und wird im Fokus weiterführender Arbeiten liegen.

Zwei weitere Proteine des Repressor-Komplexes, die DEAD-Box-ATPase Me31B und Trailer Hitch, bilden einen Komplex, welcher superstöchiometrisch zur RNA im Repressor-Komplex vorliegt. Möglicherweise bildet sich ein RNA-abhängiges Oligomer aus, das die RNA unzugänglich verpackt. Diese Oligomerisierung entlang der RNA könnte bislang ungeklärte Eigenschaften des Repressor-Komplexes wie z.B. die ATP-Abhängigkeit der Repressor-Komplexbildung erklären. Die Protein-Protein-Interaktion innerhalb des Me31B•Tral-Komplexes konnte durch chemisches *Cross-Linking* in Kombination mit Massenspektrometrie zudem bestätigt werden.

Für die Analyse solcher *Cross-Linking*/MS-Experimente werden im zweiten Teil zwei Programme (StavroX und MeroX) vorgestellt. Die Kombination aus chemischer Quervernetzung (*Cross-Linking*) und Massenspektrometrie (MS) erlaubt es, neben Informationen zu Protein-Protein-Interaktionen auch niedrig-aufgelöste Strukturinformationen von Proteinen

oder Proteinkomplexen zu erhalten. Dabei werden Seitenketten von Proteinen durch ein chemisches Reagenz mit zwei reaktiven Gruppen (*Cross-Linker*) kovalent miteinander verknüpft. Die genauen Verknüpfungsstellen der beteiligten Proteine können dann massenspektrometrisch analysiert werden. Dabei stellt die bioinformatische Analyse der komplexen Datensätze die größte Hürde dar.

Die beiden vorgestellten Programme reduzieren den Aufwand einer *Cross-Linking*/MS-Auswertung dramatisch. StavroX und MeroX wurden für unterschiedliche Arten chemischer *Cross-Links* entworfen und führen eine automatisierte Analyse von massenspektrometrischen Daten durch. Jeder identifizierte *Cross-Link* wird durch die Programme bewertet und es wird automatisch eine Falscherkennungsrate abgeschätzt. Zudem bieten die einfach zu bedienenden grafischen Oberflächen in den Programmen die Möglichkeit, erhaltene Ergebnisse der automatischen Analyse im Detail manuell zu prüfen. In der vorliegenden Arbeit wird die Funktionsweise der Programme und die drastische Vereinfachung der Analyse von *Cross-Linking*/MS-Experimenten durch die Programme aufgezeigt.

StavroX und MeroX wurden in mehreren Studien für die Analyse verschiedenster biologischer Fragestellungen eingesetzt und konnten auch im Fall der *nanos*-Regulation in der frühen Embryogenese für die Identifizierung einer biologisch relevanten Interaktionsstelle innerhalb des Repressor-Komplexes verwendet werden.

Summary

Two different topics were addressed in this thesis: the posttranscriptional regulation of *nanos* mRNA during early *Drosophila melanogaster* development and the bioinformatic analysis of cross-linking/MS studies.

The expression of Nanos is tightly controlled and very important during early development of the fruit fly. Nanos is required for the anterior-posterior axis formation as well as stem cell formation and maintenance. Its regulation relies on posttranscriptional mechanisms that depend on the protein Smaug and others. Association of Smaug with the *nanos* mRNA leads to the formation of an RNA-protein complex that induces transcript degradation and translational repression at the same time. In this thesis the composition of the Smaug-dependent repressor complex was analysed in detail by mass spectrometry and the stoichiometry of complex components was determined in vitro. All previously known components of the repressor complex could be confirmed in this analysis, and one unanticipated component, the DEAD-box-ATPase/helicase Belle, was identified. Genetic data indicated a role for Belle in the regulation of *nanos* in vivo. The exact role of Belle remains unknown and will be in the focus of subsequent analyses.

Two other proteins of the repressor complex, Me31B and Trailer Hitch, form a complex that binds in a superstoichiometric manner to the repressed RNA. This RNA-dependent oligomerization might cause packaging of the repressed RNA into an inaccessible particle and could account for several unexplained features of the repressor complex, like its ATP-dependent formation. The protein-protein interaction within this Me31B•Tral complex was confirmed using chemical cross-linking in combination with mass spectrometry (MS).

In the second part of this thesis two software tools (StavroX and MeroX) are presented that have been designed for the analysis of such cross-linking/MS experiments. The combination of cross-linking and mass spectrometry not only allows for the identification of protein-protein interactions but also for low-resolution structural probing of proteins or protein complexes. For that purpose, two adjacent protein side chains are covalently linked by a chemical cross-linker. The cross-links are then analysed by mass spectrometric techniques, allowing the identification of cross-linked sites with residue resolution. The bioinformatic analyses of the complex data sets present the bottleneck of this technique.

The two software tools presented simplify the time-consuming analysis of a cross-linking/MS study dramatically. StavroX and MeroX were designed for different kinds of chemical cross-linking reagents and perform an automated analysis of mass spectrometric data. Each identified cross-link is evaluated by the software and a false discovery rate is automatically estimated. The easy-to-use graphical user interface permits a fast manual inspection of identified cross-links. This thesis presents the algorithms applied and functionalities of both tools and emphasizes the drastic improvements of cross-linking/MS analyses made possible by StavroX and MeroX.

Both programs have been used in several studies for the analysis of diverse biological problems. Also in the case of *nanos* regulation during early fly development, a biological relevant interaction within the repressor complex could be confirmed by cross-linking and analysis using MeroX.

Teil I.

Regulation der *nanos*-mRNA in
der frühen Embryogenese von
Drosophila melanogaster

1. Einleitung

Für alle lebenden Zellen, egal, ob es sich um einen einzelligen Organismus oder einen komplexen multizellulären Organismus wie den Menschen handelt, ist es äußerst wichtig, dass verschiedene Gene zum richtigen Zeitpunkt am richtigen Ort und in der richtigen Dosis exprimiert werden. Die Regulation der Genexpression kann an vielen Stellen erfolgen: angefangen von der Transkription der DNA zu RNA über die Translation zu Proteinen bis hin zum Abbau der Proteine. Für jedes Gen können gleichzeitig verschiedene Mechanismen greifen, die in ihrer Gesamtheit die korrekte Genexpression hervorrufen. In dieser Arbeit soll ein solcher Regulationsmechanismus während der frühen Embryogenese der Taufliege (*Drosophila melanogaster*) näher untersucht werden. Die Identifizierung beteiligter Faktoren soll ein besseres Verständnis vom Mechanismus der Regulation liefern.

1.1. *Drosophila melanogaster* als Modellorganismus

Nach der Befruchtung eines *Drosophila*-Weibchens sammelt dieses die Spermien in einer Spermathek, um die gebildeten Eizellen direkt zu befruchten. Jedes *Drosophila*-Weibchen kann dann bis zu zwanzig Tage lang jeden Tag bis zu hundert befruchtete Embryonen ablegen. Die Embryonalentwicklung dauert insgesamt 22 - 24 h, durchläuft 14 unterscheidbare Stadien (Bownes, 1975) und endet mit dem Schlüpfen der ersten Larven. Nach drei Larvenstadien kommt es zur Verpuppung und insgesamt 10 - 12 Tage nach der Befruchtung schlüpfen adulte Fliegen. Weitere zwei Tage später sind die Weibchen geschlechtsreif und können nach der Befruchtung wiederum Embryonen ablegen.

Neben der kurzen Generationsdauer von zwei Wochen ist die Möglichkeit, leicht genetische Manipulationen durchführen zu können, ein Grund, weshalb die Taufliege als Modellorganismus ihre Berühmtheit erlangt hat.

1.1.1. Oogenese und frühe Embryonalentwicklung

Oozyten werden in *Drosophila melanogaster* in zwei Ovarien mit jeweils etwa 18 Ovariolen gebildet. Eine Ovariole enthält dabei mehrere Follikel (Abb. 1), die jeweils eine Oozyte enthalten. Dabei sind die Follikel einer Ovariole aneinandergereiht und enthalten vom proximalen bis zum distalen Ende des Ovariums Oozyten aller Entwicklungsstadien von der Keimbahn-Stammzelle bis zum reifen Ei.

Am proximalen Ende der Ovariole kommt es zur asymmetrischen Teilung einer Stammzelle in eine Zyste und eine weitere Stammzelle. Die Zyste durchläuft vier mitotische Teilungen, wobei die 16 entstehenden Zellen voneinander nicht getrennt werden, sondern durch Ringkanäle verbunden bleiben. Eine der 16 Zellen wird die spätere Oozyte, die als einzige in die Meiose eintritt. Die restlichen 15 Zellen werden Nährzellen, die die Oozyte später mit zytoplasmatischen Komponenten (z.B. mRNAs und Proteine) und Nährstoffen versorgen. Die Nährzellen und die Oozyte werden von Follikelzellen umgeben, die die äußeren Hüllen des späteren Embryos synthetisieren.

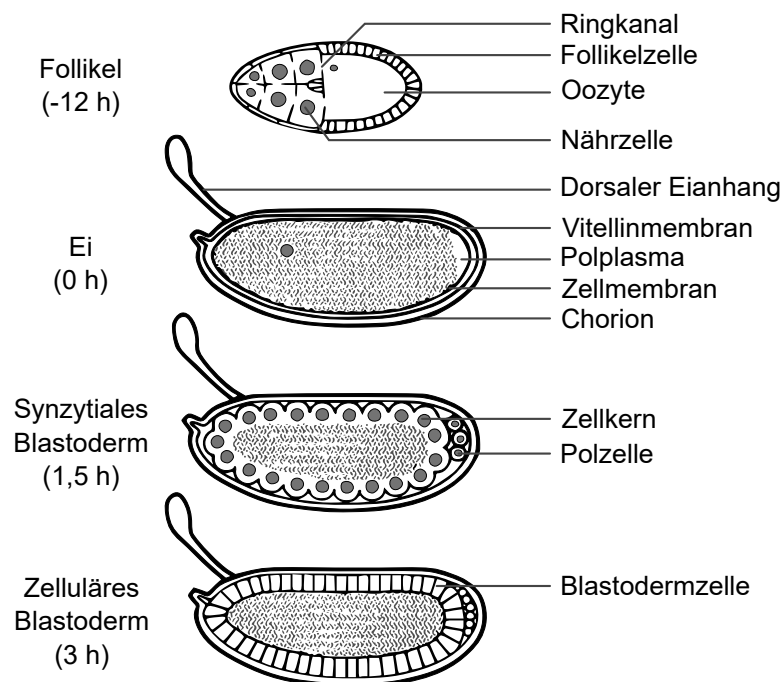


Abbildung 1: Frühe Entwicklung von *Drosophila melanogaster* - Schematische Darstellung von vier verschiedene Stadien der Entwicklung von *Drosophila*, von der reifenden Oozyte (12 h vor Befruchtung) bis zum zellulären Blastoderm (3 h nach Befruchtung). Modifiziert nach St Johnston u. Nüsslein-Volhard (1992).

Nach der Befruchtung des gebildeten Eis kommt es im Embryo zu 13 schnellen, synzytialen Kernteilungen, bei denen die Kerne etwa alle 10 min geteilt werden, ohne dass die Kerne durch eine Zellmembran voneinander getrennt werden. Nach der achten Kernteilung wandert der Großteil der Zellkerne zur Peripherie des Embryos. Während des synzytialen Blastoderms (Teilungen 10 - 13) werden die Polzellen vom Synzytium abgetrennt und bilden später die Keimbahn. Mit der 14. Kernteilung, nach etwa 3 h, kommt es zur Vereinzelung aller Kerne im zellulären Blastoderm (Abb. 1; zusammengefasst in Campos-Ortega u. Hartenstein, 1997; St Johnston u. Nüsslein-Volhard, 1992; Bastock u. St Johnston, 2008).

1.1.2. Übergang von maternalen zu zygotischer Regulation der Genexpression

Bis etwa zur Phase des synzytialen Blastoderms ist der Embryo transkriptionell inaktiv. Der Embryo ist also währenddessen auf maternale Proteine und RNAs angewiesen. Alle Regulationsmechanismen sind demnach posttranskriptionell, auch wenn neu-synthetisiert Transkripte einiger weniger Gene schon recht früh in der Embryogenese detektierbar sind (Ali-Murthy *et al.*, 2013). Die ersten zygotischen Transkripte, z.B. des *hunchback*-Gens¹, treten etwa nach der achten Teilung auf (Garcia *et al.*, 2013). Ab der achten Teilung wird auch die Rate der Kernteilungen geringer und die Teilungen geschehen asynchron. Damit beginnt die *midblastula transition* (MBT), welche der erste Prozess in der Entwicklung von *Drosophila* ist, der abhängig von der zygotischen Genom-Aktivierung (ZGA) abläuft. Ab der 14. Teilung werden die Kerne durch Membranen voneinander getrennt und der Embryo ist vollständig abhängig von der zygotischen Transkription. Der Übergang von maternalen Kontrolle der Genexpression zur zygotischen Kontrolle wird als MZT (*maternal-to-zygotic transition*) bezeichnet und stellt einen Meilenstein in der Entwicklung des Organismus dar (Harrison u. Eisen, 2015).

Die MZT ist einerseits abhängig vom Abbau der maternalen Faktoren (RNA und Proteine) und andererseits von der Synthese zygotischer RNAs, wobei diese Prozesse nicht unabhängig voneinander agieren. Der Abbau der maternalen mRNAs kann sowohl von maternalen als auch von zygotischen Komponenten abhängig sein (Bashirullah *et al.*, 1999). Die Destabilisierung der mRNAs wird initiiert durch die Deadenylierung, den Abbau des Poly(A)-Schwanzes. Für den spezifischen Abbau der maternalen Transkripte zum richtigen Zeitpunkt sind meist regulatorische Elemente im 3'-UTR der jeweiligen RNA nötig, die mit verschiedenen trans-Faktoren wie miRNAs oder RNA-Bindeproteinen interagieren.

Ein Beispiel für ein solches RNA-Bindeprotein ist das maternale Protein Smaug, welches im 3'-UTR von RNAs bindet und die Deadenylierungsmaschinerie rekrutiert (Semotok *et al.*, 2005). Der Abbau eines Großteils der maternalen Transkripte ist abhängig von der Anwesenheit des Smaug-Proteins (Tadros *et al.*, 2007). Mit dem Einsetzen der Transkription werden auch Faktoren (z.B. miRNAs oder RNA-Bindeproteine) synthetisiert, die den Abbau weiterer maternalen Transkripte einleiten und damit den zygotischen RNA-Abbau kontrollieren (Walser u. Lipshitz, 2011).

Nicht nur auf der RNA-Ebene vollzieht sich ein Übergang zwischen maternalen und zygotischen Genprodukten, sondern auch auf der Proteinebene. Vor allem maternale Effektoren,

¹Nomenklatur in *Drosophila*: Gene und mRNA-Transkripte dieser Gene werden klein und kursiv geschrieben. Das zugehörige Protein-Produkt wird groß geschrieben. Beispielhaft: *hunchback*-Gen, *hunchback*-mRNA, *hunchback*-Deletionsmutante und Hunchback-Protein.

wie z.B. die Proteine Smaug, Oskar oder Ypsilon Schachtel, werden während der MZT de-gradiert und zygotische Proteine synthetisiert, welche für die weitere Differenzierung des Embryos nötig sind (Gouw *et al.*, 2009).

1.1.3. Deadenylierung von Transkripten durch den Ccr4-Not-Komplex

Normalerweise beginnt der mRNA-Abbau mit der Verkürzung des Poly(A)-Schwanzes. Anschließend wird die RNA entweder nach einem *decapping*-Schritt durch 5'-3'-Exonukleasen wie Xrn1 oder in 3'-5'-Richtung durch das Exosom abgebaut (Abb. 2A; Garneau *et al.*, 2007). Der Ccr4-Not-Komplex stellt dabei die Hauptaktivität der Deadenylasen dar. In *Drosophila* besteht der konservierte Ccr4-Not-Komplex aus insgesamt acht Komponenten. Not1 ist mit etwa 280 kDa das größte Protein des Komplexes und bildet das Gerüst für die anderen Proteine des Komplexes (Abb. 2B). Die Proteine Twin (Ccr4) und Pop2 (Caf1) zeigen *in vitro* Deadenylase-Aktivität, wobei Pop2 *in vivo* die Hauptaktivität aufweist (Temme *et al.*, 2014). In *Xenopus* konnte gezeigt werden, dass Caf1 die Translation einer Reporter-RNA unabhängig von der Deadenylierung reprimieren kann (Cooke *et al.*, 2010; Mathys *et al.*, 2014; Waghay *et al.*, 2015).

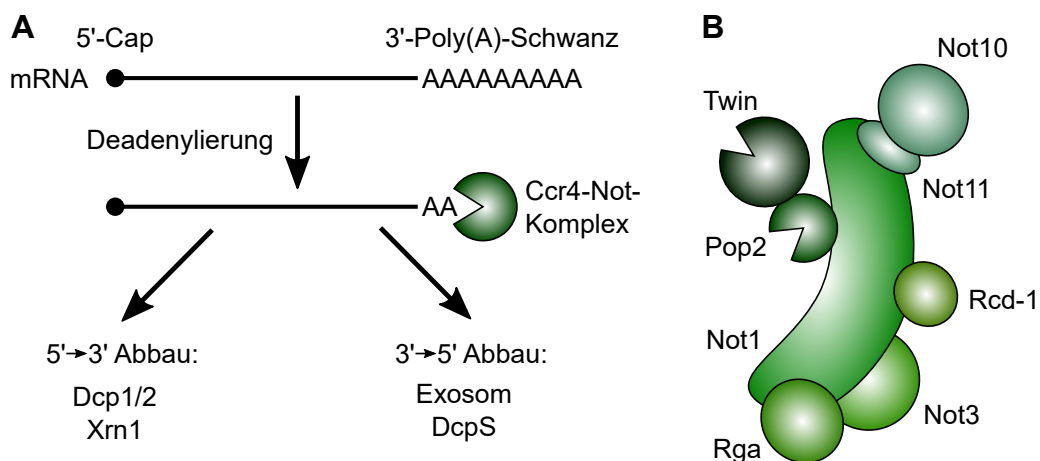


Abbildung 2: mRNA-Abbau - (A) Die Degradation einer mRNA wird durch den Abbau des Poly(A)-Schwanzes (Deadenylierung) initiiert. Anschließend wird entweder das 5'-Cap durch Dcp1/Dcp2 entfernt und die RNA durch Xrn1 abgebaut (5'-3'-Abbau) oder durch das Exosom in 3'-5'-Richtung abgebaut und das verbleibende Cap durch DcpS hydrolysiert (Garneau *et al.*, 2007). (B) Schematische Darstellung des Ccr4-Not-Komplexes, der die Deadenylierung bewirkt (Temme *et al.*, 2010).

Der Ccr4-Not-Komplex kommt in Oozyten, Embryonen und allen weiteren Entwicklungsstufen bis zum adulten Organismus vor und ist dabei zytoplasmatisch lokalisiert (Temme *et al.*, 2010). Die Deletion von Komponenten des Ccr4-Not-Komplexes führt zu längeren Poly(A)-Schwänzen in adulten Fliegen. Ein Knock-down einzelner Komponenten des Komplexes in *Drosophila*-S2-Zellen führt ebenso zur Verlängerung der Poly(A)-Schwänze und

außerdem zur Reduktion der Proteinmenge weiterer Untereinheiten des Ccr4-Not-Komplexes (Temme *et al.*, 2010). Dies zeigt, dass die Ccr4-Not-Proteine als Komplex agieren und für die Deadenylierung der mRNAs einer Zelle nötig sind.

Der Deadenylase-Komplex kann durch verschiedene RNA-Bindeproteine zu spezifischen Transkripten rekrutiert werden, um diese zu deadenylieren. Beispiele dafür sind Smaug (Semotok *et al.*, 2005), Nanos (Kadyrova *et al.*, 2007) oder Ago1/Gawky (Behm-Ansmant *et al.*, 2006). Die Rekrutierung durch Ago1 basiert dabei auf der Bindung von miRNAs im 3'-UTR des Transkriptes.

1.1.4. Achsenbildung im frühen Embryo

Bereits während der Oogenese werden die späteren Körperachsen der Larve und der adulten Fliege festgelegt. Durch Lokalisierung von *bicoid*-mRNA zum anterioren Pol und *oskar*-mRNA zum posterioren Pol wird die anterior-posterior-Achse (A/P-Achse) bereits in der Oozyte vordefiniert (Ephrussi *et al.*, 1991). Die Lokalisierung dieser mRNAs beruht dabei auf der Bewegung entlang des polarisierten Zytoskeletts. Aufgrund der jeweiligen Lokalisierung werden die Proteine Bicoid anterior und Oskar posterior synthetisiert.

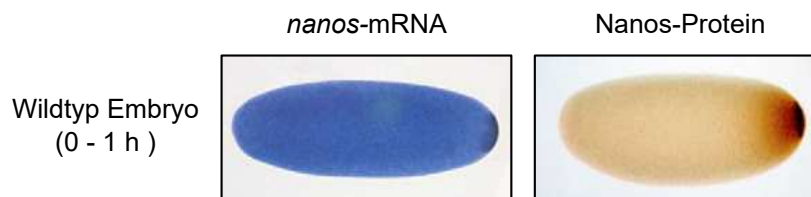


Abbildung 3: Verteilung der *nanos*-mRNA und des Nanos-Proteins im Embryo - Die *nanos*-mRNA lokalisiert im frühen Embryo zum posterioren Pol. Ein Großteil der RNA (blau gefärbt durch *In-situ*-Hybridisierung) liegt allerdings unlokalisiert vor (Bergsten u. Gavis, 1999). Das Nanos-Protein wird am posterioren Pol synthetisiert und bildet einen Gradienten vom posterioren Pol in Richtung des anterioren Pols aus (Antikörperfärbung). Unlokalisierte RNA wird translationell reprimiert und deadenyliert (Abbildungen aus Zaessinger *et al.*, 2006).

Die *nanos*-mRNA, deren Regulation im Fokus dieser Arbeit liegt, wird ebenfalls zum posterioren Pol transportiert. Die *nanos*-Lokalisierung ist allerdings nicht effizient. Nur etwa 4% der *nanos*-mRNA befinden sich später am posterioren Pol angereichert, die restliche RNA kommt im gesamten Embryo gleichmäßig verteilt vor und ist translationell reprimiert (Smibert *et al.*, 1996; Bergsten u. Gavis, 1999). Diese Repression wird am posterioren Pol vom Oskar-Protein aufgehoben, so dass Nanos nur am posterioren Pol synthetisiert wird und einen Proteingradienten von posterior zu anterior ausbildet (Abb. 3; Gavis u. Lehmann, 1994; Dahanukar u. Wharton, 1996; Zaessinger *et al.*, 2006; Jeske *et al.*, 2011). Bicoid aktiviert am anterioren Pol die Expression des *hunchback*-Gens, während Nanos am

posterioren Pol die Translation der *hunchback*-mRNA reprimiert (Wharton u. Struhl, 1991). Darauf folgen weitere Regulationsmechanismen, die z.B. zur Segmentierung des Abdomens führen. Während die späteren Mechanismen abhängig von zygotischen Genen sind, beruht die A/P-Achsenbildung auf maternalen Faktoren.

Die Fehlregulation des *nanos*-Gens führt zu schweren Defekten in der A/P-Achsenbildung. Wird das Nanos-Protein am anterioren Pol exprimiert, entwickeln die Embryonen zwei Abdomen und keine Kopfstrukturen (Gavis u. Lehmann, 1992). Die Translation der *nanos*-mRNA muss daher sehr strikt reguliert sein, um eine ektopische Expression von Nanos zu verhindern.

1.2. Regulation der *nanos*-mRNA während der frühen Entwicklung von *Drosophila*

Die *nanos*-mRNA wird während der frühen Entwicklung durch verschiedenen Mechanismen reguliert. Während der Oogenese wird *nanos* von Nährzellen synthetisiert und in die Oozyte transportiert.

Das Nanos-Protein kann in Nährzellen und während der frühen Oogenese detektiert werden und ist für die Erhaltung der Stammzellen notwendig (Lehmann u. Nüsslein-Volhard, 1991). Nanos ist ein konservierter Translationsrepressor, der die Expression von Differenzierungsfaktoren wie z.B. dem Brat-Protein (Brain Tumor) in der Keimbahn verhindert (Slaidina u. Lehmann, 2014). In der Oozyte kann das Nanos-Protein zu keiner Zeit detektiert werden (Wang *et al.*, 1994). Im frühen Embryo wird die Translation der *nanos*-mRNA am posterioren Pol aktiviert und ist im Polplasma (siehe Abb. 1) wichtig für die Bildung der Keimbahn der nächsten Generation. RNA, die nicht am posterioren Pol lokalisiert (unlokalisiert) ist, wird reprimiert (Smibert *et al.*, 1996). Die Regulation dieser komplexen Lokalisierung und Translationsregulation wird vom *nanos*-3'-UTR vermittelt.

Im 3'-UTR der *nanos*-mRNA kommen ein Lokalisierungssignal, ein Translationskontrollelement (TCE) und Bindestellen für regulatorische RNAs vor (Abb. 4A/B). Das TCE (Nukleotide 1 - 158) liegt innerhalb des Lokalisierungssignals (Nukleotide 1 - 547). Die Lokalisierung ist allerdings nicht abhängig von der Aktivität des TCE. Im TCE kann eine RNA-Struktur aus drei *stem-loops* gebildet werden (I-III). Die *loop*-Regionen von *stem-loop* II und einem weiteren *stem-loop* in 3'-Richtung werden vom Protein Smaug (Smg) gebunden (Abb. 4C, rot). Diese *stem-loops* werden als *Smaug recognition elements* (SRE) bezeichnet. Der doppelsträngige Bereich in *stem-loop* III wird vom Protein Glorund (Glo) gebunden (grün). Glorund vermittelt die Repression der *nanos*-mRNA während der Oogenese in der Oozyte. Smaug reguliert anschließend die Translation und die Deadenylierung der *nanos*-mRNA im Embryo.

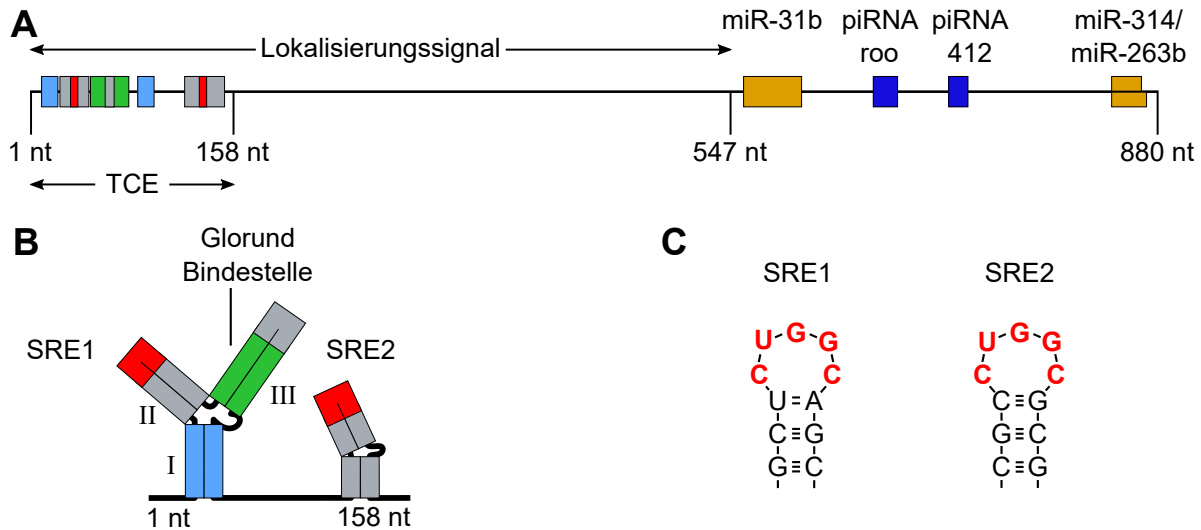


Abbildung 4: Regulatorische Sequenzen im 3'-UTR der *nanos*-mRNA - (A) Das gesamte 3'-UTR ist schematisch dargestellt und beinhaltet das Translationskontrollelement (TCE; 1 - 158), das Lokalisierungssignal (1 - 547), zwei Bindestellen für piRNAs (659 - 732) und drei potentielle Bindestellen für miRNAs (miRBase-Vorhersage in Rouget *et al.*, 2010). (B) Schematische Darstellung des TCE mit RNA-Sekundärstrukturen. Das TCE enthält zwei Bindestellen für Smaug (SRE1 und SRE2, rot) und eine Bindestelle für Glorund (grün). (C) Die beiden *Smaug recognition elements* (SRE) weisen die gleiche *loop*-Sequenz, CUGGC, auf. Die *stem*-Regionen unterscheiden sich dahingegen in ihrer Sequenz und sind nur teilweise dargestellt (Gavis *et al.*, 1996b,a; Smibert *et al.*, 1996; Kalifa *et al.*, 2006; Rouget *et al.*, 2010).

Die Smaug-vermittelte Deadenylierung und Translationsrepression agieren unabhängig voneinander und wirken auf den unlokalisierten Teil der *nanos*-mRNA (Dahanukar u. Wharton, 1996; Jeske *et al.*, 2006). Smaug ist in der Lage, den Ccr4-Not-Komplex zur RNA zu rekrutieren und damit die Deadenylierung der gebundenen RNA einzuleiten (Semotok *et al.*, 2005). Die Deletion von piRNA-Bindestellen abseits des TCEs verhindert die Deadenylierung. Dieses und weitere Experimente deuten an, dass die Deadenylierung der *nanos*-mRNA nicht ausschließlich von Smaug reguliert wird, sondern zusätzlich auch von piRNAs abhängig ist (Rouget *et al.*, 2010). Die Rolle der vorhergesagten miRNA-Bindestellen in der *nanos*-mRNA ist nicht untersucht (Abb. 4A).

Die Bindung von Smaug verhindert die Initiation der Translation der gebundenen RNA. Smaug interagiert mit dem eIF4E-bindenden Protein Cup, wodurch eine direkte Verbindung von Smaug zu einem essentiellen Translationsinitiationsfaktor hergestellt ist (Nelson *et al.*, 2004). Nach dem vorgeschlagenen Modell verhindert Cup durch seine Bindung an eIF4E die Bildung des eIF4F-Komplexes, infolge dessen die Bildung des 48S-Komplexes und schließlich die Translation der RNA (Abb. 5).

Allerdings konnte in Gradientenzentrifugationen eine Kosedimentation der *nanos*-mRNA mit Polysomen beobachtet werden (Clark *et al.*, 2000). Eine solche Beobachtung lässt üblicherweise auf eine aktive Translation der mRNA schließen, was für die *nanos*-mRNA nicht zutreffen kann. Die Autoren postulieren einen zusätzlichen Mechanismus der Repression, der an einem Schritt nach der Translationsinitiation möglicherweise kotranslational wirkt.

Es wurde zudem beschrieben, dass die Repression von Smaug auf der Assoziation mit dem Argonaut-Protein Ago1 beruht (Pinder u. Smibert, 2012). Das Argonaut-1-Protein ist essentiell für die miRNA-abhängige Regulation der Translation und Deadenylierung und ist diejenige Untereinheit des RISC-Komplexes, welche die Interaktion zwischen miRNA und Ziel-mRNA ermöglicht (Meister u. Tuschl, 2004). Im Fall der *nanos*-mRNA wird Ago1 ungewöhnlicherweise unabhängig von miRNAs durch Smaug an die *nanos*-mRNA rekrutiert. Ein Ago1-Deletion bewirkt eine ektopische Expression von Nanos, ähnlich wie in einer *smaug*-Mutante. Der Abbau der *nanos*-mRNA während der MZT ist in einer Ago1-Mutante allerdings nicht beeinflusst (Pinder u. Smibert, 2012).

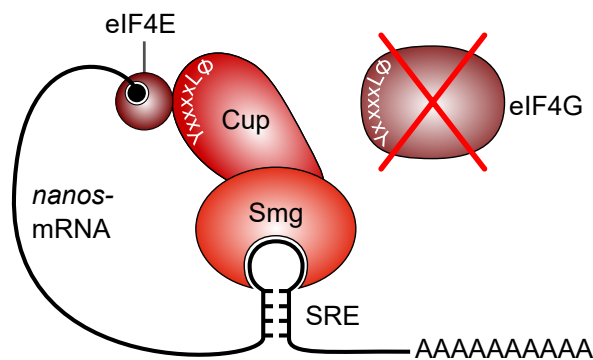


Abbildung 5: Cup-Modell der Smaug-abhängigen Repression - Smaug bindet spezifisch an SREs im 3'UTR der *nanos*-mRNA und bindet an das Cup-Protein. Cup wiederum bindet an das Cap-bindende Protein eIF4E. Cup verdrängt eIF4G, da es dieselbe eIF4E-Bindestelle wie eIF4G aufweist (YxxxxLΦ; wobei x einer beliebigen Aminosäure und Φ einer hydrophoben Aminosäure entspricht; Marcotrigiano *et al.*, 1999). Dadurch verhindert Cup die Bildung des eIF4F-Komplexes, der durch Bindung von eIF4A und eIF3 die Translation der gebundenen RNA initiiert. Der Komplex aus Smaug-Cup-eIF4E ist nach diesem Modell in der Lage, Translation zu reprimieren (Nelson *et al.*, 2004).

Am posterioren Pol wird die Smaug-abhängige Regulation, sowohl die Deadenylierung als auch die Translationsrepression, durch Oskar aufgehoben (Smith *et al.*, 1992; Dahanukar *et al.*, 1999; Zaessinger *et al.*, 2006). Der dem zugrunde liegende Mechanismus ist noch unklar, aber es ist beschrieben, dass Smaug und Oskar miteinander interagieren (Dahanukar *et al.*, 1999). Oskar ist zudem in der Lage, die *nanos*-mRNA zu binden (Jeske *et al.*, 2015), so dass eine Konkurrenz von Oskar mit der Bindung von Smaug an die RNA denkbar ist.

Weder die Rekrutierung von Ago1, noch das Cup-Modell sind ausreichend, um alle Beobachtungen zur Repression hinreichend zu beschreiben. Die Smaug-vermittelte Repression hat verschiedene ungewöhnliche Eigenschaften, die nur schwer mit den bisherigen Modellen erklärt werden können (Jeske *et al.*, 2011):

- **Langsame Bildung eines reprimierten RNA-Protein-Komplexes** - Es dauert etwa 20 - 40 Minuten, bis das volle Ausmaß der Repression erreicht ist. Die RNA wird in einen Komplex aus RNA und Protein inkorporiert (RNP - Ribonukleoprotein Partikel). Der assoziierte Proteinkomplex, der die Repression der RNA hervorruft wird im Folgenden als Repressor-Komplex beschrieben.
- **ATP-Abhängigkeit der Repression** - Die Bildung dieses Repressor-Komplexes ist abhängig von der Anwesenheit von ATP. Dies legt die Beteiligung eines ATP-bindenden Proteins nahe. Weder Smaug, Cup, Ago1 noch eIF4E binden ATP.
- **ungewöhnliche kinetische Stabilität** - Die Halbwertszeit der Dissoziation des Komplexes liegt im Bereich von 3 - 4 Stunden. Bei einer Dissoziationsratenkonstante $k_{off} \approx 5 \cdot 10^{-5} \frac{1}{s}$ ($t_{1/2} \approx 4$ h) und einer maximalen Assoziationsrate von $k_{on} \approx 10^{10} \frac{1}{Ms}$ (Limit durch Diffusion von Makromolekülen; Berg u. von Hippel, 1985) müsste eine Dissoziationskonstante von 5 fM für eine einfache Protein-Ligand-Interaktion angenommen werden. Die hohe Stabilität ist demnach nicht durch ein einfaches Dissoziationsgleichgewicht zu erklären.
- **Cap-Unabhängigkeit** - Das Cup-Modell sagt eine Abhängigkeit der Repression vom 5'-Cap voraus. Die Repression kann allerdings auch mit RNAs ohne Cap und mit einer internen Ribosomenbindestelle (IRES) beobachtet werden. Verschiedene Mechanismen müssen demnach zur Regulation der Translation beitragen.

1.3. Komponenten des reprimierten RNPs

Einige Proteine, die an der Repression beteiligt sind, konnten bereits durch *pull-down*-Experimente nachgewiesen werden (Nelson *et al.*, 2004; Jeske *et al.*, 2011; Pinder u. Smibert, 2012). Dazu zählt Smaug, als die zentrale Komponente des Repressor-Komplexes, sowie die Proteine Cup, Me31B und Trailer Hitch. Die Verläufe der mRNA-Expressionslevel während der Entwicklung von *Drosophila* sind für diese vier Proteine sehr ähnlich. Die höchsten Expressionslevel treten während der ersten vier Stunden der Embryonalentwicklung auf (Fly-Base.org, modENCODE - Gelbart u. Emmert, 2013). Danach ist die RNA-Menge deutlich reduziert. Die Proteinexpression folgt einem ähnlichen Schema: alle Proteine sind spätestens

im Embryo exprimiert und werden spezifisch während der MZT degradiert (Gouw *et al.*, 2009 mit einem Vergleich der Proteinmengen vor und nach der MZT). Die möglichen Funktionen der bekannten Komponenten Smaug, Cup, Me31B, Trailer Hitch und dem Deadenylase-Komplex sollen hier kurz eingeleitet werden.

1.3.1. Smaug

Smaug (Smg) ist ein RNA-bindendes Protein mit 999 Aminosäuren. Es wurde nach dem Drachen aus J. R. R. Tolkiens Hobbit benannt, der den Schatz der Zwerge (*nanos*) bewacht. Die *smaug*-mRNA wird während der Oogenese synthetisiert und in der Oozyte deponiert. Die Synthese des Smaug-Proteins wird erst während der Ei-Aktivierung durch die PAN GU-Kinase initiiert (Tadros *et al.*, 2007). Smaug kommt dann im gesamten Embryo vor. Nach der Zellularisierung kann Smaug kaum noch detektiert werden (Dahanukar *et al.*, 1999; Benoit *et al.*, 2009).

Mutationen im *smaug*-Gen sind bereits früh in der Embryogenese letal. Die Embryonen sterben vor dem Stadium des zellulären Blastoderms aufgrund von irregulären Kernteilungen ab der 10. synzytialen Teilung. Dieser Phänotyp ist nicht von der Regulation der *nanos*-mRNA abhängig, da Mutationen im *nanos*-Gen erst später zu Defekten in der Abdominalentwicklung führen (Dahanukar *et al.*, 1999). Smaug ist zusätzlich zur *nanos*-Regulation essentiell für die Aktivierung des zygotischen Genoms und zur Aktivierung von Replikationskontrollpunkten während der MZT (Benoit *et al.*, 2009). Etwa zwei Drittel der maternalen mRNAs, die während der MZT degradiert werden, werden direkt oder indirekt durch Smaug destabilisiert (Tadros *et al.*, 2007).

Smaug wurde mittels UV-Cross-Linking an SRE-haltige RNAs identifiziert (Smibert *et al.*, 1996). Smaug weist eine ungewöhnliche RNA-Bindedomäne auf: eine SAM-Domäne (*sterile alpha motif*), die normalerweise Protein-Protein-Interaktionen bewirkt, ist für die sequenz- und strukturabhängige Bindung der SRE-*stem-loops* nötig (Aviv *et al.*, 2003; Green *et al.*, 2003). Smaug erkennt in der *nanos*-mRNA die *loop*-Regionen zweier *stem-loops* mit der *loop*-Sequenz CUGGC (siehe Abb. 4C). Die für die Bindung benötigte Sequenz muss nicht exakt der Sequenz CUGGC entsprechen. Smaug kann *loops* mit einer Länge von 4 - 7 Nukleotiden binden. Das erste Nukleotid muss ein Pyrimidin-Nukleotid sein, dass mit dem 4. Nukleotid eine Watson-Crick Basenpaarung eingeht. Das dritte Nukleotid muss ein Guanosin sein (Konsensussequenzen: CNGG(N)₀₋₃ oder UNGA(N)₁₋₃; Aviv *et al.*, 2006).

Diese *Stem-loop*-Sequenz kann in zahlreichen 3'-UTRs von mRNAs gefunden werden, so dass Smaug eine große Anzahl an RNAs regulieren kann (bis zu 1600 in Tadros *et al.*, 2007). Schätzungen belaufen sich auf bis zu 3000 mRNAs, die von Smaug reguliert werden (Chen *et al.*, 2014a). Ob alle identifizierten Ziel-mRNAs auch wirklich direkt von Smaug

reguliert werden, bleibt bislang offen. Bei einer unabhängigen Suche nach SRE-*stem-loops* in 3'-UTRs von *Drosophila melanogaster* wurden etwa 1800 verschiedene Transkripte, die wenigstens ein mögliches SRE aufweisen, identifiziert², darunter auch die *nanos*-mRNA.

Ein anderes von Smaug reguliertes Transkript ist die *hsp83*-mRNA. Im 3'-UTR des *hsp83*-Transkriptes kann jedoch kein SRE identifiziert werden. Dafür enthält die kodierende Sequenz insgesamt acht mögliche SRE-*stem-loops*, von denen sieben ebenfalls in der unabhängigen Sequenzanalyse (siehe Abschnitt M.2.12, S. 138) identifiziert wurden. Ein weiterer Unterschied zu *nanos* ist, dass die Translation der *hsp83*-mRNA nicht durch die SREs beeinflusst wird. Smaug bewirkt lediglich den Abbau der *hsp83*-mRNA – wie auch für *nanos* – durch die Deadenylierung des Transkriptes und die Rekrutierung des Ccr4-Not-Komplexes (Semotok *et al.*, 2005, 2008). Dieses Beispiel zeigt, dass Smaug verschiedene Mechanismen auslösen kann.

1.3.2. Cup

Das Cup-Protein wird bereits in der Oogenese benötigt und ist als Translationsrepressor beschrieben. Cup kommt im Zytoplasma der Oozyte und des Embryos vor. Nach der Zellularisierung (Teilung 14) ist Cup, ähnlich wie Smaug, im Embryo nicht länger zu detektieren (Keyes u. Spradling, 1997). Cup wurde als eine Komponente in reprimierten RNPs der *oskar*-mRNA während der Oogenese beschrieben. Es wird von Bruno, einem regulatorischen Protein, welches im 3'-UTR der *oskar*-mRNA bindet, an die *oskar*-mRNA rekrutiert (Nakamura *et al.*, 2004) und interagiert selbst direkt mit dem Lokalisierungsprotein Barentz sowie eIF4E. Damit trägt Cup zur Lokalisierung und Repression der *oskar*-mRNA in der Oozyte bei (Wilhelm *et al.*, 2003).

Cup weist insgesamt zwei eIF4E-Bindestellen auf: eine, wie sie für 4E-bindende Proteine (4EBPs) beschrieben wurde, mit der Konsensussequenz YxxxxLΦ (siehe Abb. 5) und eine weitere Bindestelle mit geringerer Affinität für eIF4E (Nelson *et al.*, 2004). Beide Bindestellen konkurrieren mit eIF4G um die Bindung an eIF4E. Die Proteine eIF4E und eIF4G sind Bestandteile des eIF4F-Komplexes, der essentiell für die Cap-abhängige Initiation der Translation ist. Wenn eIF4G von Cup verdrängt wird, kommt es zur Repression der jeweiligen mRNA, da die Translationsinitiation blockiert wird.

Eine Deletion des *cup*-Gens führt zu Defekten früh in der Oogenese, so dass Effekte auf die *oskar*- oder *nanos*-Regulation nicht beobachtet werden können. Bei einer N-terminalen Deletion von Cup (*cup*^{Δ212}), bei der eine der eIF4E-Bindestellen deletiert ist, treten die frühen Defekte in der Oogenese nicht auf. Die *oskar*-mRNA wird in *cup*^{Δ212}-Mutanten nicht

²Dazu wurde ein einfacher Suchalgorithmus programmiert (SREfinder), der *stem-loops* mit der beschriebenen *loop*-Sequenz in einer Sequenzdatenbank identifiziert; siehe Abschnitt M.2.12, S. 138.

reprimiert, obwohl die RNA noch korrekt lokalisiert wird. Cup hat demnach neben der eIF4E-abhängigen Translationsrepression noch weitere Funktionen (Nakamura *et al.*, 2004). In *Drosophila* existiert ein paraloges Protein: der eIF4E-Transporter (4E-T). Beide Proteine besitzen eine *Cup-homology-domain* (CHD; Kamenska *et al.*, 2014), die vermutlich für die Bindung an Me31B benötigt wird (Ozgur *et al.*, 2015; Nishimura *et al.*, 2015; Kamenska *et al.*, 2016). Das 4E-T-Protein ist gut konserviert, während das Cup-Protein ein *Drosophila*-spezifisches Protein ist.

1.3.3. Me31B

Me31B (*maternally expressed at 31B*) ist ebenso ein maternales Protein. Es gehört zur Klasse der DEAD-Box-Helikasen (de Valoir *et al.*, 1991). DEAD-Box-Helikasen sind ATP-abhängige RNA-Helikasen, die für viele verschiedene RNA-abhängige Prozesse in der Zelle benötigt werden. Die Helikaseaktivität ist dabei nicht als prozessive Aufschmelzung eines langen RNA-Bereichs zu verstehen, sondern eher als Dissoziation kurzer RNA-Doppelstränge, ohne eine Translokation der Helikase nach der Aufschmelzung (Yang *et al.*, 2007), auch wenn durch die Bindung von Kofaktoren eine prozessive Aufschmelzung erreicht werden kann (García-García *et al.*, 2015). Der Auftrennung eines RNA-Doppelstranges geht die Bindung von ATP voraus. Die Hydrolyse des ATPs ist für die Dissoziation des Enzyms von der RNA nötig (Liu *et al.*, 2008). DEAD-Box-Enzyme weisen keine Sequenzspezifität für ihre Substrate auf und müssen daher durch Kofaktoren zur Substrat-RNA rekrutiert werden (Jankowsky, 2011).

Ein außergewöhnliches Beispiel für DEAD-Box-Helikasen stellt eIF4AIII dar. Während des Spleißens von prä-mRNAs wird ein stabiler Komplex in der Nähe der Exon-Exon-Grenzen deponiert. Dieser Proteinkomplex (EJC - *exon junction complex*) ist fest an die RNA gebunden und vermittelt z.B. den Export von mRNAs sowie den Abbau aberanter mRNAs (NMD - *nonsense mediated decay*; Gehring *et al.*, 2005). Die stabile Bindung des EJC beruht auf der RNA-Bindung von eIF4AIII (Ballut *et al.*, 2005). Die DEAD-Box-Helikase eIF4AIII wird im RNA-gebundenen Zustand durch die Interaktion mit den Proteinen Y14 und MAGOH „eingefroren“ (Andersen *et al.*, 2006; Bono *et al.*, 2006). Diese feste Bindung wird auch als *RNA-clamping* bezeichnet (Linder u. Jankowsky, 2011) und wird auch für andere DEAD-Box-Helikasen als mögliche Funktionsweise postuliert (Liu *et al.*, 2014). Damit sind DEAD-Box-Helikasen naheliegende Kandidaten, um sowohl die ATP-Abhängigkeit der Komplexbildung, als auch die hohe kinetische Stabilität des Repressor-Komplexes zu vermitteln.

Für die Homologe von Me31B konnten verschiedene Funktionen in der RNA-Prozessierung gezeigt werden. Dhh1p (*Saccharomyces cerevisiae*) wurde als *decapping*-Aktivator beschrieben und interagiert ebenso mit dem Ccr4-Not-Komplex (Coller *et al.*, 2001). Vereint werden

die Me31B-Homologe durch ihre Funktion in der Repression der Translation (Minshall *et al.*, 2001; Collier u. Parker, 2005; Tritschler *et al.*, 2009; Mathys *et al.*, 2014). Me31B und den homologen Proteinen ist demnach eine zentrale Rolle in der Stilllegung von mRNAs, ob nun durch Repression oder Abbau, zuzuschreiben. Die Homologe werden in Zellen häufig in granulären Strukturen, wie *P-bodies*, gefunden und deshalb auch als *P-body*-Marker verwendet. *P-bodies* enthalten Faktoren des mRNA-Abbaus (*decapping*-Faktoren und Deadenylassen) sowie verschiedene RNAs, die translationell reprimiert sind (Parker u. Sheth, 2007), in Übereinstimmung mit einer Rolle von Me31B in der Repression von mRNAs.

Nakamura *et al.* (2001) zeigten, dass Me31B bereits in *Drosophila*-Oozyten mit translationell reprimierten RNAs assoziiert. Me31B spielt eine essentielle Rolle in der Repression der *oskar*-mRNA. In der frühen Oogenese von *Drosophila* kolokalisiert Me31B mit Cup und der *oskar*-mRNA (Nakamura *et al.*, 2004). Eine *me31B*-Deletion führt zur frühzeitigen Synthese von Oskar und damit zu schweren Defekten, noch bevor die Oogenese beendet ist. Die Lokalisierung der *oskar*-mRNA ist in einer *me31B*-Mutante nicht beeinflusst, so dass eine Rolle in der Lokalisierung reprimierter RNAs unwahrscheinlich ist (Nakamura *et al.*, 2001). Im adulten Organismus kommt Me31B hauptsächlich in den Ovarien der Weibchen vor, aber z.B. auch im zentralen Nervensystem, wo Me31B ebenfalls an der translationalen Repression von mRNAs beteiligt ist (Hillebrand *et al.*, 2010).

Me31B interagiert mit verschiedenen Proteinen, die teilweise an posttranskriptionalen Regulationen beteiligt sind. Einige der Protein-Protein-Interaktionen sind ebenfalls konserviert und die Interaktionsstellen kartiert. Es konnte z.B. gezeigt werden, dass Me31B mit Edc3 (*enhancer of decapping*) über die C-terminale Domäne interagiert (Tritschler *et al.*, 2008). Dieselbe Bindestelle (*patch1*; Ozgur *et al.*, 2015) dient auch als Bindestelle für die Proteine Trailer Hitch (Tral), Cup, 4E-T und Pat1, die in Translationsrepression und RNA-Abbau involviert sind (Abb. 6). Demnach kann auch nur eines der Proteine an Me31B gebunden vorliegen (Tritschler *et al.*, 2009). Das Me31B-Homolog, DDX6, interagiert ebenfalls direkt mit der zentralen Komponente des Ccr4-Not-Komplexes, Not1, über eine Bindestelle abseits von *patch1* (Abb. 6; Chen *et al.*, 2014b; Mathys *et al.*, 2014; Rouya *et al.*, 2014).

Die spezifische Assoziation von Me31B mit der *nanos*-mRNA, die vielfältigen Interaktionspartner und das Potential von DEAD-Box-Helikasen, als Translationsrepressoren zu dienen, erlauben den Schluss, dass Me31B im Zentrum der *nanos*-Repression steht. Es ist jedoch nicht klar, welche exakte Rolle Me31B in dem Repressor-Komplex übernimmt.

1.3.4. Trailer Hitch

Trailer Hitch (Tral) gehört ebenfalls zu einer Gruppe konservierter Proteine, von denen RAP55 (*RNA interacting protein of 55 kDa*) als erstes in *Xenopus* beschrieben wurde

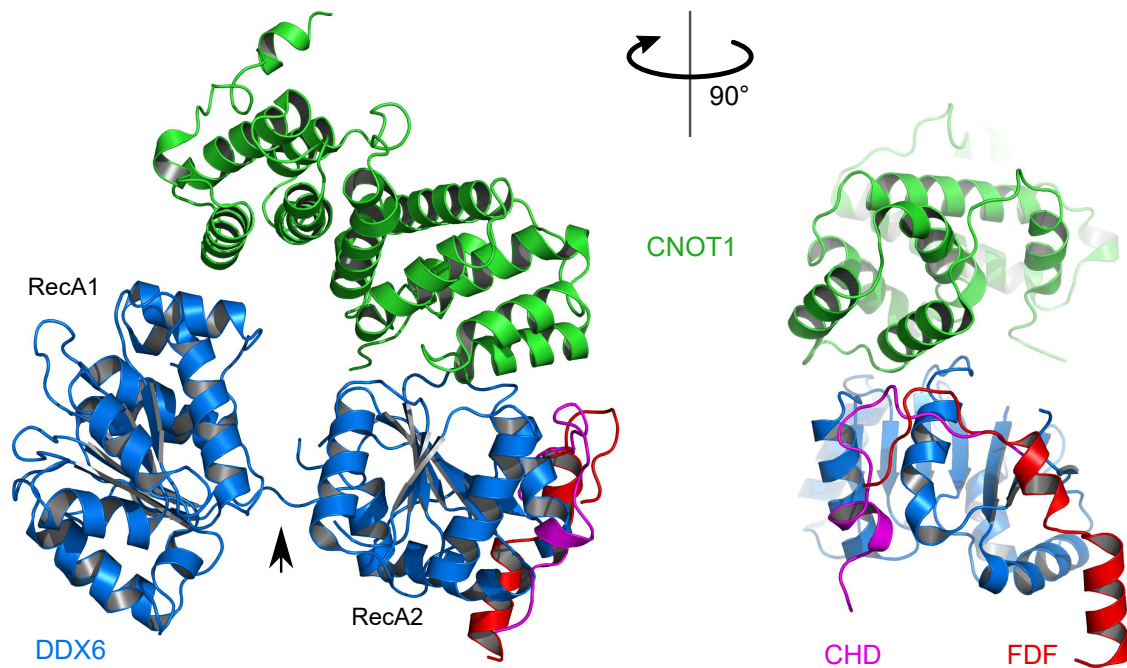


Abbildung 6: Strukturmodell der DDX6/Me31B-Interaktionen - Das Me31B Homolog (DDX6 - blau) nimmt eine für DEAD-Box-Helikasen übliche Struktur an. Die beiden RecA Domänen werden durch ein Gelenk verbunden (Pfeil links), das eine Bewegung der beiden Domänen zueinander zulässt. Um die Interaktionen zu verschiedenen Proteinen gleichzeitig darzustellen, wurden zwei Kristallstrukturen überlagert (PDB IDs: 5ANR - Ozgur *et al.*, 2015 und 2WAX - Tritschler *et al.*, 2009). Das Not1-Protein (zentrales Protein des Ccr4-Not-Komplexes) ist grün dargestellt, die *cup-homology-domain* (CHD) von 4E-T in Magenta und das FDF-Peptid von Edc3 in rot. Das FDF-Peptid und die CHD binden an dieselbe Oberfläche im C-Terminus (*patch1*; Ozgur *et al.*, 2015), während Not1 an beide RecA-Domänen abseits der FDF-Bindestelle bindet.

(Lieb *et al.*, 1998). Alle Homologe weisen ein N-terminale LSm14-Domäne auf, auf die eine Serin-/Threoninreiche Region folgt. Weiterhin enthalten die Homologe ein FDF-Peptid, eine FFD-TFG Region und RGG-Wiederholungen. Die N-terminale LSm-Domäne ist in Protein-Protein-Interaktionen involviert (Tritschler *et al.*, 2008). Das FDF-Motiv wird für die Interaktion mit Me31B benötigt und ist strukturell und funktionell mit der FFD-TFG-Region verbunden. Die RGG-Wiederholungen werden sowohl für die Repression als auch für die Lokalisierung in *P-bodies* benötigt (Marnef *et al.*, 2009; Nissan *et al.*, 2010). Während im Menschen, in *Xenopus* und in der Maus zwei Paraloge von RAP55 vorkommen (RAP55a und RAP55b), gibt es in *Drosophila* nur eines.

Trailer Hitch und seine Homologe wurden mehrfach als Translationsregulatoren beschrieben. Scd6p, das Hefe-Homologe, reprimiert die Translation durch Blockierung der 48S-Komplexbildung (Rajyaguru *et al.*, 2012), wie sie für die *nanos*-mRNA gezeigt wurde (Jeske *et al.*, 2011). Diese Repression ist abhängig von den RGG-Wiederholungen

(Nissan *et al.*, 2010). Wird das *Xenopus*-Ortholog xRAP55 an eine RNA gekoppelt, kommt es zu einer translationalen Repression (Tanaka *et al.*, 2006). In humanen Zellkulturen werden 4E-T, DDX6 und hRAP55 für die Bildung und Erhaltung von *P-bodies* benötigt (Ayache *et al.*, 2015). Im Gegensatz dazu zeigt eine Depletion von Trailer Hitch in *Drosophila*-S2-Zellen keinen Einfluss auf *P-bodies*, obwohl es deutlich in *P-bodies* angereichert ist (Eulalio *et al.*, 2007). Neben seiner Funktion in *P-bodies* zeigt Trailer Hitch einen Einfluss auf die Sekretion verschiedener Proteine am endoplasmatischen Retikulum. Der Komplex aus Me31B und Trailer Hitch reguliert beispielsweise zusammen mit Bic-C die Sekretion von Gurken (Kugler u. Lasko, 2009).

Die Rolle von Trailer Hitch im Repressor-Komplex ist unklar. Aber auch für Tral deutet die Anreicherung im reprimierten RNP (Jeske *et al.*, 2011), die Interaktion mit dem Repressor Me31B sowie die Repressionsfunktion der homologen Proteine auf eine Beteiligung an der SRE-vermittelten Repression hin.

1.4. Zielstellung

Ziel dieser Arbeit war es, den Mechanismus der posttranskriptionalen Regulation der *nanos*-mRNA während der frühen Embryogenese von *Drosophila* besser zu verstehen. Die untersuchte Regulation ist essentiell für die Embryogenese, da Manipulationen beteiligter Faktoren häufig entweder zur Letalität oder Unfruchtbarkeit führen. Die Regulation der *nanos*-mRNA ist vielschichtig und ist sowohl vom Abbau der mRNA durch Deadenylierung als auch von translationaler Kontrolle abhängig. Verschiedene Modelle sind postuliert, von denen keines die beobachteten Eigenschaften der Regulation vollständig erklären kann. Die Modelle sind zum Teil auch gegensätzlich aufzufassen. So wurde gezeigt, dass die Smaug-Bindung im 3'-UTR ausreicht, um die Deadenylierung einzuleiten (Semotok *et al.*, 2005), doch *in vivo* werden zusätzlich piRNAs für die Deadenylierung der *nanos*-mRNA benötigt (Rouget *et al.*, 2010). Es wurde postuliert, dass die Smaug-abhängige Repression entweder durch Interaktion mit Cup oder Ago1 vermittelt wird (Nelson *et al.*, 2004; Pinder u. Smibert, 2012). Die Rolle anderer Bestandteile des reprimierten RNP-Komplexes ist noch unklar. Ein weiteres Modell postuliert eine kotranslationale Repression der Translation (Clark *et al.*, 2000).

Um die beteiligten Mechanismen besser verstehen zu können, sollte in dieser Arbeit eine detaillierte Analyse des reprimierten RNA-Protein-Komplexes durchgeführt werden. Die Zusammensetzung des Komplexes sollte dazu mittels Massenspektrometrie analysiert und die Beteiligung neu identifizierte Komponenten validiert werden. Interessant wäre es dann, die Rolle der einzelnen Proteine und den Aufbau des Komplexes im Detail zu untersuchen.

Dazu sollte versucht werden, in einem *Cross-Linking*-Experiment die Interaktionen zwischen den einzelnen Untereinheiten des Smaug-RNPs zu charakterisieren.

Eine exzellente Möglichkeit, den Mechanismus der Repression zu untersuchen, stellt die *In-vitro*-Rekonstitution des Systems dar. In einem rekonstituierten System lässt sich vergleichbar einfach die Rolle einzelner Proteine überprüfen. So kann z.B. auch die Relevanz einzelner Domänen oder gar einzelner Aminosäuren der beteiligten Proteine durch den Austausch mit mutierten Varianten getestet werden. Das Wissen über die Zusammensetzung und die Stöchiometrie der Komplexkomponenten zueinander ist dabei äußerst wichtig, so dass diese Arbeit einen signifikanten Beitrag zur Rekonstitution der Smaug-abhängigen Repression und Deadenylierung liefern soll.

2. Ergebnisse

An SRE-haltigen mRNAs assembliert ein Proteinkomplex, der die translationale Repression der mRNA hervorruft. Um die Beteiligung einzelner Faktoren an der Repression aufzuklären, wäre die Rekonstitution des Repressor-Komplexes aus seinen Bestandteilen sehr aussichtsreich. Die genaue Zusammensetzung sowie die Funktionsweise aller Bestandteile im Komplex sind allerdings bis dato nicht bekannt und damit ist eine Rekonstitution nur schwer zu bewerkstelligen.

Die Identifizierung der beteiligten Faktoren sollte in dieser Arbeit durch die Aufreinigung des Komplexes aus *Drosophila*-Embryoextrakt (DEE) und anschließende massenspektrometrische Analyse erfolgen. Der zentrale Bestandteil des Komplexes, Smaug, bindet an einen *stem-loop* in der RNA und rekrutiert daraufhin direkt oder indirekt alle restlichen Komponenten an die gebundene RNA. Diese Komplexbildung kann im Extrakt nachvollzogen werden. *In-vitro*-synthetisierte RNAs, die ein SRE enthalten, werden in *Drosophila*-Embryoextrakt sowohl spezifisch deadenyliert als auch translationell reprimiert. Daraus ergibt sich die Möglichkeit, den Komplex, der an einer extern zugegebenen RNA ausgebildet wurde, mittels Affinitätsreinigung der RNA zu präparieren. Für eine sehr spezifische Anreicherung wurde dazu *in-vitro*-synthetisierte RNA verwendet, die an zufälligen Positionen Biotin trägt (*Bait*-RNA). Biotin wird mit einem K_d von etwa 10^{-15} M von Streptavidin gebunden (Green, 1963, 1966). Dies ist die stärkste nicht-kovalente Bindung, die bekannt ist.

2.1. Reinigung des Repressor-Komplexes - Vorüberlegungen

Verschiedene Aspekte müssen bei der Reinigung aus *Drosophila*-Embryoextrakt beachtet werden. Die Kapazität des Extraktes, eine RNA zu reprimieren, ist optimal im niedrigen nanomolaren Bereich (Jeske *et al.*, 2011). Es kann also nur eine geringe Menge der biotinylierten RNA zugegeben werden, im Bereich bis etwa 10 nM. Wenn Smaug an die RNA bindet, wird sowohl die Translationsrepression als auch die Deadenylierung dieser RNA eingeleitet. Für die Reinigung wurden RNAs mit einem internen Poly(A)-Schwanz verwendet (Jeske *et al.*, 2011). Die Translation wird von einem internen Poly(A)-Schwanz stimuliert, dieser kann allerdings nicht von Deadenylassen degradiert werden, so dass die Transkripte nicht SRE-spezifisch destabilisiert werden.

Weiterhin werden RNAs im Extrakt unspezifisch von RNasen degradiert, so dass eine RNA-Affinitätsreinigung erschwert wird. Nach einer halbstündigen Inkubation einer *in-vitro*-synthetisierten RNA in Embryoextrakt werden bis zu 90% der RNAs abgebaut, so dass eine Möglichkeit gefunden werden musste, die *Bait*-RNAs im Extrakt zu stabilisieren.

Die *in-vitro*-synthetisierten RNAs können sowohl SRE-spezifisch als auch -unspezifisch von Proteinen gebunden werden. Um SRE-spezifische Interaktionspartner zu identifizieren, ist eine geeignete Negativkontrolle nötig. Punktmutationen in der *loop*-Sequenz des SREs verhindern die RNA-Bindung von Smaug beinahe vollständig (Smibert *et al.*, 1996; Aviv *et al.*, 2006). Als Spezifitätskontrolle wurde daher in allen Experimenten die eingesetzte RNA mit einer RNA mit Punktmutationen in den beiden SREs verglichen. Proteine, die unspezifisch oder unabhängig vom SRE mit der RNA interagieren, sollten beide RNAs gleichmäßig binden. In dieser Arbeit werden diese RNAs mit WT und MUT, für RNAs mit Wildtyp oder mutierten SREs, abgekürzt.

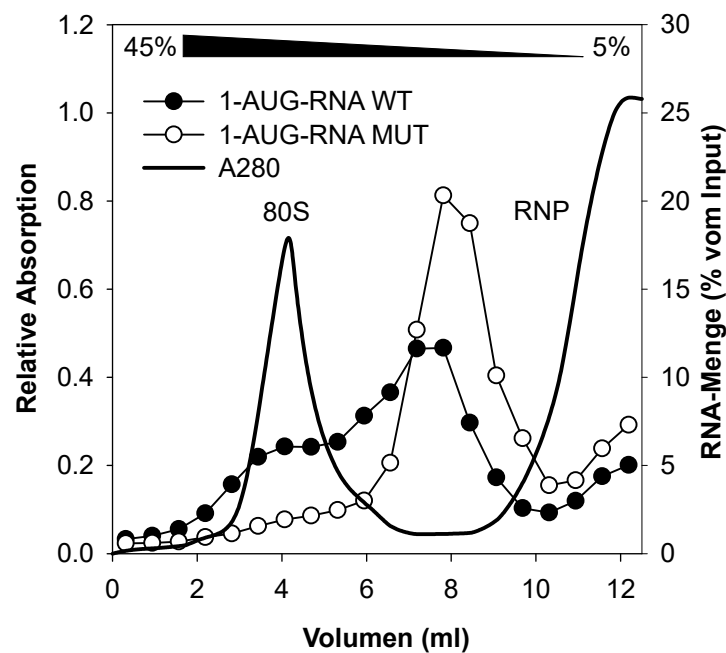


Abbildung 7: Analytische Sedimentation von 1-AUG-RNA - Radioaktiv markierte 1-AUG-RNA (10 nM; siehe Abschnitt 2.2) wurde in DEE (60%) inkubiert und 200 μ l wurden durch eine Saccharosegradientenzentrifugation (5% - 45% Saccharose) fraktioniert. Das UV-Signal (280 nm) wurde während der Fraktionierung aufgezeichnet. Die Positionen des 80S-Signals sowie der freien RNPs sind angedeutet. Die RNA-Menge je Fraktion wurde durch Szintillationszählung gemessen. RNAs mit einem Wildtyp-SRE (WT) sedimentieren deutlich schneller als RNAs mit einem mutierten SRE (MUT).

Drosophila-Embryoextrakt enthält eine Vielzahl verschiedener Proteine, von denen vermutlich nur ein kleiner Teil eine Rolle in der Regulation der *nanos*-mRNA spielt. Um die Effizienz der Reinigung zu steigern, sollte ein weiterer Reinigungsschritt durchgeführt wer-

den. Der gebildete RNP-Komplex sedimentiert in einem Saccharosegradienten von 5% - 45% ähnlich wie die sehr kompakten Ribosomen (Abb. 7). Mit der Gradientenzentrifugation als Reinigungsschritt ließen sich kleinere mRNPs sowie monomere Proteine oder kleinere Proteinkomplexe effizient abtrennen, da diese kaum in den Gradienten eindringen. Die Affinitätsreinigung mittels Biotin kann im Anschluss aus Fraktionen mit reprimierten RNAs durchgeführt werden.

2.2. Reinigung des Repressor-Komplexes

Für die Reinigung des reprimierten RNPs wurden biotinylierte *Bait*-RNAs mit einer Länge von 630 nt verwendet. Diese 1-AUG-RNAs sind Transkripte mit einem offenen Leseraster (102 Aminosäuren), das dem Luziferase-Transkript entstammt und ein einziges Startcodon (AUG) enthält. Zudem enthielten die *Bait*-RNAs das TCE des *nanos*-3'-UTRs (siehe Abb. 4) und einen internen Poly(A)-Schwanz (wie in Abb. 9A). Die Struktur der verwendeten RNA ist in Abb. 44B im Anhang auf S. XLVIII dargestellt.

2.2.1. Stabilisierung der *Bait*-RNAs

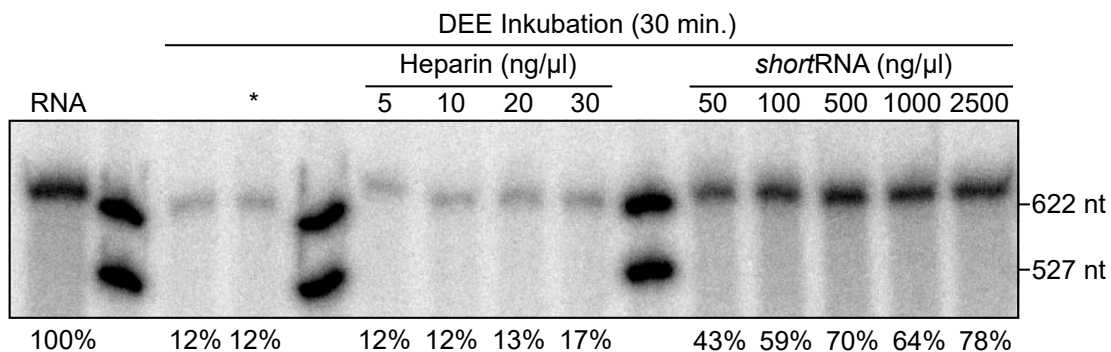


Abbildung 8: RNA-Abbau in *Drosophila*-Embryoextrakt - Radioaktiv markierte 1-AUG-RNA wurde in *Drosophila*-Embryoextrakt für 30 Minuten inkubiert und nach Proteinase K-Verdau und Ethanol-fällung mittels Harnstoff-PAGE analysiert. Die relative Menge an intakter RNA nach der Inkubation wurde durch Vergleich mit der eingesetzten RNA-Menge (Spur 1) densitometrisch bestimmt und unterhalb der Spuren notiert. Die mit * markierte Probe wurde in DEE inkubiert, der zuvor für 30 min bei 20.000 g zentrifugiert wurde. Steigende Mengen Heparin oder *short*RNAs (siehe Abschnitt M.2.2, S. 128) wurden der Inkubation zugegeben, um dem RNA-Abbau entgegenzuwirken.

In *Drosophila*-Embryoextrakten sind verschiedene RNasen enthalten, die während der Vorinkubation von 20 bis 30 Minuten zu einer beinahe vollständigen Degradation der *Bait*-RNAs führten (Abb. 8; 12% intakte RNA nach 30 min Inkubation). Die Zugabe von Heparin zeigte keine deutliche Stabilisierung der RNA. Ein Gemisch kurzer RNAs (< 30 Nukleotide [nt]; siehe Abschnitt M.2.2, S. 128) führte bereits bei einer Konzentration von 50 ng/μl zu einer deutlichen Stabilisierung. Nach Zugabe von 100 ng/μl *short*RNAs waren noch etwa 60%

der RNA nach einer 30-minütigen Inkubation intakt. In allen weiteren Reaktionen wurden *short*RNAs mit einer Konzentration von 200 ng/ μ l zugegeben. Dadurch konnten die Ausbeuten der späteren Präzipitationen der RNP-Komplexe um einen Faktor von etwa sechs erhöht werden. Die Zugabe von *short*RNAs zu Embryoextrakten erhöhte deren Translationseffizienz für Luziferase-RNAs und hatte nur einen geringen Einfluss auf die SRE-spezifische Translationsrepression (siehe Abb. 38 im Anhang S. XLI).

2.2.2. Reinigung des Repressor-Komplexes durch Gradientenzentrifugation

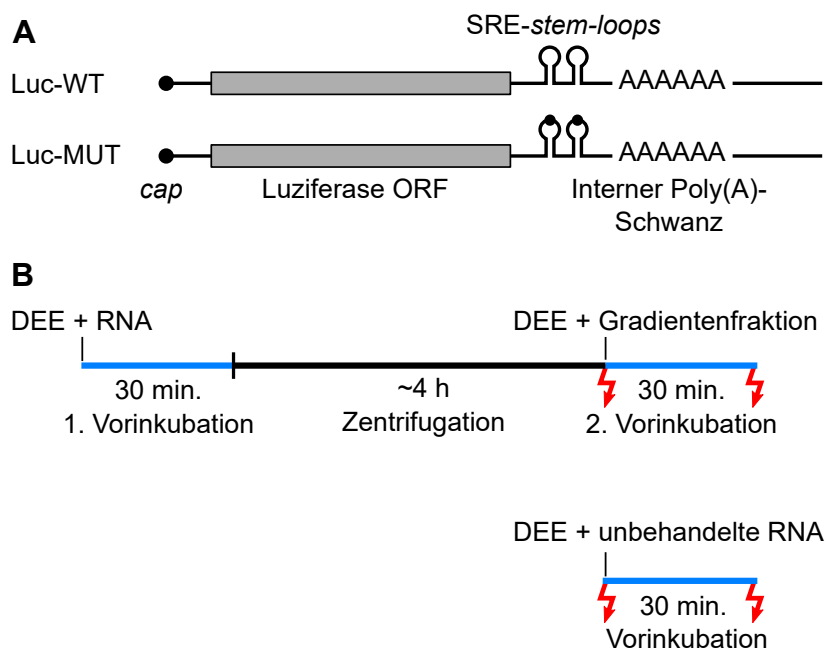


Abbildung 9: Experimentierschema: Stabilität der Repression - (A) Schema der verwendeten RNAs: Luziferase-RNAs mit den ersten 158 nt des *nanos*-3'-UTRs und einem internen Poly(A)-Schwanz. RNAs mit Mutationen in den beiden SRE-*stem-loops* (MUT) werden in DEE nicht reprimiert. (B) Schema des Experiments: RNA wird in Extrakt inkubiert und anschließend mittels Gradientenzentrifugation fraktioniert. Der Einfluss einer zweiten Vorinkubation auf die Translationsrepression wird durch Translationsexperimente (Blitze) getestet, indem RNAs mit oder ohne vorherige Inkubation und Zentrifugation verglichen werden.

Die Gradientenzentrifugation kann als Reinigungsschritt in Betracht gezogen werden, wenn die RNA nach der Zentrifugation noch reprimiert vorliegt. Nur dann kann man sicher sein, dass keine der benötigten Komponenten des stabilen Komplexes dissoziiert sind. Um den reprimierten Zustand nachzuweisen, wurden SRE-haltige Luziferase-RNAs nach Vorinkubation im Extrakt mittels Saccharosegradienten fraktioniert, anschließend wurde die Translationsrepression gemessen. Dafür wurde die Translation einer RNA mit WT-SRE mit der Translation der MUT-RNA verglichen (Abb. 9A). Um den Einfluss der Zentrifugation zu beobachten, wurden Translationsrepressionsexperimente entweder direkt

durchgeführt – oder nach vorangegangener Inkubation in Embryoextrakt mit anschließender Gradientenzentrifugation (Abb. 9B). Mit der Luziferase-RNA ist der Unterschied in der Sedimentation zwischen WT- und MUT-RNAs nicht mehr zu erkennen (vgl. Abb. 7/10A). Die Auflösung von Gradientenzentrifugationen ist geringer für längere RNAs, vermutlich aufgrund unspezifisch bindender Proteine (Gray u. Hentze, 1994).

In Abb. 9B ist der Ablauf des Experiments dargestellt. Zu verschiedenen Zeitpunkten werden Proben für ein Translationsexperiment entnommen, wobei die Translation durch Zugabe eines ATP-regenerierenden Systems gestartet wird (rote Blitze). Wenn der stabile Komplex auf der RNA verbleibt, sollte die Repression auch ohne eine zweite Vorinkubation hoch sein (Abb. 9B oberes Experiment). Bei einer direkten Translationsmessung mit unbehandelter RNA im Extrakt erwartet man einen deutlichen Effekt der Vorinkubation auf den Repressionsfaktor (Abb. 9B unteres Experiment, ohne Gradientenzentrifugation).

Die Saccharosegradientenfraction mit der höchsten Luziferase-RNA-Menge (Abb. 10A - Fraktion 4, mit einem * markiert) wurde für die Regressionstests verwendet. Wie erwartet muss eine *in-vitro*-synthetisierte RNA mit Extrakt vorinkubiert werden, um stark reprimiert zu werden (Abb. 10B vgl. direkte Messung und Messung mit Vorinkubation für unbehandelte RNA). Eine weitere Vorinkubation von Fraktion 4 hatte dagegen keinen signifikanten Einfluss auf die Repression (Abb. 10B). Die Proteine, die für die stabile Repression an die RNA binden, sind demnach auch nach der Gradientenzentrifugation noch mit der RNA assoziiert.

Der gemessene Unterschied der Translation für WT- und MUT-RNAs ist keine Folge unterschiedlicher Stabilitäten der RNAs. Beide RNAs sind in gleichen Mengen in den jeweiligen Gradientenfractionen enthalten (Abb. 10D). RNAs (WT- und MUT-RNA), die aus gleichen Volumen der Gradientenfraction präpariert wurden, zeigten bei einer Translation in Retikulozytenlysate aus Kaninchen (RRL) keinen signifikanten Unterschied (Abb. 10C). Dies zeigt, dass gleiche Mengen an WT- und MUT-RNA gereinigt wurden, da die Translation in RRL nicht von SRE-*stem-loops* in der RNA beeinflusst wird (Smibert *et al.*, 1999).

2.2.3. Präparation des reprimierten RNPs zur massenspektrometrischen Analyse

Für die Präparation des Repressor-Komplexes aus *Drosophila*-Embryoextrakt wurden insgesamt 12 Gradienten je RNA (WT und MUT) mit jeweils 1 ml Reaktionsansatz beladen, zentrifugiert und fraktioniert. Der Reaktionsansatz enthielt, neben anderen Bestandteilen, 60% DEE sowie 10 nM biotinylierte RNA (1-AUG). Bei einer hundertprozentigen Ausbeute könnten daraus mehr als 60 μ g des reprimierten RNPs gereinigt werden (bei einem geschätzten Gesamtmolekulargewicht der gebundenen Proteine von mehr als 500 kDa).

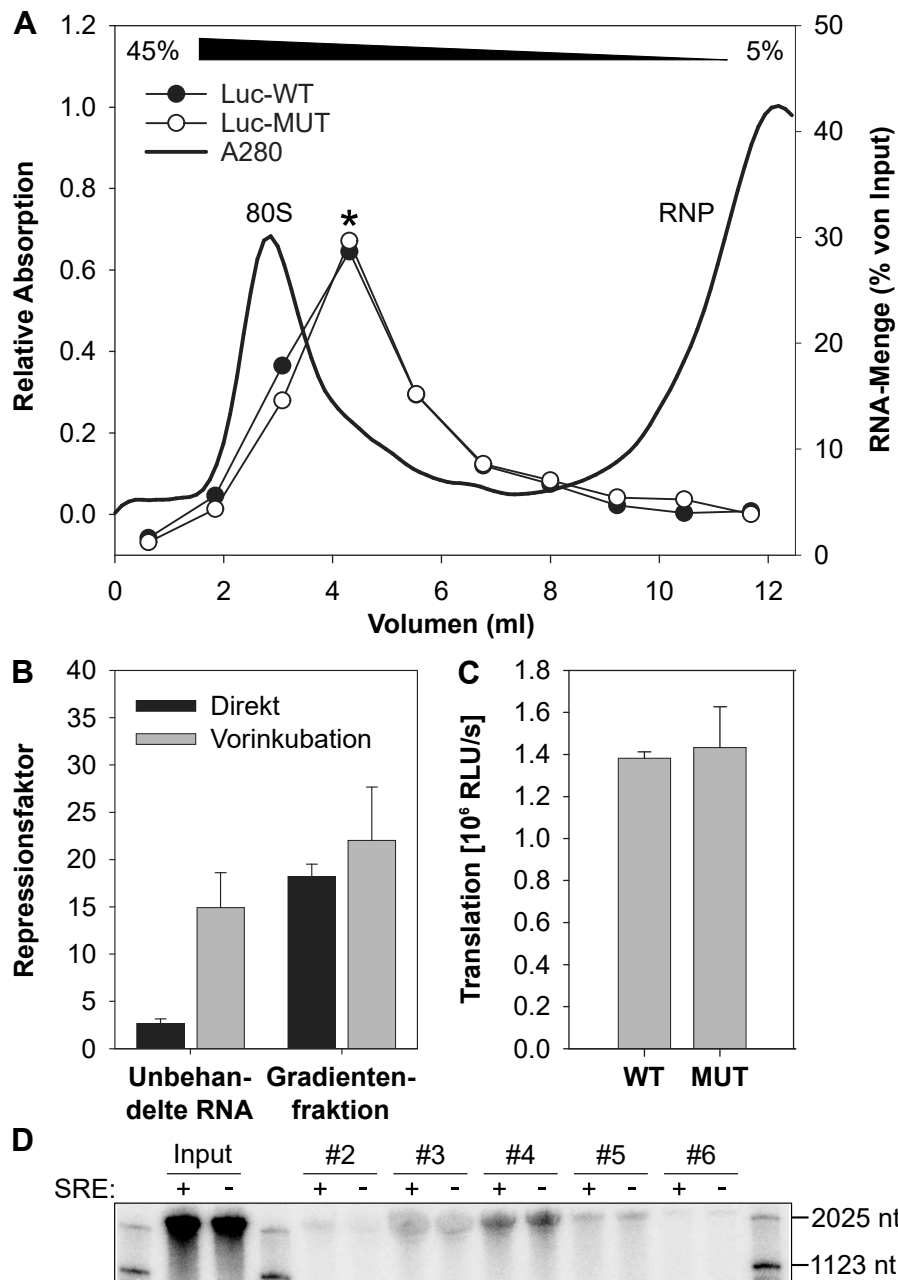


Abbildung 10: RNA bleibt während Gradientenzentrifugation reprimiert - (A) Radioaktiv markierte Luziferase-RNAs wurden in Embryoextrakt inkubiert und der gebildete Repressor-Komplex mittels Gradientenzentrifugation (5% - 45% Saccharose) fraktioniert (siehe Schema in Abb. 9). Die Verteilung der RNAs sowie das UV-Profil wurden aufgezeichnet, wie in der Legende zu Abb. 7 beschrieben. Fraktion 4 ist durch ein * markiert. (B) Translationsrepressionsmessung (direkt oder mit Vorinkubation) von Fraktion 4 im Vergleich zu 0,5 nM unbehandelter Luziferase-RNA. (C) RNAs aus Fraktion 4 wurden mittels Phenol/Chloroform-Extraktion und Ethanol-fällung gereinigt und anschließend in Retikulozytenlysate aus Kaninchen retranslatiert. (D) RNAs aus den Fraktionen 2 - 6 wurden einzeln präpariert und in einem Harnstoffgel analysiert. Die Fraktionen enthielten jeweils vergleichbare Mengen an WT- und MUT-RNA.

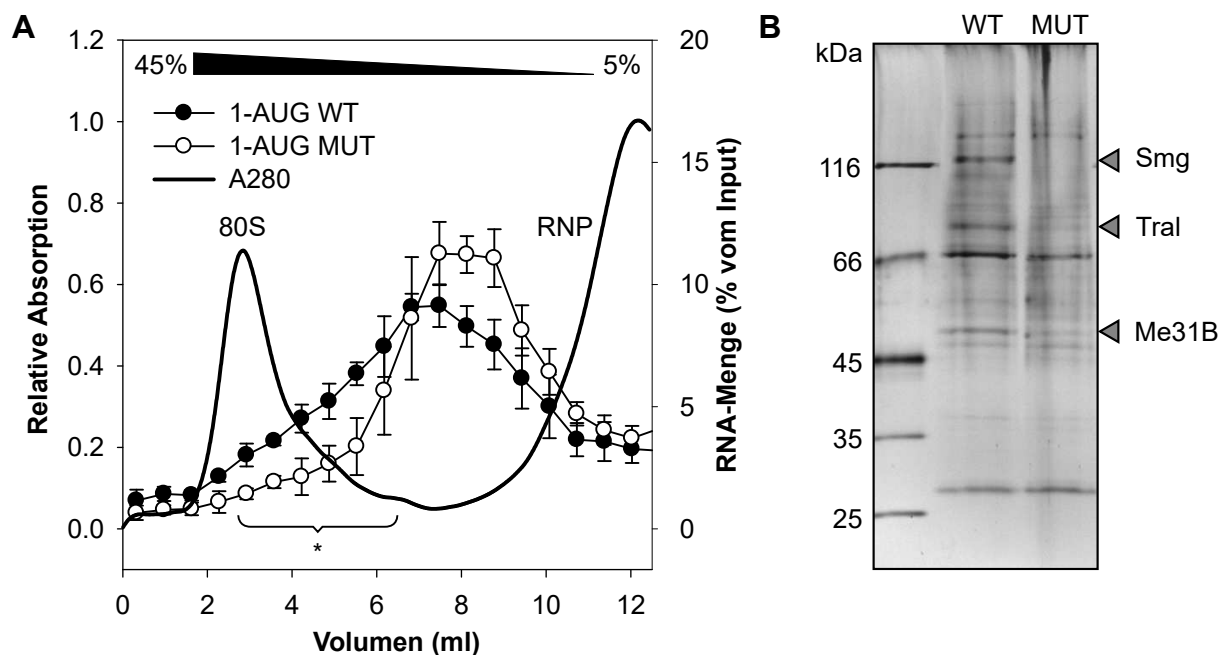


Abbildung 11: Präparation des reprimierten RNP - (A) Radioaktiv markierte, biotinylierte RNAs wurden mit Embryoextrakt inkubiert, um den Repressor-Komplex zu bilden. Durch Gradientenzentrifugation (5% - 45% Saccharose, durch den Keil angedeutet) wurden die gebildeten RNPs gereinigt. Die Verteilung der RNAs im Gradienten wurde durch Szintillation gemessen (n=4). Das UV-Profil wurde aufgezeichnet und beschriftet wie in Abb. 7. Die vereinigten Fraktionen für die weitere Reinigung sind durch ein * markiert. (B) Silber-gefärbte SDS-PAGE der Repressor-Komplexpräparation nach Affinitätsreinigung. Die molekularen Massen von Smaug, Me31B und Trailer Hitch korrespondieren mit deutlich angereicherten Banden in der SDS-PAGE.

Die Präparation wurde in vier Durchgängen durchgeführt. Dazu wurden jeweils am ersten Tag drei Reaktionsansätze je RNA via Gradientenzentrifugation fraktioniert und die benötigten Fraktionen vereint (Abb. 11A). Die gesammelten Fraktionen (Fraktionen 5 - 10 von insgesamt 20) wurden am nächsten Tag aufkonzentriert und anschließend mithilfe des inkorporierten Biotins durch Streptavidin-Affinitätschromatografie gereinigt. Die Elution der gebundenen Proteine erfolgte denaturierend in einem SDS-haltigen Puffer bei 80°C für 10 min. Unter diesen Bedingungen werden Proteine von der Matrix eluiert, während die biotinylierte RNA an der Matrix verbleibt, da Streptavidin, anders als die meisten Proteine, unter diesen Bedingungen nicht denaturiert wird. Etwa ein Drittel der eingesetzten RNA befand sich in den ausgewählten Gradientenfraktionen. Die Ausbeute des RNA-pull-downs war im Schnitt 30%, so dass eine Gesamtausbeute von $\approx 9 - 10\%$ der reprimierten RNA und des gebundenen Komplexes erhalten wurde.

Die Elutionsfraktionen der vier Reinigungsdurchgänge wurden vereint und durch SDS-Polyacrylamid-Gelelektrophorese analysiert (Abb. 11B). In einem weiteren, präparativen Gel wurde jede Spur in zwölf Stücke geschnitten und diese wurden getrennt voneinander für die massenspektrometrische Analyse (MS-Analyse) prozessiert.

2.3. MS-Analyse

Die Identifizierung angereicherter Proteine erfolgte an einem Orbitrap Fusion Tribrid-Massenspektrometer (Thermo, siehe Teil II Abschnitt 1.2.2, S. 76). Die qualitative und quantitative Analyse wurde mit der frei erhältlichen Software MaxQuant (Cox u. Mann, 2008) durchgeführt. Insgesamt wurden 2182 Proteine identifiziert, die direkt oder indirekt mit den analysierten RNAs interagierten. Die große Anzahl an Proteinen ist auf die hohe Sensitivität der Analyse und die noch zahlreich vorhandenen Verunreinigungen zurückzuführen. Die Vielzahl an Identifizierungen hat aber den entscheidenden Vorteil, dass eine statistische Auswertung quantitativer Daten möglich ist, da für den Großteil der Proteine angenommen werden kann, dass sie unspezifisch mit der RNA interagieren. Die spezifisch angereicherten Proteine sollten sich signifikant von den unspezifisch gebundenen Proteinen unterscheiden.

$$I_{\text{kor.}} = \log_2 \left\{ \frac{I \cdot PSM}{MW} \right\} \quad (1)$$

$$\text{Anreicherung} = \frac{I_{WT} \cdot PSM_{WT}}{I_{MUT} \cdot PSM_{MUT}} \quad (2)$$

Mit:

- $I_{\text{kor.}}$ - korrigierte Intensität
- I - Intensität (MaxQuant)
- PSM - *peptide spectrum matches*
- MW - Molekulare Masse des Proteins in kDa

MaxQuant gibt für jedes identifizierte Protein einen Intensitätswert (I) an, der auf der MS-Signalintensität zugehöriger Peptide beruht. Die Anzahl an Spektren, die dem jeweiligen Protein zugeordnet werden können (PSM - *peptide spectrum match*), wird häufig als einfache Quantifizierung von MS-Daten verwendet (Bantscheff *et al.*, 2012). Für eine optimierte Quantifizierung wurde der Intensitätswert mit der Anzahl an PSM multipliziert. Die Signalintensität der Peptidspektren ist unabhängig vom Molekulargewicht des zugehörigen Proteins. Die Anzahl möglicher Peptide und damit auch die PSM eines Proteins sind jedoch abhängig von dessen Größe. Deshalb wurde der PSM -Wert um die molekulare Masse des Proteins (MW) korrigiert. Für die Auftragung im Diagramm wurde der binäre Logarithmus des Intensitätswertes errechnet (Gleichung 1). Die Anreicherung eines Proteins im Vergleich von WT zu MUT wird mit Gleichung 2 errechnet.

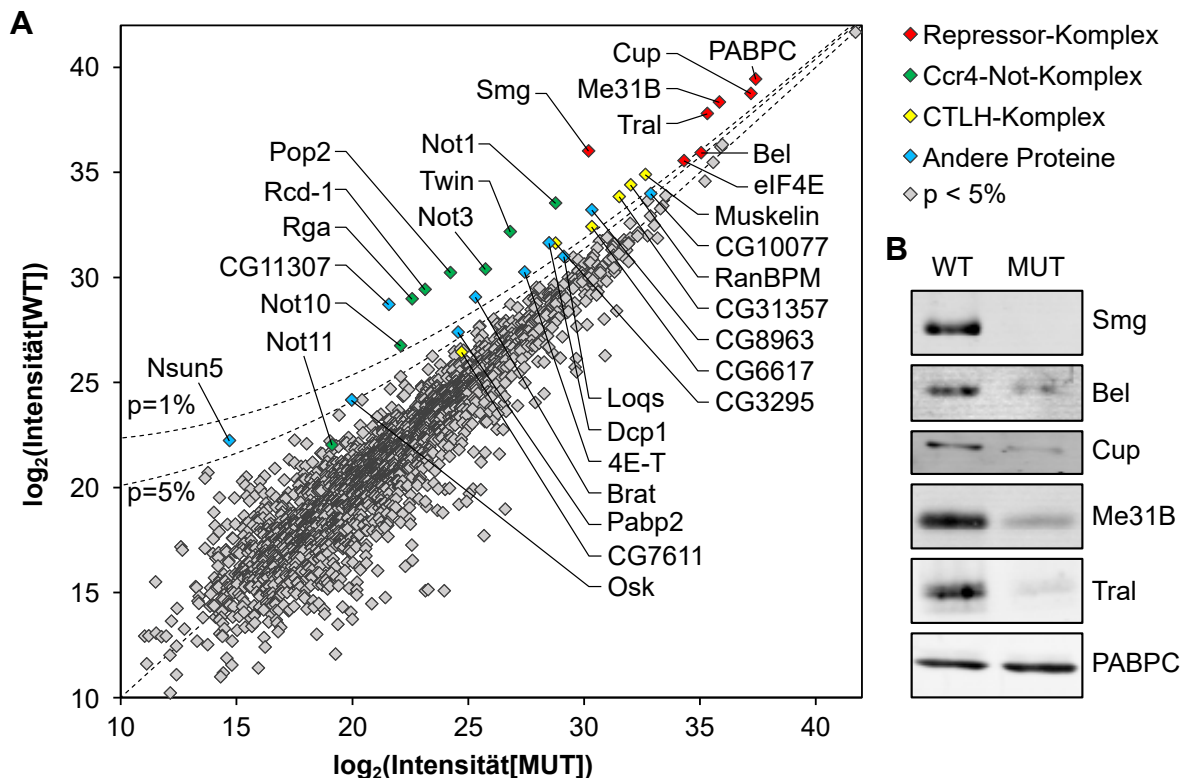


Abbildung 12: Quantitative Analyse der RNA-gebundenen Proteine - (A) Die Intensitäten für alle identifizierten Proteine in der Reinigung des WT- und MUT-RNPs wurden gegeneinander aufgetragen. Unspezifisch gebundene Proteine werden unabhängig vom SRE gebunden und liegen auf der Diagonalen. Spezifisch gebundene Proteine liegen oberhalb der Diagonalen. Proteine oberhalb der eingezeichneten Linien sind mit p-Werten kleiner als 5% bzw. 1% angereichert. (B) Die Anreicherung der Komplexkomponenten wurde mittels *Western-blot*-Analyse gezeigt.

Die Intensitäten für den WT- und MUT-*pull-down* wurden gegeneinander aufgetragen (Abb. 12). Vergleicht man die Intensitäten der Proteine im WT- und MUT-*pull-down*, lässt sich eine eindeutige Korrelation feststellen (Pearson-Koeffizient = 0.93). Dies bestätigt die Annahme, dass die meisten Proteine SRE-unspezifisch an die RNA binden. Die unspezifisch gebundenen Proteine sollten entlang der eingezeichneten Diagonalen im Diagramm auftreten. Im Wildtyp angereicherte Proteine sollten sich oberhalb dieser Diagonalen befinden.

Unter der gleichen Annahme wie zuvor wird erwartet, dass die Verteilung der Datenpunkte um die Diagonale in etwa einer Normalverteilung entspricht. Um angereicherte Proteine zu identifizieren, sollten „Ausreißer“ dieser Verteilung identifiziert werden. Dazu wurden 33 gleich große Bereiche entlang der Diagonalen getrennt voneinander betrachtet, die Abstände der Datenpunkte von der Diagonalen mit Gleichung 3 bestimmt und die Varianz der Verteilung der Datenpunkte um die Diagonale innerhalb jedes Bereichs berechnet (siehe Abb. 39, Anhang S. XLII). Gleichung 4 wurde an die erhaltenen Varianzen angepasst.

$$d(I_{korr,WT}; I_{korr,MUT}) = \frac{I_{korr,WT} - I_{korr,MUT}}{\sqrt{2}} \quad (3)$$

$$f(d) = a \cdot b^{c \cdot (45-d)} \quad (4)$$

$$p(I_{korr,WT}; I_{korr,MUT}) = \frac{1}{2} \cdot \operatorname{erfc} \left(\frac{I_{korr,WT} - I_{korr,MUT}}{2 \cdot a \cdot b^{c \cdot \left(\frac{\sqrt{2}}{2} \cdot (I_{korr,WT} + I_{korr,MUT}) - 45 \right)}} \right) \quad (5)$$

Mit:

d	- Abstand zur Diagonalen
f	- Anzupassende Funktion
p	- Signifikanz (p-Wert)
a, b, c	- angepasste Funktionsvariablen
I_{korr}	- korrigierte Intensität von WT oder MUT
$\operatorname{erfc}(x)$	- komplementäre Gaußsche Fehlerfunktion

Bei dieser Anpassung wurde die Anzahl an Datenpunkten je Bereich als Gewichtung mit einbezogen. Mit den erhaltenen Werten für a , b und c lassen sich mithilfe von Gleichung 5 die p-Werte für die Anreicherung der Proteine aus den bestimmten Intensitäten im WT- und MUT-*pull-down* errechnen. Zudem lassen sich mit Gleichung 4 Grenzfunktionen im Diagramm einzeichnen, die jeweils einem p-Wert von 1% bzw. 5% entsprechen. Dieser Analyse entsprechend konnten 29 Proteine als SRE-spezifische Interaktionspartner identifiziert werden ($p < 5\%$).

Die spezifisch angereicherten Proteine ($p < 5\%$) sind in Abb. 12A hervorgehoben und in Tabelle 1 zusammen mit orthologen Proteinen aus Mensch und Hefe aufgeführt. In der Analyse konnten drei Komplexe identifiziert werden. Angereicherte Proteine mit einer ähnlichen oder größeren Abundanz als Smaug werden im Folgenden als Repressor-Komplex bezeichnet (Abb. 12A, rot). Alle Proteine des Ccr4-Not-Komplex waren stark angereichert und sind grün eingezeichnet. Fünf angereicherte Proteine zeigten eine vergleichbare Domänenstruktur: eine Lis-Homologie-Domäne (LisH), gefolgt von einer CTLH-Domäne (C-terminal zu LisH). Diese können einem konservierten Komplex, dem CTLH-Komplex zugeordnet werden (gelb). Die Proteine beider Komplexe wurden mit einer geringeren Abundanz als die Repressor-Komplexproteine in der MS-Analyse identifiziert.

Die Anreicherung und deren Signifikanz lassen sich in einem Vulkandiagramm darstellen (Abb. 13). Smaug weist eine hohe Anreicherung (57x) und den kleinsten p-Wert aller angereicherten Proteine in der Analyse auf ($1,846 \cdot 10^{-22}$). Sowohl der Ccr4-Not-Komplex (grün)

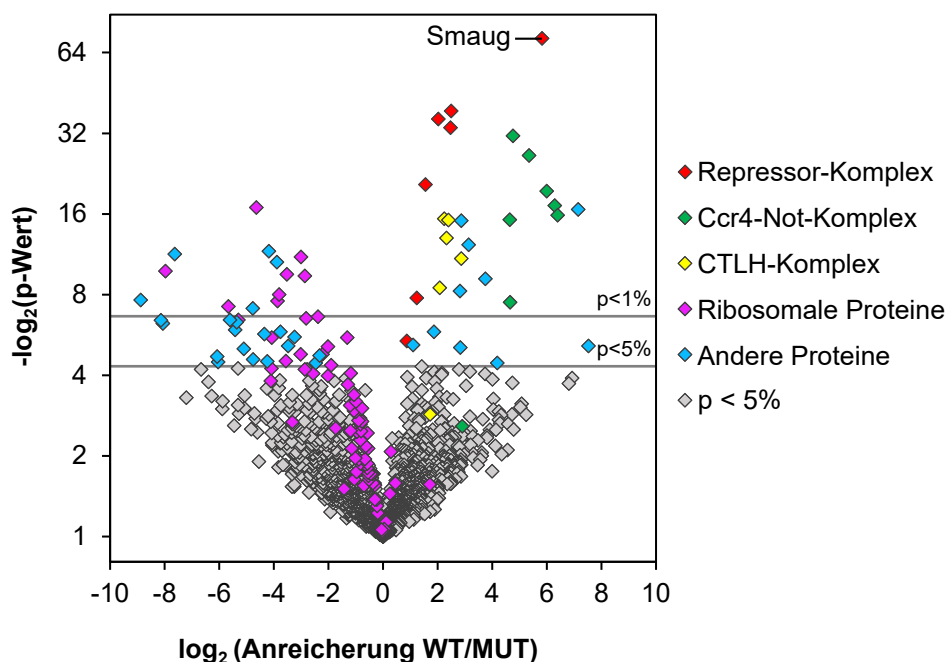


Abbildung 13: Vulkandiagramm der identifizierten Proteine - Die Signifikanz der Anreicherung der Proteine im Vergleich von WT- und MUT-*pull-down* ist aufgetragen gegen die Anreicherung. Die 1%- und 5%-Grenzen für p-Werte sind als horizontale Linien eingezeichnet. Unter den angereicherten Proteinen sind der Repressor-Komplex, der Ccr4-Not-Komplex sowie der CTLH-Komplex. Ribosomale Proteine (lila) sind deutlich und z.T. signifikant abgereichert. Tabelle 18, im Anhang auf S. XLIII, enthält alle signifikant angereicherten Proteine, wobei ribosomale Proteine hervorgehoben wurden.

als auch der CTLH-Komplex (gelb) bilden im Vulkandiagramm eine distinkte Gruppe. Die Proteine des Ccr4-Not-Komplexes sind mit einem Faktor von durchschnittlich 50x ähnlich stark angereichert wie Smaug. Der CTLH-Komplex wies mit einem Faktor von etwa 5x eine geringere Anreicherung auf. Die Proteine des Repressor-Komplexes sind im Diagramm weiter voneinander entfernt, als das für die beiden anderen Komplexe der Fall ist. Aber auch bei geringeren Anreicherungsfaktoren wiesen die Proteine des Repressor-Komplexes eine hohe Signifikanz auf.

Neben den angereicherten Proteinen können auch die abgereicherten Proteine wichtige Informationen über den reprimierten RNP preisgeben. Unter den 39 signifikant abgereicherten Proteinen sind z.B. 18 ribosomale Proteine enthalten (siehe Abb. 13 bzw. Abb. 40 und Tab. 18 im Anhang S. XLIII). Dies ist in guter Übereinstimmung mit dem reprimierten Status der RNA.

Tabelle 1: Im RNA-pull-down identifizierte Proteine - Alle signifikant angereicherten Proteine ($p < 5\%$) sind mit Anreicherungsfaktor (Gleichung 2) und der Signifikanz (Gleichung 5) aufgeführt. Drei Proteinkomplexe wurden dabei unterschieden. Die jeweiligen Orthologe im Menschen (z.T. auch Hefe) sind gelistet.

	Protein	Orthologe	Anreicherung	I_{korrr}	p-Wert
Repressor-Komplex	Smg	SAMD4A (Vts1p)	56,8	36,03	$1,846 \cdot 10^{-22}$
	Me31B	DDX6 (Dhh1p)	5,66	38,35	$2,150 \cdot 10^{-12}$
	PABPC	PABPC (Pab1p)	4,08	39,45	$1,298 \cdot 10^{-11}$
	Tral	LSm14 (Scd6p)	5,55	37,81	$7,827 \cdot 10^{-11}$
	Cup	-	2,95	38,77	$6,343 \cdot 10^{-7}$
	eIF4E	EIF4E (eIF4E)	2,36	35,57	$4,587 \cdot 10^{-3}$
	Bel	DDX3 (Ded1p)	1,83	35,94	0,02421
Ccr4-Not-Komplex	Not1	CNOT1 (Not1p)	27,1	33,55	$3,879 \cdot 10^{-10}$
	Twin	CNOT6 (Ccr4p)	41,0	32,18	$1,132 \cdot 10^{-8}$
	Pop2	CNOT7/8 (Caf1p)	63,9	30,24	$1,413 \cdot 10^{-6}$
	Rcd-1	CNOT9 (Caf40p)	78,6	29,44	$6,841 \cdot 10^{-6}$
	Rga	CNOT2 (Not2p)	84,4	28,98	$1,730 \cdot 10^{-5}$
	Not3	CNOT3 (Not3/5p)	25,1	30,40	$2,674 \cdot 10^{-5}$
	Not10	CNOT10	25,4	26,76	$5,618 \cdot 10^{-3}$
CTLH-Komplex	Muskelin	MKLN1	4,75	34,91	$2,397 \cdot 10^{-5}$
	RanBPM	RANBP9/10 (Gid1p)	5,28	34,42	$2,711 \cdot 10^{-5}$
	CG31357	MAEA (Gid9p)	5,03	33,86	$1,227 \cdot 10^{-4}$
	CG3295	RMND5A (Gid2p)	7,29	31,65	$5,251 \cdot 10^{-4}$
	CG6617	GID8 (Gid8p)	4,23	32,42	$2,835 \cdot 10^{-3}$
	CG11307	-	141	28,73	$9,946 \cdot 10^{-6}$
	CG8963	PAIP1	7,33	33,22	$2,890 \cdot 10^{-5}$
	Dcp1	DCP1A/B (Dcp1p)	8,85	31,65	$2,054 \cdot 10^{-4}$
	Brat	TRIM3 (Psh1p)	13,5	29,08	$1,769 \cdot 10^{-3}$
	4E-T	4E-T	7,10	30,27	$3,306 \cdot 10^{-3}$
	Loqs	PRKRA	3,65	31,01	0,01794
	Nsun5	NSUN5 (Rcm1p)	185	22,24	0,02842
	Pabp2	PABPN1	7,14	27,40	0,02979
	CG10077	DDX5/DDX17 (Dbp2p)	2,16	34,00	0,02747
	Osk	-	18,3	24,18	0,04589

2.3.1. Identifizierte Proteine

Repressor-Komplex

In den Abbildungen 12 und 13 sind die Proteine rot markiert, die zum Repressor-Komplex gezählt werden können. Diese Proteine waren signifikant angereichert und wiesen eine ähnliche oder höhere Intensität als Smaug auf. Smaug als zentrale Komponente rekrutiert die Proteine Cup, Me31B und Trailer Hitch, wie bereits beschrieben (Nelson *et al.*, 2004; Jeske *et al.*, 2011). Das Cap-bindende Protein eIF4E war im Wildtyp leicht angereichert. Die verwendeten RNAs trugen kein 5'-Cap, so dass eine direkte Bindung von eIF4E an die RNA nicht der Grund für die Assoziation mit dem Repressor-Komplex oder für die beobachtete Anreicherung gewesen sein sollte. Viel wahrscheinlicher ist, dass eIF4E durch die spezifische Interaktion mit Cup im Wildtyp angereichert war. Das zytoplasmatische Poly(A)-bindende Protein (PABPC) war ebenfalls angereichert, obwohl beide RNAs einen gleich langen, internen Poly(A)-Schwanz enthielten. Im Gegensatz dazu wurde bei RNA-Affinitätsreinigungen direkt aus Embryoextrakt eine gleichmäßige Bindung von PABPC sowohl an die mutierte als auch an die Wildtyp-RNA beobachtet (Abb. 12B). Der größere Zeitaufwand der Reinigung könnte eine Stabilisierung der PABPC-Bindung durch den Repressor-Komplex andeuten. Dies stimmt mit der Beobachtung überein, dass die Repression SRE-haltiger RNA von der Anwesenheit eines Poly(A)-Schwanzes in der Reporter-RNA erhöht wird (Jeske *et al.*, 2011).

Ccr4-Not-Komplex

Smaug rekrutiert den Ccr4-Not-Komplex an eine gebundene RNA und bewirkt deren Deadenylierung. In Semotok *et al.* (2005) und Pinder (2012), wurde die Interaktion von Smaug mit den Ccr4-Not-Komponenten Not1, Ccr4 und Caf1 durch Reinigung von TAP-Smaug gezeigt. In der in dieser Arbeit gezeigten Analyse konnten nun alle Kernkomponenten des Ccr4-Not-Komplexes als spezifische Interaktionspartner des reprimierten RNPs identifiziert werden. Alle Ccr4-Not-Komponenten sind grün markiert (Abb. 12 und Abb. 13). Die Proteine des Ccr4-Not-Komplexes bilden eine distinkte Gruppe mit geringerer Intensität als Smaug. Die sehr spezifische Anreicherung der Ccr4-Not-Kernkomponenten verdeutlicht die Spezifität der Analyse. Lediglich Not11, eine entbehrliche, periphere Untereinheit des Ccr4-Not-Komplexes (Temme *et al.*, 2014), ist in der Analyse nicht signifikant angereichert ($p = 16,8\%$).

CTLH-Komplex

Neben den Repressor-Komplexkomponenten und dem Ccr4-Not-Komplex wurden fünf Untereinheiten eines konservierten Komplexes im reprimierten RNP identifiziert und in den Abbildungen 12A und 13 gelb markiert. Diese Proteine haben jeweils eine LisH- und CTLH-Domäne gemeinsam (Regelmann *et al.*, 2003; Francis *et al.*, 2013) und sind als CTLH-Komplex, oder auch in Hefe als GID-Komplex (*glucose-induced degradation deficient*), beschrieben (Kobayashi *et al.*, 2007). Dieser Komplex ist in Hefe und in *Xenopus*

als E3-Ubiquitinligase aktiv (Regelmann *et al.*, 2003; Pfirrmann *et al.*, 2015). In Hefe wird die Fruktose-1,6-bisphosphatase (FBPase), ein Enzym der Glukoneogenese, abgebaut, wenn die Hefe in glukosehaltiges Medium überführt wird. Der GID-Komplex ist für die Ubiquitinierung der FBPase verantwortlich, die letztlich zum proteasomalen Abbau führt.

Unter den fünf identifizierten Proteinen ist auch das Ortholog der katalytisch aktiven Untereinheit Rmnd5/Gid2 (CG3295 in *Drosophila*) enthalten. Eine sechste Untereinheit scheint in *Drosophila* konserviert zu sein, lag in dieser Analyse allerdings nur mit geringerer Abundanz und Signifikanz vor (CG7611, $p=18,7\%$). Das Hefe-Homolog zu CG7611 (GID7) bindet unabhängig von allen anderen Untereinheiten des GID-Komplexes an das RanBPM-Homolog GID1 (Menssen *et al.*, 2012).

Unter den Proteinen, deren Menge während der *maternal-to-zygotic transition* (MZT) deutlich reduziert wird, befinden sich unter anderen Smaug, Cup, Me31B, Trailer Hitch, eIF4E und PABPC (Gouw *et al.*, 2009) und damit sechs der sieben Komponenten des Repressor-Komplexes. Der CTLH-Komplex könnte demnach als E3-Ubiquitinligase für die Ubiquitinierung und den schnellen Abbau der maternalen Proteine während der MZT von Bedeutung sein.

Die Analyse mittels Massenspektrometrie erlaubt die Identifizierung von Ubiquitin-modifizierten Proteinen. Ubiquitin wird durch eine Isopeptidbindung mit einem Lysin im Zielprotein verknüpft. Die Proteinsequenz von Ubiquitin endet auf Arg-Lys-Arg-Gly-Gly. Trypsin spaltet die Peptidbindung C-terminal von Lysin und Arginin. Nach einer tryptischen Proteolyse verbleibt die Isopeptidbindung zu einem Diglycin (+114,04 u) an der ursprünglich ubiquitinierten Stelle. Es kann auch vorkommen, dass die letzten vier Aminosäuren von Ubiquitin (LRGG; +383,23 u) am Peptid verbleiben. Diese Modifizierungen können mit gängigen Proteom-Suchalgorithmen massenspektrometrisch identifiziert werden.

Ubiquitin wurde zudem als sehr abundantes Protein in der Analyse identifiziert (siehe Abb. 40 im Anhang S. XLIV). Die erhaltenen Datensätze wurden erneut mit MaxQuant auf mögliche Ubiquitinierungsstellen untersucht. Dazu wurden Lysin-Modifizierungen von 114,04 u und 383,23 u zugelassen. Da die meisten ubiquitinierten Proteine vom Proteasom degradiert werden, wird für die Identifizierung von Ubiquitinierungsstellen normalerweise das Proteasom inhibiert. Obwohl dies hier nicht der Fall war, konnten in der massenspektrometrischen Analyse Ubiquitinierungsstellen identifiziert werden (in Me31B, Cup und PABPC). Das beste Beispiel einer solchen Ubiquitinierungsstelle war in Me31B zu finden. Zusätzlich zu einer GG-Modifizierung konnte das gleiche Peptid mit einer LRGG-Modifizierung identifiziert werden. In Abb. 14 ist das Fragmentionen-Spektrum dieses Peptids mit einer Diglycin-Modifizierung gezeigt. Die Annotation des Spektrums

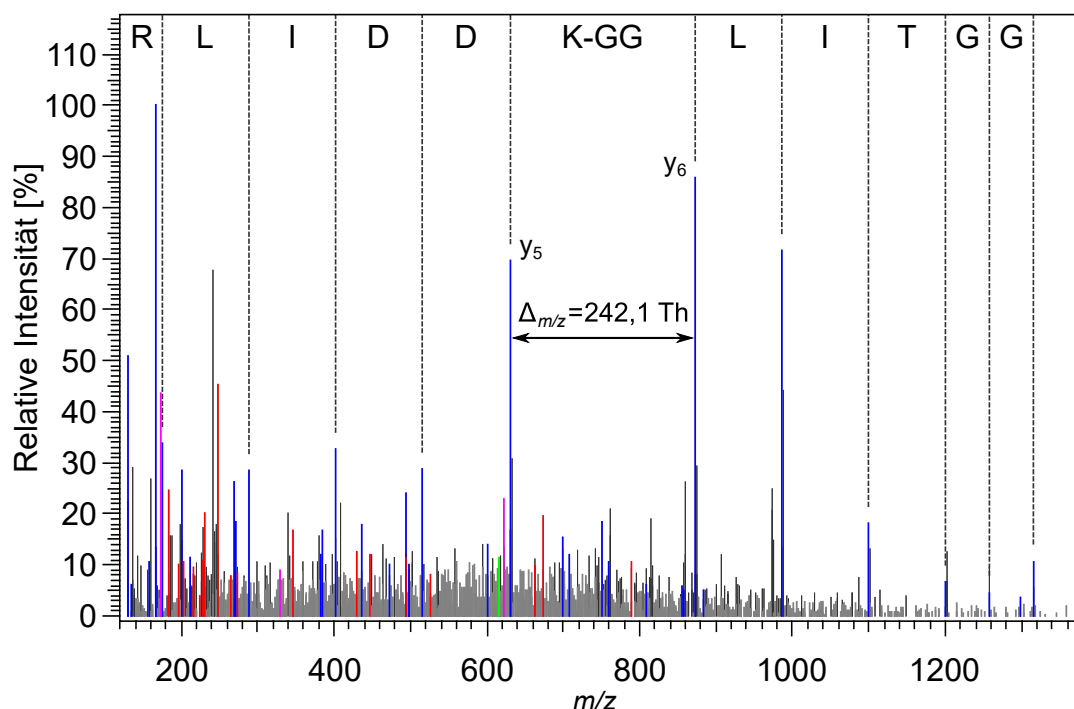


Abbildung 14: Ubiquitinierung von Me31B - Dargestellt ist ein Ausschnitt des Fragmentionen-Spektrums (Vorläufer: $621,3399^{3+}$) eines Me31B-Peptids (Aminosäuren 157 - 172, VMVTTGGTILKDDILR), mit einer Modifizierung von 114,04 u an Lysin 167. Diese Modifizierung kann als Überrest einer Ubiquitinierung an einem Lysin gemessen werden, da nach der Proteolyse die beiden C-terminalen Glycine am Lysin verbleiben. Zur Annotation des Spektrums wurde StavroX genutzt (für Details siehe Teil II, Abschnitt 2.4.3, S. 110). Ionen im Spektrum sind nach Fragmentionentyp eingefärbt (y-Ionen sind blau, b-Ionen rot und Vorläuferionen sind grün dargestellt). Aus dem Spektrum lässt sich aus aufeinanderfolgenden Abständen der y-Ionen ein Teil der Peptidsequenz ablesen (GGTILK[-GG]DDILR). Diese markante y-Ionenserie ist im Spektrum hervorgehoben und annotiert (K-GG für das Diglycin-modifizierte Lysin).

erfolgte mit der *Cross-Linking*/MS-Software StavroX (siehe Teil II, Abschnitt 2.4.3, S. 110). Die Annotation durch eine *Cross-Linking*/MS-Software war in diesem Fall leicht möglich, da eine Ubiquitinierung im Prinzip einem Protein-Protein-*Cross-Link* entspricht. Die insgesamt drei erhaltenen Spektren dieses modifizierten Peptids sind in den Abbildungen 41 und 42 im Anhang ab Seite XLV so dargestellt, wie sie aus StavroX exportiert wurden.

In Abb. 14 ist die einfach positiv geladene y-Ionenserie (siehe Teil II Abschnitt 1.2.2, S. 76) hervorgehoben, anhand derer deutlich die Modifizierung des Lysins 167 in Me31B erkannt werden kann. Die Abstände zweier aufeinanderfolgender Ionen gleichen Typs entsprechen der Masse der zwischen den Fragmentierungsstellen liegenden Aminosäure. Im Falle der einfach geladenen Ionen y_5 und y_6 entspricht diese Differenz 242,1 Th und damit der Masse eines Lysin-Restes (128,1 u) mit einer Diglycin-Modifizierung (114,0 u).

Ob diese Ubiquitinierung von Me31B für den Proteinabbau während der MZT relevant ist und ob diese Modifizierung vom CTLH-Komplex abhängig ist, kann mit den hier gezeigten Experimenten nicht belegt werden.

Die DEAD-Box-ATPase Belle

Die DEAD-Box-Helikase Belle (DDX3) ist die einzige bislang unbekannte Komponente des Repressor-Komplexes, die in dieser Analyse identifiziert wurde. Die Anreicherung von Belle in der MS-Analyse ist mit 1,8x vergleichsweise gering. Die spezifische Assoziation der Proteine Smaug, Cup, Me31B, Trailer Hitch und auch Belle mit SRE-haltiger RNA konnte mit *pull-down*-Experimenten und *Western-blot*-Analysen bestätigt werden (siehe Abb. 12B, S. 27).

In Kooperation mit der Arbeitsgruppe von Martine Simonelig wurde die Beteiligung von Belle an der posttranskriptionalen Regulation der *nanos*-mRNA untersucht. Die nachfolgenden Analysen von *bel*-Mutanten wurden von Jérémy Dufourt und Nagraj Sambrani (Institut de Génétique Humaine in Montpellier, Frankreich) durchgeführt. Die genetischen Experimente wurden zusammen mit den biochemischen Experimenten dieser Arbeit zur Publikation eingereicht (Götze *et al.*, eingereicht). Zwei Fliegenstämme (*bel*⁶ und *bel*^{L4740}; Ihry *et al.*, 2012; Johnstone *et al.*, 2005) mit Mutationen im *bel*-Gen sollten dafür während der frühen Embryogenese analysiert werden. Die Fliegenstämme zeigten jedoch einen homozygot-letalen Phänotyp noch während des Larvenstadiums. Es schlüpften keine adulten Fliegen, so dass ein maternaler Effekt auf die Embryogenese nicht beobachtet werden konnte. Um dennoch Weibchen zu erhalten, bei denen beide *bel*-Allele mutiert sind, wurden die genannten Stämme mit einem hypomorphen *bel*-Stamm gekreuzt (*bel*^{neo30}). Die aus diesen Kreuzungen entstandenen transheterozygoten Weibchen trugen jeweils eine unterschiedliche Mutation in den beiden Allelen des *bel*-Gens, waren lebensfähig und produzierten Eier.

In den Embryonen von diesen transheterozygoten *bel*-Mutanten kann ein maternaler Effekt auf die Embryogenese beobachtet werden, wenn diese mit Wildtyp-Männchen gekreuzt werden. Die von *bel*^{neo30}/*bel*⁶- oder *bel*^{neo30}/*bel*^{L4740}-Weibchen abgelegten Embryonen zeigten eine Beeinträchtigung in der *nanos*-Regulation: In *bel*^{neo30}/*L4740*-Embryonen konnte das Nanos-Protein im gesamten Embryo und nicht, wie im Wildtyp, ausschließlich am posterioren Pol detektiert werden. In beiden *bel*-Stämmen lag in 2 - 4 h alten Embryonen signifikant mehr *nanos*-mRNA vor. Die Deadenylierung der *nanos*-mRNA war in beiden Mutanten beeinträchtigt. Die Identifizierung von Belle als Bestandteil des SRE-abhängigen RNPs *in vitro* konnte, trotz der geringen Anreicherung in der MS-Analyse und *pull-downs* durch dessen Einfluss auf die posttranskriptionale *nanos*-Regulation *in vivo* bestätigt werden.

Andere Proteine

Neben den bereits genannten Proteinen aus den drei beschriebenen Komplexen wurden zehn weitere Proteine identifiziert, die keinem der Komplexe zugeordnet werden konnten. Die Proteine sind hier, wie auch in Tabelle 1, nach p-Werten sortiert. Alle zehn Proteine wiesen eine deutlich geringere Abundanz als der Repressor-Komplex auf. Eine Rolle dieser Proteine im stabilen mRNP ist damit fraglich. Einige der Proteine wurden vermutlich aufgrund ihrer Interaktionen mit Komplexkomponenten in dieser Analyse identifiziert.

CG11307 ist nur in Insekten konserviert (NCBI-blastp) und zeigt Homologien zu einer Glykosyltransferase (FlyBase.org). Das CG11307-Transkript ist hauptsächlich während der ersten 4 h der Embryonalentwicklung exprimiert. Das Protein war ähnlich stark angereichert wie der Ccr4-Not-Komplex und ebenfalls in einer Smaug-Immunpräzipitation angereichert (siehe Abschnitt 2.6, S. 44).

CG8963 ist das *Drosophila*-Ortholog zum PABPC-interagierenden-Protein PAIP1. Es enthält eine MIF4G-ähnliche Domäne und eine Armadillo-*repeat*-Domäne. Die Identifizierung von CG8963 in dieser Analyse beruht vermutlich auf seiner Interaktion mit dem Poly(A)-Bindepotein, welches mit einer hohen Intensität in der Analyse identifiziert wurde.

Dcp1, das *decapping*-Aktivatorprotein wurde als direkter Interaktionspartner von Trailer Hitch beschrieben (Tritschler *et al.*, 2008). Zusammen mit Me31B, Edc3, Trailer Hitch und weiteren Faktoren kommt Dcp1 in *P-bodies* vor. Ein bekannter Interaktionspartner, das *decapping*-Enzym Dcp2, konnte in der MS-Analyse nicht identifiziert werden.

Brat (Brain Tumor) ist ein als Translationsrepressor beschriebenes Protein (Sonoda u. Wharton, 2001). Die Bindesequenz des Brat-Proteins ist bekannt (UUGUUG/A), war aber nicht in den verwendeten *Bait*-RNAs enthalten. Die *nanos*-mRNA gehört zudem nicht zu den identifizierten Zielgenen von Brat (Laver *et al.*, 2015; Loedige *et al.*, 2015).

4E-T, ein Cup-Paralog, weist ebenfalls ein eIF4E-Bindemotiv auf (YxxxxLΦ). Es wäre denkbar, dass Cup von 4E-T im Repressor-Komplex ersetzt werden kann. Das humane Me31B-Homologe (DDX6) interagiert mit 4E-T (Ozgur *et al.*, 2015; Nishimura *et al.*, 2015; Kamenska *et al.*, 2016), wodurch ebenfalls die Anreicherung im Smaug-RNP erklärt werden könnte.

Loqs, das Loquacious-Protein, ist in die Biogenese von miRNAs involviert (Jakob *et al.*, 2016). Es kann doppelsträngige RNA binden und interagiert mit Dicer-1, welches in der MS-Analyse nicht angereichert vorlag. Auch andere Proteine der miRNA-Biogenese und -Regulation waren nicht angereichert (z.B. Argonaut-Proteine oder Gawky; Ha u. Kim, 2014).

Nsun5 ist, Sequenzhomologien folgend, eine S-Adenosylmethionin-abhängige RNA-Methyltransferase. Die *nsun5*-mRNA kommt hauptsächlich in den ersten 12 h der Embryogenese vor. Nsun5 war das Protein mit dem geringsten Intensitätswert unter den angereicherten Proteinen. Dies weist höchstens auf eine wenig stabile, substöchiometrische Assoziation des Proteins mit Smaug oder SRE-haltiger RNA hin.

Pabp2, das nukleäre Poly(A)-Bindepotein, wurde ebenfalls identifiziert. Da die genutzte RNA zu Embryoextrakt gegeben wurde, ist es wahrscheinlich, dass Pabp2 mit PABPC um den Poly(A)-Schwanz konkurriert und so auch als Interaktionspartner identifiziert wird. In *Drosophila*-Embryonen ist Pabp2 auch zytoplasmatisch lokalisiert und besitzt eine Funktion in der Poly(A)-Längenkontrolle während der Embryogenese (Benoit *et al.*, 2005).

CG10077 ist, wie auch Me31B und Belle, eine DEAD-Box-Helikase. Die Assoziation von Smaug und CG10077 wurde bereits in einer Tandem-Affinitätsreinigung (TAP) von Smaug aus Embryoextrakt gezeigt (Pinder, 2012). In derselben Arbeit wurde eine weitere DEAD-Box-Helikase identifiziert, die mit TAP-Smaug interagiert: CG7878. Dieses Protein lag im RNA-*pull-down* allerdings nicht angereichert vor.

Osk (Oskar) bewirkt am posterioren Pol des Embryos die Aufhebung der Smaug-vermittelten Translationsrepression. Oskar interagiert einerseits direkt mit Smaug, besitzt aber auch eine RNA-Bindedomäne (Jeske *et al.*, 2015) und könnte demnach durch eine der beiden Interaktionen als spezifischer Interaktionspartner des SRE-spezifischen RNPs identifiziert worden sein.

2.4. Quantifizierung der identifizierten Repressor-Komplexkomponenten

Für einige Komponenten des Repressor-Komplexes lagen sowohl spezifische Antikörper als auch rekombinante Proteine in ausreichender Menge vor, um eine Quantifizierung der Proteine mittels *Western-blot*-Analysen durchzuführen.

2.4.1. Konzentration der identifizierten Proteine im Embryoextrakt

In Tab. 2 sind die Embryoextraktkonzentrationen der identifizierten Proteine zusammengefasst. Die Konzentrationen im Embryo sind wegen der Verdünnung mit Extraktpuffer mindestens um einen Faktor von zwei höher. Um die Konzentrationen im Embryo genau zu bestimmen, müssten ganze Embryonen denaturierend aufgeschlossen werden und im *western blot* mit Proteinstandards verglichen werden. Smaug wies eine Konzentration von etwa 80 nM im Extrakt auf. Nur eine geringe Konzentration SRE-haltiger RNA

(≈ 10 nM) kann in einem *in-vitro*-Translationsexperiment effizient reprimiert werden. Darüber hinaus verringert sich der gemessene Repressionsfaktor. In einem Translationsexperiment werden 40% Embryoextrakt eingesetzt und eine Reporter-RNA besitzt zwei Smaug-Bindestellen. Für 10 nM Luc-RNA ergibt sich daraus ein Überschuss von Smaug über RNA-Bindestellen von etwa 1,5x, was Smaug als die mögliche limitierende Komponente identifiziert. Die Komponenten Cup (≈ 2 μ M), Me31B (≈ 4 μ M) und Trailer Hitch (≈ 8 μ M) liegen in deutlichem Überschuss über Smaug im Extrakt vor. Belle ist mit einer Konzentration von 0,8 μ M in etwa zehnfachem Überschuss über Smaug vorhanden.

Tabelle 2: Proteinkonzentrationen in Embryoextrakten - Die Konzentrationen der Proteine in Embryoextrakt wurde durch quantitative *Western-blot*-Analysen bestimmt und in pmol/mg Protein im Extrakt sowie in μ M angegeben. Die Proteine wurden in verschiedenen Extrakten quantifiziert (n=8), die im Mittel eine Proteinkonzentration von 46 g/l aufwiesen.

Protein	Konzentration (pmol/mg)	Konzentration (μ M)
Smaug	1,7	0,08
Me31B	75	3,5
Trailer Hitch	166	7,6
Belle	17	0,77
Cup (n=3)	43	2

2.4.2. Stöchiometrie der Komponenten im Repressor-Komplex

Um die Stöchiometrie der Komponenten zueinander im Repressor-Komplex zu messen, wurden RNA-*pull-downs* durchgeführt und die Menge an gebundenem Protein anhand quantitativer *Western-blot*-Analysen quantifiziert. Die im *pull-down* gebundene Menge an RNA konnte durch Szintillationszählung der radioaktiv markierten RNA bestimmt werden. Dadurch ist nicht nur die Stöchiometrie der einzelnen Proteine zueinander, sondern auch deren Stöchiometrie zur RNA bestimmbar. Ein Beispiel für einen solchen quantitativen *western blot* ist in Abb. 15 gezeigt. Die Konzentration der Proteine aus verschiedenen *pull-downs* wurden durch den Vergleich mit rekombinantem Protein densitometrisch bestimmt (Abb. 16).

Me31B lag lediglich als Maltosebindeprotein-Fusionsprotein vor. Dadurch konnte es nicht auf gleicher Höhe im *western blot* detektiert werden wie das Wildtyp-Protein. Die Konzentration von Me31B im Extrakt wurde in einem separaten Experiment bestimmt. Die Quantifizierung von Me31B im Repressor-Komplex basiert dann auf dem Vergleich mit Embryoextrakt-Proben mit bekannter Me31B-Konzentration.

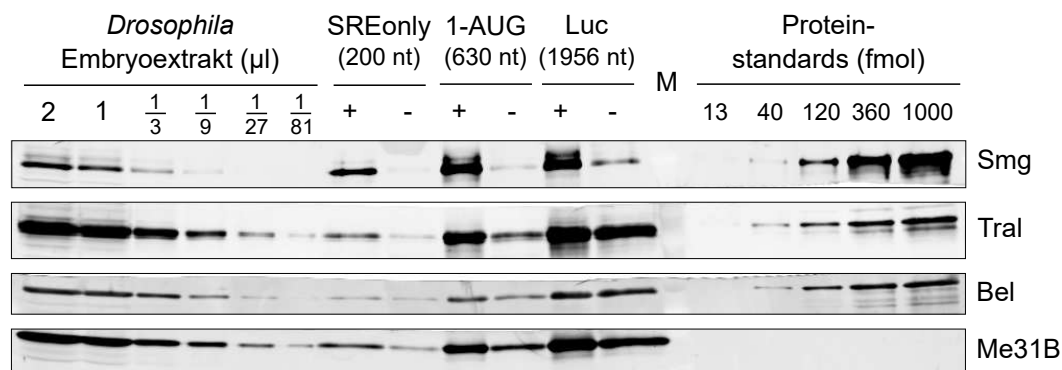


Abbildung 15: Quantitative Western-blot-Analyse - Die Proteine des Repressor-Komplexes wurden mittels RNA-Affinitätsreinigung aus Extrakt gereinigt. Dazu wurden RNAs mit unterschiedlichen Längen verwendet, die jeweils zwei SREs (WT [+] oder MUT [-]) enthielten. Es wurde die Menge an Protein geladen, die an 500 fmol RNA gebunden hatte. Verschiedene Mengen an Embryoextrakt und rekombinanten Proteinstandards wurden zum Vergleich aufgetragen und durch eine SDS-PAGE und anschließenden *western blot* analysiert.

In Abb. 16A-C und Tab. 3 ist jeweils die gemessene Anzahl an Proteinen je RNA-Molekül angegeben. Die Stöchiometrie wurde dazu mit drei unterschiedlich langen RNAs bestimmt. Alle RNAs enthielten einen internen Poly(A)-Schwanz (siehe Abb. 9). Die SREonly-RNA (200 nt; Abb. 16A) enthielt die beiden SRE-*stem-loops*, die mit 9 nt Abstand voneinander kloniert wurden. Die 1-AUG-RNA (630 nt; Abb. 16B) enthielt ein kurzes, offenes Leseraster (102 Aminosäuren) sowie die ersten 158 nt des *nanos*-3'-UTR. Die Luc-RNA (1959 nt; Abb. 16C) enthielt das offene Leseraster für die Luziferase sowie den gleichen Teil des *nanos*-3'-UTR wie die 1-AUG-RNA. Die *nanos*-mRNA, wie sie in Embryonen vorkommt, ist mit einer Länge von 2349 nt (ohne Poly(A)-Schwanz) noch länger als die längste hier analysierte RNA.

Die Quantifizierung der einzelnen Komponenten ist beispielhaft für Smaug in Abb. 16D dargestellt. Für die SREonly-RNA binden etwa 500 fmol Smaug an 500 fmol WT-RNA. Sowohl für die 1-AUG-RNA als auch die Luc-RNA binden etwa 1000 fmol Smaug an 500 fmol RNA. In Grün sind die Spezifitätskontrollen mit mutiertem SRE eingezeichnet. Die unspezifische Bindung lag bei etwa 100 fmol (≈ 200 fmol für die Luc-RNA). Die Menge an gebundenem Smaug-Protein ist demnach unabhängig von der Länge der verwendeten RNA. Da je SRE ein Smaug Protein gebunden werden kann, ist zu erwarten, dass zwei Moleküle Smaug je Molekül RNA präzipitiert werden. Das konnte für die 1-AUG und die Luc-RNA beobachtet werden. Für die SREonly-RNA wurde nur die Bindung eines Smaug-Moleküls beobachtet. Das könnte allerdings auf die artifizielle Struktur der SREonly-RNA zurückzuführen sein. Es könnte, aufgrund der räumlichen Nähe, zur Kompetition der beiden SREs untereinander kommen. Die SREonly-RNA ist funktional, da sie SRE-spezifisch deadenyliert wird (Jeske *et al.*, 2006). Dazu genügt aber vermutlich ein einziges funktionales SRE (Tadros *et al.*, 2007). Die Quantifizierung könnte noch einmal

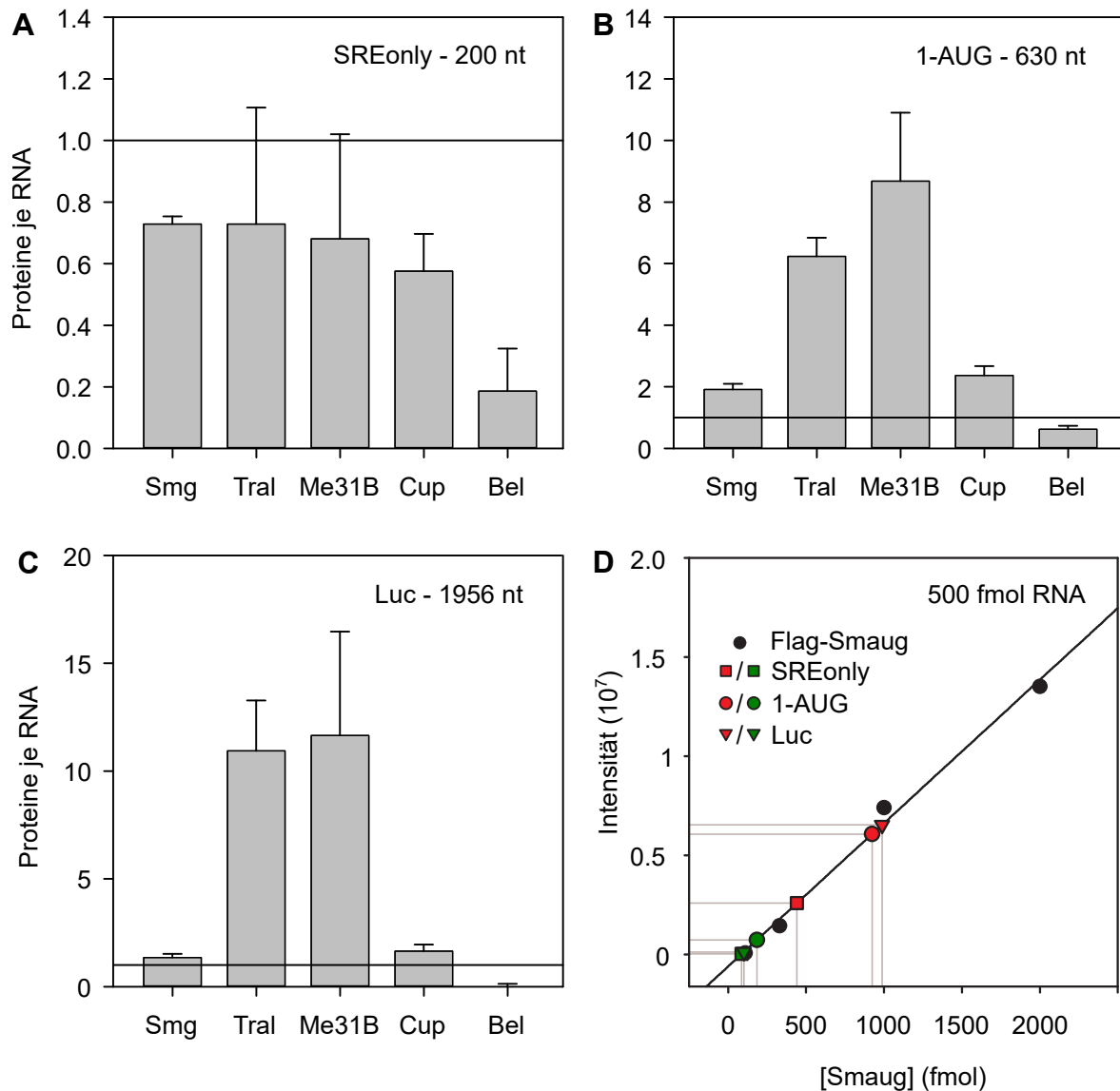


Abbildung 16: Längenabhängige Stöchiometrie der Komponenten im Repressor-Komplex - (A-C) Die Komponenten des Repressor-Komplexes wurden wie in Abb. 15 für drei verschieden lange RNAs quantifiziert. Die Proteinmengen wurden auf die eingesetzte RNA-Menge relativiert. Der spezifische Anteil der Bindung (die Differenz zwischen WT und MUT) wurde für die drei RNAs dargestellt (A - SREonly, B - 1-AUG, C - Luc). In (D) ist die Quantifizierung beispielhaft für Smaug dargestellt. Die Intensität der Banden im *western blot* wurde densitometrisch bestimmt. Verschiedene Mengen an Flag-Smaug wurden als Standard aufgetragen, quantifiziert und anschließend wurde eine lineare Regression durchgeführt. Die Menge an gebundenem Protein wurde von der erhaltenen Geraden abgelesen (WT - rot, MUT - grün).

mit Transkripten durchgeführt werden, die eine, bzw. mehr als zwei SREs enthalten, um den Einfluss der SRE-Anzahl auf die Stöchiometrie der Komplexkomponenten zu überprüfen.

Das Cup-Protein zeigt eine sehr ähnliche Stöchiometrie zur RNA wie Smaug. Je Molekül Smaug bindet ein Molekül Cup, was mit der beschriebenen Rekrutierung von Cup durch Smaug sehr gut übereinstimmt (Nelson *et al.*, 2004).

Belle zeigt eine weniger deutliche Anreicherung. Für die Luc-RNA ist die Bindung an die MUT-RNA sogar höher als die spezifische Bindung an die WT-RNA, so dass Belle nicht mehr spezifisch gebunden wurde. Für die SREonly- und die 1-AUG-RNA zeigt sich eine substöchiometrische Bindung im Vergleich zu Smaug oder Cup. An je zwei Moleküle Smaug konnte im Komplex etwa ein Molekül Belle spezifisch gebunden werden.

Für die beiden Proteine Me31B und Trailer Hitch konnte eine deutliche Abhängigkeit der Stöchiometrie von der Länge der RNA beobachtet werden. Für die SREonly-RNA wurde eine Bindung von jeweils einem Molekül gefunden. An die 630 nt lange 1-AUG-RNA assoziierten bereits sechs bzw. acht Moleküle Me31B und Trailer Hitch spezifisch. Die Anzahl erhöht sich noch einmal auf jeweils elf Moleküle pro RNA für die Luziferase-RNA. Für alle hier diskutierten Werte wurde die Bindung an die MUT-RNA jeweils abgezogen. Die unspezifische Bindung an die RNAs stieg ebenfalls mit der Länge der RNAs, so dass z.B. für Trailer Hitch tatsächlich eine Assoziation von 22 Molekülen an Luc-RNA bestimmt wurde (die Quantifizierung von Proteinen je RNA, ohne Abzug von Werten für die unspezifische Bindung ist in Abb. 43 im Anhang S. XLVII dargestellt).

Tabelle 3: RNA-Längenabhängigkeit der Stöchiometrie der Repressor-Komplexkomponenten - Verschieden lange RNAs wurden in *pull-downs* verwendet, um eine mögliche Längenabhängigkeit der Proteinassoziation zu beobachten. Die Proteinmengen wurden auf die Menge an gebundener RNA relativiert. WT listet die Menge an Proteinen, die je Molekül 1-AUG-WT-RNA gebunden wurde. In WT-MUT wurde die Proteinmenge, die unspezifisch an die MUT-RNA gebunden wurde, abgezogen. N gibt die Anzahl an unabhängigen Experimenten an, für die die Quantifizierung durchgeführt wurde.

Protein	SREonly		1-AUG		Luc		N
	WT	WT-MUT	WT	WT-MUT	WT	WT-MUT	
Smaug	0,9	0,7	2,1	1,9	1,8	1,4	3
Tral	1,0	0,7	7,6	6,2	22,8	10,9	3
Me31B	1,1	0,7	10,2	8,7	19,1	11,6	3
Cup	0,9	0,6	3,0	2,4	4,6	1,6	5
Belle	0,5	0,2	1,3	0,6	1,6	-0,2	5

2.5. Sedimentationsverhalten reprimierter RNA

Wie in Abb. 7 gezeigt, sedimentiert SRE-haltige RNA, die in *Drosophila*-Embryoextrakt inkubiert wurde, ungewöhnlich schnell. Die reprimierte RNA kosedimentierte dabei in etwa mit monomeren Ribosomen (80S). Die bloße Assoziation von Proteinen an die RNA sollte nicht ausreichen, um eine solche Mobilität im Gradienten hervorzurufen. Die scheinbare Größe des Komplexes könnte z.B. durch Assoziation mit Ribosomen hervorgerufen werden oder auch durch Bildung eines kompakten Partikels.

2.5.1. Assoziation mit Ribosomen

Eine nicht-produktive Assoziation der RNA mit Ribosomen könnte das Sedimentationsverhalten der reprimierten RNA erklären. Tatsächlich wurde die Assoziation von *nanos*-mRNA mit Ribosomen kontrovers diskutiert. Es wurde beschrieben, dass etwa 50% der *nanos*-mRNA mit Polysomen assoziiert (Clark *et al.*, 2000; Chen *et al.*, 2014a), obwohl nur etwa 4% der RNA am posterioren Pol lokalisiert und translatiert wird. Durch Ribosomen-*footprinting* wurde aber gezeigt, dass das Ribosom nicht an die *nanos*-mRNA bindet (Kronja *et al.*, 2014; Dunn *et al.*, 2013). In Abb. 13 ist für den Smaug-RNP eine deutliche Abreicherung ribosomaler Proteine zu erkennen, was ebenfalls darauf hin deutet, dass die RNA nicht mit Ribosomen assoziiert (siehe auch Tab. 18 und Abb. 40 im Anhang ab S. XLIII).

2.5.2. Oligomerisierung der RNAs

Eine ähnliche, schnelle Sedimentation konnte auch für die *oskar*-mRNA beobachtet werden (Chekulaeva *et al.*, 2006). Die *oskar*-mRNA wird durch das Bruno-Protein translationell reprimiert. Ähnlich wie für die *nanos*-mRNA wird das Cup-Protein von Bruno rekrutiert und durch die Konkurrenz von Cup und eIF4G um eIF4E wird die Translationsinitiation inhibiert. Zusätzlich dazu wurde beobachtet, dass die reprimierten RNAs Bruno-vermittelte Oligomere bilden, die in Saccharose-Gradienten mit Ribosomen kosedimentieren (50S-80S). Diese Oligomerisierung führt unabhängig von der Bruno-Cup-eIF4E-Interaktion zur Repression, da die RNA auch in Extrakten aus *cup*^{Δ212}-Mutanten, bei denen Cup nicht mehr in der Lage ist mit eIF4E zu interagieren, reprimiert wird.

Um eine Oligomerisierung der *nanos*-mRNA zu überprüfen, wurde ein ähnliches Experiment wie in Chekulaeva *et al.* (2006) durchgeführt (Abb. 17A): Zwei, in ihrer Länge unterscheidbare, SRE-haltige RNAs wurden in Embryoextrakt inkubiert. Die kürzere der beiden RNAs war mit Biotin markiert. Alle Kombinationen von WT- und MUT-RNAs wurden miteinander inkubiert. Für eine RNA-RNA-Interaktion würde man erwarten, dass beide RNAs, die biotinylierte und die unbiotinylierte, an die Matrix binden. Für eine SRE-abhängige Oligomerisierung sollten die beiden RNAs nur dann gleichzeitig binden, wenn beide ein WT-SRE tragen.

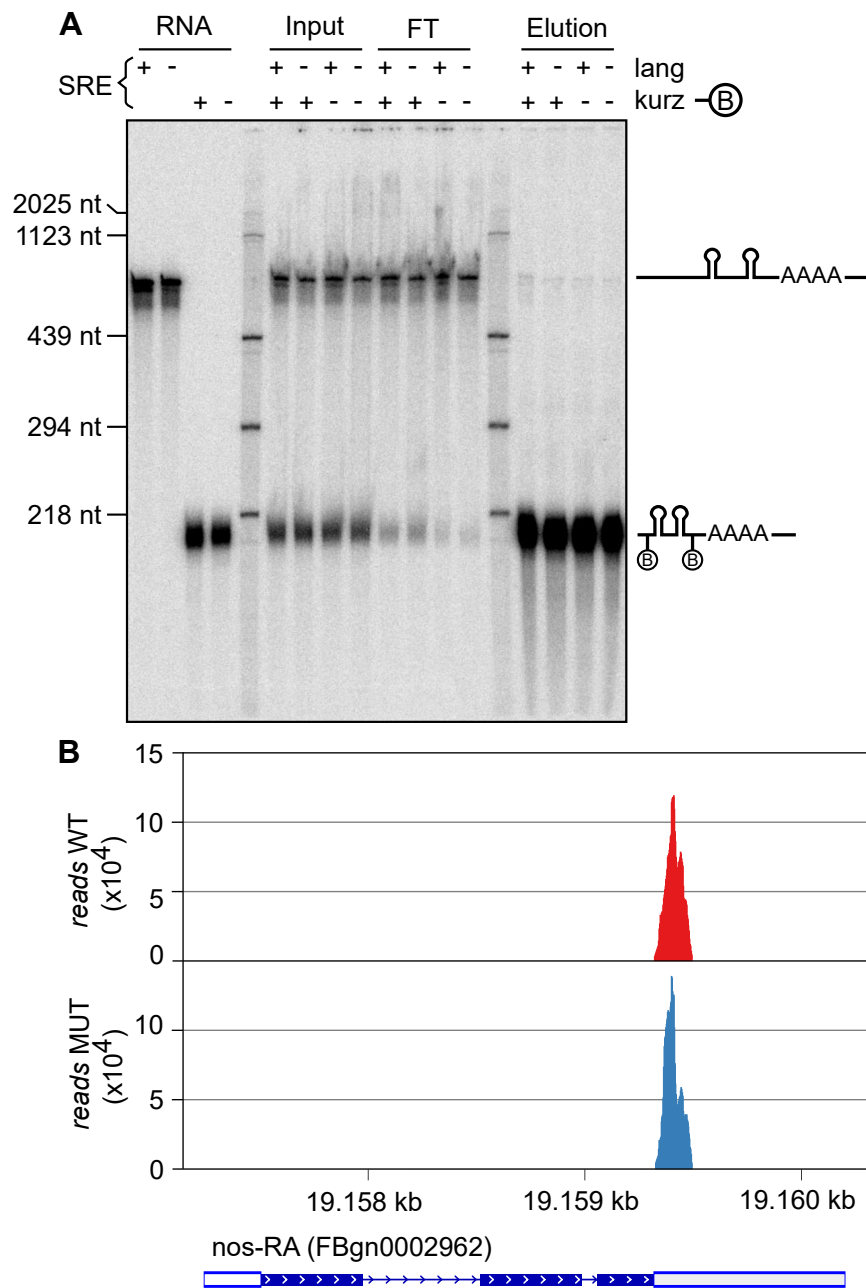


Abbildung 17: Keine Oligomerisierung SRE-haltiger RNAs - (A) Zwei SRE-haltige RNAs von unterschiedlicher Länge wurden in Embryoextrakt inkubiert. Nur die kürzere RNA war mit Biotin markiert (siehe Schema rechts). In einem Streptavidin-*pull-down* wurde nur die biotinylierte RNA präzipitiert. In den mit „RNA“ markierten Spuren wurden die eingesetzten RNAs aufgetragen. (B) RNA aus einem Streptavidin-*pull-down* wurde durch *Deep sequencing* analysiert. Es konnten *reads* auf das Genom im Bereich des *nanos*-Gens kartiert werden. Die *reads* beschränken sich dabei auf den Teil des 3'-UTRs, der in den biotinylierten RNAs enthalten ist (siehe Schema des Gens: gefüllte Box = Exon; Linie = Intron; leere Box = UTR).

In Abb. 17A sind Streptavidin-Affinitätsreinigungen für alle Kombinationen von WT und MUT in beiden RNAs dargestellt. In den ersten vier Spuren sind die eingesetzten RNAs einzeln aufgetragen. Ferner wurden RNAs aus gleichen Volumina von Input und Überstand (FT) gereinigt und aufgetragen. Die Menge der biotinylierten RNA ist im Überstand deutlich reduziert. Für die Elution wurde die Streptavidin-Sepharose in Formamid-Ladepuffer gekocht, zentrifugiert und der Überstand auf das Gel aufgetragen. Die biotinylierte RNA wurde in allen Proben gleichmäßig gebunden. Die nicht biotinylierte RNA wurde dagegen in keinem der *pull-downs* präzipitiert. Damit konnte weder eine SRE-abhängige noch eine unspezifische Oligomerisierung der Reporter-RNA im Embryoextrakt gezeigt werden. Eine Oligomerisierung der RNAs ist demnach wahrscheinlich nicht der Grund für die beobachtete, schnelle Sedimentation der RNPs.

2.5.3. Identifizierung regulatorischer RNAs

Es ist denkbar, dass neben Smaug auch regulatorische RNAs, wie miRNAs oder piRNAs an der SRE-spezifischen Regulation beteiligt sind (Rouget *et al.*, 2010). Um diese zu detektieren, wurde eine *Deep-sequencing*-Analyse eines *pull-downs* durchgeführt. Dazu wurden wiederum biotinylierte RNAs (1-AUG) mit WT- oder MUT-SREs verwendet.

Eine quantitative Darstellung aller kartierten *reads* ist in Abb. 44A im Anhang (S. XLVIII) gezeigt. Zwei miRNAs waren im WT-*pull-down* stärker als 1,5x angereichert und wiesen mehr als 1% an *reads* verglichen zur *Bait*-RNA (annotiert als *nanos-reads*) auf: mir-9b (2,9x) und mir-7 (1,8x) (Abb. 44B/C). Keine der miRNAs weist eine potentielle Bindestelle direkt am SRE auf. mir-9b kann zwischen den beiden SREs und mir-7 im offenen Leseraster der 1-AUG-RNA hybridisieren (Abb. 44B). Die Vorhersage der Bindestellen in der *Bait*-RNA erfolgte mit RNAhybrid (Abb. 44C; Krüger u. Rehmsmeier, 2006). Eine spezifische Funktion dieser RNAs für die WT-RNA ist damit eher unwahrscheinlich. Um die Anreicherung zu überprüfen, könnten *pull-downs* durchgeführt werden und die miRNAs im Anschluss z.B. mittels TaqMan[®]-PCR (Applied Biosystems) quantifiziert werden. Die erhaltenen Werte könnten dann um die *Bait*-RNA-Menge korrigiert werden.

Es wurden Effekte von piRNAs auf die SRE-abhängige Deadenylierung *in vivo* beobachtet. Diese konnten mit der durchgeführten *deep-sequencing*-Analyse nicht nachvollzogen werden, da die dafür nötigen Bindestellen außerhalb der verwendeten 3'-UTR-Sequenzen lagen (siehe Abb. 4, S. 9 und Abb. 44B, S. XLVIII).

Obwohl die Sequenzierung auf kurze RNAs abzielte, konnten dennoch *reads* auf das *nanos*-Gen kartiert werden. Die *reads* beschränkten sich dabei aber auf den Bereich des 3'-UTRs, der in den verwendeten 1-AUG-RNAs (630 nt) enthalten ist (Abb. 17B). In Kombination zeigen die vorgestellten Experimente, dass eine Oligomerisierung nicht der Grund für die

stabile Smaug-vermittelte Repression sein kann und vermutlich keine regulatorischen RNAs an der stabilen Repression beteiligt sind.

2.6. Interaktionspartner von Smaug

Es ist beschrieben, dass Smaug eine Vielzahl von maternalen Transkripten reguliert. Zwei dieser Transkripte sind dabei detaillierter analysiert: die *nanos*-mRNA und die *hsp83*-mRNA. Während die *nanos*-mRNA sowohl SRE-spezifisch deadenyliert als auch translationell reprimiert wird, kommt es bei der *hsp83*-mRNA lediglich zur Deadenylierung. Ein offensichtlicher Unterschied zwischen beiden RNAs ist die Positionierung der SREs in der RNA: während die beiden SREs in der *nanos*-mRNA im 3'-UTR liegen, befinden sich die insgesamt acht SREs ausschließlich im offenen Leseraster der *hsp83*-mRNA (Semotok *et al.*, 2005, 2008). Die Positionierung der SREs scheint aber keinen generellen Einfluss auf die Folgen der Regulation zu haben (Chen *et al.*, 2014a). Wenn die Bindung von Smaug an RNAs unterschiedliche Konsequenzen hat, ist es wahrscheinlich, dass unterschiedlich zusammengesetzte mRNPs maßgeblich für die Unterschiede in der Regulation verschiedener RNAs sind.

Um einen Einblick zu gewinnen, welche Proteine mit Smaug interagieren, wurde eine Immunpräzipitation (IP) von Smaug aus Embryoextrakt durchgeführt und ebenfalls massenspektrometrisch analysiert. Als Spezifitätskontrolle wurde zusätzlich eine IP mit Präimmunserum durchgeführt. Der Extrakt wurde nicht mit RNase behandelt, so dass auch RNA-abhängige, indirekte Interaktionen identifiziert werden können. In Vorversuchen konnte beobachtet werden, dass NonidetTM P-40 (NP-40), welches routinemäßig in Immunpräzipitationen eingesetzt wird, um unspezifische Interaktionen zu unterdrücken, auch spezifische Interaktionen (beispielsweise mit Me31B und Trailer Hitch) verhinderte. Zusätzlich konnte gezeigt werden, dass NP-40 einen negativen Einfluss auf die SRE-spezifische Deadenylierung hat (Jeske *et al.*, 2011), so dass es hier nicht verwendet wurde.

Insgesamt wurden in beiden IPs 1103 Proteine identifiziert (Abb. 18). Die Auswertung erfolgte dabei wie für Analyse des RNA-*pull-downs* (siehe Abschnitt 2.3, S. 27). 79 Proteine wiesen dabei einen p-Wert für die Anreicherung in der Smaug-IP unterhalb von 5% auf (gelistet in Tab. 19). Unter den angereicherten Proteinen befanden sich viele der im RNA-*pull-down* identifizierten Proteine. Die Überschneidung der beiden Experimente ist in Abb. 19 dargestellt. Dazu wurden die angereicherten Proteine aus beiden Experimenten (mit $p < 5\%$) verglichen. 21 der im RNA-*pull-down* identifizierten Interaktionspartner sind auch in der Smaug-IP signifikant angereichert. Von den drei beschriebenen Komplexen (Repressor-Komplex, CTLH-Komplex, Ccr4-Not-Komplex) sind nur die Proteine Belle und CG3295 mit höheren p-Werten als 5% (9% und 15%) in der Smaug-IP angereichert, alle anderen Komponenten waren in der Smaug-IP signifikant angereichert und abundant.

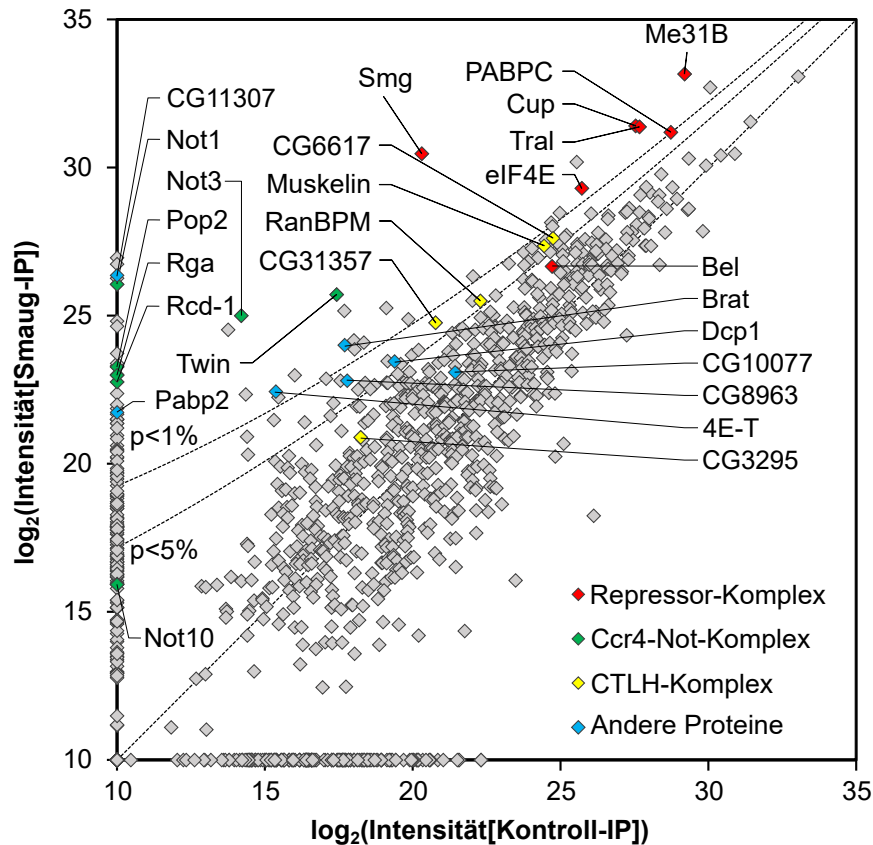


Abbildung 18: Smaug-Immunpräzipitation - Quantitative Auswertung der Smaug-Immunpräzipitation (IP). Die Smaug-IP wurde mit einer Präimmunsereum-IP verglichen, um spezifische und unspezifische Interaktionspartner voneinander unterscheiden zu können (siehe Tab. 19, Anhang S. XLIX). Die zuvor identifizierten Komplexkomponenten (Abb. 12) wurden farbig markiert. p-Werte wurden bestimmt wie auf S. 27 beschrieben und die p-Wert-Grenzen für 1% und 5% wurden eingezeichnet.

Betrachtet man die Proteine mit der höchsten Signifikanz der Anreicherung beider Analysen ($p < 1\%$), erhält man eine Liste von zwölf Proteinen: die Kernkomponenten des Repressor-Komplexes (Smaug, Me31B, Trailer Hitch, Cup, eIF4E und PABPC), einige Komponenten des CCR4/NOT-Komplexes (Not1, Not3 und Twin) sowie die Proteine 4E-T, CG11307 und Brat. Alle Proteine des CTLH-Komplexes sind mit p-Werten zwischen 5% und 1% in der Smaug-IP angereichert und deshalb nicht in dieser Liste enthalten.

Um weitere Informationen zu Smaug-interagierenden Proteinen zu erhalten, könnte die Immunpräzipitation mit und ohne RNase-Behandlung analysiert werden. Direkt Interaktionspartner könnten dann *in vitro* durch *Western-blot*-Analysen verifiziert werden, wie z.B. das Ypsilon-Schachtel-Protein, welches ebenfalls ein maternales Protein ist, welches in granulären Strukturen in Zellen gefunden wird.

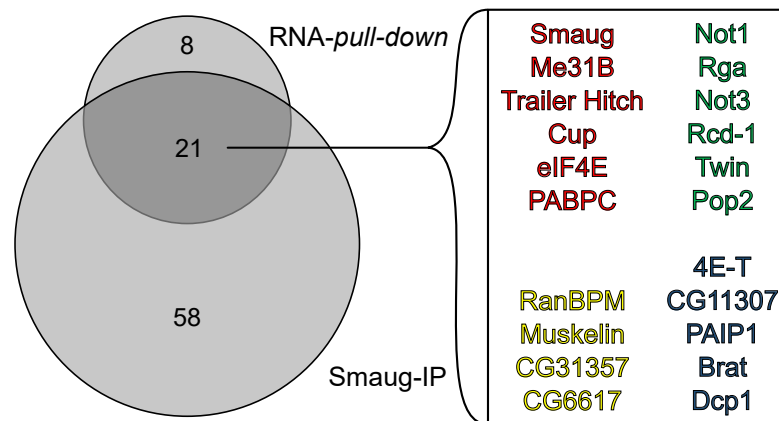


Abbildung 19: Vergleich der angereicherten Proteine im RNA-pull-down und der Smaug-IP - 21 Proteine konnten in beiden Analysen identifiziert werden und sind rechts farblich hervorgehoben. Von den drei beschriebenen Komplexen (Repressor-Komplex, Ccr4-Not-Komplex und CTLH-Komplex) sind lediglich drei Proteine nicht signifikant angereichert. Belle und CG3295 sind mit höheren p-Werten als 5% (9% und 15%) in der Smaug-IP angereichert. Not10 lag nicht angereichert vor.

2.7. *Cross-Linking* der angereicherten Proteine

In einer weiteren Smaug-Immunpräzipitation wurden die angereicherten Proteine chemisch quervernetzt, um mögliche Interaktionsstellen zu kartieren. Dabei stellt die Auswertung der Daten das größte Problem dar. Die Komplexität der Probe erlaubt theoretisch eine Vielzahl an möglichen Interaktionen, die, wenn möglich, auf den Aminosäurerest genau analysiert werden sollen. Für dieses bioinformatische Problem wurden die Programme StavroX und MeroX entwickelt, auf die im zweiten Teil dieser Dissertation näher eingegangen werden soll (Ab S. 71).

In dem Datensatz konnten zwei *Cross-Links* identifiziert werden: ein intermolekularer *Cross-Link* zwischen Me31B und Trailer Hitch und ein intramolekularer *Cross-Link* in PABPC. Die *Cross-Links* konnten anhand von bereits gelösten Protein-Strukturen validiert werden. Für die Interaktion von Me31B und Trailer Hitch wurde die Kristallstruktur des DDX6-FDF Dimers verwendet (Tritschler *et al.*, 2009). DDX6 ist das humane Homologe von Me31B und bindet die Orthologe von Edc3 und Trailer Hitch an derselben Bindestelle (FDF-Peptid in Edc3 und Trailer Hitch). Der identifizierte *Cross-Link* erstreckt sich zwischen Lysin 412 von Me31B zu Lysin 429 in Trailer Hitch (Abb. 20A). Nach Sequenzvergleichen entsprechen diese Aminosäuren den Positionen von E451 in DDX6 und I222 im FDF-Peptid von Edc3. Der C $_{\alpha}$ -C $_{\alpha}$ -Abstand dieser beiden Aminosäuren beträgt etwa 13 Å und liegt damit in sehr guter Übereinstimmung mit der erwarteten Distanz vernetzbarer Aminosäuren für den Urea-*Cross-Linker* (max. 27 Å).

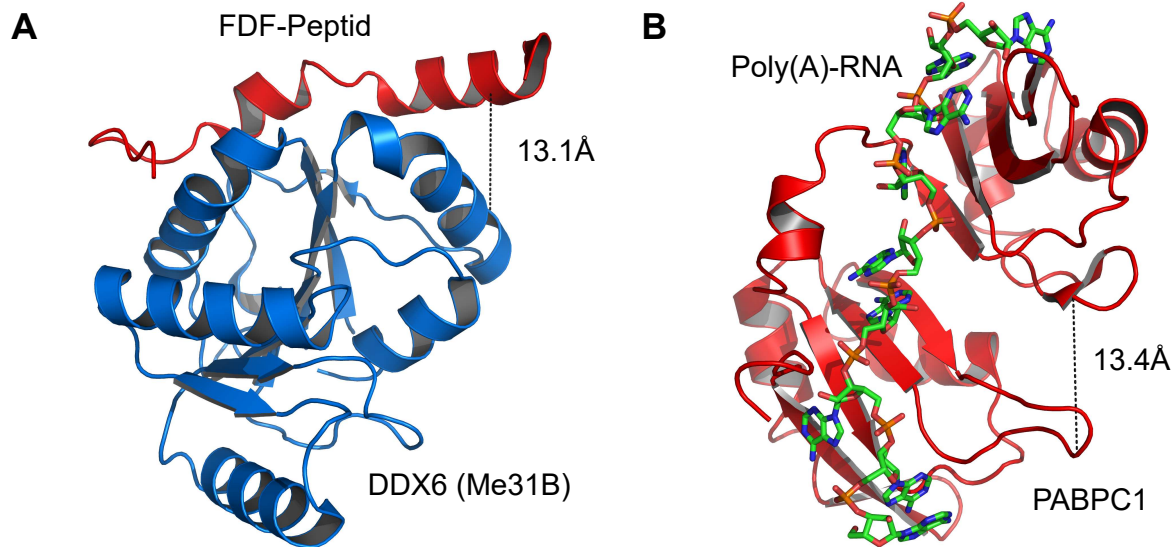


Abbildung 20: Cross-Linking-Analyse der Immunpräzipitation - Nach der Smaug-Immunpräzipitation wurde ein chemischer *Cross-Linker* zu der Matrix gegeben, um interagierende Proteine zu vernetzen und anschließend massenspektrometrisch analysiert. Zwei *Cross-Links* konnten identifiziert werden. (A) Der *Cross-Link* zwischen Me31B und Trailer Hitch wurde auf die 3D Struktur von DDX6 (blau) und dem FDF-Peptid (rot) von EDC3 (PDB-ID:2WAX) übertragen. Der C $_{\alpha}$ -C $_{\alpha}$ -Abstand der verknüpften Aminosäuren wurde zu 13,1 Å bestimmt. Das Fragmentationenspektrum zu diesem *Cross-Link* ist als Beispiel für die Annotation durch MeroX in Abb. 35 auf S. 111 dargestellt und erläutert. (B) Ein intramolekularer *Cross-Link* im PABPC wurde auf die Struktur des humanen PABPC1 (PDB-ID:1CVJ) übertragen und weist einen C $_{\alpha}$ -C $_{\alpha}$ -Abstand von 13,4 Å auf. Zusätzlich zum PABPC (rot) ist eine gebundene Poly(A)-RNA als Stabmodell gezeigt. Die Darstellung und Abstandsmessung erfolgte mit PyMol (DeLano, 2006).

Innerhalb des zytosolischen Poly(A)-Bindepoteins konnte ebenfalls ein *Cross-Link* identifiziert werden, der auf die Struktur des humanen Orthologs PABPC1 übertragen wurde (Deo *et al.*, 1999). Die vernetzten Aminosäuren (K71 und K126) entsprechen den Positionen von K80 und N135 im humanen Ortholog. Der C $_{\alpha}$ -C $_{\alpha}$ -Abstand dieser beiden Aminosäuren beträgt erneut etwa 13 Å und liegt damit innerhalb des erwarteten Distanzbereiches.

Alle drei Proteine liegen vermutlich auch in der Immunpräzipitation in einem Überschuss über Smaug vor, so dass die Identifizierung von *Cross-Links* mit Me31B, Trailer Hitch und PABPC durchaus zu erwarten war. Insgesamt wurden aber sehr wenige *Cross-Links* identifiziert. Etwa 0,3% aller Spektren (insgesamt 262 von 82851 Spektren) enthielten eine für den verwendeten *Cross-Linker* spezifische Signatur (siehe Abschnitt 1.2.3, S. 82). Die möglichen *Cross-Links* verteilen sich dabei vermutlich auf die große Zahl an identifizierten Proteinen, so dass nur wenige *Cross-Links* mit Repressor-Komplexkomponenten überhaupt vorhanden sind. Die restlichen Spektren ohne die *Cross-Link*-spezifische Signatur entspringen vermutlich größtenteils Peptiden, die nicht quervernetzt wurden.

3. Diskussion

Bei der Analyse des SRE-abhängigen mRNPs konnten etwa zwei Dutzend Proteine identifiziert werden, die wahrscheinlich an der Regulation der *nanos*-mRNA *in vivo* beteiligt sind (siehe Abb. 12 und Tab. 1, S. 27/30). Anhand dieser Liste von Proteinen kann nun eine detaillierte Analyse der molekularen Details dieser Regulation erfolgen.

Eines der wichtigsten Ziele ist, die stabile Repression der *nanos*-mRNA *in vitro* zu rekonstituieren. Dies hätte den Vorteil, dass die Rolle einzelner Faktoren durch einfache Manipulationen der Zusammensetzung des Komplexes aufgeschlüsselt werden könnte. Die identifizierten, spezifisch angereicherten Proteine sollten ausreichend sein, um den reprimierten RNP zu bilden. Die beteiligten Proteine sind zum Teil gut konserviert. Eine Rekonstitution kann demnach nicht nur Aufschluss über die *nanos*-Regulation geben, sondern auch ein Modell liefern, welches auf andere regulatorische Mechanismen in anderen Organismen, auch im Menschen, übertragbar sein kann. Die folgende Diskussion soll die einzelnen Faktoren und deren mögliche Rolle in der Regulation der *nanos*-mRNA beleuchten und einen minimalen Komplex postulieren, der für die Regulation nötig ist.

3.1. Subkomplexe im SRE-abhängigen RNP

Es konnten drei Subkomplexe unter den angereicherten Proteinen ausgemacht werden: der Repressor-Komplex mit den sieben abundantesten, spezifisch angereicherten Proteinen, der Ccr4-Not-Komplex und der CTLH-Komplex.

Die Rekrutierung des Ccr4-Not-Komplexes durch Smaug wurde bereits mehrfach beschrieben und charakterisiert (Semotok *et al.*, 2005; Jeske *et al.*, 2006; Zaessinger *et al.*, 2006). Alle Kernuntereinheiten des Ccr4-Not-Komplexes wurden mit ähnlichen Anreicherungsfaktoren (durchschnittlich 50x) identifiziert. Die Proteine eines weiteren konservierten Komplexes – des CTLH-Komplexes – wiesen ebenfalls ähnliche Anreicherungsfaktoren auf (durchschnittlich 5x). Diese Konsistenz innerhalb der identifizierten Komplexe spricht für die Qualität der durchgeführten Analyse.

Mehr als 70% der identifizierten Proteine konnten auch in einer unabhängigen Smaug-Immunpräzipitation (siehe Abb. 18/19, S. 45/46) als Interaktionspartner des Smaug-RNPs

identifiziert werden und somit als mögliche Komponenten des regulierenden Komplexes bestätigt werden. Von den Proteinen der drei Subkomplexe konnten sogar 90% in der Smaug-IP wiedergefunden werden.

Im Anschluss sollen mögliche Rollen der identifizierten Subkomplexe und einzelner Komponenten des Repressor-Komplexes sowie mögliche Funktionen weiterer identifizierter Proteine in der Regulation der *nanos*-mRNA diskutiert werden.

3.1.1. Repressor-Komplex

Die *nanos*-mRNA wird durch die Bindung von Smaug translationell reprimiert. Dazu wird ein stabiler Komplex an die RNA rekrutiert. Bei durchgeführten Experimenten blieb dieser auch während einer mehrstündigen Gradientenzentrifugation stabil und aktiv (siehe Abb. 10, S. 24). Das bedeutet, dass die identifizierten Komponenten Bestandteile eines repressionsaktiven Komplexes sind. Da die Translation und damit auch der reprimierte Zustand des gereinigten RNPs nur durch eine erneute Inkubation in Embryoextrakt gemessen werden kann, ist es möglich, dass weitere essentielle Faktoren, die hier nicht identifiziert wurden, mit dem Komplex assoziieren. Da aber keine weitere Vorinkubation nötig ist, um das volle Ausmaß der Repression zu erreichen, müssen alle Faktoren, die während des geschwindigkeitsbestimmenden Schrittes in der ersten Vorinkubation assoziieren, im vorgereinigten Komplex vorliegen. Die identifizierten Proteine gehören demnach zum stabilen Komplex, der an SRE-Strukturen in RNAs bindet.

Die zentralen Komponenten des Repressor-Komplexes sind vermutlich Smaug, Cup, Belle, Me31B und Trailer Hitch. Das Protein mit der höchsten Signifikanz der Anreicherung im RNA-*pull-down* war Smaug. Die verwendeten RNAs enthielten zwei Smaug-Bindestellen, die in der Negativkontrolle so mutiert wurden, dass die Bindung von Smaug an die RNA stark reduziert wurde. Die sehr spezifische Bindung von Smaug an die verwendete SRE-haltige RNA konnte bereits gezeigt (Jeske *et al.*, 2011) und auch in *Western-blot*-Analysen bestätigt werden. Smaug ist die zentrale Komponente der *nanos*-Regulation in der Embryogenese von *Drosophila melanogaster*, die dem gebundenen Repressor-Komplex die Spezifität für RNA-Substrate verleiht. Die zentrale Rolle von Smaug in dieser Regulation ist mit biochemischen und genetischen Daten belegt (Smibert *et al.*, 1996; Dahanukar *et al.*, 1999; Semotok *et al.*, 2005; Jeske *et al.*, 2011) und konnte auch in den hier gezeigten Analysen bestätigt werden. Die Repression einer Reporter-RNA in Embryoextrakten ist auf wenige fmol/ μ l begrenzt. Die meisten Proteine des Repressor-Komplexes weisen eine deutlich höhere Konzentration auf (siehe Tab. 2, S. 37). Smaug ist mit ≈ 80 nM im Extrakt und ≈ 30 nM in Translationsexperimenten nur in einem geringen Überschuss über der Reporter-RNA enthalten (etwa 3x bei 10 nM RNA). Smaug bindet stöchiometrisch an die SREs (siehe Abb. 16, S. 39), so dass der Überschuss an Smaug bei einer RNA mit zwei SREs auf etwa 1,5x reduziert wird. Die

Kapazitätsgrenze der Translationsrepression von etwa 10 nM RNA spricht dafür, dass der Extrakt hauptsächlich freies Smaug enthält, das nicht in einem kinetisch stabilen Komplex gebunden ist. Es wurde gezeigt, dass Smaug eine Vielzahl von Transkripten reguliert oder regulieren kann (Tadros *et al.*, 2007; Chen *et al.*, 2014a). Das lässt sich nur schwer mit der Konzentration von Smaug im Extrakt erklären. Möglicherweise ist stabil gebundenes Smaug ähnlich wie die Reporter-RNAs zu kompakten Partikeln verpackt (siehe Abb. 7 S. 20) und während der Zentrifugationsschritte der Embryoextrakt-Präparation abgetrennt worden.

Die RNA-Bindedomäne von Smaug bindet an das eIF4E-bindende Protein Cup, so dass eine direkte Verbindung von Smaug zum Translationsinitiationsfaktor eIF4E hergestellt ist (siehe Abb. 5, S. 10). eIF4E-bindende Proteine (4EBPs) wie Cup spielen eine wichtige Rolle in der Regulation der Translation von Eukaryoten (Richter u. Sonenberg, 2005). So konnte eine Beteiligung von Cup an der Bruno-abhängigen Regulation der *oskar*-mRNA gezeigt werden (Chekulaeva *et al.*, 2006). Cup trägt auch zur Repression der *nanos*-mRNA bei, indem eIF4E gebunden und die Bindung von eIF4E an eIF4G verhindert wird (Nelson *et al.*, 2004). Sowohl für die *oskar*- als auch die *nanos*-mRNA konnten allerdings Hinweise auf Cup-unabhängige Repressionsmechanismen gezeigt werden (Chekulaeva *et al.*, 2006; Jeske *et al.*, 2011).

Im RNA-*pull-down* konnten die Annahmen des 4EBP-Modells der Repression durch Cup dennoch bestätigt werden. Cup und eIF4E interagieren spezifisch mit SRE-haltiger RNA, während eIF4G abgereichert ist. Ob Cup tatsächlich essentiell für die stabile Repression ist, kann in Embryoextrakten nur schwer nachvollzogen werden. Mutationen im *cup*-Gen führen zu verschiedenen Defekten früh in der Embryogenese, so dass Effekte auf die *nanos*-Regulation in *cup*-Mutanten keine eindeutigen Aussagen zulassen (Nelson *et al.*, 2004).

Das zytoplasmatische Poly(A)-Bindeprotein (PABPC) war ebenfalls signifikant im SRE-abhängigen RNP angereichert, obwohl die beiden verglichenen *Bait*-RNAs einen gleichlangen Poly(A)-Schwanz trugen. Diese spezifische Anreicherung könnte eine Rolle von PABPC in der Repression nahelegen. Tatsächlich ist die Repression *in vitro* zum Teil abhängig von der Anwesenheit eines Poly(A)-Schwanzes in der RNA (Jeske *et al.*, 2011).

Die möglichen Rollen der weiteren Repressor-Komplexproteine Belle, Me31B und Trailer Hitch werden ausführlich in den Abschnitten 3.2 und 3.3 diskutiert.

3.1.2. Ccr4-Not-Komplex

Alle Kernkomponenten des Ccr4-Not-Komplexes waren mit dem SRE-spezifischen RNP assoziiert. Die Proteine waren ähnlich stark angereichert wie Smaug ($\approx 50\times$). Der Ccr4-Not-Komplex kann unabhängig von dessen Deadenylase-Aktivität eine Translationsrepression

vermitteln (Cooke *et al.*, 2010; Waghay *et al.*, 2015). Die beschriebene Repression der Ccr4-Not-Komponente Pop2 ist allerdings abhängig von der Interaktion zwischen Not1 und den Me31B-Homologen Xp54 bzw. DDX6 (Waghay *et al.*, 2015; Kuzuoğlu-Öztürk *et al.*, 2016). Im Vergleich zu den Komponenten des Repressor-Komplexes wiesen die Abundanzwerte aller Ccr4-Not-Komplex-Untereinheiten zudem geringere Werte auf. Eine vom Repressor-Komplex unabhängige Beteiligung des Ccr4-Not-Komplexes am geschwindigkeitsbestimmenden Schritt der Repression ist unwahrscheinlich. Einerseits hängt die Ccr4-Not-vermittelte Reaktion von Repressor-Komplexkomponenten ab (Mathys *et al.*, 2014) und andererseits sollte eine substöchiometrische Assoziation des Komplexes an RNA nicht ausreichend sein, um die hohen Repressionsfaktoren zu erreichen.

Tabelle 4: Not1-assoziierte Proteine des SRE-RNPs - Liste der Proteine, die in einer Not1-IP (Temme *et al.*, 2010) identifiziert wurden und im RNA-*pull-down* angereichert waren. Die Anzahl an identifizierten PSM (*peptide spectrum matches*) je Protein wurde für die Not1-IP zur Abschätzung der Proteinabundanz herangezogen. Die Proteine sind nach den beschriebenen Komplexen sortiert.

Repressor-Komplex	PSM	Ccr4-Not-Komplex	PSM	CTLH-Komplex und andere	PSM
Smg	32	Not1	100	Brat	27
Cup	26	Not3	40	CG6617 (CTLH)	5
Me31B	24	Twin	33	RanBPM (CTLH)	4
Tral	22	Not10	27	CG8963	4
PABPC	19	Rcd-1	18	CG31357 (CTLH)	1
		Rga	18	CG11307	1
		Pop2	12		

Vergleicht man die in der vorliegenden Arbeit gezeigten Analyse des SRE-RNPs mit einer Not1-Immunpräzipitation (Temme *et al.*, 2010), ergibt sich ebenfalls eine recht große Überlappung der Datensätze: 18 von 29 im RNA-*pull-down* angereicherten Proteine wurden auch in einer Not1-IP identifiziert (Tab. 4). Interessant ist dabei, dass die Proteine des Repressor-Komplexes (Smaug, Cup, PABPC, Me31B und Trailer Hitch) in der Not1-IP ähnlich abundant wie die Komponenten des Ccr4-Not-Komplexes waren. Die ebenfalls enthaltenen Proteine des CTLH-Komplexes waren mit 1 - 5 PSM eher substöchiometrisch enthalten. Gleiches gilt für CG11307, welches aber sowohl im RNA-*pull-down*, der Smaug-IP als auch der Not1-IP identifiziert wurde. Ebenso kam der Translationsrepressor Brat in allen drei Analyse angereichert vor. Die Anreicherung der Proteine CG11307 und Brat wird in Abschnitt 3.1.4 genauer diskutiert.

3.1.3. CTLH-Komplex

Der CTLH-Komplex wurde zunächst in Hefe als GID-Komplex (*glucose-induced degradation deficient*) identifiziert und ist für den Abbau eines Enzyms der Glukoneogenese (FBPase) verantwortlich, wenn die Zellen auf ein glukosehaltiges Medium überführt werden und die Glykolyse aktiviert wird (Regelmann *et al.*, 2003). Dabei wirkt der CTLH-Komplex als E3-Ubiquitin-Ligase mit zwei aktiven Ligase-Untereinheiten (GID2 und GID9) mit degenerierten RING-Finger-Domänen (Santt *et al.*, 2008; Braun *et al.*, 2011). Der CTLH-Komplex ist in Eukaryoten gut konserviert (Francis *et al.*, 2013) und auch das *Xenopus*-Homolog von GID2 (RMND5) ist eine aktive E3-Ubiquitin-Ligase (Pfirrmann *et al.*, 2015).

Fünf der sechs konservierten Proteine des CTLH-Komplexes in *Drosophila* konnten als Bestandteil des SRE-RNPs identifiziert werden (siehe Abb. 12, S. 27). Mit Ausnahme von CG3295 war der CTLH-Komplex auch in der Smaug-IP angereichert (siehe Abb. 18 und 19, S. 45/46).

Es fällt auf, dass beinahe alle Proteine des Repressor-Komplexes (Smaug, Cup, Me31B, Tral, eIF4E und PABPC) während der MZT stark degradiert werden (Gouw *et al.*, 2009). Für Smaug ist der zeitliche Verlauf der Proteinexpression im Detail untersucht (Dahanukar *et al.*, 1999; Benoit *et al.*, 2009). Die Synthese von Smaug wird zu Beginn der Embryogenese initiiert und die Proteinmenge steigt bis zum Maximum während der 14. Kernteilung an. Danach wird das Protein rapide degradiert. Nach 4 - 5 Stunden ist Smaug nicht mehr zu detektieren. Die Tatsache, dass Proteine, die spezifisch abgebaut werden, zusammen mit dem CTLH-Komplex, als E3-Ubiquitin-Ligase, in einem Komplex vorkommen, lässt die Annahme zu, dass dieser für den spezifischen Abbau des Repressor-Komplexes während der MZT relevant ist.

In Übereinstimmung mit dieser Annahme konnte Ubiquitin mit einer ähnlich hohen Abundanz wie der Repressor-Komplex identifiziert werden (hier annotiert als RpS27A, eines der Ubiquitin-Fusionsproteine aus *Drosophila melanogaster*). Außerdem konnten, ohne dass der proteasomale Abbau blockiert wurde, Ubiquitinierungsstellen in Me31B und Cup identifiziert werden.

Ubiquitinierung von Me31B

Für Me31B konnte ein modifiziertes Peptid identifiziert werden, welches auf eine Ubiquitinierung des Proteins hinweist. Nach einer tryptischen Proteolyse eines ubiquitinierten Peptids verbleiben die beiden C-terminalen Glycine des Ubiquitins kovalent an einem Lysin im Peptid verknüpft (Diglycin-Modifizierung). Die gezeigte Ubiquitinierungsstelle im Me31B (siehe Abb. 14, S. 33) liegt innerhalb der ersten RecA-Domäne, zwischen Motiv Ia

und Ib, welche an der RNA-Bindung von DEAD-Box-Helikasen beteiligt sind (Linder u. Jankowsky, 2011; Weston u. Sommerville, 2006). Die Proben für die massenspektrometrische Analyse wurden mit Iodacetamid behandelt, um die Bildung von Disulfiden zwischen cysteinhaltigen Peptiden zu verhindern. Dabei werden die Cysteine zunächst reduziert und anschließend alkyliert. Iodacetamid kann neben Cysteinen aber auch Lysine modifizieren. Eine doppelte Alkylierung eines Lysins ist massenspektrometrisch nur schwer von einer Diglycin-Modifizierung unterscheidbar (Nielsen *et al.*, 2008; Kim *et al.*, 2015). In der Analyse des RNA-*pull-downs* konnten viele Peptide mit einer Modifizierung von 114,04 u identifiziert werden. Viele dieser Peptide konnten aber auch mit einer Modifikation von 57,02 u identifiziert werden, welche einer einfachen Alkylierung einer Lysin-Seitenkette durch Iodacetamid entspricht und nicht nach einer vorherigen Ubiquitinierung beobachtet wird. Für diese Peptide sollten keine Rückschlüsse auf mögliche Modifizierung durch Ubiquitin getroffen werden, da die Modifizierung auch ein Iodacetamid-induziertes Artefakt sein kann.

Für die Ubiquitinierung von Me31B wurde allerdings zusätzlich zur Diglycin-Modifizierung eine LRGG-Modifizierung (383,23 u) an derselben Stelle im Protein identifiziert (siehe Abb. 41, Anhang S. XLV). Diese Modifizierung tritt auf, wenn bei der Proteolyse die letzte Spaltstelle im Ubiquitin nicht gespalten wird. Die LRGG-Modifizierung kann kein Iodacetamid-induziertes Artefakt sein und zeigt eindeutig eine Ubiquitinierung des Peptids. Um die Ubiquitinierung der Faktoren im Detail massenspektrometrisch zu untersuchen, sollte einerseits das Proteasom inhibiert werden (z.B. mit Lactacystin oder MG132; Lee u. Goldberg, 1998), um eine größere Ausbeute ubiquitinierten Proteine zu erhalten und andererseits sollten andere Alkylierungsagenzien eingesetzt werden, die spezifisch Cysteine alkylieren und nicht Lysine (z.B. Chloracetamid; Nielsen *et al.*, 2008). Dann ließe sich die mögliche Ubiquitinierung des Repressor-Komplexes durch den CTLH-Komplex massenspektrometrisch zeigen (z.B. in Embryoextrakten aus Mutanten des CTLH-Komplexes). Wenn der CTLH-Komplex relevant für den Abbau der Repressor-Komplexproteine ist, würde man in Mutanten der CTLH-Komplexproteine zudem eine Stabilisierung maternaler Proteine, wie z.B. Smaug, über die 14. Kernteilung hinweg erwarten, die in Embryonen z.B. durch *western blots* oder Antikörperfärbung beobachtet werden könnte. Eine Verzögerung der Zellularisierung wäre auch denkbar, wenn der Abbau maternaler Proteine inhibiert würde.

3.1.4. Weitere identifizierte Proteine

Zusätzlich zu den Proteinen, die den drei Komplexen zugeordnet werden konnten, wurden im RNA-*pull-down* zehn weitere Proteine signifikant angereichert. Alle zehn Proteine wiesen in der massenspektrometrischen Analyse eine geringere Intensität als der Repressor-Komplex auf. Fünf der Proteine zeigten einen p-Wert unterhalb von einem Prozent (4E-T, Dcp1, Brat, CG11307 und CG8963). Genau dieselben fünf Proteine waren auch in der gezeigten Smaug-Immunpräzipitation signifikant angereichert ($p < 5\%$), während die anderen fünf

Proteine (Loqs, Nsun5, Pabp2, CG10077, Oskar) in der Smaug-IP nicht angereichert waren oder nicht identifiziert wurden. Eine Beteiligung dieser zehn Proteine an der stabilen Repression ist aufgrund der substöchiometrischen Menge im RNA-*pull-down* eher unwahrscheinlich.

Das Cup-Paralog, 4E-T, konnte vermutlich aufgrund der spezifischen Interaktion mit Me31B (siehe Abb. 6, S. 16) bzw. Trailer Hitch im reprimierten RNP identifiziert werden. Die Interaktion dieser drei Proteine wurde mit den homologen Proteinen aus *Xenopus* und Mensch mehrfach gezeigt (Waghray *et al.*, 2015; Nishimura *et al.*, 2015; Ozgur *et al.*, 2015; Kamenska *et al.*, 2016). 4E-T könnte mit Cup um die Präsenz im Repressor-Komplex kompetieren. Es wäre sogar denkbar, dass 4E-T die Funktion von Cup im Repressor-Komplex übernehmen kann, auch wenn es dafür noch keine Evidenz gibt.

Das *decapping*-Protein Dcp1 interagiert mit Trailer Hitch an dessen LSM-Domäne (Tritschler *et al.*, 2008). Das Hefe-Homolog von Me31B, Dhh1p, stimuliert das *decapping* von mRNAs (Fischer u. Weis, 2002), so dass eine Anreicherung von Dcp1 durch die Interaktion mit dem Me31B·Tral-Heterodimer im Repressor-Komplex wahrscheinlich ist. In Übereinstimmung mit den hier gezeigten Daten wurden Smaug und Dcp1 zusammen in granulären Strukturen im Zytoplasma von Embryonen beobachtet (Zaessinger *et al.*, 2006). Die Cap-Struktur von mRNAs wird in *Drosophila*-Embryoextrakt hauptsächlich nach erfolgtem 3'-5'-Abbau durch DcpS abgebaut und nicht durch *decapping* und anschließenden 5'-3'-Abbau (Buschmann *et al.*, 2013). Der *decapping*-abhängige Abbau verschiedener maternalen Transkripte beginnt etwa zwei Stunden nach der Befruchtung des Embryos (Lin *et al.*, 2006) und überlappt damit kaum mit dem Alter der Embryonen, die für die gezeigten Experimente verwendet wurden. In Übereinstimmung damit konnte die aktive Untereinheit des *decapping*-Komplexes, Dcp2, weder im RNA-*pull-down* noch in der Smaug-IP identifiziert werden, auch nicht als unspezifisch gebundenes Protein.

Das Protein CG8963 kann durch Sequenzvergleiche als das *Drosophila*-Homolog von PAIP1 (Poly(A)-Bindeprotein-interagierendes Protein) identifiziert werden und interagiert wahrscheinlich mit dem Smaug-reprimierten RNP durch die spezifische Bindung an PABPC (Craig *et al.*, 1998). Da PABPC sowohl in dem RNA-*pull-down* als auch in der Smaug-IP angereichert war, ist auch die Anreicherung von CG8963 in beiden Experimenten eine logische Folge. Im RNA-*pull-down* konnte zusätzlich das nukleäre Poly(A)-Bindeprotein (Pabp2) identifiziert werden. In *Drosophila*-Embryonen kommt Pabp2 im Zytoplasma vor und besitzt eine Rolle in der Verkürzung der Poly(A)-Schwänze spezifischer Transkripte (Benoit *et al.*, 2005). Pabp2 konkurriert im Zytoplasma mit PABPC um die Bindung des Poly(A)-Schwanzes, so dass eine Bindung von Pabp2 an den Poly(A)-Schwanz in den verwendeten *Bait*-RNAs wahrscheinlich zu dessen Identifizierung führte.

Brat (Brain Tumor) bindet an eine definierte Sequenz im 3'-UTR von RNAs. Die *nanos*-mRNA ist nicht unter den identifizierten Ziel-Transkripten des Brat-Proteins (Loedige *et al.*, 2015; Laver *et al.*, 2015). Sie besitzt zwar eine potentielle Bindestelle im 3'-UTR (Nukleotide 270 - 275: UUGUUA), dieser folgt allerdings keine U-reiche Sequenz. Außerdem war diese Sequenz nicht in den verwendeten *Bait*-RNAs enthalten. Eine Anreicherung von Brat erfolgte demnach vermutlich nicht durch eine direkte RNA-Bindung. Es ist eher wahrscheinlich, dass Brat durch eine Interaktion mit dem Ccr4-Not-Komplex (Tab. 4, S. 52) in den MS-Analysen identifiziert wurde.

Über das CG11307-Protein, dass in beiden MS-Analysen dieser Arbeit und in der Not1-IP identifiziert wurde, ist nur wenig bekannt. CG11307 enthält eine *domain of unknown function* (DUF4745) und ist als mögliche Glykosyltransferase annotiert (FlyBase.org). Das Protein ist lediglich innerhalb des Insektenreiches konserviert (NCBI-blastp). Das CG11307-Transkript ist am stärksten in Oozyten und während der ersten vier Stunden der Embryogenese exprimiert (FlyBase.org, modENCODE; Gelbart u. Emmert, 2013) und könnte damit durchaus eine Rolle in der frühen Entwicklung von *Drosophila* spielen. Im RNA-*pull-down* wurde CG11307 mit einer ähnlichen Abundanz, Signifikanz und Anreicherung wie die Proteine des Ccr4-Not-Komplexes identifiziert. Dies könnte eine mögliche Interaktion von CG11307 mit dem Ccr4-Not-Komplex nahelegen, auch wenn in der Not1-IP nur ein Spektrum detektiert wurde, das von CG11307 stammt (Tab. 4).

Mit geringerer Signifikanz wurden noch die Proteine Oskar und CG10077 angereichert. Oskar hebt die Smaug-abhängige Repression der *nanos*-mRNA am posterioren Pol auf und kann einerseits an Smaug binden (Dahanukar *et al.*, 1999) und andererseits auch an die *nanos*-mRNA (Jeske *et al.*, 2015). Eine Interaktion mit Smaug oder der *Bait*-RNA ist also wahrscheinlich der Grund für die signifikante Anreicherung von Oskar. CG10077 ist, wie auch Belle und Me31B, eine DEAD-Box-Helikase. Die Interaktion von Smaug mit CG10077 wurde in einer TAP-Smaug Reinigung bereits beobachtet (Pinder, 2012).

Für die Proteine Nsun5 und Loquacious (Loqs) konnte kein Zusammenhang mit anderen Komponenten des RNPs hergestellt werden. Nsun5 wurde mit der geringsten Abundanz aller angereicherten Proteine identifiziert und spielt deshalb vermutlich keine Rolle im stabil-reprimierten RNP. Loqs ist beteiligt an der miRNA-Biogenese (Jakob *et al.*, 2016). Da kein weiterer Faktor angereichert wurde, der mit miRNAs assoziiert oder an deren Biogenese beteiligt ist, ist anzunehmen, dass es sich bei Loqs nicht um eine relevante Komponente des reprimierten RNPs handelt.

3.2. Die Rolle von Belle im Repressor-Komplex

Das neu identifizierte Protein Belle ist eine DEAD-Box-Helikase der Ded1p/DDX3-Familie. Das Protein ist in Eukaryoten konserviert, wobei der zentrale DEAD-Box-Helikasekern eine höhere Konservierung zeigt als die Termini der Proteine (Sharma u. Jankowsky, 2014). Belle ist mit einer Länge von 801 Aminosäuren eines der größten Proteine der Ded1p/DDX3-Familie. Der Größenunterschied beruht hauptsächlich auf Insertionen im N-Terminus des Proteins. Dieser enthält zwei konservierte Elemente, denen eine Funktion zugeschrieben werden kann: Ein Kernexportsignal (Askjaer *et al.*, 1999) und ein eIF4E-Bindemotiv, wie es auch in Cup oder 4E-T vorkommt (YxxxxLΦ; Shih *et al.*, 2008; Yarunin *et al.*, 2011; Gao *et al.*, 2016).

Die Deletion des Hefe-Homologs *ded1* ist letal (Chang *et al.*, 1990) und führt zu einer globalen Verringerung der Translation, da Ded1p für die Initiation der Translation benötigt wird (De La Cruz *et al.*, 1997). Die *ded1*-Mutation kann durch die Expression von Ded1p und seine Homologe komplementiert werden (Chuang *et al.*, 1997; Senissar *et al.*, 2014). Diese Experimente legen eine mögliche Rolle der Homologe in der Translationsinitiation nahe. Die Komplementation durch Belle war dabei allerdings am wenigsten effektiv.

Eine Interaktion mit Komponenten des eukaryotischer Initiationsfaktoren wie eIF3 und eIF4F wurde für einige der Belle-Homologe gezeigt (Lee *et al.*, 2008; Hilliker *et al.*, 2011; Soto-Rifo *et al.*, 2012, 2013; Senissar *et al.*, 2014; Gao *et al.*, 2016). Dies legt ebenfalls eine Funktion in der Regulation der Translation nahe. Tatsächlich konnte gezeigt werden, dass Belle-Homologe translationsstimulierende (Beckham *et al.*, 2008; Geissler *et al.*, 2012; Ihry *et al.*, 2012), aber auch translationsreprimierende (Shih *et al.*, 2008; Worringer *et al.*, 2009; Yarunin *et al.*, 2011) Aktivitäten aufweisen können. Die Effekte wurden dabei häufig durch globale Änderungen in der Translationsaktivität beobachtet. Nur in wenigen Fällen wurde die Regulation einer bestimmten RNA gezeigt. Eine solche spezifische Regulation wurde z.B. für die Repression der *bruno*-Translation diskutiert (Yarunin *et al.*, 2011). Die translationsaktivierende und -reprimierende Funktion wird von unterschiedlichen Teilen des Proteins vermittelt (Hilliker *et al.*, 2011). So werden im Hefe-Homolog Ded1p die eIF4E-Bindestelle und Sequenzen aus dem C-Terminus für die Vermittlung der Repression benötigt und der Helikasekern sowie die katalytische Aktivität der Helikase sind essentiell für die Aktivierung der Translation. Dabei bewirkt Ded1p zusammen mit eIF4A – einer weiteren DEAD-Box-Helikase im eIF4F-Komplex – die Entwindung strukturierter RNA-Bereiche.

Ded1p ist biochemisch sehr gut charakterisiert. So konnten die ATPase- und Helikaseaktivität sowie die RNA-Bindung *in vitro* gezeigt werden (Iost *et al.*, 1999; Yang u. Jankowsky, 2005). Ded1p ist in der Lage, RNA-Protein-Komplexe zu remodellieren (Bowers *et al.*, 2006) und bildet in der Gegenwart von bestimmten ATP-Analogen äußerst langlebige Komplexe

(Halbwertszeit von über 40 h; Liu *et al.*, 2014). Ded1p funktioniert dabei *in vitro* und *in vivo* als Trimer (Putnam *et al.*, 2015), wobei die einzelnen Proteine mithilfe ihrer C-terminale Domäne miteinander interagieren. Die Oligomerisierung anderer Homologe wurde bislang nicht gezeigt.

Für das Belle-Homolog Laf-1 aus *C. elegans* wurde *in vitro* die Bildung von flüssigen Tropfen (*liquid droplets*) gezeigt (Elbaum-Garfinkle *et al.*, 2015). Dabei kommt es zur Bildung einer Phasengrenze zwischen der Lösung und einem proteinhaltigen Tropfen. Laf-1 ist ein essentieller Bestandteil der *P-granules* in der *C. elegans*-Keimbahn. Die erhöhte Viskosität dieser zytoplasmatischen Partikel und auch anderer Partikel wie *P-bodies* könnte dabei auf eine ähnliche Phasentrennung zurückzuführen sein (Jonas u. Izaurralde, 2013). Intrinsisch unstrukturierte Proteinbereiche sind an dieser Phasentrennung maßgeblich beteiligt. Der N-Terminus in Laf-1 ist intrinsisch unstrukturiert und reicht aus, um *liquid droplets* zu bilden (Elbaum-Garfinkle *et al.*, 2015). Der vorhergesagte, intrinsisch unstrukturierte Bereich im N-Terminus von Belle ist 50% länger als in Laf-1 (Xue *et al.*, 2010; siehe Abb. 46 im Anhang, S. LIII). Ob diese Laf-1-vermittelte Phasentrennung auch durch Belle vermittelt werden kann und ob diese eine Rolle in der Regulation der *nanos*-mRNA spielt, ist unklar.

Im *Drosophila*-Embryo ist Belle hauptsächlich zytoplasmatisch lokalisiert und über den gesamten Embryo gleichmäßig verteilt. Auch eine Lokalisierung in granulären Strukturen wie z.B. *P-bodies* oder ähnlichen Strukturen in Neuronen wurde gezeigt (Beckham *et al.*, 2008; Hilliker *et al.*, 2011). Ded1p interagiert mit Dhh1p, dem Me31B-Homolog, welches ebenfalls in *P-bodies* vorkommt (Tseng-Rogenski *et al.*, 2003; Drummond *et al.*, 2011; Senissar *et al.*, 2014).

In dieser Arbeit wurde eine mögliche Beteiligung von Belle an der Regulation der *nanos*-mRNA durch die spezifische Anreicherung im SRE-abhängigen RNP beschrieben (siehe Abb. 12, S. 27). Auch *in vivo* ist Belle an der *nanos*-Regulation beteiligt. Embryonen von Müttern mit Mutationen im *belle*-Gen zeigten deutliche Defekte in der Deadenylierung als auch in der Translationsrepression der *nanos*-mRNA (in Kooperation mit Martine Simonelig; Götze *et al.*, eingereicht).

Es ist allerdings möglich, dass Belle nicht am geschwindigkeitsbestimmenden Schritt der Repression beteiligt ist. Die geringe Anreicherung von Belle, die aus den MS-Daten zu erkennen war (1,8x), bestätigte sich in quantitativen *Western-blot*-Analysen von *pull-down*-Experimenten. Die Belle-Menge im Komplex war im Vergleich zu Smaug oder Cup sogar substöchiometrisch (siehe Abb. 16, S. 39). Zusätzlich war keine Anreicherung von Belle in der Smaug-IP zu beobachten. Für einen stabil assoziierten, essentiellen Bestandteil des Repressor-Komplex würde man wenigstens eine stöchiometrische Bindung relativ zur RNA erwarten, da ansonsten die starke Repression der RNA ($\approx 20\times$) nicht zustande käme.

Welche Funktion Belle in der Regulation übernimmt, bleibt noch unklar. Rekonstitutionsexperimente der *nanos*-Regulation sollten Aufschluss darüber geben, ob Belle für die Assemblierung des stabilen Repressor-Komplexes benötigt wird oder ob Belle für einen anderen Schritt in der Regulation nötig ist.

3.3. Me31B und Trailer Hitch bedecken die reprimierte RNA

Die Quantifizierung der Repressor-Komplexproteine zeigte eine deutliche Abhängigkeit der gebundenen Me31B- und Trailer-Hitch-Menge von der Länge der eingesetzten RNA (siehe Abb. 16 und Tab. 3, S. 39/40). Es ist anzunehmen, dass Me31B und Trailer Hitch ein Heterodimer bilden. Dementsprechend binden beide Proteine in der Quantifizierung stets in einem äquimolaren Verhältnis zueinander. Zudem konnte die Interaktion der beiden Proteine *in vitro* gezeigt werden (Kopräzipitation durchgeführt von Dr. Christiane Rammelt; Götze *et al.*, eingereicht). Während an einer etwa 200 nt langen RNA nur jeweils ein Me31B•Tral-Komplex gebunden war, konnten bis zu zwanzig Me31B•Tral-Heterodimere auf einer etwa 2000 nt langen RNA gebunden werden. Dieser Umstand lässt ein Modell zu, bei dem eine RNA durch die multiple Bindung des Me31B•Tral-Heterodimers maskiert wird, so dass sie schwer zugänglich für Translationsinitiationsfaktoren oder das Ribosom wird (*masking*; Spirin, 1994). Verschiedene Aspekte deuten auf eine Rolle dieser Oligomerisierung in der Translationsrepression der *nanos*-mRNA hin:

- **Konservierte Interaktion von Me31B und Trailer Hitch.**

In den meisten Eukaryoten konnten Homologe von Me31B und Trailer Hitch identifiziert werden, die dann auch als Interaktionspartner beschrieben wurden (Boag *et al.*, 2005; Audhya *et al.*, 2005; Wilhelm *et al.*, 2005; Tanaka *et al.*, 2006; Tritschler *et al.*, 2009; Ayache *et al.*, 2015; Bish *et al.*, 2015). Beide Proteine kommen in granulären Strukturen wie *P-bodies*, *stress granules* oder granulären Strukturen in speziellen Zelltypen wie der Keimbahn (z.B. *P-granules*) oder Nervenzellen vor (Nakamura *et al.*, 2001; Tseng-Rogenski *et al.*, 2003; Audhya *et al.*, 2005; Souquere *et al.*, 2009; Naarmann *et al.*, 2010; Hillebrand *et al.*, 2010; Hubstenberger *et al.*, 2013; Zampedri *et al.*, 2016). Dies spricht für eine allgemeine Rolle von Me31B und Trailer Hitch in der Translationsrepression bzw. dem mRNA-Abbau, die diese granulären Struktur häufig innehaben.

- **Me31B und Trailer Hitch sind Repressoren der Translation**

Sowohl Me31B als auch Trailer Hitch wurden in verschiedenen Organismen als Translationsrepressoren beschrieben (Ladomery *et al.*, 1997; Lieb *et al.*, 1998; Nakamura *et al.*, 2001; Minshall *et al.*, 2001; Collier u. Parker, 2005; Chu u. Rana, 2006; Marnef *et al.*, 2009; Nissan *et al.*, 2010; Carroll *et al.*, 2011; Presnyak u. Collier, 2013). Dabei

können sowohl globale Effekte auf die Translation nach Überexpression, als auch Effekte auf Reporter-Transkripte beobachtet werden, an die eines der Proteine gekoppelt wurde (*tethering assay*). Das Trailer-Hitch-Homologe Scd6p aus Hefe ist zum Beispiel in der Lage, einen Schritt vor der 48S-Komplexbildung zu inhibieren, wie es auch für die *nanos*-Repression beschrieben wurde (Nissan *et al.*, 2010; Jeske *et al.*, 2011).

- **Clamping durch DEAD-Box-Helikasen**

DEAD-Box-Helikasen sind in der Lage, sehr langlebige Komplexe mit RNA einzugehen. Für dieses sogenannte *clamping* wird die RNA-Bindung der Helikase durch die Verhinderung der Adenosin-Nukleotid-Dissoziation im aktiven Zentrum so stabilisiert, dass der RNA-Protein-Komplex sehr langsam dissoziiert (Linder u. Jankowsky, 2011; Liu *et al.*, 2014; Xiol *et al.*, 2014). Dies konnte *in vitro* mit ATP-Analogen oder Protein-Mutanten gezeigt werden und wurde schon zuvor in einem biologisch relevanten Kontext, am Beispiel der DEAD-Box-Helikase eIF4AIII beschrieben (Ballut *et al.*, 2005; Andersen *et al.*, 2006; Bono *et al.*, 2006). Dieser Modus der RNA-Bindung könnte sowohl die ATP-Abhängigkeit der Komplexbildung als auch die hohe kinetische Stabilität des Komplexes erklären.

- **RNA-Bindung von Me31B**

Einige Me31B-Homologe binden mit hoher Affinität, aber ohne Sequenzspezifität an RNA (Dutta *et al.*, 2011; Ernoult-Lange *et al.*, 2012). Die RNA-Bindung von Trailer Hitch und homologen Proteinen wurde bislang nur anhand von UV-*Cross-Linking* untersucht (Lieb *et al.*, 1998; Tanaka *et al.*, 2006). Kokristalle von DEAD-Box-Helikasen mit RNA zeigen eine Bedeckung von nur wenigen Nukleotiden (z.B. für eIF4AIII; Bono *et al.*, 2006). Für eine hohe Affinität der RNA-Bindung (K_D im niedrigen nanomolaren Bereich) von Dhh1p werden etwa 10 - 12 Nukleotide benötigt (Dutta *et al.*, 2011). Für die in dieser Arbeit gemessene Stöchiometrie von RNA zu Protein müsste ein Me31B•Tral-Heterodimer für eine vollständige Bedeckung der RNA etwa 100 nt binden. Bei einer massenspektrometrischen Analyse von UV-*Cross-Linking*-Experimenten mit Dhh1p (Sharif *et al.*, 2013), dem Hefe-Homolog von Me31B, wurden *Cross-Link*-Stellen über einen großen Bereich der RecA2-Domäne beobachtet, die eine Bindung von deutlich mehr als zehn Nukleotiden je Molekül zulassen würde. Wenn diese vergrößerte Bindefläche auch nicht ausreichen sollte, um 100 nt zu bedecken, ist es dennoch gut vorstellbar, dass die Bedeckung von großen Teilen der RNA bereits die Assoziation von Ribosomen mit der RNA behindert. Für die Quantifizierung der Proteinmengen im Repressor-Komplex wurden Streptavidin-*pull-downs* mit anschließender *Western-blot*-Analyse durchgeführt. Dabei wurde die Streptavidin-Sepharosematrix mehrfach gewaschen, um unspezifisch gebundene Proteine zu entfernen. Es kann demnach sein, dass die beobachtete Menge von Me31B und Trailer Hitch im Komplex noch unterschätzt wird und damit auch die Bedeckungslänge je Heterodimer (≈ 100 nt) überschätzt wird.

- **RNA-abhängige Oligomerisierung von Me31B und homologen Proteinen**

Hinweise auf die Oligomerisierung von Me31B-Homologe wurden in *Xenopus*-Oozyten und menschlichen Zelllinien (Xp54 und DDX6) beobachtet (Minshall u. Standart, 2004; Ernoult-Lange *et al.*, 2012). Die Bildung der Oligomere ist dabei zum Teil abhängig von RNA (Minshall u. Standart, 2004). Diese Oligomerisierung konnte z.B. durch Immungold-Markierung von DDX6 innerhalb granulärer Strukturen in elektronenmikroskopischen Schnitten von Zellen beobachtet werden (Ernoult-Lange *et al.*, 2012).

Die hier gelisteten Eigenschaften von DEAD-Box-Helikasen und speziell der Me31B-Homologe sind in sehr guter Übereinstimmung mit der Oligomerisierung von Me31B•Tral an der *nanos*-mRNA. Die Spezifität der RNA-Bindung muss von Smaug vermittelt werden, welches als Nukleationskeim für die Bindung aller weiteren Komponenten dient.

3.3.1. Sedimentationsverhalten als Folge der Me31B•Tral-Oligomerisierung

Bei der Analyse des SRE-abhängigen mRNPs wurde die Bildung sehr schnell sedimentierender Komplexe beobachtet, die in einer Gradientenzentrifugation zu einem deutlichen Sedimentationsunterschied zwischen WT- und MUT-RNA führte (siehe Abb. 7, S. 20). Ähnliche Eigenschaften wurden auch für die *nanos*-mRNA *in vitro* beobachtet und es wurde eine nicht-produktive Interaktion der *nanos*-mRNA mit Ribosomen postuliert (Clark *et al.*, 2000). Die signifikante Abreicherung von ribosomalen Proteinen in der massenspektrometrischen Analyse des RNA-*pull-downs* spricht jedoch gegen diese Hypothese (siehe Abb. 13, S. 29 sowie Tab. 18 und Abb. 40 im Anhang ab S. XLIII). Zudem wurde mit Polysomen-*profiling*- und Ribosomen-*footprinting*-Experimenten gezeigt, dass die *nanos*-mRNA im Embryo nicht an Ribosomen gebunden ist (Dunn *et al.*, 2013; Kronja *et al.*, 2014).

Für die *oskar*-mRNA konnte, nach einer ähnlichen Beobachtung schnell sedimentierender Partikel, eine mRNA-Oligomerisierung, vermittelt durch die Proteine Bruno und PTB, gezeigt werden (Chekulaeva *et al.*, 2006; Besse *et al.*, 2009). In der vorliegenden Arbeit konnte in verschiedenen *pull-down*-Experimenten allerdings keine Oligomerisierung SRE-haltiger RNAs nachgewiesen werden, so dass die Reporter-RNA-Oligomerisierung wahrscheinlich nicht die Ursache der beobachteten schnellen Sedimentation ist (siehe Abb. 17, S. 42). Durch Einzelmolekülanalysen wurde in zwei unabhängigen Arbeiten zudem gezeigt, dass unlokalisierte *nanos*-mRNA in Partikeln mit etwa einer RNA je Partikel vorliegt, was mit den hier gezeigten Daten gut übereinstimmt (Little *et al.*, 2015; Trcek *et al.*, 2015).

Die multiple Bindung von Me31B und Trailer Hitch an die Reporter-RNA könnte den Unterschied im Sedimentationsverhalten der WT- und MUT-RNA erklären. Es ist gezeigt,

dass Me31B-Homologe RNA-abhängig oligomerisieren können und zur Bildung hochmolekularer Komplexe beitragen (Minshall u. Standart, 2004; Ernoult-Lange *et al.*, 2012). Ob die vielfache Bindung von Me31B•Tral ausreicht, die hohe Mobilität des Komplexes in der Gradientenzentrifugation zu erklären, bleibt noch unklar, stellt aber eine sehr wahrscheinliche Möglichkeit dar, die weiter verfolgt werden soll.

Eine ähnliche Beobachtung konnte in Extrakten aus *Drosophila*-Embryonen für miR2-reprimierte RNAs gemacht werden. Die miR2-abhängig gebildeten Komplexe wurden aufgrund ihrer Dichte und der Abwesenheit von Ribosomen als Pseudopolysomen beschrieben (Thermann u. Hentze, 2007). Me31B ist an der *nanos*-Repression (Jeske *et al.*, 2011) und Homologe von Me31B sind an der miRNA-abhängigen Repression beteiligt (Chu u. Rana, 2006; Mathys *et al.*, 2014; Kamenska *et al.*, 2016; Kuzuoğlu-Öztürk *et al.*, 2016). Die Bildung hochmolekularer Komplexe könnte demnach eine Eigenschaft der Me31B-Homologe sein, die zur Translationsrepression gebundener RNAs beiträgt und damit möglicherweise auch zur miRNA-abhängigen Regulation, wie z.B. im Fall der Bildung von Pseudopolysomen.

3.4. Modell der *nanos*-mRNA-Regulation während der frühen Embryogenese

Verschiedene Modelle zur Regulation der *nanos*-mRNA-Translation und -Stabilität wurden postuliert. Dabei wird z.B. angenommen, dass die Initiation der Translation entweder durch Cup oder Ago1 beeinträchtigt (Nelson *et al.*, 2004; Pinder u. Smibert, 2012) oder die Repression durch die Blockade eines späteren Schrittes in der Translation bewirkt wird (Clark *et al.*, 2000).

Die Zusammensetzung des stabilen Smaug-RNP legt die Beteiligung von Cup an der Repression (Nelson *et al.*, 2004) nahe. Die Abwesenheit von Ago1 vom stabil reprimierten mRNP macht eine Beteiligung von Ago1 an der Smaug-vermittelten, stabilen Repression der *nanos*-mRNA unwahrscheinlich. Auch konnte die Kopräzipitation von Argonaut-Proteinen (Ago1 - 3, Aubergine) mit Smaug oder dem *nanos*-mRNP (Pinder u. Smibert, 2012; Rouget *et al.*, 2010) mit den hier gezeigten Experimenten nicht rekapituliert werden. Da die potentiellen miRNA- und piRNA-Bindestellen der *nanos*-mRNA nicht in den verwendeten Reporter-Konstrukten enthalten waren, ist eine Beteiligung von regulatorischen RNAs und assoziierten Proteinen mit den hier gezeigten Analysen aber auch nicht auszuschließen.

Das hier postulierte Modell für die Repression basiert auf der Rekrutierung verschiedener Proteine an die *nanos*-mRNA durch Smaug und ist in Abb. 21 dargestellt. Dabei umfasst die Abbildung drei verschiedene Mechanismen: (I) Die unlokalisierte *nanos*-mRNA wird im Embryo spezifisch vom Ccr4-Not-Komplex deadenyliert; damit wird die Translationseffizienz

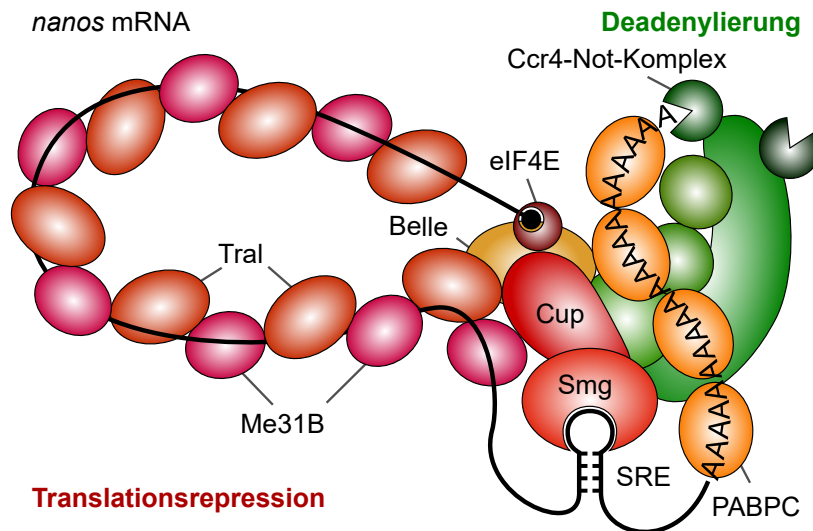


Abbildung 21: Modell des Repressor-Komplexes - schematische Darstellung der reprimierten *nanos*-mRNA während der frühen Embryogenese von *Drosophila melanogaster*. Die SREs im 3'-UTR werden von Smaug gebunden, das spezifisch sowohl die Deadenylierung durch den Ccr4-Not-Komplex einleitet, als auch die translationale Stilllegung der RNA bewirkt. Dieses Modell der Repression erweitert das Cup-Modell (Nelson *et al.*, 2004) um die Proteine Belle, Me31B und Trailer Hitch. Das gezeigte Modell postuliert, dass eine Bedeckung der RNA durch Me31B und Trailer Hitch zentral für die starke Repression der *nanos*-mRNA ist.

reduziert. Die Rekrutierung des Deadenylase-Komplexes konnte durch die Anreicherung aller Kernkomponenten des Ccr4-Not-Komplexes überzeugend demonstriert werden. (II) Cup verhindert durch die Inhibierung der eIF4F-Komplexbildung die Cap-abhängige Translationsinitiation. (III) Die Oligomerisierung von Me31B und Trailer Hitch führt zur Maskierung der RNA, so dass diese unzugänglich für das Ribosom oder Translationsinitiationsfaktoren wird. Diese drei verschiedenen Ebenen der Regulation zusammen bewirken dann die starke Repression der unlokalisierten *nanos*-mRNA, die nötig ist, um eine ektopische Expression von Nanos zu verhindern.

Die Anordnung und Stöchiometrie der Proteine in Abb. 21 beruht dabei auf den Daten dieser Arbeit und bekannten Interaktionen zwischen den identifizierten Proteinen des Repressor-Komplexes. Die einzelnen beschriebenen Interaktionen sind in Abb. 45 und Tab. 20 im Anhang auf S. LII zusammengefasst.

3.4.1. *Cis*-Interaktion des 5'- und 3'-Endes einer mRNA

Das hier vorgestellte Modell kann auch eine Erklärung für einen bisher wenig untersuchten Sachverhalt liefern: Wie können Sequenzelemente im 3'-UTR das zum Teil mehrere Tausend Nukleotide entfernte 5'-Ende derselben RNA beeinflussen? Häufig enthalten 3'-UTR-Bereiche Sequenzmotive, die die Translationseffizienz, z.B. durch Modulation der

eIF4F-Komplexbildung, manipulieren. Diese Interaktion muss in *cis* geschehen, da sonst RNAs reguliert werden können, die das regulierende Element nicht enthalten. Für die *oskar*-mRNA konnte eine Regulation in *trans* beobachtet werden (Macdonald *et al.*, 2016), so dass eine ausschließliche Regulation von RNAs in *cis* nicht gegeben sein muss.

Eine Möglichkeit, Interaktionen auf andere mRNA-Spezies zu verhindern, ist, die RNA als einziges RNA-Molekül oder eine homogene RNA-Population in einem reprimierten Partikel zu lokalisieren. Dadurch kann die Repression nur in *cis* oder zwischen gleichartigen RNAs stattfinden. Tatsächlich befindet sich die *nanos*-mRNA entweder als einzelnes RNA-Molekül in einem Partikel (unlokalisierter Anteil) oder in homotypischen Partikeln im Polplasma (Little *et al.*, 2015; Trcek *et al.*, 2015).

Alternativ dazu kann die Oligomerisierung eines Faktors entlang der RNA eine direkte Verbindung zwischen den Enden der mRNAs herstellen. Dadurch kann nur die gebundene RNA reguliert werden und nicht eine andere RNA in *trans*. Zudem könnte die Oligomerisierung auch eine räumliche Nähe des 3'- und 5'-Endes herstellen. Eine Verpackung einzelner RNA-Moleküle kann demnach sicherstellen, dass die Regulation ausschließlich in *cis* wirkt.

3.4.2. *Cross-Linking*/MS-Analyse des Smaug-RNP

Um mögliche Interaktionen innerhalb des Smaug-RNPs zu kartieren, wurde eine *Cross-Linking*-Analyse durchgeführt, die mittels Massenspektrometrie analysiert wurde.

Es konnten zwei verschiedene *Cross-Links* identifiziert werden (siehe Abb. 20, S. 47). Im Vergleich zur Anzahl an identifizierten Proteinen ist diese Zahl recht gering. Eine Ursache für die geringe Ausbeute kann natürlich die geringe Menge an tatsächlich vorhandenen *Cross-Links* in der Analyse sein. Da das *Cross-Linking* und die anschließende Proteolyse direkt an der Affinitätsmatrix erfolgte, wurde die Effektivität der *Cross-Link*-Reaktion nicht überprüft. Bei der bioinformatischen Analyse der *Cross-Links* wurden außerdem ausschließlich die Proteine des Repressor-Komplexes betrachtet. Die Smaug-Immunpräzipitation enthielt aber eine Vielzahl von Proteinen, die ebenfalls mit dem *Cross-Linker* reagieren konnten, so dass nur eine kleine Zahl an theoretisch möglichen *Cross-Links* tatsächlich in Betracht gezogen wurde. Die beiden identifizierten *Cross-Links* wurden zudem zwischen den drei superstöchiometrischen Komponenten des Repressor-Komplexes (Me31B, Trailer Hitch und PABPC; siehe Abb. 21) identifiziert und stellen wahrscheinlich die abundantesten Interaktionsstellen dar, so dass die geringe Anzahl identifizierter *Cross-Links* wahrscheinlich eine Frage der Abundanz war.

Um die Anzahl an identifizierbaren *Cross-Links* zu erhöhen, sollte eine andere Reinigungsmethode für den Smaug-RNP in Betracht gezogen werden, da z.B. der verwendete Antikörper

sowie unspezifisch gebundene Proteine um den *Cross-Linker* konkurrieren und die Analyse erschweren. Smaug kann z.B. mit kurzen, synthetischen und biotinylierten RNAs, die lediglich ein SRE-*stem-loop* enthalten, spezifisch angereichert werden (nicht gezeigt). Die kurze RNA kann nur wenige unspezifische Interaktionen vermitteln, wodurch weniger unspezifisch angereicherte Proteine während des *Cross-Linking* vorhanden sein sollten. Zudem wird kein Antikörperserum benötigt, welches zudem unspezifisch bindende Proteine enthalten kann.

3.5. Ausblick

Um die Smaug-abhängige Regulation der *nanos*-mRNA zu verstehen, soll die Rolle der einzelnen Proteine im Repressor-Komplex weiter im Detail untersucht werden. Konzeptionell ist die Zugabe von rekombinanten Proteinen zum Extrakt mit anschließender Translationsmessung eine einfache Möglichkeit, den Einfluss eines Proteins oder einer Proteinvariante auf die Repression zu testen. Die Hoffnung wäre, dass z.B. eine zugegebene Mutante einen dominant negativen Effekt auf die Translation hat oder die Zugabe eines limitierenden Faktors (vermutlich Smaug) die Kapazität eines Extraktes für die Translationsrepression erhöht.

Diese Art von Experimenten kann jedoch zu einigen Problemen führen: Die Translation in *Drosophila*-Embryoextrakt ist anfällig für Änderungen im Puffer. So führen z.B. Detergenzien zu einer Reduktion der Translationsaktivität des Extraktes und die Repression nicht mehr beobachtet werden kann. Zudem sind einige Proteine in hohen Konzentrationen im Extrakt vorhanden (z.B. Me31B und Trailer Hitch), so dass vermutlich eine große Menge an Protein zugegeben werden muss, um signifikante Effekte zu beobachten. Die Unterschiede in der Translation sind dann nur schwer mit den einzelnen zugegebenen Proteine in Korrelation zu bringen. Die Immundepletion eines Extraktes führt vermutlich zur Kodepletion von Repressor-Komplexkomponenten und kann auch die Translationseffizienz beeinflussen, so dass auch hier eine Aussage schwierig ist.

Um z.B. die Rolle von Belle in der Repression zu definieren, könnte versucht werden, ausreichende Mengen eines Extraktes aus Embryonen von *bel*-Mutanten zu gewinnen, um *In-vitro*-Translationsexperimente durchzuführen. Der Extrakt könnte dann z.B. mit rekombinantem Protein komplementiert werden, um die Funktion einzelner Bereiche (z.B. des eIF4E-Bindungsmotivs oder des gesamten unstrukturierten N-Terminus) des Proteins zu untersuchen. Ebenso könnten Extrakte von Mutanten anderer Gene hergestellt werden, um deren Rolle zu analysieren – wie z.B. für *cup*^{Δ212}, einer Mutante des Cup-Proteins, bei der der N-Terminus von Cup inklusive der eIF4E-Bindestelle deletiert ist (Nakamura *et al.*, 2004). Mutanten der Proteine des Repressor-Komplexes zeigen allerdings häufig noch vor der Embryogenese Defekte, die die Interpretation der Experimente erschweren können.

Um einigen der beschriebenen Problemen zu entgehen, müsste zunächst in einem modifizierten Extrakt oder aus gereinigten Komponenten der Repressor-Komplex auf einer RNA assembliert werden. Da der Repressor-Komplex eine Gradientenzentrifugation übersteht, bietet sich diese als Reinigungsmethode an, wenn auch nur wenige Proben gleichzeitig analysiert werden können. Im Anschluss können die Gradientenfraktionen auf Translationsrepression in unbehandeltem WT-Extrakt getestet werden. Da sich die Zusammensetzung der Fraktionen dann nur noch in den zugegebenen Komponenten unterscheidet, sollte dieser experimentelle Ansatz vergleichbare, interpretierbare Ergebnisse liefern.

Nach dem vorgestellten Modell sollten Smaug, Me31B und Trailer Hitch ausreichen, um die RNA in einen stabilen, schnell sedimentierenden Komplex zu verpacken. Dies sollte *in vitro* leicht zu überprüfen sein, indem man die gereinigten Komponenten zusammen mit 1-AUG-RNA (WT oder MUT) und ATP inkubiert und anschließend die Verteilung der RNA in einer Gradientenzentrifugation untersucht. Sollte sich ein solcher Komplex *in vitro* bilden, könnte man diesen mit Elektronenmikroskopie sichtbar machen (z.B. mithilfe von GrafiX-Gradienten; Kastner *et al.*, 2008). Fraktionen einer solchen Gradientenzentrifugation mit Luc-RNA-Komplexen könnten zudem mit Embryoextrakt gemischt werden und auf Translationsrepression getestet werden.

Die Rekonstitution der Repression aus rekombinanten Faktoren stellt die beste Möglichkeit dar, die Rolle einzelner Faktoren in der Translationsrepression zu überprüfen. Es sollte versucht werden, alle hier identifizierten Komponenten des Repressor-Komplexes rekombinant darzustellen, um die Rekonstitution zu bewerkstelligen. Dabei sollte ebenso vorgegangen werden wie zuvor beschrieben und eine Reinigung gebildeter Komplexe mittels Gradientenzentrifugation durchgeführt werden, bevor die Repression gemessen wird.

Neben der Smaug-abhängigen Translationsrepression ist auch die Analyse der Smaug-abhängigen Deadenylierung von Interesse. Es ist unklar, ob der Deadenylase-Komplex direkt von Smaug rekrutiert werden kann oder ob dafür eine überbrückende Interaktion durch Cup oder Me31B benötigt wird. Die MS-Analyse legt eine sehr spezifische Interaktion von Smaug und dem Ccr4-Not-Komplex nahe. Die einzelnen Komponenten des Ccr4-Not-Komplexes sind für die baculovirale Expression verfügbar, so dass auch eine Rekonstitution der Deadenylierung angestrebt werden sollte. So könnte auch eine mögliche deadenylierungsunabhängige Rolle des Ccr4-Not-Komplexes in der *nanos*-Repression überprüft werden.

Neben der Deadenylierung und der Repression ist es noch weitestgehend unklar, wie die Derepression der *nanos*-mRNA am posterioren Pol des Embryos funktioniert. Der stabile Repressor-Komplex müsste von der RNA gelöst werden, um eine Translation der *nanos*-mRNA zu ermöglichen. Rekombinantes Oskar allein ist nicht in der Lage, einen

einmal gebildeten Repressor-Komplex aufzulösen (Jeske *et al.*, 2011). Möglicherweise werden dazu weitere Proteine der posterioren Gruppe benötigt wie z.B. Vasa oder Tudor (Ephrussi u. Lehmann, 1992). Diese Proteine könnten rekombinant hergestellt und deren Einfluss auf die Komplexbildung oder Stabilität des Komplexes untersucht werden. Alternativ ist es möglich, dass Oskar lediglich die Assoziation von Smaug mit der RNA am posterioren Pol verhindert und auch *in vivo* nicht in der Lage ist, einen gebildeten Smaug-RNP aufzulösen. Das würde implizieren, dass der zeitliche Ablauf der Expression von Oskar und Smaug essentiell für die Translation von Nanos am posterioren Pol ist. Um dies zu überprüfen, könnte man versuchen einen Fliegenstamm zu generieren, der Smaug während der Oogenese exprimiert – vor Stadium acht der Oogenese, wenn die Oskar-Synthese am posterioren Pol einsetzt; Kim-Ha *et al.*, 1995. Wenn die Nanos-Synthese auf der zeitlich abgestimmten Expression von Oskar und Smaug beruht, sollte es nicht zur Synthese von Nanos kommen, wenn die zeitliche Reihenfolge der Expression der beiden Proteine verändert wird.

Die Ergebnisse dieser Arbeit liefern ein Modell, mit dessen Hilfe die posttranskriptionale Regulation einer spezifischen mRNA besser nachvollzogen werden kann. Die Rekonstitution des Systems kann nun Aufschluss über die Rollen der einzelnen Komponenten liefern, um die Regulation der *nanos*-mRNA und möglicherweise auch anderer reprimierter mRNAs in verschiedenen Modellorganismen detaillierter zu verstehen.

Teil II.

Bioinformatische Analyse von
Cross-Linking/MS-Analysen

1. Einleitung

Die Protein-Strukturaufklärung spielt eine wichtige Rolle in der Entschlüsselung biologischer und biochemischer Prozesse. Die meisten Proteinstrukturen wurden mittels Röntgen-Kristallografie und NMR gelöst. Mit diesen Methoden kann allerdings nur ein Teil aller Proteine erfasst werden. Einige Proteine können nicht stark genug aufkonzentriert werden oder bilden keine Kristalle. Außerdem können dynamische Änderungen der Proteinstruktur nur schwer mit der Röntgenkristallografie nachvollzogen werden. Protein-Strukturinformationen können alternativ z.B. durch die Kryo-Elektronenmikroskopie, *small-angle X-ray scattering* (SAXS) oder Methoden der Massenspektrometrie (MS) gewonnen werden. Die Kombination verschiedener Ansätze kann dabei ähnlich akkurate Modelle liefern, wie es die Röntgen-Kristallografie vermag und kann darüber hinaus Informationen über die Dynamik von Strukturen liefern (Fischer *et al.*, 2015; Lipfert u. Doniach, 2007; Kondrat *et al.*, 2015; Benda *et al.*, 2014).

1.1. Massenspektrometrische Methoden der strukturellen Proteomik

In einer massenspektrometrischen Analyse kann das Masse-zu-Ladung-Verhältnis (m/z) von ionisierten Molekülen gemessen werden. Durch spezielle Strategien in der Probenvorbereitung, Ionisierung oder der eigentlichen Analyse ist es möglich, mithilfe der Massenspektrometrie niedrig aufgelöste Proteinstrukturinformationen zu gewinnen.

So kann z.B. die Lösungsmittelzugänglichkeit der Seitenketten eines Proteins durch kovalente Modifizierungen bestimmt werden, indem die genaue Position der Modifizierung in der Proteinsequenz massenspektrometrisch identifiziert wird (Mendoza u. Vachet, 2009; Luchini *et al.*, 2014). Eine weitere Methode, die Lösungsmittelzugänglichkeit von Proteinbereichen zu analysieren, ist der Wasserstoff-Deuterium-Austausch. Amid-Wasserstoff-Atome, die nicht in Wasserstoff-Brückenbindungen innerhalb von Proteinen fixiert sind oder im Inneren eines Proteins liegen, können mit Deuterium aus schwerem Wasser (D_2O) austauschen. Der Massenunterschied vor und nach dem Austausch kann massenspektrometrisch nachgewiesen werden. So kann z.B. die Entfaltung von Proteinen untersucht werden,

da einige Amid-Wasserstoff-Atome von Proteinen im ungefalteten, aber nicht im gefalteten Zustand am Austausch teilnehmen können (Katta u. Chait, 1993; Wales u. Engen, 2006).

Mit der Elektrospray-Ionisation (ESI; Dole *et al.*, 1968; Iribarne u. Thomson, 1976) ist es möglich, ganze Proteine oder Proteinkomplexe aus der Lösung heraus zu ionisieren (Fenn *et al.*, 1989), um diese in einem Massenspektrometer zu analysieren. Die Proteinionen können dann zusätzlich durch eine gasgefüllte Röhre (*drift tube*) geführt werden, wobei Ionen mit einem kleineren Radius (*collision cross section*) schneller passieren können als größere Proteine (Jurneczko u. Barran, 2010). Neben dem m/z -Wert und der Driftzeit kann die Anzahl an Ladungen, die ein Protein trägt, Informationen zu dessen Faltungszustand preisgeben. So kann z.B. ein globuläres Protein weniger Ladungen an der Oberfläche akkumulieren als ein entfaltetes Protein (Beveridge *et al.*, 2014)

Chemisches *Cross-Linking* von Proteinen kann benutzt werden, um nicht kovalente Interaktionen zwischen Proteinen zu fixieren. Dabei werden zwei Aminosäureseitenketten kovalent durch einen *Linker* verbunden, wenn diese zuvor in einem ausreichend geringem Abstand zueinander waren. Die Vernetzungsstellen können mittels massenspektrometrischer Methoden auf den Aminosäurerest genau bestimmt werden (Chen *et al.*, 1999; Sinz, 2014). Mit einer ausreichend großen Anzahl an Quervernetzungsstellen in einem Protein ist es möglich, niedrig aufgelöste Strukturen anhand der erhaltenen Abstandsbeschränkungen zu modellieren (Young *et al.*, 2000). Die *Cross-Linking*-Informationen können darüber hinaus auch Aufschluss über die Interaktionen in komplexeren Protein-Protein-Netzwerken liefern (Chen *et al.*, 2010).

Um die gemessenen massenspektrometrischen Rohdaten, also Massenspektren all dieser Ansätze zu interpretieren, sind komplexe und aufwändige Algorithmen nötig, die von Computerprogrammen übernommen werden müssen. In dieser Arbeit wurden zwei Softwareprogramme entwickelt, die die Auswertung von *Cross-Linking*/MS-Experimenten (XL-MS) einfach und schnell ermöglichen sollen.

1.2. *Cross-Linking* in Kombination mit Massenspektrometrie

Der Ablauf eines XL-MS-Experimentes ist schematisch in Abb. 22 dargestellt. Eine Proteinprobe (Abb. 22A) wird mit einem Überschuss an *Cross-Linker* versetzt, so dass nah beieinander liegende Seitenketten kovalent verknüpft werden (Abb. 22B). Nach anschließender Proteolyse wird das Gemisch aus vernetzten, modifizierten und unmodifizierten Peptiden (Abb. 22C) mittels HPLC aufgetrennt. Die Flüssigkeitschromatografie wird mittels Nano-Elektrospray-Ionisierung (ESI) direkt massenspektrometrisch analysiert. Jeweils nach der Messung eines Massenspektrums werden Vorläuferionen (Abb. 22D, rotes Signal) für eine

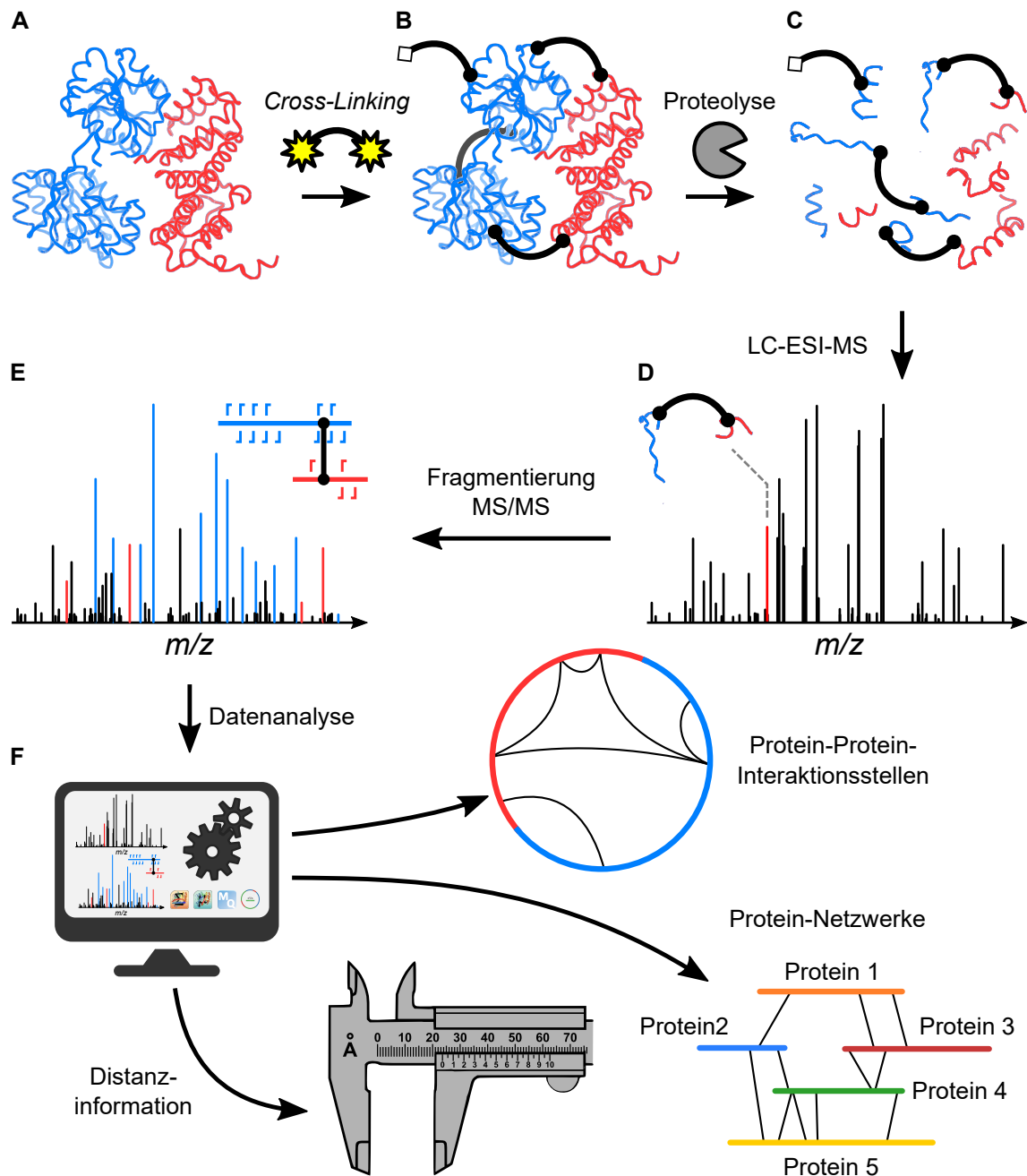


Abbildung 22: Chemisches *Cross-Linking* in Kombination mit Massenspektrometrie - Schematischer Ablauf eines *Cross-Linking*/MS-Experimentes.

Fragmentierung ausgewählt. Das erhaltene Fragmentationen-Spektrum (MS/MS-Spektrum) eines *Cross-Links* enthält Fragmente beider Peptide (Abb. 22E, rote und blaue Signale). Nach der bioinformatischen Analyse der Rohdaten (Abb. 22F) erhält man Informationen über die Vernetzungsstellen der eingesetzten Proteine. Diese Informationen können zur Aufklärung von Protein-Protein-Interaktionen oder Interaktionsnetzwerken, sowie als Abstandsbeschränkungen für die Erstellung von Proteinmodellen verwendet werden. Dieses Reaktionsschema kann auf alle Proteinkomplexe angewendet werden, die *in vitro* gereinigt werden können, kann aber auch für *Cross-Linking*-Experimente *in vivo* adaptiert werden.

Bei der Vernetzung von Proteinen kann es zur Bildung verschiedener *Cross-Link*-Typen kommen (Abb. 23). Bei Typ 0 *Cross-Links* (*dead-end*) reagierte der *Cross-Linker* nur auf einer Seite mit einem Peptid. Die zweite reaktive Gruppe wurde hydrolysiert oder reagierte mit anderen Bestandteilen des Lösungsmittels (z.B. freien Aminen im Puffer). Bei Typ 1 *Cross-Links* haben beide reaktive Gruppen mit demselben Peptid reagiert und bildeten einen intrapeptidalen *Cross-Link*. Typ 2 entspricht einem *Cross-Link* zweier unterschiedlicher Peptide. *Cross-Links* vom Typ 0 und 1 liefern kaum relevante Abstandsinformationen. Sie können aber herangezogen werden, um die Lösungsmittel-Zugänglichkeit der modifizierten Proteinregionen zu bestätigen. Weitere Kombinationen dieser *Cross-Link*-Typen können gleichzeitig vorkommen. Die verschiedenen *Cross-Link*-Typen tragen maßgeblich zur Komplexität der zu analysierenden Probe bei.

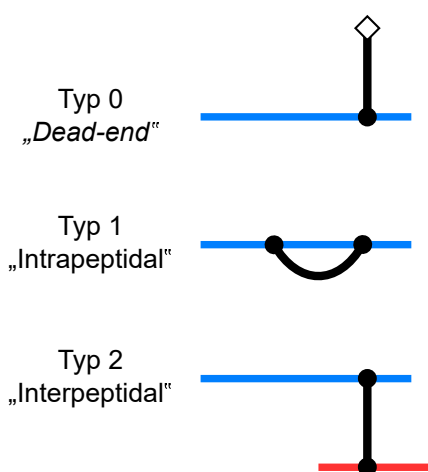


Abbildung 23: *Cross-Link*-Typen
- Unterschiedliche Typen von *Cross-Links*.

Bei der Analyse von *Cross-Linking*/MS-Datensätzen ergibt sich aufgrund der Tatsache, dass zwei Peptide miteinander kombiniert werden, eine sehr große Anzahl an theoretisch möglichen *Cross-Links*. Deren Anzahl steigt quadratisch mit der Anzahl an Peptiden in der analysierten Datenbank möglicher Peptide (Gleichung 6). Bei kleinen Sequenzdatenbanken mit wenigen Proteinsequenzen für die Analyse von einzelnen Proteinen fällt dieser Zusammenhang nicht ins Gewicht. Bei größeren Sequenzdatenbanken mit mehreren Hundert Proteinsequenzen kommt es zu sehr vielen möglichen Kombinationen, die die spätere bioinformatische Analyse deutlich komplizierter gestalten.

$$N_{Xl} = \frac{(N_{Pep})^2 + N_{Pep}}{2} \quad (6)$$

Mit:

N_{Pep} Peptidanzahl
 N_{Xl} Anzahl theoretischer *Cross-Link*-Kombinationen

1.2.1. *Cross-Linking*-Reagenzien

Üblicherweise besteht ein chemischer *Cross-Linker* aus zwei reaktiven Gruppen, die kovalent miteinander durch einen *Linker* verknüpft sind. Dabei können die reaktiven Gruppen gleich (homobifunktioneller *Cross-Linker*) oder unterschiedlich (heterobifunktioneller *Cross-Linker*) sein. Die Länge der *Linker*-Region legt letztlich den maximalen Abstand zweier Seitenketten fest, den der *Cross-Linker* überspannen kann. Ein sehr langer *Linker* hat eine größere Wahrscheinlichkeit, an zwei Stellen zu reagieren, liefert aber dafür nur lockere Beschränkungen z.B. für die Modellierung von Strukturen. Ein *Cross-Linker* mit einer sehr kurzen *Linker*-Region liefert restriktive Beschränkungen, wird aber auch weniger Typ 2-*Cross-Links* ausbilden. Es gibt *Cross-Links*, die ohne *Linker* vorkommen (*zero-length Cross-Links*), wenn z.B. eine Isopeptidbindung zwischen einem Amin und einer Säure-Gruppe gebildet wird. Disulfidbrücken sind demnach ebenfalls *zero-length Cross-Links* zwischen zwei Cysteinen, die mit der hier vorgestellten XL-MS-Strategie und den entsprechenden Software-Lösungen identifiziert werden können.

Reaktive Gruppen

Für die Verknüpfung von Proteinen stehen verschiedene Möglichkeiten zur Verfügung (Sinz, 2006), von denen hier nur einige gezeigt werden (Abb. 24). Bei der Reaktion eines NHS-Esters greift ein primäres Amin (ϵ -Amin des Lysins oder N-Terminus) am Carbonyl-Kohlenstoff des NHS-Esters an, wobei unter Abspaltung von N-Hydroxysuccinimid eine Amidbindung gebildet wird (Abb. 24A). Diese Reaktion findet bevorzugt bei höheren pH-Werten statt, wenn das ϵ -Amin weniger stark protoniert vorliegt ($pK_S(\epsilon - NH_2) = 10,53$). Bei niedrigen pH-Werten (pH 6 - 7) fallen Nebenreaktionen der NHS-Ester mit Hydroxylgruppen aus Serin, Threonin und Tyrosin stärker ins Gewicht (Mädler *et al.*, 2009). Sulfhydrylgruppen aus Cysteinen können mit Maleinsäureimiden durch eine Additionsreaktion kovalent verknüpft werden (Abb. 24B). In C und D sind zwei mögliche Reaktionen mit fotoaktivierbaren Gruppen gezeigt. Bei Diazirinen (Abb. 24C) wie dem Foto-Leucin kommt es durch Abspaltung von Stickstoff zur Bildung eines reaktiven Carbens, welches unspezifisch mit anderen Aminosäuren reagieren kann. Im para-Benzoyl-Phenylalanin (Abb. 24D) kommt es durch Belichtung zur Bildung eines Diradikals, welches mit CH-Gruppen, bevorzugt im Methionin, reagiert (Wittelsberger *et al.*, 2006). Säuregruppen im Protein (Asparaginsäure, Glutaminsäure oder

am C-Terminus) können für eine Reaktion mit Aminen durch Carbodiimide wie EDC (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimid) aktiviert werden. Das kurzlebige Intermediat reagiert mit einem Amin und bildet eine Isopeptidbindung, die einem *zero-length Cross-Link* entspricht (Abb. 24E; Nakajima u. Ikada, 1995).

1.2.2. Massenspektrometrische Analyse

Für die massenspektrometrische Analyse von *Cross-Links* ist es wichtig, auch geringe Mengen einer Spezies mit hoher Auflösung analysieren zu können. Für diese Arbeit stand ein Orbitrap Fusion Tribrid-Massenspektrometer (Thermo; Abb. 25) zur Verfügung. Die Ionisierung der Moleküle erfolgte mittels Nano-Elektrospray-Ionisierung (ESI; Fenn *et al.*, 1989). Dazu wird ein Potential zwischen einer probeführenden Kapillare und einer Elektrode am Spektrometer angelegt. Geladene Ionen bewegen sich entlang des elektrischen Feldes. An der Spitze der Kapillare bildet sich ein Flüssigkeitskegel aus (Taylor-Kegel), an dessen Spitze ein *jet* austritt. Dieser zerfällt zu einem feinen Elektrospray aus kleinen geladenen Tröpfchen, wenn nach Verdampfung des Lösungsmittels die Tröpfchen einen bestimmten Radius unterschreiten (Rayleigh Limit). Die Abstoßung der akkumulierten Ladungen führt dann zum Kollabieren der Tröpfchen. Letztlich gelangen desolvatisierte Ionen in das Massenspektrometer (Dole *et al.*, 1968). ESI ist eine vergleichsweise sanfte Ionisierungsmethode, bei der auch nicht-kovalente Interaktionen erhalten bleiben können. Somit können auch größere Biomoleküle intakt ionisiert werden, die dann meist mehrfach geladen sind.

Das Orbitrap Fusion-Massenspektrometer besitzt drei Massenanalysatoren (Abb. 25), von denen zwei als Detektoren Verwendung finden. Der Quadrupol-Massenfilter dient nicht der Detektion, sondern der Selektion von Ionen (z.B. zur späteren Fragmentierung). In der linearen Ionenfalle können Ionen mit niedriger Auflösung detektiert sowie Ionen für weitere Fragmentierungen selektiert werden (MS^n). Im Orbitrap-Massenanalysator können hoch aufgelöste Massenspektren aufgenommen werden. Aufgrund der Organisation der einzelnen Analysatoren ist es möglich, während eines *MS-scans* im Orbitrap-Analysator Ionen im Quadrupol-Massenfilter für die spätere Fragmentierung zu selektieren, so dass die Geschwindigkeit der Analyse durch die Parallelisierung der Vorgänge sehr hoch ist (Senko *et al.*, 2013).

Fragmentierung von *Cross-Links*

Während der massenspektrometrischen Analyse ist es möglich, eine definierte Spezies von Ionen (selektiert nach m/z) zu sammeln und anschließend zu fragmentieren. Dabei wird den Ionen Energie zugeführt und es kommt zu Bindungsbrüchen innerhalb des Moleküls. Die erhaltenen Fragmentionen werden anschließend in einen Massenanalysator

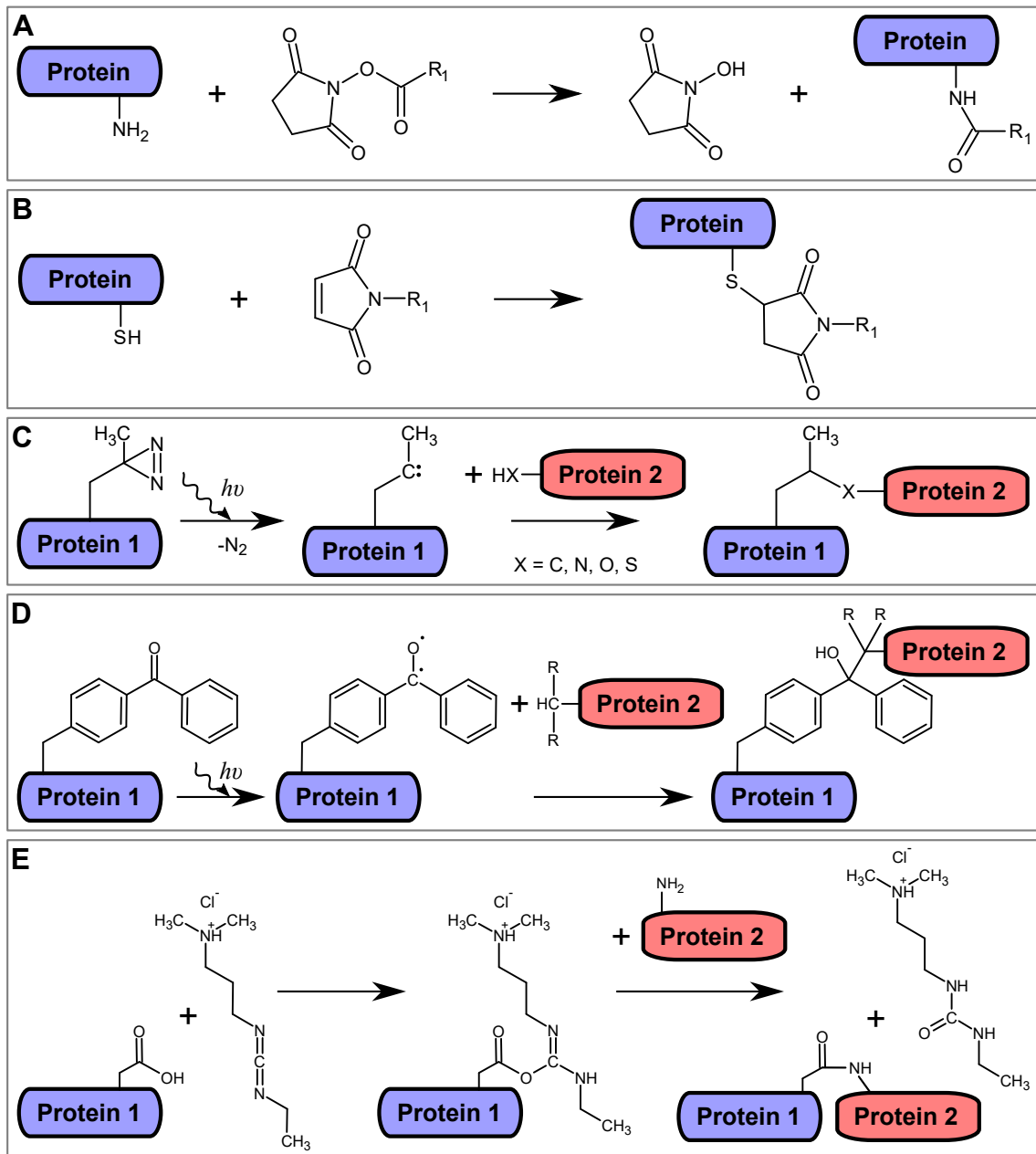


Abbildung 24: Chemische Reaktionen für die Protein-Quervernetzung - Darstellung möglicher Reaktionen zur Quervernetzung von Proteinen (R_1 zeigt den Rest des *Cross-Linkers* an, der eine zweite reaktive Stelle besitzt). (A) Reaktion von einem N-Hydroxysuccinimidester (NHS) mit einem Amin wie am N-Terminus oder in Lysinen. (B) Reaktion eines Maleinsäureimides mit einer Sulfhydrylgruppe aus einem Cystein. (C) *Cross-Linking*-Reaktionen mit fotoaktivierbaren Aminosäuren wie Foto-Leucin und (D) p-Benzoyl-Phenylalanin (p-Bpa). (E) *Zero-length-Cross-Link* durch die Reaktion einer Carboxylgruppe mit einem Amin nach Aktivierung durch EDC.

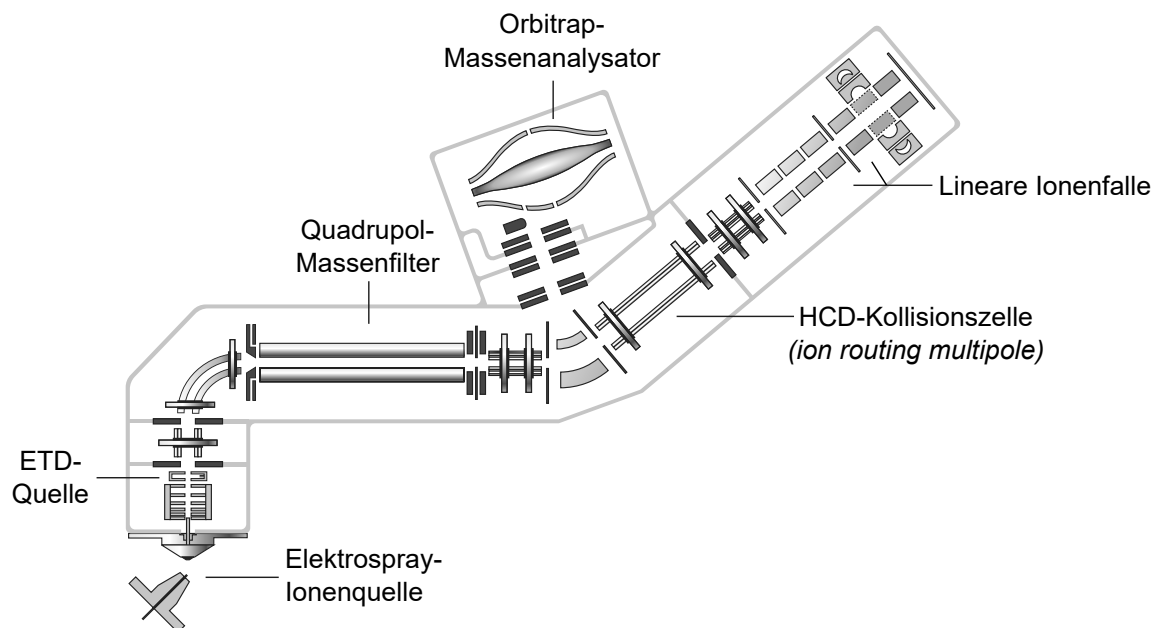


Abbildung 25: Orbitrap Fusion Tribrid-Massenspektrometer - Schematischer Aufbau des Orbitrap Fusion Tribrid-Massenspektrometers (Abbildung modifiziert nach dem Orbitrap Fusion Datenblatt von <http://www.thermofisher.com/> und Senko *et al.*, 2013).

(lineare Falle oder Orbitrap) überführt und es wird ein Fragmentionen-Spektrum (MS/MS-Spektrum) detektiert. In Peptiden sind dabei vor allem Fragmentierungen innerhalb des Peptidrückgrates interessant, da diese an jeder Aminosäure im Peptid vorkommen können (Abb. 26A). So kann bei aufeinanderfolgenden Bindungsbrüchen die Sequenz eines Peptids aus dem Fragmentionen-Spektrum bestimmt werden (vgl. Abb. 14, S. 33). Bei der häufig-angewendeten CID-Fragmentierung (*collision induced dissociation*) kommt es vor allem zu Bindungsbrüchen an der Peptidbindung, die zur Bildung von b- und y-Ionen führen (Abb. 26A).

Bei *Cross-Links* überlagern sich die Spektren zweier Peptide, die zudem durch die Modifizierung mit dem *Cross-Linker* verschoben sind, so dass *Cross-Link*-Spektren von vornherein komplexer sind. Um diese komplexen Spektren eindeutig zu annotieren, wurde in dieser Arbeit und auch den vorgestellten Programmen die Nomenklatur für Peptidfragmentionen angewendet und erweitert (Roepstorff u. Fohlman, 1984; Schilling *et al.*, 2003). Dazu wird das jeweils schwerere Peptid als α -Peptid und das leichtere als β -Peptid bezeichnet (Abb. 26B). Im Unterschied zu einfachen Peptidfragmenten können *Cross-Link*-Fragmente zusätzlich den *Cross-Linker* und das zweite Peptid enthalten. Diese Ionen werden in dieser Arbeit als verzweigte Fragmentionen bezeichnet.

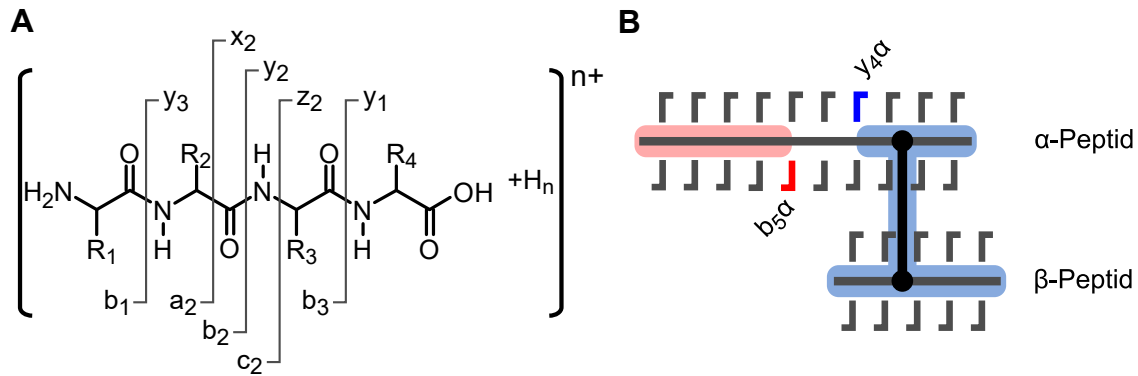


Abbildung 26: Fragmentierung von Peptiden und *Cross-Links* - (A) Mögliche Fragmentionen nach Fragmentierung eines Peptids an den verschiedenen Bindungen im Peptidrückgrat. Die Nomenklatur der Fragmente erfolgte dabei nach Roepstorff u. Fohlman (1984) (B) Schematische Darstellung von Fragmentierungen eines *Cross-Links*. Die Nomenklatur wird um die Peptidbezeichnung α oder β erweitert (Schilling *et al.*, 2003). Das $b_5\alpha$ -Ion entspricht einem linearen Ion (rot hervorgehoben und hinterlegt), wie es auch in Peptiden vorkommt. Das $y_4\alpha$ -Ion ist ein verzweigtes Ion (blau hervorgehoben und hinterlegt). Dieses enthält zusätzlich den *Cross-Linker* (schwarz) und das β -Peptid.

Fragmentierungsarten

Das Orbitrap Fusion-Massenspektrometer stellt verschiedene Fragmentierungsarten zur Verfügung:

- *Collision-induced dissociation* (CID): Bei der Kollisions-induzierten Dissoziation kommt es durch Anlegung eines Potentials zu einer resonanten Anregung der selektierten Ionen in der Ionenfalle, die anschließend mit einem neutralen Gas (Helium) kollidieren. Dabei kommt es bei Peptiden, wie auch bei Peptid-Peptid-*Cross-Links*, zu Bindungsbrüchen an den Peptidbindungen und hauptsächlich zur Bildung von b- und y-Ionen. Bindungen zu Peptid-Modifizierungen (z.B. Phosphorylierungen) fragmentieren bei CID bevorzugt, so dass es zu Neutral-Verlusten des Vorläuferions kommt und die Position der Modifikation nur schwer bestimmt werden kann (McLafferty *et al.*, 1973; Wells u. McLuckey, 2005).
- *Higher-energy C-trap dissociation* (HCD): Bei der HCD-Fragmentierung (auch *higher-energy collisional dissociation*) werden Ionen mit elementarem Stickstoff in der HCD-Kollisionszelle kollidiert (Olsen *et al.*, 2007). Die Anregung erfolgt dabei durch einen Radiofrequenzpuls. Die Fragmentierung der Ionen erfolgt ebenfalls an der Peptidbindung, so dass b- und y-Ionen entstehen. Im Vergleich zur CID sind die Fragmentierungsereignisse bei HCD gleichmäßiger entlang des Peptidrückgrates verteilt. Zusätzlich kann es zu sekundären Bindungsbrüchen kommen, so dass Neutralverluste beobachtet werden können (z.B. Bildung von a-Ionen durch einen CO-Verlust von b-Ionen). Häufig kann auch das nicht fragmentierte Vorläuferion noch im MS/MS-Spektrum beobachtet werden (Michalski *et al.*, 2012).

- *Electron transfer dissociation* (ETD): Für die ETD-Fragmentierung wird auf die Vorläuferionen ein einzelnes Elektron durch ein radikalisches Anion übertragen. Bei Peptiden führt die nicht-ergodische, radikal-induzierte Spaltung zu Brüchen der C_α-N-Bindung im Peptidrückgrat, wodurch c- und z[•]-Ionen gebildet werden. Im Gegensatz zu CID bleiben Modifizierungen erhalten, so dass ETD verwendet wird, um Phosphorylierungsstellen in Peptiden zu bestimmen (Syka *et al.*, 2004). Häufig werden ein ladungsreduziertes Vorläuferion ($[M+H_n]^{n-1}$) oder Vorläuferionen mit einem Protonen-Verlust im Spektrum beobachtet ($[M+H_{n-1}]^{n-1}$) (Hart-Smith, 2014).

Die CID- und ETD-Fragmentierungen finden beim Orbitrap Fusion-Massenspektrometer in der linearen Ionen-Falle statt, während die HCD-Fragmentierung in der HCD-Kollisionszelle (*ion-routing multipole* - IRM) stattfindet. Es sind auch Kombinationen von ETD mit CID (ETciD) oder HCD (EThcD) möglich, die dann zu einer größeren Anzahl verschiedener Fragmentionentypen führen.

1.2.3. Vereinfachung der Auswertung von *Cross-Linking*/MS-Datensätzen

Die Identifizierung von *Cross-Links* kann aufgrund der Komplexität der Probe und der gemessenen Spektren sehr kompliziert sein. Durch geschickte Modifizierung einzelner Schritte des *Cross-Linking*/MS-Ablaufs kann die spätere bioinformatische Auswertung stark vereinfacht werden. Dazu können z.B. quervernetzte Peptide chromatografisch angereichert werden oder so modifiziert werden, dass man diese in den Massenspektren deutlich von anderen Spezies unterscheiden kann.

Anreicherung quervernetzter Peptide

Cross-Links von tryptischen Peptiden können im Säuren von vornherein mindestens vier positive Ladungen tragen, zwei an den beiden N-Termini und zwei an den beiden C-terminalen Aminosäuren (Arginin oder Lysin). Zudem weisen quervernetzte Peptide eine größere molekulare Masse als unmodifizierte Peptide auf. Diese Eigenschaften können ausgenutzt werden, um *Cross-Links* von kleinen, niedrig geladenen Peptiden zu trennen. So können *Cross-Links* z.B. effizient mittels Kationen-Austausch-Chromatografie angereichert werden (Fritzsche *et al.*, 2012).

Neben bifunktionellen *Cross-Link*-Reagenzien sind trifunktionelle Reagenzien verfügbar. Diese besitzen zwei reaktive Gruppen, die mit Aminosäuren reagieren können und eine dritte chemische Gruppe, die die Anreicherung von *Cross-Links* ermöglicht, wie z.B. Biotin, das an einer Streptavidin-Matrix angereichert werden kann oder ein Alkin, das mit Aziden reagieren kann, um kovalent an eine Matrix gekoppelt zu werden (Petrochenko *et al.*, 2011; Kaake *et al.*, 2014).

Auflösung von MS/MS-Spektren

Eine verbesserte Auflösung von Massenspektren lässt später eine deutlich stringendere Zuordnung von möglichen *Cross-Links* zu. Dies trifft nicht nur auf die Genauigkeit der Vorläufermasse, sondern auch auf die Fragmentionen-Spektren zu. In hochaufgelösten MS/MS-Spektren lässt sich nicht nur das Masse/Ladungsverhältnis genauer bestimmen, sondern zusätzlich auch die Ladungszustände der Ionen. Dieser kann durch die Abstände der Isotopensignale eines Ions, die erst bei höheren Auflösungen getrennt voneinander beobachtet werden können, bestimmt werden. Bei unbekanntem Ladungszustand müssen in der Analyse für ein gemessenes Ion alle möglichen Ladungszustände angenommen werden. Das kann unter Umständen zu falschen Zuweisungen und damit falsch-positiven *Cross-Links* führen.

Isotopen-markierte *Cross-Linker*

Von einigen *Cross-Linking*-Reagenzien gibt es isotoopenmarkierte Derivate. Dabei werden mehrere Atome durch ihre schwereren Isotope ersetzt (z.B. kann Wasserstoff durch Deuterium ersetzt werden). Der kommerziell verfügbare DSS-*Cross-Linker* (Disuccinimidylsuberat) ist als Variante mit vier oder zwölf Deuteriumatomen verfügbar. Unmarkierte und markierte *Cross-Linker* werden zu gleichen Teilen gemischt (d0/d4 oder d0/d12) und anschließend wird wie in Abb. 22 weiter verfahren. Ein gebildeter *Cross-Link* sollte mit beiden *Cross-Link*-Reagenzien in etwa zu gleichen Teilen vorkommen. Im Massenspektrum ist der Unterschied von etwa 4 u oder 12 u deutlich zu erkennen. Alle Vorläuferionen, die ein entsprechendes Muster aufweisen, enthalten demnach sehr wahrscheinlich den *Cross-Linker*, wobei auch Typ 0 und 1 *Cross-Links* möglich sind (siehe Abb. 23). Die Massendifferenz zeigt sich zusätzlich auch im Fragmentionen-Spektrum, wenn ein Fragmention den *Cross-Linker* enthält und beide Varianten gleichzeitig selektiert werden. Isotoopen-markierte *Cross-Linker* erlauben demnach eine gezielte Fragmentierung von Vorläuferionen, die ein entsprechendes Isotopenmuster aufweisen, sowie eine sicherere Zuordnung von Fragmentionen. Damit kommt es zu deutlich weniger falsch-positiven Identifizierungen.

MS/MS-spaltbare *Cross-Linker*

Eine weitere Möglichkeit, MS/MS-Spektren von linearen Peptiden und *Cross-Links* zu unterscheiden, ist durch MS/MS-spaltbare *Cross-Linker* gegeben. Bei der Fragmentierung eines solchen *Cross-Links* kommt es zu Brüchen definierter Bindungen innerhalb der *spacer*-Region des *Cross-Linkers*. Dadurch entstehen charakteristische Paare von Fragmentionen, die den beiden vernetzten Peptiden mit den jeweils unterschiedlichen *Cross-Linker*-Fragmenten als Modifikation entsprechen (Reporterionen). In Abb. 27 sind drei verschiedene MS/MS-spaltbare *Cross-Linker* gezeigt.

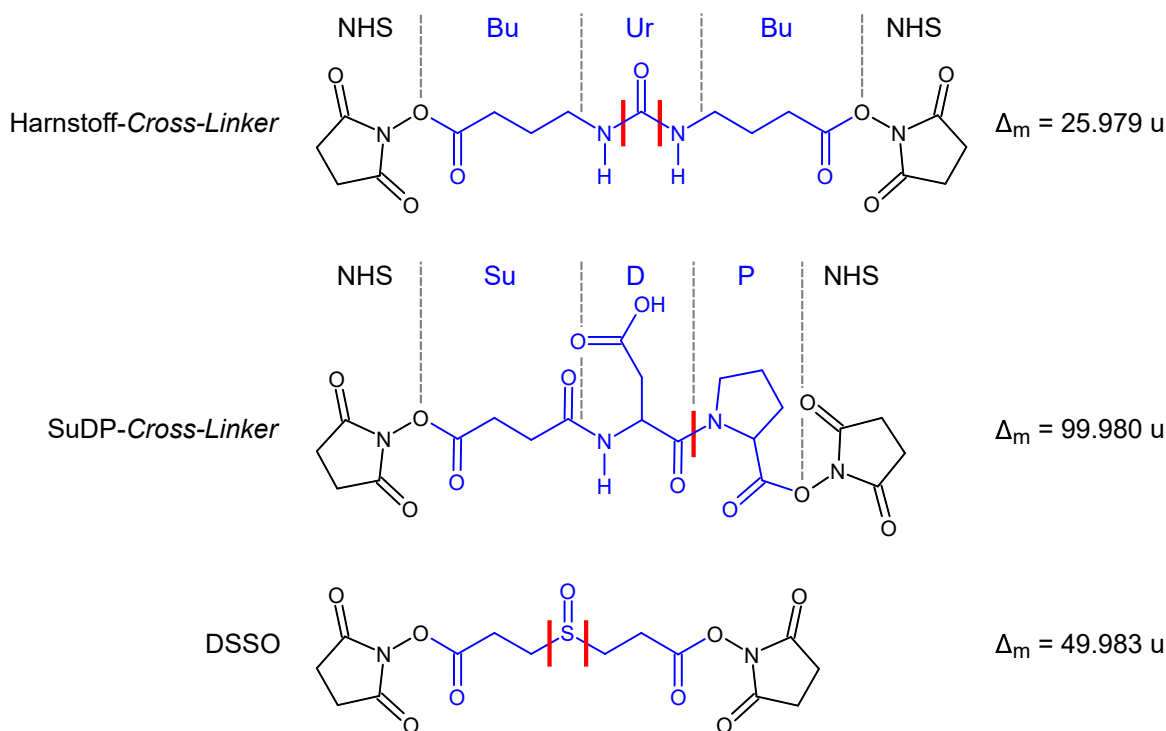


Abbildung 27: Beispiele für MS/MS-spaltbare *Cross-Linker* - Strukturen MS/MS-spaltbarer *Cross-Linker* (Müller *et al.*, 2010; Soderblom u. Goshe, 2006; Kao *et al.*, 2010). In Schwarz sind die N-Hydroxysuccinimidester dargestellt. In Blau ist die *Linker*-Region mit den Spaltstellen in Rot dargestellt. Bei der Fragmentierung des *Cross-Linkers* können zwei verschiedene Modifikationen an einem Peptid verbleiben, diese können als spezifische Reporterionen im Fragmentationen-Spektrum detektiert werden. Die Massendifferenz Δ_m der Reporterionen eines Peptids ist jeweils rechts notiert.

Alle drei *Cross-Linker* sind homobifunktionell und aminreaktiv. Der Harnstoff-*Linker* und DSSO besitzen jeweils zwei symmetrisch angeordnete Spaltstellen (rote Linien), so dass es bei der Fragmentierung zu zwei verschiedenen Modifikationen an einem Peptid kommen kann. Die gebildeten Reporterionenpaare eines Peptids weisen eine spezifische Massendifferenz auf (Δ_m). Diese Massendifferenz ist insofern ungewöhnlich, als sie einen Massendefekt aufweist. Dieser Pseudomassendefekt entsteht allerdings nicht durch die Kernzusammensetzung der Atome in den Fragmenten, sondern durch die Differenz der chemischen Zusammensetzung der beiden Reporterionen. Für den Harnstoff-*Linker* entspricht diese Differenz CO-H_2 (25,979 u). Mit modernen Massenanalysatoren können diese kleinen Unterschiede detektiert werden und sind hilfreich für die Identifizierung von *Cross-Links*. Der SuDP-*Cross-Linker* ist asymmetrisch und fragmentiert bevorzugt an der Peptidbindung zwischen Asparaginsäure (D) und Prolin (P) in der *Linker*-Region (blau). Da der *Cross-Linker* homobifunktionell ist, kann dieser theoretisch in beiden Ausrichtungen mit den Peptiden reagieren, so dass auch SuDP ein spezifisches Reporterionenpaar liefert.

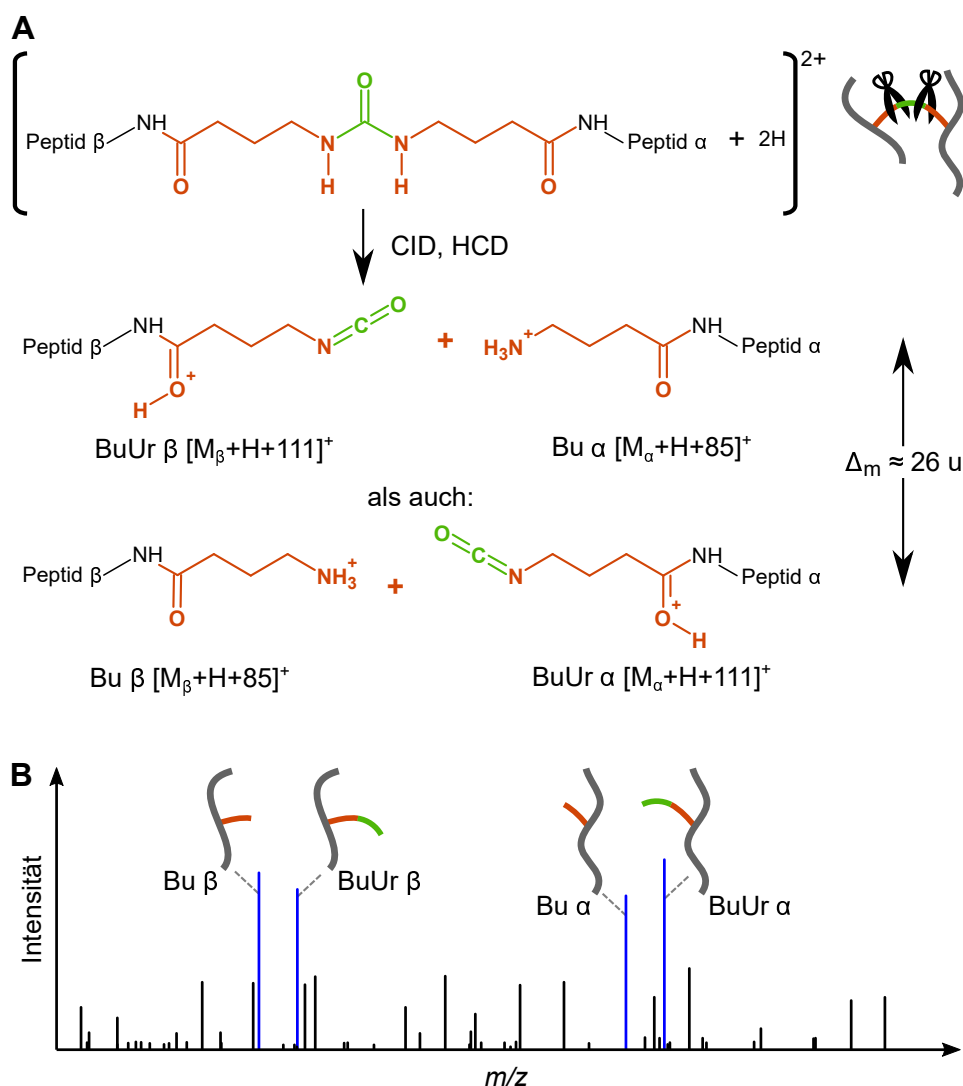


Abbildung 28: Fragmentierung des Harnstoff-Linkers - (A) Vorgeschlagener Mechanismus der Harnstoff-Linker Fragmentierung (Müller *et al.*, 2010). Die Aminobuttersäure-Teile des *Cross-Linkers* sind orange dargestellt, die Carbonylgruppe ist grün dargestellt. Nach der Fragmentierung wird der *Cross-Linker* in 4-Aminobuttersäure (Bu-Fragment) und ein 4-Isocyanato-Buttersäure-Fragment (BuUr-Fragment) gespalten. Die positive Ladung kann dabei an einem der beiden Fragmente verbleiben. (B) Schematische Darstellung eines Fragmentionen-Spektrums nach *Cross-Linker*-Fragmentierung. In Blau sind die einfach geladenen Fragmente der jeweiligen Peptide mit einem Cartoon der Fragmentionen dargestellt.

Harnstoff-Linker: Die Fragmentierung des Harnstoff-*Linkers* ist in Abb. 28A dargestellt. Rechts ist der *Cross-Link* zusätzlich als Cartoon in den gleichen Farben wie im Fragmentierungsschema dargestellt. An jedem Peptid können nach der Fragmentierung des *Cross-Linkers* ein 4-Aminobuttersäure-Fragment oder das komplementäre Isocyanat-Fragment verbleiben. Diese werden, entsprechend der Beschriftung in Abb. 27, als Bu- und BuUr-Fragment bezeichnet. Bei Typ 0-*Cross-Links* kommt es zu einer sehr ähnlichen Fragmentierung, bei der die Bildung nur eines Reporterionenpaares durch zwei Neutralverluste beobachtet wird. Diese Neutralverluste können bei einem *Cross-Link* zweier Peptide nicht auftreten und erlauben die Diskriminierung von Typ 0 und Typ 2 *Cross-Links*. Bei intrapeptidalen *Cross-Links* (Typ 1) kommt es nach Fragmentierung des *Cross-Linkers* nicht zur Spaltung des Vorläuferions, da die Fragmentierung innerhalb einer ringartigen Struktur stattfindet. Es kann allerdings zu sekundären Fragmentierungsereignissen kommen, die zur Abspaltung des Bu-Fragmentes vom Vorläuferion führen ($-C_4H_7NO$). Dieser Vorläufer-Neutralverlust kann nur bei einem Typ 1-*Cross-Link* auftreten und kann daher als Reporterion für die bioinformatische Analyse herangezogen werden. Mögliche Fragmentationen nach der Fragmentierung verschiedener *Cross-Link*-Typen sind schematisch in Abb. 47 im Anhang auf S. LIV dargestellt.

Dadurch kann mithilfe der spaltbaren *Cross-Linker* eindeutig zwischen den drei *Cross-Link*-Typen und unmodifizierten Peptiden unterschieden werden³. Durch eine weitere Fragmentierung der gebildeten Reporterionen (MS³) kann man die an *Cross-Links* beteiligten Peptide einzeln fragmentieren und mit gängigen Peptid-Suchalgorithmen identifizieren.

Für den Harnstoff-*Cross-Linker* konnte jedoch beobachtet werden, dass zusätzlich zu den Reporterionen auch Fragmentierungen im Peptidrückgrat der beiden Peptide auftreten. Damit enthält ein MS/MS-Spektrum sowohl die Information, dass es sich um einen *Cross-Link* handelt (Reporterionen), als auch Sequenzinformationen der beiden Peptide. Ein MS/MS-Spektrum ist demnach ausreichend für die Identifizierung eines *Cross-Links* (Müller *et al.*, 2010; Götze *et al.*, 2015).

1.3. Bioinformatische Auswertung von XL-MS-Experimenten

Die Auswertung von *Cross-Linking*/MS-Experimenten stellt nach wie vor einen komplizierten Schritt in der Analyse dar. Tausende von Spektren müssen automatisiert und zuverlässig nach möglichen *Cross-Links* durchsucht werden. Für die sichere Identifizierung eines *Cross-Links* reicht nur eine Übereinstimmung der Vorläuferionenmasse nicht aus. Im besten Fall sollten

³Dies ist hilfreich, um z.B. einen Typ 0-*Cross-Link* eines Peptids mit einer übersprungenen Proteolysestelle von einem Typ 2-*Cross-Link* zweier Peptide zu unterscheiden, die in der Proteinsequenz direkt aufeinanderfolgen. Beide Varianten weisen die exakt gleiche Masse auf und können dennoch eindeutig durch das Auftreten von Reporterionen unterschieden werden.

beide Peptidsequenzen aus einem Fragmentionen-Spektrum zu erkennen sein. Da es beim *Cross-Linking* zu einer sehr großen Anzahl möglicher Kombinationen von Peptiden kommen kann, sind schnelle und effiziente Algorithmen nötig, um die meist wenig abundanten *Cross-Links* in komplexen Proben zu identifizieren. Viele Arbeitsgruppen haben dazu Programme entwickelt, die dies ermöglichen sollen. Tabelle 5 stellt einige der publizierten Programme und Algorithmen zusammen, die während der letzten 15 Jahre entwickelt wurden.

1.3.1. Verfügbare *Cross-Linking*/MS-Software

Programme zur Auswertung von XL-MS-Daten haben sich mit der Zeit weit entwickelt. So wurden mit den ersten Programmen hauptsächlich experimentell erhaltene Massen mit theoretisch möglichen *Cross-Links* verglichen (GPMAW, MS-Bridge, CLPM). Später wurden Programme entwickelt, die zusätzlich die MS/MS-Spektren von *Cross-Link*-Kandidaten mit deren theoretischem MS/MS-Spektrum verglichen. Mit xQuest (Rinner *et al.*, 2008) wurde ein automatisierter *score*, eine Abschätzung der *False-Discovery-Rate* und eine einfache grafische Oberfläche eingeführt, wodurch die Analyse und die Handhabung stark verbessert wurden, allerdings beruhte die Analyse ausschließlich auf der Mischung von isotope markierten *Cross-Link*-Reagenzien. Das in dieser Arbeit vorgestellte Programm, StavroX entstand zwischen 2009 und 2010. Die erste Version von StavroX (1.5) wurde 2010 fertiggestellt, weiter optimiert und 2011 veröffentlicht (Version 2.0.5; Götze *et al.*, 2012). Diese Version war bereits auf eine Vielzahl von *Cross-Linking*-Reagenzien anwendbar. Kurz darauf wurde ein weiteres wertvolles Programm zur Auswertung von *Cross-Links*, pLink (Yang *et al.*, 2012), veröffentlicht. Seitdem sind eine Vielzahl weiterer Suchstrategien, *scoring*-Ansätze und Algorithmen veröffentlicht worden (z.B. XLPM, Kojak oder SIM-XL).

Der Einsatz MS/MS-spaltbarer *Cross-Linker* hat zur Entwicklung des Programms MeroX geführt. Für die Analyse von MS/MS-spaltbaren *Cross-Links* stehen bislang nur drei Programme zur Verfügung (MeroX, DXMSMS und XlinkX). Einige der gelisteten Programme werden nach wie vor gewartet und auf neue Bedürfnisse abgestimmt, so auch StavroX und MeroX.

Die Vielfalt an Programmen hat allerdings einen entscheidenden Nachteil: Ergebnisse, die mit unterschiedlichen Programmen erhalten wurden, sind nur schwer vergleichbar. Es gibt bislang nur wenig Bemühungen, ein einheitliches Dateiformat für die *Cross-Link*-Daten festzulegen. Vorgeschlagen wurden XML-basierende Datenformate wie proXML (Riffle *et al.*, 2016) oder pepXML (Hoopmann *et al.*, 2016). Für das proXML-Format wurde bereits eine Konvertierung von StavroX-Ergebnissen zu proXML implementiert (Riffle *et al.*, 2016). Ein allgemein anwendbares Dateiformat sollte flexibel genug sein, um die Vielfalt von *Cross-Linking*/MS-Experimenten sinnvoll repräsentieren zu können.

Tabelle 5: Programme zur *Cross-Linking*/MS-Auswertung - Eine Auswahl von Programmen, sortiert nach dem Publikationszeitpunkt. Die Programme StavroX und MeroX sind Ergebnisse dieser Arbeit. Häufig verwendeten Programme sind mit einem * markiert (nach Zitationen).

Programm	Referenz
GPMAW	Peri <i>et al.</i> (2001)
Protein Prospector (MS-Bridge)*	Clauser <i>et al.</i> (1999); Chu <i>et al.</i> (2004)
CLPM	Tang <i>et al.</i> (2005)
Pro-Crosslink	Gao <i>et al.</i> (2006)
xQuest*	Rinner <i>et al.</i> (2008)
MassMatrix	Xu <i>et al.</i> (2008)
CrossSearch	Nadeau <i>et al.</i> (2008)
DBond	Choi <i>et al.</i> (2010)
Crux	McIlwain <i>et al.</i> (2010)
DXMSMS	Petrotschenko u. Borchers (2010)
Crosswork/MassAI	Rasmussen <i>et al.</i> (2011)
XlinkIdentifier	Du <i>et al.</i> (2011)
StavroX*	Götze <i>et al.</i> (2012)
pLink*	Yang <i>et al.</i> (2012)
FindX	Söderberg <i>et al.</i> (2012)
Find_XL MatLab Skript	Kaliskan <i>et al.</i> (2012)
Hekate	Holding <i>et al.</i> (2013)
MeroX	Götze <i>et al.</i> (2015)
XLPM	Jaiswal <i>et al.</i> (2014)
MXDB	Wang <i>et al.</i> (2014)
AnchorMS	Mayne u. Patterson (2014)
Link-Hunter	Kaake <i>et al.</i> (2014)
Kojak*	Hoopmann <i>et al.</i> (2015)
xLinkX	Liu <i>et al.</i> (2015)
SIM-XL	Lima <i>et al.</i> (2015)
CrossFinder	Mueller-Planitz (2015)
Mass Spec Studio	Sarpe <i>et al.</i> (2016)
xilmass	Yilmaz <i>et al.</i> (2016)
Quantifizierung mit MaxQuant	Chen <i>et al.</i> (2016)

Visualisierung von identifizierten *Cross-Links*

Die zunehmende Größe an *Cross-Link*-Datensätzen erschwert zudem die übersichtliche Darstellung der erhaltenen Daten auf einfache Weise, so dass spezielle Programme für die visuelle Darstellung entwickelt wurden wie z.B. xVis (Grimm *et al.*, 2015), xiNET (Combe *et al.*, 2015) oder ProXL (Riffle *et al.*, 2016). Um die erhaltenen Distanzbeschränkungen auch auf die Modellierung von Proteinstrukturen anwenden zu können, wurde XL-MOD (Ferber *et al.*, 2016) entwickelt, welches die erhaltenen *Cross-Links* verwendet, um mögliche Strukturmodelle eines Proteins zu errechnen.

1.3.2. Angewendete Algorithmen zur Identifizierung von *Cross-Links*

In einer einzigen *Cross-Linking*/MS-Analyse werden mehrere Tausend MS/MS-Spektren aufgenommen, die auf *Cross-Links* untersucht werden. Die hier beschriebenen Identifizierungs- und *scoring*-Algorithmen beziehen sich jeweils auf die Analyse eines MS/MS-Spektrums einer Spezies von Vorläuferionen. Diese Algorithmen werden demnach in einer Datenanalyse mehrere tausendmal nacheinander durchgeführt.

Definitionen zum Algorithmus

Such-Algorithmen für *Cross-Links* dienen dazu, die Anzahl an Kandidaten zu reduzieren, die im anschließenden Validierungsschritt (*scoring*) untersucht werden müssen. Das Suchkriterium ist dabei für die meisten Algorithmen sehr simpel: Stimmt die gemessene Masse des Ions mit der Masse eines theoretischen *Cross-Links* aus zwei Peptiden überein?

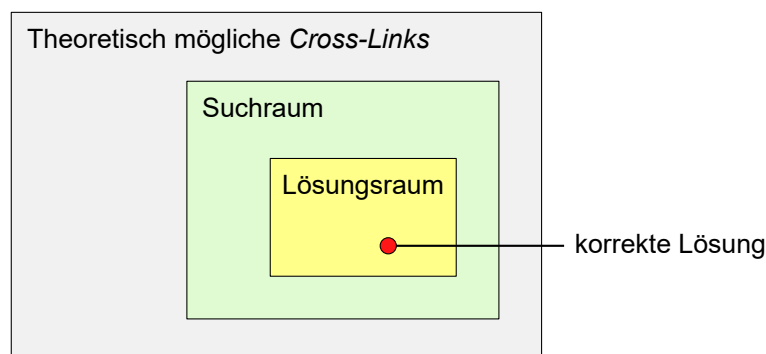


Abbildung 29: Definitionen zum *Cross-Link*-Identifizierungs-Algorithmus

- Vereinfachte Darstellung eines Identifizierungsproblems. Für die *Cross-Link*-Identifizierung wird Folgendes angenommen: Aus einer Protein-Sequenzdatenbank werden alle theoretisch möglichen *Cross-Links* errechnet (grau). Der Suchraum des Algorithmus (grün) umfasst alle durchsuchten Kombinationen von Peptiden. Der Lösungsraum (gelb) umfasst alle Peptidkombinationen, die dem Suchkriterium entsprechen. Die eigentlich Lösung (rot) befindet sich innerhalb des Lösungsraums.

Einige Begriffe, die für die Erläuterung des Suchalgorithmus in der vorliegenden Arbeit benutzt werden, sind in Abb. 29 in Beziehung zueinander gebracht. Alle Kombinationen aus zwei Peptiden, die aus einer Sequenzdatenbank errechnet werden, entsprechen den theoretisch möglichen *Cross-Links* (grau). Mit intelligenten Suchalgorithmen müssen tatsächlich nicht alle Kombinationen durchsucht werden, so dass der Suchraum (grün) nur einen Teil aller möglichen *Cross-Links* abdeckt. Die Anzahl an theoretisch möglichen *Cross-Links* und die Größe des Suchraums steigen quadratisch mit der Anzahl an möglichen Peptiden. Alle Peptidkombinationen, die dem Suchkriterium entsprechen, werden hier als Lösungsraum definiert (gelb). Auch der Lösungsraum wird quadratisch größer, da Peptidkombinationen zufällig die korrekte Masse aufweisen können. Die richtige Kombination aus zwei Peptiden ist in Rot als die korrekte Lösung dargestellt. In einer realen Analyse kennt man die korrekte Lösung nicht und diese kann sogar außerhalb des Suchraums liegen (z.B. ein *Cross-Link* mit einem unbekannten Peptid oder einer nicht angenommenen Modifikation). Der Lösungsraum enthält damit sowohl den korrekten Kandidaten als auch falsch-positive Lösungen. Diese müssen später im Validierungsschritt durch das *scoring* voneinander unterschieden werden.

Die Größe des Suchraums lässt sich durch bestimmte Algorithmen verkleinern, dabei bleibt die Größe des Lösungsraums jedoch gleich. Der Lösungsraum kann nur verkleinert werden, wenn man das Suchkriterium konkretisiert.

Der positiven Identifizierung liegt bei allen Algorithmen zugrunde, dass die Peptiddatenbank, auf der die Analyse beruht, so vollständig wie möglich, aber nur so groß wie nötig ist. Dies kann erreicht werden, indem man zunächst eine Datenbanksuche durchführt und nur identifizierte Proteine in die eigentliche *Cross-Link*-Suche einbezieht (Chu *et al.*, 2010). Zusätzlich können Spektren, die offensichtlich von linearen Peptiden oder Typ 0-*Cross-Links* stammen, ignoriert werden, um die Analyse zu beschleunigen. Für die Identifizierung von *Cross-Links* können drei verschiedene Vorgehensweisen unterschieden werden:

Verketteten von Peptiden zu einer *Cross-Link*-Datenbank

Bei dieser Methode (verwendet in Maiolica *et al.*, 2007) werden alle Kombinationen von *Cross-Links* durchsucht. Es wird eine Peptiddatenbank erzeugt, die später mit einem Standard-Peptid-Suchalgorithmus analysiert werden kann. Damit ein Eintrag in die Peptiddatenbank einen *Cross-Link* aus zwei Peptiden (A und B) widerspiegeln kann, werden die beiden Peptidsequenzen einfach hintereinander zusammengeführt. Dies macht man in beiden Ausrichtungen (AB und BA). Das verdoppelt zwar die Größe der Datenbank, ist aber nötig, um alle theoretischen Fragmentationen eines *Cross-Links* mit zwei Peptiden darstellen zu können. Anschließend wird diese Datenbank (XDB) mit etablierten Suchalgorithmen (MASCOT in Perkins *et al.*, 1999 oder SEQUEST in Eng *et al.*, 1994) analysiert. Dabei wird der *Cross-Linker* als variable Modifikation des verketteten Peptids angenommen. Der Vorteil

liegt darin, dass kein *scoring* implementiert werden muss, sondern lediglich die Erzeugung der Datenbank notwendig ist. Dieser Ansatz ist demnach sehr einfach einzusetzen. Durch die artifizielle Art eines *Cross-Link*-Eintrages in der Datenbank werden allerdings auch Fragmentationen mit dem experimentellen Spektrum verglichen, die in einem realen *Cross-Link* nicht vorkommen können. Auch die Verknüpfungsstelle der Peptide kann nicht direkt bestimmt werden.

Suche von Peptiden mit unbekannter Modifikation

Bei diesem Algorithmus werden zwei Suchdurchgänge je Spektrum durchlaufen. Im ersten Suchlauf wird angenommen, dass es sich bei dem gemessenen Ion um ein Peptid aus der Datenbank mit einer unbekannten Modifikation handelt. In der zweiten Suche werden Peptide gesucht, die die entsprechende Massendifferenz zu dem zuvor identifizierten Peptid erklären können (z.B. verwendet in pLink oder Kojak). Für beide Peptide wird dann ein *score* erhalten, der zu einem *Cross-Link-score* verrechnet wird. Der *Cross-Link-score* ist abhängig von dem Peptid mit dem schlechteren *score* (Trnka *et al.*, 2014). Ansonsten käme es zu deutlich mehr falsch-positiven Identifizierungen mit nur einem richtigen Peptid im *Cross-Link*. Bei dieser Methode wird vermieden, alle möglichen Kombinationen von *Cross-Links* zu durchsuchen, weshalb es dazu kommen kann, dass bestimmte Kombinationen von Peptiden nicht identifiziert werden. Dieser Algorithmus bietet sich besonders für große Sequenzdatenbanken an, da nicht alle möglichen Kombinationen von Peptiden durchsucht werden müssen.

Durchsuchen aller möglichen Kombinationen

Von den drei Varianten ist diese die aufwändigste Möglichkeit. Für jedes Vorläuferion müssen alle möglichen Kombinationen von vernetzten Peptiden durchsucht werden, um im Anschluss bewertet zu werden. Für identifizierte Kandidaten kann ein theoretisches Spektrum errechnet und mit dem gemessenen Spektrum verglichen werden. Bei dieser Methode wird keine Peptidkombination von vornherein ausgeschlossen. Die Stärken dieser Methode werden bei kleinen Datenbankgrößen deutlich. Mit optimierten Algorithmen ist diese Variante aber auch für große Datenbanken anwendbar.

StavroX und MeroX, die Programme, die im Folgenden vorgestellt werden sollen, verwenden eine Variante des zuletzt genannten Algorithmus. In dieser Arbeit sollen die wichtigsten Funktionen und verwendeten Algorithmen vorgestellt werden, die für die effiziente Auswertung von *Cross-Linking*/MS-Analysen nötig sind und in beide Programme integriert sind.

2. Ergebnisse

Für die Analyse massenspektrometrischer Daten sind anwendungsspezifische Programme unverzichtbar. Keine Software kann effizient das breite Spektrum an möglichen Anwendungen der Massenspektrometrie abdecken. Für die Kombination aus Protein-Protein-*Cross-Linking* und Massenspektrometrie wurden in dieser Arbeit zwei Programme entwickelt, die eine effiziente Analyse gemessener Datensätze ermöglichen sollen.

Cross-Linking-Experimente selbst können sehr variabel sein und sich z.B. in dem verwendeten *Cross-Linker*, den Proteasen, posttranslationalen Modifizierungen (PTM), dem verwendeten Massenspektrometer oder der angewendeten MS-Methode deutlich unterscheiden. Bei der Entwicklung der Programme wurde darauf geachtet, dieser Variabilität gerecht zu werden. Die beiden Programme wurden StavroX (griech. σταυρός = Kreuz) und MeroX (griech. μέρος = Teil) genannt, wobei das X für *Cross-Linking* steht.

2.1. Weiterentwicklung von StavroX und MeroX

StavroX wurde ursprünglich, bis zu Version 2 (Götze *et al.*, 2012), in der Sprache *Object Pascal* in Delphi 4 (Borland) entwickelt. Die darauf folgenden Versionen 3.X wurden komplett neu in Java implementiert. Java bietet den Vorteil, dass das Programm ohne Installation auf verschiedenen Plattformen (Windows, Linux, Mac) ausführbar ist. Das bedeutet auch, dass das Programm bei Bedarf auf einem Computer-*Cluster* verwendbar ist. Bei der Überführung zu Java wurde die Software erheblich in Bezug auf Geschwindigkeit, Benutzerfreundlichkeit, Kompatibilität und die Qualität der *Cross-Link*-Validierung verbessert. Zusätzlich entstand das Programm MeroX (Götze *et al.*, 2015), welches speziell für die Identifikation von *Cross-Links* mit spaltbarem *Cross-Linker* entwickelt wurde. Beide Programme ähneln sich stark in ihrer Funktionsweise und der grafischen Benutzeroberfläche (GUI - *graphical user interface*). Wenige, aber gravierende Unterschiede lassen sich in Algorithmen zur Identifizierung von *Cross-Links*, der Berechnung des *scores* und der grafischen Darstellung der Fragmente in den MS/MS-Spektren finden.

Im Folgenden sollen die Algorithmen zur automatischen Identifizierung und Validierung von *Cross-Links* sowie die Möglichkeiten der manuellen Validierung mit den grafischen Oberflächen von StavroX und MeroX erläutert werden.

2.2. Suchalgorithmen in StavroX und MeroX

Es ist prinzipiell denkbar, alle möglichen Kombinationen zweier Peptide als *Cross-Link*-Datenbank anzulegen (siehe Abschnitt 1.3.2, S. 89). Tatsächlich nutzten die ersten Versionen von StavroX genau diesen Ansatz, um Kandidaten zu identifizieren. Die Massen aller Peptidkombinationen wurden zwischengespeichert und dann mit der jeweils gemessenen Molekülmasse verglichen. Für komplexere Probleme mit mehreren Proteinen, zahlreichen Protein-Modifikationen oder mehreren Proteasen erhöht sich die Anzahl an theoretisch möglichen *Cross-Links* quadratisch mit der Anzahl der Peptide (Gleichung 6, S. 75). Dadurch erhöht sich auch die Größe der Datenbank, weshalb diese Vorgehensweise bei komplexen Analysen äußerst unpraktisch ist.

2.2.1. Standard-Suchalgorithmus von StavroX und MeroX

Der von StavroX und auch MeroX verwendete Algorithmus beruht auf einer Methode zur Identifizierung von Disulfiden mittels Massenspektrometrie (Choi *et al.*, 2010; Petrotchenko u. Borchers, 2014). Um festzustellen, ob es sich bei einem gemessenen Ion um einen möglichen *Cross-Link* handelt, werden zwei ineinander liegende Programmschleifen durchlaufen. In jeder der beiden Schleifen wird ein Peptid ausgewählt und dann die theoretische Masse eines *Cross-Links* der beiden Peptide mit der experimentell bestimmten Masse des Vorläuferions verglichen.

Das in Schleife 1 ausgewählte Peptid wird im Folgenden jeweils mit i nummeriert und das Peptid aus Schleife 2 mit j . Die Anzahl an theoretisch möglichen Peptiden (nach *in-silico*-Proteolyse) entspricht n . Die einfachste Möglichkeit wäre, zwei Schleifen zu benutzen, die jeweils alle Peptide umfassen. Dadurch könnten die Massen aller Peptidkombinationen mit der experimentellen Masse verglichen werden. Hierbei würde man genau n^2 Kombinationen durchsuchen müssen. Ein Problem dabei wäre, dass die Suche doppelte Kombinationen von Peptiden enthielte. Die Peptidkombination $\text{Peptid}_{i=5}$ mit $\text{Peptid}_{j=18}$ würde ebenso untersucht wie die Kombination $\text{Peptid}_{i=18}$ mit $\text{Peptid}_{j=5}$, obwohl beide Kombinationen identisch sind, da die Reihenfolge der beiden Peptide für die Masse des theoretischen *Cross-Links* irrelevant ist. Das Problem lässt sich umgehen, indem man die zweite Schleife nicht bei $j = 1$ startet, sondern bei $j = i$. Damit reduziert sich die Anzahl an möglichen *Cross-Links* auf ungefähr die Hälfte (Gleichung 6 und Abb. 30B). Doch das mathematische Problem ist immer noch quadratisch.

Durch einen einfachen Schritt lässt sich die Anzahl der durchsuchten Kombinationen zweier Peptide drastisch reduzieren: Die Liste der errechneten Peptide wird entsprechend der Masse der Peptide aufsteigend sortiert.

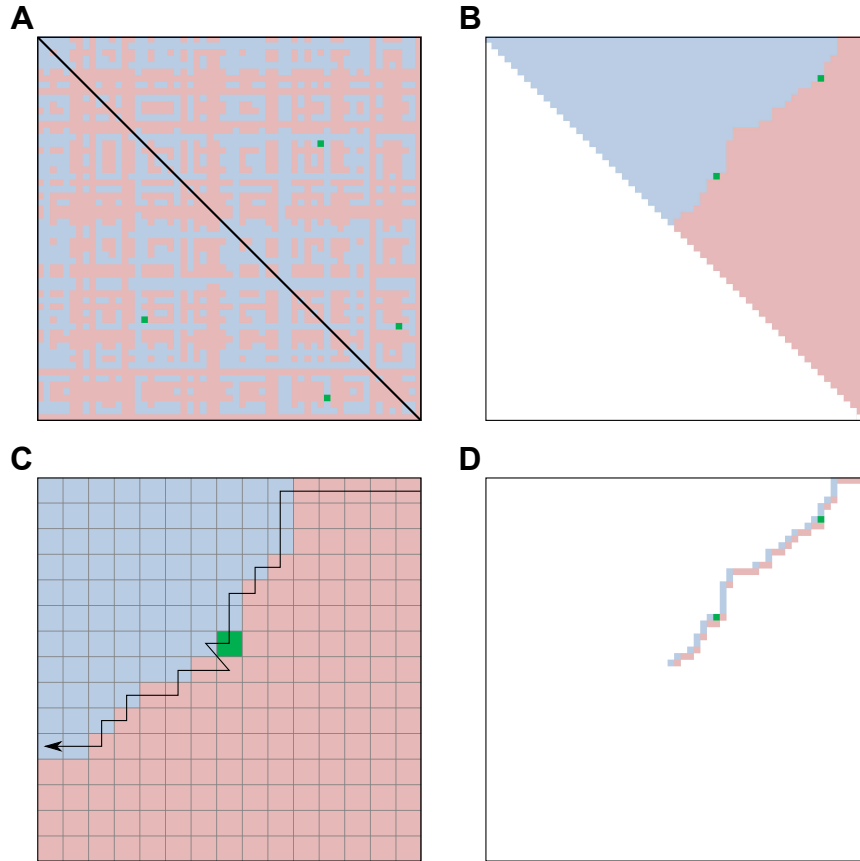


Abbildung 30: Identifizierungs-Algorithmus - Dargestellt sind symmetrische 60×60 Matrizen, die aus einer Liste zufällig generierter Peptidmassen erzeugt wurden. Jedes Feld in dieser 60×60 Matrix enthält die Summe der Massen der Peptide an der x- und y- Koordinate (m). In diesen Matrizen wurde nach einer zufälligen Masse m_x gesucht. Grüne Felder enthalten die gesuchte Masse ($m = m_x$). Die Masse in blauen Feldern ist geringer ($m < m_x$) und die Masse in roten Feldern ist größer als die gesuchte Masse ($m > m_x$). (A) Alle Kombinationen der Peptidmassen in einer unsortierten Matrix (Symmetrieachse eingezeichnet). (B) Sortierte Matrix, die doppelte Kombinationen ausschließt. (C) Detail der sortierten Matrix. Der Pfeil folgt den Feldern, die vom verwendeten Algorithmus tatsächlich geprüft werden, um *Cross-Links* zu identifizieren. (D) Matrix, die nur die Felder enthält, die vom Algorithmus tatsächlich verglichen werden (modifiziert nach Götze *et al.*, 2012).

Der Vorteil dieser Methode wird in Abbildung 30 deutlich. Dargestellt sind Summen-Matrizen aus 60 zufälligen Peptidmassen, die jeweils die x- und y-Koordinate der Matrix bilden. In jedem Feld ist die Summe der beiden zugehörigen Massen an der x- und y-Koordinate enthalten⁴ (m). In den Matrizen wurde eine zufällige Masse m_x gesucht. Stimmt die Masse m_x mit der Summe zweier Peptide in der Matrix m überein, wird das entsprechende Feld grün dargestellt. Ist die Summe größer, wird das Feld rot dargestellt, ist sie kleiner, wird es blau dargestellt. Felder sind weiß dargestellt, wenn sie in einem bestimmten Algorithmus nicht erst berechnet werden müssen. Matrix A (Abb. 30) entspricht einem Algorithmus, der $n^2 = 3600$ Kandidaten durchsucht. Man erkennt in der Matrix eine Spiegelachse, da es rechnerisch gleich ist, ob Peptid₅ mit Peptid₁₈ quervernetzt wird, oder andersherum. Ein Algorithmus, der doppelte Identifizierungen ausschließt, betrachtet nur eine Seite der Matrix. Damit wird die Anzahl auf 1830 Kandidaten reduziert (entsprechend Gleichung 6). Sortiert man zusätzlich die Peptide nach ihrer Masse, wird Matrix B erzeugt. Mögliche *Cross-Links* müssen sich an der Grenze zwischen dem blauen und roten Bereich befinden. Der von StavroX verwendete Algorithmus bewegt sich genau entlang dieser Grenze (Matrix C und D).

Die erste Programmschleife startet rechts in der obersten Zeile ($i = 1$ und $j = n$) (Abb. 30C). Der Algorithmus bewegt sich dann in der Matrix so lange nach links, bis eine Kombination von Peptiden erreicht ist, die eine geringere Masse aufweist als die gesuchte Masse m_x (blau). Die erste Programmschleife bewegt sich nun zur nächsten Zeile ($i = 2$). Die zweite Schleife startet dann aber nicht wieder ganz rechts, sondern am letzten Feld (j), das durchsucht wurde. Dieser Vorgang wird bis zur Spiegelachse der Matrix durchgeführt. Darüber hinaus würde es wieder zu doppelten Identifizierungen kommen. Im Falle einer Übereinstimmung (grün) wird sowohl im Feld links als auch unter dem identifizierten *Cross-Link*-Kandidaten nach einer weiteren möglichen Übereinstimmung gesucht. Dieser Algorithmus muss in dem hier zufällig erzeugten Beispiel nur 61 von 1830 verschiedenen Kombinationen prüfen und identifiziert dennoch dieselben Kandidaten (Matrix D). Damit wurde der Suchraum deutlich verkleinert, ohne die Größe des Lösungsraums zu verringern.

Dieser Algorithmus kann sowohl für homobifunktionelle als auch heterobifunktionelle *Cross-Linker* verwendet werden. Die Programme erzeugen zur Analyse eine Liste an Peptiden die wenigstens eine Aminosäure enthalten, die mit dem *Cross-Linker* reagieren kann. Alle anderen Peptide, die nicht vernetzt werden können, werden nicht weiter betrachtet. Bei einem identifizierten Kandidaten wird anhand der eingestellte *Cross-Linker*-Spezifität geprüft, ob der *Cross-Link* der identifizierten Peptide zulässig ist.

⁴Dies soll die Situation in einem *Cross-Link* simulieren. In der realen Analyse müssen selbstverständlich exakte Peptidmassen, Protonenmasse und die Masse des *Cross-Linkers* beachtet werden.

2.2.2. MeroX-spezifischer Suchalgorithmus: RISE-Modus

MeroX ist als Neuentwicklung aus StavroX hervorgegangen und für die Analyse von *Cross-Linking*-Experimenten mit einem MS/MS-spaltbaren *Cross-Linkern*, wie zum Beispiel dem Harnstoff-*Linker* (siehe Abschnitt 1.2.3, S. 82; Müller *et al.*, 2010) oder auch anderen spaltbaren *Linkern* wie DSSO (Kao *et al.*, 2010), geeignet. Ursprünglich wurde MeroX entwickelt, um die Validierung der erhaltenen Spektren zu vereinfachen und automatisch die zusätzlichen Fragmente, welche bei der Verwendung MS/MS-spaltbarer *Cross-Linker* entstehen (siehe 1.2.3), zu identifizieren. Diese charakteristischen Reporterionen können außerdem zur Identifizierung von möglichen *Cross-Links* benutzt werden. Die zusätzlichen *Cross-Linker*-Fragmente erzeugen ein für Peptide unnatürliches Muster im Spektrum, das bei optimaler Fragmentierung zwei Kriterien erfüllt (Abb. 31):

1. Für jedes der beiden Peptide entsteht ein Signalduplett mit einer Massendifferenz spezifisch für den *Cross-Linker* (z.B. 25.979 u für den Harnstoff-*Cross-Linker*).
2. Die beiden korrespondierenden Peptidmassen ergeben zusammen mit der Masse der *Cross-Linker*-Modifikation die Masse des Vorläuferions.

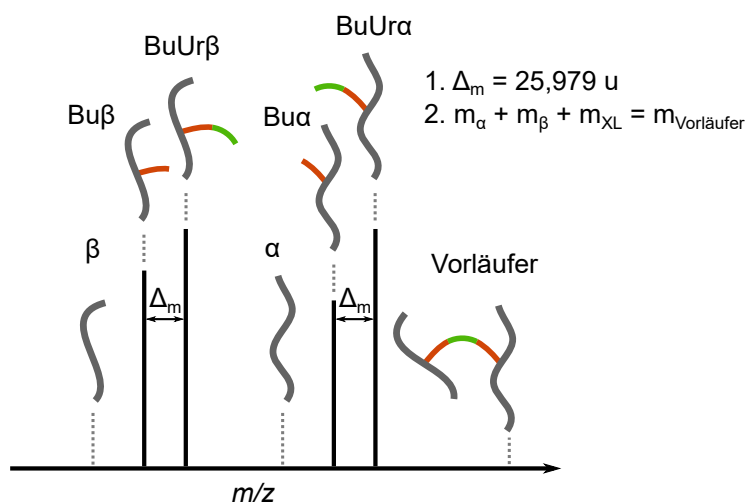


Abbildung 31: RISE-Modus - Schematische Darstellung eines Fragmentationsspektrums eines *Cross-Links* mit dem Harnstoff-*Linker* (siehe Abschnitt 1.2.3, S. 82). Lediglich die vier Reporterionen der *Cross-Linker*-Fragmentierung sind eingezeichnet. Der Abstand zwischen den Duplett-Signalen beträgt etwa 26 u. Die Fragmentationen entsprechen dem Peptid plus einem Teil des *Cross-Linkers* (Bu oder BuUr). Aus den Massen der Fragmente lassen sich die Peptidmassen von beiden Peptiden α und β ableiten. Wenn die Summe der Peptid-Massen und der *Cross-Linker*-Masse der des Vorläufers entspricht, wird dieses Spektrum als mögliches *Cross-Link*-Spektrum weiter analysiert.

Diese Regeln werden von MeroX als zusätzliches Suchkriterium verwendet und reduzieren damit drastisch die Größe des Lösungsraums. Im RISE-Modus (*Reporter Ion Scan Event*,

beschrieben in Arlt *et al.*, 2016) werden alle Spektren nach diesem spezifischen Muster durchsucht. Nur Spektren, die ein solches Muster enthalten, werden weiter untersucht. Wird ein solches Muster in einem Spektrum identifiziert, sind zugleich die exakten Massen der beiden vernetzten Peptide bekannt, da diese direkt aus den Massen der Duplett-Signale hervorgehen. Somit muss der Algorithmus nicht mehr alle Kombinationen von Peptiden absuchen. Der Algorithmus muss nach Auffinden zweier zusammengehöriger Dupletts lediglich das Vorhandensein der beiden identifizierten Peptidmassen überprüfen. Das bedeutet, dass auch der Suchraum stark verkleinert wird. Im RISE-Modus können deshalb deutlich größere Sequenzdatenbanken analysiert werden als im Standard-Modus. Mit einem ähnlichen Algorithmus wurden mit dem spaltbaren *Cross-Linker* DSSO bereits proteomweite Analysen ausgewertet (Liu *et al.*, 2015).

2.3. Validierung von *Cross-Links*

In einer einzigen Analyse können mehrere Tausend Signale anhand ihrer Masse als mögliche *Cross-Links* identifiziert werden. Doch nur ein kleiner Teil davon repräsentiert tatsächlich *Cross-Links*. Deshalb müssen die Kandidaten einem *scoring* unterzogen werden. Identifizierte *Cross-Link*-Kandidaten werden an einen eigenständigen Programmstrang (*thread*) weitergegeben, der parallelisiert ausgeführt werden kann (*multithreading*). Auf diese Weise können mehrere *Cross-Link*-Kandidaten gleichzeitig validiert werden. Die maximale Anzahl an gleichzeitigen *threads* ist dabei nur begrenzt durch die Anzahl an verfügbaren logischen Prozessoren des verwendeten Computers. Als Eingabe erhält der *thread* das gemessene MS/MS-Spektrum sowie den identifizierten Kandidaten aus zwei Peptiden und dem *Cross-Linker*. Der primäre Rückgabewert ist der *score*. Die Berechnung des *scores* basiert auf dem Vergleich gemessener Fragmentationen im MS/MS-Spektrum und der vorhergesagten Fragmentierung des *Cross-Link*-Kandidaten. Der *score* reflektiert dabei die Wahrscheinlichkeit dafür, dass der identifizierte Kandidat nicht zufällig vorkam.

2.3.1. *Score*-Berechnung

Der *score* ist einer der bedeutendsten Bestandteile der Programme StavroX und MeroX. Erst ein gutes *scoring* ermöglicht die Identifizierung von tatsächlichen *Cross-Links*. Die Berechnung des *scores* in beiden Programmen ist sehr ähnlich. Hier soll die Berechnung anhand des MeroX-*scores* erläutert werden. Für StavroX werden ähnliche oder gar dieselben Formeln benutzt, allerdings mit unterschiedlichen Werten für die verwendeten Variablen. StavroX und MeroX beziehen auch Signalintensitäten in den *score* ein. MeroX verwendet dazu die Intensitäten der Reporterionen, während StavroX die Intensitäten aller identifizierten Peptidfragmente einbezieht.

Zur Berechnung des *scores* werden zunächst alle möglichen Fragmentationen berechnet. Dabei werden von StavroX und MeroX lineare und vernetzte Fragmente entsprechend der Vernetzungsstellen der beiden Peptide theoretisch berechnet (siehe Abb. 26B). Es können verschiedene Ionentypen in den Einstellungen der Software gewählt werden. Zur Auswahl stehen a,b,c,x,y und z-Ionen (siehe Abb. 26A). Zwei weitere Ionentypen können bei Bedarf frei definiert werden. Verschiedene Fragmentierungsarten können in einem *Quick Setup Guide* direkt ausgewählt werden, um eine geeignete Auswahl an Ionentypen zu erhalten. Zusätzlich kann nach typischen Neutralverlusten der Ionen, zum Beispiel der Abspaltung von Wasser oder Ammoniak, als auch nach benutzerdefinierten Neutralverlusten gesucht werden.

Im Spektrum können die Ionen in verschiedenen Ladungszuständen vorliegen, so dass die Masse-zu-Ladung-Verhältnisse (m/z) aller Ionentypen für alle möglichen Ladungszustände (von 1+ bis zur Ladung des Vorläuferions) errechnet werden.

Das gemessene MS/MS-Spektrum wird vor der Analyse von Hintergrundsignalen (Rauschen) befreit. Dazu wird das Spektrum in 3 - 10 Abschnitte mit mindestens fünfzig Signalen eingeteilt und für jeden Bereich eine unabhängige Rauschberechnung durchgeführt. Zu Anfang wird der Mittelwert der Intensitäten aller Signale in einem Bereich gebildet. Die Signale, die über der benutzerdefinierten Signal-Rausch-Grenze – häufig einem Signal-zu-Rausch-Verhältnis (S/N) von mindestens zwei – liegen, werden als tatsächliche Signale angesehen und für die weitere Berechnung des Hintergrunds ignoriert. Mit allen übrigen Signalen wird der gleiche Vorgang so lange wiederholt, bis keines der gemittelten Signale mehr über der Signal-Rausch-Grenze liegt. Signale, die als Rauschen identifiziert wurden, werden nicht in die Berechnung einbezogen, im Spektrum aber als graue Signale eingezeichnet.

Alle möglichen Fragmentationen werden anschließend mit dem entrauschten Spektrum verglichen und verschiedene Faktoren (f_n) betrachtet, die zur Berechnung des *scores* herangezogen werden.

1. $f_{Hits/N}$ - Anzahl identifizierter Ionen relativ zur Anzahl an Datenpunkten
2. $f_{Hits/Ionen}$ - Anzahl identifizierter Ionen relativ zur Anzahl an theoretisch möglichen Fragmentationen
3. f_{Serie} - Ionenserien der einzelnen Peptide
4. f_{Xl} - Intensität von Reporterionen des *Cross-Linkers*

Anzahl der identifizierten Ionen ($f_{Hits/N}$ und $f_{Hits/Ionen}$)

Die Anzahl aller identifizierter Ionen wird entweder zur Anzahl an errechneten Ionen ($f_{Hits/Ionen}$) oder zu einer korrigierte Größe S des Spektrums ($f_{Hits/N}$) relativiert. Die Korrektur wird mit Gleichung 7 durchgeführt. Dies wirkt sich vor allem auf Spektren mit wenigen Signalen aus. Für eine kleine Anzahl an Datenpunkten (D) erhält man eine Spektrumsgröße (S) von ≈ 200 . Bei großen D wird der Zusammenhang zwischen D und S linear. Die Abhängigkeit der Spektrumsgröße S von der Anzahl an Datenpunkten D ist in Abb. 48 (Anhang S. LV) gezeigt. Zur Berechnung der Faktoren $f_{Hits/N}$ und $f_{Hits/Ionen}$ werden die Gleichungen 8 und 9 verwendet.

$$S = D - 400 + \sqrt{(400 - D)^2 + 200 \cdot D} \quad (7)$$

$$f_{Hits/N} = \frac{H}{S} \quad (8)$$

$$f_{Hits/Ionen} = \frac{H}{N_{Ionen}} \quad (9)$$

Mit:

S	Spektrumsgröße
D	Anzahl an Datenpunkten
H	Anzahl identifizierter Signale
N_{Ionen}	Anzahl an theoretisch möglichen Ionen

Ionenserien (f_{Serie})

Ionenserien sind mehrere Ionen gleichen Typs (z.B. γ -Ionen gleicher Ladung), die durch Spaltung an aufeinanderfolgenden Bindungen im Peptidrückgrat entstehen. Lange Ionenserien (ab einer Länge von ≈ 5) sind ein starkes Indiz für das Vorhandensein des jeweiligen Peptids. Sowohl die Länge (l) als auch die Anzahl von Ionenserien (n) fließen in den Faktor ein. Die Längen der Ionenserien werden quadriert, um lange Serien deutlich besser zu bewerten. Die Ionenserien werden für beide Peptide (α - und β -Peptid) getrennt betrachtet

$$f_{Serie} = \left(\frac{p_{\alpha}}{\sum_i^n (l_{\alpha,i})^2} + \frac{p_{\beta}}{\sum_i^n (l_{\beta,i})^2} \right)^{-1} \quad (10)$$

Mit:

l	Länge einer Ionenserie i des jeweiligen Peptids (α oder β)
n	Anzahl an Ionenserien insgesamt
p	Länge des jeweiligen Peptids (α oder β)

und auf die Länge des jeweiligen Peptids relativiert. Die Serien beider Peptide werden dann in f_{Serie} nach Gleichung 10 zusammengefasst.

Reporterionen der *Cross-Linker*-Fragmentierung (f_{Xl})

Die durch die Fragmentierung des *Cross-Linkers* entstehenden Reporterionen sind ein wichtiger Teil des MeroX-scores. Die Berechnung von f_{Xl} ist am Beispiel des Harnstoff-*Linkers* mit den Fragmentbezeichnungen Bu und BuUr in Gleichung 11 dargestellt. In allen Ladungszuständen (z) wird das Intensitätsprodukt von allen Reporterionen (Term 11a) mit den Intensitätsprodukten der beiden Duplett-Signale vom α -Peptid (Term 11b) und dem β -Peptid (Term 11c) addiert.

Das Fehlen eines Reporterions, z.B. des BuUr α -Ions, setzt die beiden Terme 11a und 11b auf null. Folglich ist f_{Xl} bei einem inkompletten Reporterionenmuster stark reduziert.

$$f_{Xl} = \sum_{z=1}^{z_{max}} (I_{Bu\alpha,z} \cdot I_{BuUr\alpha,z} \cdot I_{Bu\beta,z} \cdot I_{BuUr\beta,z}) \quad (11a)$$

$$+ I_{Bu\alpha,z} \cdot I_{BuUr\alpha,z} \quad (11b)$$

$$+ I_{Bu\beta,z} \cdot I_{BuUr\beta,z}) \quad (11c)$$

Mit:

z	Ladungszustand des Ions
I	Intensität des Ions

Vereinigung der Faktoren zum finalen *score*

Zur Verfeinerung des *scores* werden zusätzlich zum eigentlichen Spektrum noch verschobene Spektren verglichen. Dies entspricht einer vereinfachten Korrelationsanalyse, wie sie in der Bild- oder Tonanalyse angewendet wird, nur dass in StavroX und MeroX deutlich weniger Verschiebungen untersucht werden. Die Erwartung ist, dass für einen korrekten *Cross-Link* die beste Korrelation bei einem nicht verschobenen Spektrum auftritt.

Um die Korrelationsanalyse von MS/MS-Spektren zu verdeutlichen, wurde diese für zwei Spektren mit insgesamt 500 Verschiebungen um 0,01 Th durchgeführt. In Abb. 32 sind diese Analysen für $f_{Hits/N}$ eines guten (Abb. 32A, *score* von 161) und eines schlechten Kandidaten (Abb. 32B, *score* von 6) gezeigt. Es ist deutlich zu erkennen, dass in Abb. 32A der höchste Wert für $f_{Hits/N}$ bei einer Verschiebung von 0 Th auftritt. Nebenmaxima, die aus der natürlichen Isotopenverteilung resultieren, kommen bei ganzzahligen Verschiebungen vor. In Abb. 32B tritt zwar ein Maximum bei 0 Th auf, dieses ist allerdings weitaus schlechter

vom Hintergrund unterscheidbar als in Abb. 32A. Diese Analyse ist vor allem dann hilfreich, wenn ein Spektrum sehr viele Signale aufweist, die u.U. zufällig annotiert werden können, denn dann sollte es in allen Verschiebungen zu zahlreichen zufällig identifizierten Signalen kommen.

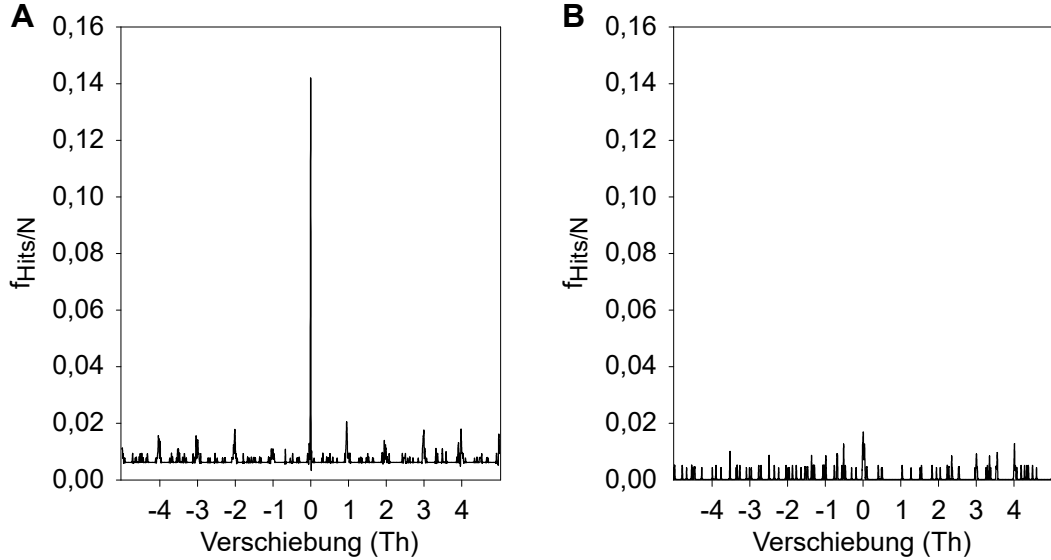


Abbildung 32: Korrelationsanalyse verschobener Spektren - Jeweils ein hochaufgelöstes MS/MS-Spektrum wurde mit einem *Cross-Link*-Kandidaten durch eine modifizierte Version von MeroX verglichen. Anstelle von vier verschobenen Spektren wurden jeweils 499 verschobene Spektren ($m/z_{shift} = 0,01 \text{ Th}$) analysiert. Das Ergebnis für $f_{Hits/N}$ (Gleichung 8) wurde gegen die Verschiebung aufgetragen. (A) Analyse für einen Kandidaten mit hohem *score* (161). (B) Kandidat mit einem niedrigen *score* (6). Die unterschiedliche Höhe der Basislinie in beiden Diagrammen ist darauf zurückzuführen, dass Vorläuferionen im Fragmentionen-Spektrum (z.B. Neutralverluste) von der Verschiebung ausgeschlossen werden, da diese für die Bewertung des *Cross-Links* irrelevant sind.

In StavroX und MeroX wird diese Methode vereinfacht angewendet, indem ein Spektrum bis zu sechsmal um einen fixen Wert ($m = 1.9 \text{ Th}$) verschoben wird ($-3m, -2m, -m, +m, +2m, +3m$). Der Wert von 1.9 Th wurde ursprünglich gewählt, um nicht auf eines der Nebenmaxima zu treffen und um außerhalb des Messfehlers einer linearen Ionenfalle zu liegen ($\approx \pm 0,8 \text{ Th}$), da sonst dasselbe Ion in zwei Verschiebungen gleichzeitig zugeordnet werden könnte. Für jedes dieser Spektren werden die zuvor beschriebenen Faktoren f_n errechnet. Aus den verschobenen Spektren wird der Mittelwert ($\bar{f}_n$) und die Standardabweichung (σ_n) errechnet und mit dem Wert für das ursprüngliche Spektrum ($f_{n,0}$) zu x_n verrechnet (Gleichung 12).

$$x_n = f_{n,0} - (\bar{f}_n + \sigma_n) \quad (12)$$

Alle *scoring*-Faktoren sind so angelegt, dass ein größerer Wert einen positiven Einfluss auf den finalen *score* hat. Um aus den absoluten x_n -Werten einen probabilistischen *score* zu erhalten, sollen diese in *subscores* (P_n) umgerechnet werden, welche eine Wahrscheinlichkeit dafür angeben, dass die Unterschiede zwischen den tatsächlich gemessenen und den verschobenen Spektren nicht zufällig sind. Um diese Wahrscheinlichkeiten errechnen zu können, wurde ein Datensatz mit charakterisierten *Cross-Links* analysiert und die einzelnen x_n -Werte für alle Spektren aufgezeichnet. Zufällige Identifizierungen wurden mit einer *Decoy*-Analyse simuliert (dazu mehr in siehe Abschnitt 2.3.2). Das Verhältnis von falschen zu korrekten Identifizierungen oberhalb eines bestimmten x_n -Wertes wurde über den Wertebereich von x_n errechnet und entweder an Gleichung 13a oder 13b angepasst. Die Ergebnisse dieser Anpassung sind in Tab. 6 zusammengefasst und bilden die Grundlage für eine effiziente Unterscheidung von falsch-positiven und richtig-positiven *Cross-Links*. Die Berechnung der *subscores* dient der Vereinheitlichung aller Faktoren auf eine besser vergleichbare Wahrscheinlichkeitsvariable.

$$P_n = e^{-a \cdot x_n} + l \quad (13a)$$

$$P_n = \frac{1}{1 + (a \cdot x_n)^b} + l \quad (13b)$$

Mit:

P_n	<i>subscore</i>
x_n	Korrigierter Faktor für die <i>Cross-Link</i> -Bewertung
a, b	Anzupassende Variablen
l	Limit zur Beschränkung des maximalen <i>score</i>

Tabelle 6: Faktoren zur Berechnung der *subscores* - Zur Berechnung der Signifikanz einer *Cross-Link*-Identifikation werden *subscores* auf Basis der Gleichung 13a oder 13b mit den hier angegebenen Faktoren a , b und l errechnet.

Symbol	Faktor	Gleichung	a	b	l
$f_{Hits/N}$	Anzahl identifizierter Ionen	(13a)	31,782	-	0,0001
$f_{Hits/Ionen}$		(13b)	3,448	6,667	0,0001
f_{Serie}	Ionenserien	(13b)	1,202	8,489	0,05
f_{Xl}	Intensität der Reporterionen	(13a)	35,008	-	0,0001

Je besser das Spektrum zum *Cross-Link*-Kandidaten passt, desto kleiner werden die *subscores*. Um den maximalen *score* sowie den Einfluss einzelner Faktoren auf den *score* zu limitieren, enthalten die Gleichungen 13a und 13b einen Term l , der das Minimum des *subscores* darstellt. Der eigentliche *score* eines *Cross-Link*-Kandidaten wird aus diesen vier Wahrscheinlichkeiten errechnet (Gleichung 14). Aufgrund der verschiedenen Limits in den Anpassungen (Gleichung 13) ergibt sich ein maximaler *score* von 200. Für *Cross-Links*, bei denen alle

subscores gleich 0.01 wären (vor Addition der Limits l), ergäbe sich ein *score* von 108 (bei *subscores* von 0.05 ergäbe sich ein *score* von 74).

$$Score = -15 \cdot \log_{10} \left(P_{Hits/N} \cdot P_{Hits/Ionen} \cdot P_{Serie} \cdot P_{Xl} \right) \quad (14)$$

2.3.2. Abschätzung der *False Discovery Rate* (FDR)

Ziel eines *scoring* der *Cross-Link*-Analyse muss sein, falsch-positive von richtig-positiven *Cross-Links* zu unterscheiden. Daraus folgt die Notwendigkeit einer hohen Sensitivität (richtig-positiv-Rate) bei hinreichend geringer FDR (*False Discovery Rate* - falsch-positiv-Rate). Die Sensitivität lässt sich nur bestimmen, wenn man alle richtig-positiven *Cross-Links* in der Analyse kennt. Daher lässt sich die Sensitivität in einer realen Analyse nur schwer automatisiert bestimmen. Für die Abschätzung der FDR hingegen wird in StavroX und MeroX automatisch eine *Decoy*-Analyse durchgeführt.

Falsch-positive *Cross-Links* können in massenspektrometrischen Analysen auf verschiedene Weise auftreten. Es ist z.B. möglich, dass die Kombination zweier Peptide zufällig die gleiche Masse eines tatsächlichen *Cross-Links* aufweist. Bei diesen falschen *Cross-Links* ist eine Übereinstimmung des Fragmentationenspektrums mit der theoretischen Fragmentierung allerdings unwahrscheinlich. Zudem kann es zur Identifizierung eines *Cross-Links* kommen, bei dem nur eines der beiden Peptide tatsächlich falsch und das andere korrekt ist. Bei kurzen Peptiden ist die Wahrscheinlichkeit recht hoch, dass ein isobares Peptid in der Sequenzdatenbank vorkommt, dass zur Identifizierung eines falsch-positiven *Cross-Links* führen kann.

Die *Decoy*-Analyse soll falsch-positive *Cross-Links* generieren und die erhaltenen *scores* für diese sicher falsch-positiven *Cross-Links* mit den *scores* der eigentlichen Analyse vergleichen. Um sicher zu gehen, dass falsch-positive Kandidaten erzeugt werden, wird zusätzlich eine Sequenzdatenbank mit falschen Proteinsequenzeinträgen generiert. Die Sequenzdatenbank mit den korrekten Einträgen wird als *Target*-Datenbank und jene mit falschen Proteinsequenzen als *Decoy*-Datenbank bezeichnet. Wenn ein *Cross-Link* zu einem Protein aus der *Decoy*-Datenbank identifiziert wird, kann man sicher davon ausgehen, dass dieser falsch-positiv ist. Die *Decoy*-Datenbank wird aus der eigentlichen *Target*-Datenbank erzeugt, um in etwa gleich viele, ähnlich zusammengesetzte Peptide zu erhalten. Dazu werden in StavroX und MeroX drei Möglichkeiten zur Verfügung gestellt:

1. Die Sequenz jedes Proteins wird invertiert.
2. Die Sequenz jedes Proteins wird gemischt.
3. Die Sequenz jedes Proteins wird gemischt, die Proteaseschnittstellen verbleiben aber an derselben Position wie in der ursprünglichen Sequenz.

Bei allen drei Optionen werden „falsche“ Sequenzen erzeugt. Das Mischen erfolgt dabei durch einen Pseudozufallsgenerator mit einem Startwert (*seed*), der aus der Proteinsequenz selbst berechnet wird. Das führt dazu, dass ein bestimmtes Protein in der Datenbank immer dieselbe *Decoy*-Sequenz erzeugt und somit die *Decoy*-Analyse reproduzierbar ist.

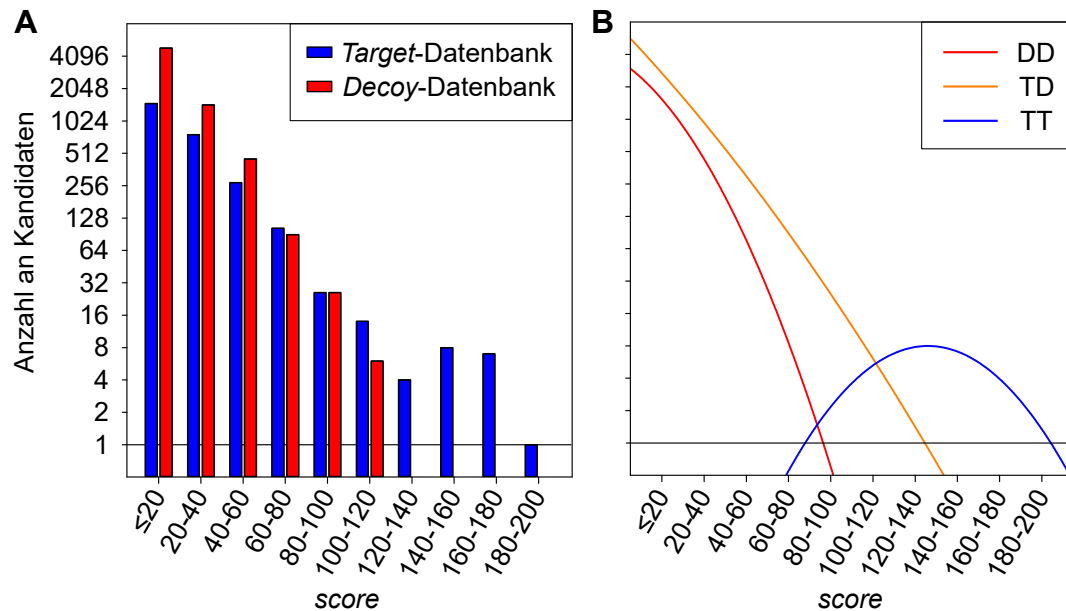


Abbildung 33: Automatische *Decoy*-Analyse - Parallel zur eigentlichen *Target*-Datenbank wird eine *Decoy*-Sequenzdatenbank analysiert, die falsche Proteinsequenzen enthält, so dass identifizierte *Cross-Links* mit Peptiden dieser *Decoy*-Datenbank als falsch-positiv angenommen werden können. (A) Histogramm der erhaltenen *scores* für die *Target*- und *Decoy*-Kandidaten. Für die hier gezeigte Analyse wurde ein *Cross-Link*-Datensatz des p53-Proteins (von Christian Arlt) mit aktiviertem RISE-Modus verwendet. Zur *Target*-Datenbank wurden außerdem 200 irrelevante *E. coli*-Proteinsequenzen hinzugefügt. Aus den Verteilungen wurde für eine $FDR < 5\%$ automatisch eine *score*-Untergrenze von 94 bestimmt. (B) Angepasste Verteilungen für die drei verschiedenen möglichen Kombinationen aus *Target*- (T) und *Decoy*-Peptiden (D).

Beide Datenbanken werden vereint und nach möglichen *Cross-Links* zwischen allen errechneten Peptiden gesucht. In Analysen von *Cross-Link*-Experimenten kommt es demnach zu drei verschiedenen Kandidatentypen:

1. **Full decoy (DD)** - Beide Peptide entstammen der *Decoy*-Datenbank.
2. **Mixed decoy (TD)** - In einem gemischten *Decoy*-Kandidaten stammt genau eines der beiden Peptide aus der *Target*-Datenbank und eines aus der *Decoy*-Datenbank.
3. **Full target (TT)** - Beide Peptide kommen in der *Target*-Datenbank vor.

Für die Anzahl der identifizierten Kandidaten in diesen Klassen erwartet man für eine zufällige Verteilung ein Verhältnis von 1:2:1. Für die Abschätzung der FDR für einen *score* s

wird Gleichung 15 angewendet. N entspricht dabei der Anzahl an Kandidaten der verschiedenen Typen, deren $score \geq s$ ist. Die Anzahl theoretischer *Decoy*-Kandidaten ($N_{DD} + N_{TD}$) ist etwa dreimal größer als N_{TT} . Um die Anzahl an falsch-positiven unter den *Full-target*-Kandidaten abzuschätzen, wird die Anzahl der identifizierten *Decoy*-Kandidaten für die FDR-Abschätzung in StavroX und MeroX gedrittelt.

$$FDR(score \geq s) = \frac{N_{DD} + N_{TD}}{3 \cdot N_{TT}} \quad (15)$$

Auch wenn für einen *Full-target*-Kandidaten beide Peptide aus der *Target*-Datenbank stammen, können dennoch Kandidaten mit einem oder auch zwei falschen Peptiden vorkommen. Die Verteilung der *Full-target*-Kandidaten entspricht demnach einer Überlagerung der drei genannten Kandidatentypen.

In Abb. 33B sind die *score*-Verteilungen dieser verschiedenen Kandidaten-Typen dargestellt. Diese wurden durch Anpassung der *score*-Verteilungen an eine Normalverteilung erzeugt. Bei Betrachtung der Verteilungen in Abb. 33B fällt auf, dass die *Full-decoy*-Verteilung (DD - *decoy-decoy*) einen Bereich mit deutlich geringeren *scores* abdeckt als die *Full-target*-Verteilung (TT - *target-target*) und beide Verteilungen kaum überlappen. Die Verteilung der gemischten Spezies (TD - *target-decoy*) verläuft zwischen diesen beiden Verteilungen. Dies ist auch so zu erwarten, da diese Kandidaten ein Peptid enthalten können, das tatsächlich in einem gemessenen *Cross-Link* vorkommt. Damit kann auch die Hälfte der möglichen Fragmentationen in einem Spektrum korrekt zugeordnet werden und der *score* ist höher als für *Full-decoy*-Kandidaten. Die *Full-target*-Verteilung (TT) zeigt ein Maximum bei einem *score* von etwa 150 und deckt einen Bereich von 80 - 200 ab. In diesem Bereich können richtig-positive *Cross-Links* erwartet werden, die aber zum Teil noch von falsch-positiven überlagert sein können. Im Bereich oberhalb von einem *score* von 120 - 140 kann man sehr wahrscheinlich ausschließlich richtig-positive Kandidaten identifizieren. Dies stimmt mit einer *score*-Untergrenze von 114 für eine FDR von 1% sehr gut überein.

Nach der erfolgten Analyse wird von StavroX und MeroX automatisch das Ergebnis der *Decoy*-Analyse in einem eigenen Fenster angezeigt. Darin ist die Verteilung der *scores* für die eigentliche Analyse und die *Decoy*-Analyse als Histogramm dargestellt. Ein solches Histogramm ist in Abb. 33A dargestellt. Für verschiedene *score*-Bereiche wurde die Anzahl an Kandidaten mit wenigstens einem *Decoy*-Peptid rot und mit zwei *Target*-Peptiden blau aufgetragen. Diese Analyse kann von Experiment zu Experiment stark variieren und hängt von sehr verschiedenen Faktoren ab. Eine deutliche Abhängigkeit ist z.B. für die Größe der Sequenzdatenbank zu beobachten. Der analysierte Datensatz (Arlt *et al.*, 2016) enthielt *Cross-Links* des p53-Tumorsuppressor-Proteins und wurde gegen eine Sequenzdatenbank analysiert, welche die p53-Sequenz und 200 weitere irrelevante Proteinsequenzen aus

E. coli enthielt. Die hinzugefügten Proteinsequenzen erhöhen die Wahrscheinlichkeit für falsch-positive Identifizierungen und sollten die Fähigkeit des *Scoring*-Algorithmus testen, zwischen falsch-positiven und richtig-positiven Identifizierungen zu unterscheiden. Ab einem Score von 114 konnten lediglich *Full-target*-Kandidaten identifiziert werden, so dass man ab diesem *score* von einer sehr geringen Falsch-Positiv-Rate ausgehen kann. Über einem *score* von 94 liegt die FDR bei 5%. Über einem *score* von 57 ist etwa jeder 4. Kandidat falsch-positiv.

Derselbe Datensatz, der für die *Decoy*-Analyse in Abb. 33 verwendet wurde, wurde mit verschiedenen Sequenzdatenbanken untersucht. Diese enthielten neben dem p53-Protein eine unterschiedliche Anzahl an *E. coli*-Proteinsequenzen und wurden mit verschiedenen Einstellungen durchgeführt (mit/ohne RISE-Modus oder mit/ohne die p53-Sequenz in der Datenbank). Die erhaltenen Ergebnisse sind in Tabelle 7 zusammengefasst. Aufgrund der quadratischen Steigerung der Anzahl möglicher Kandidaten hat die Größe der Sequenzdatenbank einen drastischen Einfluss auf die Sensitivität und FDR.

Tabelle 7: Einfluss der Größe der Sequenzdatenbank - Daten eines p53-*Cross-Link*-Experimentes (Arlt *et al.*, 2016) wurden gegen Sequenzdatenbanken mit MeroX analysiert. Die Datenbanken enthielten die Sequenz des p53-Proteins, sowie eine unterschiedliche Anzahl weiterer, irrelevanter Proteinsequenzen aus *E. coli*. Für jede Analyse ist die benötigte Berechnungszeit, die Anzahl theoretisch möglicher *Cross-Links*, die Größe des Lösungsraums, die *score*-Untergrenze bei einer FDR von 5% und die identifizierten *Cross-Links* mit einer FDR < 5% aufgeführt. In Klammern sind die identifizierten *Cross-Links* zu *E. coli*-Proteinen aufgezählt, die als falsch-positive *Cross-Links* angenommen werden können. -p53 zeigt an, dass die Datenbank kein p53 enthielt. -RISE zeigt an, dass diese Analyse nicht mit dem RISE-Modus durchgeführt wurde.

zusätzliche Sequenzen	Berechnungs- zeit (h:m:s)	Anzahl theor. <i>Cross-Links</i>	Größe des Lösungsraums	<i>score</i> - Grenze	ident. <i>Cross-Links</i>
0	00:00:15	128.896	78	0	78
0 -RISE	00:01:55	128.896	4.045	11	346
10	00:00:25	$\approx 2,26 \cdot 10^8$	127	0	87 (12)
10 -RISE	01:25:11	$\approx 2,26 \cdot 10^8$	526.740	72	98 (9)
100	00:00:33	$\approx 5,45 \cdot 10^9$	1.127	82	26 (2)
200	00:01:07	$\approx 3,14 \cdot 10^{10}$	2.687	94	25 (1)
500	00:08:58	$\approx 1,88 \cdot 10^{11}$	13.990	100	25 (2)
500 -p53	00:08:03	$\approx 1,86 \cdot 10^{11}$	12.241	77	2 (2)
1000	00:52:53	$\approx 5,48 \cdot 10^{11}$	41.865	117	18 (0)
2000	09:46:14	$\approx 2,04 \cdot 10^{12}$	153.069	119	17 (0)

Die Berechnungszeit bleibt bei kleineren Datenbanken sehr gering, im Bereich unter einer Minute. Bei der geringen Größe des Lösungsraumes für hundert zusätzliche Proteinsequenzen, bei aktiviertem RISE-Modus, ist die Identifizierung von möglichen Kandidaten limitierend für die Geschwindigkeit. Bei größeren Datenbanken steigt die Berechnungszeit überproportional an. Bei einer Verdopplung der Sequenzanzahl von 1000 auf 2000 Sequenzen wird die Berechnungszeit sogar verzehnfacht. Ab einer gewissen Größe der Datenbank wird die Berechnung des *scores* für alle Kandidaten des Lösungsraumes zum limitierenden Faktor für die Berechnungszeit.

Die wachsende Datenbankgröße hat außerdem einen negativen Einfluss auf die Anzahl der identifizierten *Cross-Links* mit einer $FDR < 5\%$. Die *score*-Untergrenze wurde immer weiter nach oben korrigiert, um diese FDR-Grenze zu erreichen, da mehr zufällige *Cross-Links* identifiziert wurden. Es fällt aber auch auf, dass weniger falsch-positive *Cross-Links* unter den identifizierten *Cross-Links* vorkamen (Tab. 7 Werte in Klammern). Die Verbesserung der Analyse durch den RISE-Modus wird ebenfalls deutlich: Bereits bei elf Proteinen in der Sequenzdatenbank dauerte die Analyse ohne den RISE-Modus etwa 1,5 Stunden (mit RISE-Modus nur 25 s). Analysen mit mehreren Hundert Proteinen in der Sequenzdatenbank sind ohne den RISE-Modus daher kaum sinnvoll möglich. Durch den RISE-Modus werden allerdings auch mögliche *Cross-Links* ignoriert. Das betrifft vor allem Spektren, bei denen nicht alle Reporterionen vorhanden sind und dennoch beide Peptide durch das Spektrum gut erklärt werden können (siehe Diskussion, Abb. 37, S. 120).

Verwendet man eine Sequenzdatenbank, die das analysierte Protein nicht enthält, erwartet man, keine sinnvollen *Cross-Links* zu erhalten. In einer Analyse mit 500 Proteinsequenzen, aber ohne die p53-Sequenz (500 -p53), wurden zwei *Cross-Links* mit *scores* von 81 und 83 mit einer $FDR < 5\%$ identifiziert. Betrachtet man die automatische generierte *Decoy*-Analyse (siehe Abb. 49 im Anhang S. LVI), fällt allerdings auf, dass diese *Cross-Links* vermutlich keine Relevanz haben.

MeroX ist also in der Lage, dank des RISE-Modus große Sequenzdatenbanken zu analysieren und im Bezug auf die angestrebte FDR konsistente Ergebnisse zu liefern. Es wird aber auch deutlich, dass mit sehr großen Datenbanken die Unsicherheit der Identifizierung von *Cross-Links* stark ansteigt, was zu einem Anstieg der *score*-Untergrenze führt, die das FDR-Kriterium gerade noch erfüllt. Es ist somit ratsam, die Sequenzdatenbank auf die in der Analyse tatsächlich enthaltenen Proteine zu limitieren. Im RISE-Modus von MeroX kann das allerdings dazu führen, dass kein einziger falsch-positiver Kandidat identifiziert wird (Dies war z.B. der Fall, wenn keine zusätzliche Sequenz hinzugefügt wurde, siehe Tab. 7 erste Zeile). In MeroX ist es deshalb möglich, zusätzlich zur eigentlichen Datenbank eine Sequenzdatenbank häufiger Proteinkontaminationen massenspektrometrischer Analysen (cRAP - *common Re-*

pository of Adventitious Proteins mit 116 Proteinsequenzen von www.thegpm.org/cRAP) zu verwenden. Dadurch wird sichergestellt, dass eine ausreichende Anzahl an *Decoy*-Sequenzen in der Datenbank vorhanden sind und eine zuverlässige FDR-Berechnung durchgeführt wird.

2.4. Beschreibung der Programmfunktionen

StavroX und MeroX sind frei verfügbar auf <http://www.StavroX.com/>. Die Programme besitzen eine detailliert ausgearbeitete grafische Oberfläche (GUI), so dass die gesamte Analyse direkt mit StavroX bzw. MeroX durchgeführt werden kann. Zudem werden auf der Webseite eine ausführliche Online Hilfe, ein *Quick User Guide* sowie Testdatensätze zur Verfügung gestellt, die den Einstieg in die Programme vereinfachen sollen. Neue Versionen werden mit regelmäßigen Updates auf der Webseite veröffentlicht. Beim Start prüfen die Programme automatisch, ob eine neue Version verfügbar ist.

2.4.1. Programmeinstellungen

StavroX und MeroX besitzen jeweils ein eigenes Dateiformat für die Einstellungsdateien (siehe Abschnitt 2.4.5). Im Vergleich zu anderen Programmen ist die Syntax von StavroX ungewöhnlich. Um Ungenauigkeiten bei der Eingabe von Aminosäuremassen oder *Cross-Linker*-Massen zu verhindern, müssen z.B. alle Moleküle mit ihrer chemischen Zusammensetzung angegeben werden. Die Programme errechnen dann aus genauen Elementmassen die Masse des eingegebenen Moleküls. Den Einstellungen können einfach Elemente (z.B. bisher nicht eingetragene Isotope) hinzugefügt werden.

Aminosäuren und Modifizierungen

StavroX und MeroX sind in der Lage, verschiedenste Aminosäuremodifizierungen zu beachten. Dazu muss die modifizierte Aminosäure neu definiert werden. Dies soll am Beispiel der Oxidation von Methionin erklärt werden: In der Liste verfügbarer Aminosäuren ist Methionin mit dem Einbuchstabencode 'M' (C_5H_9NOS) enthalten. Für oxidiertes Methionin wurde der Code 'm' ($C_5H_9NO_2S$) eingeführt. Auf diese Weise können alle denkbaren Modifizierungen von Aminosäuren definiert werden. Um diese auf die Analyse anzuwenden, wählt man die zusammengehörende unmodifizierte und modifizierte Aminosäure aus und fügt diese Kombination zur Liste der möglichen Modifizierungen hinzu. Dabei kann man zwischen statischen und variablen Modifizierungen wählen. Eine häufige statische Modifizierung ist die Alkylierung von Cysteinen, um die Bildung von Disulfidbrücken zu verhindern. Diese betrifft dann alle Cysteine in allen Proteinen. Bei variablen Modifizierungen kann man hingegen festlegen, wie viele Modifizierungen eines Typs maximal je Peptid vorkommen dürfen.

Proteaseschnittstellen

Ein großer Vorteil von StavroX und MeroX ist, dass beliebige Kombinationen von Proteasen definiert werden können. Dazu wird jede Proteaseschnittstelle einzeln mit der jeweils spezifischen Aminosäure oder Aminosäuresequenz definiert. Ein Fragezeichen deutet die Bindung an, die proteolytisch gespalten wird. Für Trypsin entspricht die Definition der Schnittstellen 'K?' und 'R?', da die Proteolyse C-terminal von Lysinen und Argininen stattfindet. Für AspN, eine Protease, die N-terminal von sauren Aminosäuren spaltet, wäre die Definition '?D' und '?E'. Für die PreScissionTM Protease (GE Healthcare) wäre die Definition entsprechend 'LEVLFQ?GP'. Für jede Proteolysestelle kann die maximale Anzahl an übersprungenen Proteolysestellen je Peptid eingestellt werden (*missed cleavage*). Bei Verwendung von unspezifischen Proteasen, wie z.B. Proteinase K, kann auch *in silico* eine unspezifische Proteolyse angewendet werden, bei der lediglich die minimale und die maximale Peptidlänge definiert werden muss. Auch eine Kombination aus spezifischer und unspezifischer Proteolyse ist möglich. Dabei muss wenigstens ein Terminus jedes Peptids durch eine sequenzspezifische Proteolyse entstanden sein.

Cross-Linker-Einstellungen

Ebenso können *Cross-Linker* frei definiert werden. Dazu wird die chemische Zusammensetzung der eingeführten *Cross-Linker*-Modifikation in einem Peptid-Peptid-*Cross-Link* im Vergleich zu den beiden freien Peptiden angegeben. Für die *Cross-Linker* DSS oder BS³ entspricht diese Zusammensetzung beispielsweise $C_8H_{10}O_2$ (138,068 u). Für *Zero-length-Cross-Links* wie Disulfide kann auch eine negative chemische Zusammensetzung angegeben werden: $-H_2$ (-2,016 u). Neben der Zusammensetzung müssen die Aminosäuren definiert werden, mit denen der *Cross-Linker* reagieren kann. Diese *Cross-Link*-Spezifität wird für die beiden reaktiven Gruppen unabhängig voneinander eingegeben, so dass auch heterobifunktionelle oder sequenzunspezifische (z.B. fotoaktivierbare) *Cross-Linker* betrachtet werden können.

Eingabe der Einstellungen in StavroX und MeroX

Viele weitere Details können in den Programmen definiert werden, so dass eine Analyse beinahe aller *Cross-Links* mit verschiedensten Geräten ermöglicht wird. Die grafische Oberfläche erlaubt die einfache Eingabe der verschiedenen Einstellungen in verschiedenen Tabellen und Textfeldern. Bewegt man die Maus über ein Feld (z.B. eine Checkbox oder eine Tabellenüberschrift), wird ein kurzer Text angezeigt, der die Funktion des Feldes erklärt. Genauere Informationen sind zusätzlich in der Online-Hilfe erklärt. Bei falscher Eingabe wird der Nutzer durch verschiedenste hilfreiche Fehlermeldungen auf den möglichen Fehler hingewiesen. Ist man mit den Einstellungen von StavroX oder MeroX noch nicht vertraut, kann man einen *Quick Setup Guide* verwenden. Dieser bietet eine Vorauswahl häufig verwendeter Proteasen

und *Cross-Linker* sowie Fragmentierungsarten und erstellt automatisch eine Einstellungsdatei, die später im Detail modifiziert werden kann, aber für die meisten Standardanalysen genügt.

2.4.2. Kommandozeilenoptionen

Für die Durchführung mehrerer Analysen ist es komfortabler, die eigentliche Analyse nicht mehrfach mithilfe der grafischen Oberfläche zu starten, sondern aus der Kommandozeile oder im *Batch*-Verfahren. Dabei werden Eingabedaten (Sequenzdatenbank, MS/MS-Daten und Einstellungen) sowie die Ergebnisdateien festgelegt, so dass im Anschluss eine automatisierte Abarbeitung aller eingegebenen Analysen erfolgt. Die Ergebnisse der einzelnen Analysen werden jeweils in separaten Ergebnisdateien gespeichert, so dass die Analyse im Nachhinein auch an einem anderen Computer validiert werden kann. Der Aufbau und der Inhalt einer Ergebnisdatei sind detaillierter in Abschnitt 2.4.5 beschrieben. Die automatisierte Durchführung von Analysen bietet sich vor allem dann an, wenn Analysen auf einem Computer-*Cluster* ausgelagert werden sollen und die Auswertung an einem Desktop-PC stattfindet.

2.4.3. Grafische Benutzeroberfläche (GUI)

Der Quelltext zu MeroX umfasst mehr als 23.000 Zeilen mit Programmanweisungen. Mehr als 11.000 davon werden allein für die grafische Darstellung und interaktive Benutzeroberfläche benötigt (Zum Vergleich: die Berechnung des *scores* umfasst etwa 700 Zeilen Quelltext). Die manuelle Validierung der von StavroX und MeroX ausgewerteten Daten ist nach wie vor ein zeitaufwändiger und unverzichtbarer Schritt bei der Analyse von XL-MS-Datensätzen. Die Programmoberfläche ist deshalb so gestaltet, dass man schnell einschätzen kann, ob es sich bei einem Kandidaten um einen *Cross-Link* handeln könnte oder nicht. Das Hauptfenster (Abb. 34) ist dafür in drei Hauptkomponenten gegliedert:

1. Eine Tabelle aller Kandidaten mit verschiedenen individuellen Informationen
2. Das annotierte Fragmentionen-Spektrum eines ausgewählten *Cross-Links*
3. Eine Fragmentierungsansicht, welche die Fragmentierungsstellen andeutet

Die Tabelle im Hauptfenster enthält alle wichtigen Informationen zu dem identifizierten *Cross-Link*-Kandidaten. Dazu zählen: Die errechnete und massenspektrometrische bestimmte Masse⁵ sowie die Abweichung voneinander, die beteiligten Peptide und deren Herkunft (Protein und Position im Protein), als auch Informationen zur Messung (Retentionszeit, Nummer des Spektrums). Man kann bestimmte Kandidaten mithilfe der

⁵sowohl das Masse-zu-Ladung-Verhältnis (m/z) und Ladung des Vorläuferions als auch die Masse im einfach geladenen Zustand

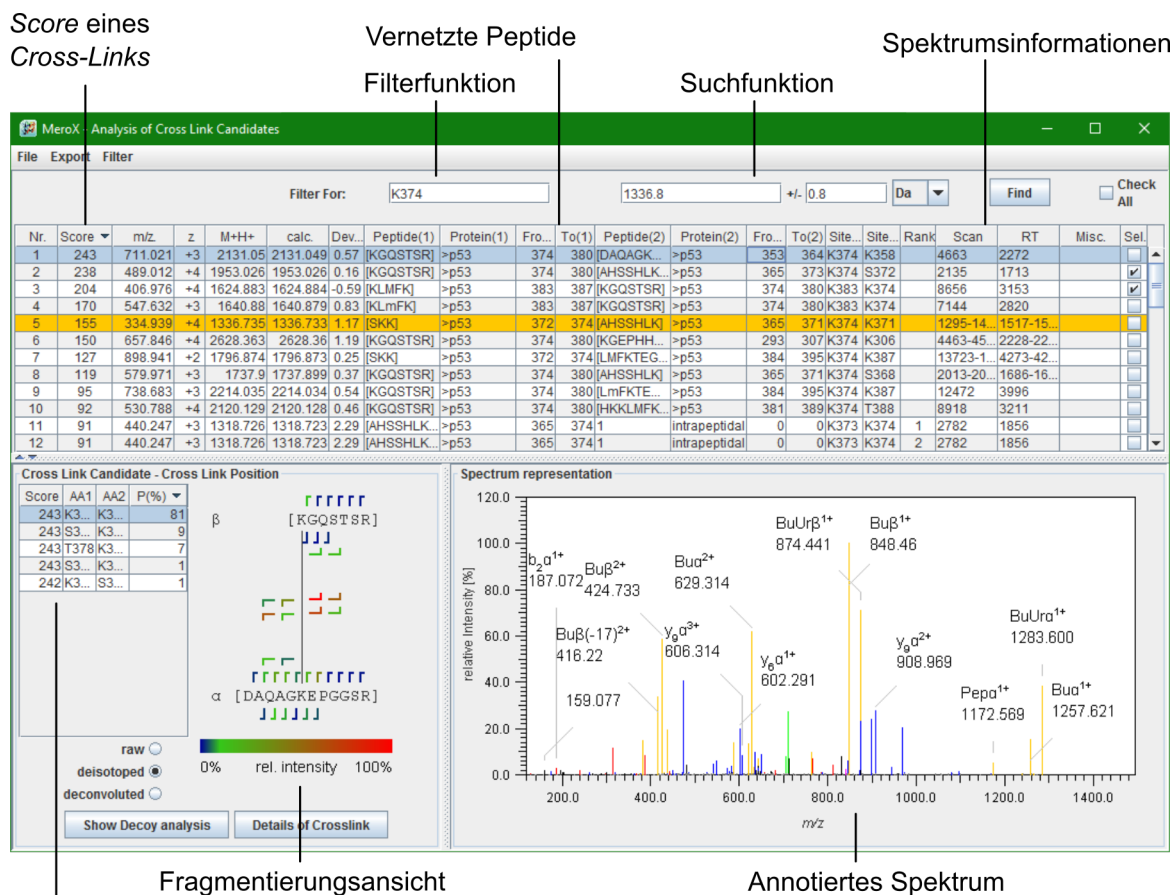


Tabelle der verschiedenen Kombinationen von *Cross-Link*-Stellen

Abbildung 34: MeroX-Hauptfenster - Screenshot des Hauptfensters nach der Analyse eines Datensatzes. Die Ansicht ist in drei Bereiche geteilt: Eine Tabelle aller Kandidaten (oben), eine Tabelle aller *Cross-Linking*-Stellen-Kombinationen mit der jeweiligen Fragmentierungsansicht (unten links) und das annotierte Spektrum (unten rechts).

Filter- und Suchfunktion hervorheben lassen. Alle Tabellen in StavroX und MeroX können gedruckt oder in andere Programme kopiert werden (z.B. Excel). In den letzten beiden Spalten können Kommentare zu einzelnen *Cross-Links* abgespeichert werden und einzelne Kandidaten mit einem Häkchen versehen werden, z.B. um diesen Kandidaten als möglichen *Cross-Link*-Kandidaten zu markieren.

Für zwei Peptide können sich verschiedene mögliche Kombinationen an *Cross-Link*-Stellen zwischen den beiden Peptiden ergeben. Die Programme errechnen *scores* für alle Kombinationen und diese können in der Tabelle links unten (Abb. 34) ausgewählt werden, so dass die annotierten Spektren verschiedener Kombinationen von *Cross-Linking*-Stellen unabhängig voneinander validiert werden können.

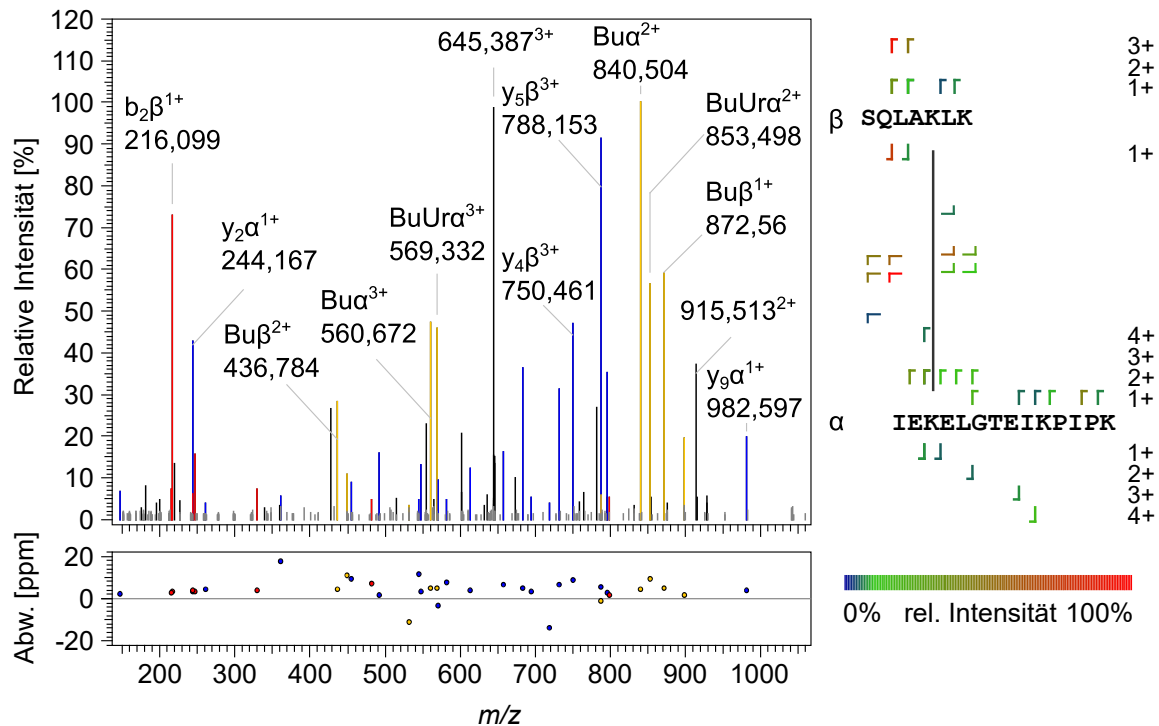


Abbildung 35: Annotiertes *Cross-Link*-Spektrum - Spektrum des identifizierten *Cross-Links* zwischen Peptiden aus Me31B (IEKELGTEIKPIPK) und Trailer Hitch (SQLAKLK) (siehe Abb. 20, S. 47). Die Signale im Spektrum sind annotiert und farbig markiert. *Cross-Linker*-Fragmente sind gelb dargestellt, N-terminale Peptid-Fragmente, wie b-Ionen sind rot und C-terminale Fragmente, wie y-Ionen, sind blau dargestellt. Unter dem Spektrum werden Abweichungen der identifizierten Signale zur berechneten Masse des Ions in *parts per million* (ppm) angezeigt. Rechts neben dem Spektrum zeigt die Fragmentierungsansicht einerseits die verknüpften Aminosäuren der beiden Peptide, als auch die identifizierten Fragmentierungsstellen. Die Fragment-Indikatoren ([und]) sind entsprechend der Intensität des korrespondierenden Signals eingefärbt (Die Farbskala ist unterhalb gezeigt).

Im unteren rechten Teil des Hauptfensters wird das Spektrum des aktuell ausgewählten *Cross-Links* und der ausgewählten Kombination an *Cross-Link*-Stellen dargestellt. Mit dem „Details of Crosslink“-Button gelangt man zu einer Detailansicht des ausgewählten *Cross-Links*, die neben einer tabellarischen Zusammenfassung der identifizierten Ionen ein detaillierteres Spektrum zeigt. In Abb. 35 ist ein von MeroX annotiertes Spektrum dargestellt.

Identifizierte Ionen werden im Spektrum farbig markiert, N-terminale Fragmente (a-, b- und c-Ionen) in Rot und C-terminale Fragmente (x-, y- und z-Ionen) in Blau. Das Vorläuferion (*precursor*) sowie dessen Neutralverluste (Verlust von Wasser oder Ammoniak) werden grün markiert. Nicht identifizierte Signale sind schwarz dargestellt und werden ebenfalls mit ihrem *m/z*-Wert (wenn bekannt, auch Ladung) beschriftet. Wenn ein Signal durch verschiedene Fragmentionentypen erklärt werden kann, wird es pink angezeigt.

Dieser Farbcode ist Standard für alle Spektren in StavroX. In MeroX gibt es zusätzlich gelb markierte Signale, die für Fragmentierungen innerhalb des *Cross-Linkers* verwendet werden. Es ist allerdings auch möglich, dieses Farbschema in den Einstellungen zu verändern. Farben für alle Ionentypen sind frei wählbar. Außerdem kann man die Fragmentfarbe entweder in Abhängigkeit vom Ionentyp oder nach ihrem Peptidursprung (α - oder β -Peptid) einstellen.

Spektrumsausschnitte können intuitiv mit dem Mausrad oder einem gezogenen Zoom-Fenster vergrößert werden. Durch einen Doppelklick gelangt man zurück zur ursprünglichen Vergrößerung. Unterhalb des Spektrums sind die Abweichungen aller gemessenen Signale dargestellt. Diese Informationen können hilfreich sein, um falsch zugeordnete Signale zu identifizieren oder auch, um einen Eindruck von der Qualität der Kalibrierung des Massenspektrometers zu erhalten.

Bei hochaufgelösten MS/MS-Spektren können die Spektren zudem deisotopiert werden. Beim Deisotopieren werden primär die schweren Isotopensignale der einzelnen Ionen entfernt, die hauptsächlich durch die zufällige Verteilung von ^{13}C -Isotopen in den Ionen auftreten. Die Intensitäten der Isotopensignale werden dann zum monoisotopischen Signal (das monoisotopische Ion enthält keine schweren Isotope) addiert. Aus dem Abstand der Isotopensignale lässt sich zusätzlich die Ladung des Ions ableiten. Dabei entspricht die Ladung des Ions etwa dem Kehrwert des Abstands aufeinanderfolgender Isotopensignale. Wenn MS/MS-Spektren vor der Analyse mit StavroX oder MeroX deisotopiert werden, geht die Ladungsinformation verloren, da diese in einigen Dateiformaten nicht implementiert wurde, demnach ist es von Vorteil, unbehandelte Datensätze zur Analyse zu verwenden. Bei der ebenfalls verfügbaren Dekonvolutierung werden die m/z -Werte aller Ionen mit bekannter Ladung auf eine einfache Ladung umgerechnet. In der grafischen Oberfläche ist es dann möglich, zwischen dem Originalspektrum, deisotopiertem und dekonvolutiertem Spektrum umzuschalten. Dadurch kann z.B. die automatische Ladungszuordnung nachträglich überprüft werden.

Rechts neben dem Spektrum wird eine detaillierte Fragmentierungsansicht angezeigt. Diese Anzeige ist eine schematische Präsentation der *Cross-Link*-Stellen sowie der Fragmentierungsstellen in den vernetzten Peptiden. Dafür werden beide Peptidsequenzen untereinander geschrieben und die beiden vernetzten Aminosäuren mit einer senkrechten Linie, die den *Cross-Linker* repräsentiert, verbunden. Fragmentationen werden dann mit [(C-terminale Ionen) und] (N-terminale Ionen) oberhalb und unterhalb der Sequenz dargestellt. Für StavroX (ab Version 3) und MeroX wurde diese Ansicht weiterentwickelt. Die Fragmentierungsindikatoren werden zusätzlich nach ihrer Intensität im Spektrum eingefärbt. Die Farbskala ist unten rechts dargestellt und verläuft von blau (minimale Intensität) über grün (relative Intensität von 10%) bis rot (rel. Intensität bis 100%). In MeroX werden

zusätzlich Indikatoren für die Spaltung des *Linkers* an der Linie zwischen den beiden Peptiden angezeigt. Für den Harnstoff-*Linker* entsprechen diese den Bu- bzw. BuUr-Fragmenten⁶.

Im Hauptfenster (siehe Abb. 34) ist eine vereinfachte Variante der Fragmentierungsansicht dargestellt, die in der Detailansicht eines *Cross-Link*-Kandidaten erweitert wird. In der vereinfachten Darstellung kann man nicht zwischen den Ladungszuständen der einzelnen Ionen differenzieren. In der Detailansicht (Abb. 35) sind die Ionen zusätzlich zur Farbkodierung noch in mehreren Reihen über- bzw. untereinander angeordnet, die jeweils einem der möglichen Ladungszustände der Ionen entsprechen. Die zugehörige Ladung eines Fragmentions ist jeweils rechts angezeigt.

Die Kombination aus annotiertem Spektrum und der farbig markierten Fragmentierungsansicht erleichtert es einerseits, Kandidaten von vornherein als *Cross-Link* auszuschließen, wenn z.B. keine oder nur wenige Ionen identifiziert wurden und keine Ionenserien erkennbar sind, andererseits kann man mögliche Kandidaten leichter erkennen, um sie später in der Detailansicht zu verifizieren.

2.4.4. Integration von xVis

Weder StavroX noch MeroX haben eine eigenständige Funktion zur visuellen Darstellung der identifizierten *Cross-Links* in der Proteinsequenz. In Grimm *et al.* (2015) wurde mit xVis ein Online-Tool vorgestellt, welches Diagramme zur Darstellung von *Cross-Linking*/MS-Daten erzeugt. xVis benötigt dazu lediglich eine korrekt formatierte Tabelle, die direkt aus StavroX oder MeroX (Ab Version 3.6 bzw. 1.6) heraus exportiert werden kann, so dass eine Darstellung der mit StavroX oder MeroX identifizierten *Cross-Links* mit xVis einfach möglich ist.

2.4.5. Dateiformate

StavroX und MeroX sind in der Lage, die gängigsten Eingabe-Dateiformate zu verarbeiten. So wird z.B. für Sequenzinformationen das Standard Fasta-Format verwendet. Für MS-Daten gibt es eine große Anzahl an verschiedenen Dateiformaten, nicht zuletzt aufgrund der Vielzahl an Herstellern von MS-Geräten. Rohdatenformate der meisten MS-Gerätehersteller (z.B. Thermo .RAW-Dateien) können zu vereinheitlichten Datenformaten konvertiert werden (MGF, mzXML und mzML). Diese Formate können für die Analyse verwendet werden, wodurch StavroX und MeroX Datensätze von verschiedensten MS-Geräten handhaben können. Für die Einstellungen von StavroX und MeroX wurde ein eigenes Datenformat entwickelt (SSF und MXF für StavroX bzw. MeroX). Darin sind alle

⁶Zusätzlich ist in Abb. 35 ein weiterer Fragmentionentyp am *Cross-Linker* angedeutet, welcher der Fragmentierung an der Amidbindung zwischen dem *Cross-Linker* und dem vernetzten Lysin entspricht.

Elemente, Aminosäuren, Proteasen, *Cross-Linker* usw. definiert. Mit einer abgespeicherten Einstellungsdatei kann man Datensätze unter reproduzierbaren Bedingungen analysieren.

Für die Ausgabe von *Cross-Link*-Daten gibt es bisher noch kein allgemein akzeptiertes Standardformat, wie z.B. mzIdentML für Proteomanalysen. Die meisten Programme exportieren die identifizierten *Cross-Links* in eine oder mehrere Textdateien. Im Fall von StavroX und MeroX wird eine einzige Ergebnisdatei erzeugt (Abb. 36). Die Ergebnisdatei enthält allerdings alle benötigten Informationen (Ergebnisse, Einstellungen und Spektren), um die Daten auch auf einem anderen Computer auswerten zu können oder auch die angewendete Suchkriterien nachzuvollziehen. Zur Verringerung des benötigten Speichers sind die Ergebnisdateien komprimiert (in einem ZIP-Archiv). Die Ergebnisdateien enthalten eine Dateistruktur, die mehrere Dokumente umfasst. Jedes einzelne Spektrum, für das ein *Cross-Link*-Kandidat identifiziert werden konnte, wird binär kodiert und mit einer einzigartigen ID versehen (UUID – *universally unique identifier*). So muss ein Spektrum nur einmal abgespeichert werden, auch wenn verschiedene Kandidaten für dasselbe Spektrum identifiziert werden, so dass der benötigte Speicherplatz für Spektren in der Ergebnisdatei reduziert wird. Die Ergebnistabelle enthält alle *Cross-Link*-Kandidaten mit den Informationen, die im Hauptfenster angezeigt werden. In jeder Ergebnisdatei werden die verwendeten Suchkriterien als Einstellungsdatei abgespeichert. Diese Funktion stellt sicher, dass immer die richtigen Einstellungen geladen werden, wenn eine Analyse im Programm geöffnet wird. Außerdem ist dies hilfreich, um zu überprüfen, unter welchen Bedingungen eine Analyse durchgeführt wurde oder um weitere Analysen unter den gleichen Bedingungen durchzuführen. Weiterhin werden die gesammelten Informationen für die FDR-Berechnung (*Decoy-Analyse*) und ein kurzer Report, der bei einer möglichen Fehlersuche helfen kann, abgespeichert. Die Proteinlängendatei ist für den xVis-Export nötig (siehe Abschnitt 2.4.4).

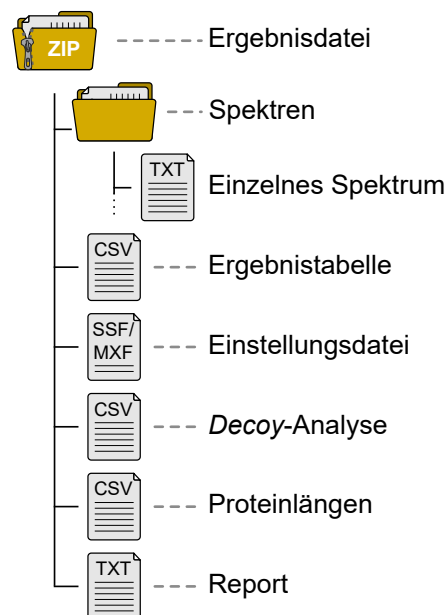


Abbildung 36: Struktur einer Ergebnisdatei

Sobald ein allgemein akzeptiertes Datenformat für XL-MS-Datensätze zur Verfügung steht, soll auch eine Funktion zur Konvertierung zu diesem Format in die Programme integriert werden. Dann könnten im Gegenzug auch Daten anderer Suchalgorithmen (z.B. Kojak) mit StavroX ausgewertet werden.

3. Diskussion

In den letzten 15 Jahren wurde eine Vielfalt von Software-Lösungen für die Auswertung von *Cross-Linking*/MS-Analysen vorgestellt. Anhand von Tabelle 5 wird diese Vielfalt deutlich. Beinahe jede Arbeitsgruppe, die an der Entwicklung von *Cross-Linking*/MS-Strategien arbeitet, hat auch eine eigene Software entwickelt, um der komplexen bioinformatischen Aufgabenstellung gerecht zu werden. Die meisten Programme führen eine Analyse durch und geben die gefundenen *Cross-Links* in Form einer Tabelle aus. Die einzige Möglichkeit, diese Daten zu validieren, ist der aufwändige Abgleich mit massenspektrometrischen Rohdaten, andernfalls muss man sich auf die angegebenen *Cross-Links* verlassen. Nur wenige Programme bieten sowohl die Analyse als auch ein Werkzeug zur Auswertung der *Cross-Links* an. Die in dieser Arbeit vorgestellten Programme sollen eine Plattform liefern, die einerseits einfach zu verwenden ist, zuverlässige und reproduzierbare Ergebnisse liefert, eine detaillierte Analyse zulässt und andererseits sehr flexibel ist.

StavroX wurde seit der ersten Version deutlich verbessert: während eine Analyse eines kleinen Datensatzes (≈ 1000 MS/MS-Spektren und *Cross-Links* eines Proteins) mit der ersten Version von StavroX noch eine halbe Stunde Berechnungszeit benötigte, ist die Analyse mit der aktuellen Version in wenigen Sekunden abgeschlossen. Auch eine detaillierte Validierung eines *Cross-Links* ist mittlerweile innerhalb weniger Minuten durchzuführen. Beide Programme sollen auch weiterhin gepflegt und optimiert werden, um die Auswertung weiter zu vereinfachen und der voranschreitenden Entwicklung der *Cross-Linking*/MS-Strategien gerecht zu werden.

Das Programm MeroX ist eines von drei Programmen (xLinkX und DXMSMS), welches die spezifischen Eigenschaften von spaltbaren *Cross-Link*-Reagenzien beachten und ausnutzen kann. Auch bei MeroX wurde auf eine hohe Flexibilität geachtet. So ist es einfach möglich, neue Reagenzien in die Einstellungen des Programms aufzunehmen und diese Substanzen zu testen (wie z.B. in Argo *et al.*, 2015). MeroX ist dabei nicht auf ein bestimmtes Spektrometer oder einen spezifischen Aufnahmemodus angewiesen, wie z.B. xLinkX (Liu *et al.*, 2015).

Der *scoring*-Algorithmus und vor allem die automatische FDR-Abschätzung sind wertvolle Bestandteile beider Programme. Der *score* ist in der Lage, falsch-positive Kandidaten von

richtig-positiven Kandidaten effizient zu unterscheiden, so dass nicht Tausende *Cross-Link*-Kandidaten manuell validiert werden müssen. Die Ergebnisse können nach einer geforderten FDR gefiltert werden, so dass nur die besten Übereinstimmungen angegeben werden, die mit hoher Wahrscheinlichkeit einem tatsächlichen *Cross-Link* entsprechen. Die grafische Benutzeroberfläche liefert dem Benutzer viele Möglichkeiten, die Daten im Detail zu untersuchen. Viele Komfortmerkmale, wie Sortier- und Filterfunktionen, einfaches Zoomen in Spektren, die automatische Annotation von Signalen oder der Export der Ergebnisse zu xVis erleichtern das Arbeiten mit großen Datensätzen ungemein. Bereits mit ein wenig Übung kann man aus der Kombination von annotiertem Spektrum und der eigens für StavroX und MeroX entwickelten Fragmentierungsansicht die Qualität eines *Cross-Link*-Spektrums beurteilen. Viele andere Programme zur *Cross-Link*-Analyse müssen mithilfe der Kommandozeile gestartet und ohne grafische Benutzeroberfläche mittels Einstellungsdateien konfiguriert werden, was für viele Benutzer häufig eine zu hohe Hürde darstellt. StavroX und MeroX besitzen hingegen einen „*Quick Setup Guide*“, mit dem man komfortabel die verwendeten Proteasen, *Cross-Linker* und Fragmentierungsarten einstellen kann. Ungeübte Nutzer können den Umgang mit der Software und den zahlreichen Funktionen durch die intuitive Benutzeroberfläche rasch erlernen.

Dieser Dissertation wurde eine CD beigelegt, die die aktuellen Versionen von StavroX (3.6.1) und MeroX (1.6.1) sowie Testdatensätze beinhaltet. Für diese Testdatensätze sind sowohl Ergebnisdateien zur sofortigen Betrachtung, als auch alle nötigen Eingabe- und Einstellungsdateien vorhanden, um die Analyse unabhängig zu wiederholen.

3.1. Anwendungsbeispiele der Software-Tools

Seit seiner Veröffentlichung 2011 wurde StavroX von verschiedenen Gruppen für diverse *Cross-Linking*/MS-Experimente angewendet. Eine der Stärken von StavroX ist die Freiheit, eine Vielzahl verschiedener *Cross-Linker* und *Cross-Link*-Strategien auswerten zu können.

So können z.B. gerichtet oder auch ungerichtet fotoaktivierbare Aminosäuren in Proteine eingebaut werden, um *Cross-Links* zu interagierenden Proteinen zu erzeugen, welche mit StavroX identifiziert werden können (Schwarz *et al.*, 2013; Konovalova *et al.*, 2014; Pettelkau *et al.*, 2014; Clavier *et al.*, 2015; Dehling *et al.*, 2016).

Eine der häufigsten Anwendungen stellt die Analyse von *Cross-Links* mit Isotopenmarkierten, homobifunktionellen, aminreaktiven *Cross-Linkern* wie z.B. DSS oder BS²G dar. Die gleichzeitige Analyse von leichtem und schwerem *Cross-Linker* steht derzeit nur für StavroX zur Verfügung, soll aber auch in MeroX integriert werden. So wurde zum Beispiel

eine Mischung aus DSS-d₀ und DSS-d₁₂ (mit 12 Deuteriumatomen anstelle von Wasserstoffatomen) verwendet, um die Interaktionsstelle von PetP (eine regulatorische Untereinheit des Cytochrom b₆f-Komplexes) mit dem Cytochrom b₆f-Dimer zu identifizieren. Ein *Cross-Link* von PetP zum N-Terminus der b₆f-Untereinheit VI konnte identifiziert werden. PetP bindet demnach an die zytoplasmatische Seite des membranständigen b₆f-Komplexes (Rexroth *et al.*, 2014). In einem sehr ähnlichen Experiment konnte in Kombination mit Kryo-Elektronenmikroskopie die Positionierung von Raf1 (einem *Rubisco assembly chaperone*) im Rubisco-Komplex während dessen Formierung aufgeklärt werden (Hauser *et al.*, 2015). Zudem wurde StavroX verwendet, um mögliche therapeutische Antikörper zu charakterisieren (Patentantrag Rosenthal *et al.*, 2016). Die Interaktion zwischen Epitop und Antikörper wurde dazu mit DSS-d₀/d₁₂ analysiert.

Neben der Isotopen-Markierung der *Cross-Linker* kann es auch nötig sein, Isotopen-markierte Proteine zu untersuchen. So kann mit einer Mischung aus markierten und unmarkierten Proteinen unterschieden werden, ob es sich bei einem *Cross-Link* um einen intramolekularen oder intermolekularen *Cross-Link* in einem Homo-Oligomer handelt (Taverner *et al.*, 2002). So wurden z.B. die Interaktionen innerhalb der Oligomere von ¹⁴N-p53/¹⁵N-p53 (Arlt *et al.*, 2015) und ¹⁴N-GCAP2/¹⁵N-GCAP2 (Pettelkau *et al.*, 2013) analysiert.

StavroX ist nicht beschränkt auf die Analyse von chemischen Quervernetzungen, sondern ist auch in der Lage, natürlich vorkommende *Cross-Links*, wie z.B. Disulfidbrücken, zu identifizieren. Eine weitere Gruppe natürlicher *Cross-Links* kann beispielsweise durch die Transglutaminase hervorgerufen werden, wobei es zur Bildung einer Isopeptidbindung zwischen Glutamin und Lysin unter Freisetzung von Ammoniak kommt. In Stamnaes *et al.* (2015) konnte die kovalente Selbstassoziation von Transglutaminase 2 (TG2) im Detail mit StavroX analysiert werden, wobei die Aminosäuren identifiziert wurden, die an der Selbstassoziation von TG2 beteiligt sind.

MeroX (frei verfügbar seit 18. November 2013) war die erste publizierte Software, die für die Analyse von *Cross-Linking*/MS-Experimenten mit einem MS/MS-spaltbaren *Cross-Linker* verwendet werden konnte. Seitdem wurden zwei weitere Programme publiziert: DXMSMS (Petrotschenko *et al.*, 2014), welches zusätzlich Isotopenmarkierungen berücksichtigt und XlinkX (Liu *et al.*, 2015). MeroX wurde erfolgreich für die Analyse verschiedener anderer spaltbarer *Cross-Linker*, wie z.B. SuDP (Disuccinimidyl-succinamyl-aspartyl-prolin in Argo *et al.*, 2015) und ähnlicher Derivate oder zusätzlich zur Analyse des radikalisch spaltenden TEMPO-*Linkers* (Hage *et al.*, 2016) verwendet. *Cross-Links* mit dem TEMPO-*Linker* konnten sowohl im Positiv- als auch im Negativ-Modus mit MeroX automatisch annotiert werden.

StavroX und MeroX erfüllen damit die Erwartung, für das breite Spektrum von möglichen *Cross-Linking*/MS-Anwendungen kompatibel zu sein. Beide Programme können einen wichtigen Beitrag zum Feld der strukturellen Proteomik liefern und machen die XL-MS-Strategie für ein breites Spektrum an Anwendern verfügbar.

3.2. Limitationen der *Cross-Link*-Analyse

Cross-Linking in Kombination mit Massenspektrometrie ist eine vielfältige Methode. Viele der Aminosäureseitenketten von Proteinen können durch spezifische und unspezifische Reagenzien chemisch vernetzt werden (siehe Abschnitt 1.2.1, S. 75). Verschiedene Proteasen und deren Kombinationen finden zudem Anwendung, um spezifische Proteolysen durchzuführen (Trypsin, GluC, AspN und weitere). Auch unspezifische Proteasen wurden bereits für XL-MS-Experimente verwendet (Petrotschenko *et al.*, 2012). Um eine gute Sequenzabdeckung beider vernetzter Peptide zu erhalten, können verschiedene Fragmentierungsmodi angewendet werden, zum Teil auch in Kombination miteinander (siehe Abschnitt 1.2.2, S. 79). Die Kombination verschiedener experimenteller und analytischer Ansätze kann dabei ausreichend komplementäre Daten liefern, um aussagekräftige Informationen über die meisten Proteine oder Proteinkomplexe zu sammeln.

Die *Cross-Linking*-Strategie wird hauptsächlich durch die Datenanalyse beschränkt. Die Limitation erfolgt dann, wenn zu viele Peptide nach der *in-silico*-Proteolyse kombiniert werden und der Such- und Lösungsraum zu groß werden. Dies kann z.B. durch eine große Sequenzdatenbank, die Verwendung unspezifischer Proteasen, die Einbeziehung vieler variabler Modifikationen oder einem sequenzunspezifischen *Cross-Linker* verursacht werden. Mit der Entwicklung spaltbarer *Cross-Linker*, den entsprechenden Workflows und spezialisierter Software, wie MeroX oder xLinkX, sind dennoch sehr komplexe Analysen realisierbar (Liu *et al.*, 2015).

In einer *Cross-Link*-Analyse machen tatsächliche *Cross-Links* nur einen kleinen Anteil der gemessenen Spezies aus. Es ist realistisch anzunehmen, dass ein Protein, für das ein *Cross-Link* identifiziert werden kann, auch durch mehrere unmodifizierte Peptide in der Analyse vertreten ist. Ein erhaltener Datensatz sollte deshalb zunächst mit Standard-Proteomsuchalgorithmen (z.B. MASCOT, SEQUEST) auf die enthaltenen Proteine untersucht werden. Für die *Cross-Linking*-Analyse sollte es dann ausreichen, die identifizierten Proteine oder nur die abundantesten Proteine als Sequenzdatenbank zu verwenden.

Eine weitere Limitation der Analyse mit StavroX und MeroX wird durch die *Cross-Link*-Typen vorgegeben. Neben den drei genannten Typen 0, 1 und 2 (siehe Abb. 23, S. 74) sind auch Kombinationen dieser Typen denkbar. *Cross-Links* mit mehr als zwei Peptiden kommen

sicherlich ebenso vor, werden aber von den Suchalgorithmen ignoriert. Eine Software, die die Suche nach drei vernetzten Peptiden erlaubt, ist MassAI (Rasmussen *et al.*, 2011). Dazu kann man MassAI auf Peptide limitieren, die auch als Typ 0 *Cross-Links* vorlagen, da ansonsten der Such- und Lösungsraum zur dritten Potenz anwächst.

3.3. Mögliche Optimierungen von StavroX und MeroX

Das erklärte Entwicklungsziel für *Cross-Linking*-Software ist es in naher Zukunft, so automatisiert wie möglich komplexe Datensätze auszuwerten. Dazu wird auch in den nächsten Jahren versucht werden, die Algorithmen so gut wie möglich zu optimieren, um die manuelle Auswertung auf das nötige Mindestmaß zu reduzieren. Dabei liegt ein Hauptaugenmerk bei der weiteren Entwicklung von StavroX und MeroX darauf, komplexere System mit mehr Proteinen analysieren zu können und den *score* weiter zu verbessern.

Die Fortschritte bei der Entwicklung von MS-Instrumenten und *Cross-Link*-Reagenzien haben ebenfalls einen großen Einfluss auf die zukünftigen Optimierungen. Die Möglichkeit, hochgenaue MS/MS-Spektren aufnehmen zu können, hat z.B. zur Einbeziehung von Deisotopierungsalgorithmen in StavroX und MeroX geführt. Beide Programme können aus den Rohdaten den Ladungszustand vieler Ionen durch das Isotopenmuster bestimmen, wodurch die Fehleranfälligkeit stark reduziert wurde, da nicht nur der m/z -Wert, sondern die tatsächliche Masse des Ions berechnet werden kann.

3.3.1. Proteomweite Analysen

MeroX und StavroX waren ursprünglich nicht für eine proteomweite Analyse von *Cross-Links* konzipiert. Deshalb ist es mit der aktuellen Version von MeroX nur sinnvoll, Proteinsequenzdatenbanken mit maximal ≈ 2000 Einträgen⁷ zu analysieren.

Diese Beschränkung ist derzeit limitiert durch die *in silico*-Proteolyse, da die Programme ursprünglich nicht für mehr als ≈ 100 Proteine konzipiert waren. Alle möglichen Peptide werden zunächst zwischengespeichert und verbrauchen in ihrer Summe viel Speicher, noch bevor die eigentliche Analyse beginnt. Das kann dazu führen, dass die zulässigen Speichergrenzen für die JAVA-Umgebung überschritten werden und das Programm abstürzt. Um proteomweite Analysen zu ermöglichen, müsste hier eine effizientere Speicherung der generierten Peptide implementiert werden. Dazu sollen zunächst externe Peptiddatenbanken in StavroX und MeroX integriert werden. Um letztlich proteomweite Analysen zu ermöglichen, müssen weitere Optimierungen mit realen Datensätzen durchgeführt werden, so dass verlässliche Ergebnisse erhalten werden können.

⁷Diese Angabe gilt bei aktiviertem RISE-Modus. Ohne RISE-Modus, in MeroX und auch StavroX, sollten maximal ≈ 50 Sequenzeinträge in der Datenbank vorhanden sein.

3.3.2. Optimierung der Suchalgorithmen

Auch wenn die Analyse von *Cross-Link*-Datensätzen im Vergleich zu früheren Versionen von StavroX schon stark verbessert wurde, gibt es verschiedene Ansätze, die verfolgt werden könnten, um die Analyse zu beschleunigen. Häufig werden bei *Cross-Link*-Analysen Messungen mit und ohne *Cross-Linker* durchgeführt. Diese könnten auch von der *Cross-Linking*-Software verwendet werden, um Spektren, die in beiden Proben vorkommen, von vornherein auszuschließen. Das Programm würde zunächst beide Datensätze ($\pm$ *Cross-Linker*) miteinander vergleichen. Zum Vergleich sollten der m/z -Wert und die Ladung des Vorläuferions, die Retentionszeit und auch das Spektrum selbst herangezogen werden. Wenn beide Analysen ein beinahe identisches Spektrum enthalten, kann es sich dabei nicht um ein Spektrum eines *Cross-Links* handeln. Dadurch sollte eine Vielzahl an irrelevanten Spektren aussortiert werden können, die Analyse beschleunigt werden und die Gefahr, falsch-positive Kandidaten zu identifizieren, minimiert werden.

Ein ähnlicher Effekt könnte erreicht werden, wenn man der *Cross-Linking*-Auswertung eine Peptid-Analyse vorschaltet und alle Spektren außer Betracht lässt, die eindeutig einem linearen Peptid zugeordnet werden können.

Der RISE-Modus in MeroX ist, so wie er hier beschrieben ist, sehr restriktiv. Denn es ist nötig, dass von beiden Peptiden beide Reporterionenpaare vorhanden sind. In Abb. 37 ist ein Spektrum eines p53-*Cross-Links* mit einem hohen *score* von 143 gezeigt. Dieses konnte allerdings nicht mit dem RISE-Modus identifiziert werden, da das intakte Reporterionenpaar des α -Peptids fehlt. Hier ist der Grund eine insgesamt zu starke Fragmentierung, so dass es zu Mehrfachfragmentierungen gekommen ist. Interne Fragmentionen des α -Peptids sind im Spektrum rot hervorgehoben. Eine Fragmentierung geschah im *Cross-Linker* und eine zweite im Peptidrückgrat, so dass es zur Bildung von y -Ionen mit den beiden *Cross-Linker*-Fragmenten Bu und BuUr kam.

Im RISE-Modus werden die Peptidmassen der beiden Peptide anhand der Reporterionenpaare berechnet. Für die Berechnung würde ein Reporterionenpaar allerdings ausreichen (siehe Gleichungen in Abb. 31). Der RISE-Modus wird um die Option erweitert werden, nur eines der beiden Reporterionenpaare zwingend zu benötigen. Die Größe des Such- und Lösungsraumes sollte dadurch nur geringfügig ansteigen, da der exakte Abstand der beiden Fragmente von 25,979 u keiner Aminosäure oder üblichen Modifikation entspricht und demnach nur bei *Cross-Link*-Spektren vorkommen sollte.

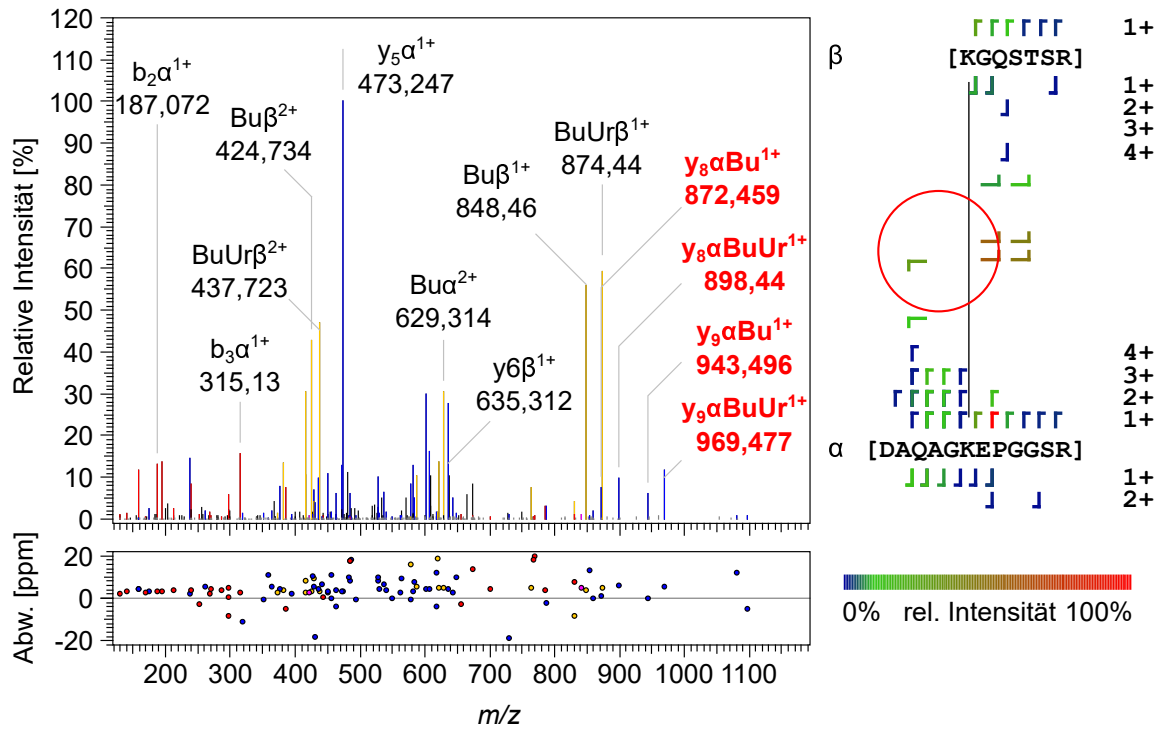


Abbildung 37: Verbesserung des RISE-Algorithmus - Gezeigt ist ein *Cross-Link*-Spektrum einer Quervernetzung zweier p53-Peptide (Daten von Christian Arlt). Das Spektrum wurde annotiert wie auf S. 112 beschrieben. Das gezeigte Spektrum wurde im RISE-Modus nicht identifiziert, da nur eines der beiden Reporterionenpaare vorhanden ist und das α -Ionenpaar fehlt (roter Kreis). Ursache ist in diesem Fall eine zu starke Fragmentierung, da es zur Bildung mehrerer interner Fragmente gekommen ist (in Rot hervorgehoben). Diese *Cross-Linker*-Fragmente fragmentierten zusätzlich im Peptidrückgrat (y-Ionen). Sowohl das Bu/BuUr-Paar für das y8-Ion als auch für das y9-Ion wurden automatisch von MeroX annotiert. Um ähnliche Spektren im RISE-Modus nicht zu verlieren, könnte man den RISE-Modus so verändern, dass nur eines der beiden Paare vorhanden sein muss, um als *Cross-Link* in Betracht gezogen zu werden.

3.3.3. Standard-Datenformate

Die *Proteomics Standard Initiative* (PSI; Orchard *et al.*, 2003) der *Human Proteome Organization* (HUPO) hat das Ziel, Standards für die Datenspeicherung im Feld der Proteomik zu entwickeln. Dies hat z.B. zur Entwicklung der offenen Datenformate mzML (Martens *et al.*, 2011) oder mzIdentML (Jones *et al.*, 2012) geführt. Obwohl es Vorschläge für ein einheitliches Ergebnisformat für *Cross-Linking*-Daten gibt (Riffle *et al.*, 2016; Hoopmann *et al.*, 2016), gibt es noch kein akzeptiertes Standardformat. Sobald ein Standardformat von der PSI akzeptiert wurde, soll auch die Möglichkeit des Datenexport und Import zu dem Standardformat implementiert werden.

3.4. Zusammenfassung

Die hier vorgestellten Programme StavroX und MeroX erleichtern die Analyse der komplexen Datensätze mannigfaltiger Varianten von *Cross-Linking*/MS-Experimenten erheblich. Die Programme sind öffentlich verfügbar und gut dokumentiert, so dass der Umgang mit den Programmen schnell erlernt werden kann. Als zwei von vielen *Cross-Linking*-Programmen zeichnen sich StavroX und MeroX besonders durch die einfache, intuitive Bedienung bei maximaler Flexibilität aus. Es konnte gezeigt werden, dass auch die Identifizierung von tatsächlichen *Cross-Links* durch den mittlerweile weit entwickelten *score* sehr gut funktioniert. Der FDR-Filter erlaubt es, nur die besten Übereinstimmungen auszugeben, so dass weniger Expertise benötigt wird, um *Cross-Link*-Datensätze auszuwerten. StavroX und MeroX stellen dem Nutzer eine Vielfalt an Funktionen zur detaillierten Auswertung zur Verfügung, so dass Datensätze ebenso erschöpfend ausgewertet werden können. Damit können StavroX und MeroX als verlässliche und einfach zu bedienende Programme mit hoher Flexibilität mit den vielen anderen Programmen konkurrieren. Nicht zuletzt wurden StavroX und MeroX seit der Veröffentlichung von StavroX.com über 5.700 mal aus 50 verschiedenen Ländern weltweit heruntergeladen (davon mehr als 4.500 mal von außerhalb Deutschlands).

Material & Methoden

M.1. Material

M.1.1. Chemikalien

Chemikalien wurden von der Firma Merck bezogen. Ausnahmen sind in der folgenden Tabelle 8 gelistet.

Tabelle 8: Liste verwendeter Chemikalien

Chemikalien und Feinchemikalien	
Chemikalie	Hersteller
4-(2-Aminoethyl)benzensulfonylfluorid (AEBSF)	AppliChem
β -Mercaptoethanol	Sigma-Aldrich
Aceton	Roth
Acrylamid 30% (37.5:1)	Serva
Acrylamid 40% (19:1)	Serva
Ameisensäure	Sigma-Aldrich
Aminosäuremix (je 1 mM)	Promega
Ammoniumhydrogencarbonat	Sigma-Aldrich
Coomassie G250	Serva
Dithiothreitol (DTT)	Gerbu
Essigsäure	Roth
Ethanol	Diagonal
Formaldehyd (37%)	Sigma-Aldrich
Formamid, deionisiert	Roth
Glycerin	Roth
Glycin	Serva
Guanidinthiocyanat	Roth

Fortsetzung auf der nächsten Seite

Tabelle 8 – Fortsetzung

Harnstoff	Roth
Harnstoff- <i>Cross-Linker</i>	Dr. Francesco Falvo (Universität zu Köln)
Hefe-RNA (total)	Roche
Heparin	Sigma-Aldrich
Iodacetamid	Sigma-Aldrich
IRGASAFE PLUS TM	Perkin Elmer
Isopropanol	Roth
Kreatinphosphat	Sigma-Aldrich
Magermilchpulver	Sucofin
Maltose	AppliChem
Natriumchlorid	Roth
Natriumdodecylsulfat (SDS)	Serva
Natriumhypochlorit (12% Lösung)	Roth
Natriumthiosulfat	Sigma-Aldrich
Nonidet TM P-40 (NP-40)	Sigma-Aldrich
Phenol	Roth
Transkriptionspuffer (5x)	Agilent
Trifluoressigsäure (TFA)	Sigma-Aldrich
tRNAs (Hefe)	Roche

Nukleotide

Chemikalie	Hersteller
α -[³² P]-UTP	Perkin Elmer
3'-O-Methyl-7mG(ppp)G Cap Analog	NEB
Biotin-16-UTP	Jena Biosciences
Nukleotid-Triphosphate	Sigma-Aldrich

Chromatografie-Material

Material	Hersteller
Amylose-Matrix	NEB
MonoQ Chromatographiesäule (1 ml)	Pharmacia

Fortsetzung auf der nächsten Seite

Tabelle 8 – Fortsetzung

Ni-NTA-Agarose	Qiagen
Protein A-Sepharose	GE Healthcare
Sephadex G50- <i>Fine</i>	GE Healthcare
Streptavidin-Sepharose	GE Healthcare

Proteine, Enzyme und Extrakte	
Protein/Enzym/Extrakt	Hersteller/Bereitgestellt von
Flag-Cup (0,8 μ M)	Dr. Claudia Temme
Flag-Smaug (1,1 μ M)	Dr. Claudia Temme
His-Belle (8,8 μ M)	Florian Kluge
His-Trailer Hitch (14 μ M)	Florian Kluge
Kaninchen-Retikulozytenlysats (RRL)	Promega
Kreatinkinase (10 g/l)	Sigma-Aldrich
MBP-PP7coat-BP (9,5 g/l)	Dr. Juliane Buschmann
Methyliertes BSA (5,3 g/l)	Dr. Anke Liepelt
PreScission TM -Protease	Dr. Claudia Temme
Pyrophosphatase (0,01 U/ μ l)	Fermentas
RNase Inhibitor (40 U/ μ l)	Dr. Bodo Moritz
SUMO-Protease (Ulp1)	Dr. Bodo Moritz
T3-RNA-Polymerase (20 U/ μ l)	Fermentas
Trypsin	Promega

Molekularbiologische Kits	
Kit	Hersteller
Luziferase Assay Reagenz (LAR)	Promega
MEGAscript T3 Kit	Life Technologies
ONE-Glo TM -Reagenz	Promega

M.1.2. Geräte

Spezielle Instrumente, die für die Analysen dieser Arbeit benötigt wurden, sind in Tabelle 9 aufgeführt.

Tabelle 9: Liste verwendeter Geräte

Gerät	Hersteller
Gradient Master (Gradienten-Mischer)	BioComp
Odyssey Classic (Fluoreszenz-Scanner)	LI-COR
Orbitrap Fusion Tribrid-Massenspektrometer	Thermo Fisher Scientific
Sirius <i>Single-tube</i> -Luminometer	Berthold
TriStar ² Multimodus-Lesegeräte	Berthold
Ultimate3000 RSLC Nano-HPLC-System	Dionex
Ultrazentrifuge L8-70M (Rotor SW40-Ti)	Beckman

M.1.3. Software

Für die Analysen der Arbeit und die Darstellung von Daten wurden verschiedene Software-Lösungen verwendet (Tab. 10). Vier der verwendeten Programme sind Teil dieser Arbeit, wobei StavroX und MeroX frei verfügbar sind (<http://www.StavroX.com/>).

Tabelle 10: Angewendete Software

Software	Verwendung
MaxQuant	Quantifizierung von MS-Daten (Cox u. Mann, 2008)
Xcalibur	Anzeige von Massenspektren (Thermo Fisher Scientific)
ImageQuant	Quantifizierung von <i>western blots</i> (Molecular Dynamics)
PyMol	Darstellung von Protein-3D-Strukturen (DeLano, 2006)
PONDR	Vorhersage unstrukturierter Proteinbereiche (Xue <i>et al.</i> , 2010)
In dieser Arbeit entstanden:	
StavroX	Auswertung von Peptid- <i>Cross-Links</i> (Götze <i>et al.</i> , 2012)
MeroX	Auswertung von Peptid- <i>Cross-Links</i> mit MS/MS-spaltbaren <i>Cross-Linkern</i> (Götze <i>et al.</i> , 2015)
SREfinder	Suche von SRE- <i>stem-loops</i> in Sequenzdatenbanken (siehe M.2.12)
GradientScan	Digitalisierung von UV-Signalen und Fraktionen der Gradientenzentrifugation

M.1.4. Plasmide

Zur rekombinanten Produktion von Proteinen und für die *In-vitro*-Synthese von RNA-Transkripten wurden verschiedene Plasmide verwendet (Tab. 11). Für die rekombinante Produktion von Me31B für die Antikörper-Herstellung wurde dazu das pET-SUMO-Me31B-MBP-Plasmid erzeugt (siehe Abschnitt M.2.1).

Tabelle 11: Liste verwendeter Plasmide - Plasmide von Mandy Jeske sind veröffentlicht in ^aJeske *et al.* (2006) und ^bJeske *et al.* (2011).

Plasmid	kodiertes Produkt	Erhalten/Hergestellt von
pET-SUMO-Me31B-MBP	His-SUMO-Me31B-MBP	M. Götze (Diese Arbeit)
pBSK-SREonly-nos	SREonly-WT-RNA	M. Jeske ^a
pBSK-SREonly-mut	SREonly-MUT-RNA	M. Jeske ^a
pBSK-1-AUG-nos	1-AUG-WT-RNA	M. Jeske ^b
pBSK-1-AUG-mut	1-AUG-MUT-RNA	M. Jeske ^b
pBSK-Luc-SRE-nos	Luc-WT-RNA	M. Jeske ^a
pBSK-Luc-SRE-mut	Luc-MUT-RNA	M. Jeske ^a

M.1.5. Antikörper

Für *Western-blot*-Analysen wurden in dieser Arbeit verschiedene primäre Antikörper verwendet (Tab. 12). Alle primären Antikörper wurden in TBST-Puffer mit 5% Milchpulver entsprechend der angegebenen Verdünnung verwendet. Zur Signaldetektion wurden fluoreszenzmarkierte sekundäre Antikörper verwendet (Tab. 13).

Tabelle 12: Primäre Antikörper - Die verwendeten Antikörper wurden veröffentlicht in ^aChartier *et al.* (2015); ^bHarnisch *et al.* (2016); ^cGötze *et al.*, eingereicht; ^dDuncan *et al.* (2009); ^eNakamura *et al.* (2004) und ^fVerrotti u. Wharton (2000).

Antigen	Organismus	Verdünnung	Erhalten/Hergestellt von
Smaug	Kaninchen	1:500	B. Moritz ^a
Me31B	Kaninchen	1:1000	M. Götze (Diese Arbeit) ^b
Tral	Ratte	1:1000	F. Kluge ^c
Belle	Kaninchen	1:1000	F. Kluge ^c
PABPC	Kaninchen	1:8000	M. Hentze ^d
Cup	Maus	1:5000	A. Nakamura ^e
	Kaninchen	1:1000	R. Wharton ^f

Tabelle 13: Sekundäre Antikörper - Sekundäre Antikörper wurden als fluorophormarkierte Antikörper (IRDye) von der Firma LI-COR bezogen. Sekundäre Antikörper wurden in TBST verdünnt und im Dunkeln mit den Membranen inkubiert, um ein Ausbleichen des Fluorophors zu verhindern.

Antikörper	Organismus	Verdünnung
α -Kaninchen 800CW	Esel	1:15000
α -Ratte 800CW	Ziege	1:15000
α -Kaninchen 680RD	Esel	1:15000
α -Maus 680	Esel	1:15000

M.2. Methoden

Biochemische Standardmethoden wurden nach Sambrook u. Russel (2001) durchgeführt. Im Folgenden sollen die Methoden im Detail beschrieben werden, die in dieser Arbeit optimiert oder etabliert wurden.

M.2.1. Erzeugung eines Me31B-spezifischen Antikörpers

Für die Herstellung des Me31B-Antikörpers wurde Me31B als Fusionsprotein kloniert. Dazu wurde die Me31B-cDNA in einen pET-SUMO-Vektor kloniert (pET-SUMO-Me31B, hergestellt von B. Moritz) und vor das Stoppcodon eine PreScissionTM-Schnittstelle, gefolgt von der kodierenden Sequenz des Maltose-Bindeproteins (MBP), eingefügt. Um möglichst native Termini des Proteins zu erhalten, konnten das SUMO-Protein mit der SUMO-Protease (Ulp1) und das Maltose-Bindeprotein durch die PreScissionTM-Protease abgespalten werden. Nach Expression und Aufschluss des Fusionsproteins wurden zwei aufeinanderfolgende Affinitätschromatografien durchgeführt (Ni-NTA- und Amylose-Chromatografie), die Fusionsanteile des Proteins abgespalten und das Protein durch eine Anionenaustauschchromatografie (Mono Q) weiter gereinigt. Die Antikörperproduktion wurde von der Firma Eurogentec mit dem gereinigten Protein durchgeführt. Die beiden erhaltenen Seren (#465 und #466) enthielten Me31B-spezifische Antikörper.

M.2.2. Präparation von *short*RNAs

Die Zugabe kurzer RNA-Fragmente (*short*RNAs) zum *Drosophila*-Embryoextrakt zeigte eine deutliche protektive Wirkung auf zugegebene RNAs (siehe Abb. 8, S. 21), vermutlich indem unspezifische RNasen abgesättigt wurden. Zur Präparation dieser *short*RNAs wurden 80 mg Hefe-RNA (total) in 5 ml Tris-HCl (20 mM, pH8) gelöst. Zu der RNA-Lösung wurden 200 μ l Natronlauge (2,5 M) zugegeben und für 50 min bei 40°C inkubiert. Die alkalische Hydrolyse wurde mit 200 μ l Salzsäure (5 M) für 10 min bei 40°C gestoppt und der Ansatz anschließend mit 600 μ l Natriumacetat (3 M, pH5) versetzt. Die RNAs wurden mittels Phenol-Chloroform

extrahiert und mit Isopropanol gefällt und bei einer Konzentration von 10 g/l bei -20°C gelagert. Die *short*RNAs (< 30 nt Länge) wurden zur Stabilisierung *in-vitro*-synthetisierter RNAs in einer Konzentration von 200 ng/ μ l in verschiedenen Reaktionen eingesetzt.

M.2.3. *In-vitro*-Synthese von RNAs

Um große Menge unmodifizierter RNAs zu erhalten, wurde die *In-vitro*-Synthese von RNAs mit dem MEGAscript® T3 Kit durchgeführt. Die Menge der einzelnen Nukleotide wurde dann an die Nukleotidverteilung in der jeweiligen RNA angepasst. Zur Synthese biotinylierter, bzw. auch radioaktiv markierter RNAs wurde der folgende Reaktionsansatz verwendet:

Tabelle 14: *In-vitro*-Transkriptionsreaktion - Zusammensetzung einer *In-vitro*-Transkriptionsreaktion. In Klammern sind Reagenzien angegeben, die zur Markierung der RNA Verwendung finden und nur eingesetzt wurden, wenn eine RNA biotinyliert oder radioaktiv markiert werden sollte.

Reaktionsansatz	
Transkriptionspuffer (5x)	1x
DTT	30 mM
NTP-Mix (ohne UTP)	5 mM
Magnesiumchlorid	10 mM
RNase Inhibitor	0,8 U/ μ l
Pyrophosphatase	0,01 U/Reaktion
T3-RNA-Polymerase	1,6 U/ μ l
Plasmid-DNA	60 ng/ μ l
UTP	siehe Text
(Biotin-16-UTP)	20 μ M
(α -[32 P]-UTP)	$\approx$ 10 μ Ci/Reaktion

Um eine vergleichbare Anzahl an Biotin- oder 32 P-Markierungen je RNA für die unterschiedlich langen RNAs zu erhalten, wurde die Menge an unmarkiertem UTP entsprechend dem zu synthetisierenden Transkript angepasst (SRE:0,16 mM 1-AUG:1 mM Luc:3,2 mM). Die Reaktion wurde anschließend für 2 - 4 h bei 37°C inkubiert und durch Inkubation mit DNaseI für 15 min bei 37°C abgestoppt. Zum Entfernen nicht eingebauter Nukleotide wurde entweder eine Sephadex G50-Chromatografie mit anschließender Phenol/Chloroform-Extraktion und Ethanol-fällung durchgeführt oder das Transkript gelelektrophoretisch aufgereinigt. Dazu wurde die RNA gefällt und auf ein 5%-Polyacrylamidgel (mit 8 M Harnstoff in 1xTBE) geladen. Nach der Elektrophorese wurde die Position der RNA im Gel autoradiografisch detektiert und das entsprechende Gelstück ausgeschnitten. Zur Gelelution wurde das Gelstück in 500 mM Ammoniumacetat, 10 mM Magnesiumacetat, 0,1 mM EDTA,

0,5% SDS über Nacht bei Raumtemperatur inkubiert und anschließend mit Ethanol gefällt. Die Quantifizierung der gereinigten RNA erfolgte entweder UV-spektrometrisch oder bei radioaktiv-markierten RNAs durch Vergleich von inkorporiertem radioaktivem UMP mit der spezifischen Aktivität des eingesetzten UTPs.

Synthese 5'-Cap-markierter RNAs

Für den Einbau von 5'-Cap-markierter RNA wurden 7 mM Cap-Analog (3'-O-Methyl-m⁷G(5')ppp(5')G) zu einem GTP-freien Reaktionsansatz gegeben. Die komplette Reaktion wurde 1 min bei 37°C inkubiert und anschließend durch die Zugabe von 2,5 mM GTP gestartet. Nach 45 min wurde erneut T3-RNA-Polymerase zugegeben, 1 min inkubiert und wieder mit 2,5 mM GTP gestartet. Die Reaktion wurde weitere 2 - 4 h bei 37°C inkubiert und anschließend wie zuvor beschrieben aufgereinigt und quantifiziert.

M.2.4. Präparation von *Drosophila*-Embryoextrakt (DEE)

Für die Präparation von Embryoextrakten mit hohen Translations- und Repressionsaktivitäten wurden Embryonen verwendet, die 15 min bis 2 h 15 min alt waren. Die Embryonen wurden durch drei aufeinanderfolgende Siebe gespült, um adulte Fliegen und Körperteile von Fliegen abzutrennen, und anschließend in ein kleineres Sieb überführt. Durch Zugabe von etwa 6%igem Natriumhypochlorit wurden die gesammelten Embryonen für 1 - 2 min dechorionisiert, bis sie deutlich klumpten. Anschließend wurden sie mit Wasser gespült, bis das Hypochlorit vollständig entfernt war. Die Embryonen wurden kurz mit destilliertem Wasser gewaschen und intensiv getrocknet. Im Anschluss wurden sie in einen Douncer-Homogenisator (1 ml; Wheaton mit *tight*-Pistill) überführt und je Gramm Embryonen wurde 1 ml kalter Extraktpuffer (30 mM Hepes pH7,4; 100 mM Kaliumacetat; 2 mM Magnesiumacetat; 0,5 mM AEBSF und 5 mM DTT) zugegeben und die Embryonen auf Eis aufgeschlossen, bis das Pistill bis zum Boden des Homogenisators reichte. Der erhaltene Rohextrakt wurde auf 400 - 500 µl Aliquots aufgeteilt und anschließend für 45 min zentrifugiert (20.000 g, 4°C). Der klare Extrakt unterhalb der Fettschicht wurde in ein frisches Röhrchen überführt und zusammen mit den ursprünglichen Aliquots erneut für 45 min zentrifugiert (20.000 g, 4°C). Ausschließlich der klare Überstand nach der zweiten Zentrifugation wurde vereint, passend aliquotiert, in flüssigem Stickstoff eingefroren und bei -80°C gelagert. Ein Aliquot jedes Extraktes wurde zum Testen des Extraktes verwendet. Getestet wurde die Translationsausbeute und die Repression von 2 nM Luziferase-SRE-RNA nach 20-minütiger Vorinkubation im Extrakt und 30-minütiger Translation.

Werden Extrakte aufgetaut, kann der klare Überstand nach einer halbstündigen Zentrifugation (20.000 g, 4°C) vereint, aliquotiert und erneut eingefroren werden, ohne dass ein Aktivitätsverlust zu beobachten ist. Extrakte eines Präparationstages oder verschiedener

Präparationen wurden daher vereint, um mehrere Experimente unter vergleichbaren Bedingungen durchführen zu können. Bevor Extrakte vereint wurden, wurden diese auf ihre Aktivität in Translation und Repression (siehe Abschnitt M.2.5) getestet. Ein typischer Extrakt wies eine Proteinkonzentration von 40 - 50 g/l und eine Repression der Translation nach Vorinkubation von mindestens 10 - 20× auf.

M.2.5. *In-vitro*-Translations- und Repressionsmessungen

Für Translationsmessungen wurden Luziferase-mRNAs in translationsaktiven Extrakten (aus *Drosophila*-Embryonen oder Kaninchen-Retikulozyten) inkubiert und die gebildete Luziferase durch den Umsatz von ATP und Luciferin luminometrisch bestimmt. Dazu wurden zwei verschiedene kommerzielle Substrate an zwei verschiedenen Luminometern verwendet. Für Reaktionen mit nur geringen Mengen an Translationsprodukten wurden 3 - 5 µl einer Reaktion mit 25 - 50 µl LAR (Luziferase Assay Reagenz, Promega) gemischt und die Lumineszenz sofort am Berthold Sirius *Single-tube*-Luminometer gemessen. Für alle anderen Reaktionen wurden 5 µl der Reaktion mit 30 µl ONE-GloTM-Reagenz (Promega) in einer Mikrotiterplatte gemischt und an einem Berthold TriStar² Multimodus-Lesegeräte vermessen. Die Berechnung des Repressionsfaktors aus den Translationswerten erfolgte anhand von Gleichung 16. Bei einem Repressionsfaktor von z.B. 20 würde man erwarten, dass 95% aller RNAs reprimiert werden und die restliche Translation auf die nicht reprimierten 5% zurückzuführen ist.

$$R_F = \frac{TL_{MUT}}{TL_{WT}} \quad (16)$$

Mit:

R_F	Repressionsfaktor
TL_{WT}	Translationsausbeute der Luc-WT-RNA in RLU/s
TL_{MUT}	Translationsausbeute der Luc-MUT-RNA in RLU/s

Translation in *Drosophila*-Embryoextrakt (DEE)

Für Messungen der Translation in *Drosophila*-Embryoextrakt wurden ein Mastermix und ein ATP-regenerierendes System (ARS) hergestellt (Tab. 15 - DEE). Das ARS wurde für die Aktivierung der Translation benötigt und beinhaltete Aminosäuren und als essentielle Komponente des ATP-regenerierenden Systems Kreatinphosphat. Die restlichen Komponenten des ATP-regenerierenden Systems, ATP und Kreatinkinase, waren bereits im Mastermix enthalten. Üblicherweise wurden 2 nM Luziferase-RNA in einer Translationsreaktion zum Mastermix gegeben und für 20 min bei 25°C inkubiert. Die Translationsreaktion wurde dann mit der Zugabe von ARS gestartet und nach einer 30-minütigen Inkubation bei 25°C durch Abkühlen auf Eis gestoppt.

Tabelle 15: Translationsreaktionen - Angegeben sind finale Konzentrationen der einzelnen Komponenten im Translationsreaktionsansatz in DEE bzw. RRL.

Mastermix (DEE)		Mastermix (RRL)	
DEE	40%	RRL	50%
Hepes pH7,4	16 mM	Kaliumacetat	79 mM
Kaliumacetat	50 mM	Magnesiumacetat	0,5 mM
Magnesiumacetat	1 mM	Kreatinphosphat	20 mM
tRNAs	0,25 g/l	Kreatinkinase	0,08 g/l
ATP	0,8 mM	Aminosäuremix	20 μ M
Kreatinkinase	0,08 g/l	DTT	2 mM
<i>short</i> RNAs	0,2 g/l	RNase Inhibitor	0,4 U/ μ l
RNase Inhibitor	0,4 U/ μ l	tRNAs	50 mg/l
ARS (DEE)			
Kreatinphosphat	20 mM		
Aminosäuremix	20 μ M		

Translation in Kaninchen-Retikulozytenlysate (RRL)

In Retikulozytenlysaten aus Kaninchen kann keine SRE-abhängige Regulation der Translation beobachtet werden (Smibert *et al.*, 1999). Die Translation in RRL eignet sich demnach, um die Menge und Qualität der WT- und MUT-RNAs zu vergleichen. Für eine *In-vitro*-Translation in RRL wurden 2 nM Luziferase-RNA in den angegebenen Reaktionsansatz (Tab. 15 - RRL) gegeben, für 30 - 60 min bei 30°C inkubiert und die Reaktion auf Eis abgestoppt.

M.2.6. Reinigung des reprimierten RNPs

Für die Reinigung des SRE-abhängigen RNPs wurden biotinylierte 1-AUG-RNAs (WT und MUT) synthetisiert (siehe Abschnitt M.2.3) und in Translationspuffer (TL-Puffer; Tab. 16) mit 200 ng/ μ l *short*RNAs, 1 mM DTT und 0,8 U/ μ l RNase-Inhibitor für 25 min bei 25°C inkubiert. Die dabei gebildeten RNA-Protein-Komplexe wurden anschließend entweder direkt einer Streptavidin-Affinitätsreinigung unterzogen oder zunächst durch Saccharosegradientenzentrifugation vorgereinigt.

Saccharosegradientenzentrifugation

Es wurden analytische sowie präparative Saccharosegradienten angefertigt. Diese hatten jeweils ein Gradientenvolumen von 12 ml und ein Konzentrationsgefälle von 5% - 45% (w/v) Saccharose in TL-Puffer (Tab. 16). Nach Vorinkubation der RNA in Embryoextrakt

Tabelle 16: Puffer für die Reinigung des reprimierten RNPs

Translationspuffer (TL-Puffer)		Wasch-Puffer (WB)	
Hepes pH7,4	16 mM	Hepes pH7,4	50 mM
Kaliumacetat	50 mM	Kaliumchlorid	150 mM
Magnesiumacetat	1 mM	Kaliumacetat	54 mM
ATP	0,8 mM	Magnesiumacetat	1 mM
		Heparin	30 $\mu\text{g/ml}$

wurden verschiedene Mengen des Ansatzes (0,2 - 1 ml) auf den Gradienten geladen. Die Zentrifugation erfolgte in einem SW40-Ti Rotor (Beckman) bei 40.000 rpm (285.000 g) für 3 h bei 4°C. Um ein Vermischen der Gradienten zu verhindern, wurde zunächst mit minimaler Beschleunigung eine Geschwindigkeit von 170 rpm eingestellt (Einstellung Acc-1) und dann bis 40.000 rpm normal beschleunigt. Nach der Zentrifugation wurde auf 170 rpm aktiv gebremst und dann die Zentrifugation ohne aktive Bremsung beendet (Einstellung Dec-1).

Die Gradienten wurden mit einer Glaskanüle vom Boden des Röhrchens her abgesaugt und fraktioniert (500 - 700 μl je Fraktion). Während des Fraktionierens wurde das UV-Profil ($\lambda = 280 \text{ nm}$) des Gradienten aufgezeichnet, um die Position der 80S-Ribosomen im Gradienten zu lokalisieren. Zur Digitalisierung wurde ein 12 bit USB-A/D-Wandler an den analogen Schreiber (Pharmacia) adaptiert, der neben der Absorption auch den Wechsel zwischen Fraktionen aufzeichnet. Die zur Digitalisierung nötige Software (GradientScan) wurde in Java programmiert. Bei Proben mit radioaktiven RNAs wurden Aliquots der Fraktionen durch Szintillationszählung (mit IRGASAFE PLUSTM) quantifiziert und zur eingesetzten Radioaktivitätsmenge relativiert. Die Fraktionen wurden in flüssigem Stickstoff eingefroren und bei -80°C für weitere Experimente gelagert.

Streptavidin-Biotin-*pull-down*

Zur Anreicherung der RNA-gebundenen Proteine wurden Streptavidin-Biotin-Affinitätsreinigungen (*pull-downs*) durchgeführt. Es wurden dafür stets Protein-LoBind-Röhrchen (Eppendorf) verwendet, um eine unspezifische Bindung der Proteine an die Wandungen der Röhrchen zu minimieren. Wenn nicht anders beschrieben, wurde die Matrix mit TL-Puffer (Tab. 16) gewaschen. Streptavidin-Sepharose (GE Healthcare) wurde dreimal gewaschen und mit 0,1 g/l Hefe-RNA (total) und 0,1 g/l methyliertem BSA in TL-Puffer für 30 min blockiert. Anschließend wurde die Sepharosematrix drei weitere Male gewaschen. Die Matrix wurde mit der biotinylierten RNA (in Embryoextrakt oder aufkonzentrierten Saccharosegradientenfraktionen) für eine halbe Stunde bei Raumtemperatur unter

Rotation inkubiert. Vor der Inkubation wurden 0,1 g/l Hefe-RNA (total) zugegeben, um unspezifische Protein-RNA-Interaktionen zu minimieren. Die Matrix wurde anschließend dreimal gewaschen und durch ein 30%iges Saccharosekissen (in TL-Puffer) zentrifugiert. Die Streptavidin-Sepharose wurde anschließend in ein neues Röhrchen überführt und ein weiteres Mal gewaschen. Anschließend wurde die Matrix mit Waschpuffer (Tab. 16; +0,1 g/l Hefe-RNA) für 15 min unter Rotation gewaschen. Nach einem letzten Waschschritt in Waschpuffer wurden die gebundenen Proteine in 10 mM Tris-HCl pH8 und 0,5% SDS für 10 min bei 80°C eluiert. Das Eluat wurde mittels Vakuumzentrifugation eingengt und in SDS-Ladepuffer aufgenommen.

Zur Analyse der gebundenen Proteine wurden anschließend SDS-Polyacrylamid-Gelelektrophoresen durchgeführt, die entweder mittels Coomassie (Wondrak, 2003), Silberfärbung (Nesterenko *et al.*, 1994) oder *western blot* (siehe Abschnitt M.2.9) analysiert wurden.

Pull-down zur RNA-Analyse

Für die Analyse SRE-spezifisch gebundener RNAs wurden 1-AUG-RNAs wie zuvor in Embryoextrakt inkubiert und ein Streptavidin-*pull-down* durchgeführt, ohne dass Hefe-RNA als Kompetitor zugegeben wurde. Für die Gelanalyse wurde die Matrix in Formamid-Ladepuffer für 5 min gekocht und der Überstand auf das Gel aufgetragen. Zur Präparation gebundener RNAs wurde die Matrix mit Trizol für 10 min bei 80°C inkubiert, anschließend wurden die RNAs extrahiert und mit Isopropanol präzipitiert.

M.2.7. *Deep-sequencing-Analyse*

Für die Analyse SRE-spezifisch gebundener RNAs wurden 500 ng der präparierten RNA eines RNA-*pull-downs* in die Synthese einer *Deep-sequencing*-Bibliothek eingesetzt. Die Synthese der Bibliothek sowie die Sequenzierung der RNAs wurden von Dr. Knut Krohn am interdisziplinären Zentrum für klinische Forschung Leipzig (IKZF) durchgeführt. Die Bibliothek-Synthese wurde mit dem TruSeqTM Small RNA sample prekit v2 (Illumina) durchgeführt. Die Sequenzierung von 100 nt langen Sequenzen erfolgte wie in Stokowy *et al.* (2014) beschrieben. Die Kartierung der erhaltenen *reads* und Quantifizierung identifizierter Transkripte wurde von Stéphanie Pierson vom Institut de Génétique Humaine (Montpellier, Frankreich) durchgeführt. Dazu wurden die 3'-*Linker*-Sequenzen mit Cutadapt (Martin, 2011) entfernt und die erhaltenen Sequenzen auf das *Drosophila melanogaster*-Genom mit dem Bowtie-Algorithmus (Langmead *et al.*, 2009) kartiert.

M.2.8. MS-Analyse

Für die Analyse des reprimierten RNPs wurden Streptavidin-*pull-downs* aus jeweils zwölf Gradientenzentrifugationsexperimenten (WT und MUT) vereint und mittels SDS-PAGE aufgetrennt. Die weitere Probenvorbereitung und massenspektrometrische Analyse wurde von Dirk Tänzler und Dr. Christian Ihling am Institut für Pharmazie der Martin-Luther-Universität Halle-Wittenberg durchgeführt. Die Gelspuren (vom WT- und MUT-*pull-down*) wurden in zwölf Stücke geschnitten und einer In-Gel-Proteolyse unterzogen (Shevchenko *et al.*, 2006). Zur Reduktion von Disulfiden wurden die Proben mit DTT inkubiert und anschließend mit Iodacetamid alkyliert.

Die erhaltenen Peptide wurden mittels Flüssigkeits-Umkehrphasenchromatografie (Ultimate3000 RSLC Nano-HPLC) getrennt und direkt mit der massenspektrometrischen Analyse durch eine Nano-Elektrospray-Ionenquelle gekoppelt. Die MS/MS-Analyse erfolgte an einem Orbitrap Fusion Tribrid-Massenspektrometer. Die Proben wurden auf eine Vorsäule geladen (Acclaim PepMap C8, 300 $\mu\text{m} \times 5 \text{ mm}$, 5 μm , 100 Å) und für 15 min mit 0,1% Trifluoressigsäure (TFA) bei einem Fluss von 30 $\mu\text{l/min}$ gewaschen. Anschließend wurden die Peptide von der Vorsäule auf eine äquilibrierte Trennsäule eluiert (Acclaim PepMap C18, 75 $\mu\text{m} \times 250 \text{ mm}$, 2 μm , 100 Å, äquilibriert in Fließmittel A: 0,1% Ameisensäure). Die Peptide wurden durch einen linearen Gradient von 0% - 35% des Fließmittels B (100% Acetonitril, 0,08% Ameisensäure) innerhalb von 90 min bei 40°C und einer Flussrate von 300 nl/min eluiert. Vorläuferionen-Spektren wurden im Orbitrap-Massenanalysator bei einer Auflösung von $R=60.000$ aufgenommen. Für 5 s wurden die intensivsten Vorläuferionen nacheinander selektiert, mittels HCD (30% normalisierte Kollisionsenergie) fragmentiert und in der linearen Ionenfalle detektiert (Top 5s-Methode).

Auswertung der massenspektrometrischen Daten

Die erhaltenen Rohdaten der massenspektrometrischen Analyse wurden mit MaxQuant (Version 1.5.2.8; Cox u. Mann, 2008) ausgewertet. Zur Identifizierung und markierungsfreien Quantifizierung von Proteinen im WT- und MUT-*pull-down* wurden die Rohdaten mit dem Proteom von *Drosophila melanogaster* (20.042 individuelle Proteineinträge, Zugriff vom 19.01.2015) abgeglichen. Als Maß für die Falsch-positiv-Identifizierung wurde eine *Decoy*-Analyse mit einer invertierten Sequenzdatenbank automatisch durchgeführt. Als maximale Abweichung von Ionen wurden 20 ppm für Vorläuferionen (Orbitrap-Analysator) und 0,5 Da für Fragmentionen (lineare Ionenfalle) verwendet. Es wurden die Alkylierung von Cysteinen als statische Modifizierung, die Oxidation von Methionin sowie eine N-terminale Acylierung als variable Modifizierungen angenommen. Neben den Proteinen des *Drosophila*-Proteoms wurde auch gegen bekannte Kontaminationen massenspektrometrischer Analysen automatisch durch MaxQuant abgeglichen. Diese Kontaminationen wurden allerdings nicht

in den Abbildungen 12A, 13, 18 und 40 einbezogen, da sie keine biologische Relevanz in der *nanos*-Regulation aufweisen.

Um Ubiquitinierungsstellen in den Repressor-Komplexproteinen zu identifizieren, wurden die Rohdaten in einer weiteren MaxQuant-Analyse lediglich mit den Sequenzen der sieben Repressor-Komplexproteine (Smaug, Me31B, Trailer Hitch, Belle, eIF4E und PABPC) abgeglichen. Die Analyse erfolgte wie beschrieben mit drei posttranslationalen Modifikationen an Lysinen: Diglycin-Modifizierung (+114,04 u), LRGG (+383,23 u) und Alkylierung (+57,02 u). Zur Darstellung der Spektren ubiquitinierten Peptide wurde das Programm StavroX verwendet.

M.2.9. *Western blotting*

Proteine in SDS-Polyacrylamid-Gele wurden über Nacht in 25 mM Tris und 192 mM Glycin auf methanol-aktivierte PVDF-Membranen transferiert. Die Membranen wurden in TBST mit 5% Milchpulver blockiert und für mindestens eine Stunde mit den jeweiligen primären Antikörpern inkubiert (siehe Tab. 12). Nach fünfmaligem Waschen wurden die Membranen mit sekundären, fluoreszenzmarkierten Antikörpern (siehe Tab. 13) inkubiert. Nach dreimaligem Waschen wurden die Membranen an einem Fluoreszenz-Scanner (LI-COR Odyssey Classic) eingelesen.

Quantitative Auswertung von *Western-blot*-Analysen

Zur Quantifizierung von *Western-blot*-Analysen wurden zum Vergleich Proteinstandards verschiedener Proteine mit bekannten Konzentrationen auf das SDS-PA-Gel geladen (Flag-Smaug, Flag-Cup, His-Belle, His-Tral). Die Konzentration von Me31B wurde in einem Extrakt durch Vergleich mit Me31B-MBP bestimmt. Dieser Extrakt wurde als Me31B-Standard für weitere Quantifizierungen verwendet. Die Konzentration der Proteinstandards wurde densitometrisch in Coomassie-gefärbeten SDS-PA-Gele bestimmt. Dazu wurden die Bandenintensitäten mit einem Standard bekannter Konzentration (MBP-PP7coat-BP oder BSA) verglichen.

Fünf verschiedene Mengen jedes Proteinstandards wurden auf ein SDS-PA-Gel aufgetragen und die Intensität der Signale nach erfolgter *Western-blot*-Analyse mit dem Programm ImageQuant densitometrisch bestimmt. Die erhaltenen Intensitäten der Standards wurden wie in Abb. 16D aufgetragen und nach einer linearen Anpassung die Konzentration der analysierten Probe aus deren Intensität bestimmt, wenn diese im linearen Bereich der Analyse lag.

M.2.10. Immunpräzipitation von Smaug

Protein A-Sepharosematrix ($3 \times 60 \mu\text{l}$) wurde dreimal mit TL-Puffer gewaschen und anschließend mit Anti-Smaug-Serum (zwei Ansätze) oder Präimmunserum ($20 \mu\text{l}$ Serum in $500 \mu\text{l}$ TL-Puffer, Tab. 16) 2 h im Kühlraum inkubiert. Nach zwei Waschschritten wurden zu allen Ansätzen $500 \mu\text{l}$ DEE und 1,5 ml TL-Puffer zugegeben und für weitere 2 h im Kühlraum inkubiert. Die Matrix wurde in Protein-LoBind-Röhrchen (Eppendorf) überführt und mehrmals gewaschen (zweimal TL-Puffer, dreimal mit WB-Puffer und noch zweimal mit TL-Puffer). Die gebundenen Proteine der Smaug-IP und Präimmun-IP wurden bis zur weiteren Verwendung auf Eis gelagert. Ein weiterer Ansatz wurde für eine *Cross-Linking*-Reaktion eingesetzt. Dazu wurde die Matrix in $500 \mu\text{l}$ TL-Puffer mit $0,5 \mu\text{M}$ Harnstoff-*Cross-Linker* inkubiert, um *Cross-Links* zwischen interagierenden Proteinen zu erhalten.

Die gebundenen Proteine wurden anschließend für alle drei Ansätze in 8 M Harnstoff und 0,4 M Ammoniumhydrogencarbonat denaturiert, Disulfidbrückenbindungen mit DTT (10 mM) reduziert und mit Iodacetamid (55 mM) alkyliert. Anschließend wurden die Ansätze zehnfach mit Wasser verdünnt und mit Trypsin ($0,5 \mu\text{g}$) über Nacht bei 37°C proteolysiert. Die noch enthaltene Sepharosematrix wurde abzentrifugiert, der Überstand eingengt und mit TFA (0,1%) angesäuert. Die massenspektrometrische Analyse erfolgte wie zuvor beschrieben (siehe Abschnitt M.2.8) mit der Ausnahme, dass Fragmentationen im Orbitrap-Massenanalysator ($R=60.000$) mit einer Top 5 s-Methode detektiert wurden. Die chemisch quervernetzte Probe wurde zusätzlich mit einer stufenweisen HCD-Fragmentierung analysiert ($30\% \pm 5\%$ normalisierte Kollisionsenergie). Die Identifizierung und Quantifizierung der Proteine in der Immunpräzipitation erfolgte wie zuvor beschrieben (siehe Abschnitt M.2.8).

M.2.11. Identifizierung von *Cross-Links* im Smaug-RNP

Zur Analyse möglicher *Cross-Links* in der Smaug-Immunpräzipitation wurden zwei Datensätze mit unterschiedlichen Fragmentierungsmodi aufgezeichnet: HCD mit 30% normalisierter Kollisionsenergie (50.653 Spektren) und *stepped*-HCD mit steigender normalisierter Kollisionsenergie (32.411 Spektren). Beide Datensätze wurden mit MeroX auf *Cross-Links* innerhalb des Repressor-Komplexes untersucht. Dabei wurden folgende Einstellungen verwendet: maximal drei übersprungene Proteaseschnittstellen für Trypsin, Cysteinalkylierung als statische Modifizierung, Methioninoxidation als variable Modifizierung, der Harnstoff-*Linker* als *Cross-Linker* mit Reaktivität gegenüber Lysinen und N-Terminalen Aminen, maximal 5 ppm Abweichung für Vorläuferionen und 10 ppm Abweichung für Fragmentationen (b- und y-Ionen) bei deaktiviertem RISE-Modus.

M.2.12. Suche nach SRE-*stem-loops* in 3'-UTRs von mRNAs

Um mögliche SREs in 3'-UTRs zu identifizieren, wurde ein Java-Programm geschrieben, welches nach den Konsensussequenzen des *stem-loops* CNGG oder UNGA in den 3'-UTR-Sequenzen von *Drosophila*-mRNAs sucht (Sequenzen von FlyBase.org, Revision 6.12 April 2016). Im Anschluss wird die mögliche Bildung von *stem-loops* getestet. Dazu wurden maximal 3 Lücken zugelassen. G-C Basenpaarungen werden mit einem Wert von 3 gezählt, A-U sowie G-U Basenpaarungen werden mit einem Wert von 2 gezählt. Die Werte für alle Basenpaarungen in einem möglichen *stem-loop* werden summiert. Für Fehlbasenpaarung wird ein Wert von 2 abgezogen. Für jede weitere Fehlbasenpaarung wird jeweils ein Wert von 4, 6, 8 usw. abgezogen. Es wurden *stem-loops* mit einer *loop*-Länge von 5 und 6 zugelassen.

Es wurden 1746 RNAs mit wenigstens einem SRE im 3'-UTR mit einer Bewertung von wenigstens 9 gefunden. Die beiden *stem-loops* in der *nanos*-mRNA wurden identifiziert mit Bewertungen von 22 und 13. Auch sieben der acht beschriebenen SRE-*stem-loops* im offenen Leseraster von *hsp83* konnten mit der beschriebenen Methode identifiziert werden.

M.2.13. Verwendete Bibliotheken in StavroX und MeroX

Zum Export von Grafiken, sowohl Pixelgrafiken als auch Vektorgrafiken, wurde die FreeHep VectorGraphics-Bibliothek in StavroX und MeroX integriert. Diese ist unter der GNU *lesser general public license*, Version 2.1 frei verwendbar. Diverse Anregungen und Quelltextzeilen wurden dem Forum auf Stackoverflow.com entlehnt. Alle auf Stackoverflow.com veröffentlichten Quelltexte sind der *Public Domain* verschrieben und unter der *creative common*-Lizenz ebenfalls frei verwendbar. StavroX und MeroX sind als *Freeware* lizenziert. Die Quelltexte von StavroX und MeroX sind nicht öffentlich gemacht.

M.2.14. Verwendete Datensätze zur Analyse mit MeroX

Zur Erstellung der Abbildungen 32 (Korrelationsanalyse), 33 (*Decoy*-Analyse und *score*-Verteilung), 37 (Optimierung des RISE-Modus) wurde ein Datensatz eines *Cross-Linking*/MS-Experimentes mit dem p53-Protein von Christian Arlt verwendet. Die Probenvorbereitung und Datenaufnahme sind in Arlt *et al.* (2016) beschrieben. Beim verwendeten Datensatz (4stepped.MGF, 14.762 MS/MS-Spektren) wurde als Fragmentierungsmethode eine stufenweise HCD-Fragmentierung angewendet (*stepped*-HCD).

Literaturverzeichnis

- [Ali-Murthy *et al.* 2013] ALI-MURTHY, Zehra ; LOTT, Susan E. ; EISEN, Michael B. ; KORNBERG, Thomas B.: An Essential Role for Zygotic Expression in the Pre-Cellular *Drosophila* Embryo. In: *PLOS Genet.* 9 (2013), Nr. 4, S. e1003428
- [Andersen *et al.* 2006] ANDERSEN, Christian B. F. ; BALLUT, Lionel ; JOHANSEN, Jesper S. ; CHAMIEH, Hala ; NIELSEN, Klaus H. ; OLIVEIRA, Cristiano L. P. ; PEDERSEN, Jan S. ; SÉRAPHIN, Bertrand ; LE HIR, Hervé ; ANDERSEN, Gregers R.: Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. In: *Science* 313 (2006), Nr. 5795, S. 1968–1972
- [Argo *et al.* 2015] ARGO, Andrew S. ; SHI, Chunxiao ; LIU, Fan ; GOSHE, Michael B.: Performing protein crosslinking using gas-phase cleavable chemical crosslinkers and liquid chromatography-tandem mass spectrometry. In: *Methods* 89 (2015), S. 64–73
- [Arlt *et al.* 2016] ARLT, Christian ; GÖTZE, Michael ; IHLING, Christian H. ; HAGE, Christoph ; SCHÄFER, Mathias ; SINZ, Andrea: Integrated Workflow for Structural Proteomics Studies Based on Cross-Linking/Mass Spectrometry with an MS/MS Cleavable Cross-Linker. In: *Anal. Chem.* 88 (2016), Nr. 16, S. 7930–7937
- [Arlt *et al.* 2015] ARLT, Christian ; IHLING, Christian H. ; SINZ, Andrea: Structure of full-length p53 tumor suppressor probed by chemical cross-linking and mass spectrometry. In: *Proteomics* 15 (2015), Nr. 16, S. 2746–2755
- [Askjaer *et al.* 1999] ASKJAER, Peter ; BACHI, Angela ; WILM, Matthias ; BISCHOFF, F. R. ; WEEKS, Daniel L. ; OGNIIEWSKI, Vera ; OHNO, Mutsuhito ; NIEHRS, Christof ; KJEMS, Jørgen ; MATTAJ, Iain W. ; FORNEROD, Maarten: RanGTP-Regulated Interactions of CRM1 with Nucleoporins and a Shuttling DEAD-Box Helicase. In: *Mol. Cell. Biol.* 19 (1999), Nr. 9, S. 6276–6285
- [Audhya *et al.* 2005] AUDHYA, Anjon ; HYNDMAN, Francie ; MCLEOD, Ian X. ; MADDOX, Amy S. ; YATES, John R. ; DESAI, Arshad ; OEGEMA, Karen: A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in *Caenorhabditis elegans*. In: *J. Cell Biol.* 171 (2005), Nr. 2, S. 267–279

- [Aviv *et al.* 2006] AVIV, Tzvi ; LIN, Zhen ; BEN-ARI, Giora ; SMIBERT, Craig A. ; SICHERI, Frank: Sequence-specific recognition of RNA hairpins by the SAM domain of Vts1p. In: *Nat. Struct. Mol. Biol.* 13 (2006), Nr. 2, S. 168–176
- [Aviv *et al.* 2003] AVIV, Tzvi ; LIN, Zhen ; LAU, Stefanie ; RENDL, Laura M. ; SICHERI, Frank ; SMIBERT, Craig A.: The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. In: *Nat. Struct. Biol.* 10 (2003), Nr. 8, S. 614–621
- [Ayache *et al.* 2015] AYACHE, Jessica ; BÉNARD, Marianne ; ERNOULT-LANGE, Michèle ; MINSHALL, Nicola ; STANDART, Nancy ; KRESS, Michel ; WEIL, Dominique: P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes. In: *Mol. Biol. Cell* 26 (2015), Nr. 14, S. 2579–2595
- [Ballut *et al.* 2005] BALLUT, Lionel ; MARCHADIER, Brice ; BAGUET, Aurélie ; TOMASETTO, Catherine ; SÉRAPHIN, Bertrand ; LE HIR, Hervé: The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. In: *Nat. Struct. Mol. Biol.* 12 (2005), Nr. 10, S. 861–869
- [Bantscheff *et al.* 2012] BANTSCHEFF, Marcus ; LEMEER, Simone ; SAVITSKI, Mikhail M. ; KUSTER, Bernhard: Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. In: *Anal. Bioanal. Chem.* 404 (2012), Nr. 4, S. 939–965
- [Bashirullah *et al.* 1999] BASHIRULLAH, Arash ; HALSELL, Susan R. ; COOPERSTOCK, Ramona L. ; KLOC, Malgorzata ; KARAIKAKIS, Angelo ; FISHER, William W. ; FU, Weili ; HAMILTON, Jill K. ; ETKIN, Laurence D. ; LIPSHITZ, Howard D.: Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in *Drosophila melanogaster*. In: *EMBO J.* 18 (1999), Nr. 9, S. 2610–2620
- [Bastock u. St Johnston 2008] BASTOCK, Rebecca ; ST JOHNSTON, Daniel: *Drosophila* oogenesis. In: *Curr. Biol.* 18 (2008), Nr. 23, S. R1082–R1087
- [Beckham *et al.* 2008] BECKHAM, Carla ; HILLIKER, Angela ; CZIKO, Anne-Marie ; NOUEIRY, Amine ; RAMASWAMI, Mani ; PARKER, Roy: The DEAD-box RNA helicase Ded1p affects and accumulates in *Saccharomyces cerevisiae* P-bodies. In: *Mol. Biol. Cell* 19 (2008), Nr. 3, S. 984–993
- [Behm-Ansmant *et al.* 2006] BEHM-ANSMANT, Isabelle ; REHWINKEL, Jan ; DOERKS, Tobias ; STARK, Alexander ; BORK, Peer ; IZAURRALDE, Elisa: mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. In: *Genes Dev.* 20 (2006), Nr. 14, S. 1885–1898

- [Benda *et al.* 2014] BENDA, Christian ; EBERT, Judith ; SCHELTEMA, Richard A. ; SCHILLER, Herbert B. ; BAUMGÄRTNER, Marc ; BONNEAU, Fabien ; MANN, Matthias ; CONTI, Elena: Structural Model of a CRISPR RNA-Silencing Complex Reveals the RNA-Target Cleavage Activity in Cmr4. In: *Mol. Cell* 56 (2014), Nr. 1, S. 43–54
- [Benoit *et al.* 2009] BENOIT, Beatrice ; HE, Chun H. ; ZHANG, Fan ; VOTRUBA, Sarah M. ; TADROS, Wael ; WESTWOOD, J. T. ; SMIBERT, Craig A. ; LIPSHITZ, Howard D. ; THEURKAUF, William E.: An essential role for the RNA-binding protein Smaug during the *Drosophila* maternal-to-zygotic transition. In: *Development* 136 (2009), Nr. 6, S. 923–932
- [Benoit *et al.* 2005] BENOIT, Béatrice ; MITOU, Géraldine ; CHARTIER, Aymeric ; TEMME, Claudia ; ZAESSINGER, Sophie ; WAHLE, Elmar ; BUSSEAU, Isabelle ; SIMONELIG, Martine: An essential cytoplasmic function for the nuclear poly (A) binding protein, PABP2, in poly (A) tail length control and early development in *Drosophila*. In: *Dev. Cell* 9 (2005), Nr. 4, S. 511–522
- [Berg u. von Hippel 1985] BERG, Otto G. ; HIPPEL, Peter H.: Diffusion-controlled macromolecular interactions. In: *Annu. Rev. Biophys. Biophys. Chem.* 14 (1985), Nr. 1, S. 131–158
- [Bergsten u. Gavis 1999] BERGSTEN, Sherri E. ; GAVIS, Elizabeth. R.: Role for mRNA localization in translational activation but not spatial restriction of nanos RNA. In: *Development* 126 (1999), Nr. 4, S. 659–669
- [Besse *et al.* 2009] BESSE, Florence ; QUINTO, Sonia L. ; MARCHAND, Virginie ; TRUCCO, Alvar ; EPHRUSSI, Anne: *Drosophila* PTB promotes formation of high-order RNP particles and represses oskar translation. In: *Genes Dev.* 23 (2009), Nr. 2, S. 195–207
- [Beveridge *et al.* 2014] BEVERIDGE, Rebecca ; COVILL, Sam ; PACHOLARZ, Kamila J. ; KALAPOTHAKIS, Jason M. D. ; MACPHEE, Cait E. ; BARRAN, Perdita E.: A Mass-Spectrometry-Based Framework To Define the Extent of Disorder in Proteins. In: *Anal. Chem.* 86 (2014), Nr. 22, S. 10979–10991
- [Bish *et al.* 2015] BISH, Rebecca ; CUEVAS-POLO, Nerea ; CHENG, Zhe ; HAMBARDZUMYAN, Dolores ; MUNSCHAUER, Mathias ; LANDTHALER, Markus ; VOGEL, Christine: Comprehensive Protein Interactome Analysis of a Key RNA Helicase: Detection of Novel Stress Granule Proteins. In: *Biomolecules* 5 (2015), Nr. 3, S. 1441–1466
- [Boag *et al.* 2005] BOAG, Peter R. ; NAKAMURA, Akira ; BLACKWELL, T. K.: A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in *C. elegans*. In: *Development* 132 (2005), Nr. 22, S. 4975–4986

- [Bono *et al.* 2006] BONO, Fulvia ; EBERT, Judith ; LORENTZEN, Esben ; CONTI, Elena: The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. In: *Cell* 126 (2006), Nr. 4, S. 713–725
- [Bowers *et al.* 2006] BOWERS, Heath A. ; MARONEY, Patricia A. ; FAIRMAN, Margaret E. ; KASTNER, Berthold ; LÜHRMANN, Reinhard ; NILSEN, Timothy W. ; JANKOWSKY, Eckhard: Discriminatory RNP remodeling by the DEAD-box protein DED1. In: *RNA* 12 (2006), Nr. 5, S. 903–912
- [Bownes 1975] BOWNES, Mary: A photographic study of development in the living embryo of *Drosophila melanogaster*. In: *Development* 33 (1975), Nr. 3, S. 789–801
- [Braun *et al.* 2011] BRAUN, Bernhard ; PFIRRMANN, Thorsten ; MENSSEN, Ruth ; HOFMANN, Kay ; SCHEEL, Hartmut ; WOLF, Dieter H.: Gid9, a second RING finger protein contributes to the ubiquitin ligase activity of the Gid complex required for catabolite degradation. In: *FEBS Lett.* 585 (2011), Nr. 24, S. 3856–3861
- [Buschmann *et al.* 2013] BUSCHMANN, Juliane ; MORITZ, Bodo ; JESKE, Mandy ; LILIE, Hauke ; SCHIERHORN, Angelika ; WAHLE, Elmar: Identification of *Drosophila* and Human 7-Methyl GMP-specific Nucleotidases. In: *J. Biol. Chem.* 288 (2013), Nr. 4, S. 2441–2451
- [Campos-Ortega u. Hartenstein 1997] CAMPOS-ORTEGA, José A. ; HARTENSTEIN, Volker: *The embryonic development of Drosophila melanogaster*. Springer-Verlag Berlin Heidelberg GmbH, 1997
- [Carroll *et al.* 2011] CARROLL, Johanna S. ; MUNCHEL, Sarah E. ; WEIS, Karsten: The DExD/H box ATPase Dhh1 functions in translational repression, mRNA decay, and processing body dynamics. In: *J. Cell Biol.* 194 (2011), Nr. 4, S. 527–537
- [Chang *et al.* 1990] CHANG, Tien-Hsien ; ARENAS, Jaime ; ABELSON, John: Identification of five putative yeast RNA helicase genes. In: *PNAS* 87 (1990), Nr. 4, S. 1571–1575
- [Chartier *et al.* 2015] CHARTIER, Aymeric ; KLEIN, Pierre ; PIERSON, Stéphanie ; BARBEZIER, Nicolas ; GIDARO, Teresa ; CASAS, François ; CARBERRY, Steven ; DOWLING, Paul ; MAYNADIER, Laurie ; BELLEC, Maëlle ; OLOKO, Martine ; JARDEL, Claude ; MORITZ, Bodo ; DICKSON, George ; MOULY, Vincent ; OHLENDIECK, Kay ; BUTLER-BROWNE, Gillian ; TROLLET, Capucine ; SIMONELIG, Martine: Mitochondrial Dysfunction Reveals the Role of mRNA Poly(A) Tail Regulation in Oculopharyngeal Muscular Dystrophy Pathogenesis. In: *PLOS Genet.* 11 (2015), Nr. 3, S. e1005092
- [Chekulaeva *et al.* 2006] CHEKULAEVA, Marina ; HENTZE, Matthias W. ; EPHRUSSI, Anne: Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles. In: *Cell* 124 (2006), Nr. 3, S. 521–533

- [Chen *et al.* 2014a] CHEN, Linan ; DUMELIE, Jason G. ; LI, Xiao ; CHENG, Matthew H. ; YANG, Zhiyong ; LAVER, John D. ; SIDDIQUI, Najeeb U. ; WESTWOOD, J. T. ; MORRIS, Quaid ; LIPSHITZ, Howard D. ; SMIBERT, Craig A.: Global regulation of mRNA translation and stability in the early *Drosophila* embryo by the Smaug RNA-binding protein. In: *Genome Biol.* 15 (2014), Nr. 1, S. R4
- [Chen *et al.* 1999] CHEN, Xiaohui ; CHEN, Yong H. ; ANDERSON, Vernon E.: Protein Cross-Links: Universal Isolation and Characterization by Isotopic Derivatization and Electrospray Ionization Mass Spectrometry. In: *Anal. Biochem.* 273 (1999), Nr. 2, S. 192–203
- [Chen *et al.* 2014b] CHEN, Ying ; BOLAND, Andreas ; KUZUOĞLU-ÖZTÜRK, Duygu ; BAWANKAR, Praveen ; LOH, Belinda ; CHANG, Chung-Te ; WEICHENRIEDER, Oliver ; IZAURRALDE, Elisa: A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. In: *Mol. Cell* 54 (2014), Nr. 5, S. 737–750
- [Chen *et al.* 2016] CHEN, Zhuo A. ; FISCHER, Lutz ; COX, Jürgen ; RAPPSILBER, Juri: Quantitative Cross-linking/Mass Spectrometry Using Isotope-labeled Cross-linkers and MaxQuant. In: *Mol. Cell. Proteomics* 15 (2016), Nr. 8, S. 2769–2778
- [Chen *et al.* 2010] CHEN, Zhuo A. ; JAWHARI, Anass ; FISCHER, Lutz ; BUCHEN, Claudia ; TAHIR, Salman ; KAMENSKI, Tomislav ; RASMUSSEN, Morten ; LARIVIERE, Laurent ; BUKOWSKI-WILLS, Jimi-Carlo ; NILGES, Michael ; CRAMER, Patrick ; RAPPSILBER, Juri: Architecture of the RNA polymerase II–TFIIF complex revealed by cross-linking and mass spectrometry. In: *EMBO J.* 29 (2010), Nr. 4, S. 717–726
- [Choi *et al.* 2010] CHOI, Seonhwa ; JEONG, Jaeho ; NA, Seungjin ; LEE, Hyo S. ; KIM, Hwa-Young ; LEE, Kong-Joo ; PAEK, Eunok: New algorithm for the identification of intact disulfide linkages based on fragmentation characteristics in tandem mass spectra. In: *J. Proteome Res.* 9 (2010), Nr. 1, S. 626–635
- [Chu u. Rana 2006] CHU, Chia-ying ; RANA, Tariq M.: Translation Repression in Human Cells by MicroRNA-Induced Gene Silencing Requires RCK/p54. In: *PLOS Biol.* 4 (2006), Nr. 7, S. e210
- [Chu *et al.* 2010] CHU, Feixia ; BAKER, Peter R. ; BURLINGAME, Alma L. ; CHALKLEY, Robert J.: Finding chimeras: a bioinformatics strategy for identification of cross-linked peptides. In: *Mol. Cell. Proteomics* 9 (2010), Nr. 1, S. 25–31
- [Chu *et al.* 2004] CHU, Feixia ; SHAN, Shu-ou ; MOUSTAKAS, Demetri T. ; ALBER, Frank ; EGEA, Pascal F. ; STROUD, Robert M. ; WALTER, Peter ; BURLINGAME, Alma L.: Unraveling the interface of signal recognition particle and its receptor by using chemical cross-linking and tandem mass spectrometry. In: *PNAS* 101 (2004), Nr. 47, S. 16454–16459

- [Chuang *et al.* 1997] CHUANG, Ray-Yuan ; WEAVER, Paul L. ; LIU, Zheng ; CHANG, Tien-Hsien: Requirement of the DEAD-Box Protein Ded1p for Messenger RNA Translation. In: *Science* 275 (1997), Nr. 5305, S. 1468–1471
- [Clark *et al.* 2000] CLARK, Ira E. ; WYCKOFF, David ; GAVIS, Elizabeth R.: Synthesis of the posterior determinant Nanos is spatially restricted by a novel cotranslational regulatory mechanism. In: *Curr. Biol.* 10 (2000), Nr. 20, S. 1311–1314
- [Clauser *et al.* 1999] CLAUSER, Karl R. ; BAKER, Peter ; BURLINGAME, Alma L.: Role of accurate mass measurement (+/-10 ppm) in protein identification strategies employing MS or MS/MS and database searching. In: *Anal. Chem.* 71 (1999), Nr. 14, S. 2871–2882
- [Clavier *et al.* 2015] CLAVIER, Séverine ; BOLBACH, Gérard ; SACHON, Emmanuelle: Photocross-Linked Peptide-Protein Complexes Analysis: A Comparative Study of CID and ETD Fragmentation Modes. In: *J. Am. Soc. Mass Spectrom.* 26 (2015), Nr. 6, S. 1014–1026
- [Coller u. Parker 2005] COLLER, Jeff ; PARKER, Roy: General translational repression by activators of mRNA decapping. In: *Cell* 122 (2005), Nr. 6, S. 875–886
- [Coller *et al.* 2001] COLLER, Jeffery M. ; TUCKER, Morgan ; SHETH, Ujwal ; VALENCIA-SANCHEZ, Marco A. ; PARKER, Roy: The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. In: *RNA* 7 (2001), Nr. 12, S. 1717–1727
- [Combe *et al.* 2015] COMBE, Colin W. ; FISCHER, Lutz ; RAPPSILBER, Juri: xiNET: cross-link network maps with residue resolution. In: *Mol. Cell. Proteomics* 14 (2015), Nr. 4, S. 1137–1147
- [Cooke *et al.* 2010] COOKE, Amy ; PRIGGE, Andrew ; WICKENS, Marvin: Translational repression by deadenylases. In: *J. Biol. Chem.* 285 (2010), Nr. 37, S. 28506–28513
- [Cox u. Mann 2008] COX, Jürgen ; MANN, Matthias: MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. In: *Nat. Biotechnol.* 26 (2008), Nr. 12, S. 1367–1372
- [Craig *et al.* 1998] CRAIG, Andrew W. ; HAGHIGHAT, Ashkan ; YU, Annie T. K. ; SONENBERG, Nahum: Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. In: *Nature* 392 (1998), Nr. 6675, S. 520–523
- [Dahanukar *et al.* 1999] DAHANUKAR, Anupama ; WALKER, James A. ; WHARTON, Robin P.: Smaug, a novel RNA-binding protein that operates a translational switch in *Drosophila*. In: *Mol. Cell* 4 (1999), Nr. 2, S. 209–218

- [Dahanukar u. Wharton 1996] DAHANUKAR, Anupama ; WHARTON, Robin P.: The Nanos gradient in Drosophila embryos is generated by translational regulation. In: *Genes Dev.* 10 (1996), Nr. 20, S. 2610–2621
- [De La Cruz *et al.* 1997] DE LA CRUZ, Jesús ; IOST, Isabelle ; KRESSLER, Dieter ; LINDER, Patrick: The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent translation in *Saccharomyces cerevisiae*. In: *PNAS* 94 (1997), Nr. 10, S. 5201–5206
- [Dehling *et al.* 2016] DEHLING, Eva ; VOLKMANN, Gerrit ; MATERN, Julian C. J. ; DÖRNER, Wolfgang ; ALFERMANN, Jonas ; DIECKER, Julia ; MOOTZ, Henning D.: Mapping of the communication-mediating interface in nonribosomal peptide synthetases using a genetically encoded photocrosslinker supports an upside-down helix-hand motif. In: *J. Mol. Biol.* 428 (2016), Nr. 21, S. 4345–4360
- [DeLano 2006] DELANO, Warren L.: *The PyMOL Molecular Graphics System v0.99rc6* (<http://www.pymol.org/>). 2006
- [Deo *et al.* 1999] DEO, Rahul C. ; BONANNO, Jeffrey B. ; SONENBERG, Nahum ; BURLEY, Stephen K.: Recognition of polyadenylate RNA by the poly(A)-binding protein. In: *Cell* 98 (1999), Nr. 6, S. 835–845
- [Dole *et al.* 1968] DOLE, Malcolm ; MACK, Lawrence L. ; HINES, Robert L. ; MOBLEY, Ralph C. ; FERGUSON, Lowell D. ; ALICE, Martin B.: Molecular Beams of Macroions. In: *J. Chem. Phys.* 49 (1968), Nr. 5, S. 2240–2249
- [Drummond *et al.* 2011] DRUMMOND, Sheona P. ; HILDYARD, John ; FIRCUK, Helena ; REAMTONG, Onrapak ; LI, Ning ; KANNAMBATH, Shichina ; CLAYDON, Amy J. ; BEYNON, Robert J. ; EYERS, Claire E. ; MCCARTHY, John E. G.: Diauxic shift-dependent relocation of decapping activators Dhh1 and Pat1 to polysomal complexes. In: *Nucl. Acids Res.* 39 (2011), Nr. 17, S. 7764–7774
- [Du *et al.* 2011] DU, Xiuxia ; CHOWDHURY, Saiful M. ; MANES, Nathan P. ; WU, Si ; MAYER, M. U. ; ADKINS, Joshua N. ; ANDERSON, Gordon A. ; SMITH, Richard D.: Xlink-identifier: an automated data analysis platform for confident identifications of chemically cross-linked peptides using tandem mass spectrometry. In: *J. Proteome Res.* 10 (2011), Nr. 3, S. 923–931
- [Duncan *et al.* 2009] DUNCAN, Kent E. ; STREIN, Claudia ; HENTZE, Matthias W.: The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA. In: *Mol. Cell* 36 (2009), Nr. 4, S. 571–582
- [Dunn *et al.* 2013] DUNN, Joshua G. ; FOO, Catherine K. ; BELLETIER, Nicolette G. ; GAVIS, Elizabeth R. ; WEISSMAN, Jonathan S.: Ribosome profiling reveals pervasive and regulated stop codon readthrough in *Drosophila melanogaster*. In: *eLife* 2 (2013), S. e01179

- [Dutta *et al.* 2011] DUTTA, Arnob ; ZHENG, Suting ; JAIN, Deepti ; CAMERON, Craig E. ; REESE, Joseph C.: Intermolecular interactions within the abundant DEAD-box protein Dhh1 regulate its activity in vivo. In: *J. Biol. Chem.* 286 (2011), Nr. 31, S. 27454–27470
- [Elbaum-Garfinkle *et al.* 2015] ELBAUM-GARFINKLE, Shana ; KIM, Younghoon ; SZCZEPANIAK, Krzysztof ; CHEN, Carlos Chih-Hsiung ; ECKMANN, Christian R. ; MYONG, Sua ; BRANGWYNNE, Clifford P.: The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. In: *PNAS* 112 (2015), Nr. 23, S. 7189–7194
- [Eng *et al.* 1994] ENG, Jimmy K. ; MCCORMACK, Ashley L. ; YATES, John R.: An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. In: *J. Am. Soc. Mass Spectrom.* 5 (1994), Nr. 11, S. 976–989
- [Ephrussi *et al.* 1991] EPHRUSSI, Anne ; DICKINSON, Laura K. ; LEHMANN, Ruth: oskar organizes the germ plasm and directs localization of the posterior determinant nanos. In: *Cell* 66 (1991), Nr. 1, S. 37–50
- [Ephrussi u. Lehmann 1992] EPHRUSSI, Anne ; LEHMANN, Ruth: Induction of germ cell formation by oskar. In: *Nature* 358 (1992), Nr. 6385, S. 387–392
- [Ernoul-Lange *et al.* 2012] ERNOULT-LANGE, Michèle ; BACONNAIS, Sonia ; HARPER, Maryannick ; MINSHALL, Nicola ; SOUQUERE, Sylvie ; BOUDIER, Thomas ; BÉNARD, Marianne ; ANDREY, Philippe ; PIERRON, Gérard ; KRESS, Michel ; STANDART, Nancy ; CAM, Eric I. ; WEIL, Dominique: Multiple binding of repressed mRNAs by the P-body protein Rck/p54. In: *RNA* 18 (2012), Nr. 9, S. 1702–1715
- [Eulalio *et al.* 2007] EULALIO, Ana ; BEHM-ANSMANT, Isabelle ; SCHWEIZER, Daniel ; IZAURRALDE, Elisa: P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. In: *Mol. Cell. Biol.* 27 (2007), Nr. 11, S. 3970–3981
- [Fenn *et al.* 1989] FENN, John B. ; MANN, Matthias ; MENG, Chin K. ; WONG, Shek F. ; WHITEHOUSE, Craig M.: Electrospray ionization for mass-spectrometry of large biomolecules. In: *Science* 246 (1989), Nr. 4926, S. 64–71
- [Ferber *et al.* 2016] FERBER, Mathias ; KOSINSKI, Jan ; ORI, Alessandro ; RASHID, Umar J. ; MORENO-MORCILLO, María ; SIMON, Bernd ; BOUVIER, Guillaume ; BATISTA, Paulo R. ; MÜLLER, Christoph W. ; BECK, Martin ; OTHERS: Automated structure modeling of large protein assemblies using crosslinks as distance restraints. In: *Nat. Meth.* 13 (2016), Nr. 6, S. 515–520
- [Fischer u. Weis 2002] FISCHER, Nicole ; WEIS, Karsten: The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. In: *EMBO J.* 21 (2002), Nr. 11, S. 2788–2797

- [Fischer *et al.* 2015] FISCHER, Niels ; NEUMANN, Piotr ; KONEVEGA, Andrey L. ; BOCK, Lars V. ; FICNER, Ralf ; RODNINA, Marina V. ; HOLGER STARK: Structure of the E. coli ribosome-EF-Tu complex at $< 3 \text{ \AA}$ resolution by Cs-corrected cryo-EM. In: *Nature* 520 (2015), Nr. 7548, S. 567–570
- [Francis *et al.* 2013] FRANCIS, Ore ; HAN, Fujun ; ADAMS, Josephine C.: Molecular phylogeny of a RING E3 ubiquitin ligase, conserved in eukaryotic cells and dominated by homologous components, the muskelin/RanBPM/CTLH complex. In: *PLoS One* 8 (2013), Nr. 10, S. e75217
- [Fritzsche *et al.* 2012] FRITZSCHE, Romy ; IHLING, Christian H. ; GÖTZE, Michael ; SINZ, Andrea: Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. In: *Rap. Commun. Mass Spectrom.* 26 (2012), Nr. 6, S. 653–658
- [Gao *et al.* 2006] GAO, Qiuxia ; XUE, Song ; DONEANU, Catalin E. ; SHAFFER, Scott A. ; GOODLETT, David R. ; NELSON, Sidney D.: Pro-CrossLink. Software tool for protein cross-linking and mass spectrometry. In: *Anal. Chem.* 78 (2006), Nr. 7, S. 2145–2149
- [Gao *et al.* 2016] GAO, Zhaofeng ; PUTNAM, Andrea A. ; BOWERS, Heath A. ; GUENTHER, Ulf-Peter ; YE, Xuan ; KINDSFATHER, Audrey ; HILLIKER, Angela K. ; JANKOWSKY, Eckhard: Coupling between the DEAD-box RNA helicases Ded1p and eIF4A. In: *eLife* 5 (2016), S. e16408
- [Garcia *et al.* 2013] GARCIA, Hernan G. ; TIKHONOV, Mikhail ; LIN, Albert ; GREGOR, Thomas: Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. In: *Curr. Biol.* 23 (2013), Nr. 21, S. 2140–2145
- [García-García *et al.* 2015] GARCÍA-GARCÍA, Cuauhtémoc ; FRIEDA, Kirsten L. ; FEOKTISTOVA, Kateryna ; FRASER, Christopher S. ; BLOCK, Steven M.: Factor-dependent processivity in human eIF4A DEAD-box helicase. In: *Science* 348 (2015), Nr. 6242, S. 1486–1488
- [Garneau *et al.* 2007] GARNEAU, Nicole L. ; WILUSZ, Jeffrey ; WILUSZ, Carol J.: The highways and byways of mRNA decay. In: *Nat. Rev. Mol. Cell Biol.* 8 (2007), Nr. 2, S. 113–126
- [Gavis *et al.* 1996a] GAVIS, Elizabeth R. ; CURTIS, Daniel ; LEHMANN, Ruth: Identification of *cis*-Acting Sequences That Control *nanos* RNA Localization. In: *Dev. Biol.* 176 (1996), Nr. 1, S. 36–50
- [Gavis u. Lehmann 1992] GAVIS, Elizabeth R. ; LEHMANN, Ruth: Localization of *nanos* RNA controls embryonic polarity. In: *Cell* 71 (1992), Nr. 2, S. 301–313

- [Gavis u. Lehmann 1994] GAVIS, Elizabeth R. ; LEHMANN, Ruth: Translational regulation of nanos by RNA localization. In: *Nature* 369 (1994), Nr. 6478, S. 315–318
- [Gavis *et al.* 1996b] GAVIS, Elizabeth R. ; LUNSFORD, Lynn ; BERGSTEN, Sherri E. ; LEHMANN, Ruth: A conserved 90 nucleotide element mediates translational repression of nanos RNA. In: *Development* 122 (1996), Nr. 9, S. 2791–2800
- [Gehring *et al.* 2005] GEHRING, Niels H. ; KUNZ, Joachim B. ; NEU-YILIK, Gabriele ; BREIT, Stephen ; VIEGAS, Marcelo H. ; HENTZE, Matthias W. ; KULOZIK, Andreas E.: Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. In: *Mol. Cell* 20 (2005), Nr. 1, S. 65–75
- [Geissler *et al.* 2012] GEISLER, Rene ; GOLBIK, Ralph P. ; BEHRENS, Sven-Erik: The DEAD-box helicase DDX3 supports the assembly of functional 80S ribosomes. In: *Nucl. Acids Res.* 40 (2012), Nr. 11, S. 4998–5011
- [Gelbart u. Emmert 2013] GELBART, W.M. ; EMMERT, D.B.: FlyBase High Throughput Expression Pattern Data. (2013)
- [Götze *et al.* 2015] GÖTZE, Michael ; PETTELKAU, Jens ; FRITZSCHE, Romy ; IHLING, Christian H. ; SCHÄFER, Mathias ; SINZ, Andrea: Automated Assignment of MS/MS Cleavable Cross-Links in Protein 3D-Structure Analysis. In: *J. Am. Soc. Mass Spectrom.* 26 (2015), Nr. 1, S. 83–97
- [Götze *et al.* 2012] GÖTZE, Michael ; PETTELKAU, Jens ; SCHAKS, Sabine ; BOSSE, Konstanze ; IHLING, Christian H. ; KRAUTH, Fabian ; FRITZSCHE, Romy ; KÜHN, Uwe ; SINZ, Andrea: StavroX – a software for analyzing crosslinked products in protein interaction studies. In: *J. Am. Soc. Mass Spectrom.* 23 (2012), Nr. 1, S. 76–87
- [Götze u. Wahle 2014] GÖTZE, Michael ; WAHLE, Elmar: Smaug destroys a huge treasure. In: *Genome Biol.* 15 (2014), Nr. 1, S. 101
- [Gouw *et al.* 2009] GOUW, Joost W. ; PINKSE, Martijn W. ; VOS, Harmjan R. ; MOSHKIN, Yuri ; VERRIJZER, C. P. ; HECK, Albert J. ; KRIJGSVELD, Jeroen: In vivo stable isotope labeling of fruit flies reveals post-transcriptional regulation in the maternal-to-zygotic transition. In: *Mol. Cell. Proteomics* 8 (2009), Nr. 7, S. 1566–1578
- [Gray u. Hentze 1994] GRAY, Nicola K. ; HENTZE, Matthias W.: Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs. In: *EMBO J.* 13 (1994), Nr. 16, S. 3882–3891
- [Green *et al.* 2003] GREEN, Justin B. ; GARDNER, Cary D. ; WHARTON, Robin P. ; AGGARWAL, Aneel K.: RNA recognition via the SAM domain of Smaug. In: *Mol. Cell* 11 (2003), Nr. 6, S. 1537–1548

- [Green 1963] GREEN, N. M.: Avidin. 1. The use of [14C] biotin for kinetic studies and for assay. In: *Biochem. J.* 89 (1963), Nr. 3, S. 585
- [Green 1966] GREEN, N. M.: Thermodynamics of the binding of biotin and some analogues by avidin. In: *Biochem. J.* 101 (1966), Nr. 3, S. 774–780
- [Grimm *et al.* 2015] GRIMM, Maximilian ; ZIMNIAK, Tomasz ; KAHRAMAN, Abdullah ; HERZOG, Franz: xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. In: *Nucl. Acids Res.* 43 (2015), Nr. W1, S. W362–369
- [Ha u. Kim 2014] HA, Minju ; KIM, V. N.: Regulation of microRNA biogenesis. In: *Nat. Rev. Mol. Cell Biol.* 15 (2014), Nr. 8, S. 509–524
- [Hage *et al.* 2016] HAGE, Christoph ; IHLING, Christian H. ; GÖTZE, Michael ; SCHÄFER, Mathias ; SINZ, Andrea: Dissociation Behavior of a TEMPO-Active Ester Cross-Linker for Peptide Structure Analysis by Free Radical Initiated Peptide Sequencing (FRIPS) in Negative ESI-MS. In: *J. Am. Soc. Mass Spectrom.* (2016), S. 1–13
- [Harnisch *et al.* 2016] HARNISCH, Christiane ; CUZIC-FELTENS, Simona ; DOHM, Juliane C. ; GÖTZE, Michael ; HIMMELBAUER, Heinz ; WAHLE, Elmar: Oligoadenylation of 3' decay intermediates promotes cytoplasmic mRNA degradation in *Drosophila* cells. In: *RNA* 22 (2016), Nr. 3, S. 428–442
- [Harrison u. Eisen 2015] HARRISON, Melissa M. ; EISEN, Michael B.: Transcriptional Activation of the Zygotic Genome in *Drosophila*. In: LIPSHITZ, Howard D. (Hrsg.): *Current Topics in Developmental Biology* Bd. 113. Academic Press, 2015, S. 85–112
- [Hart-Smith 2014] HART-SMITH, Gene: A review of electron-capture and electron-transfer dissociation tandem mass spectrometry in polymer chemistry. In: *Anal. Chim. Acta* 808 (2014), S. 44–55
- [Hauser *et al.* 2015] HAUSER, Thomas ; BHAT, Javaid Y. ; MILIČIĆ, Goran ; WENDLER, Petra ; HARTL, F. U. ; BRACHER, Andreas ; HAYER-HARTL, Manajit: Structure and mechanism of the Rubisco-assembly chaperone Raf1. In: *Nat. Struct. Mol. Biol.* (2015)
- [Hillebrand *et al.* 2010] HILLEBRAND, Jens ; PAN, Kanyu ; KOKARAM, Anil ; BARBEE, Scott ; PARKER, Roy ; RAMASWAMI, Mani: The Me31B DEAD-box helicase localizes to postsynaptic foci and regulates expression of a CaMKII reporter mRNA in dendrites of *Drosophila* olfactory projection neurons. In: *Drosophila* 4 (2010), S. 121
- [Hilliker *et al.* 2011] HILLIKER, Angela ; GAO, Zhaofeng ; JANKOWSKY, Eckhard ; PARKER, Roy: The DEAD-box protein Ded1 modulates translation by the formation and resolution of an eIF4F-mRNA complex. In: *Mol. Cell* 43 (2011), Nr. 6, S. 962–972

- [Holding *et al.* 2013] HOLDING, Andrew N. ; LAMERS, Meindert H. ; STEPHENS, Elaine ; SKEHEL, J. M.: Hekate: software suite for the mass spectrometric analysis and three-dimensional visualization of cross-linked protein samples. In: *J. Proteome Res.* 12 (2013), Nr. 12, S. 5923–5933
- [Hoopmann *et al.* 2016] HOOPMANN, Michael R. ; MENDOZA, Luis ; DEUTSCH, Eric W. ; SHTEYNBERG, David ; MORITZ, Robert L.: An Open Data Format for Visualization and Analysis of Cross-Linked Mass Spectrometry Results. In: *J. Am. Soc. Mass Spectrom.* (2016), S. 1–7
- [Hoopmann *et al.* 2015] HOOPMANN, Michael R. ; ZELTER, Alex ; JOHNSON, Richard S. ; RIFFLE, Michael ; MACCOSS, Michael J. ; DAVIS, Trisha N. ; MORITZ, Robert L.: Kojak: efficient analysis of chemically cross-linked protein complexes. In: *J. Proteome Res.* 14 (2015), Nr. 5, S. 2190–2198
- [Hubstenberger *et al.* 2013] HUBSTENBERGER, Arnaud ; NOBLE, Scott L. ; CAMERON, Cris-tiana ; EVANS, Thomas C.: Translation repressors, an RNA helicase, and developmental cues control RNP phase transitions during early development. In: *Dev. Cell* 27 (2013), Nr. 2, S. 161–173
- [Igreja u. Izaurrealde 2011] IGREJA, Catia ; IZAURREALDE, Elisa: CUP promotes deadenylation and inhibits decapping of mRNA targets. In: *Genes Dev.* 25 (2011), Nr. 18, S. 1955–1967
- [Ihry *et al.* 2012] IHRY, Robert J. ; SAPIRO, Anne L. ; BASHIRULLAH, Arash: Translational control by the DEAD Box RNA helicase belle regulates ecdysone-triggered transcriptional cascades. In: *PLOS Genet.* 8 (2012), Nr. 11, S. e1003085
- [Iost *et al.* 1999] IOST, Isabelle ; DREYFUS, Marc ; LINDER, Patrick: Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase. In: *J. Biol. Chem.* 274 (1999), Nr. 25, S. 17677–17683
- [Iribarne u. Thomson 1976] IRIBARNE, J. V. ; THOMSON, B. A.: On the evaporation of small ions from charged droplets. In: *J. Chem. Phys.* 64 (1976), Nr. 6, S. 2287–2294
- [Jaiswal *et al.* 2014] JAISWAL, Mihir ; CRABTREE, Nathaniel ; BAUER, Michael A. ; HALL, Roger ; RANEY, Kevin D. ; ZYBailov, Boris L.: XLPM: efficient algorithm for the analysis of protein-protein contacts using chemical cross-linking mass spectrometry. In: *BMC Bioinformatics* 15 Suppl 11 (2014), S. S16
- [Jakob *et al.* 2016] JAKOB, Leonhard ; TREIBER, Thomas ; TREIBER, Nora ; GUST, Alexander ; KRAMM, Kevin ; HANSEN, Kerrin ; STOTZ, Mathias ; WANKERL, Ludwig ; HERZOG, Franz ; HANNUS, Stefan ; GROHMANN, DINA ; MEISTER, GUNTER: Structural and functional insights into the fly microRNA biogenesis factor Loquacious. In: *RNA* 22 (2016), Nr. 3, S. 383–396

- [Jankowsky 2011] JANKOWSKY, Eckhard: RNA helicases at work: binding and rearranging. In: *Trends Biochem. Sci.* 36 (2011), Nr. 1, S. 19–29
- [Jeske *et al.* 2015] JESKE, Mandy ; BORDI, Matteo ; GLATT, Sebastian ; MÜLLER, Sandra ; RYBIN, Vladimir ; MÜLLER, Christoph W. ; EPHRUSSI, Anne: The crystal structure of the *Drosophila* germline inducer Oskar identifies two domains with distinct Vasa helicase-and RNA-binding activities. In: *Cell Rep.* 12 (2015), Nr. 4, S. 587–598
- [Jeske *et al.* 2006] JESKE, Mandy ; MEYER, Sylke ; TEMME, Claudia ; FREUDENREICH, Dorian ; WAHLE, Elmar: Rapid ATP-dependent deadenylation of nanos mRNA in a cell-free system from *Drosophila* embryos. In: *J. Biol. Chem.* 281 (2006), Nr. 35, S. 25124–25133
- [Jeske *et al.* 2011] JESKE, Mandy ; MORITZ, Bodo ; ANDERS, Alexander ; WAHLE, Elmar: Smaug assembles an ATP-dependent stable complex repressing nanos mRNA translation at multiple levels. In: *EMBO J.* 30 (2011), Nr. 1, S. 90–103
- [Johnstone *et al.* 2005] JOHNSTONE, Oona ; DEURING, Renate ; BOCK, Ronald ; LINDER, Patrick ; FULLER, Margaret T. ; LASKO, Paul: Belle is a *Drosophila* DEAD-box protein required for viability and in the germ line. In: *Dev. Biol.* 277 (2005), Nr. 1, S. 92–101
- [Jonas u. Izaurralde 2013] JONAS, Stefanie ; IZAURRALDE, Elisa: The role of disordered protein regions in the assembly of decapping complexes and RNP granules. In: *Genes Dev.* 27 (2013), Nr. 24, S. 2628–2641
- [Jones *et al.* 2012] JONES, Andrew R. ; EISENACHER, Martin ; MAYER, Gerhard ; KOHLBACHER, Oliver ; SIEPEN, Jennifer ; HUBBARD, Simon J. ; SELLEY, Julian N. ; SEARLE, Brian C. ; SHOFSTAHL, James ; SEYMOUR, Sean L. ; OTHERS: The mzIdentML data standard for mass spectrometry-based proteomics results. In: *Mol. Cell. Proteomics* 11 (2012), Nr. 7, S. M111–014381
- [Jurneczko u. Barran 2010] JURNECZKO, Ewa ; BARRAN, Perdita E.: How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. 136 (2010), Nr. 1, S. 20–28
- [Kaake *et al.* 2014] KAAKE, Robyn M. ; WANG, Xiaorong ; BURKE, Anthony ; YU, Clinton ; KANDUR, Wynne ; YANG, Yingying ; NOVTISKY, Eric J. ; SECOND, Tonya ; DUAN, Jicheng ; KAO, Athit ; GUAN, Shenheng ; VELLUCCI, Danielle ; RYCHNOVSKY, Scott D. ; HUANG, Lan: A New in Vivo Cross-linking Mass Spectrometry Platform to Define Protein–Protein Interactions in Living Cells. In: *Mol. Cell. Proteomics* 13 (2014), Nr. 12, S. 3533–3543
- [Kadyrova *et al.* 2007] KADYROVA, Lyudmila Y. ; HABARA, Yasuaki ; LEE, Tammy H. ; WHARTON, Robin P.: Translational control of maternal Cyclin B mRNA by Nanos in the *Drosophila* germline. In: *Development* 134 (2007), Nr. 8, S. 1519–1527

- [Kalifa *et al.* 2006] KALIFA, Yossi ; HUANG, Tao ; ROSEN, Lynne N. ; CHATTERJEE, Seema ; GAVIS, Elizabeth R.: Glorund, a Drosophila hnRNP F/H homolog, is an ovarian repressor of nanos translation. In: *Dev. Cell* 10 (2006), Nr. 3, S. 291–301
- [Kalisman *et al.* 2012] KALISMAN, Nir ; ADAMS, Christopher M. ; LEVITT, Michael: Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. In: *PNAS* 109 (2012), Nr. 8, S. 2884–2889
- [Kamenska *et al.* 2014] KAMENSKA, Anastasiia ; LU, Wei-Ting ; KUBACKA, Dorota ; BROOMHEAD, Helen ; MINSHALL, Nicola ; BUSHELL, Martin ; STANDART, Nancy: Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. In: *Nucl. Acids Res.* 42 (2014), Nr. 5, S. 3298–3313
- [Kamenska *et al.* 2016] KAMENSKA, Anastasiia ; SIMPSON, Clare ; VINDRY, Caroline ; BROOMHEAD, Helen ; BÉNARD, Marianne ; ERNOULT-LANGE, Michèle ; LEE, Benjamin P. ; HARRIES, Lorna W. ; WEIL, Dominique ; STANDART, Nancy: The DDX6–4E-T interaction mediates translational repression and P-body assembly. In: *Nucl. Acids Res.* 44 (2016), Nr. 13, S. 6318–6334
- [Kao *et al.* 2010] KAO, Athit ; CHIU, Chi-li ; VELLUCCI, Danielle ; YANG, Yingying ; PATEL, Vishal R. ; GUAN, Shenheng ; RANDALL, Arlo ; BALDI, Pierre ; RYCHNOVSKY, Scott D. ; HUANG, Lan: Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. In: *Mol. Cell. Proteomics* 10 (2010), Nr. 1, S. M110.002212
- [Kastner *et al.* 2008] KASTNER, Berthold ; FISCHER, Niels ; GOLAS, Monika M. ; SANDER, Bjoern ; DUBE, Prakash ; BOEHRINGER, Daniel ; HARTMUTH, Klaus ; DECKERT, Jochen ; HAUER, Florian ; WOLF, Elmar ; UCHTENHAGEN, Hannes ; URLAUB, Henning ; HERZOG, Franz ; PETERS, Jan M. ; POERSCHKE, Dietmar ; LÜHRMANN, Reinhard ; STARK, Holger: GraFix: sample preparation for single-particle electron cryomicroscopy. In: *Nat. Meth.* 5 (2008), Nr. 1, S. 53–55
- [Katta u. Chait 1993] KATTA, Viswanatham ; CHAIT, Brain T.: Hydrogen/deuterium exchange electrospray ionization mass spectrometry: a method for probing protein conformational changes in solution. In: *J. Am. Chem. Soc.* 115 (1993), Nr. 14, S. 6317–6321
- [Keyes u. Spradling 1997] KEYES, Linda N. ; SPRADLING, Allan C.: The Drosophila gene fs (2) cup interacts with otu to define a cytoplasmic pathway required for the structure and function of germ-line chromosomes. In: *Development* 124 (1997), Nr. 7, S. 1419–1431. – PMID 9118812
- [Kim *et al.* 2015] KIM, Min-Sik ; ZHONG, Jun ; PANDEY, Akhilesh: Common errors in mass spectrometry-based analysis of post-translational modifications. In: *Proteomics* 16 (2015), Nr. 5, S. 700–714

- [Kim-Ha *et al.* 1995] KIM-HA, Jeongsil ; KERR, Karen ; MACDONALD, Paul M.: Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. In: *Cell* 81 (1995), Nr. 3, S. 403–412
- [Kinkelin *et al.* 2012] KINKELIN, Kerstin ; VEITH, Katharina ; GRÜNWALD, Marlene ; BONO, Fulvia: Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. In: *RNA* 18 (2012), Nr. 9, S. 1624–1634
- [Kobayashi *et al.* 2007] KOBAYASHI, Nobuaki ; YANG, Jun ; UEDA, Atsuhisa ; SUZUKI, Takeyuki ; TOMARU, Kouji ; TAKENO, Mitsuhiro ; OKUDA, Kenji ; ISHIGATSUBO, Yoshiaki: RanBPM, Muskelein, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8 α and ARMC8 β are components of the CTLH complex. In: *Gene* 396 (2007), Nr. 2, S. 236–247
- [Kondrat *et al.* 2015] KONDRAT, Frances D. L. ; STRUWE, Weston B. ; BENESCH, Justin L. P.: Native Mass Spectrometry: Towards High-Throughput Structural Proteomics. In: OWENS, Raymond J. (Hrsg.): *Structural Proteomics* Bd. 1261. Springer New York, 2015, S. 349–371
- [Konovalova *et al.* 2014] KONOVALOVA, Anna ; PERLMAN, David H. ; COWLES, Charles E. ; SILHAVY, Thomas J.: Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β -barrel proteins. In: *PNAS* 111 (2014), Nr. 41, S. E4350–4358
- [Kronja *et al.* 2014] KRONJA, Iva ; YUAN, Bingbing ; EICHHORN, Stephen W. ; DZEYK, Kristina ; KRIJGSVELD, Jeroen ; BARTEL, David P. ; ORR-WEAVER, Terry L.: Widespread Changes in the Posttranscriptional Landscape at the Drosophila Oocyte-to-Embryo Transition. In: *Cell Reports* 7 (2014), Nr. 5, S. 1495–1508
- [Krüger u. Rehmsmeier 2006] KRÜGER, Jan ; REHMSMEIER, Marc: RNAhybrid: microRNA target prediction easy, fast and flexible. In: *Nucl. Acids Res.* 34 (2006), Nr. suppl 2, S. W451–W454
- [Kugler u. Lasko 2009] KUGLER, Jan-Michael ; LASKO, Paul: Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during Drosophila oogenesis. In: *Fly* 3 (2009), Nr. 1, S. 15–28
- [Kuzuoğlu-Öztürk *et al.* 2016] KUZUOĞLU-ÖZTÜRK, Duygu ; BHANDARI, Dipankar ; HUNTINGER, Eric ; FAUSER, Maria ; HELMS, Sigrun ; IZAURRALDE, Elisa: miRISC and the CCR4–NOT complex silence mRNA targets independently of 43S ribosomal scanning. In: *EMBO J.* 35 (2016), Nr. 11, S. 1186–1203

- [Ladomery *et al.* 1997] LADOMERY, Michael ; WADE, Eleanor ; SOMMERVILLE, John: Xp54, the *Xenopus* homologue of human RNA helicase p54, is an integral component of stored mRNP particles in oocytes. In: *Nucl. Acids Res.* 25 (1997), Nr. 5, S. 965–973
- [Langmead *et al.* 2009] LANGMEAD, Ben ; TRAPNELL, Cole ; POP, Mihai ; SALZBERG, Steven L.: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. In: *Genome Biol.* 10 (2009), Nr. 3, S. R25
- [Laver *et al.* 2015] LAVER, John D. ; LI, Xiao ; RAY, Debashish ; COOK, Kate B. ; HAHN, Noah A. ; NABEEL-SHAH, Syed ; KEKIS, Mariana ; LUO, Hua ; MARSOLAIS, Alexander J. ; FUNG, Karen Y. ; HUGHES, Timothy R. ; WESTWOOD, J. T. ; SIDHU, Sachdev S. ; MORRIS, Quaid ; LIPSHITZ, Howard D. ; SMIBERT, Craig A.: Brain tumor is a sequence-specific RNA-binding protein that directs maternal mRNA clearance during the *Drosophila* maternal-to-zygotic transition. In: *Genome Biol.* 16 (2015), Nr. 1, S. 94
- [Lee *et al.* 2008] LEE, Chung-Sheng ; DIAS, Anusha P. ; JEDRYCHOWSKI, Mark ; PATEL, Arvind H. ; HSU, Jeanne L. ; REED, Robin: Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. In: *Nucl. Acids Res.* 36 (2008), Nr. 14, S. 4708–4718
- [Lee u. Goldberg 1998] LEE, Do H. ; GOLDBERG, Alfred L.: Proteasome inhibitors: valuable new tools for cell biologists. In: *Trends Cell Biol.* 8 (1998), Nr. 10, S. 397–403
- [Lehmann u. Nüsslein-Volhard 1991] LEHMANN, Ruth ; NÜSSLEIN-VOLHARD, Christiane: The maternal gene nanos has a central role in posterior pattern formation of the *Drosophila* embryo. In: *Development* 112 (1991), Nr. 3, S. 679–691
- [Lieb *et al.* 1998] LIEB, Bernhard ; CARL, Marina ; HOCK, Robert ; GEBAUER, Dagmar ; SCHEER, Ulrich: Identification of a Novel mRNA-Associated Protein in Oocytes of *Pleurodeles waltlandi* and *Xenopus laevis*. In: *Exp. Cell Res.* 245 (1998), Nr. 2, S. 272–281
- [Lima *et al.* 2015] LIMA, Diogo B. ; LIMA, Tatiani B. ; BALBUENA, Tiago S. ; NEVES-FERREIRA, Ana Gisele C. ; BARBOSA, Valmir C. ; GOZZO, Fábio C. ; CARVALHO, Paulo C.: SIM-XL: A powerful and user-friendly tool for peptide cross-linking analysis. In: *J. Proteomics* 129 (2015), S. 51–55
- [Lin *et al.* 2006] LIN, Ming-Der ; FAN, Shih-Jung ; HSU, Wei-Shan ; CHOU, Tze-Bin: *Drosophila* Decapping Protein 1, dDcp1, Is a Component of the oskar mRNP Complex and Directs Its Posterior Localization in the Oocyte. In: *Dev. Cell* 10 (2006), Nr. 5, S. 601–613
- [Linder u. Jankowsky 2011] LINDER, Patrick ; JANKOWSKY, Eckhard: From unwinding to clamping - the DEAD box RNA helicase family. In: *Nat. Rev. Mol. Cell Biol.* 12 (2011), Nr. 8, S. 505–516

- [Lipfert u. Doniach 2007] LIPFERT, Jan ; DONIACH, Sebastian: Small-angle X-ray scattering from RNA, proteins, and protein complexes. In: *Annu. Rev. Biophys. Biomol. Struct.* 36 (2007), S. 307–327
- [Little *et al.* 2015] LITTLE, Shawn C. ; SINSIMER, Kristina S. ; LEE, Jack J. ; WIESCHAUS, Eric F. ; GAVIS, Elizabeth R.: Independent and coordinate trafficking of single *Drosophila* germ plasm mRNAs. In: *Nat. Cell Biol.* 17 (2015), Nr. 5, S. 558–568
- [Liu *et al.* 2015] LIU, Fan ; RIJKERS, Dirk T. S. ; POST, Harm ; HECK, Albert J. R.: Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. In: *Nat. Meth.* 12 (2015), Nr. 12, S. 1179–1184
- [Liu *et al.* 2008] LIU, Fei ; PUTNAM, Andrea ; JANKOWSKY, Eckhard: ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. In: *PNAS* 105 (2008), Nr. 51, S. 20209–20214
- [Liu *et al.* 2014] LIU, Fei ; PUTNAM, Andrea A. ; JANKOWSKY, Eckhard: DEAD-box helicases form nucleotide-dependent, long-lived complexes with RNA. In: *Biochemistry* 53 (2014), Nr. 2, S. 423–433
- [Loedige *et al.* 2015] LOEDIGE, Inga ; JAKOB, Leonhard ; TREIBER, Thomas ; RAY, Debashish ; STOTZ, Mathias ; TREIBER, Nora ; HENNIG, Janosch ; COOK, Kate B. ; MORRIS, Quaid ; HUGHES, Timothy R. ; ENGELMANN, Julia C. ; KRAHN, Michael P. ; MEISTER, Gunter: The crystal structure of the NHL domain in complex with RNA reveals the molecular basis of *Drosophila* brain-tumor-mediated gene regulation. In: *Cell Rep.* 13 (2015), Nr. 6, S. 1206–1220
- [Luchini *et al.* 2014] LUCHINI, Alessandra ; ESPINA, Virginia ; LIOTTA, Lance A.: Protein painting reveals solvent-excluded drug targets hidden within native protein-protein interfaces. In: *Nat. Commun.* 5 (2014), S. 4413
- [Macdonald *et al.* 2016] MACDONALD, Paul M. ; KANKE, Matt ; KENNY, Andrew: Community effects in regulation of translation. In: *eLife* 5 (2016), S. e10965
- [Mädler *et al.* 2009] MÄDLER, Stefanie ; BICH, Claudia ; TOUBOUL, David ; ZENOBI, Renato: Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. In: *J. Mass Spectrom.* 44 (2009), Nr. 5, S. 694–706
- [Maiolica *et al.* 2007] MAIOLICA, Alessio ; CITTARO, Davide ; BORSOTTI, Dario ; SENNELS, Lau ; CIFERRI, Claudio ; TARRICONE, Cataldo ; MUSACCHIO, Andrea ; RAPPSILBER, Juri: Structural Analysis of Multiprotein Complexes by Cross-linking, Mass Spectrometry, and Database Searching. In: *Mol. Cell. Proteomics* 6 (2007), Nr. 12, S. 2200–2211

- [Marcotrigiano *et al.* 1999] MARCOTRIGIANO, Joseph ; GINGRAS, Anne-Claude ; SONENBERG, Nahum ; BURLEY, Stephen K.: Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. In: *Mol. Cell* 3 (1999), Nr. 6, S. 707–716
- [Marnef *et al.* 2009] MARNEF, Aline ; SOMMERVILLE, John ; LADOMERY, Michael R.: RAP55: Insights into an evolutionarily conserved protein family. In: *Int. J. Biochem. Cell Biol.* 41 (2009), Nr. 5, S. 977–981
- [Martens *et al.* 2011] MARTENS, Lennart ; CHAMBERS, Matthew ; STURM, Marc ; KESSNER, Darren ; LEVANDER, Fredrik ; SHOFSTAHL, Jim ; TANG, Wilfred H. ; RÖMPP, Andreas ; NEUMANN, Steffen ; PIZARRO, Angel D. ; OTHERS: mzML-a Community Standard for Mass Spectrometry Data. In: *Mol. Cell. Proteomics* 10 (2011), Nr. 1, S. R110.000133
- [Martin 2011] MARTIN, Marcel: Cutadapt removes adapter sequences from high-throughput sequencing reads. In: *EMBnet.journal* 17 (2011), Nr. 1, S. pp. 10–12
- [Mathys *et al.* 2014] MATHYS, Hansruedi ; BASQUIN, Jérôme ; OZGUR, Sevim ; CZARNOCKI-CIECIURA, Mariusz ; BONNEAU, Fabien ; AARTSE, Aafke ; DZIEMBOWSKI, Andrzej ; NOWOTNY, Marcin ; CONTI, Elena ; FILIPOWICZ, Witold: Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression. In: *Mol. Cell* 54 (2014), Nr. 5, S. 751–765
- [Mayne u. Patterton 2014] MAYNE, Shannon L. N. ; PATTERTON, Hugh-G.: AnchorMS: a bioinformatics tool to derive structural information from the mass spectra of cross-linked protein complexes. In: *Bioinformatics* 30 (2014), Nr. 1, S. 125–126
- [McIlwain *et al.* 2010] MCILWAIN, Sean ; DRAGHICESCU, Paul ; SINGH, Pragya ; GOODLETT, David R. ; NOBLE, William S.: Detecting cross-linked peptides by searching against a database of cross-linked peptide pairs. In: *J. Proteome Res.* 9 (2010), Nr. 5, S. 2488–2495
- [McLafferty *et al.* 1973] MCLAFFERTY, F. W. ; BENTE, P. F. ; KORNFELD, Richard ; TSAI, Shih-Chuan ; HOWE, Ian: Collisional activation spectra of organic ions. In: *J. Am. Chem. Soc.* 95 (1973), Nr. 7, S. 2120–2129
- [Meister u. Tuschl 2004] MEISTER, Gunter ; TUSCHL, Thomas: Mechanisms of gene silencing by double-stranded RNA. In: *Nature* 431 (2004), Nr. 7006, S. 343–349
- [Mendoza u. Vachet 2009] MENDOZA, Vanessa L. ; VACHET, Richard W.: Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. In: *Mass Spectrom. Rev.* 28 (2009), Nr. 5, S. 785–815
- [Menssen *et al.* 2012] MENSSEN, Ruth ; SCHWEIGGERT, Jörg ; SCHREINER, Jens ; KUŠEVIĆ, Denis ; REUTHER, Julia ; BRAUN, Bernhard ; WOLF, Dieter H.: Exploring the topology of the Gid complex, the E3 ubiquitin ligase involved in catabolite-induced degradation of gluconeogenic enzymes. In: *J. Biol. Chem.* 287 (2012), Nr. 30, S. 25602–25614

- [Michalski *et al.* 2012] MICHALSKI, Annette ; NEUHAUSER, Nadin ; COX, Jürgen ; MANN, Matthias: A systematic investigation into the nature of tryptic HCD spectra. In: *J. Proteome Res.* 11 (2012), Nr. 11, S. 5479–5491
- [Minshall u. Standart 2004] MINSHALL, Nicola ; STANDART, Nancy: The active form of Xp54 RNA helicase in translational repression is an RNA-mediated oligomer. In: *Nucl. Acids Res.* 32 (2004), Nr. 4, S. 1325–1334
- [Minshall *et al.* 2001] MINSHALL, Nicola ; THOM, George ; STANDART, Nancy: A conserved role of a DEAD box helicase in mRNA masking. In: *RNA* 7 (2001), Nr. 12, S. 1728–1742
- [Mueller-Planitz 2015] MUELLER-PLANITZ, Felix: Crossfinder-assisted mapping of protein crosslinks formed by site-specifically incorporated crosslinkers. In: *Bioinformatics* 31 (2015), Nr. 12, S. 2043–2045
- [Müller *et al.* 2010] MÜLLER, Mathias Q. ; DREIOCKER, Frank ; IHLING, Christian H. ; SCHÄFER, Mathias ; SINZ, Aandrea: Cleavable cross-linker for protein structure analysis: reliable identification of cross-linking products by tandem MS. In: *Anal. Chem* 82 (2010), Nr. 16, S. 6958–6968
- [Naarmann *et al.* 2010] NAARMANN, Isabel S. ; HARNISCH, Christiane ; MÜLLER-NEWEN, Gerhard ; URLAUB, Henning ; OSTARECK-LEDERER, Antje ; OSTARECK, Dirk H.: DDX6 recruits translational silenced human reticulocyte 15-lipoxygenase mRNA to RNP granules. In: *RNA* 16 (2010), Nr. 11, S. 2189–2204
- [Nadeau *et al.* 2008] NADEAU, Owen W. ; WYCKOFF, Gerald J. ; PASCHALL, Justin E. ; ARTIGUES, Antonio ; SAGE, Jessica ; VILLAR, Maria T. ; CARLSON, Gerald M.: CrossSearch, a user-friendly search engine for detecting chemically cross-linked peptides in conjugated proteins. In: *Mol. Cell. Proteomics* 7 (2008), Nr. 4, S. 739–749
- [Nakajima u. Ikada 1995] NAKAJIMA, Naoki ; IKADA, Yoshito: Mechanism of Amide Formation by Carbodiimide for Bioconjugation in Aqueous Media. In: *Bioconjugate Chem.* 6 (1995), Nr. 1, S. 123–130
- [Nakamura *et al.* 2001] NAKAMURA, Akira ; AMIKURA, Reiko ; HANYU, Kazuko ; KOBAYASHI, Saturo: Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during *Drosophila* oogenesis. In: *Development* 128 (2001), Nr. 17, S. 3233–3242
- [Nakamura *et al.* 2004] NAKAMURA, Akira ; SATO, Keiji ; HANYU-NAKAMURA, Kazuko: *Drosophila* Cup Is an eIF4E Binding Protein that Associates with Bruno and Regulates oskar mRNA Translation in Oogenesis. In: *Dev. Cell* 6 (2004), Nr. 1, S. 69–78

- [Nelson *et al.* 2004] NELSON, Meryl R. ; LEIDAL, Andrew M. ; SMIBERT, Craig A.: Drosophila Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. In: *EMBO J.* 23 (2004), Nr. 1, S. 150–159
- [Nesterenko *et al.* 1994] NESTERENKO, Michael V. ; TILLEY, Michael ; UPTON, Steve J.: A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in polyacrylamide gels. In: *J. Biochem. Biophys. Meth.* 28 (1994), Nr. 3, S. 239–242
- [Nielsen *et al.* 2008] NIELSEN, Michael L. ; VERMEULEN, Michiel ; BONALDI, Tiziana ; COX, Jürgen ; MORODER, Luis ; MANN, Matthias: Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. In: *Nat. Meth.* 5 (2008), Nr. 6, S. 459–460
- [Nishimura *et al.* 2015] NISHIMURA, Tamiko ; PADAMSI, Zoya ; FAKIM, Hana ; MILETTE, Simon ; DUNHAM, Wade H. ; GINGRAS, Anne-Claude ; FABIAN, Marc R.: The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs. In: *Cell Rep.* 11 (2015), Nr. 9, S. 1425–1436
- [Nissan *et al.* 2010] NISSAN, Tracy ; RAJYAGURU, Purusharth ; SHE, Meipei ; SONG, Haiwei ; PARKER, Roy: Decapping activators in *Saccharomyces cerevisiae* act by multiple mechanisms. In: *Mol. Cell* 39 (2010), Nr. 5, S. 773–783
- [Olsen *et al.* 2007] OLSEN, Jesper V. ; MACEK, Boris ; LANGE, Oliver ; MAKAROV, Alexander ; HORNING, Stevan ; MANN, Matthias: Higher-energy C-trap dissociation for peptide modification analysis. In: *Nat. Meth.* 4 (2007), Nr. 9, S. 709–712
- [Orchard *et al.* 2003] ORCHARD, Sandra ; HERMJAKOB, Henning ; APWEILER, Rolf: The Proteomics Standards Initiative. In: *Proteomics* 3 (2003), Nr. 7, S. 1374–1376
- [Ozgur *et al.* 2015] OZGUR, Sevim ; BASQUIN, Jérôme ; KAMENSKA, Anastasiia ; FILIPOWICZ, Witold ; STANDART, Nancy ; CONTI, Elena: Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. In: *Cell Rep.* 13 (2015), Nr. 4, S. 703–711
- [Parker u. Sheth 2007] PARKER, Roy ; SHETH, Ujwal: P bodies and the control of mRNA translation and degradation. In: *Mol. Cell* 25 (2007), Nr. 5, S. 635–646
- [Peri *et al.* 2001] PERI, Suraj ; STEEN, Hanno ; PANDEY, Akhilesh: GPMAW-a software tool for analyzing proteins and peptides. In: *Trends Biochem. Sci.* 26 (2001), Nr. 11, S. 687–689
- [Perkins *et al.* 1999] PERKINS, David N. ; PAPPIN, Darryl J. C. ; CREASY, David M. ; COTRELL, John S.: Probability-based protein identification by searching sequence databases using mass spectrometry data. In: *Electrophoresis* 20 (1999), Nr. 18, S. 3551–3567

- [Petrotchenko u. Borchers 2010] PETROTCHENKO, Evgeniy V. ; BORCHERS, Christoph H.: ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. In: *BMC Bioinformatics* 11 (2010), S. 64
- [Petrotchenko u. Borchers 2014] PETROTCHENKO, Evgeniy V. ; BORCHERS, Christoph H.: Application of a fast sorting algorithm to the assignment of mass spectrometric cross-linking data. In: *Proteomics* 14 (2014), Nr. 17-18, S. 1987–1989
- [Petrotchenko *et al.* 2014] PETROTCHENKO, Evgeniy V. ; MAKEPEACE, Karl A. T. ; BORCHERS, Christoph H.: DXSMS Match Program for Automated Analysis of LC-MS/MS Data Obtained Using Isotopically Coded CID-Cleavable Cross-Linking Reagents. In: *Curr. Protoc. Bioinformatics* 48 (2014), S. 8.18.1–19
- [Petrotchenko *et al.* 2011] PETROTCHENKO, Evgeniy V. ; SERPA, Jason J. ; BORCHERS, Christoph H.: An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. In: *Mol. Cell. Proteomics* 10 (2011), Nr. 2, S. M110–001420
- [Petrotchenko *et al.* 2012] PETROTCHENKO, Evgeniy V. ; SERPA, Jason J. ; HARDIE, Darryl B. ; BERJANSKII, Mark ; SURIYAMONGKOL, Bow P. ; WISHART, David S. ; BORCHERS, Christoph H.: Use of proteinase K nonspecific digestion for selective and comprehensive identification of interpeptide cross-links: application to prion proteins. In: *Mol. Cell. Proteomics* 11 (2012), Nr. 7, S. M111.013524
- [Pettelkau *et al.* 2014] PETTELKAU, Jens ; IHLING, Christian H. ; FROHBERG, Petra ; WERVEN, Lars van ; JAHN, Olaf ; SINZ, Andrea: Reliable identification of cross-linked products in protein interaction studies by ¹³C-labeled p-benzoylphenylalanine. In: *J. Am. Soc. Mass Spectrom.* 25 (2014), Nr. 9, S. 1628–1641
- [Pettelkau *et al.* 2013] PETTELKAU, Jens ; THONDORF, Iris ; THEISGEN, Stephan ; LILIE, Hauke ; SCHRÖDER, Thomas ; ARLT, Christian ; IHLING, Christian H. ; SINZ, Andrea: Structural Analysis of Guanylyl Cyclase-Activating Protein-2 (GCAP-2) Homodimer by Stable Isotope-Labeling, Chemical Cross-Linking, and Mass Spectrometry. In: *J. Am. Soc. Mass Spectrom.* 24 (2013), Nr. 12, S. 1969–1979
- [Pfirrmann *et al.* 2015] PFIRRMANN, Thorsten ; VILLAVICENCIO-LORINI, Pablo ; SUBUDHI, Abinash K. ; MENSSEN, Ruth ; WOLF, Dieter H. ; HOLLEMANN, Thomas: RMND5 from *Xenopus laevis* is an E3 ubiquitin-ligase and functions in early embryonic forebrain development. In: *PLoS One* 10 (2015), Nr. 3, S. e0120342
- [Pinder 2012] PINDER, Benjamin D.: *Mechanisms of Smaug-mediated post-transcriptional regulation in the early Drosophila embryo*, University of Toronto, Diss., 2012

- [Pinder u. Smibert 2012] PINDER, Benjamin D. ; SMIBERT, Craig A.: microRNA-independent recruitment of Argonaute 1 to nanos mRNA through the Smaug RNA-binding protein. In: *EMBO Rep.* 14 (2012), Nr. 1, S. 80–86
- [Presnyak u. Collier 2013] PRESNYAK, Vlad ; COLLIER, Jeff: The DHH1/RCKp54 family of helicases: An ancient family of proteins that promote translational silencing. In: *Biochim. Biophys. Acta, Gene Regul. Mech.* 1829 (2013), Nr. 8, S. 817–823
- [Putnam *et al.* 2015] PUTNAM, Andrea A. ; GAO, Zhaofeng ; LIU, Fei ; JIA, Huijue ; YANG, Quansheng ; JANKOWSKY, Eckhard: Division of labor in an oligomer of the DEAD-Box RNA helicase ded1p. In: *Mol. Cell* 59 (2015), Nr. 4, S. 541–552
- [Rajyaguru *et al.* 2012] RAJYAGURU, Purusharth ; SHE, Meipei ; PARKER, Roy: Scd6 targets eIF4G to repress translation: RGG motif proteins as a class of eIF4G-binding proteins. In: *Mol. Cell* 45 (2012), Nr. 2, S. 244–254
- [Rasmussen *et al.* 2011] RASMUSSEN, Morten I. ; REFSGAARD, Jan C. ; PENG, Li ; HOUEN, Gunnar ; HØJRUP, Peter: CrossWork: software-assisted identification of cross-linked peptides. In: *J. Proteomics* 74 (2011), Nr. 10, S. 1871–1883
- [Regelmann *et al.* 2003] REGELMANN, Jochen ; SCHÜLE, Thomas ; JOSUPEIT, Frank S. ; HORAK, Jaroslav ; ROSE, Matthias ; ENTIAN, Karl-Dieter ; THUMM, Michael ; WOLF, Dieter H.: Catabolite Degradation of Fructose-1,6-bisphosphatase in the Yeast *Saccharomyces cerevisiae*: A Genome-wide Screen Identifies Eight Novel GID Genes and Indicates the Existence of Two Degradation Pathways. In: *Mol. Biol. Cell* 14 (2003), Nr. 4, S. 1652–1663
- [Rexroth *et al.* 2014] REXROTH, Sascha ; REXROTH, Dorothea ; VEIT, Sebastian ; PLOHNKE, Nicole ; CORMANN, Kai U. ; NOWACZYK, Marc M. ; RÖGNER, Matthias: Functional Characterization of the Small Regulatory Subunit PetP from the Cytochrome b6f Complex in *Thermosynechococcus elongatus*. In: *Plant Cell* 26 (2014), Nr. 8, S. 3435–3448
- [Richter u. Sonenberg 2005] RICHTER, Joel D. ; SONENBERG, Nahum: Regulation of cap-dependent translation by eIF4E inhibitory proteins. In: *Nature* 433 (2005), Nr. 7025, S. 477–480
- [Riffle *et al.* 2016] RIFFLE, Michael ; JASCHOB, Daniel ; ZELTER, Alex ; DAVIS, Trisha N.: ProXL (Protein Cross-Linking Database): A Platform for Analysis, Visualization, and Sharing of Protein Cross-Linking Mass Spectrometry Data. In: *J. Proteome Res.* 15 (2016), Nr. 8, S. 2863–2870

- [Rinner *et al.* 2008] RINNER, Oliver ; SEEBACHER, Jan ; WALZTHOENI, Thomas ; MUELLER, Lukas ; BECK, Martin ; SCHMIDT, Alexander ; MUELLER, Markus ; AEBERSOLD, Ruedi: Identification of cross-linked peptides from large sequence databases. In: *Nat. Meth.* 5 (2008), Nr. 4, S. 315–318
- [Roepstorff u. Fohlman 1984] ROEPSTORFF, Peter ; FOHLMAN, J.: Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. In: *Biol. Mass Spectrom.* 11 (1984), Nr. 11, S. 601–601
- [Rosenthal *et al.* 2016] ROSENTHAL, Arnon ; LEVITEN, Michael ; STEVENS, Beth A. ; HONG, Soyoon ; WILTON, Daniel: Methods of Treatment for Alzheimer’s Disease and Huntington’s Disease; US Patent 20160159890. (2016)
- [Rouget *et al.* 2010] ROUGET, Christel ; PAPIN, Catherine ; BOUREUX, Anthony ; MEUNIER, Anne-Cécile ; FRANCO, Bénédicte ; ROBINE, Nicolas ; LAI, Eric C. ; PELISSON, Alain ; SIMONELIG, Martine: Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. In: *Nature* 467 (2010), Nr. 7319, S. 1128–1132
- [Rouya *et al.* 2014] ROUYA, Christopher ; SIDDIQUI, Nadeem ; MORITA, Masahiro ; DUCHAINE, Thomas F. ; FABIAN, Marc R. ; SONENBERG, Nahum: Human DDX6 effects miRNA-mediated gene silencing via direct binding to CNOT1. In: *RNA* 20 (2014), Nr. 9, S. 1398–1409
- [Sambrook u. Russel 2001] SAMBROOK, Joseph ; RUSSEL, David W.: *Molecular Cloning: a laboratory manual*, Cold Spring Harbour Laboratory Press. Cold Spring Harbour, 2001
- [Santt *et al.* 2008] SANTT, Olivier ; PFIRRMANN, Thorsten ; BRAUN, Bernhard ; JURETSCHKE, Jeannette ; KIMMIG, Philipp ; SCHEEL, Hartmut ; HOFMANN, Kay ; THUMM, Michael ; WOLF, Dieter H.: The yeast GID complex, a novel ubiquitin ligase (E3) involved in the regulation of carbohydrate metabolism. In: *Mol. Biol. Cell* 19 (2008), Nr. 8, S. 3323–3333
- [Sarpe *et al.* 2016] SARPE, Vladimir ; RAFIEI, Atefeh ; HEPBURN, Morgan ; OSTAN, Nicholas ; SCHRYVERS, Anthony B. ; SCHRIEMER, David C.: High Sensitivity Crosslink Detection Coupled With Integrative Structure Modeling in the Mass Spec Studio. In: *Mol. Cell. Proteomics* 15 (2016), Nr. 9, S. 3071–3080
- [Schilling *et al.* 2003] SCHILLING, Birgit ; ROW, Richard H. ; GIBSON, Bradford W. ; GUO, Xin ; YOUNG, Malin M.: MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. In: *J. Am. Soc. Mass Spectrom.* 14 (2003), Nr. 8, S. 834–850

- [Schwarz *et al.* 2013] SCHWARZ, Rico ; TÄNZLER, Dirk ; IHLING, Christian H. ; MÜLLER, Mathias Q. ; KÖLBEL, Knut ; SINZ, Andrea: Monitoring Conformational Changes in Peroxisome Proliferator-Activated Receptor α by a Genetically Encoded Photoamino Acid, Cross-Linking, and Mass Spectrometry. In: *J. Med. Chem.* 56 (2013), Nr. 11, S. 4252–4263
- [Semotok *et al.* 2005] SEMOTOK, Jennifer L. ; COOPERSTOCK, Ramona L. ; PINDER, Benjamin D. ; VARI, Heli K. ; LIPSHITZ, Howard D. ; SMIBERT, Craig A.: Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early *Drosophila* embryo. In: *Curr. Biol.* 15 (2005), Nr. 4, S. 284–294
- [Semotok *et al.* 2008] SEMOTOK, Jennifer L. ; LUO, Hua ; COOPERSTOCK, Ramona L. ; KARAISKAKIS, Angelo ; VARI, Heli K. ; SMIBERT, Craig A. ; LIPSHITZ, Howard D.: *Drosophila* maternal Hsp83 mRNA destabilization is directed by multiple SMAUG recognition elements in the open reading frame. In: *Mol. Cell. Biol.* 28 (2008), Nr. 22, S. 6757–6772
- [Senissar *et al.* 2014] SENISSAR, Meriem ; LE SAUX, Agnès ; BELGAREH-TOUZÉ, Naïma ; ADAM, Céline ; BANROQUES, Josette ; TANNER, N. K.: The DEAD-box helicase Ded1 from yeast is an mRNP cap-associated protein that shuttles between the cytoplasm and nucleus. In: *Nucl. Acids Res.* 42 (2014), Nr. 15, S. 10005–10022
- [Senko *et al.* 2013] SENKO, Michael W. ; REMES, Philip M. ; CANTERBURY, Jesse D. ; MATHUR, Raman ; SONG, Qingyu ; ELIUK, Shannon M. ; MULLEN, Chris ; EARLEY, Lee ; HARDMAN, Mark ; BLETHROW, Justin D. ; BUI, Huy ; SPECHT, August ; LANGE, Oliver ; DENISOV, Eduard ; MAKAROV, Alexander ; HORNING, Stevan ; ZABROUSKOV, Vlad: Novel Parallelized Quadrupole/Linear Ion Trap/Orbitrap Tribrid Mass Spectrometer Improving Proteome Coverage and Peptide Identification Rates. In: *Anal. Chem.* 85 (2013), Nr. 24, S. 11710–11714
- [Sharif *et al.* 2013] SHARIF, Humayun ; OZGUR, Sevim ; SHARMA, Kundan ; BASQUIN, Claire ; URLAUB, Henning ; CONTI, Elena: Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. In: *Nucl. Acids Res.* 41 (2013), Nr. 17, S. 8377–8390
- [Sharma u. Jankowsky 2014] SHARMA, Deepak ; JANKOWSKY, Eckhard: The Ded1/DDX3 subfamily of DEAD-box RNA helicases. In: *Crit. Rev. Biochem. Mol. Biol.* 49 (2014), Nr. 4, S. 343–360
- [Shevchenko *et al.* 2006] SHEVCHENKO, Andrej ; TOMAS, Henrik ; HAVLI, Jan ; OLSEN, Jesper V. ; MANN, Matthias: In-gel digestion for mass spectrometric characterization of proteins and proteomes. In: *Nat. Protoc.* 1 (2006), Nr. 6, S. 2856–2860
- [Shih *et al.* 2008] SHIH, J. W. ; TSAI, T. Y. ; CHAO, C. H. ; LEE, YH W.: Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. In: *Oncogene* 27 (2008), Nr. 5, S. 700–714

- [Sinz 2006] SINZ, Andrea: Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. In: *Mass Spectrom. Rev.* 25 (2006), Nr. 4, S. 663–682
- [Sinz 2014] SINZ, Andrea: The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. In: *Expert Rev. Proteomics* 11 (2014), Nr. 6, S. 733–743
- [Slaidina u. Lehmann 2014] SLAIDINA, Maija ; LEHMANN, Ruth: Translational control in germline stem cell development. In: *J. Cell Biol.* 207 (2014), Nr. 1, S. 13–21
- [Smibert *et al.* 1996] SMIBERT, C A. ; WILSON, J E. ; KERR, K ; MACDONALD, P M.: *smaug* protein represses translation of unlocalized nanos mRNA in the *Drosophila* embryo. In: *Genes Dev.* 10 (1996), Nr. 20, S. 2600–2609
- [Smibert *et al.* 1999] SMIBERT, Craig A. ; LIE, Yung S. ; SHILLINGLAW, Wendy ; HENZEL, William J. ; MACDONALD, Paul M.: *Smaug*, a novel and conserved protein, contributes to repression of nanos mRNA translation in vitro. In: *RNA* 5 (1999), Nr. 12, S. 1535–1547
- [Smith *et al.* 1992] SMITH, Jeffrey L. ; WILSON, Joan E. ; MACDONALD, Paul M.: Overexpression of oskar directs ectopic activation of nanos and presumptive pole cell formation in *Drosophila* embryos. In: *Cell* 70 (1992), Nr. 5, S. 849–859
- [Söderberg *et al.* 2012] SÖDERBERG, Christopher A. G. ; LAMBERT, Wietske ; KJELLSTRÖM, Sven ; WIEGANDT, Alena ; WULFF, Ragna P. ; MÅNSSON, Cecilia ; RUTSDOTTIR, Gudrun ; EMANUELSSON, Cecilia: Detection of Crosslinks within and between Proteins by LC-MALDI-TOF/TOF and the Software FINDX to Reduce the MS/MS-Data to Acquire for Validation. In: *PLoS One* 7 (2012), Nr. 6, S. e38927
- [Soderblom u. Goshe 2006] SODERBLOM, Erik J. ; GOSHE, Michael B.: Collision-Induced Dissociative Chemical Cross-Linking Reagents and Methodology: Applications to Protein Structural Characterization Using Tandem Mass Spectrometry Analysis. In: *Anal. Chem.* 78 (2006), Nr. 23, S. 8059–8068
- [Sonoda u. Wharton 2001] SONODA, Junichiro ; WHARTON, Robin P.: *Drosophila* Brain Tumor is a translational repressor. In: *Genes Dev.* 15 (2001), Nr. 6, S. 762–773
- [Soto-Rifo *et al.* 2012] SOTO-RIFO, Ricardo ; RUBILAR, Paulina S. ; LIMOUSIN, Taran ; BREYNE, Sylvain de ; DÉCIMO, Didier ; OHLMANN, Théophile: DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. In: *EMBO J.* 31 (2012), Nr. 18, S. 3745–3756

- [Soto-Rifo *et al.* 2013] SOTO-RIFO, Ricardo ; RUBILAR, Paulina S. ; OHLMANN, Théophile: The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. In: *Nucl. Acids Res.* 41 (2013), Nr. 12, S. 6286–6299
- [Souquere *et al.* 2009] SOUQUERE, Sylvie ; MOLLET, Stéphanie ; KRESS, Michel ; DAUTRY, François ; PIERRON, Gérard ; WEIL, Dominique: Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. In: *J. Cell Sci.* 122 (2009), Nr. 20, S. 3619–3626
- [Spirin 1994] SPIRIN, Alexander S.: Storage of messenger RNA in eukaryotes: envelopment with protein, translational barrier at 5' side, or conformational masking by 3' side? In: *Mol. Reprod. Dev.* 38 (1994), Nr. 1, S. 107–117
- [St Johnston u. Nüsslein-Volhard 1992] ST JOHNSTON, Daniel ; NÜSSLEIN-VOLHARD, Christiane: The origin of pattern and polarity in the *Drosophila* embryo. In: *Cell* 68 (1992), Nr. 2, S. 201–219
- [Stamnaes *et al.* 2015] STAMNAES, Jorunn ; IVERSEN, Rasmus ; PRÉ, M. Fleur d. ; CHEN, Xi ; SOLLID, Ludvig M.: Enhanced B-Cell Receptor Recognition of the Autoantigen Transglutaminase 2 by Efficient Catalytic Self-Multimerization. In: *PLoS One* 10 (2015), Nr. 8, S. e0134922
- [Stokowy *et al.* 2014] STOKOWY, Tomasz ; ESZLINGER, Markus ; ŚWIERNIAK, Michal ; FUJAREWICZ, Krzysztof ; JARZAB, Barbara ; PASCHKE, Ralf ; KROHN, Knut: Analysis options for high-throughput sequencing in miRNA expression profiling. In: *BMC Research Notes* 7 (2014), Nr. 1, S. 1
- [Syka *et al.* 2004] SYKA, John E. P. ; COON, Joshua J. ; SCHROEDER, Melanie J. ; SHABANOWITZ, Jeffrey ; HUNT, Donald F.: Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. In: *PNAS* 101 (2004), Nr. 26, S. 9528–9533
- [Tadros *et al.* 2007] TADROS, Wael ; GOLDMAN, Aaron L. ; BABAK, Tomas ; MENZIES, Fiona ; VARDY, Leah ; ORR-WEAVER, Terry ; HUGHES, Timothy R. ; WESTWOOD, J. T. ; SMIBERT, Craig A. ; LIPSHITZ, Howard D.: SMAUG is a major regulator of maternal mRNA destabilization in *Drosophila* and its translation is activated by the PAN GU kinase. In: *Dev. Cell* 12 (2007), Nr. 1, S. 143–155
- [Tanaka *et al.* 2006] TANAKA, Kimio J. ; OGAWA, Kenji ; TAKAGI, Masatoshi ; IMAMOTO, Naoko ; MATSUMOTO, Ken ; TSUJIMOTO, Masafumi: RAP55, a cytoplasmic mRNP component, represses translation in *Xenopus* oocytes. In: *J. Biol. Chem.* 281 (2006), Nr. 52, S. 40096–40106

- [Tang *et al.* 2005] TANG, Yong ; CHEN, Yingfeng ; LICHTI, Cheryl F. ; HALL, Roger A. ; RANEY, Kevin D. ; JENNINGS, Steven F.: CLPM: a cross-linked peptide mapping algorithm for mass spectrometric analysis. In: *BMC Bioinformatics* 6 (2005), Nr. Suppl 2, S. S9
- [Taverner *et al.* 2002] TAVERNER, Thomas ; HALL, Nathan E. ; RICHARD, A. J. ; SIMPSON, Richard J.: Characterization of an antagonist interleukin-6 dimer by stable isotope labeling, cross-linking, and mass spectrometry. In: *J. Biol. Chem.* 277 (2002), Nr. 48, S. 46487–46492
- [Temme *et al.* 2014] TEMME, Claudia ; SIMONELIG, Martine ; WAHLE, Elmar: Deadenylation of mRNA by the CCR4–NOT complex in *Drosophila*: molecular and developmental aspects. In: *Front. Genet.* 5 (2014), S. 143
- [Temme *et al.* 2010] TEMME, Claudia ; ZHANG, Lianbing ; KREMMER, Elisabeth ; IHLING, Christian ; CHARTIER, Aymeric ; SINZ, Andrea ; SIMONELIG, Martine ; WAHLE, Elmar: Subunits of the *Drosophila* CCR4–NOT complex and their roles in mRNA deadenylation. In: *RNA* 16 (2010), Nr. 7, S. 1356–1370
- [Thermann u. Hentze 2007] THERMANN, Rolf ; HENTZE, Matthias W.: *Drosophila* miR2 induces pseudo-polysomes and inhibits translation initiation. In: *Nature* 447 (2007), Nr. 7146, S. 875–878
- [Trcek *et al.* 2015] TRCEK, Tatjana ; GROSCH, Markus ; YORK, Andrew ; SHROFF, Hari ; LIONNET, Timothée ; LEHMANN, Ruth: *Drosophila* germ granules are structured and contain homotypic mRNA clusters. In: *Nat. Commun.* 6 (2015), S. 7962
- [Tritschler *et al.* 2009] TRITSCHLER, Felix ; BRAUN, Joerg E. ; EULALIO, Ana ; TRUFFAULT, Vincent ; IZAURRALDE, Elisa ; WEICHENRIEDER, Oliver: Structural basis for the mutually exclusive anchoring of P body components EDC3 and Tral to the DEAD box protein DDX6/Me31B. In: *Mol. Cell* 33 (2009), Nr. 5, S. 661–668
- [Tritschler *et al.* 2008] TRITSCHLER, Felix ; EULALIO, Ana ; HELMS, Sigrun ; SCHMIDT, Steffen ; COLES, Murray ; WEICHENRIEDER, Oliver ; IZAURRALDE, Elisa ; TRUFFAULT, Vincent: Similar Modes of Interaction Enable Trailer Hitch and EDC3 To Associate with DCP1 and Me31B in Distinct Protein Complexes. In: *Mol. Cell. Biol.* 28 (2008), Nr. 21, S. 6695–6708
- [Trnka *et al.* 2014] TRNKA, Michael J. ; BAKER, Peter R. ; ROBINSON, Philip J. J. ; BURLINGAME, A. L. ; CHALKLEY, Robert J.: Matching cross-linked peptide spectra: only as good as the worse identification. In: *Mol. Cell. Proteomics* 13 (2014), Nr. 2, S. 420–434

- [Tseng-Rogenski *et al.* 2003] TSENG-ROGENSKI, Stephanie S.-I. ; CHONG, Jean-Leon ; THOMAS, Christopher B. ; ENOMOTO, Shinichiro ; BERMAN, Judith ; CHANG, Tien-Hsien: Functional conservation of Dhh1p, a cytoplasmic DExD/H-box protein present in large complexes. In: *Nucl. Acids Res.* 31 (2003), Nr. 17, S. 4995–5002
- [de Valoir *et al.* 1991] VALOIR, Tamsen de ; TUCKER, Mark A. ; BELIKOFF, Esther J. ; CAMP, Laura A. ; BOLDUC, Clare ; BECKINGHAM, Kathy: A second maternally expressed Drosophila gene encodes a putative RNA helicase of the "DEAD boxfamily. In: *PNAS* 88 (1991), Nr. 6, S. 2113–2117
- [Verrotti u. Wharton 2000] VERROTTI, Arturo C. ; WHARTON, Robin P.: Nanos interacts with cup in the female germline of Drosophila. In: *Development* 127 (2000), Nr. 23, S. 5225–5232
- [Waghray *et al.* 2015] WAGHRAY, Shruti ; WILLIAMS, Clay ; COON, Joshua J. ; WICKENS, Marvin: Xenopus CAF1 requires NOT1-mediated interaction with 4E-T to repress translation in vivo. In: *RNA* 21 (2015), Nr. 7, S. 1335–1345
- [Wales u. Engen 2006] WALES, Thomas E. ; ENGEN, John R.: Hydrogen exchange mass spectrometry for the analysis of protein dynamics. In: *Mass Spectrom. Rev.* 25 (2006), Nr. 1, S. 158–170
- [Walser u. Lipshitz 2011] WALSER, Claudia B. ; LIPSHITZ, Howard D.: Transcript clearance during the maternal-to-zygotic transition. In: *Curr. Opin. Genet. Dev.* 21 (2011), Nr. 4, S. 431–443
- [Wang *et al.* 1994] WANG, Charlotte ; DICKINSON, Laura K. ; LEHMANN, Ruth: Genetics of nanos localization in Drosophila. In: *Dev. Dyn.* 199 (1994), Nr. 2, S. 103–115
- [Wang *et al.* 2014] WANG, Jian ; ANANIA, Veronica G. ; KNOTT, Jeff ; RUSH, John ; LILL, Jennie R. ; BOURNE, Philip E. ; BANDEIRA, Nuno: Combinatorial approach for large-scale identification of linked peptides from tandem mass spectrometry spectra. In: *Mol. Cell. Proteomics* 13 (2014), Nr. 4, S. 1128–1136
- [Wells u. McLuckey 2005] WELLS, Mitchell J. ; MCLUCKEY, Scott A.: Collision-Induced Dissociation (CID) of Peptides and Proteins. In: ENZYMOLOGY, BT Methods i. (Hrsg.): *Methods in Enzymology* Bd. 402. Academic Press, 2005, S. 148–185
- [Weston u. Sommerville 2006] WESTON, Andrew ; SOMMERVILLE, John: Xp54 and related (DDX6-like) RNA helicases: roles in messenger RNP assembly, translation regulation and RNA degradation. In: *Nucl. Acids Res.* 34 (2006), Nr. 10, S. 3082–3094
- [Wharton u. Struhl 1991] WHARTON, Robin P. ; STRUHL, Gary: RNA regulatory elements mediate control of Drosophila body pattern by the posterior morphogen nanos. In: *Cell* 67 (1991), Nr. 5, S. 955–967

- [Wilhelm *et al.* 2005] WILHELM, James E. ; BUSZCZAK, Michael ; SAYLES, Suzanne: Efficient Protein Trafficking Requires Trailer Hitch, a Component of a Ribonucleoprotein Complex Localized to the ER in *Drosophila*. In: *Dev. Cell* 9 (2005), Nr. 5, S. 675–685
- [Wilhelm *et al.* 2003] WILHELM, James E. ; HILTON, Meredith ; AMOS, Quinlan ; HENZEL, William J.: Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. In: *J. Cell Biol.* 163 (2003), Nr. 6, S. 1197–1204
- [Wittelsberger *et al.* 2006] WITTELSBERGER, Angela ; THOMAS, Beena E. ; MIERKE, Dale F. ; ROSENBLATT, Michael: Methionine acts as a “magnet” in photoaffinity crosslinking experiments. In: *FEBS Lett.* 580 (2006), Nr. 7, S. 1872–1876
- [Wondrak 2003] WONDRAK, E.M.: Solution for fast visualization of protein; US Patent 6,555,382 B2. (2003). – US Patent 6,555,382
- [Worringer *et al.* 2009] WORRINGER, Kathleen A. ; CHU, Feixia ; PANNING, Barbara: The zinc finger protein Zn72D and DEAD box helicase Belle interact and control maleless mRNA and protein levels. In: *BMC Mol. Biol.* 10 (2009), S. 33
- [Xiol *et al.* 2014] XIOL, Jordi ; SPINELLI, Pietro ; LAUSSMANN, Maike A. ; HOMOLKA, David ; YANG, Zhaolin ; CORA, Elisa ; COUTÉ, Yann ; CONN, Simon ; KADLEC, Jan ; SACHIDANANDAM, Ravi ; KAKSONEN, Marko ; CUSACK, Stephen ; EPHRUSSI, Anne ; PILLAI, Ramesh S.: RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. In: *Cell* 157 (2014), Nr. 7, S. 1698–1711
- [Xu *et al.* 2008] XU, Hua ; ZHANG, Liwen ; FREITAS, Michael A.: Identification and Characterization of Disulfide Bonds in Proteins and Peptides from Tandem MS Data by Use of the MassMatrix MS/MS Search Engine. In: *J. Proteome Res.* 7 (2008), Nr. 1, S. 138–144
- [Xue *et al.* 2010] XUE, Bin ; DUNBRACK, Roland L. ; WILLIAMS, Robert W. ; DUNKER, A. K. ; UVERSKY, Vladimir N.: PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. In: *Biochim. Biophys. Acta, Proteins Proteomics* 1804 (2010), Nr. 4, S. 996–1010
- [Yang *et al.* 2012] YANG, Bing ; WU, Yan-Jie ; ZHU, Ming ; FAN, Sheng-Bo ; LIN, Jinzhong ; ZHANG, Kun ; LI, Shuang ; CHI, Hao ; LI, Yu-Xin ; CHEN, Hai-Feng ; LUO, Shu-Kun ; DING, Yue-He ; WANG, Le-Heng ; HAO, Zhiqi ; XIU, Li-Yun ; CHEN, She ; YE, Keqiong ; HE, Si-Min ; DONG, Meng-Qiu: Identification of cross-linked peptides from complex samples. In: *Nat. Meth.* 9 (2012), Nr. 9, S. 904–906
- [Yang *et al.* 2007] YANG, Quansheng ; DEL CAMPO, Mark ; LAMBOWITZ, Alan M. ; JANKOWSKY, Eckhard: DEAD-box proteins unwind duplexes by local strand separation. In: *Mol. Cell* 28 (2007), Nr. 2, S. 253–263

- [Yang u. Jankowsky 2005] YANG, Quansheng ; JANKOWSKY, Eckhard: ATP-and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. In: *Biochemistry* 44 (2005), Nr. 41, S. 13591–13601
- [Yarunin *et al.* 2011] YARUNIN, Alexander ; HARRIS, Robin E. ; ASHE, Mark P. ; ASHE, Hilary L.: Patterning of the Drosophila oocyte by a sequential translation repression program involving the d4EHP and Belle translational repressors. In: *RNA Biol.* 8 (2011), Nr. 5, S. 904–912
- [Yılmaz *et al.* 2016] YILMAZ, Şule ; DREPPER, Friedel ; HULSTAERT, Niels ; ČERNIČ, Maša ; GEVAERT, Kris ; ECONOMOU, Anastassios ; WARSCHIED, Bettina ; MARTENS, Lennart ; VANDERMARLIERE, Elie: Xilmass: A New Approach toward the Identification of Cross-Linked Peptides. In: *Anal. Chem.* 88 (2016), Nr. 20, S. 9949–9957
- [Young *et al.* 2000] YOUNG, Malin M. ; TANG, Ning ; HEMPEL, Judith C. ; OSHIRO, Connie M. ; TAYLOR, Eric W. ; KUNTZ, Irwin D. ; GIBSON, Bradford W. ; DOLLINGER, Gavin: High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. In: *PNAS* 97 (2000), Nr. 11, S. 5802
- [Zaessinger *et al.* 2006] ZAESSINGER, Sophie ; BUSSEAU, Isabelle ; SIMONELIG, Martine: Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. In: *Development* 133 (2006), Nr. 22, S. 4573–4583
- [Zampedri *et al.* 2016] ZAMPEDRI, Cecilia ; TINOCO-CUELLAR, Maryana ; CARRILLO-ROSAS, Samantha ; DIAZ-TELLEZ, Abigail ; RAMOS-BALDERAS, Jose L. ; PELEGRI, Francisco ; MALDONADO, Ernesto: Zebrafish P54 RNA helicases are cytoplasmic granules residents that are required for development and stress resilience. In: *Biol. Open* 5 (2016), Nr. 10, S. 1473–1484
- [Zappavigna *et al.* 2004] ZAPPAVIGNA, Vincenzo ; PICCIONI, Federica ; VILLAESCUSA, J. C. ; VERROTTI, Arturo C.: Cup is a nucleocytoplasmic shuttling protein that interacts with the eukaryotic translation initiation factor 4E to modulate Drosophila ovary development. In: *PNAS* 101 (2004), Nr. 41, S. 14800–14805

Abbildungsverzeichnis

Abb. 1:	Frühe Entwicklung von <i>Drosophila melanogaster</i>	4
Abb. 2:	mRNA-Abbau	6
Abb. 3:	Verteilung der <i>nanos</i> -mRNA und des Nanos-Proteins im Embryo . . .	7
Abb. 4:	Regulatorische Sequenzen im 3'-UTR der <i>nanos</i> -mRNA	9
Abb. 5:	Cup-Modell der Smaug-abhängigen Repression	10
Abb. 6:	Strukturmodell der DDX6/Me31B-Interaktionen	16
Abb. 7:	Analytische Sedimentation von 1-AUG-RNA	20
Abb. 8:	RNA-Abbau in <i>Drosophila</i> -Embryoextrakt	21
Abb. 9:	Experimentierschema: Stabilität der Repression	22
Abb. 10:	RNA bleibt während Gradientenzentrifugation reprimiert	24
Abb. 11:	Präparation des reprimierten RNP	25
Abb. 12:	Quantitative Analyse der RNA-gebundenen Proteine	27
Abb. 13:	Vulkandiagramm der identifizierten Proteine im reprimierten RNP . .	29
Abb. 14:	Ubiquitinierung von Me31B	33
Abb. 15:	Quantitative <i>Western-blot</i> -Analyse der Repressor-Komplexkomponenten	38
Abb. 16:	Längenabhängige Stöchiometrie der Komponenten im Repressor-Komplex	39
Abb. 17:	Keine Oligomerisierung SRE-haltiger RNAs	42
Abb. 18:	MS-Analyse der Smaug-Immunpräzipitation	45
Abb. 19:	Vergleich des RNA- <i>pull-downs</i> und der Smaug-IP	46
Abb. 20:	<i>Cross-Linking</i> -Analyse der Smaug-Immunpräzipitation	47
Abb. 21:	Modell des Repressor-Komplexes	63
Abb. 22:	<i>Cross-Linking</i> /MS-Strategie	73
Abb. 23:	<i>Cross-Link</i> -Typen	74
Abb. 24:	Chemische Reaktionen für die Protein-Quervernetzung	77
Abb. 25:	Orbitrap Fusion Tribrid-Massenspektrometer	78
Abb. 26:	Fragmentierung von Peptiden und <i>Cross-Links</i>	79
Abb. 27:	Beispiele für MS/MS-spaltbare <i>Cross-Linker</i>	82
Abb. 28:	Fragmentierung des Harnstoff- <i>Linkers</i>	83
Abb. 29:	Definitionen zum <i>Cross-Link</i> -Identifizierungs-Algorithmus	87
Abb. 30:	Identifizierungs-Algorithmus	93
Abb. 31:	RISE-Modus	95
Abb. 32:	Korrelationsanalyse verschobener Spektren	100

Abb. 33:	Automatische <i>Decoy</i> -Analyse	103
Abb. 34:	MeroX-Hauptfenster	110
Abb. 35:	Annotiertes <i>Cross-Link</i> -Spektrum	111
Abb. 36:	Struktur einer Ergebnisdatei	114
Abb. 37:	Verbesserung des RISE-Algorithmus	121
Abb. 38:	Einfluss der <i>short</i> RNAs auf die Translationsrepression	XLI
Abb. 39:	Anreicherung/Abreicherung der Proteine im RNA- <i>pull-down</i>	XLII
Abb. 40:	Translationsinitiationsfaktoren und ribosomale Proteine	XLIV
Abb. 41:	Spektrum eines LRGG-modifizierten Peptids	XLV
Abb. 42:	Spektren eines GG-modifizierten Peptids	XLVI
Abb. 43:	Längenabhängige Stöchiometrie der Komponenten im Komplex	XLVII
Abb. 44:	<i>Deep-sequencing</i> -Analyse des <i>pull-downs</i>	XLVIII
Abb. 45:	Beschriebene Interaktionen innerhalb des Repressor-Komplexes	LII
Abb. 46:	Intrinsisch unstrukturierte Bereiche im Belle-Protein	LIII
Abb. 47:	Mögliche Fragmentationen nach <i>Cross-Linker</i> -Fragmentierung	LIV
Abb. 48:	Korrektur der Spektrumsgröße	LV
Abb. 49:	Screenshot der automatischen <i>Decoy</i> -Analyse	LVI

Tabellenverzeichnis

Tab. 1:	Im RNA- <i>pull-down</i> identifizierte Proteine	30
Tab. 2:	Proteinkonzentrationen in Embryoextrakten	37
Tab. 3:	RNA-Längenabhängigkeit der Repressor-Komplex-Stöchiometrie	40
Tab. 4:	Not1-assoziierte Proteine des SRE-RNPs	52
Tab. 5:	Programme zur <i>Cross-Linking</i> /MS-Auswertung	86
Tab. 6:	Faktoren zur Berechnung der <i>subscores</i>	101
Tab. 7:	Einfluss der Größe der Sequenzdatenbank	105
Tab. 8:	Liste verwendeter Chemikalien	123
Tab. 9:	Liste verwendeter Geräte	126
Tab. 10:	Angewendete Software	126
Tab. 11:	Liste verwendeter Plasmide	127
Tab. 12:	Primäre Antikörper	127
Tab. 13:	Sekundäre Antikörper	128
Tab. 14:	<i>In-vitro</i> -Transkriptionsreaktion	129
Tab. 15:	Translationsreaktionen	132
Tab. 16:	Puffer für die Reinigung des reprimierten RNPs	133
Tab. 17:	Abkürzungen	XXXV
Tab. 18:	Im RNA- <i>pull-down</i> angereicherte Proteine	XLIII
Tab. 19:	Angereicherte Proteine in der Smaug-Immunpräzipitation	XLIX
Tab. 20:	Referenzen zu Interaktionen im reprimierten RNP	LII

Abkürzungsverzeichnis

Tabelle 17: verwendete Abkürzungen

Abkürzung	Bedeutung
1-AUG	Transkript mit einem einzigen Startcodon (AUG)
4EBP	eIF4E-bindendes Protein
Å	Ångström (1 Å = 0,1 nm)
A/D-Wandler	Analog-Digital-Wandler
A/P-Achse	Anterior-posterior-Achse
A280	Absorption bei 280 nm
AEBSF	4-(2-Aminoethyl)benzensulfonylfluorid
APS	Ammoniumperoxodisulfat
ARS	ATP-regenerierendes System
AspN	Endoprotease AspN (Proteolyse N-terminal von sauren Aminosäuren)
ATP	Adenosintriphosphat
Bel	Belle (<i>Drosophila</i> -DDX3-Helikase)
blastp	<i>Basic local alignment search tool for proteins</i>
Bp	Basenpaare
BS ² G	Bissulfosuccinimidylglutarat (<i>Cross-Linker</i>)
BS ³	Bissulfosuccinimidylsuberat (<i>Cross-Linker</i>)
BSA	Bovines Serumalbumin
Bu	4-Aminobuttersäure
BuUr	4-Isocyanato-Buttersäure
<i>C. elegans</i>	<i>Caenorhabditis elegans</i> – Fadenwurm

Fortsetzung auf der nächsten Seite

Tabelle 17 – Fortsetzung

Abkürzung	Bedeutung
cDNA	Komplementäre DNA (nach reverser Transkription von mRNA)
CHD	Cup Homologiedomäne
CID	Kollisions-induzierte Dissoziation (<i>collision-induced dissociation</i>)
CTLH	C-terminal zur Lis-Homologie-Domäne
d _n	Substanz enthält <i>n</i> Deuteriumatome anstelle von Wasserstoffatomen
Da	Dalton (1 Da=1 u)
DEAD	Sequenzmotiv, Einbuchstabencode
DEE	<i>Drosophila</i> -Embryoextrakt
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
DSS	Disuccinimidylsuberat (<i>Cross-Linker</i>)
DSSO	Disuccinimidylsulfoxid (<i>Cross-Linker</i>)
DTT	Dithiothreitol
DUF	Domäne mit unbekannter Funktion
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimid (zur Aktivierung von Säuregruppen)
<i>E. coli</i>	<i>Escherichia coli</i>
Edc3	<i>Enhancer of decapping 3</i>
EDTA	Ethylendiamintetraacetat
eIF	Eukaryotischer Translationsinitiationsfaktor
EJC	<i>Exon junction complex</i>
ESI	Elektrospray Ionisierung
ETciD	Kombination aus ETD und CID
ETD	Elektronentransfer Dissoziation
EThcD	Kombination aus ETD und HCD
FBPase	Fruktose-1,6-bisphosphatase
FDF	Sequenzmotiv, Einbuchstabencode
FDR	Falsch-positiv-Rate (<i>False discovery rate</i>)

Fortsetzung auf der nächsten Seite

Tabelle 17 – Fortsetzung

Abkürzung	Bedeutung
FFD-TFG	aufeinander folgende Sequenzmotive, Einbuchstabencode
FT	Durchfluss (<i>flow through</i>) einer Chromatografie, ungebundene Fraktion
GID	<i>Glucose induced degradation deficient</i>
GluC	Endoprotease GluC (Proteolyse C-terminal von sauren Aminosäuren)
GUI	Grafische Benutzeroberfläche (<i>Graphical user interface</i>)
HCD	<i>Higher-energy C-trap dissociation</i> oder <i>Higher-energy collisional dissociation</i>
HEPES	4-(2-Hydroxyethyl)-1-piperazinethansulfonsäure
HPLC	<i>High performance liquid chromatography</i>
HUPO	<i>Human proteome organisation</i>
IP	Immunpräzipitation
IRES	Interne ribosomale Eintrittsstelle
IRM	<i>Ion routing multipole</i>
kb	Kilobasenpaare
K_D	Dissoziationskonstante
LC	Flüssigchromatografie
LisH	Lis-Homologie-Domäne
LSm	<i>Like Sm</i> -Domäne
Luc	Luziferase
MBT	Midblastula Übergang
MG132	Proteasominhibitor
Me31B	<i>Maternally expressed at chromosome 31B</i> (<i>Drosophila</i> -DDX6-Helikase)
MIF4G	Mittlere Domäne von eIF4G
miRNA	mikro RNA
MLU	Martin-Luther-Universität
modENCODE	<i>Model organism ENCyclopedia Of DNA Elements</i>

Fortsetzung auf der nächsten Seite

Tabelle 17 – Fortsetzung

Abkürzung	Bedeutung
mRNA	Boten-RNA (<i>messenger RNA</i>)
MS	Massenspektrometrie
MS ⁿ	<i>n</i> aufeinanderfolgende Selektions- und Fragmentierungsereignisse
MS/MS-spaltbar	Struktur, die bei Energiezuführung während der MS-Analyse (z.B. durch CID oder HCD) fragmentiert
MS/MS-Spektrum	Fragmentionen-Spektrum, Tandemmassenspektrum
MS-Spektrum	Massenspektrum, Vorläuferspektrum
MUT	Mutante
MW	Molekulargewicht in kDa
<i>m/z</i>	Masse-zu-Ladung-Verhältnis
MZT	Maternal-zygotischer Übergang
Ni ²⁺ -NTA	Ni ²⁺ -Nitrilotriessigsäure
NCBI	<i>National center for biotechnology information</i>
NHS	N-Hydroxy-Succinimid
NMD	<i>Nonsense-mediated mRNA Decay</i> – Abbau aberanter mRNAs
NP-40	Nonidet TM P-40
nt	Nukleotide
ORF	<i>Open reading frame</i> – offenes Leseraster
Osk	Oskar
PA	Polyacrylamid
PABPC	Zytosolisches Poly(A)-Bindepotein
PAGE	Polyacrylamid Gelelektrophorese
p-Bpa	Para-Benzoyl-Phenylalanin
PD	<i>Pull-down</i>
piRNA	Piwi-interagierende RNA
Poly(A)	Polyadenylat
ppm	<i>Parts per million</i>
PSI	<i>Proteomics standard initiative</i>

Fortsetzung auf der nächsten Seite

Tabelle 17 – Fortsetzung

Abkürzung	Bedeutung
PSM	<i>Peptide spectrum match</i>
RAP55	<i>RNA interacting protein of 55 kDa</i> – Trailer-Hitch-Homolog
RGG	Sequenzmotiv, Einbuchstabencode
RING	<i>Really interesting new gene</i>
RISC	RNA-induzierter Silencer-Komplex
RISE	<i>Reporter ion scan event</i> – MeroX-Algorithmus
RNA	Ribonukleinsäure
RNP/mRNP	Ribonukleoprotein-Partikel/messenger-Ribonukleoprotein-Partikel
RpL	Ribosomales Protein der großen Untereinheit
RpS	Ribosomales Protein der kleinen Untereinheit
RRL	Kaninchen-Retikulozytenlysats
S2-Zellen	Schneider-2-Zellen
SAM	<i>Sterile alpha motif</i>
SAXS	<i>Small-angle X-ray scattering</i>
SDS	Natriumdodecylsulfat
Smg	Smaug
SRE	Smaug-Erkennungselement (<i>Smaug recognition element</i>)
SREonly	Transkript mit zwei SRE- <i>stem-loops</i>
SuDP	Disuccinimidyl-succinamyl-aspartyl-prolin (<i>Cross-Linker</i>)
SUMO	<i>Small ubiquitin-like modifier</i>
TAP	<i>Tandem affinity purification</i>
TCE	Translations-Kontrollelement
Temed	N,N,N',N'-Tetramethylethyldiamin
TFA	Trifluoressigsäure
Th	Thomson ($1 \text{ Th} = 1 \frac{u}{z}$)
Tral	Trailer Hitch
Tris	Tris(hydroxymethyl)aminoethan

Fortsetzung auf der nächsten Seite

Tabelle 17 – Fortsetzung

Abkürzung	Bedeutung
USB	<i>Universal serial bus</i> – serielle Computerschnittstelle
UTR	Untranslatierter Bereich (<i>untranslated region</i>)
UUID	<i>Universally unique identifier</i>
UV	Ultraviolett
WT	Wildtyp
XDB	<i>Cross-Link</i> -Datenbank
XL-MS	<i>Cross-Linking</i> in Kombination mit Massenspektrometrie
YxxxxL ϕ	Sequenzmotiv eines eIF4E-Bindemotivs
ZGA	Zygotische Genomaktivierung
Dateiformate	
CSV	<i>Comma separated values</i> , Tabellenformat
MXF	Einstellungsdateiformat MeroX
mzIdentML	Ergebnisdateiformat für MS-Analysen
mzML	Massenspektrometrisches Dateiformat
mzXML	Massenspektrometrisches Dateiformat
proXML	<i>Cross-Link</i> -Dateiformat
SSF	Einstellungsdateiformat StavroX
TXT	Text-Dateiformat
XML	<i>Extended markup language</i> , generelles Dateiformat
ZHRS	<i>Cross-Link</i> -Dateiformat StavroX
ZHRM	<i>Cross-Link</i> -Dateiformat MeroX
ZIP	Komprimiertes Dateiformat

Anhang

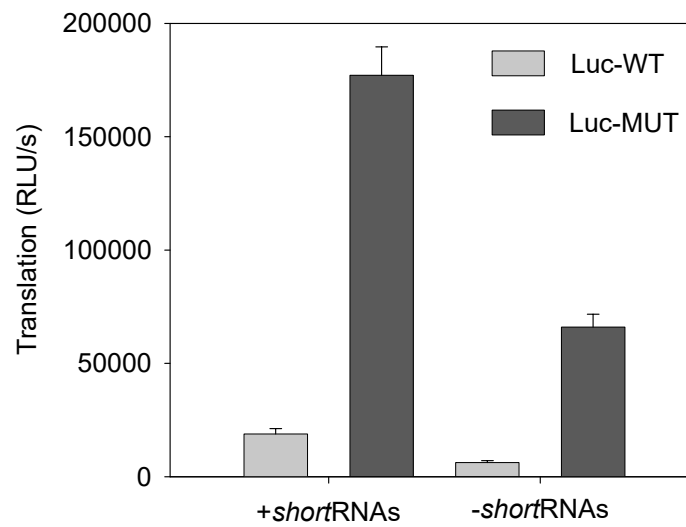


Abbildung 38: Einfluss der *shortRNAs* auf die Translationsrepression - Translationsrepressionsmessung nach Standardprotokoll in An- und Abwesenheit von *shortRNAs*. Die Zugabe von *shortRNAs* führte zu einer etwa dreifach höheren Translationsausbeute sowohl für die WT- als auch MUT-RNA. Die WT-RNA wurde in beiden Fällen etwa zehnfach reprimiert.

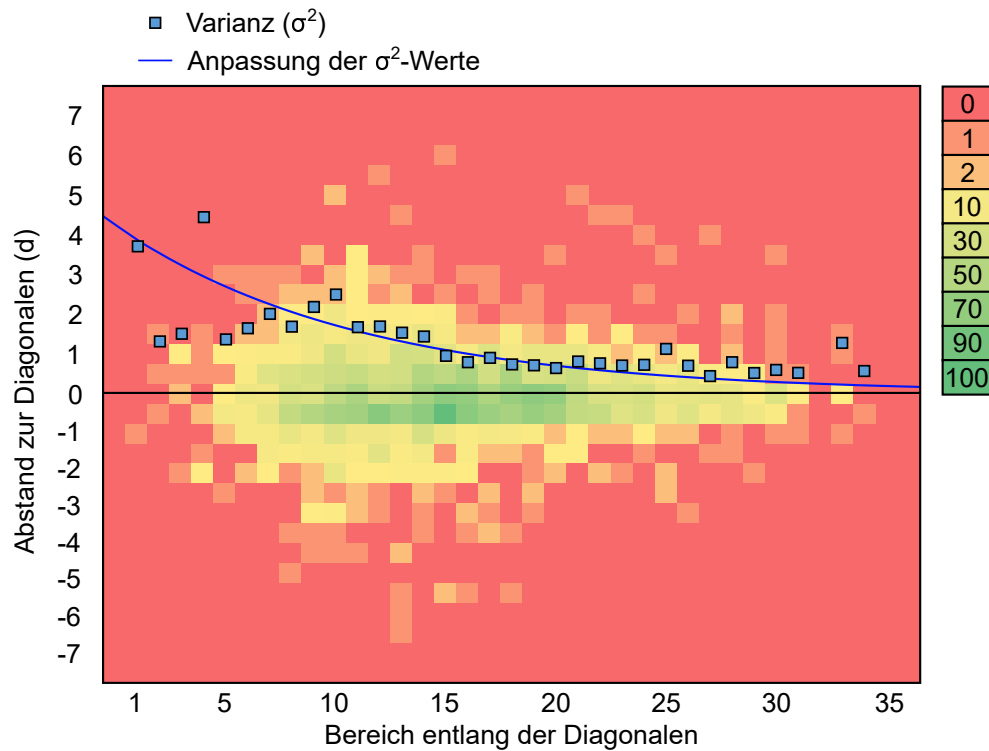


Abbildung 39: Anreicherung/Abreicherung der Proteine im RNA-pull-down
 - Statistische Auswertung der Anreicherung und Abreicherung der identifizierten Proteine im RNA-pull-down (vgl. Abb. 12). Die Daten aus Abb. 12 wurden transformiert (um 45° im Uhrzeigersinn gedreht), so dass nicht angereichert Proteine einen Wert von null erhalten. Die Achsen wurde dann in Bereiche eingeteilt, um ein zweidimensionales Histogramm der Messpunktverteilung zu erhalten. Die Anzahl an Datenpunkten je Bereich ist farblich kodiert (siehe Skala rechts). Die statistische Verteilung der identifizierten Proteine in den einzelnen Bereichen der x-Achse wurde an eine Normalverteilung angepasst und daraus der σ^2 -Wert extrahiert und aufgetragen (blaue Boxen). Die σ^2 -Werte wurden anschließend an Gleichung 4 angepasst. Aus dieser Anpassung konnten dann mit Gleichung 5 die p-Werte für die Anreicherung der identifizierten Proteine berechnet werden.

Tabelle 18: Im RNA-pull-down (PD) abgereicherte Proteine - Alle signifikant abgereicherten Proteine ($p < 5\%$) der RNA-pull-down-Analyse sind mit Intensitätswerten, Abreicherungsfaktor und Signifikanz aufgeführt. Ribosomale Proteine sind in der Tabelle hervorgehoben.

Protein	Intensität _{korrr}		p-Wert	Abreicherung
	WT-PD	MUT-PD		
RpS10b	26.29	30.92	$8.219 \cdot 10^{-6}$	24.84
l(3)72Ab	25.51	29.69	$3.188 \cdot 10^{-4}$	18.12
SmD2	18.98	26.60	$3.937 \cdot 10^{-4}$	197.0
RpL23	28.42	31.42	$4.748 \cdot 10^{-4}$	8.014
His4	25.73	29.62	$6.631 \cdot 10^{-4}$	14.82
RpS26	17.72	25.69	0.00113	251.2
RpS15Aa	26.15	29.67	0.00136	11.51
RpS18	28.02	30.89	0.00148	7.288
RpL36	24.48	28.28	0.00394	13.96
SmG	15.09	23.97	0.00504	470.8
RpL30	24.05	27.91	0.00525	14.49
RpL35	19.90	25.55	0.00674	50.44
His2B	21.60	26.36	0.00728	27.14
RpL28	27.76	30.14	0.0103	5.211
RpS25	26.16	28.98	0.0109	7.050
RpLP1	19.93	25.23	0.0116	39.26
hoip	19.34	24.94	0.0117	48.44
Nelf-A	15.14	23.28	0.0118	281.6
BcDNA.LD23634	19.71	25.06	0.0126	40.93
FLASH	15.10	23.16	0.0132	267.8
ctp	19.20	24.61	0.0167	42.51
CG31368	22.76	26.52	0.0179	13.51
His3.3A	21.22	25.56	0.0193	20.35
tyf	23.89	27.14	0.0213	9.468
RpS28b	21.67	25.74	0.0217	16.89
RpS14b	31.87	33.19	0.0218	2.481
Gem3	22.78	26.25	0.0282	11.08
RpS15	27.67	29.69	0.0290	4.054
Cbp20	18.81	23.91	0.0308	34.25
RpL32	23.63	26.65	0.0361	8.109
RpL27	26.42	28.64	0.0364	4.645
CG2941	25.97	28.29	0.0375	4.977
CG3003	16.53	22.60	0.0386	67.32
CG17737	18.92	23.68	0.0419	27.00
RpL38	21.69	25.25	0.0437	11.77
SmD3	19.98	24.21	0.0441	18.80
CG10979	16.26	22.30	0.0450	65.91
SMC3	25.00	27.46	0.0459	5.505
RpS23	27.18	29.07	0.0488	3.720

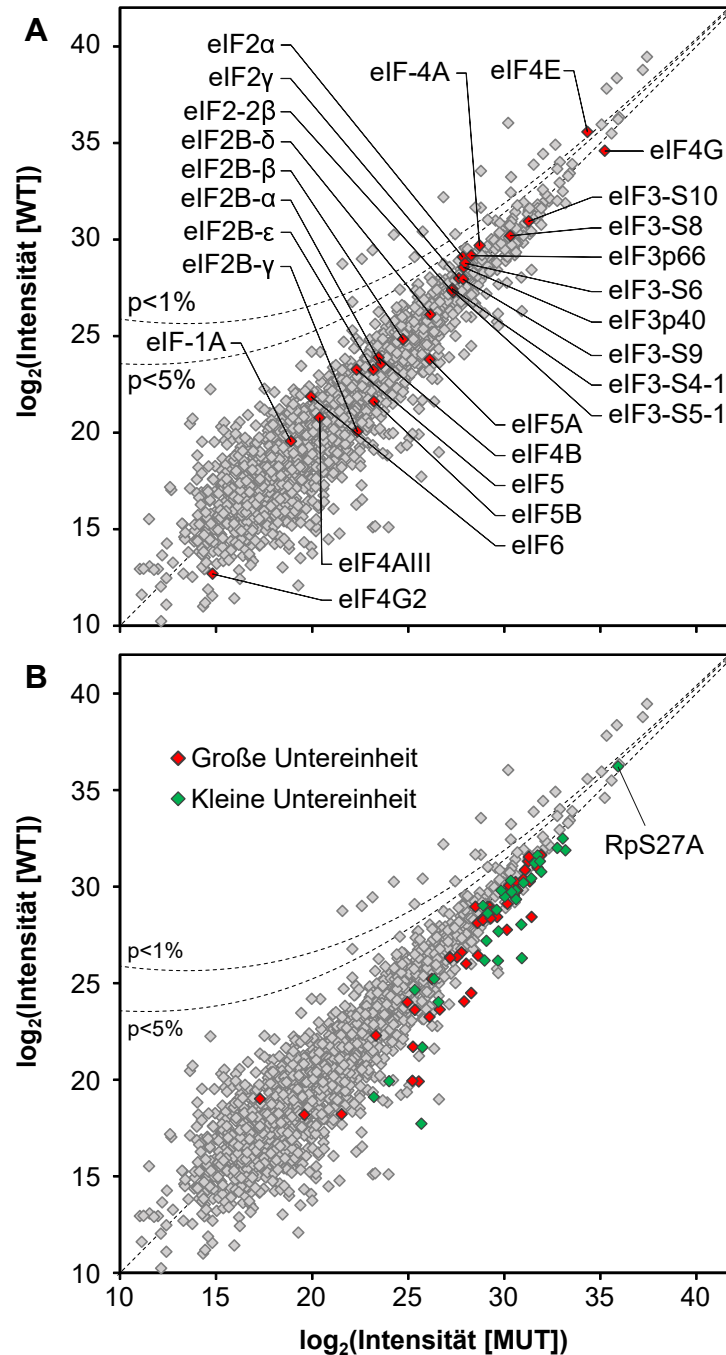


Abbildung 40: Translationsinitiationsfaktoren und ribosomale Proteine im RNA-pull-down - Die Intensitäten von Translationsinitiationsfaktoren (A) und ribosomalen Proteinen (B) im RNA-pull-down sind wie in Abb. 12 dargestellt. Verschiedene Translationsinitiationsfaktoren banden im gleichen Maß sowohl an die WT- als auch die MUT-RNA. Ribosomale Proteine waren im WT-pull-down eher abgereichert. Das ribosomale Protein RpS27A ist als Fusionsprotein aus Ubiquitin und einem ribosomalen Protein im C-Terminus in *Drosophila* kodiert. Die meisten RpS27A-Peptidspektren konnten dem Ubiquitin-Anteil zugewiesen werden. Die hohe Abundanz von RpS27A ist durch die hohe Abundanz von Ubiquitin in den Proben zu erklären.

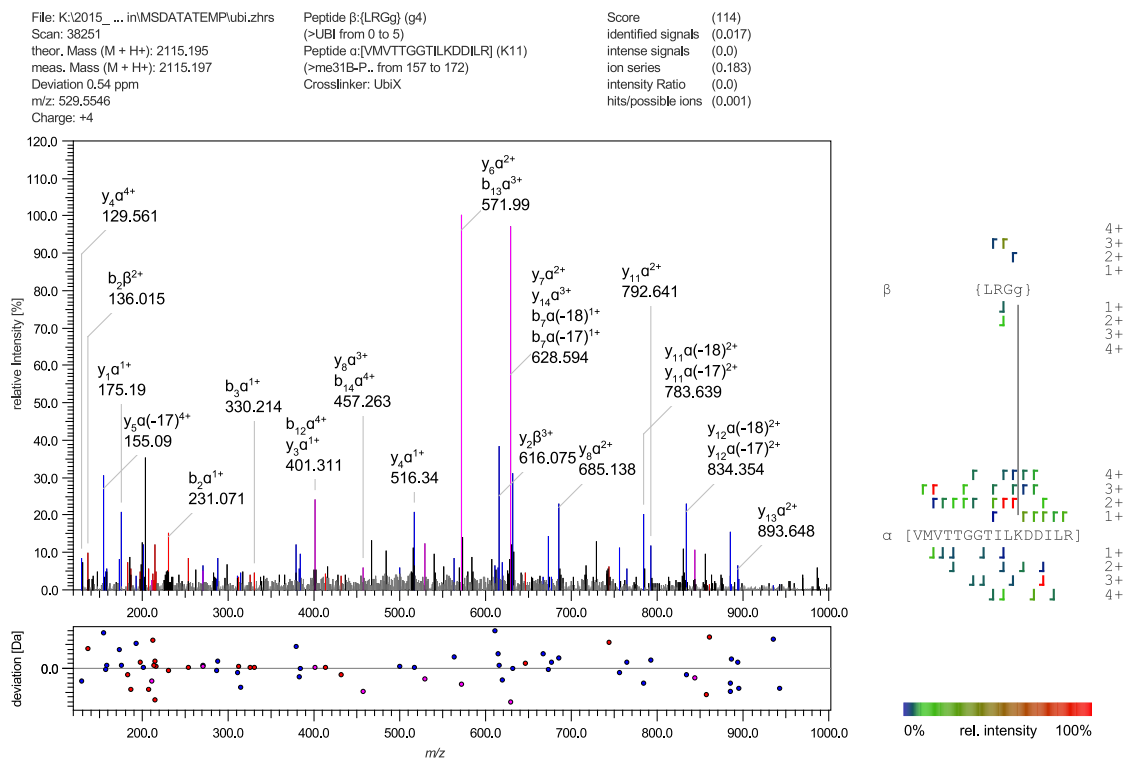


Abbildung 41: Spektrum eines LRGG-modifizierten Peptids - Es ist das annotierte Spektrum eines vierfach positiv geladenen Vorläuferions bei m/z 529,5546 gezeigt. Es wurde als LRGG-modifiziertes Peptid (VMVTTGGTILKDDILR) von Me31B identifiziert. Die Annotation erfolgte automatisch mittels StavroX (siehe auch Teil II, Abschnitt 2.4.3, S. 112). Im linken Teil ist das Fragmentionen-Spektrum und ein Abweichungsdiagramm dargestellt. Die Nomenklatur der annotierten Ionen richtet sich nach Roepstorff u. Fohlman (1984) und Schilling *et al.* (2003). Die Farben im Spektrum zeigen verschiedene Ionentypen an (N-Terminal - rot, C-Terminal - blau, Vorläuferion - grün, verschiedene Typen - magenta). Im rechten Teil ist eine Fragmentierungsansicht gezeigt. Jeder Indikator (⌈ oder ⌋) über oder unter der Peptidsequenz zeigt ein Fragmention an. Die Position des Indikators ist durch die Stelle der Fragmentierung und der Ladung des Ions (rechts angezeigt) festgelegt. Die Indikatoren sind anhand ihrer Intensität nach der Skala (unterhalb) eingefärbt. Das kleingeschriebene „g“ in der Sequenz sowie eckige oder geschweifte Klammern werden durch die Syntax von StavroX vorgegeben. Im oberen Teil sind allgemeine Informationen zu dem ubiquitinierten Peptid angegeben (bestimmte und theoretische Masse, Sequenz, Herkunft, *score*, usw.).

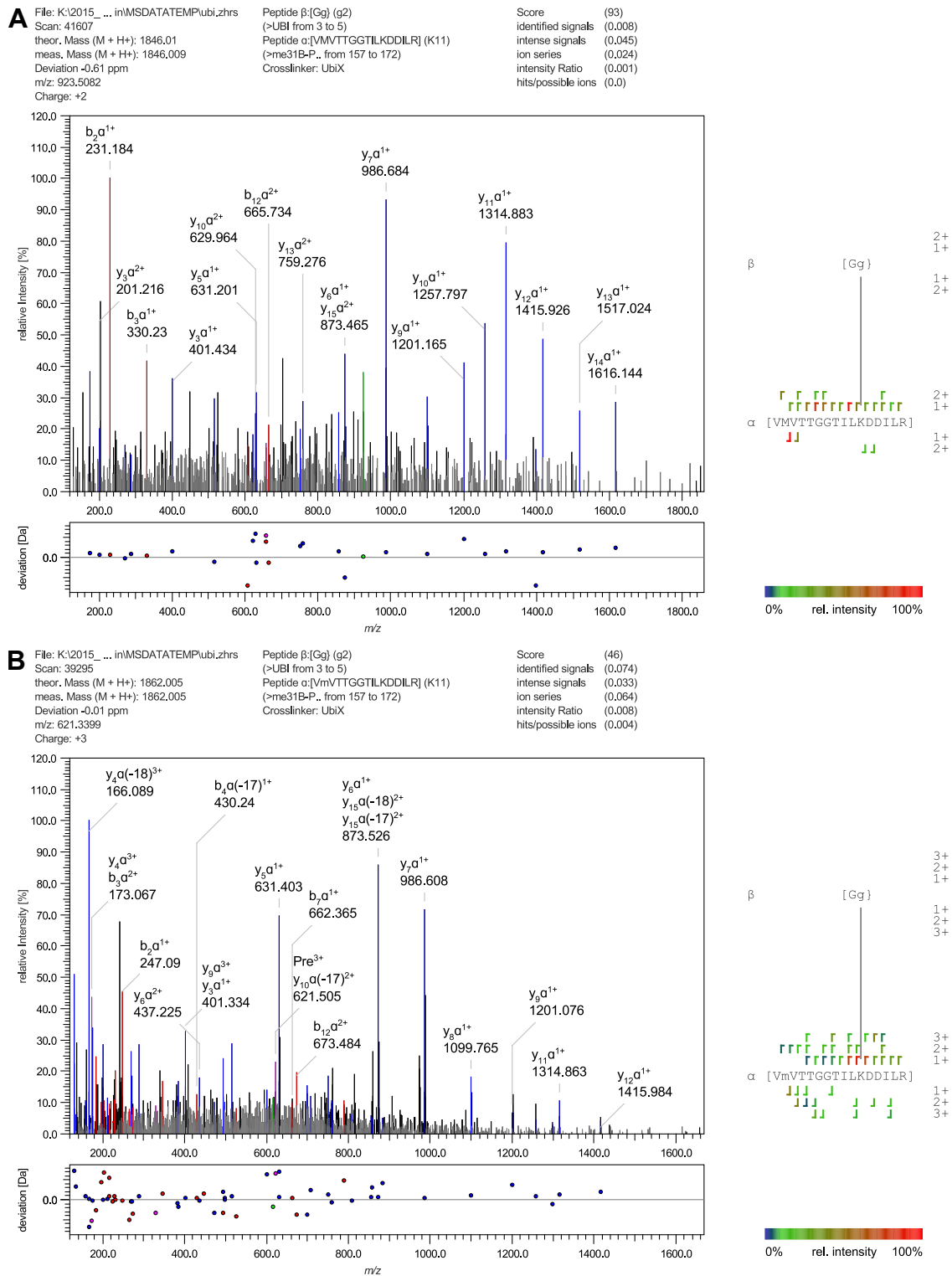


Abbildung 42: Spektren eines GG-modifizierten Peptids - Es sind zwei von StavroX annotierte Spektren eines Diglycin-modifizierten Peptides dargestellt. In (A) ist das Spektrum eines zweifach positiv geladenen Vorläuferions bei m/z 923,5082 gezeigt. In (B) ist das Spektrum eines dreifach positiv geladenen Vorläuferions bei m/z 621,3399 gezeigt. Das Peptid in (B) ist zusätzlich durch ein oxidiertes Methioin modifiziert (kleines 'm' in der Sequenz). Weitere Details zum Abbildungsinhalt können der Bildunterschrift von Abb. 41 entnommen werden.

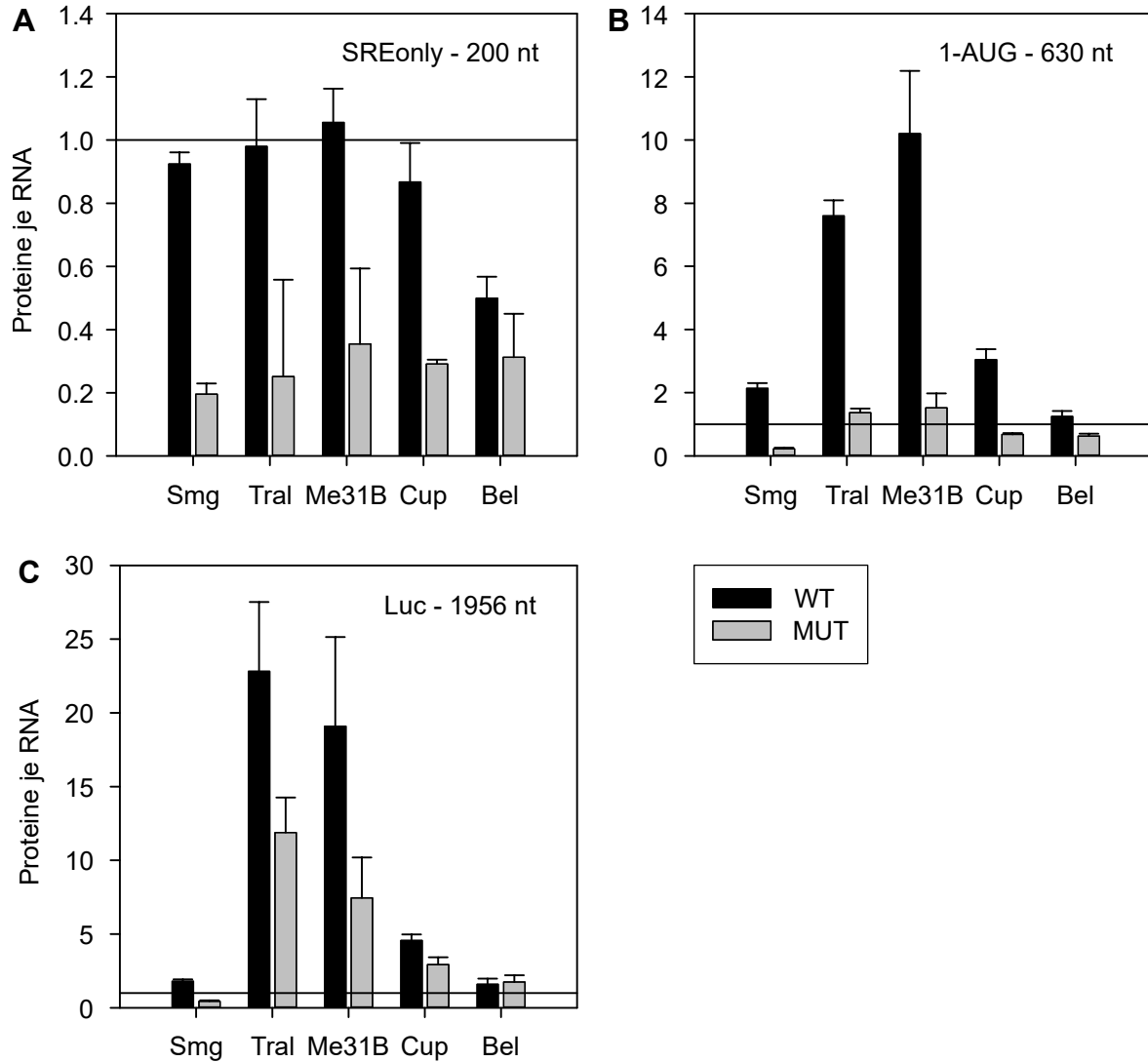


Abbildung 43: Längenabhängige Stöchiometrie der Komponenten im Komplex (unkorrigiert) - (A-C) Die Komponenten des Repressor-Komplexes wurden wie in Abb. 15 für drei verschieden lange RNAs quantifiziert. Die Proteinmengen wurden auf die eingesetzte RNA-Menge relativiert. Die Anzahl an gebundenen Proteinen je RNA wurde für drei verschiedene RNA-Längen dargestellt (A - SREonly, B - 1-AUG, C - Luc). Hier sind die absoluten Werte für die Bindung dargestellt, der korrigierte, spezifische Anteil der Bindung ist in Abb. 16 auf S. 39 dargestellt.

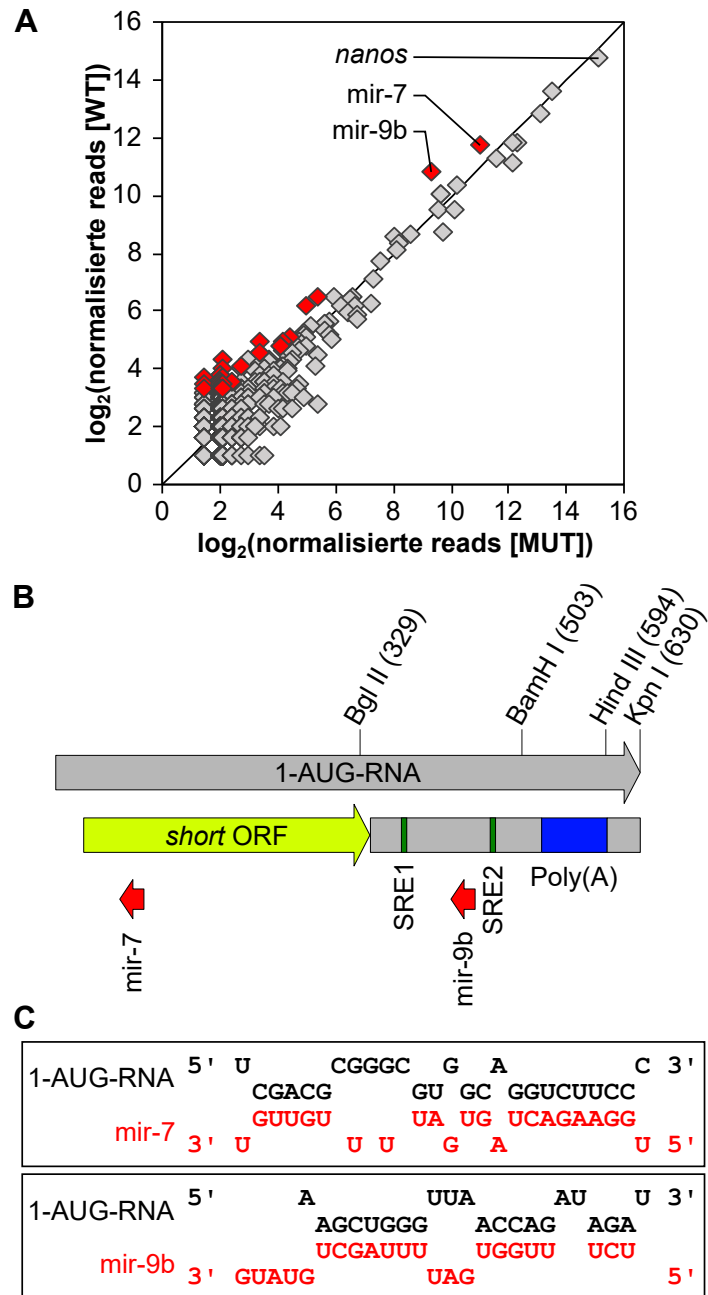


Abbildung 44: Deep-sequencing-Analyse des RNA-pull-downs - (A) Quantitativer Vergleich der *reads* für den WT- und MUT-pull-down. Die *reads* wurden normalisiert auf die Anzahl der kartierten *reads*. Alle RNAs mit mehr als 10 *reads* im WT und einer Anreicherung größer als 1,5 sind in Rot markiert. Zwei mikroRNAs (mir-7 und mir-9b) waren als einzige abundant und angereichert. (B) Schematische Darstellung der für den pull-down verwendeten 1-AUG-RNA. Die angedeuteten Restriktionsstellen können zur Synthese verschiedener Transkripte z.B. mit oder ohne SRE, bzw. mit oder ohne Poly(A)-Schwanz verwendet werden. Die vorhergesagten Bindestellen der miRNAs überlappen nicht mit den *stem-loops* der SREs im 3'-UTR. (C) Hybridisierungsstellen der miRNAs an die 1-AUG-RNA wurden mit RNAhybrid bestimmt (Krüger u. Rehmsmeier, 2006) Gepaarte Basen stehen nah beieinander, während ungepaarte Basen voneinander entfernt stehen.

Tabelle 19: Angereicherte Proteine in der Smaug-Immunpräzipitation - Quantitative Auswertung der massenspektrometrischen Untersuchung einer Smaug-Immunpräzipitation. Die Berechnung der Intensitäts- und p-Werte erfolgte wie für den RNA-*pull-down*, beschrieben in Abschnitt 2.3 auf S.27. Es sind alle angereicherten Proteine mit einem p-Wert unter 5% gelistet. Die mit * markierten p-Werte wurden abgeschätzt. Dazu wurde der Mittelwert aller Intensitätswerte (18,37) als Wert für eine unspezifische Bindung angenommen.

Protein	Intensität _{korrr}		p-Wert	Anreicherung
	Smaug-IP	Präimmun-IP		
Smaug	30,48	20,31	$3,81 \cdot 10^{11}$	1154,1
CG1440	26,96	n.d.	$*7,22 \cdot 10^6$	n.d.
Me31B	33,17	29,19	$1,33 \cdot 10^5$	15,7
CG4199	26,75	n.d.	$*1,40 \cdot 10^5$	n.d.
Not3	25,01	14,21	$1,44 \cdot 10^5$	1787,0
Ranbp21	24,52	13,76	$3,12 \cdot 10^5$	1740,4
CG11307	26,37	n.d.	$*4,18 \cdot 10^5$	n.d.
Hop	26,27	n.d.	$*5,51 \cdot 10^5$	n.d.
Twin	25,72	17,43	$7,06 \cdot 10^5$	312,7
Not1	26,07	n.d.	$*9,24 \cdot 10^5$	n.d.
Feo	30,19	25,55	$1,15 \cdot 10^4$	25,0
Cup	31,41	27,54	$2,09 \cdot 10^4$	14,6
Trailer Hitch	31,38	27,67	$3,39 \cdot 10^4$	13,1
CG11208	25,17	17,69	$3,49 \cdot 10^4$	178,0
Lola	25,26	19,11	0,00145	71,0
LManII	24,82	n.d.	*0,00151	n.d.
TBC1D5	24,65	n.d.	*0,00208	n.d.
Rudimentary	32,71	30,06	0,00216	6,3
CG9467	22,34	14,35	0,00277	254,2
eIF4E	29,31	25,72	0,00291	12,0
Brat	24,01	17,70	0,00325	79,3
Kri	24,17	18,02	0,00337	71,1
CstF-64	22,99	16,02	0,00372	125,9

Fortsetzung auf der nächsten Seite

Tabelle 19 – Fortsetzung

Protein	Intensität _{kor}		p-Wert	Anreicherung
	Smaug-IP	Präimmun-IP		
4E-T	22,45	15,37	0,00497	134,7
CG2941	24,09	18,35	0,00525	53,7
Sqd	23,87	18,03	0,00555	57,2
Nup358	28,47	24,97	0,00598	11,3
CG6767	24,88	19,86	0,00671	32,5
Glass	26,33	22,04	0,00675	19,5
Men	22,25	15,46	0,00693	110,3
Yps	28,16	24,68	0,00749	11,2
PABPC	31,20	28,73	0,00918	5,6
CstF-50	23,73	n.d.	*0,00939	n.d.
Gem3	22,88	17,07	0,0101	56,1
DCTN2-p50	28,02	24,72	0,0109	9,8
Imp	27,15	23,75	0,0145	10,5
Arp1	26,77	23,25	0,0147	11,5
CG8184	23,39	n.d.	*0,0149	n.d.
DCTN4-p62	21,48	15,30	0,0157	72,2
Slmb	22,81	17,56	0,0162	38,2
Vps16A	20,92	14,39	0,0164	92,2
CG3744	23,31	n.d.	*0,0167	n.d.
RanGAP	27,27	24,03	0,0169	9,5
Pop2	23,27	n.d.	*0,0175	n.d.
Klp3A	27,55	24,49	0,0194	8,3
CG8963	22,81	17,79	0,0194	32,5
CG31357	24,76	20,77	0,0210	15,9
CG6617	27,63	24,74	0,0241	7,4
Gek	23,02	n.d.	*0,0241	n.d.
Rga	22,99	n.d.	*0,0247	n.d.
Muskelin	27,36	24,44	0,0254	7,6

Fortsetzung auf der nächsten Seite

Tabelle 19 – Fortsetzung

Protein	Intensität _{korrr}		p-Wert	Anreicherung
	Smaug-IP	Präimmun-IP		
Crinkled	18,37	10,46	0,0255	240,6
Lark	25,52	22,02	0,0260	11,3
RpS13	23,47	19,15	0,0261	20,0
RpII215	20,32	14,43	0,0301	59,3
RpL27A	26,83	23,91	0,0310	7,6
CkIIalpha	26,71	23,78	0,0318	7,7
Rcd-1	22,77	n.d.	*0,0322	n.d.
Dcp1	23,46	19,38	0,0323	16,9
Mhcl	22,05	17,32	0,0327	26,5
AIMP2	21,41	16,35	0,0329	33,4
Rab1	28,45	26,03	0,0345	5,4
RanBPM	25,51	22,29	0,0352	9,3
CG6364-RA	22,51	18,16	0,0363	20,4
RpLP1	27,49	24,87	0,0368	6,1
Nap1	27,73	25,20	0,0376	5,8
RpL32	26,96	24,24	0,0384	6,6
Hrb27C	27,47	24,89	0,0385	6,0
CG5174	27,65	25,14	0,0395	5,7
SF2	23,68	19,97	0,0409	13,1
Pp1-87B	23,84	20,20	0,0411	12,5
DCTN3-p24	23,37	19,59	0,0426	13,7
MED4	27,16	24,65	0,0462	5,7
Cana	22,66	18,70	0,0464	15,5
Mahe	21,62	17,23	0,0471	21,0
CG31368	18,37	12,04	0,0471	80,7
RpL11	25,68	22,80	0,0477	7,4
SmD1	22,53	18,56	0,0480	15,6
CG2246	22,38	n.d.	*0,0487	n.d.

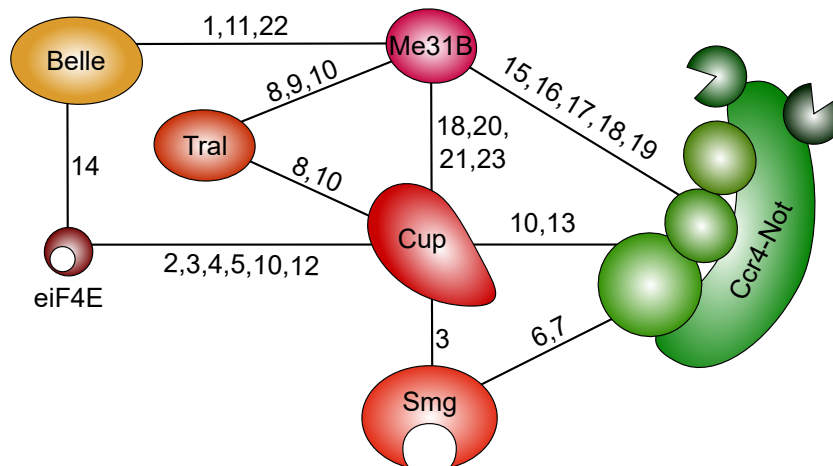


Abbildung 45: Beschriebene Interaktionen innerhalb des Repressor-Komplexes - Grafische Darstellung der beschriebenen Interaktionen zwischen den Proteinen des reprimierten mRNP. Die Ziffern entsprechen den Zitaten in Tab. 20.

Tabelle 20: Referenzen zu Interaktionen im reprimierten RNP - Zusammenfassung von Referenzen der in Abb. 45 dargestellten Interaktionen.

Nr.	Zitat	Nr.	Zitat
1	Tseng-Rogenski <i>et al.</i> (2003)	13	Kamenska <i>et al.</i> (2014)
2	Wilhelm <i>et al.</i> (2003)	14	Sharma u. Jankowsky (2014)
3	Nelson <i>et al.</i> (2004)	15	Chen <i>et al.</i> (2014b)
4	Zappavigna <i>et al.</i> (2004)	16	Rouya <i>et al.</i> (2014)
5	Nakamura <i>et al.</i> (2004)	17	Mathys <i>et al.</i> (2014)
6	Semotok <i>et al.</i> (2005)	18	Ozgur <i>et al.</i> (2015)
7	Zaessinger <i>et al.</i> (2006)	19	Waghray <i>et al.</i> (2015)
8	Tritschler <i>et al.</i> (2008)	20	Ozgur <i>et al.</i> (2015)
9	Tritschler <i>et al.</i> (2009)	21	Nishimura <i>et al.</i> (2015)
10	Igreja u. Izaurrealde (2011)	22	Bish <i>et al.</i> (2015)
11	Drummond <i>et al.</i> (2011)	23	Kamenska <i>et al.</i> (2016)
12	Kinkelin <i>et al.</i> (2012)		

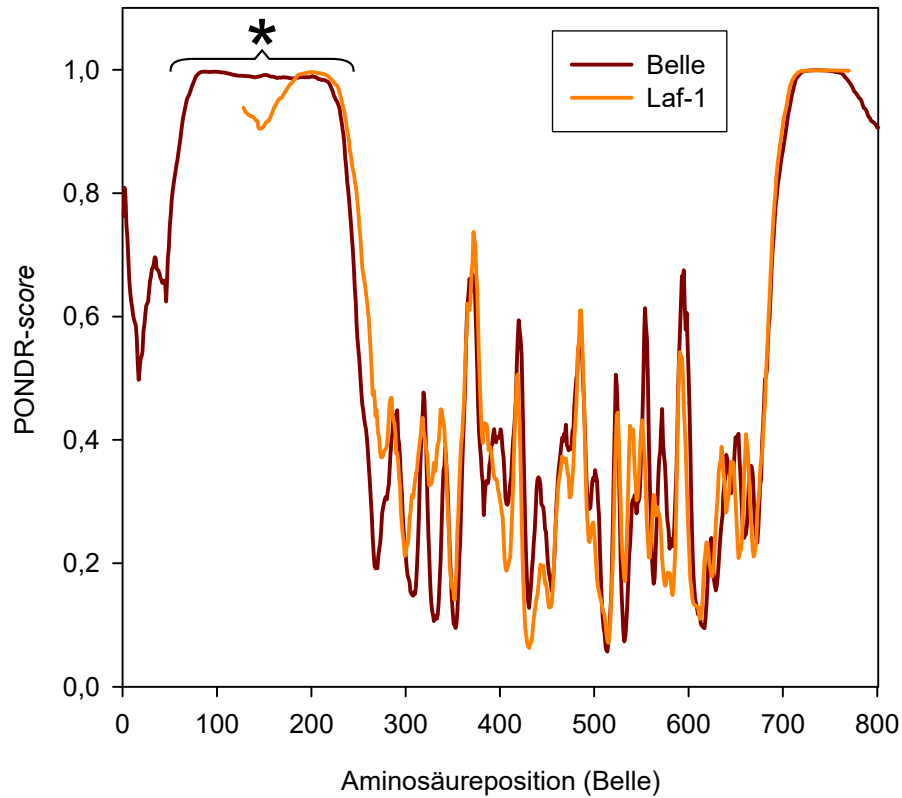


Abbildung 46: Intrinsisch unstrukturierte Bereiche im Belle-Protein - Der PONDR-Score für die DDX3-Homologe Belle (*Drosophila melanogaster*) und Laf-1 (*C. elegans*) wurde über den gesamten Sequenzbereich errechnet (Xue *et al.*, 2010). Je höher der Wert (im Bereich von 0 - 1), desto eher ist dieser Sequenzbereich intrinsisch unstrukturiert. Im N-Terminus von Belle ist ein deutlich längerer unstrukturierter Bereich zu erkennen (mit * markiert) als in Laf-1, dessen N-Terminus die Bildung von *liquid droplets* vermitteln kann (Elbaum-Garfinkle *et al.*, 2015). Die beiden PONDR-Diagramme für Laf-1 und Belle wurden an der DEAD-Sequenz im ATPase-Kern ausgerichtet und übereinander gelegt. Die Nummerierung der Aminosäuren erfolgte anhand der Sequenz von Belle.

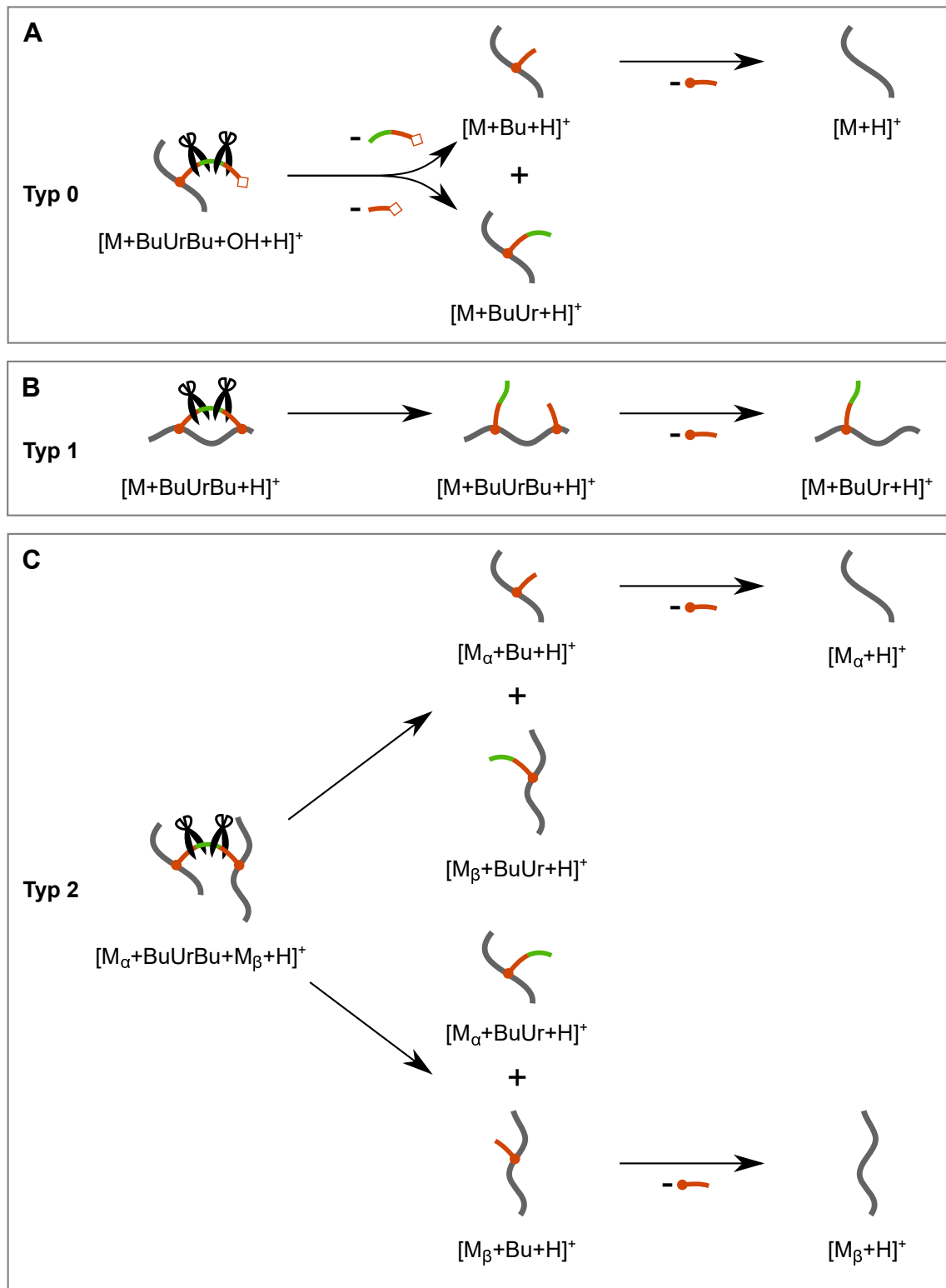


Abbildung 47: Mögliche Fragmentationen nach *Cross-Linker*-Fragmentierung
 - Darstellung möglicher Fragmentationen der verschiedenen *Cross-Link*-Typen nach der CID- oder HCD-Fragmentierung eines *Cross-Links* mit dem Harnstoff-*Linker*. (A) Typ 0 (*Dead-end-Cross-Links*). (B) Typ 1 (intra-peptidale *Cross-Links*). (C) Typ 2 (inter-peptidale *Cross-Links*)

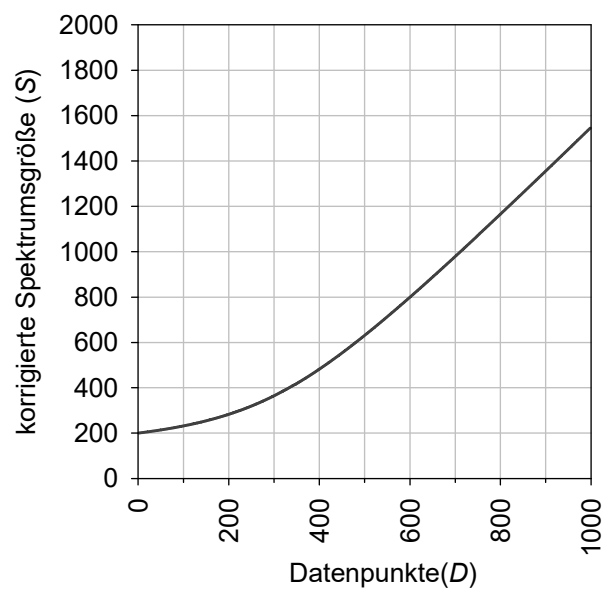


Abbildung 48: Korrektur der Spektrumsgröße - Spektren mit wenigen Signalen, können den errechneten *score* verfälschen, wenn einige der Signale zugeordnet werden können, da dann unter Umständen ein relativ großer Anteil aller Signale identifiziert wird. Gleichung 7 wird verwendet, um die Anzahl an Datenpunkten im Spektrum zu korrigieren. Die Abhängigkeit der korrigierten Spektrumsgröße von der Anzahl an Datenpunkten ist hier dargestellt.

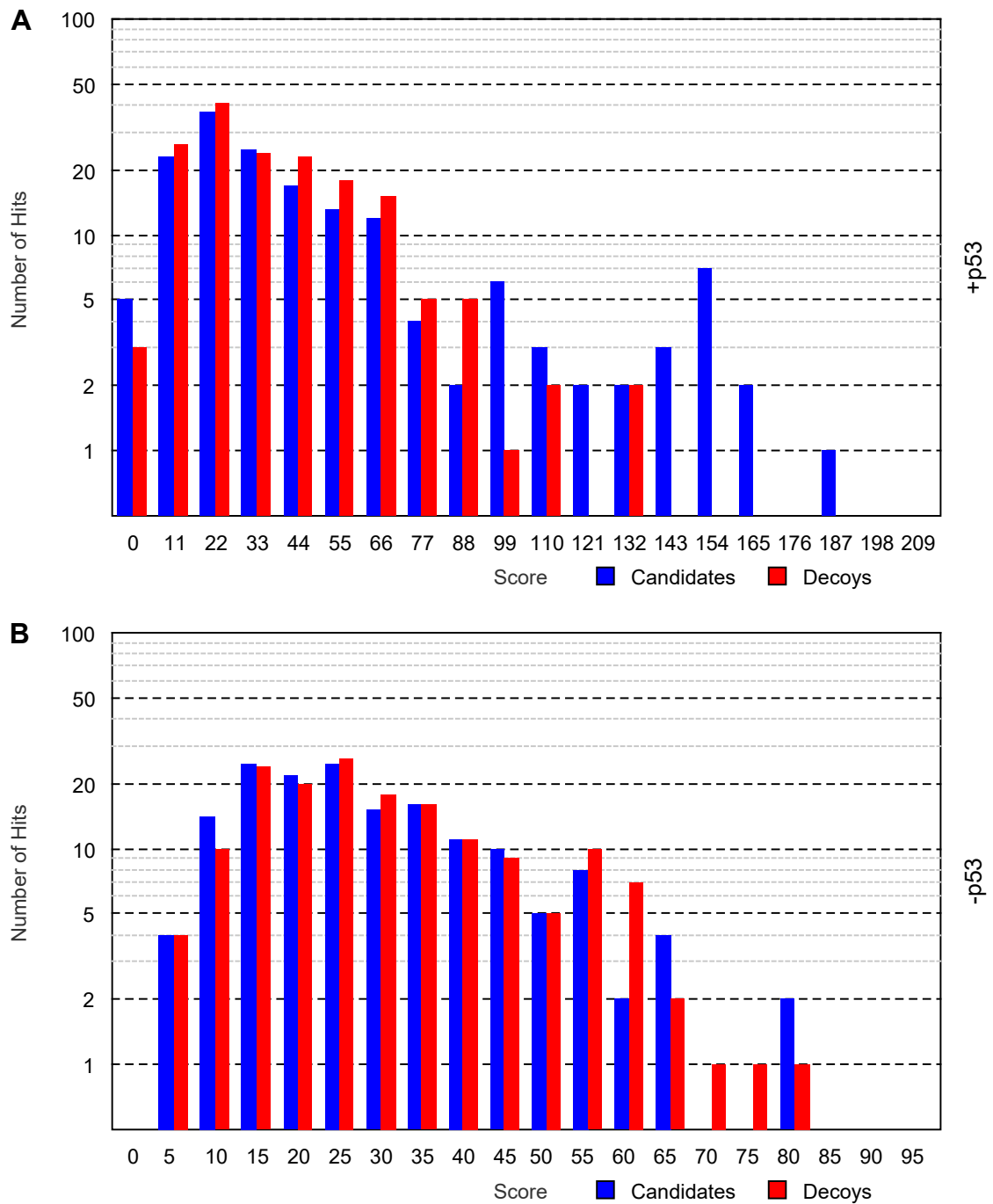


Abbildung 49: Screenshot der automatischen *Decoy*-Analyse - MeroX-Analyse eines *Cross-Link*-Experiments mit dem p53-Protein. Die verwendete Sequenzdatenbank enthielt 500 irrelevante *E. coli*-Proteinsequenzen, entweder mit der p53-Sequenz (+p53 in A) oder ohne die p53-Sequenz (-p53 in B). In (A) ist ein deutlicher Unterschied zwischen der Analyse der *Target*-Datenbank und der *Decoy*-Analyse zu erkennen, der in (B) nicht erkennbar ist. Dies ist ein deutlicher Hinweis darauf, dass tatsächlich kein *Cross-Link* in der Analyse (B) identifiziert wurde.

Publikationsliste

Originalarbeiten

GÖTZE, MICHAEL ; PETTELKAU, JENS ; SCHAKS, SABINE ; BOSSE, KONSTANZE ; IHLING, CHRISTIAN H. ; KRAUTH, FABIAN ; FRITZSCHE, ROMY ; KÜHN, UWE ; SINZ, ANDREA: StavroX – a software for analyzing crosslinked products in protein interaction studies. In: *J. Am. Soc. Mass Spectrom.* 23 (2012), Nr. 1, S. 76–87

FRITZSCHE, ROMY ; IHLING, CHRISTIAN H. ; GÖTZE, MICHAEL ; SINZ, ANDREA: Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. In: *Rap. Commun. Mass Spectrom.* 26 (2012), Nr. 6, S. 653–658

GÖTZE, MICHAEL ; WAHLE, ELMAR: Smaug destroys a huge treasure. In: *Genome Biol.* 15 (2014) Nr. 1, S. 101

GÖTZE, MICHAEL ; PETTELKAU, JENS ; FRITZSCHE, ROMY ; IHLING, CHRISTIAN H. ; SCHÄFER, MATHIAS ; SINZ, ANDREA: Automated Assignment of MS/MS Cleavable Cross-Links in Protein 3D-Structure Analysis. In: *J. Am. Soc. Mass Spectrom.* 26 (2015), Nr. 1, S. 83–97

HARNISCH, CHRISTIANE ; CUZIC-FELTENS, SIMONA ; DOHM, JULIANE C. ; GÖTZE, MICHAEL ; HIMMELBAUER, HEINZ ; WAHLE, ELMAR: Oligoadenylation of 3' decay intermediates promotes cytoplasmic mRNA degradation in Drosophila cells. In: *RNA* 22 (2016), Nr. 3, S. 428–442

HAGE, CHRISTOPH ; IHLING, CHRISTIAN H. ; GÖTZE, MICHAEL ; SCHÄFER, MATHIAS ; SINZ, ANDREA: Dissociation Behavior of a TEMPO-Active Ester Cross-Linker for Peptide Structure Analysis by Free Radical Initiated Peptide Sequencing (FRIPS) in Negative ESI-MS. In: *J. Am. Soc. Mass Spectrom.* (2016), S. 1–13

ARLT, CHRISTIAN ; GÖTZE, MICHAEL ; IHLING, CHRISTIAN H. ; HAGE, CHRISTOPH ; SCHÄFER, MATHIAS ; SINZ, ANDREA: Integrated Workflow for Structural Proteomics Studies Based on Cross-Linking/Mass Spectrometry with an MS/MS Cleavable Cross-Linker. In: *Anal. Chem.* 88 (2016), Nr. 16, S. 7930–7937

Zur Publikation eingereicht:

GÖTZE, MICHAEL ; DUFOUT, JÉRÉMY ; RAMMELT, CHRISTIANE ; PIERSON, STÉPHANIE ; NAGRAJ, SAMBRANI ; TEMME, CLAUDIA ; SINZ, ANDREA ; SIMONELIG, MARTINE ; WAHLE, ELMAR: Translational repression of the Drosophila nanos mRNA involves the RNA helicase Belle and RNA coating by Me31B and Trailer hitch. Eingereicht in: *eLife*

Vorträge

GÖTZE, MICHAEL: Rapid Analysis of Cross Linking-MS Experiments using MeroX. *3rd Symposium on Structural Proteomics* (21.-22. November 2013) in Prag (CZE)

SINZ, ANDREA ; GÖTZE, MICHAEL: Workshop: Chances and Pitfalls of Chemical Cross-Linking/Mass Spectrometry for structural proteomics. *13th Austrian Proteomic Research Symposium* (26.-28. August 2015) in Klosterneuburg (AUT)

GÖTZE, MICHAEL: Efficient Identification of Cross-Linked Peptides with StavroX and MeroX in Structural Proteomics. *49. Jahrestagung der DGMS* (28. Februar - 2. März 2016) in Hamburg (GER)

Poster-Präsentationen

GÖTZE, MICHAEL ; MORITZ, BODO ; WAHLE, ELMAR *In vitro* characterization of Me31B, a DEAD-box-helicase involved in translational repression. *International Meeting GRK1591* (15.-17. März 2012) in Halle (GER)

GÖTZE, MICHAEL ; PETTELKAU, JENS ; SCHAKS, SABINE ; BOSSE, KONSTANZE ; IHLING, CHRISTIAN H. ; KRAUTH, FABIAN ; FRITZSCHE ROMY ; KÜHN, UWE ; SINZ, ANDREA: StavroX – a software for analyzing crosslinked products in protein interaction studies. *19th International Mass Spectrometry Conference* (15.-21. September 2012) in Kyoto (JP)

GÖTZE, MICHAEL ; ARLT, CHRISTIAN ; PETTELKAU, JENS ; IHLING, CHRISTIAN H. ; SCHÄFER, MATHIAS ; SINZ, ANDREA: Efficient identification of cross-linked peptides with StavroX and MeroX in structural proteomics. *14th Human Proteome Organization World Congress* (27.-30. September 2015) in Vancouver (CAN)

GÖTZE, MICHAEL ; ARLT, CHRISTIAN ; PETTELKAU, JENS ; IHLING, CHRISTIAN H. ; SCHÄFER, MATHIAS ; SINZ, ANDREA: Efficient identification of cross-linked peptides with StavroX and MeroX in structural proteomics. *5th Symposium on Structural Proteomics* (19.-20. November 2015) in Halle (GER)

Lebenslauf

Persönliche Daten

Name: Michael Götze
Geschlecht: männlich
Geburtstag und -ort: 3. Mai 1987 in Zwenkau
Staatsangehörigkeit: deutsch
Familienstand: ledig
Adresse: Brandenburger Straße 4, 06114 Halle (Saale)

Ausbildung

1993-1997 **Grundschule**, 60. Grundschule Leipzig
1997-2005 **Gymnasium, Abitur** Johannes-Kepler-Gymnasium Leipzig
2005-2010 **Diplomstudium-Biochemie**, MLU Halle-Wittenberg
2010 **Diplom in Biochemie**, MLU Halle-Wittenberg, Institut für Biochemie und Biotechnologie, Allgemeine Biochemie (Prof. Dr. Elmar Wahle), zum Thema: „Untersuchung der Wechselwirkungen des 3'-prozessierenden Apparates“
2011-2016 **Wissenschaftlicher Mitarbeiter, Doktorand**, MLU Halle-Wittenberg, Institut für Biochemie und Biotechnologie, Allgemeine Biochemie (Prof. Dr. Elmar Wahle),

Halle (Saale), 19. April 2017

Michael Götze

Selbstständigkeitserklärung

Ich versichere hiermit, die vorliegende Dissertation selbstständig und ohne unerlaubte fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die im Literaturverzeichnis angeführten Quellen genutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet. Die vorgelegte Arbeit ist weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt worden. Ich habe keine früheren erfolglosen Promotionsversuche unternommen

Halle (Saale), 19. April 2017

Michael Götze

Danksagung

So vielen Menschen gebührt Dank, da sie mich während meiner Promotion unterstützt haben. Ein großer Dank gilt meinem Doktorvater, Elmar Wahle. Ich habe in den letzten Jahren viel von ihm gelernt und er hat mit vielen Vorschlägen für Experimente und konstruktiven Diskussionen maßgeblich zur Anfertigung dieser Promotionsarbeit beigetragen. Er hat mich stets unterstützt und mir den Besuch zahlreicher Konferenzen ermöglicht. Die Unterstützung z.B. bei der Entwicklung von StavroX und MeroX habe ich dabei nie als selbstverständlich empfunden. Vielen Dank, Elmar!

Bei Andrea Sinz möchte ich mich einerseits für die Übernahme des Zweitgutachtens und andererseits für die großartige Kooperation bedanken, die seit mehr als sechs Jahren besteht und hoffentlich auch weiterhin bestehen wird. Sie hat den zweiten Teil dieser Arbeit erst möglich gemacht und mit zahlreichen Diskussionen und Vorschlägen an der Verbesserung der hier vorgestellten Programme mitgewirkt. Ich bin dankbar dafür, dass sie mich stets dazu motivierte, meine Daten bei verschiedenen internationalen Konferenzen zu präsentieren. Ich bin ihr zudem dankbar für die vielen Messungen, die ich an den Massenspektrometern ihrer Gruppe durchführen durfte.

Ich möchte mich sehr herzlich bei Prof. Bernhard Spengler für die Übernahme des Drittgutachtens bedanken.

Ich konnte während meiner Promotion sehr viele gute wissenschaftliche Vorträge hören, was ich vor allem der Forschergruppe 855 und dem Graduiertenkolleg 1591 verdanke. Ohne die Finanzierung durch die DFG hätte ich weder diese Promotion anfertigen, noch internationale Tagungen besuchen können.

Rebekka hat mich in den letzten vier Jahren stets begleitet und meine Zeit in Halle um einiges reicher gemacht. Ein großer Dank gilt ihr auch für die Erstellung der Logos für StavroX und MeroX. Außerdem hat ihr rechtschreiberisches Gespür noch zahlreiche Fehler wie falsche Plurali in meiner Arbeit aufdecken können, danke dafür. Allen anderen, Elmar, Christiane, Christian T., Christian I., Christoph und Rico, die den Inhalt dieser Dissertation mit mir diskutierten und meine Arbeit auf Fehler durchforstet haben, gilt der gleiche Dank.

Vielen Dank an alle Mitglieder der beiden Arbeitsgruppen Wahle und Sinz! Wenn wir auch viele gute wissenschaftliche Diskussionen hatten, sind es vor allem die Mittagspausen, Kaffeepausen und Kuchen, die mir meine Promotion „versüßt“ haben! Den Mitgliedern der Gruppe Sinz danke ich vor allem für die zahlreichen *Bug-reports*! Ich hätte nie gedacht, dass man so viele Fehler in ein Programm einbauen kann. Ich möchte euch allen mit der finalen Abbildung dieser Arbeit danken:

