



# Strukturbiologische Untersuchungen zur Chaperone-vermittelten Zusammenlagerung spleißosomaler U-snRNPs

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades  
der  
Julius-Maximilians-Universität Würzburg

vorgelegt von

Diplom-Biologe  
**Jann-Patrick Pelz**

aus Salzgitter

Würzburg 2015



Eingereicht bei der Fakultät für Chemie und Pharmazie am

\_\_\_\_\_

Gutachter der schriftlichen Arbeit

1. Gutachter: \_\_\_\_\_

2. Gutachter: \_\_\_\_\_

Prüfer des öffentlichen Promotionskolloquiums

1. Prüfer: \_\_\_\_\_

2. Prüfer: \_\_\_\_\_

3. Prüfer: \_\_\_\_\_

Datum des öffentlichen Promotionskolloquiums

\_\_\_\_\_

Doktorurkunde ausgehändigt am

\_\_\_\_\_



Diese Arbeit wurde angefertigt in der Arbeitsgruppe von Prof. Dr. Utz Fischer, am Lehrstuhl für Biochemie der Julius-Maximilians-Universität Würzburg

## Erklärung

Hiermit erkläre ich an Eides statt, dass ich die Dissertation selbstständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Ich erkläre außerdem, dass diese Dissertation weder in gleicher oder anderer Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ich habe früher außer den mit dem Zulassungsgesuch urkundlich vorgelegten Graden keine weiteren akademischen Grade erworben oder zu erwerben versucht.

Würzburg, April 2015

---

Jann-Patrick Pelz

„Die Befriedigung der eigenen Neugier ist eine der größten

Quellen von Glück im Leben“

- Linus Pauling (1901-1994)

# Inhalt

1. Zusammenfassung.....	1
2. Summary.....	3
3. Einleitung.....	5
3.1 Prolog - Die zelluläre Entstehung makromolekularer Komplexe .....	5
3.2 Der Spleißprozess .....	6
3.3 Die Rolle der U-snRNPs während dem Spleißprozess .....	6
3.4 Die Protein-Komponenten der U-snRNPs .....	8
3.5 Die Biogenese der U-snRNPs der Sm-Klasse .....	11
3.6 Assistenzfaktoren während der cytosolischen Phase der U-snRNP-Biogenese .....	12
3.7 Die frühe Phase der Zusammenlagerung ist vom PRMT5-Komplex abhängig .....	13
3.8 Die späte Phase der Zusammenlagerung und die Rolle des SMN-Komplexes.....	14
3.9 Die Beteiligung individueller Komponenten des SMN-Komplex an der Sm-Core-Zusammenlagerung.....	15
4. Material & Methoden.....	17
4.1 Chemikalien .....	17
4.2 Materialien .....	17
Geräte .....	17
Verbrauchsmaterialien .....	18
Chromatographie – Materialien & vorgepackte Säulen.....	18
Medien für Bakterienkultur.....	19
Plasmide und Bakterienstämme .....	20
Enzyme .....	21
4.3 Methoden.....	21
Agarosegelelektrophorese .....	21
Heterologe Expression von rekombinanten Proteinen.....	22
Molekularbiologische und bioanalytische Methoden.....	24
Chromatographische Methoden zur Präparation von rekombinanten Proteinen .....	28
Protokolle zur Präparation der 6S- und 8S-Komplexuntereinheiten .....	31
Präparation des 6S- und 8S-Komplexes.....	34
Röntgen-Strukturaufklärung von Proteinen.....	35
Elektronenmikroskopie zur Strukturuntersuchung des 8S-Komplexes.....	40
<i>In silico</i> Methoden zur Untersuchung der Moleküldynamiken .....	41
5. Ergebnisse.....	43

Biologischer Hintergrund.....	43
5.1    Strukturanalyse der 6S- und 8S-Intermediate der Sm-Core-Zusammenlagerung .....	43
Die Kristallstruktur des 6S-Komplexes.....	43
Die Sm-Proteine innerhalb des 6S-Komplexes .....	44
Die Struktur von pICln im 6S-Komplex .....	47
Die Interaktionen von pICln mit den benachbarten Sm-Proteinen SmD1 und SmG .....	50
Physikochemische Eigenschaften von pICln im Kontext der Sm-Proteinbindung.....	52
Die Kristallstruktur des 8S-Komplexes.....	55
Die Strukturen von Gemin2 und SMNΔC in der Kristallstruktur des 8S-Komplexes .....	55
Ergänzende Strukturanalysen des 8S-Komplexes auf Basis der Elektronenmikroskopie .....	57
Vergleich der Strukturen der 6S- und 8S-Komplexe mit der Struktur des 7S-Komplexes.....	58
Einblicke in die Moleküldynamik des 8S-Komplexes.....	60
5.2    Strategie zur Lösung der Kristallstrukturen des 6S- und 8S-Komplexes .....	64
Präparation und Rekonstitution des 6S- und 8S-Komplexes.....	64
Kristallisation und Strukturlösung .....	66
6.    Diskussion.....	74
6.1    Zelluläre Zusammenlagerungsprozesse benötigen häufig spezifische Assistenzfaktoren....	74
6.2    pICln agiert als Chaperon bei der Sm-Core-Zusammenlagerung und zeigt eine molekulare Mimikry der Sm-Proteinfaltung.....	75
6.3    pICln agiert als topologischer Organisator des Sm-Core.....	77
6.4    Die elektrostatischen Eigenschaften von pICln ermöglichen eine stabile Rekrutierung der Sm-Proteine zum 6S-Komplex .....	79
6.5    Strukturelle Ursachen für die Inhibition der Interaktion zwischen snRNA und Sm-Proteinen im 6S-Komplex.....	80
6.6    Die Sm-Proteine werden im 8S-Komplex durch Gemin2 kontaktiert .....	81
6.7    Eine putative Sollbruchstelle im Sm/pICln-Ring des 8S-Komplexes könnte die Dissoziation von pICln einleiten.....	82
6.8    Die Funktion des SMN-Komplexes bei der Dissoziation von pICln legt eine Rolle der SMN-Tudor-Domäne bei diesem Vorgang nahe .....	83
6.9    Warum benötigt die Zelle ein solch komplexes System von Assistenzfaktoren? .....	84
7.    Ausblick.....	86
8.    Referenzen .....	88
9.    Anhang.....	98
9.1    Nucleoside und Nucleinbasen.....	98
9.2    Aminosäuren .....	98
9.3    Publikationen.....	98

9.4	Primärsequenz-Alignment verschiedener pICln-Orthologe .....	99
9.5	Lebenslauf .....	100
9.6	Danksagung .....	101

# 1. Zusammenfassung

Durch die Spleißreaktion werden nicht-kodierende Sequenzelemente (Introns) aus eukaryotischen Vorläufer-mRNAs entfernt und die kodierenden Sequenzelemente (Exons) miteinander zu einem offenen Leserahmen verbunden. Dieser zentrale Prozessierungsschritt während der eukaryotischen Genexpression wird durch das Spleißosom katalysiert, das aus den vier kleinen nukleären Ribonucleoproteinpartikeln (snRNPs) U1, U2, U4/U6 und U5, sowie einer Vielzahl weiterer Proteinfaktoren gebildet wird. Alle snRNPs besitzen eine gemeinsame ringförmige Kernstruktur, die aus sieben gemeinsamen Sm-Proteinen (SmB/B'-D1-D2-D3-E-F-G) besteht, die ein einzelsträngiges Sequenzmotiv auf der snRNAs binden. Während sich diese, als Sm-Core-Domäne bezeichnete Struktur *in vitro* spontan ausbilden kann, erfolgt die Zusammenlagerung *in vivo* in einem assistierten und hochregulierten Prozess. Dieser ist abhängig von insgesamt mindestens 12 trans-agierenden Faktoren, die in den PRMT5- und SMN-Komplexen organisiert sind. Der PRMT5-Komplex agiert in der frühen Phase der Zusammenlagerung, indem er die Sm-Proteine durch die Untereinheit pICln rekrutiert und die symmetrische Methylierung von Argininresten in den C-terminalen Schwänzen von SmB/B', SmD1 und SmD3 katalysiert.

Als Resultat dieser frühen Phase befinden sich die Sm-Proteine SmD1-D2-E-F-G und SmB/B'-D3 in zwei getrennten und durch pICln organisierten Komplexen. Während SmB/B'-D3-pICln am PRMT5-Komplex gebunden bleibt, existiert der zweite Komplex als freies Intermediat mit einem Sedimentationskoeffizienten von 6S. Diese Intermediate können nicht mit RNA assoziieren, sodass für die Fortsetzung des Zusammenlagerungsprozesses die Interaktion der Sm-Proteine mit pICln aufgelöst werden muss. Dies geschieht in der späten Phase der Sm-Core-Zusammenlagerung, in der die Sm-Proteine vom SMN-Komplex (bestehend aus SMN, Gemin2-8 und unrip) übernommen werden und pICln dissoziiert wird. Dadurch werden die Sm-Proteine für ihre Interaktion mit der snRNA aktiviert und können auf die Sm-Bindestelle transferiert werden, wodurch die Formierung des Sm-Core abgeschlossen wird.

Im Rahmen dieser Arbeit konnten mit Hilfe einer Kombination röntgenkristallographischer und elektronenmikroskopischer Methoden zwei wichtige Intermediate dieses Zusammenlagerungsprozesses strukturell charakterisiert werden. Bei diesen Intermediaten handelt es sich um den 6S-Komplex, sowie um ein Sm-Protein-Transferintermediat mit einem Sedimentationskoeffizienten von 8S. In diesem ist der 6S-Komplex an zwei zentrale Untereinheiten des SMN-Komplexes (SMN und Gemin2) gebunden, während pICln den Komplex noch nicht verlassen hat. Der 8S-Komplex stellt daher ein „gefangenes“ Intermediat zwischen der frühen und späten Phase der Zusammenlagerung dar.

Zunächst gelang es eine erste Kristallform des rekombinant hergestellten 8S-Komplexes zu erhalten, die jedoch keine Strukturlösung erlaubte. Durch eine kombinierte Optimierung der Kristallisationsbedingung und der verwendeten Proteine wurde eine weitere ähnliche Kristallform erhalten, mit der die Kristallstruktur des 8S-Komplexes gelöst werden konnte. Die Kristallisation des 6S-Komplexes gelang im Anschluss auf Basis der Hypothese, dass Kristalle beider Komplexe aufgrund der kompositionellen Verwandtschaft zwischen 6S und 8S auch Ähnlichkeiten in der Architektur ihrer Kristallgitter aufweisen könnten. Daher wurden innerhalb von pICln gezielt Aminosäuren substituiert, die sich innerhalb von Kristallkontakten der 8S-Kristalle befanden und konformationell eingeschränkt waren. Mit entsprechend rekonstituierten 6S-Präparationen konnten dann zwei Kristallformen erzeugt werden, die eine Strukturlösung des 6S-Komplexes ermöglichten.

Durch die Kristallstruktur des 6S-Komplexes konnte für pICln eine strukturelle Mimikry der Sm-Proteine identifiziert werden. Diese ermöglicht eine Bindung der Sm-Proteine und eine frühzeitige topologische Organisation des Sm-Pentamers D1-D2-F-E-G in einer geschlossenen hexameren Ringstruktur. Die Kristallstruktur des 8S-Komplexes zeigt, wie der SMN-Komplex über Gemin2 an das Sm-Pentamer bindet. In Kombination mit einer EM-Struktur des 8S-Komplexes gelang es weiterhin, einen plausiblen Mechanismus für die Elimination von pICln und die Aktivierung der Sm-Proteine für die snRNA-Bindung zu formulieren. Somit konnten diese Arbeiten zu einem besseren Verständnis der Funktionen von trans-agierenden Faktoren bei Zusammenlagerung von RNA-Protein-Komplexen *in vivo* beitragen.

## 2. Summary

Splicing is the process in which non-coding sequence elements (introns) are removed from eukaryotic pre-mRNAs and coding sequence elements (exons) are linked to an open reading frame. This central step in eukaryotic gene expression is catalyzed by the spliceosome, which is composed of the four small nuclear Ribonucleoproteins (snRNPs) U1, U2, U4/U6, U5 and a large number of additional protein factors. The snRNPs possess a common ring-shaped core structure that is formed by the seven Sm proteins (SmB/B'-D1-D2-D3-E-F-G) around a single-stranded sequence (Sm site) of the snRNAs. While this so-called Sm core domain forms spontaneously *in vitro*, its assembly is a highly regulated and assisted process *in vivo*. It is dependent on the action of at least 12 trans-acting factors which are organized in the PRMT5 and SMN complexes. The PRMT5 is active in the early phase of assembly and recruits the Sm proteins via its pICln subunit and catalyzes the symmetrical di-methylation of arginine residues in the C-terminal tails of SmB/B', SmD1 and SmD3.

As a result of the early phase the Sm proteins SmD1-D2-E-F-G and SmB/B'-D3 are organized by pICln in two distinct complexes. While SmB/B'-D3 remains bound to the PRMT5 complex, the second complex exists as a free intermediate with a sedimentation coefficient of 6S. These intermediates cannot associate with RNA and the interaction of the Sm proteins with pICln has to be resolved for the assembly process to be continued. This happens in the late phase of Sm core assembly in which the Sm proteins are taken over by the SMN complex and pICln is dissociated. Afterwards the Sm proteins can be transferred onto the Sm site of the snRNA and the Sm core is formed.

As part of this thesis two key intermediates of this assembly process could structurally be characterized by a combination of crystallographic and electron microscopic methods. These intermediates comprise the 6S complex and an Sm protein transfer-intermediate with a sedimentation coefficient of 8S. In this 8S complex the 6S complex is bound to two central subunits of the SMN complex (SMN and Gemin2) while pICln is still associated with the Sm proteins. Hence, this complex represents a trapped intermediate between the early and late phase of assembly.

In the beginning a first crystal form of a recombinantly prepared 8S complex was obtained that did not allow the solution of the structure. By a combined optimization of the crystallization condition and the proteins a further similar crystal form was obtained that allowed for the solution of the 8S crystal structure. The crystallization of the 6S complex could successfully be accomplished based on the hypothesis that the lattices of crystals of both complexes might show an architectural similarity because of the similar composition of the complexes. Hence, amino acids of pICln that were conformationally restricted within crystal contacts of the 8S crystals were targeted for substitution to

alanine. 6S preparations reconstituted with these proteins yielded two new crystal forms that allowed for the structure solution of the 6S complex.

Based on the crystal structure of the 6S complex a structural mimicry of Sm proteins by pICln was revealed. This enables binding of the Sm proteins by pICln which is the basis for an early topological organisation of the Sm Pentamer D1-D2-F-E-G within a closed hexameric ring structure. The crystal structure of the 8S complex revealed how the SMN complex binds to the Sm Pentamer via its Gemin2 subunit. In combination with an EM structure of the 8S complex both structures revealed a plausible mechanism for the elimination of pICln and the activation of the Sm proteins for snRNA binding. The solution of both structures helps to better understand the function of trans-acting factors during the *in vivo* assembly of RNA-protein complexes.

## 3. Einleitung

### 3.1 Prolog - Die zelluläre Entstehung makromolekularer Komplexe

Makromolekulare Komplexe entstehen durch die Bindung verschiedener Proteine miteinander, oder durch deren Interaktionen mit Nukleinsäuren. Für ihre korrekte Ausbildung ist es notwendig, dass die Komponenten zur selben Zeit und in entsprechender Stöchiometrie am richtigen Ort verfügbar sind. Da die Begegnung von Untereinheiten in erster Linie diffusionsgesteuert ist, wird sie durch die außergewöhnlich hohen Konzentrationen an Makromolekülen in der Zelle erschwert (Ellis, 2001a, b; Zimmerman and Minton, 1993). Diese nehmen zirka 30 Prozent des Zellvolumens ein, was ein Grund dafür ist, dass sich Zellsuspensionen nicht wie ideale Lösungen verhalten. Die Makromoleküle können sich aufgrund ihres hohen Molekulargewichts nur in einem begrenzten Ausschlussvolumen aufhalten, was neben der herabgesetzten Diffusionsrate auch die Effektivität ihrer Zusammenlagerung erhöht (Hall and Minton, 2003). Dies birgt die Gefahr der Ausbildung unspezifischer Interaktionen, die zur Aggregation oder Präzipitation der betroffenen Makromoleküle führen kann. Dieses physikalische Phänomen wird als makromolekulares Gedränge bezeichnet und bedingte die Evolution verschiedener zellulärer Strategien bei der Synthese und Zusammenlagerung makromolekularer Komplexe. So findet die Konstruktion komplexer zellulärer Maschinen, wie dem Ribosom, an speziellen Orten wie den Nukleoli statt (Cmarko et al., 2008). Dadurch ist die lokale Konzentration der Untereinheiten erhöht und die Begegnung kognater Bindungspartner wird wahrscheinlicher. Auch die Faltung der einzelnen Komponenten in ihre funktionale dreidimensionale Struktur, sowie die Ausbildung korrekter Interaktionen können assistiert sein. Hierbei kommen zum Teil Assistenzfaktoren zur Anwendung, die als molekulare *Chaperone* (Anstandsdamen) bezeichnet werden. Deren Aufgabe ist es inkorrekte Interaktionen zwischen Makromolekülen zu verhindern, oder solche sekundär wieder aufzulösen. Heute sind besonders faltungsassistierende Chaperone bekannt, welche meist unter Verwendung von ATP und durch die Bereitstellung abgegrenzter Faltungsräume die Ausbildung der funktionalen Konformation ihrer Substrate unterstützen. Neben diesen existieren außerdem sogenannte *Assembly*-Chaperone, die direkt bei der korrekten zellulären Zusammensetzung makromolekularer Komplexe assistieren (Ellis, 2006).

In den letzten Jahren konnte eine Vielzahl trans-agierender Faktoren identifiziert werden, die die Zusammenlagerung der RNA-Protein-Partikel (RNPs) des prä-mRNA-prozessierenden Systems der Eukaryonten steuern. Diese stellen daher ideale Modell-Substrate für die Analyse eines solchen assistierten Prozesses dar, um generelle Prinzipien für die Konstruktion zellulärer Maschinen ableiten zu können. Durch die im Folgenden beschriebenen Arbeiten konnten erste detaillierte Einblicke in die strukturellen Mechanismen eines solchen Vorgangs erhalten werden.

## 3.2 Der Spleißprozess

Die Struktur der meisten proteinkodierenden Gene in Eukaryonten ist dadurch gekennzeichnet, dass die kodierenden Sequenzen (Exons) von nicht-kodierenden intronischen Bereichen (Introns) unterbrochen sind (Gilbert, 1978). Diese Exon-Intron-Struktur wird bei der Transkription zunächst in einem primären Transkript übernommen. In einem komplexen Prozessierungsvorgang, dem sogenannten prä-mRNA Spleißen werden co- bzw. post-transkriptionell die Introns herausgeschnitten und die Exons miteinander ligiert. Dieser, durch das Spleißosom katalysierte Prozess, sichert die Generierung der reifen mRNA und damit des offenen, proteincodierenden Leserahmens. Hierbei erfolgt die Erkennung der Introns durch das Spleißosom auf Basis definierter Bereiche auf der prä-mRNA, die unter anderem durch kurze Konsensus-Sequenzen definiert sind. Im Wesentlichen sind dies die 5'-Spleißstelle (SS) mit einem invarianten GU-Dinukleotid, die 3'-SS mit einem AG-Dinukleotid, sowie die sogenannte Verzweigungsstelle (Will and Luhrmann, 2011). Chemisch gesehen handelt es sich beim Spleißprozess um zwei aufeinanderfolgende Umesterungs-Reaktionen: Im ersten Schritt greift die 2'-OH-Gruppe eines Adenosins innerhalb der Verzweigungsstelle die Phosphodiesterbindung an der 5'-SS an, was zur Freisetzung des 5'-Exons und zu einer Zirkularisierung des Introns führt. Im zweiten Schritt greift die freie 3'-OH-Gruppe des 5'-Exons an der 3'-SS an. Dies resultiert in der Verknüpfung beider Exons und der Freisetzung des zirkulären Introns (Moore and Sharp, 1993).

## 3.3 Die Rolle der U-snRNPs während dem Spleißprozess

Die Erkennung der cis-agierenden Sequenzelemente durch das Spleißosom erfolgt über dessen Hauptbestandteile, den vier small nuclear ribonucleoprotein-particles (snRNPs) U1, U2, U4/U6 und U5 (Wahl et al., 2009). Diese RNA-Proteinkomplexe sind durch den Besitz Uridylsäure-reicher kleiner nukleärer und nicht-kodierender RNAs (U-rich small nuclear RNA, U-snRNA) charakterisiert, welche mit einer Vielzahl an Proteinen assoziiert sind. Die jeweiligen Partikel haben unterschiedliche Funktionen während des umfangreichen Reifungsprozesses des Spleißosoms, in deren Folge es zur Ausbildung des katalytisch aktiven RNA-Netzwerkes kommt. Hierfür sind vielfache konformationelle Umlagerungen, sowie der koordinierte Zu- und Abgang einer Vielzahl von Faktoren notwendig. Das Spleißosom gehört damit auch zu den komplexesten und dynamischsten makromolekularen Maschinen innerhalb der Zelle (Will and Luhrmann, 2011).

Für jeden einzelnen Spleiß-Vorgang werden die trans-agierenden Spleißfaktoren auf der jeweiligen prä-mRNA zusammengelagert und aktiviert. Die initiale Lokalisierung der intronischen Sequenzmotive erfolgt durch die U1- und U2-snRNPs. Während U1 die 5'-SS bindet (Ruby and Abelson, 1988), erfolgt die Erkennung der 3'-SS und des Verzweigungspunktes durch U2 (Komplex E) (Madhani and Guthrie, 1994). Zumindest in höheren Lebewesen sind wegen der Länge der

intronischen Sequenzen außerdem umfangreiche Umlagerungen notwendig, um die 5' und 3' Elemente zueinander in räumliche Nähe zu bringen (Komplex A). Anschließend erfolgt die Bindung von U4/U6 und U5, die als zuvor gebildetes tri-snRNPs U4/U6.U5 hinzukommen (Komplex B) (Lamond et al., 1988). Dieser Komplex durchläuft weitere Umlagerungen kompositioneller und konformationeller Art. Unter anderem verlassen U1 und U4 den Komplex und es kommt zur Bildung des katalytisch aktiven Spleißosoms (Komplex B\*) (Staley and Guthrie, 1998). U1, U2, U5 und U4/U6 sind also während des Zusammenlagerungsprozesses notwendige Bestandteile, im katalytisch aktiven Spleißosom finden sich letztendlich aber nur noch U2, U5 und U6. Die eigentliche Spleißreaktion beginnt mit der ersten Umesterung, in deren Folge sich ein zirkularisiertes, lassoartiges Intermediat bildet (Komplex C). Bevor es mit dem Angriff der freien 3'-OH-Gruppe des 5'-Exons auf die 3'-SS zur zweiten Umesterung kommen kann, sind nochmals konformationelle Umlagerungen nötig. Anschließend werden die beiden Exons verknüpft, wodurch das Intron nicht mehr weiter Bestandteil der prä-mRNA ist. Das Intron wird anschließend freigesetzt und das Spleißosom zerfällt in seine Bestandteile (post-spleißosomaler Komplex) (Brow, 2002).

Der zuvor beschriebene Vorgang (illustriert in Abbildung 1, Seite 8) gilt für das sogenannte „major“ Spleißosom. Zusätzlich existiert noch ein „minor“ Spleißosom, dass für das Spleißen von atypischen Introns verantwortlich ist, die ungefähr 1% der intronischen Sequenzen im menschlichen Genom ausmachen (Steitz et al., 2008). Diese Introns besitzen mit AT und AC abweichende Erkennungssequenzen an den 5'- und 3'-SS und werden daher auch ATAC-Introns genannt. Das „minor“-Spleißosom zeigt eine funktionelle und strukturelle Ähnlichkeit zum „major“-Spleißosom, die U1, U2 und U4/U6 sind aber mit U11, U12 und U4atac/U6atac durch alternative U-snRNPs ersetzt (Montzka and Steitz, 1988).

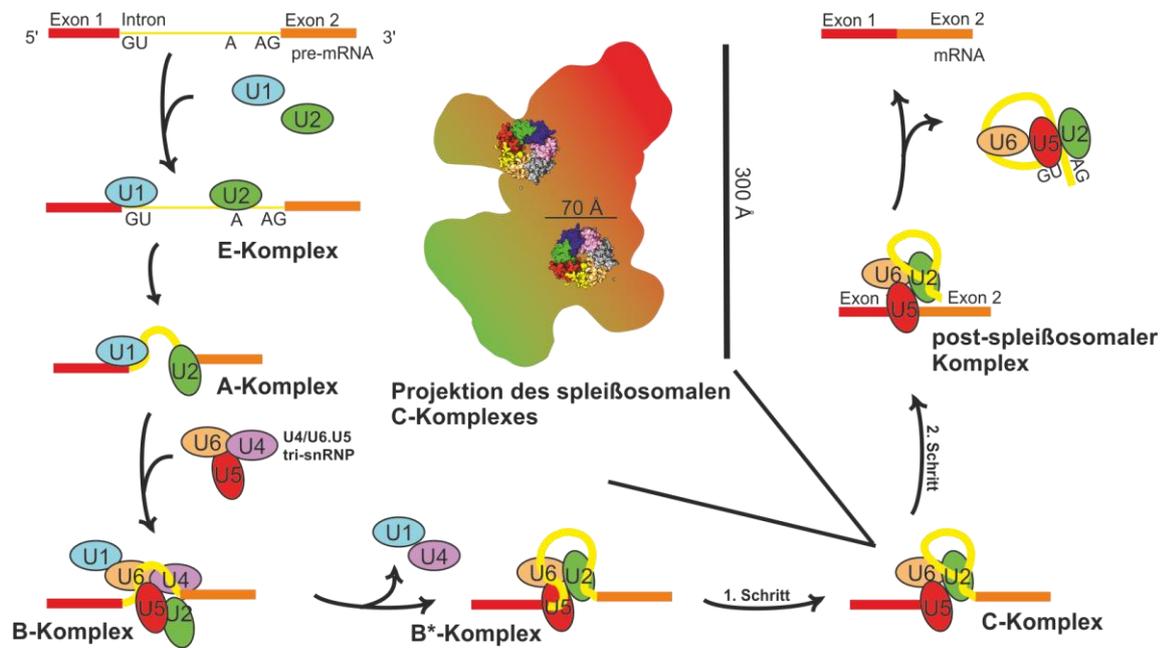


Abbildung 1 Schematische Darstellung der Zusammenlagerung des Spleißosoms und der Katalyse der Spleißreaktion.

### 3.4 Die Protein-Komponenten der U-snRNPs

Die jeweiligen Funktionen der verschiedenen U-snRNPs werden durch einen spezifischen Proteinanteil unterstützt (Abbildung 2). Zusätzlich besitzen die Partikel einen generellen Proteinanteil, der durch die sieben Sm-Proteine B/B', D1, D2, D3, E, F und G (bei B' handelt es sich um eine Spleiß-Alternative von B) repräsentiert ist (van Dam et al., 1989). Diese Proteine erkennen auf den snRNAs ein gemeinsames Sequenzmerkmal, das aus einem hochkonservierten einzelsträngigen Bereich mit dem Sequenzmotiv PuAU<sub>4-6</sub>GPU besteht (Luhrmann et al., 1990). Durch die Bindung der sieben Sm-Proteine an diese Sm-Bindungsstelle entsteht eine heptamere Ringstruktur, die als *Sm-Core*-Domäne bezeichnet wird (Kastner et al., 1990). Unter den snRNAs bildet die U6- (und U6atac-) snRNA eine Ausnahme, da sie keine Sm-Bindungsstelle besitzt und allein nicht mit Sm-Proteinen assoziiert. Das U6-snRNP besitzt jedoch mit den sieben Lsm- (*like Sm*-) Proteinen 2-8 einen Anteil strukturell verwandter Proteine. Deren Bindung an die U6-snRNA erfolgt zwar am 3'-Ende, aber ebenfalls unter Ausbildung einer heptameren Ringstruktur (Zhou et al., 2014).

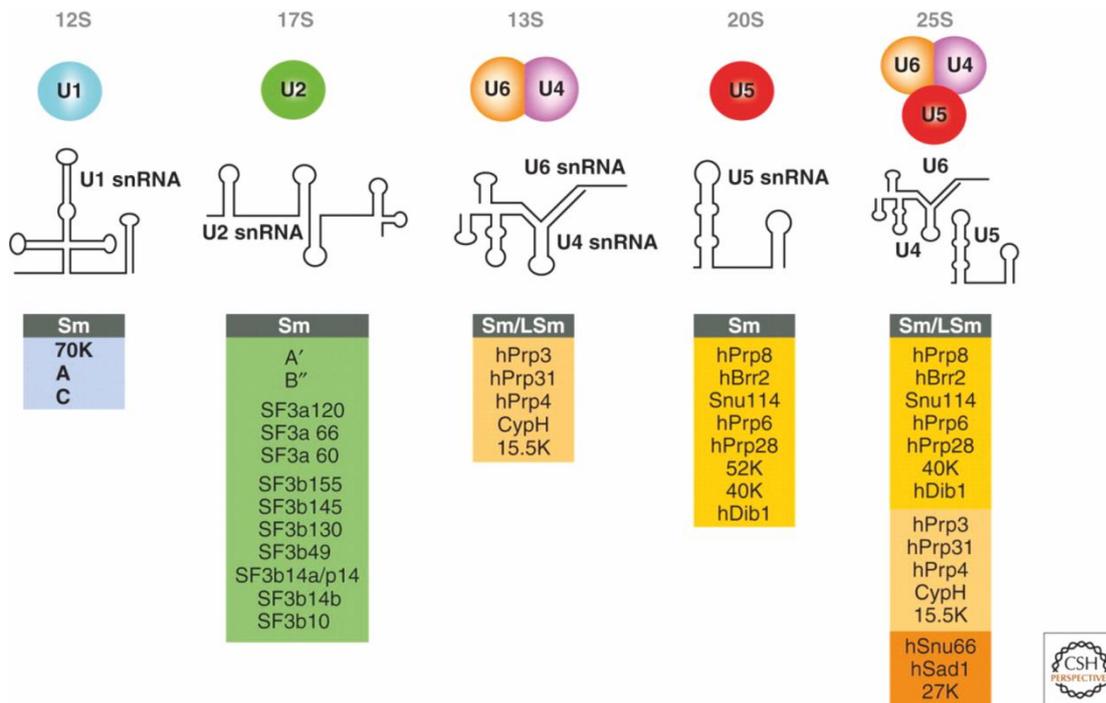
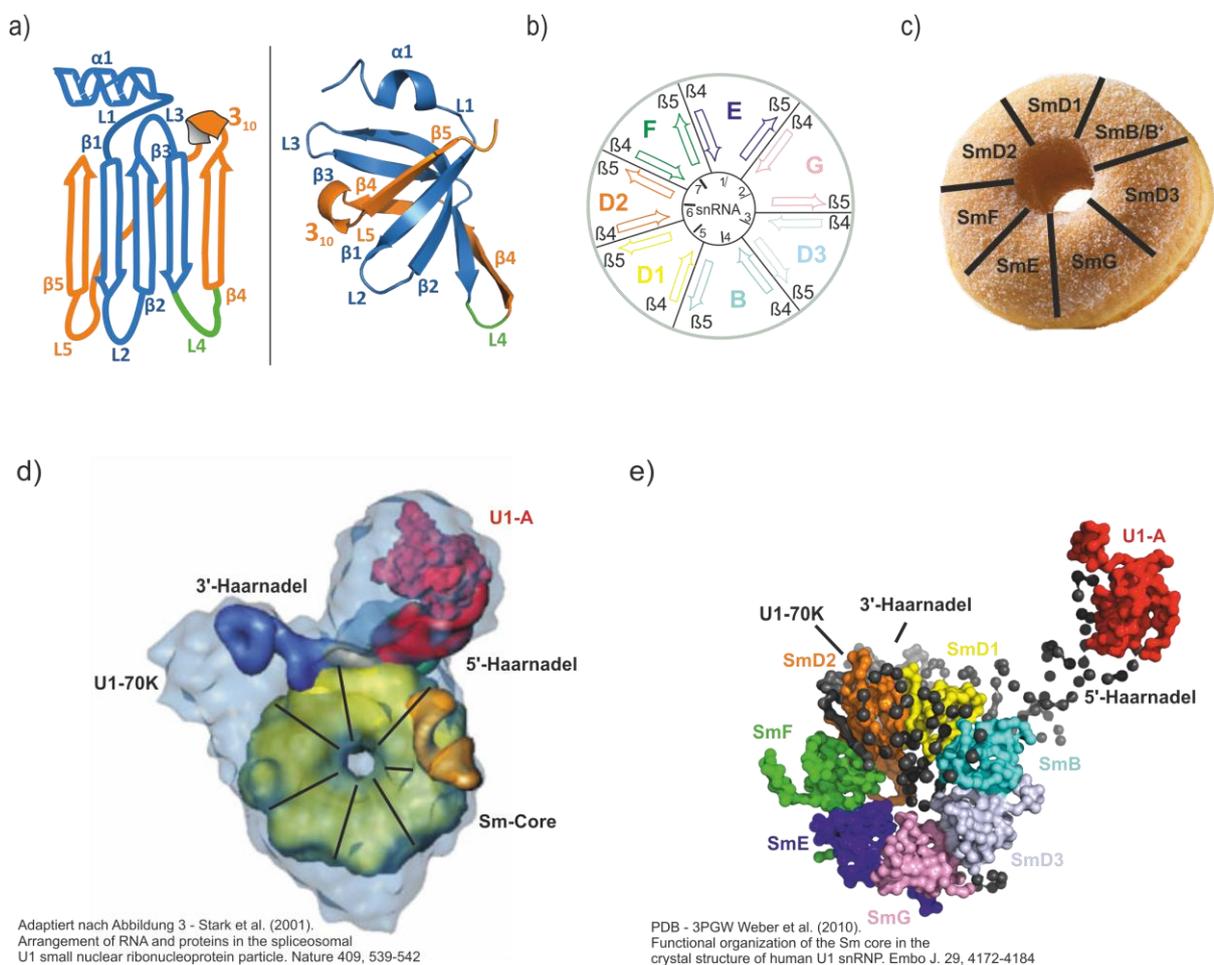


Abbildung 2 Organisation und Komponenten der U-snRNPs des Major-Spleißosoms. Mit Erlaubnis von Cold Spring Harbor Laboratory Press (nach Will, C.L. und Lührmann, R., 2011).

Sm- und Lsm-Proteine gehören zu einer evolutionär konservierten Protein-Familie, die durch ein bipartites Sequenzmotiv gekennzeichnet ist, das in die Sm-Domäne faltet (Achsel et al., 2001). Sie zeichnet sich durch eine N-terminale  $\alpha$ -Helix und ein fünfsträngiges antiparalleles  $\beta$ -Faltblatt aus (Kambach et al., 1999). Während die Paarung der  $\beta$ -Stränge  $\beta$ 1-4 zwischen  $\beta$ -Strängen erfolgt, die in der Primärsequenz benachbart sind, paart  $\beta$ 5 mit dem N-terminalen  $\beta$ 1-Strang (Abbildung 3a, links, Seite 10). Dadurch ist das  $\beta$ -Faltblatt sehr stark gekrümmt und es ergibt sich eine fassförmige Struktur (Abbildung 3a, rechts, Seite 10). Die Interaktion der Sm-Proteine miteinander beruht jeweils auf einer antiparallelen Paarung der beiden äußeren  $\beta$ -Stränge  $\beta$ 4 und  $\beta$ 5 miteinander. Dadurch besitzt jedes Sm-Protein zwei unterschiedliche Interaktionsflächen, durch welche jeweils ein benachbarter Interaktionspartner gebunden wird. Da diese Bindung durchgängig antiparallel erfolgt entsteht ein kontinuierliches antiparalleles  $\beta$ -Faltblatt, in dem die Sm-Proteine die sequentielle Abfolge SmB/B'-D1-D2-F-E-G-D3 haben, was in einer geschlossenen Ringstruktur resultiert (Abbildung 3b). In der Mitte dieser Ringstruktur bildet sich eine zentrale Pore, durch welche die snRNA verläuft. Die Erkennung der Sm-Bindungsstelle geschieht über äquivalente Positionen innerhalb der L3-Schleifen (zwischen den  $\beta$ -Strängen  $\beta$ 3 und  $\beta$ 4) und der L5-Schleifen (zwischen den  $\beta$ -Strängen  $\beta$ 3 und  $\beta$ 4) der Sm-Proteine. Hierbei kontaktiert jedes Sm-Protein spezifisch jeweils ein Nukleotid des Sequenzmotivs in 5'-3'-Richtung in der Reihenfolge SmE-G-D3-B/B'-D1-D2-SmF (Abbildung 3b, Seite 10). Die Basen der snRNA müssen hierfür in Richtung der Sm-Proteine nach außen gerichtet

werden, wofür eine ungewöhnliche Konformation des RNA-Phosphatrückgrads notwendig ist. (Leung et al., 2011; Pomeranz Krummel et al., 2009; Weber et al., 2010). Makroskopisch lässt sich diese Struktur anschaulich mit der Form eines Donuts vergleichen (Abbildung 3c). Die einzelsträngige Sm-Bindungsstelle mit dem heptameren Sm-Proteinring wird sowohl 5'- als auch 3'-seitig von jeweils einer Haarnadelstruktur flankiert (Branlant et al., 1982). Diese markanten Elemente bilden gemeinsam den strukturellen Kern der Partikel, was zur Bezeichnung Sm-Core-Domäne geführt hat. Sie ist ausschlaggebend für Stabilität und Funktion der Partikel, welche auch als Sm-Klasse der U-snRNPs bezeichnet werden (Liautard et al., 1982).

In den letzten Jahren konnten elektronenmikroskopische (Abbildung 3d, Seite 10) und kristallographische Strukturen (Abbildung 3e) der Sm-Core-Domänen der U1- und U4-snRNP Partikel gelöst werden (Leung et al., 2011; Pomeranz Krummel et al., 2009; Stark et al., 2001; Weber et al., 2010).



**Abbildung 3** Strukturelle und topologische Eigenschaften der Sm-Proteine und des Sm-Cores der U-snRNPs. a) Sekundärstrukturelemente der Sm-Faltung (links) und deren dreidimensionale Struktur am Beispiel von SmG. Elemente die zum Sm1-Sequenzmotiv gehören sind in blau, die Elemente des Sm2-Sequenzmotivs in orange dargestellt. Die L4-Schleifenregion trennt das bipartite Sequenzmotiv und ist von Sm-Protein zu Sm-Protein unterschiedlich lang. b) Topologie und Interaktion zwischen den Untereinheiten des Sm-Rings des Sm-Core. Die Sm-Bindungsstelle der snRNA

durchläuft die zentrale Pore der Ringstruktur. d) Makroskopisch kann man den Sm-Ring gut mit einem Donut vergleichen. d) Die EM-Struktur des U1-snRNP zeigt die markante Struktur des Sm-Core. Mit Erlaubnis von Nature Publishing Group (verändert nach Stark, H. et al., 2001). e) Im Vergleich zur EM-Struktur (e) ermöglichte die Kristallstruktur des U1-snRNP Sm-Core detailliertere Einblicke, wie beispielsweise die eindeutige Positionierung der Sm-Proteine innerhalb des Sm-Core.

### 3.5 Die Biogenese der U-snRNPs der Sm-Klasse

Bevor die U-snRNPs ihre jeweilige Rolle im Spleißprozess übernehmen können, durchlaufen sie einen komplexen und hochregulierten Reifungsprozess. Dieser beginnt mit der Transkription der U-snRNAs durch die RNA-Polymerase II im Nukleus. Wie auch prä-mRNAs erhalten die prä-U snRNAs eine 5'-terminale  $m^7GpppG$ -Kappenstruktur ( $m^7G$ -Cap). Das 3'-Ende hingegen unterscheidet sich bezüglich seiner cis-agierenden Elemente, als auch hinsichtlich der erforderlichen Prozessierung von mRNA-Transkripten (Chen et al., 2012; Ezzeddine et al., 2011). Es beinhaltet ein 3'-Box genanntes Prozessierungssignal, welches vom Integratorkomplex erkannt wird (Baillat et al., 2005). Daraufhin wird das naszierende Transkript endonukleolytisch gespalten und es kommt zur Termination der Transkription (Mandel et al., 2006). Trotz ihrer Funktion innerhalb des Nukleus werden die neu synthetisierten und prozessierten U-snRNAs vorübergehend ins Cytoplasma transportiert. Hierbei ist die 5'-Capstruktur entscheidend, welche vom CBC (cap-binding-complex) und den Adapterproteinen PHAX und ARS2 erkannt wird. Diese sind für snRNAs spezifisch und vermitteln den Exportin1/RanGTP-abhängigen Export durch die Kernpore. Im Cytoplasma kommt es zur Dissoziation des Exportkomplexes und zur Zusammenlagerung des Sm-Core. Es dient als Erkennungssignal für die RNA-Methyltransferase TGS1, welche das  $m^7G$  der Cap-Struktur zu einem 2,2,7-Trimethylguanosin modifiziert (Mouaikel et al., 2003). Das hypermethylierte Cap dient gemeinsam mit dem Sm-Core als kombiniertes Signal für den Kernimport (Fischer and Luhrmann, 1990). Bevor der Import der snRNPs in den Kern erfolgen kann, wird zusätzlich noch das 3'-Ende exonukleolytisch auf seine endgültige Länge verkürzt. Die nun hypermethylierte Cap-Struktur wird durch den Importfaktor Snurportin erkannt, der mit Importin- $\beta$  interagiert und die snRNPs in den Kern transportiert (Huber et al., 1998; Palacios et al., 1997). Im Nukleus akkumulieren die snRNPs vorübergehend in subnukleären Domänen, den sogenannten „Cajal Bodies“ und es kommt zur abschließenden Reifung der Partikel. Hier binden die spezifischen Proteine (Abbildung 2, Seite 9) und es werden einige Nukleotide der snRNA kovalent modifiziert (Jady et al., 2003; Nasic et al., 2004). Anschließend verlassen die gereiften Partikel die Cajal Bodies, um schließlich als Komponenten des Spleißosoms am Spleißprozess teilzunehmen (Abbildung 4, Seite 12).

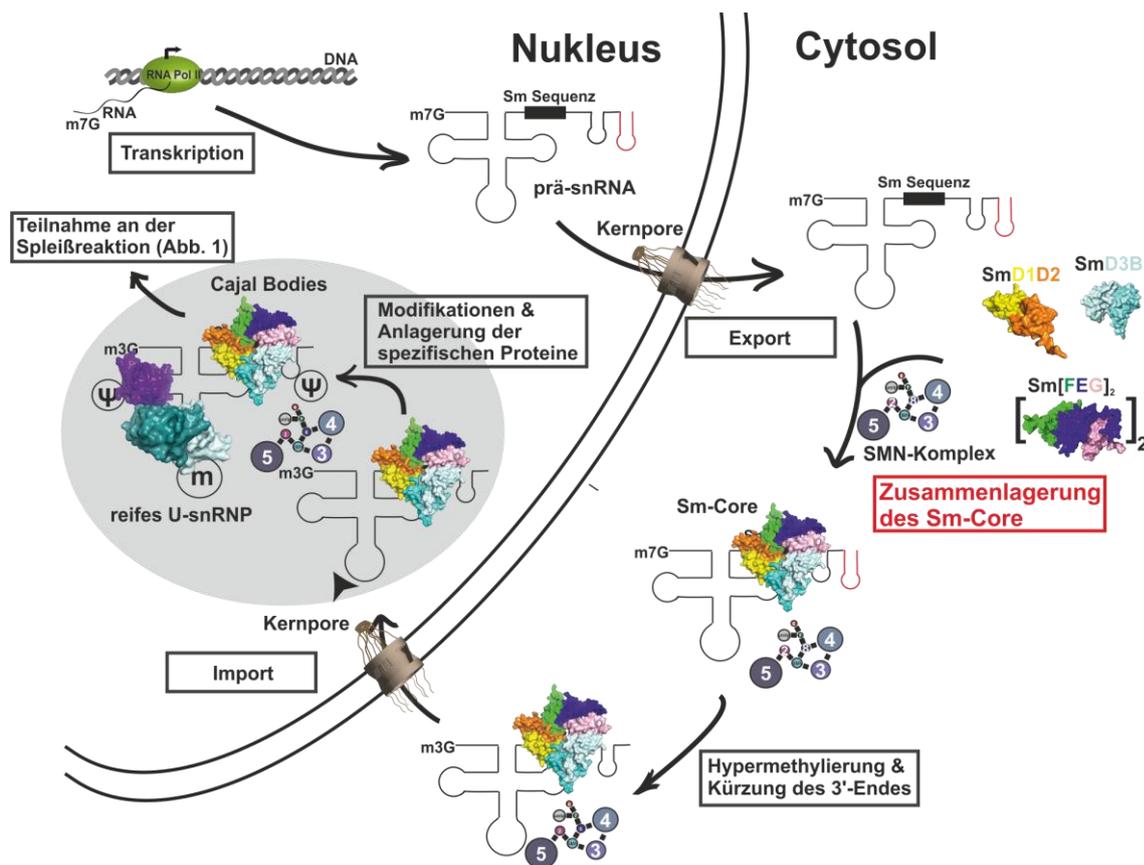


Abbildung 4 Schematischer Ablauf der Biogenese der snRNPs der Sm-Klasse.

### 3.6 Assistenzfaktoren während der cytosolischen Phase der U-snRNP-Biogenese

Die Bildung des Sm-Core stellt den zentralen Schritt der U-snRNP Biogenese dar, da sie Voraussetzung für deren weitere Reifung und den Import in den Kern dient. *In vitro* Studien führten zur Postulierung eines mehrstufigen Modells zum Ablauf dieses Prozesses. So bilden die Sm-Proteine nach ihrer Translation am Ribosom drei unterschiedliche oligomere Komplexe aus. Hierbei lassen sich zwei Dimere, SmD1-D2 und SmB/B'-D3 sowie ein dritter Komplex aus SmE, F und G charakterisieren, der als Sm(F-E-G)<sub>2</sub> Hexamer vorliegt (Plessel et al., 1997). Diese Heterooligomere bilden sich unabhängig von der snRNA und besitzen isoliert voneinander auch keine hinreichende Affinität zur snRNA (Raker et al., 1999). Liegen jedoch SmD1-D2 und SmE-F-G gemeinsam in Gegenwart der snRNA vor, bildet sich spontan eine als Sm-Subcore bezeichnete Zwischenstufe des Sm-Core. Durch Zugabe von SmB/B'-D3 kann dieses Sm-Subcore anschließend in das vollständige Sm-Core überführt werden (Raker et al., 1996). Eine solche hierarchische Abfolge der Zusammenlagerung des Sm-Core über das Sm-Subcore ließe ein pentameres Intermediat aus den Proteinen SmD1-D2-E-F-G vermuten. SmD1-D2 und SmE-F-G zeigen jedoch in Abwesenheit von RNA keine Affinität zueinander und bilden alleine keinen gemeinsamen Komplex aus.

Während sich das Sm-Core *in vitro* spontan ausbildet, gilt heute als sicher, dass dieser Prozess *in vivo* die Assistenz durch eine Vielzahl von Faktoren benötigt. So sind nach heutigem Stand mindestens 12 Proteinfaktoren an der Sm-Core-Zusammenlagerung beteiligt, die sich in zwei distinkten Einheiten organisieren, dem PRMT5- sowie dem SMN-Komplex (Fischer et al., 1997; Meister and Fischer, 2002; Pellizzoni et al., 2002). Da beide Einheiten aufeinanderfolgende Aufgaben übernehmen, ergibt sich für die Zusammenlagerung eine sequenzielle Abfolge, die in eine frühe sowie eine späte Phase eingeteilt wird.

### **3.7 Die frühe Phase der Zusammenlagerung ist vom PRMT5-Komplex abhängig**

In der frühen Phase der Zusammenlagerung des Sm-Core ist der PRMT5-Komplex aktiv, welcher die Sm-Proteine rekrutiert. Dessen zentrale namensgebende Komponente gehört zur Enzymklasse der Protein-Arginin-Methyltransferasen (PRMT) und übernimmt die Methylierung von Aminen der Guanidingruppe von Argininseitenketten innerhalb der Substratproteine. PRMT5 katalysiert hierbei symmetrische Dimethylierungen, was bedeutet, dass je eine Methylgruppe an den beiden Aminen der Guanidingruppe addiert wird (Branscombe et al., 2001). Neben der katalytisch aktiven Untereinheit gehört auch WD45 (ebenfalls als MEP50 bekannt) (Friesen et al., 2002), ein Protein der WD40-Repeat-Familie (Li and Roberts, 2001; Neer et al., 1994), zum Grundgerüst des PRMT5-Komplexes. Dieses Grundgerüst wird durch weitere substratspezifische Proteine ergänzt, die beispielsweise die Rekrutierung der Substratproteine übernehmen können (Guderian et al., 2011; Meister et al., 2001). Zu diesen Substratproteinen gehören auch die Sm-Proteine SmB/B', D1 und D3, die Arginin/ Glycin (RG)-reiche C-Termini aufweisen, an deren Argininresten die Methylierung erfolgt (Friesen et al., 2001b; Meister et al., 2001; Meister and Fischer, 2002). Welche Rolle diese Modifikation der Sm-Proteine spielt, ist jedoch noch ungeklärt.

Der zentrale rekrutierende und organisierende Faktor in dieser Phase des Prozesses stellt pICln dar. Dieses bindet die Sm-Proteine und es lassen sich aus cytosolischen Extrakten zwei pICln-assoziierte Komplexe mit Sedimentationskoeffizienten von 20S und 6S isolieren (Chari et al., 2008; Pu et al., 1999). Der 20S-Komplex beinhaltet mit PRMT5, WD45 und pICln die Komponenten des PRMT5-Komplexes, sowie alle sieben Sm-Proteine. SmE, F und G sind in diesem Komplex im Vergleich zu SmD1-D2 und SmB/B'-D3 aber unterrepräsentiert (Chari et al., 2008). Der 6S-Komplex hingegen ist frei von SmB/B'-D3, existiert unabhängig von PRMT5/WD45 und enthält neben pICln die Sm-Proteine D1, D2, E, F und G. Somit beinhaltet der 6S-Komplex genau jene Sm-Proteine, die für die Ausbildung des Sm-Subcore benötigt werden. Trotz dieser hinreichenden Komposition zeigt der 6S-Komplex keine Affinität zu snRNA. Da aus zellulären Extrakten isolierter 6S-Komplex bereits methyliertes SmD1 enthält, ist davon auszugehen, dass der 6S-Komplex während der Assoziation der

Sm-Proteine mit dem PRMT5-Komplex oder im Anschluss an diese entsteht. Daraufhin erfolgt seine Dissoziation vom PRMT5-Komplex und der 6S-Komplex wird als intermediäres Zwischenprodukt freigesetzt.

### 3.8 Die späte Phase der Zusammenlagerung und die Rolle des SMN-Komplexes

Im Anschluss an diese frühe Phase werden die an pICln gebundenen und im 6S-Komplex organisierten Sm-Proteine vom SMN-Komplex übernommen. Dessen namensgebende Komponente ist das *survival motor neuron* (SMN) Protein. Diese Bezeichnung beruht darauf, dass genetische Defekte innerhalb des kodierenden SMN1-Genes zur Spinalen Muskelatrophie (SMA) führen. Die Krankheit äußert sich durch den fortschreitenden Untergang von motorischen Vorderhornzellen im Rückenmark und ist eine der häufigsten genetischen Ursachen für Kindersterblichkeit. SMN ist Teil des gleichnamigen Komplexes, der bei Vertebraten im Cytosol aus neun verschiedenen Proteinfaktoren besteht. Neben SMN handelt es sich bei diesen Faktoren um die sieben Gemine 2-8 und ein weiteres Protein namens unrip (Abbildung 5) (Carissimi et al., 2005; Grimmeler et al., 2005; Otter et al., 2007). Obwohl die Zusammenlagerung des Sm-Core *in vitro* völlig spontan und in Abwesenheit des SMN-Komplexes erfolgen kann, ist dieser *in vivo* notwendig für den Transfer der Sm-Proteine auf die snRNA (Fischer et al., 1997; Liu et al., 1997). Hierbei wurde dem SMN-Komplex eine Kontrolllesefunktion bei der korrekten Zusammenlagerung des Sm-Core zugeschrieben (Kroiss et al., 2008). Außerdem konnte kürzlich gezeigt werden, dass der SMN-Komplex die Dissoziation von pICln aus dem 6S-Komplex bewerkstelligt, was für die Fortsetzung des Zusammenlagerungsprozesses notwendig ist (Chari et al., 2008).

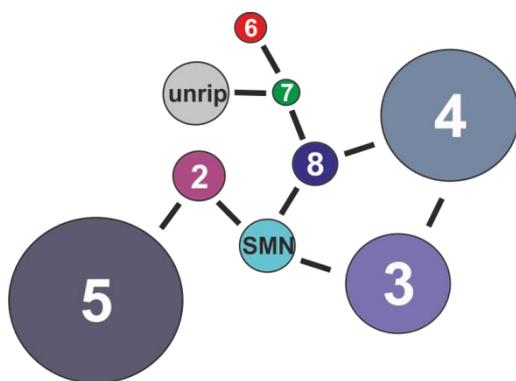


Abbildung 5 Schematische Interaktionskarte des SMN-Komplexes

### 3.9 Die Beteiligung individueller Komponenten des SMN-Komplex an der Sm-Core-Zusammenlagerung

Der SMN-Komplex besteht bei Vertebraten aus den neun Proteinen SMN, Gemin2-8 und unrip. Die einzelnen Komponenten existieren jedoch auch zeitweise in zwei SMN-unabhängigen Komplexen, die aus den Geminen3/4/(5) und Gemin6/7/(unrip) bestehen. Des Weiteren lässt sich ein Grundgerüst des SMN-Komplexes bestehend somit aus SMN, Gemin2 und Gemin8 identifizieren, an welches sich die weiteren Untereinheiten modulartig anlagern können (Battle et al., 2007). Die Entwicklung des SMN-Komplexes erfolgte zu einem frühen evolutionären Zeitpunkt und zu allen Komponenten des SMN-Komplexes sind bereits ab den Cnidariern Orthologe bekannt. Die einzige Ausnahme stellt Gemin4 dar, das bis dato erst ab dem Phylum der Echinodermata nachgewiesen werden konnte (Kroiss et al., 2008). In mehreren *Knock-out* Experimenten mit unterschiedlichen Modellsystemen wurde festgestellt, dass es sich bei SMN, Gemin2, Gemin3 und Gemin5 um essentielle Proteine handelt (Briese et al., 2009; Cauchi, 2010; Cauchi et al., 2008; Jablonka et al., 2002; Schrank et al., 1997; Shpargel et al., 2009). Daher ist es umso interessanter, dass zumindest *in vitro* bereits ein auf die Untereinheiten SMN, Gemin2 und Gemin8 reduzierter SMN-Komplex hinreichend für die Dissoziation von pICln und den Transfer der Sm-Proteine auf die Sm-Bindungsstelle der snRNA ist. Dieser scheint auch über die Kontrolllese-Funktion zu verfügen (Chari et al., 2008; Kroiss et al., 2008). Eine Analogie zu diesem reduzierten SMN-Komplex existiert allem Anschein nach, zumindest im Phylum der Dipteren, auch *in vivo*. So wurde aus *Drosophila melanogaster* ein funktionsfähiger dimerer SMN-Komplex isoliert, der nur aus SMN und Gemin2 bestand. Orthologe zu den anderen Geminen sind in Dipteren entweder nicht vorhanden oder haben im Fall von Gemin3 (Dhh1) und Gemin5 (Rigor mortis) einen Funktionsverlust erlitten (Kroiss et al., 2008).

Dieses reduzierte System eignet sich aufgrund seiner verringerten Komplexität und der Möglichkeit der rekombinanten Präparation für biochemische Studien und besonders auch für strukturelle Untersuchungen. So lassen sich auf Basis dieses reduzierten SMN-Komplexes zwei Sm-Transferintermediate mit Sedimentationskoeffizienten von 7S und 8S rekonstituieren. Beide Transferintermediate beinhalten das Sm-Pentamer SmD1-D2-E-F-G, welches an die beiden SMN-Komplex-Komponenten SMN und Gemin2 gebunden ist. pICln ist trotz der Bindung des SMN-Komplexes, die normalerweise zu seiner Dissoziation führen sollte, weiterhin stabiler Bestandteil des 8S-Komplexes. Diese Stabilisierung wird durch eine C-terminale Kürzung von SMN erreicht (Chari et al., 2008). Der 8S-Komplex kann also als frühes Intermediat angesehen werden, da pICln den Komplex noch nicht verlassen hat. pICln ist hingegen nicht Bestandteil des 7S-Komplexes. Dieser wird daher auch als spätes Intermediat angesehen, das im Anschluss an die Dissoziation von pICln entsteht (Abbildung 5, Seite 14). Durch eine Analyse der Strukturen des 6S-, 7S- und 8S-Komplexes sollte ein detaillierter Einblick in den Übergang von der frühen zur späten Phase der

Sm-Core-Zusammenlagerung erhalten werden, um die mechanistischen Grundlagen des Sm-Protein-Transfers von pICln auf den SMN-Komplex auf atomarer Ebene aufzuklären.

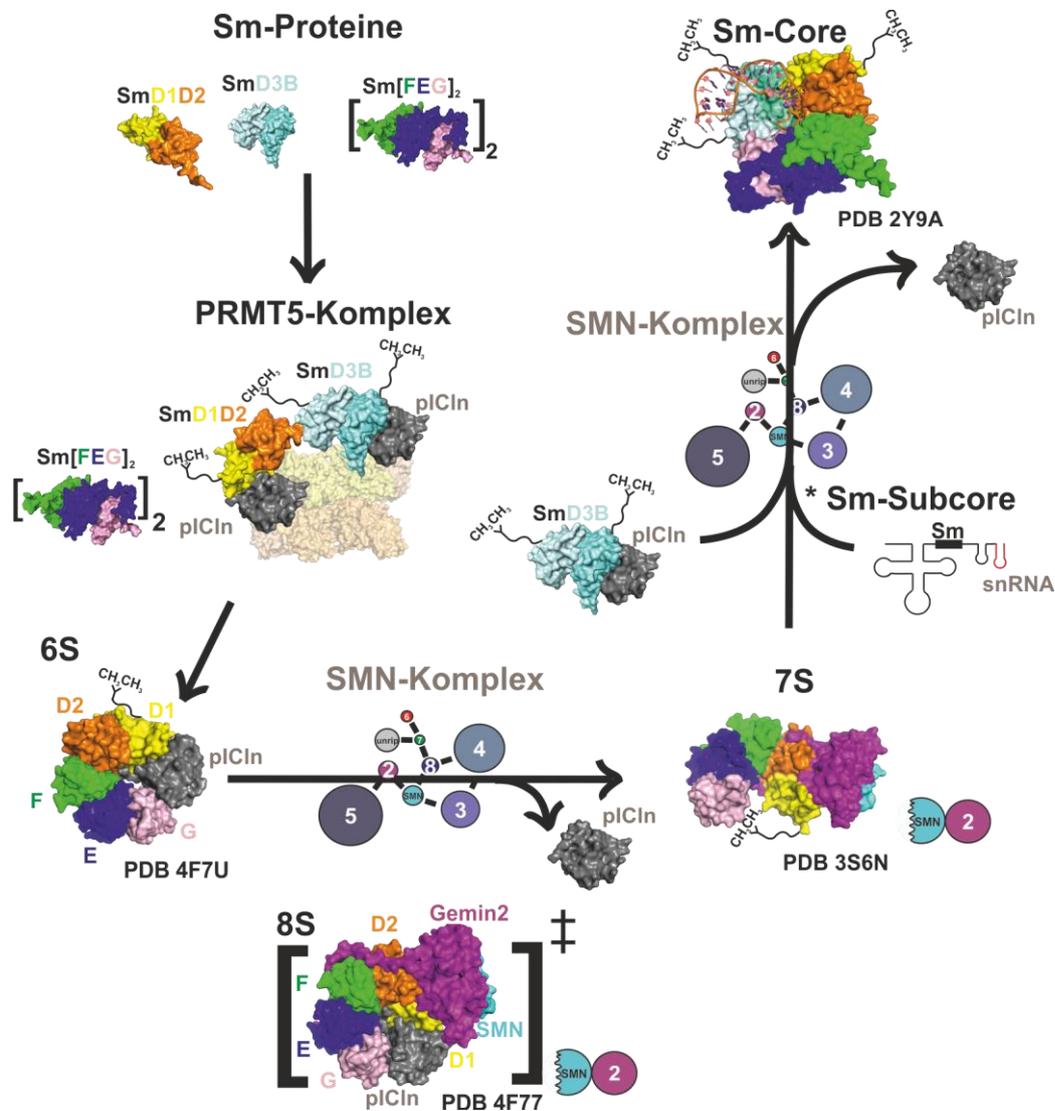


Abbildung 6 Schematischer Ablauf der Sm-Core-Zusammenlagerung. Nach ihrer Synthese bilden die sieben Sm-Proteine spontan Heterooligomere aus. Diese werden von pICln rekrutiert und am PRMT5-Komplex kommt es zur symmetrischen Dimethylierung von SmD1, SmB/B' und SmD3 sowie zur Ausbildung des 6S-Komplexes. Der SMN-Komplex dissoziiert pICln und transferiert das Sm-Pentamer D1-D2-E-F-G und SmB/B'-D3 auf die snRNA, wodurch es zur Ausbildung des Sm-Core kommt. Zumindest *in vitro* lässt sich in Abwesenheit von SmB/B'-D3 ein intermediäres Sm-Subcore erzeugen, das nur SmD1-D2-E-F-G enthält. Die durch *in vitro* Rekonstitution erzeugten 8S- und 7S-Komplexe stellen stabilisierte Sm-Transferintermediate dar, die *in vivo* nicht isolierbar sind. Der Farbcode der Untereinheiten wird über die gesamte Arbeit verwendet (SmD1 gelb, SmD2 orange, SmE blau, SmF grün, SmG rosa, pICln grau, Gemin2 magenta, SMN türkis, SmD3 blau-grau, SmB hellblau).

## 4. Material & Methoden

### 4.1 Chemikalien

Alle Chemikalien wurden kommerziell und in analytischer Qualität (p.A.) erworben. Das für die Experimente verwendete Wasser wurde in einer GenPure-Wasseraufbereitungsanlage gefiltert. Bei der Herstellung von Medien für die bakterielle Zellkultur wurde vollentsalztes (VE)-Wasser aus der Anlage des Biozentrums verwendet, um einen Verbleib von Spurenelementen in den Medien für ein besseres Wachstum der Zellen sicherzustellen.

### 4.2 Materialien

#### Geräte

Äkta prime plus	GE-Healthcare, München , Deutschland
Äkta purifier	GE-Healthcare, München , Deutschland
Äkta explorer	GE-Healthcare, München , Deutschland
Avanti® J-20-XP Zentrifuge	Beckman-Coulter, Krefeld, Deutschland
Avanti® J-HC Zentrifuge	Beckman-Coulter, Krefeld, Deutschland
BioPhotometer (uv/vis)	Eppendorf, Hamburg, Deutschland
Cartesian Honeybee Nano Dispensor	Hamilton Robotic, Martinsried, Deutschland
Dialyseschlauch ZelluTrans	Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland
Geldokumentationssystem	Peqlab QUANTUM, Erlangen Deutschland
GeneAmp® PCR System 9700	Applied Biosystems, Carlsbad, CA, USA
GenPure Wasseraufbereitungsanlage	TKA, Niederelbert, Deutschland
Innova® 44 Inkubatorschüttler	New Brunswick Scientific, Edison, NJ, USA
Mini-PROTEAN® Tetra Cell	Bio-Rad, München, Deutschland
Novaspec II (vis) Spektralphotometer	Pharmacia, Berlin, Deutschland
Optima™ L-80XP Ultrazentrifuge	Beckman-Coulter, Krefeld, Deutschland
Optima™ L-90K Ultrazentrifuge	Beckman-Coulter, Krefeld, Deutschland
pH-Meter 522	WTW, Weilheim, Deutschland
Rotoren – Ultrazentrifugen (45Ti, 70Ti, SW60Ti)	Beckman-Coulter, Krefeld, Deutschland
Rotoren – Zentrifugen (JLA 8.1000, JA 25.50)	Beckman-Coulter, Krefeld, Deutschland

Branson Sonifier 250	G. Heinemann, Schwäbisch Gmünd, Deutschland
Thermomixer compact	Eppendorf, Hamburg, Deutschland
Tischzentrifuge 5804R	Eppendorf, Hamburg, Deutschland
Überkopffrotor	Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland
Varioklav® Steam Sterilizer	H + P Labortechnik, Oberschleißheim, Deutschland

## Verbrauchsmaterialien

Additiv Screen	Hampton, Aliso Viejo, CA, USA
CrystalQuick™ Kristallisationsplatten 96 well	Greiner Bio-One, Frickenhausen, Deutschland
22 mm Deckgläschen (einfach & silikonisiert)	Jena Bioscience, Jena, Deutschland
Mounted LithoLoops	MDL, Newmarket, UK
Petrischalen	Greiner Bio-One, Frickenhausen, Deutschland
Reaktionsgefäße 15 + 50 ml	Greiner Bio-One, Frickenhausen, Deutschland
Reaktionsgefäße 0,2; 1,5; 2 ml	Eppendorf, Hamburg, Deutschland
Reaktionsgefäße ohne Deckel, 1,5 ml	Hartenstein, Würzburg, Deutschland
Slide-A-Lyzer® Dialyse Gefäße	Pierce, Rockford, IL, USA
Sterilfilter (0.22 & 0.45 µm)	Merck Millipore, Schwalbach
SuperClear™ Kristallisationsplatten 24 well	Jena Bioscience, Jena, Deutschland

## Chromatographie – Materialien & vorgepackte Säulen

Die Säulen wurden entweder vorgepackt gekauft oder im Falle der Materialien für die Affinitätschromatographie und Ionenaustauscherchromatographie entsprechend der Anleitungen der Firma GE-Healthcare manuell gepackt. Hierbei kamen XK- sowie C-Ummantelungen der Firma GE-Healthcare zur Anwendung.

Bio-Scale™ Mini CHT™	Bio-Rad, München, Deutschland
DEAE Sepharose FF	GE-Healthcare, München, Deutschland
HiTrap Heparin HP	GE-Healthcare, München, Deutschland
Mono Q™ 5/50 GL	GE-Healthcare, München, Deutschland
Ni-NTA Agarose	Qiagen, Hilden, Deutschland
Ni Sepharose HP	GE-Healthcare, München, Deutschland
Ni Sepharose 6 FF	GE-Healthcare, München, Deutschland

RESOURCE™ S	GE-Healthcare, München, Deutschland
RESOURCE™ Q	GE-Healthcare, München, Deutschland
SP Sepharose FF	GE-Healthcare, München, Deutschland
SP Sepharose HP	GE-Healthcare, München, Deutschland
Superdex 75 10/300 GL	GE-Healthcare, München, Deutschland
Superdex 75 26/60 PG	GE-Healthcare, München, Deutschland
Superdex 200 10/300 GL	GE-Healthcare, München, Deutschland
Superdex 200 26/60 PG	GE-Healthcare, München, Deutschland
Q Sepharose FF	GE-Healthcare, München, Deutschland
Q Sepharose HP	GE-Healthcare, München, Deutschland

## Medien für Bakterienkultur

### ***Terrific broth Medium (TB)***

- 1,2 % (w/v) Bacto-Trypton
- 2,4 % (w/v) Hefeextrakt
- 4 % (v/v) Glycerin
- 17 mM Kaliumdihydrogenphosphat
- 72 mM Dikaliumhydrogenphosphat

### ***Luria-Bertani-Medium (LB)***

- 1% (w/v) Bacto-Trypton
- 0,5% (w/v) Hefeextrakt
- 1% (w/v) NaCl

Für LB-Agarplatten wurden zu dem LB-Medium 1,5% (w/v) Agarose zugegeben, erhitzt, autoklaviert und in Petrischalen gegossen. Für Selektionsplatten wurde dem Medium nach dem Autoklavieren je nach Bedarf Ampicillin (100 mg/ml), Kanamycin (25 mg/ml) oder Chloramphenicol (30 mg/ml) beigemischt.

Agarose	Clontech, Heidelberg, Deutschland
Ampicillin	Carl Roth, Mannheim, Deutschland
Bacto-Trypton	Clontech, Heidelberg, Deutschland
Hefeextrakt	Clontech, Heidelberg, Deutschland

Chloramphenicol	Carl Roth, Mannheim, Deutschland
Isopropyl- $\beta$ -thiogalactosid (IPTG)	Carl Roth, Mannheim, Deutschland
Kanamycin	Carl Roth, Mannheim, Deutschland

## Plasmide und Bakterienstämme

### Plasmide

#### Leervektoren

pETM-11	prokaryotischer Expressionsvektor, kodiert für N- und optionale C-terminale His6-Markierung, $Kan^R$ , <i>lacI</i> , T7-Promoter, Schnittstelle für TEV-Protease
pETM-13	prokaryotischer Expressionsvektor, keine N-terminale Affinitätsmarkierung. Kodiert für optionale C-terminale His6-Markierung, $Kan^R$ , <i>lacI</i> , T7-Promoter, Schnittstelle für TEV-Protease
pETM-30	prokaryotischer Expressionsvektor, kodiert für N-terminale Glutathion-S-Transferase und His6-Markierung. Optionale C-terminale His6-Markierung, $Kan^R$ , <i>lacI</i> , T7-Promoter, Schnittstelle für TEV-Protease
pET15b	Prokaryotischer Expressionsvektor, kodiert für N-terminale His6-Markierung, $Amp^R$ , T7-Promoter, Schnittstelle für Thrombin
pET28a	Prokaryotischer Expressionsvektor, kodiert für N-terminales T7-Epitop und N-terminale His6-Markierung, $Kan^R$ , T7-Promoter, Schnittstelle für Thrombin
pLysSRARE2	Kodiert für 7 tRNAs, die <i>in E. coli</i> nur gering exprimiert werden. $Cam^R$ , pLysS

#### Expressionsvektoren

SmD1D2:pRK172	Dr. Christian Kambach, PSI Villingen (Kambach et al., 1999)
SmEFG:pET15b	Dr. Christian Kambach, PSI Villingen (Kambach et al., 1999)
dSMN(1-122):pETM30	Dr. Ashwin Chari, Dr. Matthias Kroiss (Chari et al., 2008; Kroiss et al., 2008)
dGemin2:pETM30	Dr. Ashwin Chari, Dr. Matthias Kroiss (Chari et al., 2008; Kroiss et al., 2008)
dmpICln:pET28a	Dr. Ashwin Chari (Chari et al., 2008)

### Bakterienstämme

<i>E. coli</i> BL21 Rosetta2 <sup>TM</sup> (DE3)pLysS	$\lambda$ DE3, kodiert für T7 RNA-Polymerase (IPTG induzierbar) und T7 Lysozym, <i>F</i> , <i>lon</i> <sup>-</sup> , <i>ompT</i> , <i>pLysSRARE2</i> ( $Cam^R$ ), <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup></i> ), <i>gal</i> , <i>dcm</i>
<i>E. coli</i> BL21 (DE3)pLysS	$\lambda$ DE3, kodiert für T7 RNA-Polymerase (IPTG induzierbar) und T7 Lysozym, <i>F</i> -, <i>lon</i> <sup>-</sup> , <i>ompT</i> , <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup></i> ), <i>gal</i> , <i>dcm</i>

<i>E. coli</i> BL21 Star <sup>TM</sup> (DE3)pLysS	λDE3, kodiert für T7 RNA-Polymerase (IPTG induzierbar) und T7 Lysozym, <i>F</i> <sup>-</sup> , <i>lon</i> <sup>-</sup> , <i>ompT</i> , <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup></i> ), <i>gal</i> , <i>dcm</i> , <i>rne131</i>  (Merck, Darmstadt, Deutschland)
<i>E. coli</i> DH5α	<i>fhuA2</i> , Δ( <i>argF-lacZ</i> )U169, <i>phoA</i> , <i>glnV44</i> , Φ80, Δ( <i>lacZ</i> )M15, <i>gyrA96</i> , <i>recA1</i> , <i>relA1</i> , <i>endA1</i> , <i>thi-1</i> , <i>hsdR17</i>  (Invitrogen, Carlsbad, CA, USA)

## Enzyme

Acc65I Restriktions-Endonuklease	Fermentas, St. Leon-Rot, Deutschland
KAPAHiFi <sup>TM</sup> Polymerase	Peqlab/ VWR, Erlangen, Deutschland
NcoI Restriktions-Endonuklease	Fermentas, St. Leon-Rot, Deutschland
Shrimp Alkaline Phosphatase (SAP)	Fermentas, St. Leon-Rot, Deutschland
T4 DNA-Ligase	Fermentas, St. Leon-Rot, Deutschland
Tabakätzvirus-Protease (TEV)	Eigenhändige rekombinante Herstellung im bakteriellen Expressionssystem

## 4.3 Methoden

### Agarosegelelektrophorese

Zur Auftrennung von DNA wurden 0,8 bis 1,5 prozentige Agarosegele auf Basis von TBE-Puffer hergestellt. Die maximale Leistung während der Auftrennung in einer horizontalen Gelelektrophoreseapparatur wurde auf 25 Watt begrenzt. Das Ergebnis wurde anschließend mit Hilfe einer Geldokumentationsanlage analysiert und dokumentiert.

#### 6x DNA-Probenpuffer

10 mM Tris/HCl, pH 8,0

30% (v/v) 86%iges Glyzerin

0,1% (w/v) XC

0,1% (w/v) BPB

GeneRuler<sup>TM</sup> 100 bp Plus DNA Ladder Fermentas, St. Leon-Rot, Deutschland

MassRuler<sup>TM</sup> DNA Ladder Mix Fermentas, St. Leon-Rot, Deutschland

Agarosegel (0,8 – 1,5%) 100 ml TBE

6 µl Ethidiumbromid

### **10x TBE-Puffer**

890 mM Tris

890 mM H<sub>3</sub>BO<sub>4</sub>

20 mM Na<sub>2</sub>EDTA

pH 8,3 mit HCl

Laufpuffer 1x TBE

## **Heterologe Expression von rekombinanten Proteinen**

### ***Das bakterielle Expressionssystem***

Bei der Strukturaufklärung von Proteinen und Proteinkomplexen aus endogener Herkunft stellen besonders die Heterogenität (beispielsweise durch Glykosilierungen) und ihre geringe Abundanz enorme Hürden dar. Daher werden die Zielproteine bevorzugt unter Verwendung heterologer Expressionssysteme etabliert. Hierbei hat sich das bakterielle Überexpressionssystem bewährt, wobei zumeist Laborstämme von *Escherichia Coli* Anwendung finden. Besonders die einfache genetische Manipulation der Bakterien, die verlässliche Kontrolle der Genexpression, sowie die vergleichsweise preiswerten Medien sind Gründe für dessen weitreichende Anwendung (Gopal and Kumar, 2013).

### ***Herstellung kompetenter Zellen und Transformation***

Die kodierende Gensequenz wird mit Hilfe eines extrachromosomalen zirkulären DNA-Molekül, dem Expressionsvektor in die bakteriellen Zellen eingebracht. Bei diesem als Transformation bezeichneten Vorgang, wird die bakterielle Fähigkeit zur spontanen DNA-Aufnahme genutzt. Eine Möglichkeit, die Bakterien für die Aufnahme von DNA kompetent zu machen, ist die Rubidium-Chlorid-Methode (Hanahan, 1983). Dafür wurden in dieser Arbeit 500 ml LB-Medium mit 1 ml einer 10 ml Übernacht-Kultur (ÜN-Kultur) angeimpft und bis zu einer optischen Dichte von 0,5 bei 310K geschüttelt. Die Zellsuspension wurde für 15 Minuten auf Eis gekühlt und anschließend abzentrifugiert (277K, 4000 g, 5 Minuten). Das Pellet wurde in 0,4 Volumen eisgekühltem Transformationspuffer Tfb I resuspendiert, 15 Minuten auf Eis gestellt und erneut zentrifugiert (277K, 4000 g, 5 Minuten). Der Zellschlag wurde in 0,04 Volumen eiskaltem Tfb II aufgenommen, 15 Minuten auf Eis gestellt, in 100 µl Aliquots aufgeteilt und entweder sofort verwendet oder in flüssigem Stickstoff schockgefroren und bei 193K für eine spätere Verwendung gelagert.

Um das gewünschte Plasmid in die kompetenten Bakterienzellen einzuschleusen, wurden 100 µl der vorbereiteten Zellen auf Eis aufgetaut und mit 50 –100 ng der Plasmid-DNA eine halbe Stunde auf Eis inkubiert. Darauf folgte eine Hitzeschockbehandlung bei 315K für 45 Sekunden im temperierten Wasserbad gefolgt von einer weiteren Minute auf Eis. Nach Zugabe von 900 µl LB-Medium wurde der

Ansatz für 1 Stunde bei 310K geschüttelt. Abschließend wurde der Ansatz nochmals zentrifugiert, der Überstand bis auf 100 µl abgenommen, das Zellpellet resuspendiert und anschließend auf einer LB-Agaroseplatte mit entsprechenden Antibiotika zur Selektion ausplattiert. Die Platten wurden über Nacht bei 310K inkubiert.

#### **Tfb I**

30 mM KOH

100 mM Rubidiumchlorid

10 mM Calciumchlorid

50 mM Manganchlorid

15% (v/v) Glycerin

pH 5,8 mit 0,1 M Essigsäure

#### **Tfb II**

75 mM Calciumchlorid

10 mM Rubidiumchlorid

15% Glycerin

pH auf 6,5 mit 0,1 M NaOH

### ***Proteinexpression***

Bei den verwendeten bakteriellen Expressionsstämmen handelt es sich in der Regel um (DE3) Derivate, welche die T7 virale RNA Polymerase besitzen, die sich durch hohe Transkriptionsraten auszeichnet und nicht mit der Wirtsgensynthese interferiert. Der häufig verwendete *E. coli* BL21-Stamm zeichnet sich zusätzlich durch die Defizienz einiger Proteasen aus, sodass die Gefahr einer proteolytischen Degradierung des Zielproteins verringert wird. Der Großteil der Expressionsvektoren verwendeten zur Kontrolle der Genexpression den Lac Promotor (Jacob and Monod, 1961), sodass die Expression durch Zugabe von Isopropylthio-Galaktosidase (IPTG) aktiviert werden kann. Dieses Laktose-Analogon lässt sich nicht metabolisieren, wodurch eine konstante Konzentration des Induktors und somit eine durchgängige Aktivierung des Promotors gewährleistet wird. Im Labormaßstab werden die Bakterienkulturen zumeist in Erlenmeyer-Kolben angesetzt, die in einem temperierbaren Schüttelinkubator inkubiert werden. Das bakterielle Expressionssystem ist im Vergleich zur eukaryotischen Zellkultur relativ simpel und reproduzierbare Ergebnisse sind bereits durch die Regulation weniger Parameter möglich. Hierbei sollte insbesondere die Kolbentemperatur zum Zeitpunkt der Induktion, als auch während der Expression kontrolliert werden. Die Induktion der Proteinexpression erfolgte stets während der logarithmischen Wachstumsphase der Bakterien, was

durch die Bestimmung der Zelldichte mit einem Novaspec II-Spektralphotometer kontrolliert wurde. Die Induktion erfolgte in der Regel bei einer, an diesem Photometer bestimmten,  $OD_{600}$  von 0,8 - 1,0 und einer Kolbentemperatur von 291K.

Eine gute Kenntnis des biologischen Systems bei der rekombinanten Erzeugung von Proteinen kann ebenfalls von Bedeutung sein. Besonders wenn das Zielprotein vornehmlich als Bestandteil makromolekularer Komplexe existiert, ist eine isolierte Expression eines solchen Proteins möglicherweise problematisch. Daher existieren vielfältige Systeme zur simultanen Co-Expression bis hin zur *in vivo* Rekonstitution vollständiger rekombinanter und heterogener Proteinkomplexe. Hierfür kann die Expression der unterschiedlichen Untereinheiten durch Co-Transformation mehrerer Expressionsvektoren mit unterschiedlichen Selektionsmarkern (und unterschiedlichen Replikationsursprüngen) erfolgen. Alternativ ist auch die Verwendung von multicistronischen Vektoren oder ein Vektoren mit mehreren Promotoren möglich. Die Expression von SmD1 und SmD2 sowie SmE, Sm und SmG erfolgte über zwei multicistronische Vektoren (Kambach et al., 1999), sodass sie als die beiden Heterooligomere SmD1-D2 und SmE-F-G präpariert werden konnten.

### ***Herstellung der Expressionsvektoren***

Das Einbringen kodierender Sequenzen in die entsprechenden Leervektoren erfolgte mit Hilfe der Restriktionsklonierung in die hierfür vorgesehenen multiplen Klonierungsstellen der verwendeten Vektoren. Der enzymatischen Verdau mit kommerziellen Restriktionsendonukleasen wurde mit 10 Units für 1  $\mu$ g DNA (Plasmidpräparation) beziehungsweise 1 U für 1  $\mu$ g DNA (Insert) in den empfohlenen Reaktionspuffern durchgeführt. Die Reaktion wurde entweder durch Zugabe von 6x DNA-Probenpuffer für anschließende Auftrennung in einem Agarosegel (Insert) oder durch Dephosphorylierung mit alkaliner Phosphatase (Plasmidpräparation) gestoppt. Anschließend erfolgte die Ligation der DNA-Fragmente in entsprechend linearisierte Vektoren mit kompatiblen DNA-Überhängen (*sticky ends*). Dabei bildet z.B. die T4-DNA-Ligase die Phosphodiesterbindung zwischen der 5'-Phosphat und der 3'-Hydroxylgruppe. Für solche Reaktionen wurde die Plasmid-DNA mit einem dreifach molaren Überschuss des Inserts sowie mit T4-DNA-Ligase in einem geeigneten Puffersystem versetzt. Nach Inkubation bei 289K wurde der Ansatz in chemisch kompetente Zellen transformiert. Die Vektoren wurden durch die Firma GATC GmbH (Konstanz, Deutschland) mit Hilfe der Didesoxy-Kettenabbruch-Sequenzierung (Sanger et al., 1977) verifiziert.

## **Molekularbiologische und bioanalytische Methoden**

### ***Polymerase-Kettenreaktion (PCR)***

Die Vervielfältigung von dsDNA-Fragmenten erfolgte mit der PCR-Methode (Mullis et al., 1986). Zu der als Vorlage dienenden *Template*-DNA werden zwei synthetische Oligonukleotid-Primer gegeben, deren Sequenzen zu den 5'-Enden der beiden Stränge des zu amplifizierenden Sequenzabschnitts

komplementär sind. Die Reaktion wird mit Hilfe von hitzestabilen Polymerasen in geeignetem Puffer katalysiert. Für die gezielte Einführung von Punktmutationen wurde die *QuickChange-Site-Directed-Mutagenesis*-Methode angewandt (Carter, 1986). Hierfür wurden Oligonukleotid-Primer verwendet, die im Bereich der zu erzeugenden Punktmutation für die gewünschte Aminosäure kodieren und sich somit von der Sequenz der *Template*-DNA unterscheiden. Bei dieser Methode kommt es zur Neusynthese des gesamten Vektors durch die Polymerase und es erfolgt keine Amplifikation der Zielsequenz. Die unmutierte *Template*-DNA hat daher einen signifikanten Anteil an der Gesamt-DNA und es empfiehlt sich diese gezielt zu entfernen. Da die *Template* DNA in der Regel bakterieller Herkunft ist und somit Dam methyliert vorliegt, kann diese im Gegensatz zur neusynthetisierten DNA spezifisch durch einen Verdau mit dem Restriktionsenzym *Dpn1* degradiert werden. Die DNA-Präparation wird anschließend in kompetente Zellen transformiert und durch die Sanger-Sequenzierung validiert (siehe Abschnitt 4.3, Seite 22). Für die Deletion von Sequenzabschnitten wurde die *overlapping-extension PCR*-Methode verwendet (Bryksin and Matsumura, 2010). Es erfolgt der Einsatz von Oligonukleotid-Primern, die für die jeweiligen 5' und 3' flankierenden Bereiche des zu deletierenden Abschnitts, jedoch nicht für diesen selbst kodieren. Die PCR-Reaktion wurde in zwei Teilschritten durchgeführt, sodass die 5'- und 3'-seitigen PCR-Produkte zuerst in getrennten Reaktionen erzeugt wurden. Anschließend wurden die PCR-Produkte des ersten Schritts vereinigt und durch eine weitere PCR-Reaktion vervollständigt. Da beide PCR-Produkte im Bereich der Oligonukleotid-Primer komplementär überlappen entsteht hierbei ein Gesamtprodukt, welches nicht mehr für den zu deletierenden Bereich des Proteins kodiert. Dieses kann anschließend durch Restriktionsenzyme verdaut und in Leervektoren ligiert werden.

### ***Diskontinuierliche Polyacrylamid-Gelelektrophorese (SDS-PAGE)***

Bei der diskontinuierlichen und denaturierenden SDS-PAGE handelt es sich um eine Variante der Gelelektrophorese, bei welcher die Analyten innerhalb eines Polyacrylamid-Gels im elektrischen Feld aufgetrennt werden (Laemmli, 1970). Die Probe wird vor der Analyse durch Erhitzen bei 368K im Probenpuffer denaturiert. Hierbei kommt es mit Hilfe des enthaltenen  $\beta$ -Mercaptoethanol zur Reduktion von Cystinen, während sich das SDS (Natriumlaurylsulfat) an die Proteine anlagert. Dies geschieht ungefähr im Verhältnis 1,4 mg SDS auf ein 1 mg Protein. Bei SDS handelt es sich um ein anionisches Tensid, sodass die Analyten durch seine Bindung eine negative Ladung tragen, die mit ihrem Molekulargewicht korreliert. In der SDS-PAGE erfolgt die Trennung der Probe daher im Endeffekt nach dem Molekulargewicht der jeweiligen Proteine.

Die Gele wurden nach dem unten genannten Rezept hergestellt und die Elektrophorese wurde in einer Bio-Rad Mini-PROTEAN Gelkammer mit einer Begrenzung der Stromleistung auf 15 Watt pro Gel durchgeführt. Zur Auftrennung der Sm-Proteine wurden Trenngele mit einem Acrylamid-Anteil

von 12-15% und hoher TEMED Konzentration hergestellt, da dies die Trenneigenschaften für kleinere Proteine verbessert (Lehmeier et al., 1990). Die Gellösung wurde auf Vorrat präpariert, bei 280K gelagert und die Quervernetzung wurde durch Zugabe von APS und TEMED bei Bedarf gestartet. Zur gezielten Auftrennung von SmE-F-G wurde, neben dem standardmäßigen Lämmli-Puffersystem auf TRIS-Glycin-Basis, zusätzlich das TRIS-Tricin-Puffersystem nach Schägger angewendet (Schagger and von Jagow, 1987). Dieses ermöglicht im Bereich zwischen 1-50 kDa bessere Trennergebnisse als das TRIS-Glycin-Puffersystem. In diesem zeigt jedoch besonders SmD1 ein ungewöhnliches Laufverhalten, sodass zumeist beide Gelsysteme parallel verwendet wurden.

#### **SDS-Proteinprobenpuffer**

60 mM TRIS-HCl pH 6.8

2% SDS

10% Glycerin

5%  $\beta$ -Mercaptoethanol

0.01% Bromphenolblau

#### **10x Lämmli-Puffer**

0,25 M TRIS

1,92 M Glycin

1% SDS

pH 8,4

#### **50 ml High-TEMED Sammelgel (5%)**

8,5 ml Acrylamid-/ Bisacrylamidlösung (37,5:1 A:BA)

13 ml 0,5 M TRIS-HCl pH 6,8

28,5 ml ddH<sub>2</sub>O

500  $\mu$ l 10% SDS

22,5  $\mu$ l APS (pro Gel; 4 ml)

11  $\mu$ l TEMED (pro Gel; 4ml)

#### **100 ml High-TEMED Trenngel (12 oder 15%)**

33,3 ml bzw. 50 ml 30% Acrylamid-/ Bisacrylamidlösung (37,5:1 A:BA)

375 mM TRIS pH 8,8

ad 100 ml ddH<sub>2</sub>O

30  $\mu$ l TEMED (pro Gel; 10 ml)

30 µl APS (pro Gel; 10 ml)

#### **10x Kathodenpuffer TRIS-Tricine Gele**

1,0M TRIS

1,0M Tricin

1% SDS, pH 8,25

#### **10x Anodenpuffer TRIS-Tricin Gele**

2,0M TRIS, pH 8,9

TRIS-Tricin Gelpuffer

3M TRIS-HCl pH 8,45

0,3% SDS

#### **TRIS-Tricin (3%); 1 Gel**

0,25 mL 30% Acrylamid-/ Bisacrylamidlösung (28,25:1 A:BA)

3,15 ml ddH<sub>2</sub>O

1,25 ml TRIS-Tricin Gelpuffer

25 µl 10% SDS

25 µl APS

12,5 µl TEMED

#### **TRIS-Tricine Trenngel (15%); 1 Gel**

6 ml 30% Acrylamid-/ Bisacrylamidlösung (28,25:1 A:BA)

2 ml ddH<sub>2</sub>O

4 ml TRIS-Tricin Gelpuffer

60 µl 10% SDS

25 µl APS

12,5 µl TEMED

### ***Quantitative Proteinassays***

Für die quantitative Analyse von Proteinlösungen kamen mit dem Bradford- (Bradford, 1976) und dem Biuret-Test (Gornall et al., 1949) zwei unterschiedliche photometrische Methoden zum Einsatz. Der Bradford-Test (Bradford, 1976) beruht auf dem Farbstoff Coomassie-Brillantblau G-250, der unter sauren Bedingungen kationische Proteinseitenketten bindet. Durch die Komplexierung verschiebt sich das Absorptionsmaximum des Farbstoffs von 470 nm auf 595 nm, sodass die

Proteinkonzentration bei dieser Wellenlänge photometrisch bestimmt werden kann. Hierbei entspricht die Absorption eines 0,1 prozentigen Proteingemisches einer  $OD_{595}$  von 1,0, wenn 20  $\mu$ l auf 1 ml Testlösung eingesetzt werden. Da die Reaktion sequenzabhängig ist, sollte für eine genaue Konzentrationsbestimmung von reinen Lösungen eines bestimmten Proteins eine Kalibrierung erfolgen. Wegen der Sequenzabhängigkeit sind Kalibrierungen mit Standardproteinen nicht geeignet. Vielmehr sind hochgereinigte Präparationen des Zielproteins notwendig und die jeweilige Kalibrierung ist auch nur für hochgereinigte Präparation gültig. Bei diesen hochreinen Proben ist jedoch alternativ die photometrische Bestimmung der Proteinkonzentration durch eine direkte Messung der UV-Absorption bei 280 nm sinnvoller. Die Bradford-Testlösung wurde kommerziell von Bio-Rad als 5x Konzentrat erworben.

Der Biuret-Test hingegen ist sequenzunabhängig, da eine Komplexierung mit dem Peptid-Rückgrat stattfindet. Die Reaktion findet im wässrig-alkalischen Milieu statt und beruht auf der Komplexierung zweiwertiger Kupferionen durch die Amidgruppe.

#### **Biuret-Testlösung**

3% (w/v) NaOH

0,15% (w/v)  $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$

0,6% (w/v) Natriumkaliumtartrat

#### **Bio-Rad 5x Protein Assay**

Bio-Rad, München, Deutschland

## **Chromatographische Methoden zur Präparation von rekombinanten Proteinen**

### ***Strategie***

Die Isolation eines bestimmten Zielproteins aus zellulären Lysaten erfordert die Ausnutzung der spezifischen physikochemischen Eigenschaften, um innerhalb möglichst weniger Schritte eine reine Präparation des Zielproteins zu erhalten. Hierfür existiert eine Vielzahl, meist flüssigchromatographischer Methoden, die eine flexible und individuelle Entwicklung von Reinigungsprotokollen ermöglichen. Eine Minimierung an Arbeitsschritten ist sowohl aus Gründen des Zeitmanagements, als auch wegen der begrenzten Rückgewinnungsraten eines jeden Reinigungsschritts empfehlenswert. Bereits eine sehr gute Rückgewinnungsrate von 90% pro Schritt führt nach 5 Schritten zu einer mäßigen Gesamtausbeute von weniger als 60%. Als Grundgerüst bei der Planung eines jeden Protokolls empfiehlt sich die CIPP- (*Capture-Intermediate-Polishing-Purification*) Strategie. Die verschiedenen Reinigungsphasen sind mit unterschiedlichen Zielsetzungen verbunden und erfordern daher auch unterschiedliche Eigenschaften der einzusetzenden chromatographischen Materialien. Hierbei lassen sich Kapazität, Geschwindigkeit, Rückgewinnung und die Auflösung als wichtige Größen nennen. Die initiale Isolation des Zielproteins aus dem

zellulären Lysat sollte mit einer spezifischen Bindung an das Säulenmaterial starten (Capture). Besonders affinitätschromatographische Methoden und grobporige Ionenaustauscher mit hoher Kapazität, (Fluss)-Geschwindigkeit und Robustheit sind hierfür geeignet. Es müssen hinsichtlich der Auflösung im Capture-Schritt Abstriche gemacht werden, da diese bei chromatographischen Materialien in erster Linie durch eine gleichmäßige Packung besonders feinporiger Materialien erreicht wird. Das Ziel der anschließenden intermediären Reinigung ist die möglichst effektive Abtrennung großer Mengen an kontaminierenden Partikeln. Die Anwendbarkeit bestimmter chromatographischer Methoden ist sowohl von spezifischen Eigenschaften des Zielproteins als auch von den Pufferbedingungen abhängig. Daher sollten einzelne Schritte konsequent aufeinander abgestimmt sein, um zeitintensive Schritte wie Pufferwechsel durch Dialyse oder Ultrafiltration zu vermeiden. Die Polishing-Phase stellt die finale Reinigungsstufe dar, bei welcher verbleibende Spuren von Verunreinigungen abgetrennt werden sollen. Die Kontaminationen weisen aufgrund ihrer Co-Reinigung mit dem Zielprotein oftmals ähnliche Eigenschaften wie dieses auf. Deshalb sollten die Materialien in dieser abschließenden Phase nach einer bestmöglichen Auflösung ausgewählt werden.

### ***Affinitätschromatographie***

Bei der Affinitätschromatographie handelt es sich um ein sehr effektives bindendes Trennverfahren, welches auf der spezifischen Interaktion des Analyten (Protein) mit einem kovalent ans Trägermaterial gekoppelten Liganden beruht. Die Methoden können auf der Immunoaffinität zwischen Antikörpern und Antigenen, Enzym-Substrat Bindungen, oder Rezeptor-Liganden Interaktionen beruhen. Für die Reinigung rekombinanter Proteine können diese gezielt mit einer Affinitätsmarkierung versehen werden. Hierbei kann es sich entweder um Fusionspartner oder kurze Sequenzmotiv-Markierungen, wie beispielsweise der Hexahistidin-Markierung handeln. Letztere zeigt eine Affinität zu Metallionen, welche über an die Trägermatrix gebundene Chelatoren komplexiert und immobilisiert sind. Die Bindung kann über Säulen oder im Batch-Verfahren erfolgen. Die Elution erfolgt im Normalfall durch die Zugabe von kompetitiven Substanzen, die den Analyten von der Säule verdrängen.

### ***Ionenaustauscher-Chromatographie***

Die Auftrennung erfolgt bei der Ionenaustauscher-Chromatographie mit Hilfe positiv oder negativ geladener Materialien. Die Bindung erfolgt aufgrund der Oberflächenladung des Proteins und ist nicht unbedingt durch die Gesamtladung, sondern durch die Verteilung der geladenen Gruppen auf der Proteinoberfläche definiert. Da Proteine zu den amphoterer Molekülen zählen ist ihre Ladung in pH-Abhängigkeit titrierbar, wodurch ihre Bindungseigenschaften an Ionenaustauscher-Materialien durch eine Veränderung des pH beeinflussbar sind. Dies macht unter Umständen ein Screening unterschiedlicher pH-Bedingungen bei der Durchführung notwendig. Die Bindung an die Säule erfolgt

im Regelfall unter Niedrigsalz-Bedingungen, während die Elution in einem Gradienten steigender Salzkonzentrationen erfolgt, um eine Auftrennung unterschiedlich stark gebundener Analyten zu ermöglichen.

### ***CHT-Chromatographie***

CHT steht für Hydroxylapatit-Keramik (ceramic Hydroxyapatite), bei dem es sich um eine natürliche Mineralform von Kalziumapatit mit einer hohen Anzahl an Hydroxylgruppen handelt. Die chemische Formel ist  $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$ , sodass die funktionellen Gruppen sowohl Paare von positiv geladenen ( $\text{Ca}^{2+}$ , C-sites) als auch Bündel von negativ geladenen Gruppen ( $\text{PO}_4^-$  und  $\text{OH}^-$ , P-sites) umfassen. Die Bindung beruht in erster Näherung auf elektrostatischen Wechselwirkungen und ist daher vergleichbar mit Ionenaustauschern. Besonders die Bindung von C-sites durch Carboxyl- und Phosphorylfunktionen beruht jedoch zusätzlich auf der Ausbildung von Koordinationskomplexen mit deutlich erhöhter Bindungsenergie. Dass es sich bei dieser chromatographischen Methode nicht um eine reine Ionenaustauscher-Chromatographie handelt zeigt sich auch dadurch, dass die Bindekapazität des Materials beispielsweise für azide Proteine bei einer Erhöhung des pH sinkt. Die Elution basischer Analyten erfolgt durch steigende Konzentrationen von Chlorid-, Sulfat- oder Phosphationen. Azide Analyten können hingegen nicht durch Natriumchlorid eluiert werden. Dies ermöglicht die klare Trennung von basischen und aziden Analyten und ist besonders effektiv in Hinsicht auf die Abtrennung kontaminierender Nukleinsäuren von basischen Zielproteinen.

### ***Größenausschlusschromatographie***

Bei der Größenausschlusschromatographie (size-exclusion chromatography, SEC) beruht die Trennung auf dem hydrodynamischen Radius der Analyten und es erfolgt im Gegensatz zu den vorher genannten chromatographischen Trennmethoden keine Bindung an das Säulenmaterial. Das Säulenmaterial ist eine poröse Matrix aus gleichgroßen kugelförmigen Partikeln, die ein *gepacktes Bett* ausbilden. Große Moleküle können nicht in die Poren eintreten und eluieren im Ausschlussvolumen. Kleinere Partikel können jedoch in die Poren eintreten, sodass diese später eluieren. Hierfür sollte das Säulenmaterial ungeladen und inert sein, sodass Interaktionen des Analyten mit dem Säulenmaterial vermieden werden. Für eine effektive Trennung darf das Probenvolumen maximal 0,5 bis 4 Prozent des Säulenvolumens betragen, sodass eine Einengung des Volumens bspw. durch Ultrafiltration notwendig ist. Da die Säulen aufgrund ihrer Porengröße und Packung anfällig für ein Verstopfen durch unlösliche Partikel und hochmolekulare Aggregate sind, eignet sich die SEC als abschließender Schritt von Reinigungsprotokollen in der Polishing-Phase.

## **Protokolle zur Präparation der 6S- und 8S-Komplexuntereinheiten**

### ***Präparation der Sm-Proteine***

Die Präparation der Subkomplexe SmD1-D2 und SmE-F-G erfolgte nach den Protokollen von Christian Kambach (Kambach, C. *et al.*, 1999).

### ***Präparation der pICln-Konstrukte***

*Drosophila Melanogaster* pICln cDNA (Genbank AF216522) wurde in den pETM13-Expressionsvektor transferiert. Sequenzmanipulationen zur Erzeugung unterschiedlicher Varianten und Fragmente von pICln entstanden auf Basis dieses Vektors mit Hilfe der *QuickChange-Site-Directed-Mutagenesis* Methode (Punktmutationen) (Carter, 1986) bzw. unter Anwendung der *overlapping-extension-PCR* (Deletionen) (Bryksin and Matsumura, 2010) (siehe Abschnitt 4.3, Seite 24). Alle Proteinvarianten enthielten C-terminal ein Hexahistidin-Sequenzmotiv, die hierfür kodierende Sequenz wurde über den jeweiligen 3'-Oligonukleotid-Primer während der PCR eingebracht. Alle pICln-Varianten wurden in BL21(DE3)pLysSRARE Zellen exprimiert. Die Expression wurde durch Zugabe von 0,5 mM IPTG induziert und erfolgte für 15 Stunden bei 290K. Die Ernte der Bakterienkulturen erfolgte per Zentrifugation (277K, 5000 g, 15 Minuten) in deren Anschluss die Pellets in einem reduziertem Volumen (25 ml pro 1 l Kultur) Lysepuffer resuspendiert wurden. Die Lyse erfolgte per Sonifikation und die Klärung des Zelllysats erfolgte per Ultrazentrifugation (277K, 45.000 rpm, 45 Minuten, 45Ti, Beckman Coulter). Die anschließende Affinitätschromatographie erfolgte im Batch-Verfahren, sodass das geklärte Lysat für 30 Minuten mit Ni-NTA-Agarose inkubiert wurde. Im Anschluss wurde die Ni-NTA-Agarose mit Hilfe einer, mit einem Frittenfilter ausgestatteten, Glassäule gesammelt und vor der Elution mit 20 Bettvolumen Waschpuffer gewaschen. Das Eluat wurde durch Ultrafiltration aufkonzentriert und durch eine SEC unter Verwendung einer Superdex 75-Säule abschließend gereinigt. Die Proteinpräparationen wurden im Anschluss durch Ultrafiltration mit Centricons auf 30 mg/ml aufkonzentriert und bei 193K gelagert.

### **Lysepuffer**

300 mM NaCl

50 mM HEPES pH 7,5

15 mM Imidazol

2 mM  $\beta$ -Mercaptoethanol

10% (v/v) Glycerin

### **Waschpuffer**

150 mM NaCl

50 mM HEPES pH 7,5

20 mM Imidazol

2 mM  $\beta$ -Mercaptoethanol

### **Elutionspuffer**

150 mM NaCl

50 mM HEPES pH 7,5

200 mM Imidazol

2 mM  $\beta$ -Mercaptoethanol

### **Gelfiltrationspuffer/ Aufbewahrungspuffer**

150 mM NaCl

10 mM HEPES pH 7,5

5 mM DTT

### ***Präparation von SMN $\Delta$ C-Gemin2***

Die kodierenden Sequenzen für SMN $\Delta$ C (1-122) und Gemin2 wurden in pETM30-Vektoren subkloniert und als N-terminale Hexahistidin-Glutathion-S-Transferase (GST)-Fusionen in BL21(DE3)pLysSRARE Zellen exprimiert. Die Expression wurde durch Zugabe von 0,5 mM IPTG induziert und erfolgte für 15 Stunden bei 290K. Nach der Ernte wurden die Bakterienpellets gemischt, in IMAC-A Puffer resuspendiert und durch Sonifikation ein Co-Lysat erzeugt. Die anschließende Affinitätschromatographie erfolgte im Batch-Verfahren, sodass das geklärte Lysat für 60 Minuten mit Ni-NTA-Agarose inkubiert wurde. Im Anschluss wurde die Ni-NTA-Agarose mit Hilfe einer, mit einem Frittenfilter ausgestatten, Glassäule gesammelt und vor der Elution mit 20 Bettvolumen IMAC-Waschpuffer gewaschen. Die Elution erfolgte mit IMAC-Waschpuffer, der mit 200 mM Imidazol versetzt wurde. Das Eluat wurde anschließend mit GSH-Sepharose versetzt und ebenfalls für 60 Minuten inkubiert. Anschließend wurde die GSH-Sepharose wiederum in einer Glassäule gesammelt und mit 20 Bettvolumen GSH-A Puffer gewaschen. Die Elution erfolgte mit GSH-B. Die Affinitätsmarkierung wurde durch proteolytischen Verdau mit TEV Protease abgespalten und durch erneute Inkubation mit Ni-NTA Agarose aus der Proteinlösung entfernt. Die abschließende Reinigung erfolgte durch eine SEC mit einer Superdex 200 Säule. Die Fraktionen wurden per SDS-PAGE analysiert und ein Pool wurde aus allen Fraktionen gebildet, die stöchiometrische Mengen an

SMNΔC und Gemin2 enthielten. Diese wurden durch Ultrafiltration unter Verwendung von Centricons auf 30mg/ml aufkonzentriert und für die weitere Verwendung bei 193K gelagert.

#### **IMAC-A Puffer**

500 mM NaCl

20 mM TRIS-HCL pH 8

10 mM Imidazol

10% Glycerin (v/v)

2 mM β-Mercaptoethanol

#### **IMAC Waschpuffer**

200 mM NaCl

20 mM TRIS-HCL pH 8

20 mM Imidazol

10% Glycerin (v/v)

2 mM β-Mercaptoethanol

#### **IMAC-B Puffer**

500 mM NaCl

20 mM TRIS-HCL pH 8

250 mM Imidazol

10% Glycerin (v/v)

2 mM β-Mercaptoethanol

#### **GSH-A**

150 mM NaCl

20 mM HEPES pH 7,5

5% Sucrose (w/v)

2 mM EDTA

2 mM β-Mercaptoethanol

#### **GSH-B**

150 mM NaCl

20 mM HEPES pH 7,5

5% Sucrose (w/v)

2 mM  $\beta$ -Mercaptoethanol

2 mM EDTA

15 mM reduziertes Glutathione

### **SEC Puffer/ Aufbewahrungspuffer**

150 mM NaCl

10 mM HEPES pH 7,5

2 mM TCEP

2 mM EDTA

### **Präparation des 6S- und 8S-Komplexes**

Die gereinigten Proteine SmD1-D2, SmE-F-G und pICln wurden in eben dieser Reihenfolge in äquimolaren Mengen unter Hochsalzbedingungen in 4x Rekonstitutionspuffer gemischt, sodass dieser im Gesamtvolumen einfach konzentriert war. Die hohe Salzkonzentration von 1 M NaCl war notwendig, da es bei physiologischen Bedingungen aufgrund der starken entgegengesetzten Ladungen der Proteine, insbesondere bei Zugabe von pICln zu SmD1-D2, zu einer spontanen Aggregation kommt, die nur teilweise reversibel ist. Nach einer Inkubationszeit von 30 Minuten auf Eis erfolgte die Dialyse in drei Schritten von jeweils 3 Stunden (1 M  $\rightarrow$  750 mM  $\rightarrow$  500 mM  $\rightarrow$  200 mM) hin zu Niedrigsalzbedingungen. Anschließend erfolgte eine SEC mit Hilfe einer Superdex 200-Säule mit anschließender Auswertung auf einem 15 prozentigen SDS-PAGE-Gel. Fraktionen, die den vollständig ausgebildeten 6S-Komplex beinhaltenen wurden gepoolt und eine Proteinkonzentration von 15 g/l wurde durch Ultrafiltration mit Hilfe von Centricon-Filtereinheiten eingestellt, bevor die Präparation bei 193K gelagert wurde. Ein typisches Chromatogramm der SEC und das dazu gehörige SDS-PAGE Gel ist in (Abbildung 20, Seite 65) zu sehen. Die Rekonstitution des 8S-Komplexes erfolgte durch Zugabe äquimolarer Mengen von SMN $\Delta$ C-Gemin2 zu diesen präparierten 6S-Präparation. Nach einer Inkubation von 30 Minuten bei 280K erfolgte anschließend eine SEC mit einer Superdex 200-Säule und eine Analyse der Fraktionen mit Hilfe von 15 prozentigen SDS-PAGE-Gelen. Fraktionen, die stöchiometrische Mengen aller acht Untereinheiten enthielten wurden gepoolt, mit Hilfe von Centricon-Filtereinheiten aufkonzentriert und für die weitere Verwendung mit einer Konzentration von 15 mg/ml bei 193K aufbewahrt. Abbildung 20 auf Seite 65 zeigt exemplarisch ein typisches Chromatogramm und das dazugehörige SDS-PAGE Gel.

#### **4x Rekonstitutionspuffer**

4 M NaCl

80 mM HEPES pH 7,5

20 mM TCEP pH 7,5

#### **Dialysepuffer**

750 -> 500 -> 200 mM NaCl

20 mM HEPES pH 7,5

0,5 mM TCEP pH 7,5

10% Glycerin (v/v)

#### **Gelfiltrationspuffer/ Aufbewahrungspuffer**

150 mM NaCl

10 mM HEPES pH 7,5

2 mM TCEP

### **Röntgen-Strukturaufklärung von Proteinen**

#### ***Kristallisation von Proteinen***

Die Röntgenkristallographie liefert bis heute mit über 90% die weitaus größte Anzahl von Proteinstrukturen (PDB statistics, März 2015). Die Methode ist aber naturgemäß von der Verfügbarkeit von Kristallen des jeweiligen Proteins abhängig. Hierbei ist die Kristallisation von Proteinen, besonders durch deren Anfälligkeit für Denaturierung methodisch im Vergleich zur Kristallisation kleiner Moleküle eingeschränkter, beruht aber natürlich auf den gleichen physikalischen Grundsätzen. Erreicht die Löslichkeit des Proteins den Punkt der Übersättigung, bleibt dieses nicht länger in Lösung, sondern es präzipitiert. Die Löslichkeit hängt beispielsweise von Faktoren wie Proteinkonzentration, Ionenstärke, Temperatur und pH ab. Beim Vorgang der Präzipitation kann ein Kristall, aber auch ein amorpher und eventuell denaturierter Feststoff gebildet werden. Weder die Kristallkontakte auf denen das Kristallgitter beruht, noch die Bedingungen unter denen es zum Kristallwachstum kommt, lassen sich bis dato vorhersagen. Deshalb repräsentiert die Kristallisation immer noch den zentralen Flaschenhals von Projekten zur Röntgenstrukturaufklärung. Der Kristallisationsprozess wird in Form eines auf Versuch und Irrtum basierenden Screenings durchgeführt. Dessen Erfolg ist von vielen Variablen, in erster Linie aber von der Natur des Zielproteins abhängig (Claverie et al., 2002; Ding et al., 2002; Sulzenbacher et al., 2002). Die gängigste Kristallisationsmethode ist die Dampfdiffusion, die alternativ nach der Methode des

hängenden oder sitzenden Tropfen durchgeführt werden kann. Hierbei wird die Proteinlösung mit einem Präzipitant gemischt und als Tropfen im Kopfraum einer Vertiefung abgesetzt, die ihrerseits mit einem größeren Volumen der Präzipitant gefüllt ist. Aufgrund der höheren Präzipitant-Konzentration der Mutterlauge verliert der Proteintropfen durch Diffusion Wasser. Dadurch steigen die Konzentrationen von Protein und Präzipitant im Proteintropfen langsam an. Im Idealfall wird langsam der Punkt der Übersättigung erreicht und es kommt zur Kristallkeimung. Durch die Kristallisation des Proteins sinkt die Proteinkonzentration in der Lösung langsam, wodurch der Zustand der Übersättigung aufgehoben wird. Dadurch wird die Keimung weiterer Kristalle verhindert, während das Wachstum vorhandener Kristalle vorangetrieben wird.

### ***Thermodynamische Basis des Kristallwachstums***

Thermodynamisch gesehen wird das Kristallwachstum durch kleine negative Veränderungen der Gibbs freien Energie angetrieben und beruht in erster Linie auf Änderungen der Entropie (Gliko et al., 2005; Paunov et al., 2001; Yau et al., 2000), während enthalpische Effekte eher vernachlässigt werden können (Vekilov, 2003; Vekilov et al., 2002). Die Anordnung von Proteinen innerhalb eines dreidimensionalen Kristallgitters führt durch einen Verlust an Dreh- und Translations-Freiheitsgraden zu einer Entropieabnahme, während die Freisetzung von koordinierten Lösemittelmolekülen zu einer Entropiesteigerung führt. Besonders der Einbau intrinsisch ungeordneter Proteinregionen leistet einen großen Beitrag zum Entropieverlust, weshalb die Präsenz ausgedehnter Termini oder Schleifenregionen die Kristallisation verhindern kann (Derewenda, 2010). Durch eine statistische Untersuchung der Sequenzen gelöster Kristallstrukturen zeigte sich außerdem, dass die Neigung eines Proteins zu kristallisieren negativ mit der Anzahl langer polarer Aminosäure-Seitenketten korreliert ist (Price 2009). Eine ähnliche Studie konnte analog zeigen, dass in Kristallkontakten kleine hydrophobe Aminosäuren bevorzugt vorkommen (Cieslik and Derewenda 2009).

In dieser Arbeit kamen unterschiedliche Strategien zur gezielten Optimierung des Kristallisationsverhaltens der Proteine zur Anwendung. Diese beschränkten sich auf pICln, da dieses Bestandteil sowohl des 8S- als auch des 6S-Komplexes war und zu Beginn der Arbeiten eine NMR-Struktur (Furst et al., 2005) des pICln-Orthologs von *Canis lupus familiaris* verfügbar war. Auf Basis dieser NMR-Struktur war es möglich ein Homologie-Modell des verwendeten *Drosophila melanogaster* pICln-Orthologs zu generieren. Auf dessen Basis wurden gezielt Deletionen von flexiblen Protein-Bereichen durchgeführt. Diese befanden sich hauptsächlich innerhalb des C-Terminus und der Schleifenregion zwischen den  $\beta$ -Strängen  $\beta_6$  und  $\beta_7$  von pICln. Außerdem wurden einige Aminosäuren auf der Oberfläche von pICln durch Alanin substituiert, um die Oberflächenentropie zu verringern. Alle auf diese Weise modifizierten Varianten von pICln ließen sich exprimieren und eluierten von SEC-Säulen ohne Ausnahme als monodisperser *Peak* bei einem

Elutionsvolumen, das einem monomeren Oligomerisierungszustand entsprach (Abbildung 21, Seite 67). Anschließend wurden diese pICln-Präparationen für die Rekonstitution von 6S- und 8S-Komplexen verwendet. Hierbei kristallisierten zwei unterschiedliche Varianten des 6S-Komplexes und eine Variante des 8S-Komplexes, die alle mit unterschiedlichen Fragmenten von pICln rekonstituiert wurden (Tabelle 1, Seite 71).

### ***Kristallisation der 6S- und 8S-Komplexe***

Das Kristallisations-Screening wurde im Hochdurchsatzverfahren unter Verwendung eines Cartesian Honeybee Nano-Dispensor automatisiert in Polypropylen-Mikrotiterplatten unter Verwendung der Methode des sitzenden Tropfens durchgeführt. Hierbei wurden jeweils 0,2 µl der Präzipitantz und der Proteinlösung gemischt. Die Reproduktion und Optimierung initialer Kristallisationsbedingungen zur Erzeugung von Kristallen für die diffraktometrische Messungen erfolgte manuell unter Anwendung der Methode des hängenden Tropfens auf Deckgläschen in Polystyrol-Platten. Hierbei umfasste das Screening ansonsten identischer Bedingungen auch unterschiedliche Protein: Präzipitantz Verhältnisse (1:1 – 10:1), Tropfengrößen (0,5 – 8 µl) und Temperaturen (293K, 289K, 277K). Die erfolgreichen Kristallisationsbedingungen sind in Tabelle 1 auf Seite 71 aufgeführt.

### ***Röntgendiffraktion an Kristallen***

Bei der Röntgendiffraktometrie werden die Beugungseffekte der Probe auf die eingehenden Röntgenstrahlen aufgezeichnet. Auf Basis der zweidimensionalen Beugungsmuster, die von unterschiedlichen Orientierungen des Kristalls aufgenommen werden, kann mit Hilfe der mathematischen Methode der Fouriertransformation die dreidimensionale Verteilung der Elektronen innerhalb des Kristalls berechnet werden. Eine Interpretation dieser Elektronendichte ermöglicht die Erstellung eines Modells der Kristallstruktur des beugenden Objektes. Das Röntgendiffraktometer besteht aus einer Röntgenquelle, optischen Instrumenten zur Erzeugung eines monochromatischen und kollimierten Strahles, dem Goniometer zur exakten Positionierung und Drehung der Probe im Strahl, sowie einem Detektor zur Aufnahme der Beugungsdaten. Die Beugung beruht auf einer elastischen Interaktion der hochenergetischen elektromagnetischen Röntgenstrahlung mit den Elektronen der Atome innerhalb des Kristalls. Zwar verursacht auch ein einzelnes Molekül Beugungseffekte, die hierdurch verursachte Streuung ist aber für eine Messung deutlich zu gering. Das periodische Gitter eines Kristalls führt zu einer systematischen Amplifikation des Signals, das bei einem typischen Proteinkristall (100 µM isotropische Kantenlänge, 100 Å Durchmesser des Moleküls) in jede Dimension um den Faktor  $10^4$  verstärkt wird. Dadurch ist das Gesamtsignal eines Kristalls ungefähr  $10^{12}$ mal so groß, wie das eines einzelnen Moleküls. Außerdem verbessert sich durch die diskrete Positionierung der Beugungspunkte massiv das Signal-zu-Rausch-Verhältnis. Diese ergeben sich aufgrund der Beschränkung der Diffraktionswinkel für die konstruktive Interferenz im

Kristallgitter, welche in der Bragg-Gleichung ( $n\lambda = 2d \sin(\theta)$ ) festgelegt ist. Diese besagt, dass die Phasenbeziehung partieller Wellen durch Reflektion an parallelen Gitterebenen nur erhalten bleibt, wenn der Gangunterschied einem ganzzahligen Vielfachen der Wellenlänge entspricht. Hierbei hat aufgrund ihres Gitterabstands jede Schar von Gitterebenen einen charakteristischen Braggwinkel, für den die Gleichung erfüllt ist. Ist die Bragg-Gleichung nicht erfüllt, kommt es aufgrund der hohen Anzahl an Atomen im Kristall statistisch zur Auslöschung der gebeugten Welle.

Die im Rahmen dieser Arbeit durchgeführten Messungen zur Bestimmung der 6S- und 8S-Kristallstrukturen wurden an Messtationen für makromolekulare Datensammlung an den europäischen Synchrotrons ESRF (European Synchrotron Radiation Facility), Grenoble, Frankreich und SLS (Swiss Light Source), Villingen, Schweiz durchgeführt. Die jeweiligen Messstationen an denen die einzelnen Datensätze gemessen wurden, die zur Strukturlösung führten, sind in Tabelle 2 auf Seite 72 vermerkt.

### ***Datensammlung, Strukturlösung und Verfeinerung***

Zur Minimierung von Strahlungsschäden während der Messung wird diese in der Regel bei kryogenen Temperaturen durchgeführt. Hierfür wurden die Kristalle in einen zur Kristallisationsbedingung ähnlichen Kryo-Puffer überführt. Bei diesem wurde der Wasseranteil teilweise durch Kryo-Protektiva ersetzt, die während dem Frieren zu einer Vitrifizierung des Puffers führen. Genaue Angaben zu den Kryo-Puffern befinden sich in Tabelle 1 auf Seite 71. Anschließend wurden die Kristalle mit Hilfe von *Litho-Loops* in flüssigen Stickstoff überführt und für den Transport zur Messstation in Dewargefäßen aufbewahrt. Zwei unterschiedliche Kristallformen des 6S-Komplexes führten zur Strukturlösung. Diese werden zur Unterscheidung im Folgenden anhand ihrer Raumgruppen-Zugehörigkeit benannt (6S-P321 und 6S-C2).

### ***8S-Komplex***

Die 8S-Kristalle wurden über fünf Zwischenschritte, die durch Mischen des Kryo-Puffers mit abnehmenden Anteilen an ursprünglicher Präzipitantlösung entstanden, in den finalen Kryo-Puffer transferiert und anschließend innerhalb von *Litho-Loops* in flüssigem Stickstoff gefroren. Während der Messung befanden sich die Kristalle innerhalb eines konstanten Stickstoff Gasstroms bei 100K. Der finale Datensatz wurde an Messtation ID14-4 des ESRF in Grenoble aufgenommen. Die Prozessierung erfolgte mit XDS (Kabsch, 2010) wobei für eine korrekte Indizierung der Reflexe der Wert für die Einstellung „STRONG\_PIXEL“ auf 4 gesetzt werden musste. Die allermeisten Datensätze, die von den 8S Kristallen gesammelt wurden, wiesen auf eine C2-Symmetrie mit den Zellparametern  $a = 257 \text{ \AA}$ ,  $b = 72 \text{ \AA}$ ,  $c = 150 \text{ \AA}$  und  $\alpha = 90^\circ$ ,  $\beta = 117,6^\circ$   $\gamma = 90^\circ$  hin. Die Beugungsdaten des finalen Datensatzes zeigten jedoch zusätzliche schwache Reflexe und nach einem Anpassen der Autoindizierungsparameter konnte die korrekte P2<sub>1</sub>-Zelle mit den Parametern  $a = 149,9 \text{ \AA}$ ,  $b = 358,3 \text{ \AA}$ ,

$c=229,6 \text{ \AA}$  und  $\alpha=90^\circ$ ,  $\beta=97,7^\circ$   $\gamma=90^\circ$  bestimmt werden. Der Molekulare Ersatz erfolgte durch Molrep (Vagin and Teplyakov, 1998, 2010) unter Verwendung der Koordinaten von PDB Eintrag 3S6N als Suchmodell (Zhang et al., 2011) und war sowohl mit dem in C2 als auch mit dem in  $P2_1$  prozessierten Datensatz erfolgreich. Ein Vergleich der Zellparameter zeigt, dass die a-Achse der  $P2_1$ -Zelle identisch zur c-Achse der C2-Zelle ist und die b-Achse der  $P2_1$  Zelle fünfmal so lang ist wie die der C2-Zelle. Dieser Faktor 5 spiegelt sich auch an der Zahl der 8S-Moleküle in der Zelle wieder. Während die C2-Zelle acht Moleküle enthielt (4 x 2 pro asymmetrische Einheit) konnten in der  $P2_1$ -Zelle 40 Moleküle (2 x 20 pro asymmetrische Einheit) gefunden werden. Die  $P2_1$ -Zelle weist translationale nicht-kristallographische Symmetrie auf, welche eine Verwandtschaft zur C2-Zelle widerspiegelt. Die Verfeinerung erfolgte in PHENIX (Adams et al., 2010) unter Verwendung der Rigid-body, Koordinaten und ADP Algorithmen des Programms. Anschließend erfolgte die manuelle Platzierung der pICln PH-Domäne aus PDB 1ZYI in den entsprechenden Differenzdichten innerhalb der 20 Komplexmoleküle. Mit dem vervollständigten Modell folgte eine weitere Runde automatisierter Verfeinerung, woraufhin fehlende und abweichende Bereiche manuell in COOT (Emsley and Cowtan, 2004) ergänzt bzw. korrigiert wurden. Nach 10 Runden manueller und automatisierter Verfeinerung konvergierten die Werte für die  $R/R_{\text{free}}$ -Faktoren. Ab diesem Zeitpunkt war ebenfalls das hochaufgelöste Strukturmodell des 6S-Komplexes verfügbar. Dieses ermöglichte die weitere Verfeinerung mit dem Programm BUSTER (Smart et al., 2012), indem die Koordinaten der 6S-Kristallstruktur als Referenzmodell dienten. Das Programm verfügt über eine automatische Erkennung von Koordinatendifferenzen, die während der Verfeinerungsrunden aktiviert wurde. Hierdurch waren ungefähr 25 % der Referenzmodell-Einschränkungen automatisch deaktiviert.

## **6S-Komplex**

### **6S-P321 Kristallstruktur**

Die 6S-P321-Kristalle wurden in einen Kryopuffer transferiert, bei dem das Natriumformiat der Kristallisationsbedingung durch Lithiumformiat ersetzt wurde (Tabelle 1, Seite 71) (Rubinson et al., 2000). Die Kristalle wurden mit Hilfe von *Litho-Loops* in flüssigen Stickstoff überführt. Während der Messung befanden sich die Kristalle innerhalb eines konstanten Stickstoff Gasstroms bei 100K. Der Datensatz wurde an Messstation PX-II des SLS in Villingen, Schweiz aufgenommen. Die Prozessierung erfolgte mit dem Programm XDS und die Strukturlösung erfolgte mit der Methode des Molekularen Ersatzes mit PHASER (McCoy et al., 2007) durchgeführt. Als Suchmodell wurde der 8S-Kristallstruktur nach Entfernung der Gemin2 und SMN Untereinheiten verwendet. Hierbei konnten 18 Moleküle des 6S-Komplexes innerhalb der Asymmetrischen Einheit gefunden werden. Die Verfeinerung mit COOT und PHENIX wurde zu einem frühen Zeitpunkt eingestellt, da eine weitere

Kristallform erhalten wurde, von der ein hochauflösender Datensatz gemessen wurde, der zur Strukturaufklärung des 6S-Komplexes führte.

### **6S-C2 Kristallstruktur**

Die 6S-C2-Kristalle wurden über einen Zwischenschritt (1:1 Mischung von Reservoir und Kryopuffer) in den endgültigen Kryopuffer (Tabelle 1, Seite 71) überführt. Die Kristalle wurden mit Hilfe von *LithoLoops* in flüssigen Stickstoff überführt. Während der Messung befanden sich die Kristalle innerhalb eines konstanten Stickstoff Gasstroms bei 100K. Der Datensatz wurde an Messstation ID14-4 des ESRF in Grenoble, Frankreich aufgenommen. Die Prozessierung erfolgte mit dem Programm XDS und die Strukturlösung erfolgte mit der Methode des Molekularen Ersatzes mit PHASER (McCoy et al., 2007) durchgeführt. Als Suchmodell wurde der 8S-Kristallstruktur nach Entfernung der Gemin2 und SMN Untereinheiten verwendet. Hierbei konnten zwei Moleküle des 6S-Komplexes innerhalb der Asymmetrischen Einheit gefunden werden. Die Verfeinerung des Modells erfolgte manuell in COOT und automatisiert in PHENIX inklusive TLS Verfeinerung. Die Werte für  $R/R_{\text{free}}$  konvergierten nach fünf Verfeinerungsrunden während denen ein PEG-Molekül (Rückgrad des verwendeten Polyetheramines Jeffamine ED-2003) und 372 Wassermoleküle platziert wurden. Durch das modellieren von alternativen Konformationen an 27 Aminosäurepositionen konnte der  $R_{\text{free}}$  nochmals um zirka 2% verringert werden.

### **Elektronenmikroskopie zur Strukturuntersuchung des 8S-Komplexes**

Die Transmissionselektronenmikroskopie hat in den letzten Jahren große Fortschritte gemacht und ist besonders für große makromolekulare Komplexe eine erfolgreiche Alternative zur Röntgenkristallographie. Dies findet häufig Anwendung wenn vergleichsweise geringe Mengen der Probe zur Verfügung stehen. Das ist besonders dann notwendig, wenn die Präparation der Probe aus endogenem Material erfolgt. Für die elektronenmikroskopischen Aufnahmen werden die Probenpartikel auf einem Stützgitter fixiert. Dieses wird oftmals mit Salzen von Schwermetallen getränkt, wodurch während der Messung eine Negativfärbung entsteht. Die Elektronen durchdringen zwar die Probe, jedoch nicht das Gitter. Um die Probe möglichst geringen Strahlungsschäden auszusetzen, erfolgt die Messung zumeist bei kryogenen Temperaturen. Außerdem können die Partikel durch die vitrifizierte Wasserschicht fixiert werden, wodurch artifizielle Effekte durch die Fixierung minimiert werden. Mit Hilfe spezieller Probenbehandlung kann bei der Kryo-EM auch auf eine Anfärbung verzichtet werden. Bei den während der Messung erzeugten Rohdaten handelt es sich um 2D-Projektionen der dreidimensionalen Partikel auf dem Gitter. Da diese aber zufällig in unterschiedlicher Orientierung positioniert sind, werden die Partikel aus verschiedenen Betrachtungswinkeln aufgenommen. Gelingt es den jeweiligen Orientierungen ihre Projektionswinkel zuzuordnen, ist es möglich ein dreidimensionales Modell der Partikel zu rekonstruieren. Um das

Signal-Rausch-Verhältnis zu verbessern, werden viele Partikel identischer Orientierung gemittelt und sogenannte 2D-Klassensummen erzeugt.

Die elektronenmikroskopischen Arbeiten wurden durch Dr. Ashwin Chari in Kooperation mit Prof. Dr. Holger Stark durchgeführt. Es gelang bereits vor Beginn der kristallographischen Arbeiten 2D-Klassensummen von 8S-Präparationen zu erzeugen, Versuche einer 3D-Rekonstruktion scheiterten jedoch. Dies gelang erst auf Basis eines Tiefpass-gefilterten Modells der 8S-Kristallstruktur. Durch die Tiefpass-Filterung werden die hochauflösenden Merkmale der Struktur gefiltert, wodurch ein niedrig aufgelöstes Strukturmodell erzeugt wird, das größere Ähnlichkeit zum elektronenmikroskopischen Strukturmodell zeigt. Mit Hilfe dieses Modells gelang es, den Klassensummen erste Projektionswinkel zuzuweisen und eine dreidimensionale EM-Struktur mit einer Auflösung von 20 Å zu erzeugen. Zur Verifizierung der 3D Rekonstruktion wurden 2-D Reprojektionen berechnet, die eine gute Übereinstimmung mit den ursprünglichen Klassensummen zeigen. Für die Lokalisierung von Strukturelementen, die in der 8S-Kristallstruktur nicht aufgeklärt werden konnten, wurden zwei weitere Varianten des 8S-Komplexes elektronenmikroskopisch untersucht.

### ***In silico* Methoden zur Untersuchung der Moleküldynamiken**

Für eine computergestützte Untersuchung der Moleküldynamiken der 6S- und 8S-Komplexe fanden die Normalmoden-Analyse, sowie molekulardynamische (MD)-Simulationen Anwendung. Während die Simulation genauere Modelle liefern kann, ist sie aufgrund der hohen Rechenintensität für Proteine und makromolekulare Komplexe sehr zeitaufwendig. Dies liegt einerseits an der Komplexität des Testsystems sowie an den langen Simulationszeiträumen, die notwendig sind um dynamische Eigenschaften von biologischen Systemen zu beschreiben. Die Normalmoden-Analyse ist weniger rechenintensiv und kann für die Beschreibung des Gesamtschwingungsverhaltens des Moleküls herangezogen werden. Hierbei wird dieses als elastisches Netzwerk angesehen, wobei bei Proteinen die einzelnen Aminosäuren als Knotenpunkte definiert werden, die jeweils durch Federn miteinander verbunden sind. Die Berechnung erfolgte mit dem eNémo-Server (Suhre and Sanejouand, 2004).

Bei der Simulation der Moleküldynamiken (MD-Simulationen) handelt es sich um eine computerbasierte Methode zur Modellierung atomarer Bewegungen mit dem Ziel ihre Interaktionen über den Simulationszeitraum beschreiben zu können. Hierfür wird ein begrenzter Simulationsraum definiert, welcher die jeweiligen Zielmoleküle enthält. Die in dieser Arbeit diskutierten MD-Simulationen wurden mit GROMACS (Van Der Spoel et al., 2005) sowohl für den 6S- als auch den 8S-Komplexe durchgeführt. Die Trajektorien umfassten jeweils 1500 Zustände, was einem Simulationszeitraum von drei Nanosekunden entspricht. Die unterschiedlichen Positionierungen der Atome in den einzelnen Zuständen lassen sich in B-Faktoren ausdrücken, welche ein Maß für die

Dislokation der Atome um einen mittleren Aufenthaltsort sind. Diese dienen normalerweise der Beschreibung der thermischen Gitterschwingungen, die bei kristallographischen Messungen zur Dämpfung der Intensität von Beugungsreflexen führen. Die Simulation wurde isotherm und isobar bei 310K und 1 Bar durchgeführt. Die Moleküle wurden in einer Wasserbox platziert, deren Nettoladung durch die Platzierung einer adäquaten Anzahl an Chlorid- oder Natriumatomen auf 0 gebracht wurde. Diese Berechnungen wurden durch den Leiter unserer kristallographischen Abteilung, Dr. Clemens Grimm, durchgeführt.

## 5. Ergebnisse

### Biologischer Hintergrund

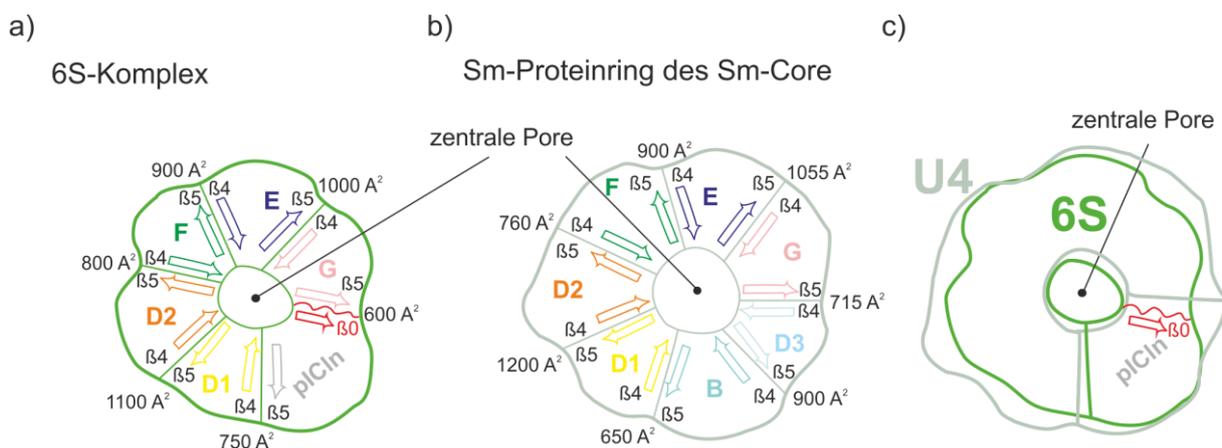
Der zentrale Schritt der U-snRNP-Biogenese ist die Ausbildung des Sm-Core. Diese strukturgebende Domäne der U-snRNP-Partikel zeichnet sich durch einen heptameren Ring aus, der durch die sieben Sm-Proteine SmB/B'-D1-D2-D3-E-F-G gebildet wird. Diese lagern sich zuerst unabhängig von der snRNA zu den Heterooligomeren SmE-F-G, SmD1-D2 und SmB/B'-D3 zusammen. Der zentrale Faktor während dieser frühen Zusammenlagerungsphase ist pICln. Dieses bindet mit SmD1-D2 und SmE-F-G zwei der Sm-Heterooligomere und sorgt damit für die Bildung eines hexameren Intermediates mit einem Sedimentationskoeffizienten von 6S. Mit der Bildung dieses 6S-Komplexes wird die frühe Phase des Zusammenlagerungsprozesses abgeschlossen. Im Anschluss wird dieser Komplex vom SMN-Komplex übernommen, wobei es zur Dissoziation von pICln und zum Transfer der Sm-Proteine auf die Sm-Bindungsstelle der snRNA kommt (Chari et al., 2008; Fisher et al., 1985; Raker et al., 1996). Um den Interaktionsmodus von pICln aufzuklären, besonders aber auch um Einblicke in die mechanistischen Grundlagen für seine spätere Dissoziation zu erhalten, wurden die Kristallstrukturen des 6S-Komplexes und eines weiteren verwandten Komplexes mit einem Sedimentationskoeffizienten von 8S gelöst. In diesem ist der 6S-Komplex an die beiden Untereinheiten SMN und Gemin2 des SMN-Komplexes gebunden. Durch eine C-terminale Kürzung des SMN Proteins (SMN $\Delta$ C) bleibt pICln Bestandteil des Komplexes und wird nicht dissoziiert. Daher ist der 8S-Komplex ein „gefangenes“ Transferintermediat zwischen der frühen und späten Phase der Sm-Core-Zusammenlagerung.

### 5.1 Strukturanalyse der 6S- und 8S-Intermediate der Sm-Core-Zusammenlagerung

#### Die Kristallstruktur des 6S-Komplexes

Der 6S-Komplex beinhaltet pICln und die beiden Sm-Heterooligomere SmD1-D2 und SmE-F-G. Die Interaktion zwischen den ursprünglichen Heterooligomeren erfolgt innerhalb des 6S-Komplexes über die benachbarten SmD2- und SmF-Untereinheiten. SmD1 und SmG werden durch pICln kontaktiert, wodurch sich die kontinuierliche Sequenz SmD1-D2-F-E-G-pICln in der topologischen Abfolge der Komplexuntereinheiten ergibt. Dadurch entsteht eine ringförmige Struktur, die in ihrer Mitte eine zentrale Pore umschließt (Abbildung 7a, Seite 44). Innerhalb dieses Sm/pICln-Rings befinden sich die Sm-Proteine zueinander bereits in der Anordnung, die sie auch im späteren Sm-Core einnehmen (Vergleiche Abbildung 7a und 7b) (Leung et al., 2011; Pomeranz Krummel et al., 2009; Weber et al., 2010). Der Proteinring des Sm-Core wird aus insgesamt sieben Sm-Proteinen gebildet, die etwa gleichgroße Sektoren des ringförmigen Komplexes einnehmen (Abbildung 7b, sowie Abbildung

3b und 3c, Seite 10). Betrachtet man hierbei die Schnittlinien der Ringebene mit den Kontaktflächen zwischen den einzelnen Untereinheiten als Schenkel eines Winkels, so spannt jedes Sm-Protein einen Winkel von ca.  $51^\circ$  (entsprechend  $360^\circ/7$ ) auf. Im 6S-Komplex nimmt hingegen pICln die topologische Position der Sm-Proteine SmB/B'-D3 ein. pICln beansprucht dabei innerhalb des Sm/pICln-Rings des 6S-Komplexes den Platz von ungefähr 1,5 Sm-Untereinheiten. Daher erscheint der 6S-Komplex leicht ellipsoid, während der heptamere Sm-Proteinring des Sm-Core eine nahezu kreisrunde äußere Form hat. Die ellipsoide Erscheinung wird außerdem noch dadurch verstärkt, dass pICln radial weiter aus dem Ring herausragt als die Sm-Proteine. Der 6S-Komplex ist somit entlang seiner SmD2-SmG-Achse mit  $65 \text{ \AA}$  etwas schmaler als das Sm-Core, während er entlang der pICln-SmF-Achse mit  $72 \text{ \AA}$  ähnlich dimensioniert ist, wie das Sm-Core. Eine Projektion der Ringstrukturen des 6S-Komplexes und des U4-Sm-Core entlang der Ringachse und eine anschließende Überlagerung dieser Projektionen veranschaulicht diesen Sachverhalt (Abbildung 7c).



**Abbildung 7** Umrisslinien (Projektion entlang der Ringachse) und Anordnung der Untereinheiten im 6S-Komplex und in der Sm-Core-Struktur. a) Schematische Darstellung des 6S-Komplexes. Die Untereinheiten sind entsprechend ihrer Farbkodierung gefärbt und die interagierenden  $\beta$ -Stränge schematisch angezeigt. Die jeweiligen Kontaktflächen sind in Quadrat-Ångström angegeben. Der  $\beta_0$ -Strang ist in rot dargestellt. b) Schematische Darstellung des Sm-Rings der U4 Sm-Core. Die Untereinheiten sind entsprechend ihrer Farbkodierung gefärbt und die interagierenden  $\beta$ -Stränge schematisch angezeigt. Die jeweiligen Kontaktflächen sind in Quadrat-Ångström angegeben. c) Schematische Überlagerung des Sm-Rings der U4-Sm-Core mit dem Sm/pICln-Ring des 6S-Komplexes.

## Die Sm-Proteine innerhalb des 6S-Komplexes

Bei den Sm-Proteinen handelt es sich um Eindomänenproteine womit sich ihre Faltungstopologie im Wesentlichen aus der Sm-Domäne ergibt, welche aus einer N-terminalen  $\alpha$ -Helix und fünf  $\beta$ -Strängen besteht. Die  $\beta$ -Stränge  $\beta_1$ -4 paaren in antiparalleler Orientierung so miteinander, dass in der Primärsequenz benachbarte  $\beta$ -Stränge auch in der Tertiärstruktur nebeneinander zu liegen kommen. Der  $\beta_5$ -Strang hingegen paart mit  $\beta_1$ , was zu einer Krümmung des  $\beta$ -Faltblatts führt, wodurch sich eine fassförmige Struktur der Domäne ergibt (Abbildung 3a, Seite 10). Die Interaktion der

Sm-Proteine miteinander beruht in erster Linie auf einer Paarung des  $\beta$ 4-Strangs mit dem  $\beta$ 5-Strang der jeweils benachbarten Sm-Untereinheit (Abbildung 7a und b, Seite 44). Die Paarung der  $\beta$ -Stränge benachbarter Untereinheiten erfolgt, wie auch schon die Paarung der  $\beta$ -Stränge innerhalb der Sm-Faltung, in antiparalleler Orientierung (Kambach et al., 1999; Leung et al., 2011). Außerdem sind die N-terminalen  $\alpha$ 1-Helices des  $\beta$ 5-Interaktors oberhalb des  $\beta$ -Faltblattes des  $\beta$ 4-Bindungspartners positioniert. Dadurch sind weitere Kontakte zwischen den Resten der  $\beta$ -Stränge  $\beta$ 3 und  $\beta$ 4 auf der einen Seite und den Resten der  $\alpha$ 1-Helices auf der anderen Seite möglich. Durch diese sich wiederholenden gleichförmigen Interaktionen weisen alle Sm-Proteine eine zueinander einheitliche Orientierung auf. So befinden sich beispielsweise die L4-Schleifen aller Sm-Proteine auf der jeweils einen Oberfläche des Sm/pICln-Rings, während sich die N-terminalen  $\alpha$ 1-Helices auf der gegenüberliegenden Oberfläche befinden. Aufgrund der sich daraus ergebenden topologischen Merkmale haben sich die Bezeichnung kegelförmige (*tapered face*) und flache Oberfläche (*flat face*) etabliert (Abbildung 8a, Seite 46) (Leung et al., 2011). Die kegelförmige Seite ist in der Kristallstruktur des 6S-Komplexes deutlich weniger stark ausgeprägt, als dies in den Kristallstrukturen der Sm-Core der Fall war (Abbildung 8b) (Leung et al., 2011; Pomeranz Krummel et al., 2009; Weber et al., 2010). In den Kristallstrukturen des Sm-Core sind die ausgedehnten L4-Schleifen von SmD2 und SmB/B' an Kontakten zur 3'-Haarnadelstruktur der snRNA beteiligt, wodurch diese stabilisiert werden und ihr Verlauf vollständig modelliert werden konnte (Abbildung 8a). In der Kristallstruktur des 6S-Komplexes ist hingegen keine Elektronendichte für die L4-Schleife von SmD2 vorhanden, was auf eine hohe Flexibilität dieser Proteinregion hinweist und die kegelförmige Ringfläche flacher erscheinen lässt (Abbildung 8b). Dies war auch in den Kristallstrukturen der Heterooligomere SmD1-D2 und SmB-D3 der Fall, in denen die L4-Schleifen der Sm-Proteine ebenfalls nicht modelliert werden konnten (Kambach et al., 1999). In den Kristallstrukturen der Sm-Core konnten innerhalb von SmD1, SmD2 und SmF zusätzliche Sekundärstrukturelemente identifiziert werden, die nicht zur kanonischen Sm-Faltung gehören. Hierbei handelt es sich um eine zusätzliche N-terminale Helix bei SmD2 (Helix  $\alpha$ 0) und zwei C-terminale Helices (Helix  $\alpha$ 2 und Helix  $\alpha$ 3) bei SmD1 und SmF. Diese waren in der Kristallstruktur des Sm-Core des U4-snRNP, ähnlich wie die L4-Schleifen, voranging in Kontakte zur snRNA oder zu snRNP-spezifischen Proteinen involviert (Leung et al., 2011). Mit Ausnahme der N-terminalen  $\alpha$ 0-Helix von SmD2 und der SmD1  $\alpha$ 2-Helix konnten diese Sekundärstrukturelemente nicht in der Kristallstruktur des 6S-Komplexes identifiziert werden (Abbildung 9, Seite 46).



## Die Struktur von pICln im 6S-Komplex

### **Die erweiterte PH-Domäne von pICln**

pICln ermöglicht durch seine Bindung der Heterooligomere SmD1-D2 und SmE-F-G, die ansonsten keine Affinität zueinander besitzen, mit der Formierung des hexameren 6S-Komplexes ein Intermediat höherer Ordnung. Innerhalb der Topologie des 6S-Komplexes besetzt pICln eine Position, die äquivalent zur Position von SmB/B'-D3 im Sm-Core ist (Vergleiche Abbildung 7a und 7b, Seite 44). Hierfür bildet pICln zwar Kontakte zu SmD1 und SmG aus, besitzt jedoch im Gegensatz zu diesen keine Sm-Faltung, sondern eine Pleckstrin-Homologie (PH)-Domäne (Furst et al., 2005; Haslam et al., 1993) (Abbildung 10, Seite 48). Diese Faltungseinheit besteht aus sieben  $\beta$ -Strängen, die sich in zwei antiparallelen  $\beta$ -Faltblättern mit zueinander orthogonaler Ausrichtung organisieren, sowie einer C-terminalen  $\alpha$ -Helix. Die PH-Domäne von pICln wird in der Sequenz des verwendeten Orthologs aus *Drosophila melanogaster* durch die Reste Met1 bis Leu159 gebildet. Die in der Primärsequenz benachbarten Stränge liegen in den beiden Falblättern nebeneinander und bilden antiparallele  $\beta$ - $\beta$ -Interaktionen miteinander aus. Die Faltungstopologie der PH-Domäne weist somit sehr große Ähnlichkeit zum Beta-Sandwich Faltungstypus auf (Woolfson et al., 1993). Bei diesem werden aus insgesamt sechs  $\beta$ -Strängen zwei antiparallele Falblätter mit jeweils vier  $\beta$ -Strängen gebildet. Hierfür gehen die äußeren  $\beta$ -Stränge  $\beta$ 1 und  $\beta$ 4 ohne Unterbrechung an zwei diagonal gegenüberliegenden Ecken in das jeweils andere Falblatt über. Dies ist in ähnlicher Weise auch bei der PH-Domäne der Fall (Abbildung 10, Seite 48). Bei dieser ist der  $\beta$ 4-Strang jedoch im Übergang zwischen den Falblättern unterbrochen, sodass die PH-Domäne insgesamt sieben  $\beta$ -Stränge besitzt. Bei pICln ist zwischen den  $\beta$ -Strängen  $\beta$ 4 und  $\beta$ 5 eine  $3_{10}$ -Helix lokalisiert, durch deren Windung die Hauptkette eine Drehung um  $90^\circ$  beschreibt und den Übergang in das C-terminale Falblatt ermöglicht. Ein weiterer Unterschied der PH-Domäne zum einfachen Beta-Sandwich ergibt sich dadurch, dass auch der  $\beta$ 2-Strang durch eine ununterbrochene Paarung zum  $\beta$ 1-Strang an der Ausbildung beider Falblätter beteiligt ist. Dadurch besteht das C-terminale Falblatt nicht aus vier, sondern aus insgesamt fünf  $\beta$ -Strängen.

Eine Besonderheit der PH-Domäne von pICln, die erst durch die Lösung der Kristallstruktur des 6S-Komplexes aufgeklärt werden konnte, ist ein zusätzlicher N-terminaler  $\beta$ -Strang, der antiparallel mit dem  $\beta$ 4-Strang gepaart ist. Für diesen zusätzlichen  $\beta$ -Strang wurde die Bezeichnung  $\beta$ 0 gewählt, um die gängige Nomenklatur für die Topologie der PH-Domäne beibehalten zu können. Diese Nomenklatur war durch die NMR-Struktur des *C. l. familiaris* Orthologs auch für pICln bereits etabliert, da der  $\beta$ 0-Strang in dieser nicht strukturiert war (Furst et al., 2005). pICln besitzt also zwei fünfsträngige  $\beta$ -Falblätter, wobei das N-terminale  $\beta$ -Falblatt aus den  $\beta$ -Strängen  $\beta$ 1,  $\beta$ 2,  $\beta$ 3,  $\beta$ 4 und  $\beta$ 0 und das C-terminale aus den  $\beta$ -Strängen  $\beta$ 5,  $\beta$ 6,  $\beta$ 7,  $\beta$ 1 und  $\beta$ 2 gebildet wird. Aufgrund dieses Unterschieds zur kanonischen PH-Domäne wird bei pICln von einer *erweiterten* PH-Domäne

gesprochen. Die orthogonal zueinander verlaufenden  $\beta$ -Stränge der beiden Faltblätter bilden eine fassförmige Struktur aus, womit eine strukturelle Analogie zur Sm-Faltung besteht. Dieses  $\beta$ -Fass besitzt zwei gegenüberliegende Öffnungen, die durch die jeweiligen äußeren  $\beta$ -Stränge der beiden  $\beta$ -Faltblätter begrenzt sind. Die durch die  $\beta$ -Stränge  $\beta 1$  und  $\beta 5$  begrenzte Öffnung („obere“ Öffnung in Abbildung 10) wird durch die zur erweiterten PH-Domäne gehörende C-terminale  $\alpha$ -Helix bedeckt.

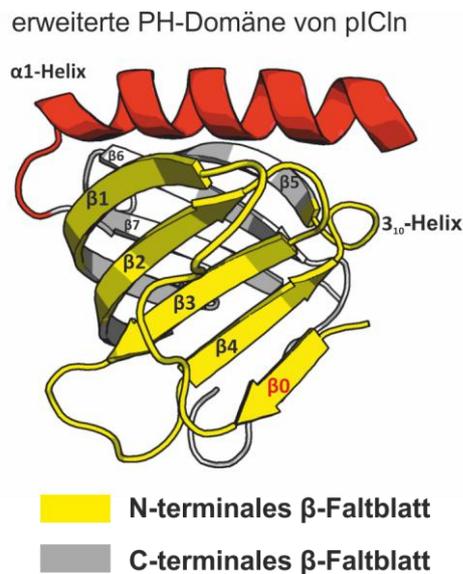


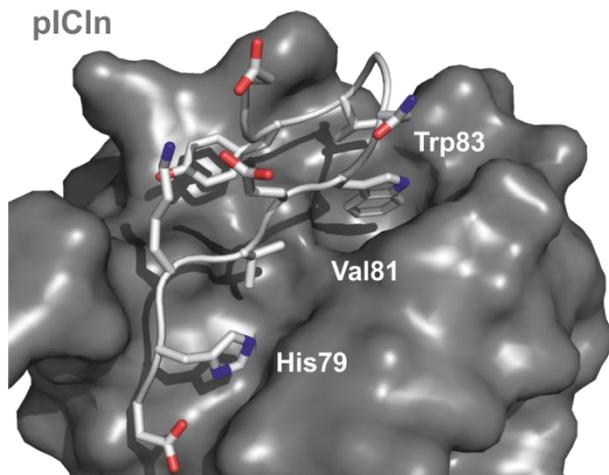
Abbildung 10 Cartoon-Darstellung der erweiterten PH-Domäne von pICln.  $\beta$ -Stränge des N-terminalen  $\beta$ -Faltblattes sind gelb und  $\beta$ -Stränge des C-terminalen  $\beta$ -Faltblattes sind grau koloriert. Die C-terminale  $\alpha 1$ -Helix ist in rot dargestellt.

### ***pICln weist ausgedehnte flexible Bereiche auf***

C-terminal von der erweiterten PH-Domäne kann für die Reste His160 bis Glu215 von pICln keine Sekundärstruktur vorhergesagt werden. Für die erfolgreiche Kristallisation und Strukturlösung des 6S-Komplexes war es notwendig, diese Reste von pICln zu deletieren. Anschließend wurden entsprechende 6S-Komplexe mit dieser pICln-Variante ( $\Delta 160-215$ ) rekonstituiert. Bei PH-Domänen kann außerdem die Länge der Schleifenregionen zwischen den einzelnen  $\beta$ -Strängen extrem variieren (Sugimoto et al., 2003). Im Fall von pICln ist besonders die Schleifenregion zwischen den  $\beta$ -Strängen  $\beta 6$  und  $\beta 7$ , die im Folgenden als L5-Schleife bezeichnet wird, stark ausgeprägt. Bereits in der NMR-Struktur des *C. I. familiaris* pICln-Orthologs zeigten die, in diesem Bereich lokalisierten Reste Ser93 bis Ser115 (entsprechend den Resten Asp78 bis Glu132 von *D. melanogaster* pICln, Abbildung 27, Seite 99 im Anhang), eine hohe Flexibilität (Furst et al., 2005). Daher wurden ebenfalls pICln-Varianten erzeugt, die Deletion in diesem Bereich aufwiesen. Diese Deletionen wurden mit den C-terminalen Kürzungen kombiniert und es wurden entsprechend modifizierte Varianten des 6S-Komplexes präpariert. Eine erste Kristallform des 6S-Komplexes wurde anschließend mit einer

6S-Präparationen erhalten, die eine solche pICln-Variante beinhaltet ( $\Delta 90-120 \Delta 160-215$ ). Auch eine weitere 6S-Kristallform konnte anschließend mit einer solchen 6S-Präparation erhalten werden, die eine ähnliche pICln-Variante ( $\Delta 90-127 \Delta 160-215$ ) enthielt, bei der zusätzlich der Rest His144 durch Alanin substituiert wurde (eine detaillierte Beschreibung der Kristallisationsstrategie folgt unter Abschnitt 5.2, ab Seite 64).

Trotz des Fehlens eines Großteils der L5-Schleifenregion, wies dieser Bereich auch in den gekürzten pICln-Varianten immer noch eine hohe Flexibilität auf und konnte nicht vollständig modelliert werden. Im finalen Modell fehlen die Reste Gly88 bis Glu132 und der N-terminale Teil der L5-Schleife konnte nur von Asp78 bis Tyr87 modelliert werden. Die hier lokalisierten Reste His79, Val81 und Trp83 binden auf der Oberfläche von pICln in hydrophobe Taschen, wodurch sie strukturell stabilisiert wird (Abbildung 11). Die drei hydrophoben Taschen befinden sich im Bereich der „unteren“ Öffnung (entsprechend der Orientierung von pICln in Abbildung 10, Seite 48) der Faszstruktur zwischen den beiden Faltblättern der erweiterten PH-Domäne. Der C-terminale Teil der L5-Schleife lässt sich in der Elektronendichte erst wieder ab einem Glutamat in Position 132 nachvollziehen, das sich in räumlicher Nähe zu His79 befindet. Damit liegen beide Enden der Schleife auf der kegelförmigen Fläche des Sm/pICln-Rings und es ist daher wahrscheinlich, dass sich die dazwischenliegenden Reste ebenfalls im Bereich dieser Ringoberfläche befinden.



**Abbildung 11** His79, Val81 und Trp83 binden in hydrophobe Taschen auf der Oberfläche von pICln wodurch der N-terminale Bereich der L5-Schleifenregion stabilisiert wird.

## Die Interaktionen von pICln mit den benachbarten Sm-Proteinen SmD1 und SmG

### Die Interaktion zwischen SmD1 und pICln

Der heptamere Proteinring des Sm-Core weist ein durchgängiges antiparalleles  $\beta$ -Faltblatt auf, das sich durch die jeweils alternierende Bindung zwischen den  $\beta$ -Strängen  $\beta 4$  und  $\beta 5$  der benachbarten Sm-Untereinheiten ergibt (Abbildung 7b). Dieser Bindungsmodus gilt auch für die Sm-Untereinheiten des 6S-Komplexes im Sm/pICln-Ring. Die Kontaktflächen zwischen den einzelnen Sm-Proteinen umfassen  $800\text{-}1100 \text{ \AA}^2$  und sind damit ähnlich dimensioniert wie im Sm-Core (vergleiche Abbildung 7a und b, Seite 44). pICln bindet die flankierenden Untereinheiten des Sm-Pentamers und kontaktiert hierbei den  $\beta 4$ -Strang von SmD1 und den  $\beta 5$ -Strang von SmG (Abbildung 7a und Abbildung 8b). Der Kontakt zu SmD1 erfolgt über den  $\beta 5$ -Strang der erweiterten PH-Domäne von pICln, bei dem es sich um einen der beiden äußeren Stränge des C-terminalen  $\beta$ -Faltblattes der Domäne handelt (Abbildung 10, Seite 48). Dieser ist in ähnlicher Weise positioniert wie die  $\beta 5$ -Stränge der Sm-Proteine und die Paarung mit dem  $\beta 4$ -Strang von SmD1 erfolgt somit ebenfalls in antiparalleler Orientierung. Innerhalb des Kontaktes werden fünf Wasserstoffbrücken zwischen den Resten Ser61<sub>pICln</sub>-Ile65<sub>pICln</sub> und Thr57<sub>SmD1</sub>-Arg61<sub>SmD1</sub> ausgebildet (Abbildung 12, Seite 51, links). Beteiligt sind hierbei auf der Seite von SmD1 die gleichen Aminosäuren, die im Sm-Core auch den Kontakt zu SmB/B' ausbilden (Abbildung 25, Seite 77). Ergänzt werden diese fünf Wasserstoffbrücken durch eine sechste Wasserstoffbrücke, die zwischen der Hydroxylgruppe der Seitenkette von Ser66<sub>pICln</sub> und dem Carbonyl-Sauerstoff von Gln54<sub>SmD1</sub> ausgebildet wird. In Richtung der kegelförmigen Ringfläche ist außerdem ein Verbund hydrophober Aminosäuren (Leu62<sub>pICln</sub>, Tyr74<sub>pICln</sub>, Met76<sub>pICln</sub>, Val133<sub>pICln</sub>; Leu19<sub>SmD1</sub>, Met45<sub>SmD1</sub>, Leu55<sub>SmD1</sub>, Ile58<sub>SmD1</sub>) zu finden (Abbildung 12, rechts). Auf der flachen Ringseite existieren weitere Kontakte von SmD1 zur C-terminalen  $\alpha 1$ -Helix von pICln. Diese ist ähnlich positioniert wie die N-terminalen  $\alpha 1$ -Helices der Sm-Proteine, wodurch sie teilweise über dem  $\beta$ -Faltblatt von SmD1 zu liegen kommt. Damit weist dieser Kontakt starke Ähnlichkeiten zum typischen Sm-Sm-Interaktionsmodus auf und auch die Kontaktfläche zwischen pICln und SmD1 ist von ihrem Umfang her vergleichbar mit denen zwischen zwei Sm-Proteinen. Mit  $750 \text{ \AA}^2$  liegt diese aber am unteren Rand dieser zwischen  $800\text{-}1100 \text{ \AA}^2$  umfassenden Interaktionsflächen (Abbildung 7a, Seite 44).

### Kontakt zwischen pICln und SmD1

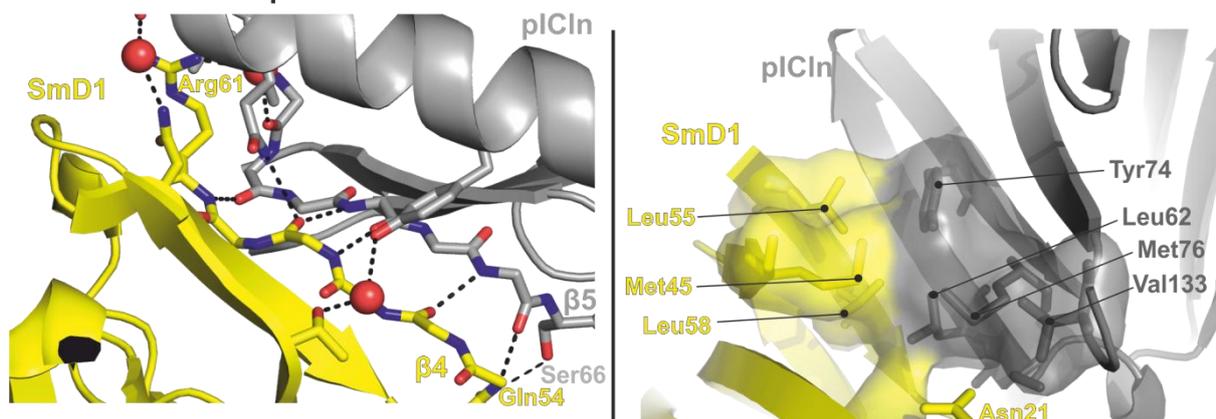


Abbildung 12 Interaktionen zwischen pICln und SmD1. Die  $\beta$ -Stränge  $\beta 4$  von SmD1 und  $\beta 5$  von pICln paaren in antiparalleler Orientierung (links). Des Weiteren bildet sich ein hydrophobes Cluster auf Höhe der kegelförmigen Ringseite (rechts).

### Die Interaktion zwischen SmG und pICln

Der Modus der Interaktion zwischen pICln und SmG unterscheidet sich deutlich von dem beschriebenen Sm-Sm-Interaktionsmodus. Der Kontakt zwischen beiden Proteinen erfolgt über den  $\beta 0$ -Strang von pICln und den  $\beta 5$ -Strang von SmG. Bei einer Nachahmung des Sm-Sm-Interaktionsmodus innerhalb des Kontakts zwischen SmG und pICln müsste der  $\beta 0$ -Strang von pICln, wegen seiner Paarung mit dem  $\beta 5$ -Strang von SmG, als Pendant zu den  $\beta 4$ -Strängen der Sm-Proteine angesehen werden. Die  $\beta 4$ -Stränge verlaufen in Richtung der zentralen Pore und ermöglichen damit eine antiparallele Paarung mit den in gegengesetzter Richtung verlaufenden  $\beta 5$ -Strängen (Abbildung 7b, Seite 44). Der  $\beta 0$ -Strang von pICln verläuft jedoch in identischer Richtung wie die  $\beta 5$ -Stränge, so dass seine Paarung mit dem  $\beta 5$ -Strang von SmG nicht in antiparalleler, sondern vielmehr in paralleler Orientierung erfolgt (Abbildung 7a und Abbildung 8b, Seite 44 bzw. Seite 46). Die Interaktion der beiden  $\beta$ -Stränge beruht auf insgesamt vier Wasserstoffbrückenbindungen, wobei die Bindung zwischen Ile<sub>4</sub><sub>pICln</sub> und Ala72<sub>SmG</sub> durch ein koordiniertes Wassermolekül vermittelt wird (Abbildung 13, links, Seite 52). Das N-terminale Methionin von pICln ist Bestandteil des  $\beta 0$ -Strangs und die N-terminale Aminogruppe wird durch ein ausgedehntes Netzwerk von Wasserstoffbrücken kontaktiert. Met1<sub>pICln</sub> ist innerhalb einer hydrophoben Tasche von SmG platziert, die durch Leu9<sub>SmG</sub>, Leu40<sub>SmG</sub>, Ile67<sub>SmG</sub> und Leu70<sub>SmG</sub> gebildet wird. Eine ähnliche Tasche bindet auf Seiten von pICln Met69<sub>SmG</sub> und wird durch Ile4<sub>pICln</sub>, Val81<sub>pICln</sub>, Trp83<sub>pICln</sub> gebildet. Hierdurch kommt es zu einer gegenseitigen, zapfenartigen Verankerung beider Proteine ineinander (Abbildung 13, rechts). Besonders auffällig ist die Positionierung des  $\beta 0$ -Strangs am äußeren Rand der erweiterten PH-Domäne. Hierdurch sind die fassförmigen Faltblätter beider Proteine deutlich weiter voneinander entfernt, als dies bei den Kontakten zwischen den Sm-Proteinen und auch beim Kontakt

zwischen pICln und SmD1 der Fall ist (Abbildung 8b). Es ist daher auch nicht möglich, dass die N-terminale  $\alpha$ 1-Helix von SmG ähnliche Kontakte zu pICln ausbildet, wie sie bei den Sm-Proteinen und im SmD1-pICln-Kontakt beobachtet werden. Das spiegelt sich auch in einer verringerten Interaktionsfläche von nur 600 Å<sup>2</sup> wider (Abbildung 7a).

### Kontakt zwischen pICln und SmG

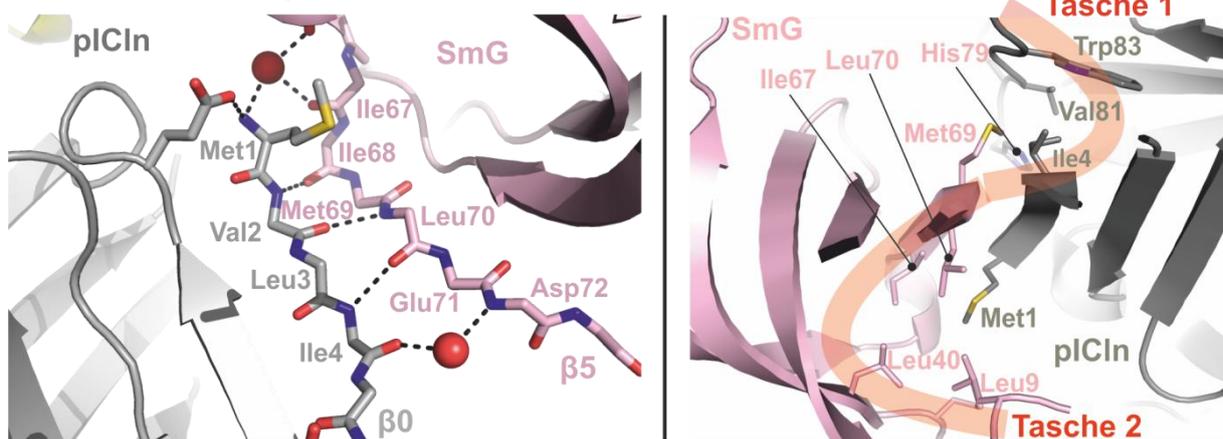


Abbildung 13 Interaktionen zwischen pICln und SmG. Die  $\beta$ -Stränge  $\beta_5$  von SmD1 und  $\beta_0$  von pICln paaren in paralleler Orientierung (links). Des Weiteren wird jeweils ein Methionin (Met1<sub>pICln</sub> und Met69<sub>SmG</sub>) beider Proteine in einer hydrophoben Tasche des jeweils anderen Proteins gebunden (rechts). Mit Erlaubnis von Cell Press (rechts verändert nach Grimm, C. et al.,2013).

### Physikochemische Eigenschaften von pICln im Kontext der Sm-Proteinbindung

Ein Sequenzvergleich von verschiedenen pICln-Orthologen zeigt, dass nur wenige Bereiche des Proteins auf Aminosäure-Ebene streng konserviert sind. Die Konservierung ist hauptsächlich auf Reste beschränkt, die sich innerhalb der Kontaktfläche zu SmD1 befinden (Abbildung 27, Seite 99 im Anhang). Interessanterweise ist mit Ausnahme des N-terminalen Methionins nicht einmal die Kontaktfläche zu SmG konserviert (Abbildung 14a).

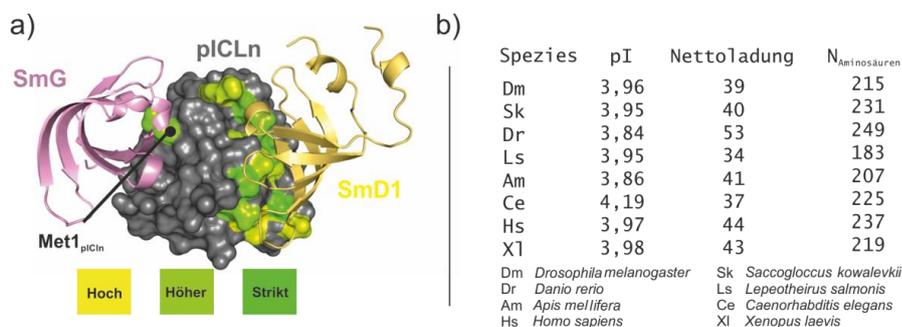


Abbildung 14 a) Konservierte Sequenzbereiche von pICln auf dessen Oberfläche. b) pI und Nettoladung unterschiedlicher pICln-Orthologe. Mit Erlaubnis von Cell Press (verändert nach Grimm, C. et al.,2013).

Beim Vergleich der physikochemischen Daten der pICln-Orthologen zeigt sich aber trotz der geringen Sequenzkonservierung eine spezieübergreifend hohe Anzahl an sauren Resten. Dadurch ergeben sich sehr ähnliche Werte sowohl für die Nettoladung als auch für den isoelektrischen Punkt (pI) aller orthologen Proteine. Diese liegen zwischen 3,8 und 4,0 für den pI und zwischen 40-50e für die Nettoladung (Abbildung 14b). Hierbei sind die sauren Reste größtenteils in drei aziden Regionen (AR) innerhalb Primärsequenz lokalisiert (siehe auch Abbildung 27) (Schedlbauer et al., 2011). AR1 befindet sich innerhalb der L5-Schleife, während AR2 sich am C-terminalen Ende der erweiterten PH-Domäne befindet und AR3 im entfernten C-Terminus lokalisiert ist. Somit enthielt die für die Kristallisation des 6S-Komplexes verwendete pICln-Variante keines dieser sauren Bereiche (siehe auch unter Abschnitt 5.2, ab Seite 64), weshalb eine außerordentlich starke Veränderung der Nettoladung und eine Erhöhung des isoelektrischen Punktes auf einen Wert von 5,5 mit den Kürzungen einhergingen. Für eine detaillierte Analyse der elektrostatischen Eigenschaften (Baker et al., 2001; Dolinsky et al., 2004) wurde daher zumindest die L5-Schleife, die AR1 beherbergt, auf Grundlage der NMR-Struktur des *C. l. familiaris* pICln-Orthologs rekonstruiert (Abbildung 15a, Seite 54) (Furst et al., 2005). Die Untersuchungen ergaben, dass das stark negative elektrostatische Potential von pICln hauptsächlich in Regionen des Proteins lokalisiert ist, die innerhalb des 6S Komplexes in Richtung der zentralen Pore des Rings ausgerichtet sind (Abbildung 15a). Durch diese Ladungsverteilung ergibt sich über das Molekül ein Dipolmoment dessen Vektor aus Richtung der zentralen Pore mit dem positiven Pol radial nach außen zeigt (Abbildung 15b). Entsprechende elektrostatische Analysen wurden ebenfalls mit dem Sm-Pentamer durchgeführt (Abbildung 15c). Mit einem pI von 9,9 und einer positiven Nettoladung von 37e hat das Sm-Pentamer ein ausgeprägt starkes positiv elektrostatisches Potential. Die hierfür verantwortlichen Ladungen sind hauptsächlich auf der zur zentralen Pore ausgerichteten Oberfläche der Sm-Proteine lokalisiert (Abbildung 15c). Hier befinden sich in erster Linie die L3- und L5-Schleifen der Sm-Proteine, welche im Sm-Core die Kontakte zur snRNA ausbilden (Leung et al., 2011) (Abbildung 15d).

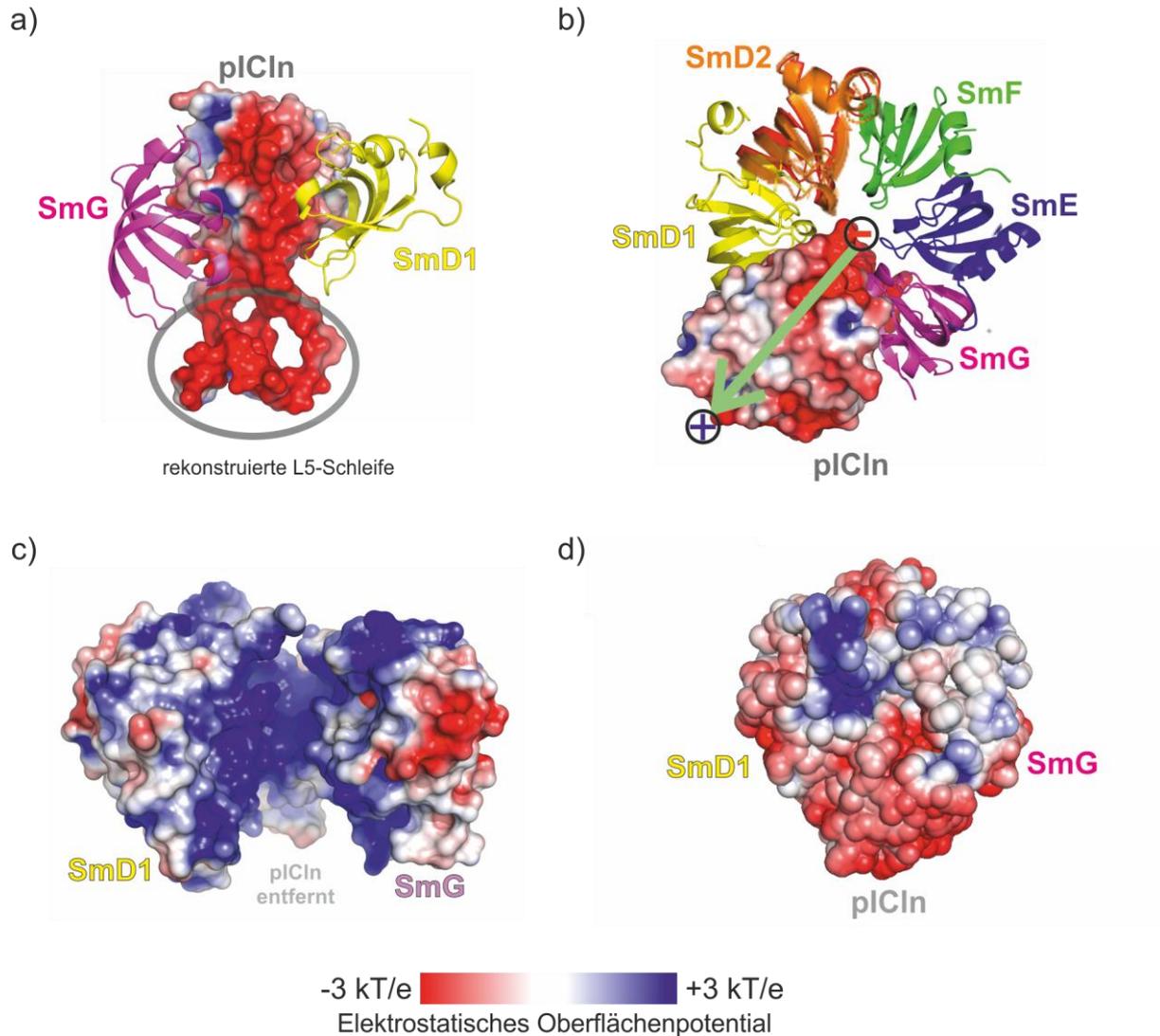


Abbildung 15 Elektrostatische Eigenschaften von pICln und des Sm-Pentamers SmD1-D2-E-F-G. a) Elektrostatisches Oberflächenpotential von pICln. Für diese Berechnungen wurde die unstrukturierte L5-Schleifenregion anhand der NMR-Struktur des *C. l. familiaris* pICln-Orthologs rekonstruiert. b) Durch die Ladungsverteilung entsteht über pICln ein Dipolmoment, dessen Vektor mit dem positiven Pol von der zentralen Pore des Sm/pICln-Rings radial nach außen zeigt. c) Elektrostatisches Oberflächenpotential des Sm-Pentamers SmD1-D2-E-F-G. d) Elektrostatisches Oberflächenpotential des 6S-Komplexes mit Blick auf die flache Ringfläche. Mit Erlaubnis von Cell Press. Teil (a), b) und c) verändert nach Grimm, C. et al.,2013).

## Die Kristallstruktur des 8S-Komplexes

Die Kristallstruktur des 6S-Komplexes ermöglicht detaillierte Einblicke in den Bindungsmodus zwischen pICln und den Sm-Proteinen. Die Kristallstruktur des 8S-Komplexes gibt zusätzlich Einblicke in den Mechanismus der SMN-Komplex-katalysierten Dissoziation von pICln.

Ein Vergleich der Sm/pICln-Ringe zwischen der 6S- und 8S-Kristallstruktur zeigt keine auffälligen Unterschiede. Der hexamere Ring, der durch pICln und das Sm-Pentamer gebildet wird, bleibt auch während der Bindung von SMNΔC-Gemin2 erhalten und es ergeben sich zwischen beiden Strukturen nur geringe Positionsabweichungen, die sich in einer geringen RMS-Koordinatendifferenz von nur 0,57 Å widerspiegeln. Damit ist der auffälligste Unterschied zwischen beiden Kristallstrukturen die Präsenz der beiden SMN-Komplexkomponenten im 8S-Komplex (Abbildung 16).

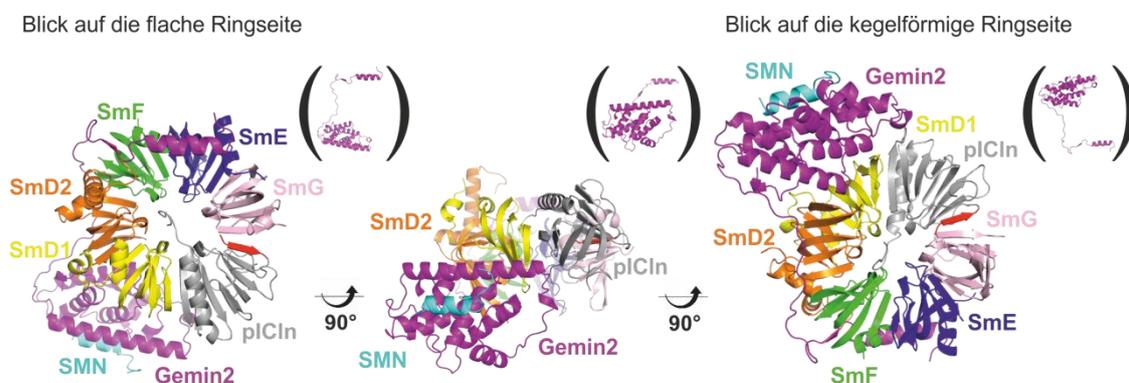


Abbildung 16 8S-Komplex in Cartoon-Darstellung mit Blick auf die flache (links) und kegelförmige (rechts) Ringseite, sowie entlang der Ringebene (mitte). Die Struktur von Gemin2 ist für jede Blickrichtung nochmals getrennt in den Klammern angezeigt.

## Die Strukturen von Gemin2 und SMNΔC in der Kristallstruktur des 8S-Komplexes

Gemin2 besitzt eine überwiegend helikale Sekundärstruktur und kann in drei unterschiedliche Regionen gegliedert werden. Aufgrund dieser Einteilung ergibt sich eine N-terminale Region (Met1<sub>G2</sub>-Ser55<sub>G2</sub>), eine C-terminale Domäne (Pro85<sub>G2</sub>-Ile245<sub>G2</sub>) und ein dazwischenliegender unstrukturierter Bereich (Ser56<sub>G2</sub>-Thr84<sub>G2</sub>). Die Sekundärstrukturelemente innerhalb des N-Terminus sind auf eine  $\alpha$ -Helix (Glu34<sub>G2</sub>-Arg46<sub>G2</sub>) und einen einzelnen  $\beta$ -Strang (Ala49<sub>G2</sub>-Ser55<sub>G2</sub>) beschränkt. Die C-terminale Domäne besitzt im Kern den Faltungstypus eines Vierhelixbündels (Abbildung 17a, helles Magenta, Seite 57), welches durch drei weitere  $\alpha$ - und zwei  $3_{10}$ -Helices ergänzt wird (Abbildung 17a, dunkles Magenta). Diese Domäne macht den größten Teil des Proteins aus und befindet sich auf der Seite der kegelförmigen Fläche des Sm/pICln-Rings auf Höhe von SmD1-D2. Hierbei wird SmD1 über die L4-Schleife (zwischen den Helices  $\alpha$ 4 und  $\alpha$ 5) und die L6-Schleife (zwischen den  $\alpha$ -Helices  $\alpha$ 6 und  $\alpha$ 7)

von Gemin2 kontaktiert. Diese Interaktionen beruhen auf der Ausbildung einiger Wasserstoffbrücken und Salzbrücken zwischen Aminosäuren, die sich in den  $\beta$ -Strängen  $\beta 3$  und  $\beta 4$  von SmD1 befinden (Abbildung 17b). Die Positionierung der L6-Schleife von Gemin2 im Kontaktbereich von SmD1 und SmD2 bedingt hier im Vergleich zu der im 6S-Komplex gefundenen Konformation ein Umklappen von Asn24<sub>SmD1</sub>. Diese Bewegung steht am Anfang einer Kaskade von Seitenkettenumlagerungen, die sich, einem Dominoeffekt ähnlich, über die gesamte Interaktionsfläche fortpflanzen (Abbildung 17c). Der unstrukturierte Bereich zwischen den Faltungselementen von Gemin2 (Ser56<sub>G2</sub>-Thr84<sub>G2</sub>) verläuft entlang der kegelförmigen Fläche, wodurch der Eindruck eines um den Ring greifenden Armes entsteht. Von der Position des C-terminalen Helixbündels in Höhe von SmD1-D2 ausgehend verläuft dieser bis hin zu SmE und SmF. Hier kontaktiert Gemin2 durch eine antiparallele Paarung seines einzelnen  $\beta$ -Strangs mit dem  $\beta 2$ -Strang von SmF eine weitere Untereinheit des Sm-Pentamers (Abbildung 17d). Zusätzlich erfolgt auch eine Bindung von SmE durch die Ausbildung zweier Salzbrücken zwischen den Resten Arg44<sub>G2</sub> und Glu52<sub>SmE</sub> sowie Glu31<sub>G2</sub> und Gln26<sub>SmE</sub>. Außerdem wird durch die Zusammenlagerung einiger Seitenketten der  $\alpha 1$ -Helices beider Proteine (Tyr36<sub>Gemin2</sub>, Leu37<sub>Gemin2</sub> und Ile18<sub>SmE</sub>, Ile21<sub>SmE</sub>, Phe22<sub>SmE</sub>, Leu25<sub>SmE</sub>), sowie Phe50<sub>SmE</sub> ein hydrophobes Cluster gebildet (Abbildung 17e). Somit kontaktiert Gemin2 das Sm-Pentamer an gegenüberliegenden Seiten des Sm/pICln-Rings und erkennt mit SmD1, D2, E und F gleichzeitig Untereinheiten, aus beiden der ursprünglichen Heterooligomere SmD1-D2 und SmE-F-G.

Von SMN konnten in der Kristallstruktur interessanterweise nur die Positionen Ala8<sub>SMN</sub>-Gly24<sub>SMN</sub> aufgeklärt werden, obwohl die N-terminalen 122 Aminosäuren Bestandteil der verwendeten C-terminal gekürzten SMN-Variante sind. Diese falten größtenteils in eine  $\alpha$ -Helix (Leu14-Val23) und binden im Bereich des C-terminalen Helixbündels an Gemin2. Hierbei bildet die SMN-Helix durch Interaktionen mit den vier  $\alpha$ -Helices  $\alpha 2$ ,  $\alpha 4$ ,  $\alpha 6$  und  $\alpha 8$  von Gemin2 ein Netzwerk aus polaren Interaktionen, welches durch eine hydrophobe Kontaktfläche ergänzt wird (Abbildung 17f). Bemerkenswerterweise sind in der Kristallstruktur keine Kontakte zwischen SMN und den Sm-Proteinen zu beobachten.

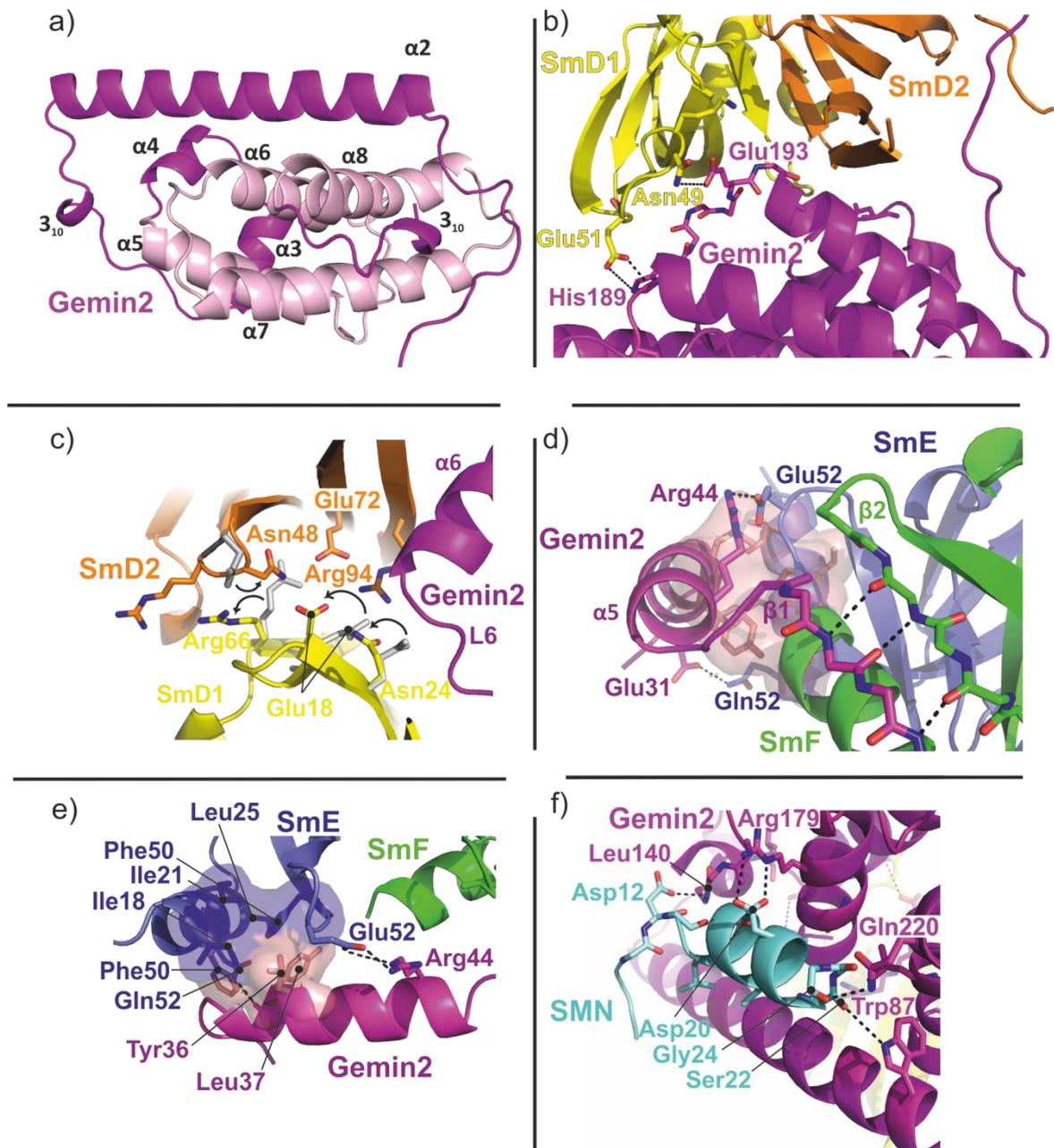


Abbildung 17 Die Struktur von Gemin2 und seine Kontakte zu SMN $\Delta$ C und den Sm-Proteinen. a) Die Helixbündel-Domäne von Gemin2. Das Vierhelixbündel ist in hellem Magenta farblich hervorgehoben. b) Gemin2 kontaktiert SmD1 und SmD2 im Bereich ihrer Kontaktfläche. c) Die Bindung von Gemin2 verursacht im Vergleich zum 6S-Komplex eine Kaskade von Seitenkettenumlagerungen in SmD1 und SmD2. d) Die N-terminale Region von Gemin2 kontaktiert SmE und SmF. Hierbei erfolgt eine antiparallele Paarung der  $\beta$ -Stränge  $\beta$ 1 von Gemin2 und  $\beta$ 2 von SmF. e) Ein hydrophobes Cluster bildet sich zwischen SmE und Resten der N-terminalen Region von Gemin2. f) Kontakte der Gemin2-bindenden Helix von SMN zu Gemin2.

## Ergänzende Strukturanalysen des 8S-Komplexes auf Basis der Elektronenmikroskopie

Die kristallographischen Arbeiten wurden durch elektronenmikroskopische Strukturanalysen des 8S-Komplexes ergänzt, die nach der erfolgreichen Lösung der Kristallstruktur insbesondere das Ziel hatten, die Tudor-Domäne von SMN zu lokalisieren. Hierbei kam die Methode der Einzelpartikel-

Negativfärbung-Elektronenmikroskopie zur Anwendung, nachdem die Partikel für die Aufnahmen mit der GraFix-Methode präpariert wurden (Kastner et al., 2008; Stark, 2010). Die elektronenmikroskopischen (EM-) Aufnahmen wurden vor Beginn der kristallographischen Arbeiten begonnen und durch Dr. Ashwin Chari in Kooperation mit Prof. Dr. Holger Stark durchgeführt. Die Rekonstruktion eines 3D-Modells aus den elektronenmikroskopischen 2D-Klassensummen gelang jedoch erst später mit Hilfe eines Tiefpass-gefilterten Modells der 8S-Kristallstruktur (siehe auch Abschnitt 4.3, Seite 40). Die Auswertung der resultierenden EM-Struktur erfolgte anschließend durch ein molekulares Docking der 8S-Kristallstruktur innerhalb der EM-Elektronendichte. Dabei ergibt sich auf Seiten der elektronenmikroskopischen Struktur eine, nicht durch das Modell der 8S-Kristallstruktur erklärbare, Differenz-Elektronendichte. Diese befindet sich auf der kegelförmigen Fläche des Sm/pICln-Rings, nahe der C-terminalen Helixbündel-Domäne von Gemin2. Innerhalb der 8S-Kristallstruktur konnten weite Sequenzbereiche von SMN $\Delta$ C (Leu25-Trp122) und auch ein Großteil der L5-Schleife von pICln nicht lokalisiert werden. Um festzustellen welches der abwesenden Sequenzbereiche für die elektronenmikroskopische Differenzdichte verantwortlich sein könnte, wurden 8S-Komplexe präpariert, in denen die entsprechenden Bereiche der beiden Proteine deletiert wurden. So entstand eine Präparation des 8S-Komplexes, die mit einem im Bereich der L5-Schleife gekürzten pICln-Konstrukt rekonstituiert wurde (8S $\Delta$ L5). Außerdem wurde eine weitere Variante des 8S-Komplexes präpariert, in dem SMN auf die Reste Ala8-Gly24 reduziert war (8S $\Delta$ Tudor). Diese modifizierten 8S-Komplex-Präparationen wurden anschließend von Dr. Ashwin Chari elektronenmikroskopisch untersucht. Bei der Analyse der 8S $\Delta$ L5 EM-Struktur fiel auf, dass zwar weiterhin eine Differenzdichte zur 8S-Kristallstruktur vorhanden ist, diese jedoch, im Vergleich zur ursprünglichen Differenzdichte, ein wesentlich verringertes Volumen aufweist. Ein Teil der im ursprünglichen Komplex identifizierten Differenzdichte verschwindet also bei einer Deletion der L5-Schleife von pICln. Das im 8S $\Delta$ L5-Komplex verbleibende Dichteelement ist hierbei ausreichend, um beispielsweise ein Tiefpass-gefiltertes Modell der SMN-Tudor-Domäne aufzunehmen. Um deren mögliche Positionierung innerhalb der verbleibenden Differenzdichte zu verifizieren, wurde außerdem die EM-Struktur des 8S $\Delta$ Tudor-Komplexes untersucht. Bei diesem war im Vergleich zur 8S-Kristallstruktur keine elektronenmikroskopische Differenzdichte mehr zu erkennen (Grimm et al., 2013).

### **Vergleich der Strukturen der 6S- und 8S-Komplexe mit der Struktur des 7S-Komplexes**

Parallel zu den Arbeiten an den 6S- und 8S-Kristallstrukturen konnte durch eine konkurrierende Arbeitsgruppe die Kristallstruktur eines weiteren Intermediates der Sm-Core-Zusammenlagerung gelöst werden (Abbildung 6, Seite 16) (Zhang et al., 2011). Dieser Komplex enthielt das Sm-Pentamer D1-D2-E-F-G, sowie Gemin2 und die Gemin2-bindende  $\alpha$ -Helix von SMN, jedoch nicht pICln. Damit

entspricht dieser Komplex im Wesentlichen dem späten 7S-Intermediat, bei dem pICln bereits dissoziiert ist (Abbildung 6, Seite 16) (Chari et al., 2008). Durch die Abwesenheit von pICln ist der Sm-Ring geöffnet, wodurch sich für diesen Komplex bei einer Projektion entlang der Ringachse ein hufeisenförmiges Aussehen ergibt (Abbildung 18b, Seite 60). Die Veröffentlichung der Koordinaten dieses Strukturmodells war essentiell für die Strukturlösung des 8S-Komplexes, die mit der Methode des Molekularen Ersatzes gelang (siehe auch Abschnitt 4.3, Seite 38 sowie Abschnitt 5.2, Seite 67). Trotz der Verfügbarkeit verschiedener Kristallstrukturen von SmD1-D2 (Kambach et al., 1999), SmE-F-G (Leung et al., 2011; Pomeranz Krummel et al., 2009; Weber et al., 2010) und der NMR-Struktur der erweiterten PH-Domäne von pICln (Furst et al., 2005), scheiterten vorher mehrfache Versuche der Strukturlösung mit dieser Methode. Für diese wurden verschiedene hexamere Modelle auf Basis des ebenfalls ringförmigen bakteriellen Sm-Proteinhomologs Hfq generiert (Sauter et al., 2003). Später zeigte sich jedoch, dass diese Modelle große Koordinatendifferenzen zur 8S-Kristallstruktur aufwiesen. Außerdem enthielten diese keine strukturellen Informationen zu SMN und Gemin2. Dies sind naheliegende Gründe für das Scheitern des Molekularen Ersatzes mit diesen Modellen.

Bei der Verfeinerung der 8S-Kristallstruktur wurde deutlich, dass auch zwischen dem 8S-Komplex und dem 7S-Komplex signifikante Koordinatendifferenzen bestehen. Diese ergeben sich durch leicht unterschiedliche äußere Abmessungen, die in ähnlicher Weise bereits beim Vergleich des Sm/pICln-Rings des 6S-Komplexes mit dem Sm-Proteinring der Sm-Core Kristallstruktur festgestellt wurden (siehe unter Abschnitt 5.1, Seite 43). Diese Differenzen ergeben sich in erster Linie dadurch, dass pICln im 6S-Komplex eine deutlich geringere radiale Fläche einnimmt, als die beiden Sm-Proteine SmB'B-D3, die im Sm-Core eine entsprechende Position einnehmen (siehe Abschnitt 5.1 und Abbildung 7c, Seite 44). Dadurch verringert sich die Distanz zwischen SmD1 und SmG, sodass der Sm/pICln-Ring entlang dieser Achse schmaler ist als das Sm-Core (Abbildung 7c). Innerhalb des 7S-Komplexes ist diese Distanz größer als in den 6S- und 8S-Komplexen, jedoch kleiner als im Sm-Core. Dieser Sachverhalt kann durch die Messung der Distanz zwischen jeweils einem Referenzatoms von SmD1 (Leu58) und SmG (Leu70) verdeutlicht werden (Abbildung 18a). Die Situation innerhalb des 7S-Komplexes muss aufgrund der Ringöffnung im Bereich von pICln beziehungsweise SmB/B'-D3 als entspannte Konformation angesehen werden. Innerhalb des 6S- und 8S-Komplexes werden die Sm-Proteine also zusammengezogen, während sie im Sm-Core auseinander gedehnt werden. Die Ringspannung herrscht jedoch nicht nur entlang der Ringachse, sondern besitzt zusätzlich auch eine Komponente, die orthogonal zur Ringachse gerichtet ist. Dies wird durch eine Überlagerung der Ringstrukturen auf Basis der Koordinaten der jeweiligen SmD1 Untereinheiten deutlich. Im Vergleich zu seiner Position innerhalb des Sm-Core und den 6S- und 8S-Komplexen ist SmG im 7S-Komplex in Richtung der flachen Ringfläche verschoben (Abbildung 18b).

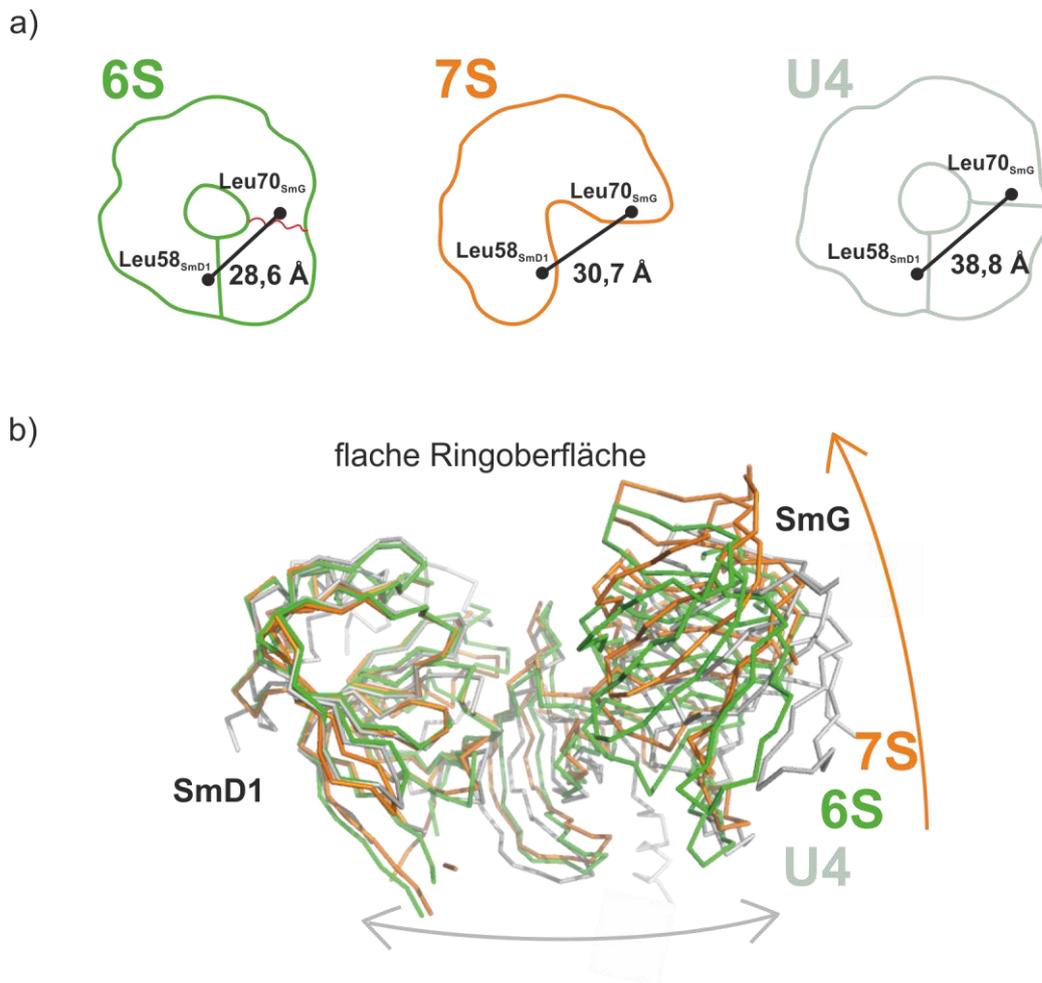


Abbildung 18 Der Ringschluss im 6S-Komplex und im Sm-Ring der Sm-Core erzeugt eine Spannung mit einer Komponente entlang der Ringebene und orthogonal zu dieser. a) Projektionen des 6S- und 7S-Komplexes, sowie des Sm-Rings des U4-snRNP entlang der Ringachse. pICln und SmB/B'-D3 nehmen unterschiedliche radiale Flächen zwischen SmD1 und SmG ein. Die sich daraus ergebenden unterschiedlichen Öffnungsweiten zwischen SmD1 und SmG sind durch die Distanz von Leu58<sub>SmD1</sub> und Leu70<sub>SmG</sub> angegeben. b) Während sich die Proteine innerhalb des Sm/pICln-Rings des 6S-Komplexes und im Sm-Proteinring des Sm-Core innerhalb einer Ringebene befinden, ist SmG bei einer Überlagerung auf SmD1 in Richtung der flachen Ringfläche aus dieser Ebene verschoben.

## Einblicke in die Moleküldynamik des 8S-Komplexes

### *Konformationelle Differenzen zwischen den 20 Kopien des 8S-Komplexes in der AU*

Die asymmetrische Einheit der P2<sub>1</sub>-Zelle der 8S-Kristalle enthielt 20 Kopien des Komplexes, die untereinander kleine konformationelle Differenzen zeigen. Die Positionsunterschiede lassen sich gut durch eine graphische Visualisierung der RMS-Positionsdifferenzen zwischen äquivalenten Positionen der 20 Moleküle veranschaulichen (Abbildung 19a, Seite 63). Hierbei wird deutlich, dass die Differenzen größtenteils im Kontaktbereich zwischen pICln und SmG kumulieren und hauptsächlich orthogonal zur Ringebene bestehen (Abbildung 19a). Der Kontakt zwischen pICln und SmG stellte mit 600 Å<sup>2</sup> bereits in der 6S-Kristallstruktur die kleinste Interaktionsfläche zwischen den Untereinheiten

dar und ist in den 20 8S-Kopien der AU nochmals um 10 bis 25% verringert. Die Kontaktfläche schwankt hierbei zwischen den einzelnen Molekülen um Werte von 450 bis 550 Å<sup>2</sup>, während für alle anderen Kontaktflächen stabile Werte ermittelt werden können (Abbildung 19b).

### ***Normalmodenanalyse und moleküldynamische Simulationen***

Die zuvor beschriebenen Differenzen treten zwischen unterschiedlichen Komplexmolekülen auf, die sich innerhalb des Kristallgitters in verschiedenen Umgebungen befinden. Daher handelt es sich nur um potentielle Hinweise auf dynamische Molekülbewegungen des 8S-Komplexes. Mit Hilfe der Normalmodenanalyse (ben-Avraham and Tirion, 1998) und moleküldynamischen (MD)-Simulationen sollten diese Beobachtungen durch *in silico* Methoden unterstützt werden. Die Normalmodenanalyse beschreibt das Schwingungsverhalten von Molekülen mit harmonischen Schwingungen. Da die Proteine in dieser Methode mit relativ einfachen mechanischen Modellen beschrieben werden können (Hinsen, 1998; Tirion, 1996), eignet sie sich besonders für die Analyse längerer Simulationszeiträume und von großen Proteinen und Proteinkomplexen (Tama et al., 2003). Aufgrund der Vermutung, dass für die Dissoziation von pICln eine Ringöffnung im Bereich des pICln-SmG-Kontaktes stattfinden würde, wurde für die Normalmodenanalyse des 8S-Komplexes der Kontakt zwischen pICln und SmG aufgebrochen und die Ringöffnung entsprechend der im 7S-Komplex gefundenen Konformation erweitert (Vergleiche Abbildung 18, Seite 60). Dies geschah durch eine Überlagerung der einzelnen Untereinheiten des 8S-Untereinheiten mit ihren jeweiligen Pendants in 3S6N, während SmD1 und pICln als gemeinsame Einheit verschoben wurden. Durch diese Verschiebung besteht kein Kontakt mehr zwischen SmG und pICln. Letzteres ist daher nur noch an SmD1 gebunden, während SmG nur durch SmE kontaktiert wird. Dadurch sind pICln und SmG entlang ihres gemeinsamen Kontaktbereichs zueinander frei beweglich. In dieser geöffneten Konformation erfolgte anschließend die Normalmodenanalyse des 8S-Komplexes unter Verwendung des éINemo-Server (Suhre and Sanejouand, 2004)

Während die Normalmoden 1-6 nur die Translations- beziehungsweise Rotationsbewegung eines Festkörpers beschreiben, repräsentieren die höheren Normalmoden elastisch deformierende Schwingungen. Hierbei zeigt bei der Analyse des 8S-Komplexes die Normalmode 7 eine Bewegung, die einen Übergang zwischen den beobachteten Positionsdifferenzen der 20 8S-Kopien beschreiben kann. Die Perturbation verläuft ebenfalls orthogonal zur Ringebene und ist im Kontaktbereich zwischen pICln und SmG maximal ausgeprägt. Die Trajektorie wird durch insgesamt 21 Zustände beschrieben. Zustand 1 zeigt eine maximale Verschiebung von SmG aus der Ringebene in Richtung der flachen Ringoberfläche, während pICln maximal in Richtung der kegelförmigen Seite verschoben ist (Abbildung 19c, Seite 63). Diese Begebenheit kehrt sich beim Verfolgen der einzelnen Zustände um, sodass bei Zustand 21 pICln in Richtung der flachen Seite und SmG in Richtung der kegelförmigen

Seite aus der Ringebene verschoben ist (Abbildung 19c). Die hierbei durch den Komplex beschriebene Bewegung weist eine große Ähnlichkeit zu den Koordinatendifferenzen innerhalb der unterschiedlichen 8S-Kopien auf (vergleiche Abbildung 19a mit Abbildung 19c).

Zusätzlich wurden durch Dr. Clemens Grimm MD-Simulationen mit dem 8S- und auch dem 6S-Komplex durchgeführt, die in eine ähnliche Richtung deuten. Der 8S-Komplex zeigt in diesen Untersuchungen im Vergleich zum 6S-Komplex deutlich verstärkte dynamische „Atmungsbewegungen“. Diese Bewegungen sind sowohl im 6S- als auch im 8S Komplex im Bereich der pICln-SmG-Kontaktfläche am stärksten ausgeprägt (Grimm et al., 2013).

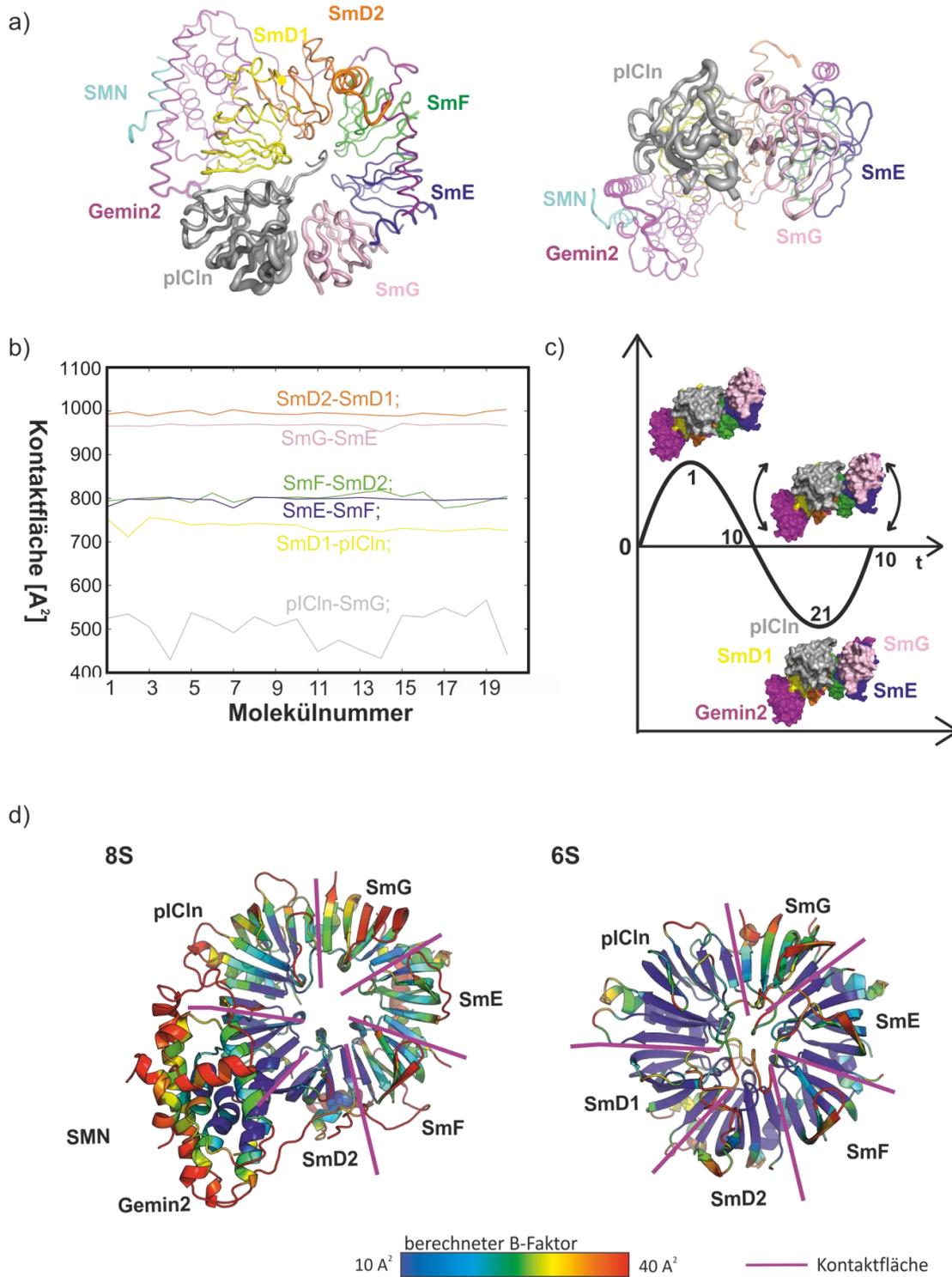
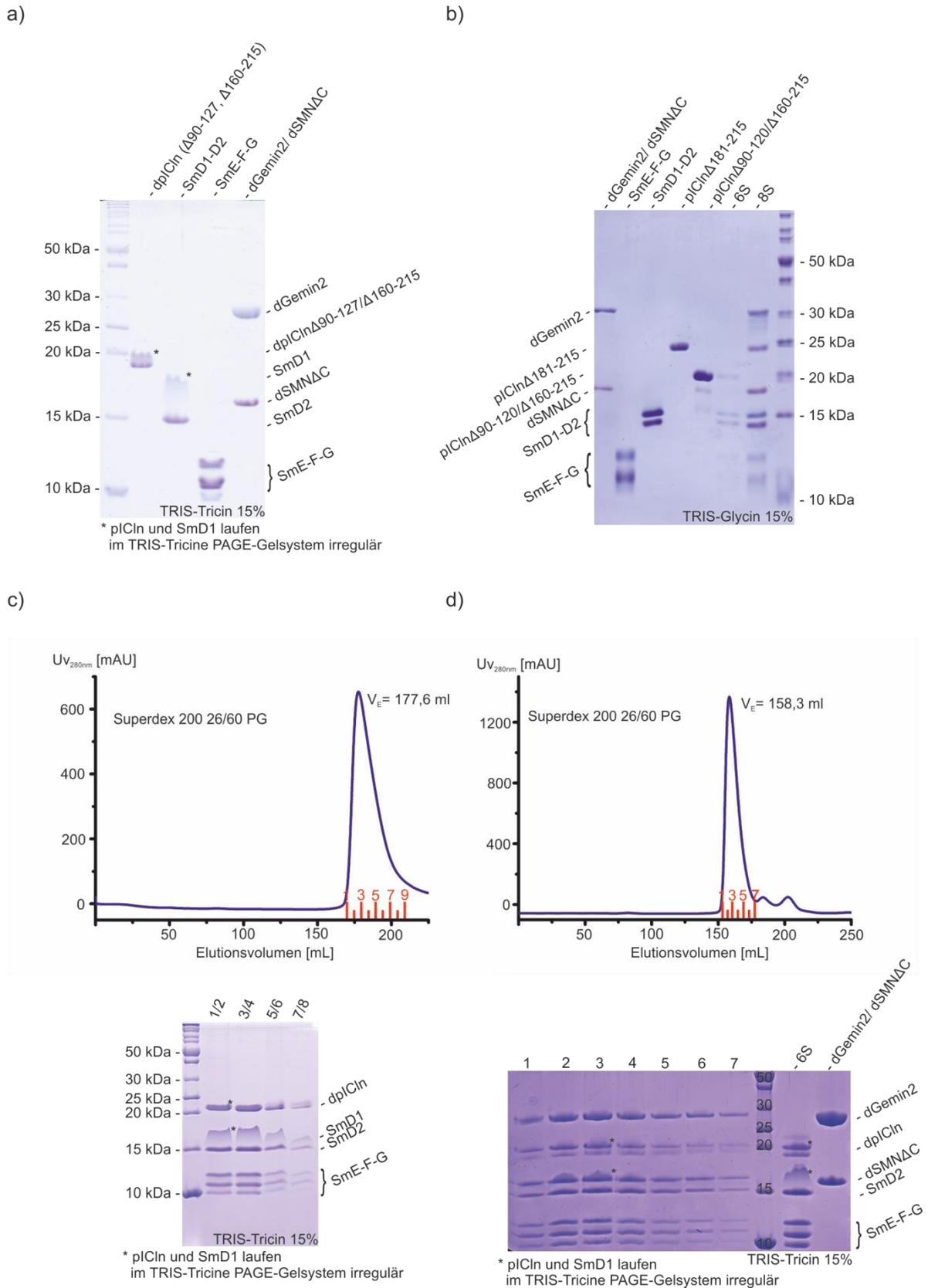


Abbildung 19 Konformationelle Unterschiede zwischen den 20 Kopien des 8S-Komplexes innerhalb der asymmetrischen Einheit der 8S-Kristalle. a) Röhren-Darstellung des 8S-Komplexes mit Blick auf die flache Ringseite (links) und entlang der Ringebene (rechts). Die RMS-Koordinatendifferenzen sind durch den Durchmesser der Röhre präsentiert. b) Kontaktflächen zwischen den Komplexuntereinheiten innerhalb der 20 Kopien des 8S-Komplexes. c) Darstellung der Trajektorie der Normalmoden Analyse. Diese umfasste 21 Zustände, wobei in Zustand 1 eine maximale Verschiebung von SmG aus der Ringebene in Richtung der flachen Ringoberfläche zeigt, während pICln maximal in Richtung der kegelförmigen Seite verschoben ist. Diese Begebenheit kehrt sich beim Verfolgen der einzelnen Zustände um, sodass bei Zustand 21 pICln in Richtung der flachen Seite und SmG in Richtung der kegelförmigen Seite aus der Ringebene verschoben ist. d) Für den 8S- und 6S- Komplex wurden MD-Simulationen durchgeführt. Die Kolorierung entspricht den aus den Koordinatendifferenzen ermittelten B-Faktoren von jeweils 1500 Zuständen (entsprechend einer Simulationszeit von 3ns). Mit Erlaubnis von Cell Press. Teil (a), b) und d) verändert nach Grimm, C. et al.,2013).

## **5.2 Strategie zur Lösung der Kristallstrukturen des 6S- und 8S-Komplexes**

### **Präparation und Rekonstitution des 6S- und 8S-Komplexes**

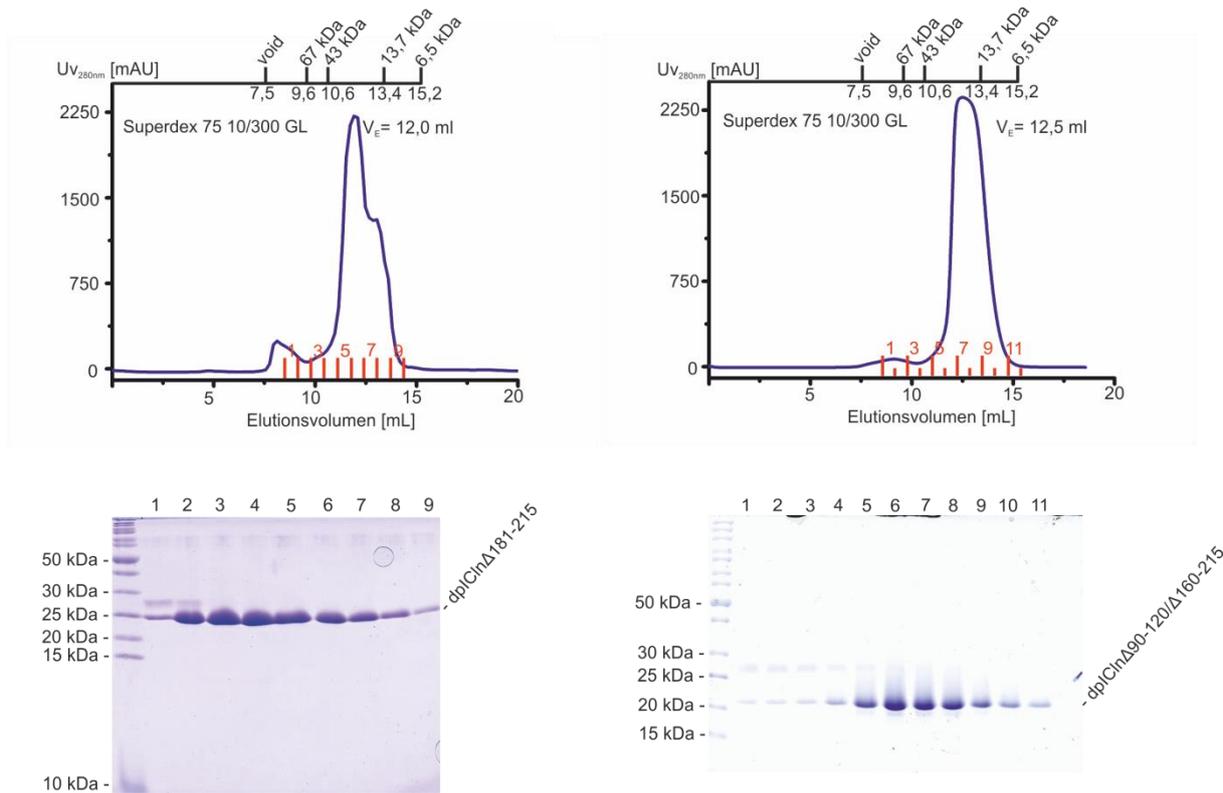
Alle Proteine wurden unter Verwendung eines bakteriellen Expressionssystem rekombinant exprimiert und präpariert. Für die Rekonstitution des 6S-Komplexes erfolgte die Präparation der Sm-Proteine in Form der Heterooligomere SmD1-D2 und SmE-F-G, während pICln getrennt von diesen aufgereinigt wurde (Abbildung 20, Seite 65). Die gereinigten Untereinheiten wurden anschließend in äquimolaren Mengen unter Hochsalzbedingungen bei 1M NaCl zusammengegeben und die Salzkonzentration wurde in drei Stufen über Nacht per Dialyse langsam auf physiologische Konzentrationen gesenkt. Für die Rekonstitution des 8S-Komplexes wurde zu den vorgefertigten 6S-Präparationen die gemeinsam aufgereinigten SMN $\Delta$ C-Gemin2 Präparationen in stöchiometrischen Mengen hinzugefügt. Die korrekte Formierung der Komplexe wurde jeweils durch Größenausschlusschromatographie (size-exclusion chromatography, SEC) mit anschließender elektrophoretischer Auftrennung der Fraktionen in SDS-PAGE Gelen überprüft (Abbildung 20, Seite 65). Eine detaillierte Ausführung der Protokolle findet sich im Methodenteil dieser Arbeit (Abschnitt 4.3, Seite 31).



**Abbildung 20** Präparationen der Komplexuntereinheiten und Rekonstitution der 6S- und 8S-Komplexe. a) 15% TRIS-Tricine SDS-PAGE Gel der piCln ( $\Delta 90-127 \Delta 160-215$ ), SmD1-D2, SmE-F-G und SMN $\Delta$ C-Gemin2 Präparationen. b) 15% TRIS-Glycin SDS-PAGE Gel der SMN $\Delta$ C-Gemin2, SmE-F-G, SmD1-D2, piCln ( $\Delta 181-215$ ), piCln ( $\Delta 90-120 \Delta 160-215$ ), 6S- und 8S-Präparationen. c) SEC-Chromatogramm und zugehöriges 15% TRIS-TRICINE SDS-PAGE Gel einer typischen 6S-Präparation. d) SEC-Chromatogramm und zugehöriges 15% TRIS-TRICINE SDS-PAGE Gel einer typischen 8S-Präparation.

## Kristallisation und Strukturlösung

Ein ausgedehntes Hochdurchsatz-Screening der 6S- und 8S-Komplexpräparationen mit kommerziellen Screens war zu Beginn nicht erfolgreich. Die hierbei getesteten Bedingungen beruhten neben verschiedenen Salzen hauptsächlich auf Polyethylenglykolen (PEG) unterschiedlicher Molekulargewichte. Um den durch das Screening abgedeckten Kristallisationsraum zu erweitern, wurde ein auf alternativen Präzipitanten beruhender Screen entwickelt (Grimm et al., 2010). Eine der in diesem Screening getesteten Bedingungen enthielt die Polyacrylsäure Sokalan CP-42, mit welcher eine erste Kristallform des 8S-Komplexes erhalten wurde (Abbildung 22a, Seite 70). Beugungsbildern dieser Kristalle, die an den Synchrotrons SLS und ESRF aufzeichnet wurden, hatten jedoch ein sehr begrenztes Beugungslimit von ungefähr  $8 \text{ \AA}$  in der äußeren Schale. Umfangreiche Bestrebungen diese Kristalle zu optimieren, zeigten keine signifikante Verbesserung der Beugungsdaten, weshalb anschließend auf eine Strategie zur Optimierung der Proteinsequenzen mit dem Ziel eines verbesserten Kristallisationsverhaltens gesetzt wurde. Hierbei fiel die Wahl auf pICln, da dieses sowohl Bestandteil des 6S- als auch des 8S-Komplexes war, sodass beide Teilprojekte potentiell von diesen Optimierungen profitieren konnten. Außerdem war zu Beginn der Arbeiten bereits die NMR-Struktur (Furst et al., 2005) eines *Canis lupus familiaris* pICln-Orthologs verfügbar, die eine strukturelle Analyse des Proteins und eine gerichtete Sequenzmanipulation ermöglichte. pICln besitzt eine erweiterte Pleckstrin-Homologie (PH)- Domäne (Haslam et al., 1993), die durch die N-terminalen Reste (entsprechend *Drosophila melanogaster* pICln Positionen 1-159) des Proteins gebildet wird. Für den C-Terminus (161-215) kann keine Sekundärstruktur vorhergesagt werden. Die NMR-Studien zeigten, dass auch die Reste 93-115 von *C. l. familiaris* pICln (entsprechend Positionen *D. melanogaster* pICln 78-132, Vergleiche mit Sequenz-Alignment, Abbildung 27, Seite 99 im Anhang) eine erhöhte Flexibilität aufweisen. Da der Einbau solcher flexibler unstrukturierter Proteinregionen in ein Kristallgitter mit einem erhöhten Entropieverlust einhergeht, kann die Präsenz ausgedehnter Termini und Schleifenregionen die Kristallisation der betreffenden Proteine verhindern (Derewenda, 2010). Die gezielte Entfernung solcher flexibler Proteinregionen kann daher die Kristallisationsbereitschaft der jeweiligen Proteine erhöhen und ist eine weithin akzeptierte Methode, die vielfach erfolgreich angewendet wurde (Chen et al., 1998; Dale et al., 1999; Martin et al., 2000; Yeh et al., 1996). Begonnen wurde mit einer Kürzung des C-Terminus von pICln um jeweils fünf Aminosäuren, bis hin zu Position 159, die das C-terminalen Ende der PH-Domäne markiert. Anschließend wurden innerhalb der hierdurch entstandenen Vektor-Konstrukte verschiedene Deletionen im Bereich der Reste 90-130 eingeführt. Alle diese Varianten von pICln ließen sich löslich exprimieren und eluierten als monodisperse *Peaks*, deren Elutionsvolumen einem monomeren Oligomerisierungsstatus entsprachen (Abbildung 21). Mit diesen pICln-Varianten rekonstituierte Präparationen des 6S- und 8S-Komplexes wurden anschließend durch umfassende Kristallisation-Screening analysiert.



**Abbildung 21** SEC-Chromatogramme und zugehörige 12% TRIS-Glycin SDS-PAGE Gele von pICln-Präparationen. Links pICln ( $\Delta 181-215$ ), das für die Präparation und Kristallisation des 8S-Komplexes verwendet wurde. Rechts pICln ( $\Delta 90-120 \Delta 160-215$ ), das für die Präparation und Kristallisation des 6S-Komplexes verwendet wurde.

### **Kristallisation und Strukturlösung des 8S-Komplexes**

Während anschließenden Kristallisations-Screenings zeigte eine 8S-Präparation, die mit einer C-terminal gekürzten Variante von pICln ( $\Delta 181-215$ ) rekonstituiert wurde, ein verändertes Kristallisationsverhalten. Mit dieser Präparation wurden Kristalle erhalten (Abbildung 22b, Seite 70), die signifikant größerer wurden als die ursprünglichen 8S-Kristalle, hierbei aber die gleiche Raumgruppe und ähnliche Zellparameter zeigten (Raumgruppe  $P2_1$ ,  $a = 150 \text{ \AA}$ ,  $b = 358 \text{ \AA}$ ,  $c = 230 \text{ \AA}$ ,  $\alpha = 90^\circ$ ,  $\beta = 97.7^\circ$ ,  $\gamma = 90^\circ$ ). Während die ursprüngliche Kristallform ausschließlich in Bedingungen mit der Polyakrylsäure Sokalan CP-42 wuchs, konnte die neue Kristallform außerdem auch mit PEG- und PEG-MME-Polymeren unterschiedlicher Molekulargewichte erhalten werden. Da diese Kristalle weiterhin eine eingeschränkte Beugungskraft hatten, war es aufgrund von Strahlungsschäden nicht möglich, vollständige Datensätze von Einkristallen mit einer Auflösung von mehr als  $5 \text{ \AA}$  zu messen. Daher wurde eine Optimierung der Kristallisationsbedingung (Tabelle 1, Seite 71) angestrebt, die ein langsames Wachstum der Kristalle zu einer Größe von mehreren 100 Mikrometern Kantenlänge ermöglichte. Um das Volumen großer Kristalle effektiv zu nutzen, ist es möglich partielle Datensätze an unterschiedlichen Positionen entlang der Spindelachse zu messen und diese anschließend zu vereinigen und zu skalieren (Bourenkov and Popov, 2010). Am ESRF existiert außerdem die

Möglichkeit vor Beginn der Messung zwei Punkte auf dem Kristall festzulegen, sodass die Rotation während der Messung dann entlang der Spindelachse in helikaler Weise erfolgt (Flot et al., 2010). Leider zeigte sich, dass unterschiedliche Bereiche der 8S-Kristalle eine zu hohe Anisotropie aufwiesen, sodass eine effektive Skalierung partieller Datensätze nicht möglich war. Datensätze, die an Messstation ID14-4 des ESRF gemessen wurden, zeigten hingegen deutlich bessere Statistiken. Diese Messstation zeichnet sich durch einen vergleichsweise defokussierten Strahl aus, der deshalb ein relativ großes Kristallvolumen belichten kann (McCarthy et al., 2009). Unter ungefähr 100 Datensätzen, die an dieser Messstation aufgezeichnet wurden, zeigte ein einzelner Datensatz verwendbares Signal bis zu einem Diffraktionslimit von 3,1 Å (Karplus and Diederichs, 2012). Diese Auflösungsgrenze galt jedoch nur entlang der reziproken  $a^*$ -Achse, sodass mit Hilfe des *Diffraction Anisotropy Server* (Strong et al., 2006) eine ellipsoide Trunkation mit den Auflösungsgrenzen von 3,1 Å, 3,8 Å und 4,0 Å entlang der  $a^*$ ,  $b^*$  und  $c^*$  -Achsen durchgeführt wurde.

Zur experimentellen Phasierung konnten erfolgreich Schweratomderivate der 8S-Kristalle unter Verwendung eines  $Ta_6Br_{12}$ -Clusters (Knablein et al., 1997) erzeugt werden. Da für die Aufnahme von anomalen Beugungsdaten an der Absorptionskante des Schwermetalls gemessen werden muss, erhöht sich auch die durch den Kristall absorbierte Strahlungsenergie. Da sich damit auch mit der Messung einhergehende Strahlungsschäden erhöhen, muss die Intensität der Röntgenstrahlung entsprechend verringert werden. Aufgrund der intrinsisch geringen Beugungskraft der 8S-Kristalle scheiterte die experimentelle Phasierung jedoch trotz der erfolgreichen Derivatisierung mit dem Tantalum-Cluster. Die Strukturlösung gelang erst mit Hilfe des Molekularen Ersatzes unter Verwendung der Koordinaten der Kristallstruktur des späten Intermediates (PDB Eintrag 3S6N) (Zhang et al., 2011). Die Suche lieferte 20 Moleküle innerhalb der asymmetrischen Einheit. Nach einer ersten Verfeinerungsrunde erfolgte die manuelle Platzierung der erweiterten PH-Domäne von pICln unter Verwendung der Koordinaten von PDB 1ZYI innerhalb der entsprechenden Differenzdichten (siehe auch Abschnitt 4.3, Seite 39).

### ***Kristallisation und Strukturlösung einer ersten 6S-Kristallstruktur***

Während eine Kürzung des C-Terminus von pICln zur erfolgreichen Kristallisation entsprechender 8S-Komplex Präparationen führte, zeigten sich bei korrespondierenden 6S-Präparationen zuerst keine vergleichbaren Erfolge. Daher wurden weitere Fragmente von pICln erzeugt, bei denen die C-terminalen Kürzungen mit einer Deletion der Reste 90-120 kombiniert wurden. Mit Präparationen des 6S-Komplexes, die mit einer solchen pICln-Variante ( $\Delta 90-120$ ,  $\Delta 160-215$ ) rekonstituiert wurden, konnte eine erste reproduzierbare Kristallform erhalten werden (Abbildung 22c, Seite 70, Tabelle 1, Seite 71). Während diese Kristallform ein homogenes Wachstumsverhalten mit einer Kantenlänge von 300-400  $\mu\text{m}$  zeigten und hierbei sehr zuverlässig zu reproduzieren waren, war ihre Beugungskraft

limitiert. Es wurden ungefähr 100 Datensätze von diesen Kristallen gemessen, wobei das Auflösungsmaximum der besten Datensätze bei ungefähr 4 Å lag. Die Kristalle gehörten zu Raumgruppe P321 mit den Zellparametern  $a = b = 203,9 \text{ \AA}$ ,  $c = 215,9 \text{ \AA}$ ,  $\alpha = \beta = 90,0^\circ$ ,  $\gamma = 120^\circ$  (Tabelle 2, Seite 72). Es gelang mit Hilfe der Methode des Molekularen Ersatzes ein vorläufiges Strukturmodell dieser Kristalle zu lösen. Hierfür wurden die Koordinaten des vorläufigen 8S-Strukturmodells verwendet, aus denen SMN $\Delta$ C und Gemin2 entfernt wurden. Die Kristalle gehörten zur Raumgruppe P321 (Kristallform 6S-P321) und die asymmetrische Einheit enthielt 18 Kopien des 6S-Komplexes. Wegen der geringen Auflösung und der hohen Anzahl an Komplexmolekülen innerhalb der asymmetrischen Einheit, war nicht gesichert, ob aus den gemessenen Datensätzen ein Strukturmodell mit zufriedenstellender Qualität verfeinert werden könnte. Daher wurde parallel weiter daran gearbeitet, eine weitere Kristallform des 6S-Komplexes zu erhalten (siehe auch Abschnitt 4.3, Seite 39).

### ***Kristallisation und Strukturlösung einer weiteren 6S-Kristallstruktur***

Für die erfolgreiche Kristallisation einer weiteren Kristallform des 6S-Komplexes, war neben den beschriebenen Deletionen flexibler Bereiche von pICln, die gezielte Modifikation eines Oberflächenrestes des Proteins notwendig (Tabelle 1, Seite 71). Hierbei wurde eine Strategie verwendet, die unter dem Namen *Surface Entropy Reduction* (SER) bekannt ist. Die Idee hinter dieser Methode ist es, den bei der Kristallisation entstehenden lokalen Entropieverlust zu verringern, indem flexible Oberflächenreste durch Aminosäuren mit rigideren Seitenketten, wie beispielsweise Alanin ersetzt werden (Longenecker et al., 2001). Basierend auf der Annahme, dass Kristalle der 8S- und 6S-Komplexe aufgrund ihrer strukturellen Verwandtschaft auch Ähnlichkeiten bei der Architektur des Kristallgitters aufweisen könnten, wurden gezielt Positionen substituiert, die in Kristallkontakte innerhalb der 8S-Kristalle involviert waren. Eine genaue Analyse dieser Kontakte zeigte, dass pICln an der Ausbildung von zwei Kontaktflächen beteiligt ist (Abbildung 23a, Seite 70). Innerhalb dieser Kontaktflächen wurden verschiedene Reste identifiziert, die konformationell eingeschränkt waren. Einer dieser Reste war His144 von pICln, (Abbildung 23b) welcher für eine Substitution durch Alanin ausgewählt wurde. Entsprechende 6S-Komplexvarianten wurden mit einem entsprechenden pICln-Konstrukt ( $\Delta 90-120$ ,  $\Delta 160-215$ , H144A) rekonstituiert und präpariert. Mit dieser Komplex-Präparation konnte eine weitere Kristallform des 6S-Komplexes erhalten werden (Abbildung 22d, Seite 70) und von einem dieser Kristalle konnte ein hochauflösender Datensatz mit einem Auflösungslimit von 1,9 Å gemessen werden. Die Kristalle gehörten zur Raumgruppe C2 und werden daher als 6S-C2 Kristallform bezeichnet (Tabelle 2, Seite 72). Die Strukturlösung gelang mit Hilfe der Methode des Molekularen Ersatzes unter Verwendung der Koordinaten des 8S-Strukturmodells, aus denen SMN $\Delta$ C und Gemin2 entfernt wurden (siehe außerdem Abschnitt 4.3, Seite 38). Die asymmetrische Einheit

enthielt zwei Moleküle des 6S-Komplexes, die strukturell weitgehend identisch sind und daher in der Strukturanalyse nicht getrennt diskutiert wurden.

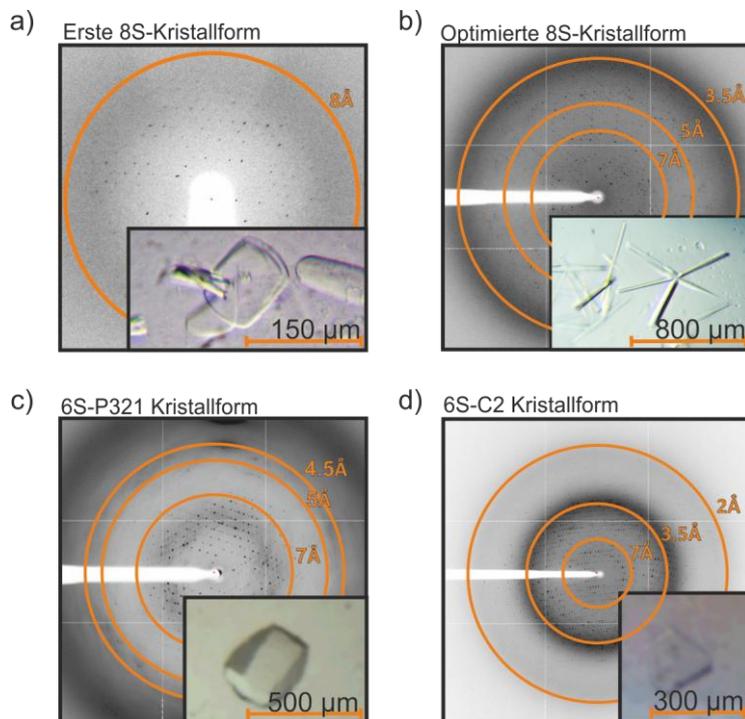


Abbildung 22 Beugungsbilder und Morphologie der erhaltenen 8S- und 6S-Kristallformen.

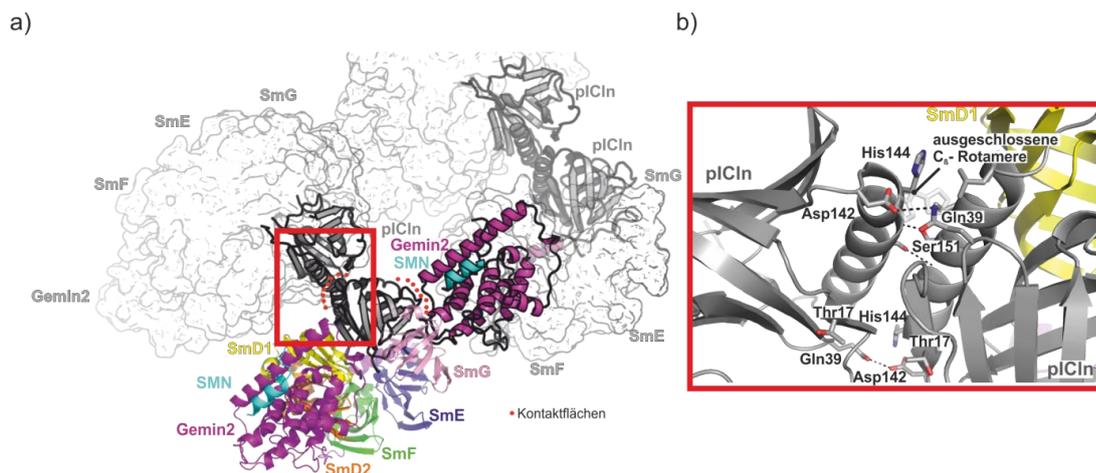


Abbildung 23 Durch pICln vermittelte Kristallkontakte innerhalb der 8S-Kristalle. a) Es sind insgesamt vier 8S-Moleküle dargestellt, von denen eines in Cartoon-Darstellung und drei weitere als Connolly-Oberflächen dargestellt sind. Die an den Kristallkontakten beteiligten Untereinheiten sind in Cartoon-Darstellung und ihrer jeweiligen Farbkodierung dargestellt. Die beiden Kontaktflächen sind durch rote Punkte markiert. Der rote Kasten markiert die Position des in Abbildung 23b gezeigten Ausschnitts. b) Der Ausschnitt zeigt den Kristallkontakt der zwischen zwei benachbarten pICln Molekülen ausgebildet wird. His144 ist konformationell eingeschränkt und die daher ausgeschlossenen C<sub>β</sub>-Rotamere sind für eines der beiden Reste in blassem grau gezeigt.

**Tabelle 1 Kristallisation**

Kristallform	8S
pICln Konstrukt	<i>D. melanogaster</i> pICln $\Delta$ 181-215
Methode	Dampfdiffusion im hängenden Tropfen
Kristallisationsplatte	SuperClear™ Polystyrolplatten
Temperatur (K)	289
Proteinkonzentration	15 g/l
Pufferkomposition Proteinlösung	150 mM NaCl; 10 mM HEPES pH 7,5; 5 mM TCEP
Kristallisationsbedingung	24% PEG 4000; 10% Ethanol; 150mM NaCl, pH 7,5
Kryo-Pufferkomposition	18% Jeffamine ED-2003 titriert auf pH 7, 10% Ethanol, 30% Glycerin
Tropfenvolumen und Verhältnis	7 $\mu$ l, 7:1 (Protein: Präzipitantz)
Volumen im Gefäßreservoir	300 $\mu$ l
<hr/>	
Kristallform	6S-C2
pICln Konstrukt	<i>D. melanogaster</i> pICln ( $\Delta$ 90-127 $\Delta$ 160-215 H144A)
Methode	Dampfdiffusion im hängenden Tropfen
Kristallisationsplatte	SuperClear™ Polystyrolplatten
Temperatur (K)	289
Proteinkonzentration	15 g/l
Pufferkomposition Proteinlösung	150 mM NaCl; 10 mM HEPES pH 7,5; 5 mM TCEP
Komposition Kristallisationsbedingung	15% Jeffamine ED-2003 titriert auf pH 7; 10% Ethanol
Kryo-Pufferkomposition	18% Jeffamine ED-2003, titriert auf pH 7; 10% Ethanol; 30% Glycerin
Tropfenvolumen und Verhältnis	1 $\mu$ l, 1:1
Volumen im Gefäßreservoir	300 $\mu$ l
<hr/>	
Kristallform	6S-P321
pICln Konstrukt	<i>D. melanogaster</i> pICln ( $\Delta$ 90-120 $\Delta$ 160-215)
Methode	Dampfdiffusion im hängenden Tropfen
Kristallisationsplatte	SuperClear™ Polystyrolplatten
Temperatur (K)	293
Proteinkonzentration	12 g/l

Pufferkomposition Proteinlösung	150 mM NaCl; 10 mM HEPES pH 7,5; 5 mM TCEP
Komposition Kristallisationsbedingung	2,7 M Natriumformiat; 0,1 M Imidazol pH 8,0
Kryo-Pufferkomposition	3 M Lithiumformiat; 0,1 M Imidazol pH 8,0
Tropfenvolumen und Verhältnis	1 $\mu$ l, 1:1
Volumen im Gefäßreservoir	300 $\mu$ l

**Tabelle 2 Datensammlung und Prozessierung**

Die Angaben in Klammern beziehen sich auf die höchste Auflösungsschale

Datensatz	8S	6S-C2	6S-P321
Anzahl Moleküle in AU	20	2	18
Röntgenquelle	ESRF Messstation ID14-4	ESRF Messstation ID14-4	SLS Messstation PXII
Wellenlänge (Å)	0,9762	0,9393	1,0
Temperatur (K)	100	100	100
Detektor	ADSC Quantum 315r	ADSC Quantum 315r	Pilatus 6M
Distanz Kristall-Detektor (mm)	478,0	262,3	262,3
Rotation pro Bild (°)	0,5	0,5	0,25
Gesamtrotation (°)	180	127,5	37,5
Belichtungszeit (s)	0,5	0,8	0,25
Raumgruppe	$P2_1$	$C2$	$P321$
$a, b, c$ (Å)	150,9, 356,8, 230,8	180,7, 65,3, 99,4	203,9, 203,9, 215,9
$\alpha, \beta, \gamma$ (°)	90,0, 97,3, 90,0	90,0, 92,5, 90,0	90,0, 90,0, 120,0
Mosaisizität	0,154	0,372	0,115
Auflösungsbereich (Å)	59,5–3,1 (3,29–3,1)	49,6–1,9 (2,01–1,90)	50,0–3,99 (4,23–3,99)
Gesamtzahl Reflektionen	1.090.980	237.308	444.855
Anzahl Reflektionen (einzigartig)	281.563	88.058	84.895
Vollständigkeit (%)	98,9 (96,7) <sup>a</sup> , 64,6 (5,8) <sup>b</sup>	96,2 (96,2)	99,4 (96,6)
Redundanz	3,9 (3,9)	2,69 (2,71)	5,24 (4,99)
$\langle I/\sigma(I) \rangle$	5,8 (2,3) <sup>a</sup> , 4,6 (1,4) <sup>b</sup>	14,3 (1,84)	12,54 (1,72)
$R_{\text{sym}}$	13,9 (50,5) <sup>a</sup> , 20,0 (92,8) <sup>b</sup>	4,4 (58,2)	9,2 (109,1)
PDB Zugriffscode	4V98	4F7U	Nicht übermittelt

a Für den Auflösungsbereich von 60,0–3,1 Å, die Angabe in Klammern bezieht sich auf die höchste Auflösungsschale von 4,23–4,0 Å.

b Für den Auflösungsbereich von 60,0–3,1 Å, ohne ellipsoide Trunkation, die Angabe in Klammern bezieht sich auf die höchste Auflösungsschale von 3,29–3,10 Å.

**Tabelle 3 Strukturlösung und Verfeinerung**

Die Angaben in Klammern beziehen sich auf die höchste Auflösungsschale

Datensatz	8S	6S-C2	6S-P321
Auflösungsbereich (Å)	59,47–3,10 (3,18–3,10)	49,63–1,90 (1,92–1,90)	/
Vollständigkeit (%)	64,6	96,1	/
Reflektionen Working-Set	281,563 (733)	88,054 (2518)	/
Reflektionen Test-Set	14,120 (44)	4,426 (117)	/
Finaler $R_{\text{cryst}}$	0,232 (0,218)	0,181 (0,3128)	/
Finaler $R_{\text{free}}$	0,256 (0,206)	0,221 (0,3502)	/
Protein (ohne H-Atome)	121.940	8.140	/
Ligand (ohne H-Atome)	50	31	/
Wasser (ohne H-Atome)	0	372	/
Total (ohne H-Atome)	121.990	8.543	/
R.m.s. Abweichungen			
Bindungen (Å)	0,010	0,012	/
Winkel (°)	1,29	1,52	/
Durchschnittliche $B$ Faktoren (Å <sup>2</sup> )			
Protein	97,3	44,7	/
Ion	/	/	
Ligand	122,9	48,8	/
Ramachandran Plot			
Bevorzugte Konformere (%)	92,4	98,3	/
Erlaubte Konformere (%)	5,2	1,8	/
PDB Zugriffscode	4V98	4F7U	nicht verfeinert

## 6. Diskussion

### 6.1 Zelluläre Zusammenlagerungsprozesse benötigen häufig spezifische Assistenzfaktoren

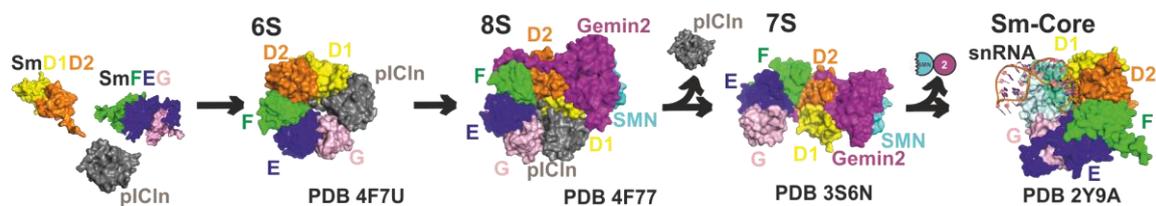
Makromolekulare Komplexe vermitteln vielfältige Funktionen in allen Lebewesen. Der Zusammenlagerung dieser Gebilde *in vivo* kommt daher eine entscheidende Bedeutung zu. So gelten die koordinierte Synthese der Untereinheiten, deren Faltung in eine funktionale Konformation und schließlich das korrekte Zusammenfügen der Untereinheiten als essentielle Schritte bei der Bildung aktiver Komplexe. Während einige Aspekte dieser „Assembly-Line“, wie die koordinierte Bildung der Untereinheiten und deren Faltung bereits detailliert untersucht werden konnten (Sprinzak et al., 2009), ist die Regulation des Zusammenlagerungs-Prozesses selbst noch wenig verstanden.

Klassische Arbeiten am monomeren Enzym Ribonuklease A zeigten, dass Polypeptidketten spontan ihre native Konformation einnehmen können. Diese Daten führten zur Formulierung des Anfinsen-Dogmas, nach dem die dreidimensionale Struktur eines Proteins grundsätzlich in seiner Primärstruktur angelegt und die Ausbildung der nativen Konformation ein thermodynamisch getriebener Prozess ist (Anfinsen, 1973). Analoge Beobachtungen konnten *in vitro* auch für die Bildung makromolekularer Komplexe gemacht werden: Partikel wie Ribosomen oder spleißosomale Untereinheiten bilden sich *in vitro* spontan, wenn deren Bestandteile unter definierten Bedingungen inkubiert werden (Dohme and Nierhaus, 1976; Fischer et al., 1993; Nomura et al., 1968; Segault et al., 1995). Es ließen sich sogar aus orthologen Komponenten chimäre Komplexe zusammenlagern (Nomura et al., 1968), was eindrucksvoll zeigte, dass die strukturelle Information für die korrekte Zusammenlagerung multimerer Komplexe in der dreidimensionalen Struktur angelegt und evolutionär konserviert ist.

Umso überraschender war daher die Beobachtung, dass ein als Nucleoplasmin bezeichneter Faktor notwendig für die erfolgreiche Zusammenlagerung von Nucleosomen ist. In Abwesenheit dieses Proteins wird die Bildung eines unlöslichen Präzipitats aus Histonen und DNA beobachtet, während in Gegenwart von Nucleoplasmin eine korrekte Formierung von Nucleosomen stattfindet (Laskey et al., 1978). Das azide Nucleoplasmin bindet an die Histone und schirmt so die starke positive Ladung der Histone ab. Dadurch werden unspezifische und verfrühte Wechselwirkungen der Histone mit Nucleinsäuren verhindert und die korrekte Formierung der Nucleosomen ermöglicht (Laskey et al., 1993). Als „Hilfsprotein“ oder „Assembly-Chaperon“ ist Nucleoplasmin kein integraler Bestandteil der reifen Nucleosomen, sondern interagiert nur transient mit den Histonen vor deren Vereinigung mit der DNA.

Ein ähnlicher Faktor ist das in dieser Arbeit strukturell untersuchte pICln-Protein, welches entscheidend an der Biogenese der U-snRNPs beteiligt ist. Biochemische Untersuchungen hatten zuvor gezeigt, dass pICln die Zusammenlagerung der Sm-Core-Domäne spleißosomaler U-snRNPs, entscheidend beeinflusst. Diese markante Struktur besteht aus den sieben Sm-Proteine, B/B'-D1-D2-D3-E-F-G, die einen heptameren Proteinring um die Sm-Bindungsstelle der snRNAs bilden.

Die Sm-Proteine lagern sich post-translational in drei definierte und RNA-freie Heterooligomere zusammen: SmD1-D2, SmE-F-G und SmB-B'-D3. Die Einschleusung dieser Heterooligomere in den „Assembly-Pathway“ und deren räumlich Anordnung zueinander erfolgt durch pICln, das eine Untereinheit des PRMT5-Komplexes ist. An diesem Komplex findet auch die symmetrische Dimethylierung von Argininresten in den Arginin- und Glycin-reichen C-Termini der Sm-Proteine B/B', D1 und D3 wobei noch unklar ist, welche Bedeutung diese post-translazionale Modifikation hat. Die Sm-Proteine verlassen schließlich den PRMT5-Komplex um in die späte (SMN-Komplex getriebene) Phase des Zusammenlagerungswegs einzutreten. Das „Hauptprodukt“ der frühen Phase ist das 6S-Intermediat, bestehend aus pICln, SmD1, D2, E, F und G. (Chari et al., 2008; Fisher et al., 1985; Friesen et al., 2001b; Sauterer et al., 1990). Die im 6S-Komplex gebundenen Sm Proteine sind aufgrund einer kinetischen Falle nicht in der Lage, mit der snRNA zu assoziieren. Für deren Aktivierung ist der SMN-Komplex zwingend notwendig. Wie diese Aktivierung mechanistisch erfolgt war zu Beginn dieser Arbeit unklar. Die von mir erhaltenen strukturellen Daten in Kombination mit biochemischen Daten anderer haben aber zu einem plausiblen Modell dieses kritischen Assembly-Schrittes geführt (Abbildung 24).



**Abbildung 24** Hierarchische Einordnung der diskutierten Sm-Protein-Transferintermediate im Verlauf der Sm-Core-Zusammenlagerung.

## 6.2 pICln agiert als Chaperon bei der Sm-Core-Zusammenlagerung und zeigt eine molekulare Mimikry der Sm-Proteinfaltung

In dieser Arbeit gelang es mit dem 6S-Komplex das zentrale Intermediat zwischen der frühen und späten Phase der Sm-Core-Zusammenlagerung auf atomarer Ebene strukturell zu charakterisieren und damit einen Einblick in den Schlüsselschritt der U-snRNP Biogenese zu erhalten. Hierbei war von besonderem Interesse, wie die Bindung der Sm-Proteine durch pICln erfolgt. Diese ermöglicht die

Erzeugung eines stabilen hexameren Intermediats höherer Ordnung aus den beiden Heterooligomeren SmD1-D2 und SmE-F-G, die alleine keine Affinität zueinander besitzen.

Mit der Sm-Faltung (Kambach et al., 1999) besitzen die Sm-Proteine, wie auch pICln mit der erweiterten PH-Domäne, eine  $\beta$ -fassartige Topologie. Die Bindung beruht in erster Linie jeweils auf einer antiparallelen Paarung der jeweiligen  $\beta$ 4- und  $\beta$ 5-Stränge benachbarter Sm-Proteine, die durch Kontakte zwischen weiteren Sekundärstrukturelemente beider Interaktionspartner ergänzt wird. Hierbei ist in erster Linie die Positionierung der N-terminalen  $\alpha$ 1-Helices der  $\beta$ 5-Interaktoren auf der flachen Ringoberfläche über dem  $\beta$ -Faltblatt der  $\beta$ 4-Interaktoren zu nennen, die weitere Kontakte zwischen den benachbarten Proteinen ermöglicht. In Analogie hierzu beruht auch die Interaktion von pICln mit den benachbarten Sm-Proteinen SmD1 und SmG im Kern auf der Paarung von zwei  $\beta$ -Strängen der erweiterten PH-Domäne mit deren  $\beta$ 4- beziehungsweise  $\beta$ 5-Strängen. Durch diese durchgängigen  $\beta$ -Paarungen ergibt sich ein über die Gesamtheit des Sm/pICln-Rings des 6S-Komplexes kontinuierliches  $\beta$ -Faltblatt.

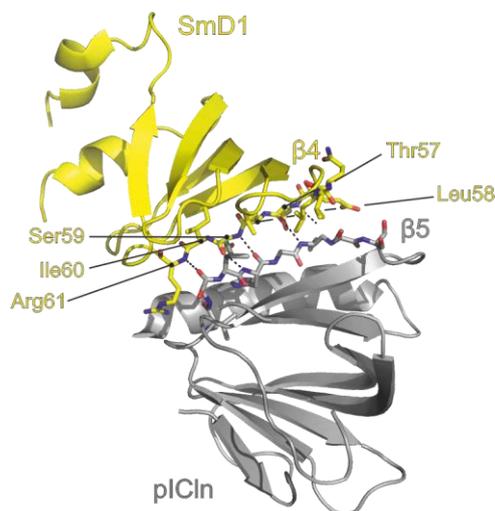
Die Interaktionen zwischen den Sm-Proteinen und pICln werden also durch eine strukturelle Verwandtschaft zwischen diesen Proteinen ermöglicht. Im Fall der Bindung zwischen pICln und SmD1 beruht diese sogar auf einer Nachahmung des typischen Sm-Sm-Interaktionsmodus. So kontaktieren sich beide Proteine über eine antiparallele Paarung der benachbarten  $\beta$ -Stränge  $\beta$ 5 von pICln und  $\beta$ 4 von SmD1. Auch die C-terminale  $\alpha$ -Helix von pICln ist, ähnlich wie die N-terminalen  $\alpha$ -Helices der Sm-Proteine, oberhalb des  $\beta$ -Faltblattes von SmD1 positioniert und ermöglicht weitere Kontakte zwischen beiden Proteinen. Im Gegensatz dazu weicht die Interaktion zwischen pICln und SmG von diesem Interaktionsmodus ab und war aus der Betrachtung der Strukturen der einzelnen Proteine nicht vorhersehbar. Die Paarung der benachbarten  $\beta$ -Stränge beider Proteine erfolgt in paralleler Orientierung und umfasste auf Seite von pICln einen zuvor unbekanntem  $\beta$ -Strang. Dieser wird als  $\beta$ 0 bezeichnet, befindet sich im N-Terminus des Proteins und bindet auf Seiten von SmG an den  $\beta$ 5-Strang. Der N-Terminus von pICln war in der NMR-Struktur von *C. I. familiaris* pICln ungeordnet und die Existenz des  $\beta$ 0-Strangs konnte erst in den Kristallstrukturen der 6S- und 8S-Komplexe nachgewiesen werden. In monomerem pICln ist die Interaktionsfläche für SmG daher vermutlich noch nicht ausgebildet. Es scheint naheliegend, dass durch die Bindung von SmD1-D2 an den N-Terminus von pICln die Ausbildung des  $\beta$ 0-Strangs induziert und somit die anschließende Bindung von SmE-F-G ermöglicht wird. Diese Vermutung steht in gutem Einklang mit biochemischen Daten, die eine schrittweise Abfolge der Zusammenlagerung des 6S-Komplexes nahelegen. Danach bindet SmD1-D2 zuerst an pICln, wodurch ein trimerer pICln-SmD1-D2 Komplex entsteht. Eine isolierte Bindung von SmE-F-G an pICln wird hingegen nicht beobachtet und erfolgt erst an das pICln-SmD1-D2 Intermediat (Chari et al., 2008). Die Bildung des 6S-Komplexes erfolgt also in einer hierarchischen Abfolge und beginnt mit der spezifischen Erkennung von SmD1 durch pICln. Daraufhin

erfolgt die Bindung von SmE-F-G, die im Anschluss an die Bindung von SmD1-D2, vermutlich durch eine induzierte Bildung des  $\beta_0$ -Stanges, ermöglicht wird.

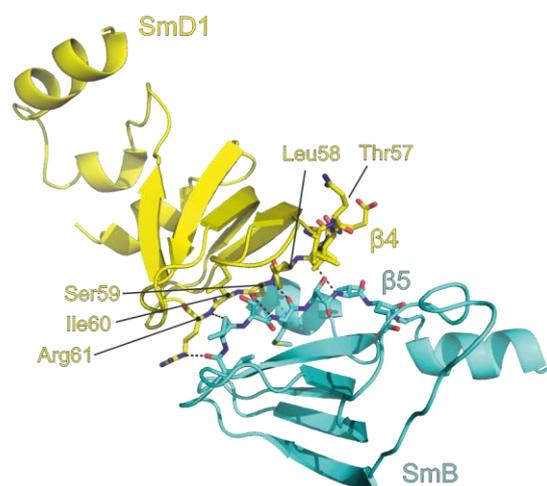
### 6.3 pICln agiert als topologischer Organisator des Sm-Core

Anhand der Kristallstruktur des 6S-Komplexes konnte nachgewiesen werden, dass sich die fünf Sm-Proteine innerhalb des Sm/pICln-Rings des 6S-Komplexes zueinander bereits in der Position befinden, die sie auch im Sm-Core der snRNPs haben. pICln nimmt innerhalb des 6S-Komplexes eine zu SmB/B'-D3 äquivalente Position ein und kontaktiert SmG und SmD1 über dieselben Kontaktflächen wie SmD3 und SmB/B' im Sm-Core (Abbildung 25).

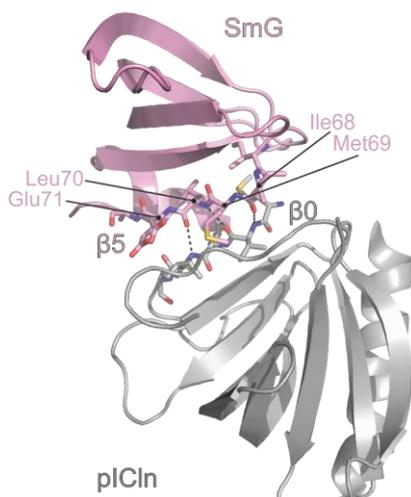
SmD1-pICln (6S-Komplex)



SmD1-SmB (U4-snRNP)



SmG1-pICln (6S-Komplex)



SmG1-SmD3 (U4-snRNP)

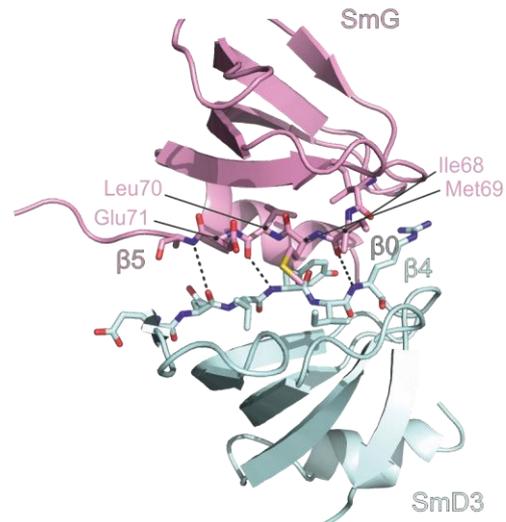


Abbildung 25 Vergleich der Kontakte von SmD1 und SmG mit ihren jeweiligen Interaktionspartnern innerhalb des 6S-Komplexes beziehungsweise innerhalb des Sm-Rings des U4-snRNP. SmD1 interagiert sowohl mit pICln, als auch mit SmB über seinen  $\beta_4$ -Strang. SmG interagiert sowohl mit pICln, als auch mit SmD3 über seinen  $\beta_5$ -Strang.

Bei der Rekrutierung der SmD1-D2 und SmE-F-G Heterooligomere wäre theoretisch auch eine Bindung von pICln an SmD2 oder SmF vorstellbar. Eine solche Bindung lässt sich jedoch aus sterischen Gründen ausschließen, da Reste dieser Proteine, die sich in äquivalenten Positionen zur SmD1-Kontaktfläche befinden, Kollisionen mit pICln verursachen würden (Abbildung 26, Seite 78). Interessanterweise ist auf der Seite von pICln die Kontaktfläche zu SmD1 hoch konserviert, während das Protein generell nur eine geringe Sequenzkonservierung zeigt (Abbildung 27, Seite 99 im Anhang). Dies spiegelt sich auch in einer Reihe an Sequenz-spezifischen Bindungen zwischen pICln und SmD1 wider, die eine ähnliche Interaktion mit anderen Sm-Proteinen ausschließt. Damit bedingt die Bildung des 6S-Komplexes bereits zu einem frühen Zeitpunkt die Abfolge der Sm-Proteine innerhalb der Ringstruktur. pICln agiert also als topologischer Organisator des Sm-Cores und verhindert dadurch eine alternative und fehlerhafte Positionierung der Untereinheiten zueinander. Dies ist für die Bindung der snRNA in der späten Phase der Zusammenlagerung wichtig. Die Interaktion zwischen den Sm-Proteinen und den Nukleotiden der Sm-Bindungsstelle erfolgt über äquivalente Positionen innerhalb der L3- und L5-Schleifen der Sm-Proteine, wobei jedes Sm-Protein jeweils eines der Nukleotide spezifisch kontaktiert. Dieser positionsabhängige Bindungsmodus erfordert sehr wahrscheinlich eine strikte Positionierung der Sm-Proteine zueinander (Leung et al., 2011), die durch die vorherige Organisation mit Hilfe von pICln gewährleistet ist.

Es ist auffällig, dass durch pICln genau die fünf Sm-Proteine, die *in vitro* die Bildung des Sm-Subcores ermöglichen, im 6S-Komplex getrennt von SmB/B'-D3 gebunden und organisiert werden. Dies könnte auf eine eventuelle zelluläre Relevanz des Sm-Subcore als Zwischenstufe einer hierarchischen Zusammenlagerung des Sm-Core hindeuten. *In vivo* konnte die Existenz des Sm-Subcore jedoch bisher nicht nachgewiesen werden.

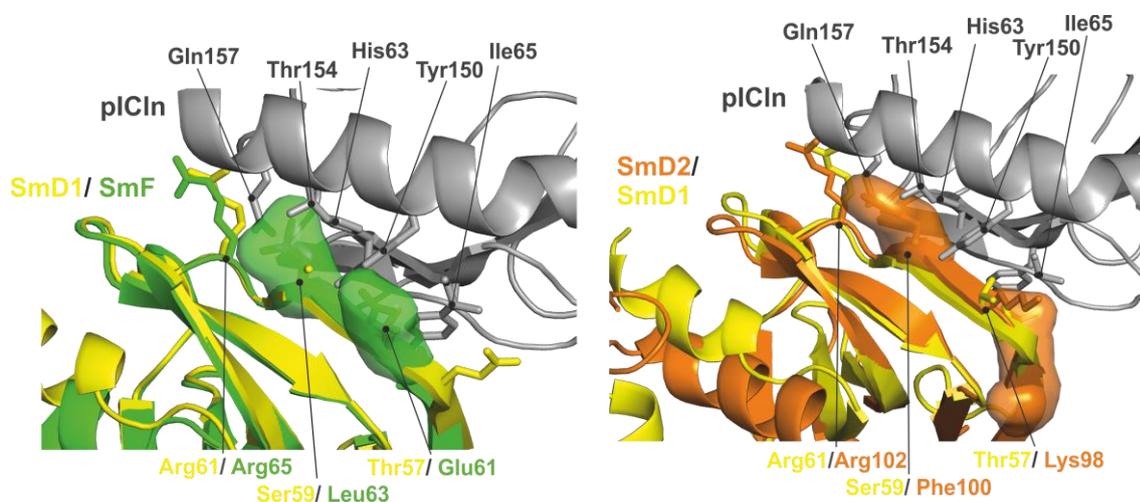


Abbildung 26 Strukturelle Überlagerung von SmF (links) und SmD2 (rechts) auf SmD1 im Kontaktbereich zwischen SmD1 und pICln. Reste von SmF und SmD2, die innerhalb des Kontaktbereichs zu Kollisionen mit pICln führen würden sind als Connolly-Oberflächen dargestellt.

## 6.4 Die elektrostatischen Eigenschaften von pICln ermöglichen eine stabile Rekrutierung der Sm-Proteine zum 6S-Komplex

pICln besitzt drei konservierte saure Regionen, die innerhalb ausgedehnter flexibler Bereiche des Proteins lokalisiert sind und die für die hohe negative Nettoladung des Proteins verantwortlich sind (Abbildung 27, Seite 99 im Anhang) (Schedlbauer et al., 2011). Im Zuge der Optimierungen des 6S-Komplexes für die Kristallisation war es jedoch notwendig, diese flexiblen Bereiche des Proteins zu deletieren. Die Herstellung entsprechender pICln-Fragmente ging daher auch mit einer signifikanten Reduktion der Nettoladung und einer dadurch bedingten Erhöhung des isoelektrischen Punktes von 4,0 auf 5,1 einher. Trotz dieser signifikanten Änderung der physikochemischen Eigenschaften war die Rekonstitution der 6S- und 8S-Komplexe mit diesen pICln-Varianten *in vitro* weiterhin möglich. Aufgrund der auffälligen Konservierung ist aber davon auszugehen, dass die elektrostatischen Eigenschaften von pICln dennoch bedeutend für seine zelluläre Funktionalität sind. Hierfür ist es wichtig, die physikalischen Grundlagen für das Zustandekommen von Bindungsereignissen zu betrachten.

Die Ausbildung spezifischer Kontakte zwischen Interaktionspartnern erfordert, dass sich diese sowohl in räumlicher Nähe, als auch in korrekter Orientierung zueinander befinden. Spezifische Wechselwirkungen wie Wasserstoffbrücken und Van-der-Waals-Kräfte wirken jedoch nur über eine kurze Distanz. Daher beruht das Zusammentreffen von Makromolekülen in erster Linie auf Diffusionsbewegungen der Moleküle, die sich in ungerichteten Translations- und Drehbewegungen äußern. Diese Phase stellt deshalb auch den geschwindigkeitsbestimmenden Schritt für die Bildung makromolekularer Komplexe dar (Schreiber et al., 2009). Da elektrostatische Anziehungs- und Abstoßungskräfte hingegen über lange Distanzen wirken, sind zwischen ungleich geladenen Untereinheiten weitreichende Anziehungskräfte vorhanden. Diese können die Ausbildung von Proteininteraktionen stark beschleunigen (Alsallaq and Zhou, 2008). Im Falle von pICln sind die Ladungen zusätzlich so über das Protein verteilt, dass sich über das Protein ein Dipolmoment ergibt. Dessen Richtung ermöglicht in Kombination mit der starken negativen Nettoladung einen elektrostatischen Lenkungseffekt (Liu et al., 2015; Zhang and Buck, 2013) und somit frühzeitig eine korrekte Ausrichtung von pICln zu den Sm-Proteinen. Besonders unter den erschwerten Bedingungen des zellulären Gedränges kommt diesen Eigenschaften von pICln möglicherweise eine entscheidende Rolle bei der Rekrutierung der Sm-Proteine zu.

Weiterhin ermöglicht die ausgeprägte negative Nettoladung von pICln wohl überhaupt erst die Interaktion zwischen SmD1-D2 und SmE-F-G in Abwesenheit von RNA. Obwohl SmD1-D2 und SmE-F-G alleine keine Affinität zueinander besitzen, umfasst der Kontakt zwischen SmD2 und SmF im

6S-Komplex spezifische Interaktionen und umfasst eine Fläche von ungefähr  $800 \text{ \AA}^2$ . Damit umfasst dieser Kontakt eine signifikant größere Kontaktfläche als der Kontakt zwischen SmG und pICln mit  $600 \text{ \AA}^2$ . Es erscheint daher auf den ersten Blick verwunderlich, dass eine Interaktion zwischen den beiden Heterooligomeren SmD1-D2 und SmE-F-G nur in Gegenwart von pICln erfolgt. Da beide Heterooligomere jedoch eine ausgeprägte positive Nettoladung tragen, kommt es in Abwesenheit von pICln höchstwahrscheinlich zur elektrostatischen Abstoßung, die eine spezifische Interaktion beider Heterooligomere verhindert. Durch seine komplementäre negative Ladung führt die Bindung von pICln zu einer Neutralisierung der Nettoladung im 6S-Komplex. Außerdem befinden sich die aziden Regionen von pICln größtenteils innerhalb der zentralen Pore des 6S-Komplexes und können durch diese Positionierung die potentiellen elektrostatischen Abstoßungseffekte zwischen den Sm-Proteinen effektiv kompensieren.

## **6.5 Strukturelle Ursachen für die Inhibition der Interaktion zwischen snRNA und Sm-Proteinen im 6S-Komplex**

Obwohl die Komposition des 6S-Komplexes mit dem Sm-Pentamer SmD1-D2-E-F-G hinreichend für die Ausbildung des Sm-Subcore wäre (Raker et al., 1996), zeigt er keine Affinität zur snRNA. pICln agiert zwar als Assembly-Chaperon, blockiert jedoch in der frühen Phase durch seine Interaktionen mit den Sm-Proteinen die Bindung an die snRNA und damit die Ausbildung des Sm-Cores (Chari et al., 2008; Pu et al., 1999; Sauterer et al., 1990). Die detaillierte Auswertung der 6S-Kristallstruktur liefert einige Anhaltspunkte, mit denen sich diese Inhibition zwanglos erklären lässt. Die Interaktion mit der Sm-Bindungsstelle der snRNA erfolgt innerhalb der zentralen Pore des Sm-Cores durch die L3- und L5-Schleifenregionen der einzelnen Sm-Proteine (Leung et al., 2011). Dies wird innerhalb des 6S-Komplexes, wie auch im 8S-Komplex, durch eine Kombination sterischer und elektrostatischer Begebenheiten verhindert.

Zum einen besetzt pICln im 6S-Komplex im Vergleich zum Sm-Core topologisch die Position von zwei Sm-Proteinen, nimmt räumlich aber nur das Volumen von ungefähr 1,5 Sm Untereinheiten ein. Dadurch verringert sich die Öffnungsweite zwischen SmD1 und SmG, was sich auch in einem verringerten Durchmesser der zentralen Pore äußert (Abbildung 7c, Seite 44). Zusätzlich wird diese durch ausgedehnte flexiblere Bereiche von pICln regelrecht verschlossen. So befinden sich innerhalb der zentralen Pore Teile des pICln-C-Terminus, welche diese von der flachen Ringfläche kommend in Richtung der kegelförmigen Ringfläche durchlaufen. Auf dieser Ringfläche befinden sich ebenfalls die in den Kristallstrukturen identifizierbaren Teile der L5-Schleifenregion von pICln. Somit sind die drei aziden Regionen (AR1-AR3) von pICln, die sich innerhalb der L5-Schleife und im C-Terminus befinden, auf der kegelförmigen Ringoberfläche im Bereich der zentralen Pore lokalisiert. Es kommt also neben einer sterischen Blockade auch zu einer Konzentration negativer Ladungen im Bereich der zentralen

Pore. Dadurch werden elektrostatische Wechselwirkungen mit dem Phosphat-Rückgrat der snRNA effektiv verhindert.

## 6.6 Die Sm-Proteine werden im 8S-Komplex durch Gemin2 kontaktiert

Die Kristallstruktur des 6S-Komplexes hat zeigen können, wie die Sm-Proteine zu einem Intermediat höherer Ordnung zusammengefügt werden. Wie jedoch die Auflösung der hierbei entstehenden kinetischen Falle erfolgt, sollte anhand der Kristallstruktur des 8S-Komplexes entschlüsselt werden. Im 8S-Komplex ist der Sm/pICln-Ring des 6S-Komplexes durch die SMN-Komplex Untereinheiten SMN und Gemin2 gebunden, sodass dieser ein „gefangenes“ Sm-Transferintermediat vom PRMT5-Komplex auf die snRNA darstellt. Die Stabilisierung des 8S-Komplexes erfolgte durch eine C-terminal gekürzte Variante des SMN-Proteins, die nur die 122 N-terminalen Reste der insgesamt 226 Reste des Proteins umfasste (Chari et al., 2008). Die Aufklärung der Kristallstruktur des 8S-Komplexes hatte zum Ziel, strukturelle Einblicke in die Rolle des SMN-Komplexes zu erhalten und dessen Mechanismus bei der Dissoziation von pICln aufzuklären. In der Kristallstruktur konnte überraschenderweise mit den Aminosäuren Ala8-Gly24 nur ein Bruchteil der Positionen des gekürzten SMN-Proteins identifiziert werden. Diese Reste bilden eine  $\alpha$ -Helix aus, über die der Kontakt zu Gemin2 hergestellt wird (Grimm et al., 2013; Liu et al., 1997; Zhang et al., 2011), während keine direkten Kontakte von SMN zu den Sm-Proteinen nachgewiesen werden konnten. Diese erfolgen interessanterweise ausschließlich über Gemin2, für das in mehreren vorhergehenden Studien eine direkte Interaktion mit den Sm-Proteinen ausgeschlossen wurde (Cauchi, 2010; Friesen et al., 2001a; Liu et al., 1997). C-terminal von der Gemin2-bindenden  $\alpha$ -Helix von SMN beinhaltet die verwendete SMN-Variante noch weitere 98 Aminosäuren. Diese sind jedoch in der 8S-Kristallstruktur nicht strukturell definiert, obwohl in diesem Sequenzabschnitt beispielsweise mit der SMN-Tudor-Domäne eine abgeschlossene Faltungseinheit lokalisiert ist (Selenko et al., 2001). Mutationen in dieser Domäne führen zu einer reduzierten Interaktion von SMN mit Sm-Proteinen und konnten in Patienten mit spinaler Muskelatrophie identifiziert werden (Buhler et al., 1999). Eine direkte Interaktion der Tudor-Domäne mit den Sm-Proteinen erfolgt über deren C-terminale RG-reichen Sequenzen, wobei präferentiell dimethylierte Arginine in SmD1, SmB/B' und SmD3 gebunden werden. Diese Interaktionen wurden in verschiedenen Studien sowohl biochemisch wie auch strukturell nachgewiesen (Brahms et al., 2001; Friesen et al., 2001a; Tripsianes et al., 2011). In den rekombinant erzeugten 8S-Präparationen lag der RG-reiche Sequenzbereich von SmD1 nicht methyliert vor. Dies könnte aufgrund der Präferenz für dimethylierte Arginine die Abwesenheit einer stabilen Bindung zwischen der SMN-Tudor-Domäne und SmD1 erklären. Dementsprechend konnten vorhergehende *in vitro* Bindungsstudien zeigen, dass die Affinität von SMN zu den Sm-Proteinen deutlich herabgesetzt ist, wenn diese unmethyliert vorliegen (Friesen et al., 2001a). Andererseits aber konnten neuere Studien auch zeigen, dass die

Methylierung der Sm-Proteine *in vitro* nicht notwendig für den erfolgreichen Transfer der Sm-Proteine auf die snRNA ist und zumindest *in vitro* keine Rolle bei der Formierung der Sm-Core spielt (Chari et al., 2008; Kroiss et al., 2008).

Gemin2 scheint somit zumindest beim Übergang von der frühen zur späten Phase der Sm-Core-Zusammenlagerung der entscheidende Faktor für die Übernahme der Sm-Proteine zu sein. Hierbei kontaktiert Gemin2 im 6S-Komplex mit SmD1, SmD2, SmE und SmF vier unterschiedliche Sm-Untereinheiten. Damit erfolgt durch Gemin2 gleichzeitig eine Bindung von Untereinheiten des ursprünglichen SmD1-D2 Dimers, als auch eine Bindung von Untereinheiten des SmE-F-G Oligomers. Da vorhergehende *in vitro* Bindungsstudien zeigen konnten, dass Gemin2 zumindest nicht an monomere Sm-Proteine bindet, ist es durchaus vorstellbar, dass die Erkennung der Sm-Proteine durch Gemin2 spezifisch für das Sm-Pentamer SmD1-D2-E-F-G ist. Diese Spezifität könnte als weiterer Kontrollpunkt für die korrekte topologische Anordnung und Stöchiometrie der Sm-Proteine vor der Zusammenlagerung des Sm-Core dienen.

## **6.7 Eine putative Sollbruchstelle im Sm/pICln-Ring des 8S-Komplexes könnte die Dissoziation von pICln einleiten.**

Sowohl innerhalb der Kristallstruktur des 6S-Komplexes, als auch in der Kristallstruktur des 8S-Komplexes, konnte der Kontakt zwischen pICln und SmG als derjenige mit der geringsten Kontaktfläche ausgemacht werden. Hierbei ist diese Kontaktfläche innerhalb der 20 Kopien des 8S-Komplexes im Vergleich zum 6S-Komplex nochmals um durchschnittlich 10% verringert. Außerdem ist auffällig, dass pICln im Kontaktbereich zu SmG keine Sequenzkonservierung aufweist und die Interaktionen größtenteils Sequenz-unspezifisch sind. Des Weiteren kumulieren die, zwischen den 20 Kopien des 8S-Komplexes innerhalb der asymmetrischen Einheit beobachteten RMS-Koordinatendifferenzen, im Bereich dieser Kontaktfläche und weisen auf eine starke strukturelle Flexibilität innerhalb dieser Region hin. Ein Vergleich des Sm/pICln-Rings mit der hufeisenförmigen Struktur des späten 7S-Intermediates legt zusätzlich die Existenz einer Spannung innerhalb des hexameren Sm/pICln-Rings nahe, welche mit der Dissoziation von pICln relaxiert zu werden scheint. Dies konnte an einer Überlagerung der Strukturen des hexameren Sm/pICln-Rings des 6S- und 8S-Komplexes mit dem geöffneten Sm-Ring des späten 7S-Intermediates verdeutlicht werden (Abbildung 18, Seite 60). In letzterem ist SmG im Gegensatz zur Situation im Sm/pICln-Ring aus der Ringebene verschoben und die Distanz zwischen SmD1 und SmG ist deutlich vergrößert. Die mit dieser Relaxation verbundene Vergrößerung der Ringöffnung ermöglicht die anschließende Bindung von SmB/B'-D3 und die hierfür notwendige Erweiterung der Ringstruktur zu dem heptameren Sm-Core.

## 6.8 Die Funktion des SMN-Komplexes bei der Dissoziation von pICln legt eine Rolle der SMN-Tudor-Domäne bei diesem Vorgang nahe

Für die Fortsetzung der Zusammenlagerung der Sm-Core, ausgehend vom 6S-Komplex, muss pICln von den Sm-Proteinen dissoziieren. Dies findet im Zuge der Bindung des 6S-Komplexes an den SMN-Komplex statt, woraufhin die Sm-Proteine für den Transfer auf die snRNA aktiviert werden. Vorhergehende Arbeiten unserer Gruppe legten nahe, dass sich der SMN-Komplex hierbei wie ein Enzym verhält. Nach dieser Hypothese würde der 6S-Komplex ein lokales energetisches Minimum auf dem Weg zur reifen Sm-Core-Struktur repräsentieren, die innerhalb dieses Modells das globale Minimum darstellt. Der Übergang vom 6S-Komplex zur reifen Sm-Core erfordere hierbei die Überwindung einer hohen Aktivierungsenergie, weshalb dieser kinetisch gehemmt wäre. Durch die Stabilisierung eines Sm-Transferintermediates könnte der SMN-Komplex dafür sorgen, dass diese Aktivierungsenergie herabgesetzt wird und die Formierung der Sm-Core stattfinden könnte. Diese Hypothese konnte durch experimentelle Daten unterstützt werden. Bei einem molaren Überschuss an SMN-Komplex kann mit dem Aufschmelzen der extrem stabilen Sm-Core die Rückreaktion beobachtet werden. Außerdem konnte gezeigt werden, dass der SMN-Komplex wiederholt und in substöchiometrischen Mengen aktiv ist und sich damit tatsächlich wie ein Enzym verhält (Chari et al., 2008).

Die Bildung des 6S-Komplexes beruht auf einer molekularen Mimikry der Faltung und des Interaktionsmodus der Sm-Proteine durch pICln. Hierbei erfolgt jedoch nur für SmD1 eine spezifische Bindung, während der Kontakt zu SmG als potentielle Bruchstelle für die spätere Dissoziation von pICln dienen könnte. Im Bereich dieses Kontakts lässt sich die höchste strukturelle Flexibilität des 6S-Komplexes feststellen, die innerhalb des 8S-Komplex nochmal deutlich verstärkt ist. Dies dürfte vor allem durch die Bindung des SMN-Gemin2 Dimers bedingt sein, das den Sm/pICln-Rings mit der C-terminalen Helixbündeldomäne und dem N-terminalen Arm von Gemin2 an zwei gegenüberliegenden Seiten kontaktiert (Abbildung 16, Seite 55). Weiterhin deutet die Verschiebung von SmG aus der Ringebene hinaus auf eine orthogonal zur Ringebene bestehende Spannung innerhalb der Ringstruktur. Diese Flexibilität könnte zu einem Aufbrechen des Kontaktes zwischen pICln und SmG führen. Diese möglicherweise nur transiente Ringöffnung könnte durch den SMN-Komplex stabilisiert werden, was seine Funktion als Katalysator der Sm-Core-Zusammenlagerung erklären könnte. Eine wichtige Rolle könnte hierbei möglicherweise die SMN-Tudor-Domäne spielen, die durch ihre fassartige Struktur, ähnlich wie auch pICln eine Mimikry der Sm-Proteinfaltung zeigt (Grimm et al., 2013; Ma et al., 2005; Selenko et al., 2001).

Differenzen zwischen der Kristall- und der EM-Struktur des 8S-Komplexes legen eine Lokalisierung der SMN-Tudor-Domäne oberhalb der kegelförmigen Ringfläche nahe (Grimm et al., 2013). Damit

wäre sie ideal positioniert um eine transiente Öffnung des pICln-SmG-Kontaktes durch eine Bindung an SmG oder pICln zu stabilisieren. Bemerkenswerterweise ist die SMN-Tudor-Domäne nicht das einzige Strukturelement des SMN-Komplexes, das eine strukturelle Homologie zu den Sm-Proteinen zeigt. Interessanterweise falten Gemin6 und Gemin7 ebenfalls in eine zur Sm-Faltung homologe Domäne und zeigen auch untereinander den bekannten Sm-Sm-Interaktionsmodus. Außerdem zeigen Gemin6 und Gemin7 *in vitro* eine hohe Affinität zu SmB/B'-D3, während unter den Sm-Proteinen des 6S-Komplexes nur zu SmE eine signifikante Affinität zu Gemin6 und Gemin7 gezeigt werden konnte (Ma et al., 2005). Es ist also wahrscheinlich, dass die strukturelle Mimikry der Sm-Faltung zu unterschiedlichen Zeitpunkten der Sm-Core-Zusammenlagerung eine wichtige Rolle spielt und sowohl die koordinierte Rekrutierung, Bindung und den Transfer der Sm-Proteine unterstützt.

## **6.9 Warum benötigt die Zelle ein solch komplexes System von Assistenzfaktoren?**

Besonders die Zusammenlagerung von RNPs stellt die eukaryotische Zelle vor die Herausforderung, dass die Synthese von Nukleinsäuren im Nukleus erfolgt, während die Synthese von Proteinen im Cytosol stattfindet. Dies erfordert, dass entweder die Nukleinsäuren oder die Proteine in das jeweils andere Kompartiment gelangen müssen, um eine Zusammenlagerung des funktionalen RNPs zu ermöglichen. Interessanterweise scheint die Konstruktion der Partikel oftmals gerade nicht in dem Kompartiment stattzufinden, indem die reifen und aktiven Partikel ihrer Funktion nachgehen. So erfolgt die Reifung der ribosomalen Untereinheiten in den Nukleoli im Nukleus (Cmarko et al., 2008), während zentralen Schritte der U-snRNP Biogenese im Cytosol stattfinden (Carissimi et al., 2005; Fischer et al., 1997; Liu et al., 1997; Meister and Fischer, 2002; Pellizzoni et al., 2002). Dies hat den Vorteil, dass Intermediate nicht am Ort der katalytischen Aktivität lokalisiert sind. Trotzdem sind die einzelnen Untereinheiten bei ihrer Synthese und während des anschließenden Transports zumindest zeitweise von einem Teil ihrer Interaktionspartner isoliert. Daher können sich Bereiche, die normalerweise in Kontaktflächen verborgen sind, in dieser Phase an zugänglichen Oberflächenbereichen befinden und potentiell in unspezifische Interaktionen verwickelt werden. Im extremsten Fall kann dies zur Aggregation und Präzipitation der betroffenen Makromoleküle führen. Die Existenz von molekularen Chaperonen, die entweder die Faltung der Untereinheiten unterstützen, an partiell entfaltete Monomere binden oder die Organisation von Untereinheiten zu Intermediaten höherer Ordnung unterstützen, ist daher ein wichtiger Faktor für die erfolgreiche zelluläre Konstruktion makromolekularer Maschinen.

Durch die kombinierte biochemische und strukturelle Charakterisierung der spleißosomalen U-snRNPs und ihrer komplexen Biogenese lassen sich die Notwendigkeit und die zugrundeliegenden

Mechanismen der an diesem Prozess beteiligten Assistenzfaktoren besser verstehen. Besonders aufgrund ihrer außergewöhnlich stark positiven Nettoladung stellen die Sm-Proteine eine potentielle Gefahr für die Zelle dar. So ermöglicht ihre Ladung nicht nur eine starke Interaktion mit der kognaten snRNA, sondern birgt auch die Gefahr unspezifischer Interaktionen mit nicht-kognaten cytosolischen RNA-Molekülen. In ähnlicher Weise stellen auch die Kontaktflächen der Sm-Proteine für ihre Interaktionen untereinander eine potentielle Gefahr dar. Diese beinhalten zum Teil ausgeprägte hydrophobe Bereiche. Besonders in den monomeren Sm-Proteinen, aber auch in den frühen Heterooligomeren, befinden sich diese Kontaktflächen, aufgrund der Abwesenheit der Interaktionspartner, teilweise an den Proteinoberflächen. Durch die Bildung des geschlossenen hexameren Sm/pICln-Rings im 6S-Intermediat wird die Gefahr solcher unspezifischer Interaktionen effektiv minimiert. pICln bindet an die freien Interaktionsflächen der Sm-Proteine und blockiert mögliche RNA-Interaktionen. Außerdem bedingt die hierarchische und frühzeitige Aktivität der im PRMT5- und SMN-Komplex organisierten Assistenzfaktoren eine definierte Stöchiometrie und korrekte topologische Abfolge der Sm-Proteine im späteren Sm-Core.

Die Problemstellungen, denen der komplexe Apparat an Assistenzfaktoren während der U-snRNP Biogenese begegnet, sind nicht allein spezifisch für diesen Prozess. Dementsprechend konnten in den letzten Jahren weitere Faktoren, die bei diversen Zusammenlagerungs-Vorgängen assistieren, identifiziert werden (Hirano et al., 2005; Hondele and Ladurner, 2011; Karbstein, 2010; Saschenbrecker et al., 2007). Diese scheinen also nicht wie anfänglich angenommen, exotische Ausnahmefälle zu sein. Die Verwendung von Assembly-Chaperonen scheint vielmehr eine generelle Strategie für einen kontrollierten und geregelten Zusammenschluss von makromolekularen Maschinen in der überfüllten zellulären Umgebung zu sein. Die biochemische und strukturelle Charakterisierung eines solchen Faktors, welche in dieser Arbeit durchgeführt wurde, half dabei, wichtige mechanistische Einblicke in die Zusammenlagerung von makromolekularen Komplexen zu erlangen.

## 7. Ausblick

Die Kristallstrukturen des 6S- und 8S-Komplexes ermöglichen in Kombination mit der Kristallstruktur des späten 7S-Intermediates detaillierte Einblicke in wichtige mechanistische Abläufe während der Zusammenlagerung der Sm-Core-Domäne. Hierbei war es möglich die Interaktion von pICln mit den Sm-Proteinen aufzuklären und eine Hypothese zum Mechanismus der Dissoziation von pICln zu formulieren. So scheint die Kontaktfläche zwischen SmG und pICln eine prädestinierte Bruchstelle im Sm/pICln-Ring darzustellen. Deren vorübergehende Öffnung könnte durch den SMN-Komplex ausgelöst und möglicherweise ebenfalls auf Basis einer strukturellen Mimikry, stabilisiert werden. Da pICln jedoch im 8S-Komplex weiterhin stabil gebunden bleibt, existiert derzeit kein eindeutiger experimenteller Nachweis für diese Annahme. Durch die C-terminale Kürzung von SMN und der Abwesenheit von Gemin8 gelang es die Bindung der Sm-Proteine durch den SMN-Komplex von der Dissoziation von pICln zu entkoppeln. Die im 8S-Komplex vorhandenen Elemente des SMN-Komplexes sind alleine also nicht in der Lage, die Dissoziation von pICln zu bewerkstelligen.

Weitere strukturelle Analysen eines vollständig aktiven SMN-Komplexes erscheinen daher sinnvoll, um seine Funktion noch besser zu verstehen. Besonders für kristallographische Studien scheinen hierbei entweder der trimere SMN-Komplex aus SMN, Gemin2 und Gemin8 oder der pentamere SMN-Komplex, der zusätzlich noch Gemin6 und Gemin7 beinhaltet, am ehesten geeignet zu sein. Beide lassen sich rekombinant herstellen und sind, zumindest *in vitro*, für alle Funktionen des SMN-Komplexes während des Sm-Core-Assembly hinreichend. Aufgrund der dynamischen Eigenschaften, die sich bereits bei der Analyse der 8S-Kristallstruktur andeuteten und möglicherweise entscheidend für die Dissoziation von pICln sind, ist eine Kombination von unterschiedlichen strukturellen Analysemethoden sinnvoll. Hierbei ist besonders die NMR-Spektroskopie zu nennen, die eine Untersuchung dynamischer Eigenschaften der beteiligten Faktoren in Lösung erlaubt. Eine Kooperation mit der Arbeitsgruppe von Prof. Dr. Michael Sattler, mit dem Ziel, die vermutete induzierte Faltung des  $\beta$ 0-Stranges von pICln und die möglicherweise transienten Interaktionen der SMN-Tudor-Domäne nachzuweisen, wurde im Rahmen dieser Doktorarbeit initiiert.

Vorhergehende biochemische Daten ließen eine zentrale Rolle von SMN bei der Bindung der Sm-Proteine vermuten, während direkte Kontakte von Gemin2 zu diesen sogar ausgeschlossen wurden (Cauchi, 2010; Friesen et al., 2001a; Liu et al., 1997). Die alleinige Bindung des 6S-Komplexes durch Gemin2 war daher ein durchaus überraschender Befund. Als zentralem Faktor bei der Sm-Protein-Bindung sollte während zukünftiger Studien daher ein größeres Augenmerk auf Gemin2 gelegt werden. Interessanterweise konnte eine Interaktion von SMN mit den Sm-Proteinen bereits durch *in vitro*-Bindungsstudien nachgewiesen werden (Brahms et al., 2001; Friesen et al., 2001a; Liu

et al., 1997). Auch die Interaktion der SMN-Tudor-Domäne mit dimethylierten Argininresten, wie sie sich innerhalb der RG-reichen Sequenzen der Sm-Proteine befinden, konnte strukturell aufgeklärt werden (Selenko et al., 2001; Tripsianes et al., 2011). Diese Interaktion steht hierbei in direktem Bezug zur spinalen Muskelatrophie, da beispielsweise die Patientenmutation E134K innerhalb der Tudor-Domäne lokalisiert ist und diese Interaktion stört (Selenko et al., 2001). Trotzdem scheint eine direkte Interaktion zwischen SMN und den Sm-Proteinen in der frühen Phase des Transfers eine untergeordnete Rolle zu spielen. Dies schließt eine wichtige Bedeutung zu einem späteren Zeitpunkt, beispielsweise bei der Dissoziation von pICln, jedoch nicht aus.

Der SMN-Komplex von Vertebraten besteht aus neun verschiedenen Untereinheiten, die zum Teil auch noch oligomerisiert vorliegen (Otter et al., 2007). Diese Komplexität steht in einem starken Kontrast zu der Beobachtung, dass *in vitro* ein Trimer aus SMN, Gemin2 und Gemin8 hinreichend für den gesamten Zusammenlagerungsprozess ist (Chari et al., 2008; Kroiss et al., 2008). Es ist daher sehr wahrscheinlich, dass die weiteren Gemine eine wichtige Rolle für eine effektive Katalyse im zellulären Umfeld spielen und beispielsweise für eine erhöhte Spezifität verantwortlich sind. Es ist ebenfalls möglich, dass diese Funktionen ausüben die, wie die Rekrutierung der snRNP-Komponenten, *in vitro* eine untergeordnete Rolle spielen. Eine solche Funktion wird beispielsweise für Gemin5 bei der Bindung der snRNA vermutet (Yong et al., 2010). Diese Hypothese muss aber, wie auch die Funktionen der anderen Gemine, noch durch weitere Studien verifiziert werden.

## 8. Referenzen

- Achsel, T., Stark, H., and Luhrmann, R. (2001). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proceedings of the National Academy of Sciences of the United States of America* 98, 3685-3689.
- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography* 66, 213-221.
- Alsallaq, R., and Zhou, H.X. (2008). Electrostatic rate enhancement and transient complex of protein-protein association. *Proteins* 71, 320-335.
- Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223-230.
- Baillat, D., Hakimi, M.A., Naar, A.M., Shilatifard, A., Cooch, N., and Shiekhattar, R. (2005). Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* 123, 265-276.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* 98, 10037-10041.
- Battle, D.J., Kasim, M., Wang, J., and Dreyfuss, G. (2007). SMN-independent subunits of the SMN complex. Identification of a small nuclear ribonucleoprotein assembly intermediate. *The Journal of biological chemistry* 282, 27953-27959.
- ben-Avraham, D., and Tirion, M.M. (1998). Normal modes analyses of macromolecules. *Physica A* 249, 415-423.
- Bourenkov, G.P., and Popov, A.N. (2010). Optimization of data collection taking radiation damage into account. *Acta crystallographica Section D, Biological crystallography* 66, 409-419.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* 72, 248-254.
- Brahms, H., Meheus, L., de Brabandere, V., Fischer, U., and Luhrmann, R. (2001). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *Rna* 7, 1531-1542.
- Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Haendler, B., and Jacob, M. (1982). U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *The EMBO journal* 1, 1259-1265.
- Branscombe, T.L., Frankel, A., Lee, J.H., Cook, J.R., Yang, Z., Pestka, S., and Clarke, S. (2001). PRMT5 (Janus kinase-binding protein 1) catalyzes the formation of symmetric dimethylarginine residues in proteins. *The Journal of biological chemistry* 276, 32971-32976.
- Briese, M., Esmaili, B., Fraboulet, S., Burt, E.C., Christodoulou, S., Towers, P.R., Davies, K.E., and Sattelle, D.B. (2009). Deletion of *smn-1*, the *Caenorhabditis elegans* ortholog of the spinal muscular atrophy gene, results in locomotor dysfunction and reduced lifespan. *Human molecular genetics* 18, 97-104.

- Brow, D.A. (2002). Allosteric cascade of spliceosome activation. *Annual review of genetics* 36, 333-360.
- Bryksin, A.V., and Matsumura, I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *BioTechniques* 48, 463-465.
- Buhler, D., Raker, V., Luhrmann, R., and Fischer, U. (1999). Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. *Human molecular genetics* 8, 2351-2357.
- Carissimi, C., Baccon, J., Straccia, M., Chiarella, P., Maiolica, A., Sawyer, A., Rappsilber, J., and Pellizzoni, L. (2005). Unrip is a component of SMN complexes active in snRNP assembly. *FEBS letters* 579, 2348-2354.
- Carter, P. (1986). Site-directed mutagenesis. *The Biochemical journal* 237, 1-7.
- Cauchi, R.J. (2010). SMN and Gemins: 'we are family' ... or are we?: insights into the partnership between Gemins and the spinal muscular atrophy disease protein SMN. *BioEssays : news and reviews in molecular, cellular and developmental biology* 32, 1077-1089.
- Cauchi, R.J., Davies, K.E., and Liu, J.L. (2008). A motor function for the DEAD-box RNA helicase, Gemin3, in *Drosophila*. *PLoS genetics* 4, e1000265.
- Chari, A., Golas, M.M., Klingenhager, M., Neuenkirchen, N., Sander, B., Englbrecht, C., Sickmann, A., Stark, H., and Fischer, U. (2008). An assembly chaperone collaborates with the SMN complex to generate spliceosomal SnRNPs. *Cell* 135, 497-509.
- Chen, G.Q., Sun, Y., Jin, R., and Gouaux, E. (1998). Probing the ligand binding domain of the GluR2 receptor by proteolysis and deletion mutagenesis defines domain boundaries and yields a crystallizable construct. *Protein science : a publication of the Protein Society* 7, 2623-2630.
- Chen, J., Ezzeddine, N., Waltenspiel, B., Albrecht, T.R., Warren, W.D., Marzluff, W.F., and Wagner, E.J. (2012). An RNAi screen identifies additional members of the *Drosophila* Integrator complex and a requirement for cyclin C/Cdk8 in snRNA 3'-end formation. *Rna* 18, 2148-2156.
- Claverie, J.M., Monchois, V., Audic, S., Poirot, O., and Abergel, C. (2002). In Search of new anti-bacterial target genes: a comparative/structural genomics approach. *Combinatorial chemistry & high throughput screening* 5, 511-522.
- Cmarko, D., Smigova, J., Minichova, L., and Popov, A. (2008). Nucleolus: the ribosome factory. *Histology and histopathology* 23, 1291-1298.
- Dale, G.E., Kostrewa, D., Gsell, B., Stieger, M., and D'Arcy, A. (1999). Crystal engineering: deletion mutagenesis of the 24 kDa fragment of the DNA gyrase B subunit from *Staphylococcus aureus*. *Acta crystallographica Section D, Biological crystallography* 55, 1626-1629.
- Ding, H.T., Ren, H., Chen, Q., Fang, G., Li, L.F., Li, R., Wang, Z., Jia, X.Y., Liang, Y.H., Hu, M.H., *et al.* (2002). Parallel cloning, expression, purification and crystallization of human proteins for structural genomics. *Acta crystallographica Section D, Biological crystallography* 58, 2102-2108.
- Dohme, F., and Nierhaus, K.H. (1976). Total reconstitution and assembly of 50 S subunits from *Escherichia coli* Ribosomes in vitro. *Journal of molecular biology* 107, 585-599.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic acids research* *32*, W665-667.

Ellis, R.J. (2001a). Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Current opinion in structural biology* *11*, 114-119.

Ellis, R.J. (2001b). Macromolecular crowding: obvious but underappreciated. *Trends in biochemical sciences* *26*, 597-604.

Ellis, R.J. (2006). Molecular chaperones: assisting assembly in addition to folding. *Trends in biochemical sciences* *31*, 395-401.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography* *60*, 2126-2132.

Ezzeddine, N., Chen, J., Waltenspiel, B., Burch, B., Albrecht, T., Zhuo, M., Warren, W.D., Marzluff, W.F., and Wagner, E.J. (2011). A subset of *Drosophila* integrator proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3'-end formation. *Molecular and cellular biology* *31*, 328-341.

Fischer, U., Liu, Q., and Dreyfuss, G. (1997). The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* *90*, 1023-1029.

Fischer, U., and Luhrmann, R. (1990). An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus. *Science* *249*, 786-790.

Fischer, U., Sumpster, V., Sekine, M., Satoh, T., and Luhrmann, R. (1993). Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *The EMBO journal* *12*, 573-583.

Fisher, D.E., Conner, G.E., Reeves, W.H., Wisniewolski, R., and Blobel, G. (1985). Small nuclear ribonucleoprotein particle assembly in vivo: demonstration of a 6S RNA-free core precursor and posttranslational modification. *Cell* *42*, 751-758.

Flot, D., Mairs, T., Giraud, T., Guijarro, M., Lesourd, M., Rey, V., van Brussel, D., Morawe, C., Borel, C., Hignette, O., *et al.* (2010). The ID23-2 structural biology microfocus beamline at the ESRF. *Journal of synchrotron radiation* *17*, 107-118.

Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A., and Dreyfuss, G. (2001a). SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Molecular cell* *7*, 1111-1117.

Friesen, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001b). The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Molecular and cellular biology* *21*, 8289-8300.

Friesen, W.J., Wyce, A., Paushkin, S., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). A novel WD repeat protein component of the methylosome binds Sm proteins. *The Journal of biological chemistry* *277*, 8243-8247.

Furst, J., Schedlbauer, A., Gandini, R., Garavaglia, M.L., Saino, S., Gschwenter, M., Sarg, B., Lindner, H., Jakab, M., Ritter, M., *et al.* (2005). ICln159 folds into a pleckstrin homology domain-like structure. Interaction with kinases and the splicing factor LSm4. *The Journal of biological chemistry* *280*, 31276-31282.

- Gilbert, W. (1978). Why genes in pieces? *Nature* 271, 501.
- Gliko, O., Neumaier, N., Pan, W., Haase, I., Fischer, M., Bacher, A., Weinkauff, S., and Vekilov, P.G. (2005). A metastable prerequisite for the growth of lumazine synthase crystals. *Journal of the American Chemical Society* 127, 3433-3438.
- Gopal, G.J., and Kumar, A. (2013). Strategies for the production of recombinant protein in *Escherichia coli*. *The protein journal* 32, 419-425.
- Gornall, A.G., Bardawill, C.J., and David, M.M. (1949). Determination of serum proteins by means of the biuret reaction. *The Journal of biological chemistry* 177, 751-766.
- Grimm, C., Chari, A., Pelz, J.P., Kuper, J., Kisker, C., Diederichs, K., Stark, H., Schindelin, H., and Fischer, U. (2013). Structural basis of assembly chaperone-mediated snRNP formation. *Molecular cell* 49, 692-703.
- Grimm, C., Chari, A., Reuter, K., and Fischer, U. (2010). A crystallization screen based on alternative polymeric precipitants. *Acta crystallographica Section D, Biological crystallography* 66, 685-697.
- Grimmler, M., Otter, S., Peter, C., Muller, F., Chari, A., and Fischer, U. (2005). Unrip, a factor implicated in cap-independent translation, associates with the cytosolic SMN complex and influences its intracellular localization. *Human molecular genetics* 14, 3099-3111.
- Guderian, G., Peter, C., Wiesner, J., Sickmann, A., Schulze-Osthoff, K., Fischer, U., and Grimmler, M. (2011). RioK1, a new interactor of protein arginine methyltransferase 5 (PRMT5), competes with pICln for binding and modulates PRMT5 complex composition and substrate specificity. *The Journal of biological chemistry* 286, 1976-1986.
- Hall, D., and Minton, A.P. (2003). Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochimica et biophysica acta* 1649, 127-139.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology* 166, 557-580.
- Haslam, R.J., Koide, H.B., and Hemmings, B.A. (1993). Pleckstrin domain homology. *Nature* 363, 309-310.
- Hinsen, K. (1998). Analysis of domain motions by approximate normal mode calculations. *Proteins* 33, 417-429.
- Hirano, Y., Hendil, K.B., Yashiroda, H., Iemura, S., Nagane, R., Hioki, Y., Natsume, T., Tanaka, K., and Murata, S. (2005). A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. *Nature* 437, 1381-1385.
- Hondele, M., and Ladurner, A.G. (2011). The chaperone-histone partnership: for the greater good of histone traffic and chromatin plasticity. *Current opinion in structural biology* 21, 698-708.
- Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M., and Luhrmann, R. (1998). Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *The EMBO journal* 17, 4114-4126.
- Jablonka, S., Holtmann, B., Meister, G., Bandilla, M., Rossoll, W., Fischer, U., and Sendtner, M. (2002). Gene targeting of Gemin2 in mice reveals a correlation between defects in the biogenesis of U

snRNPs and motoneuron cell death. *Proceedings of the National Academy of Sciences of the United States of America* *99*, 10126-10131.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology* *3*, 318-356.

Jady, B.E., Darzacq, X., Tucker, K.E., Matera, A.G., Bertrand, E., and Kiss, T. (2003). Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *The EMBO journal* *22*, 1878-1888.

Kabsch, W. (2010). Xds. *Acta crystallographica Section D, Biological crystallography* *66*, 125-132.

Kambach, C., Walke, S., Young, R., Avis, J.M., de la Fortelle, E., Raker, V.A., Luhrmann, R., Li, J., and Nagai, K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* *96*, 375-387.

Karbstein, K. (2010). Chaperoning ribosome assembly. *The Journal of cell biology* *189*, 11-12.

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. *Science* *336*, 1030-1033.

Kastner, B., Bach, M., and Luhrmann, R. (1990). Electron microscopy of snRNPs U2, U4/6 and U5: evidence for a common structure-determining principle in the major UsnRNP family. *Molecular biology reports* *14*, 171.

Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., *et al.* (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. *Nature methods* *5*, 53-55.

Knablein, J., Neufeind, T., Schneider, F., Bergner, A., Messerschmidt, A., Lowe, J., Steipe, B., and Huber, R. (1997). Ta6Br(2+)12, a tool for phase determination of large biological assemblies by X-ray crystallography. *Journal of molecular biology* *270*, 1-7.

Kroiss, M., Schultz, J., Wiesner, J., Chari, A., Sickmann, A., and Fischer, U. (2008). Evolution of an RNP assembly system: a minimal SMN complex facilitates formation of UsnRNPs in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* *105*, 10045-10050.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* *227*, 680-685.

Lamond, A.I., Konarska, M.M., Grabowski, P.J., and Sharp, P.A. (1988). Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. *Proceedings of the National Academy of Sciences of the United States of America* *85*, 411-415.

Laskey, R.A., Honda, B.M., Mills, A.D., and Finch, J.T. (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature* *275*, 416-420.

Laskey, R.A., Mills, A.D., Philpott, A., Leno, G.H., Dilworth, S.M., and Dingwall, C. (1993). The role of nucleoplasmin in chromatin assembly and disassembly. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* *339*, 263-269; discussion 268-269.

Lehmeier, T., Foulaki, K., and Luhrmann, R. (1990). Evidence for three distinct D proteins, which react differentially with anti-Sm autoantibodies, in the cores of the major snRNPs U1, U2, U4/U6 and U5. *Nucleic acids research* *18*, 6475-6484.

Leung, A.K., Nagai, K., and Li, J. (2011). Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* *473*, 536-539.

Li, D., and Roberts, R. (2001). WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cellular and molecular life sciences : CMLS* *58*, 2085-2097.

Liautard, J.P., Sri-Widada, J., Brunel, C., and Jeanteur, P. (1982). Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. Proteins interact closely with a similar structural domain of U1, U2, U4 and U5 small nuclear RNAs. *Journal of molecular biology* *162*, 623-643.

Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997). The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* *90*, 1013-1021.

Liu, Z., Leng, E.C., Gunasekaran, K., Pentony, M., Shen, M., Howard, M., Stoops, J., Manchulenko, K., Razinkov, V., Liu, H., *et al.* (2015). A novel antibody engineering strategy for making monovalent bispecific heterodimeric IgG antibodies by electrostatic steering mechanism. *The Journal of biological chemistry* *290*, 7535-7562.

Longenecker, K.L., Garrard, S.M., Sheffield, P.J., and Derewenda, Z.S. (2001). Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI. *Acta crystallographica Section D, Biological crystallography* *57*, 679-688.

Luhrmann, R., Kastner, B., and Bach, M. (1990). Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochimica et biophysica acta* *1087*, 265-292.

Ma, Y., Dostie, J., Dreyfuss, G., and Van Duyne, G.D. (2005). The Gemin6-Gemin7 heterodimer from the survival of motor neurons complex has an Sm protein-like structure. *Structure* *13*, 883-892.

Madhani, H.D., and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annual review of genetics* *28*, 1-26.

Mandel, C.R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J.L., and Tong, L. (2006). Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* *444*, 953-956.

Martin, G., Keller, W., and Doublet, S. (2000). Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *The EMBO journal* *19*, 4193-4203.

McCarthy, A.A., Brockhauser, S., Nurizzo, D., Theveneau, P., Mairs, T., Spruce, D., Guijarro, M., Lesourd, M., Ravelli, R.B.G., and McSweeney, S. (2009). A decade of user operation on the macromolecular crystallography MAD beamline ID14-4 at the ESRF. *J Synchrotron Rad* *16*, 803-812.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *Journal of applied crystallography* *40*, 658-674.

Meister, G., Eggert, C., Buhler, D., Brahms, H., Kambach, C., and Fischer, U. (2001). Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Current biology : CB* *11*, 1990-1994.

- Meister, G., and Fischer, U. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *The EMBO journal* *21*, 5853-5863.
- Montzka, K.A., and Steitz, J.A. (1988). Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proceedings of the National Academy of Sciences of the United States of America* *85*, 8885-8889.
- Moore, M.J., and Sharp, P.A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* *365*, 364-368.
- Mouaikel, J., Narayanan, U., Verheggen, C., Matera, A.G., Bertrand, E., Tazi, J., and Bordonne, R. (2003). Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO reports* *4*, 616-622.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology* *51 Pt 1*, 263-273.
- Neer, E.J., Schmidt, C.J., Nambudripad, R., and Smith, T.F. (1994). The ancient regulatory-protein family of WD-repeat proteins. *Nature* *371*, 297-300.
- Nesic, D., Tanackovic, G., and Kramer, A. (2004). A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *Journal of cell science* *117*, 4423-4433.
- Nomura, M., Traub, P., and Bechmann, H. (1968). Hybrid 30S ribosomal particles reconstituted from components of different bacterial origins. *Nature* *219*, 793-799.
- Otter, S., Grimmler, M., Neuenkirchen, N., Chari, A., Sickmann, A., and Fischer, U. (2007). A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *The Journal of biological chemistry* *282*, 5825-5833.
- Palacios, I., Hetzer, M., Adam, S.A., and Mattaj, I.W. (1997). Nuclear import of U snRNPs requires importin beta. *The EMBO journal* *16*, 6783-6792.
- Paunov, V.N., Kaler, E.W., Sandler, S.I., and Petsev, D.N. (2001). A Model for Hydration Interactions between Apoferritin Molecules in Solution. *Journal of colloid and interface science* *240*, 640-643.
- Pellizzoni, L., Yong, J., and Dreyfuss, G. (2002). Essential role for the SMN complex in the specificity of snRNP assembly. *Science* *298*, 1775-1779.
- Plessel, G., Luhrmann, R., and Kastner, B. (1997). Electron microscopy of assembly intermediates of the snRNP core: morphological similarities between the RNA-free (E.F.G) protein heteromer and the intact snRNP core. *Journal of molecular biology* *265*, 87-94.
- Pomeranz Krummel, D.A., Oubridge, C., Leung, A.K., Li, J., and Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* *458*, 475-480.
- Pu, W.T., Krapivinsky, G.B., Krapivinsky, L., and Clapham, D.E. (1999). pICln inhibits snRNP biogenesis by binding core spliceosomal proteins. *Molecular and cellular biology* *19*, 4113-4120.
- Raker, V.A., Hartmuth, K., Kastner, B., and Luhrmann, R. (1999). Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Molecular and cellular biology* *19*, 6554-6565.

- Raker, V.A., Plessel, G., and Luhrmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *The EMBO journal* *15*, 2256-2269.
- Rubinson, K.A., Ladner, J.E., Tordova, M., and Gilliland, G.L. (2000). Cryosalts: suppression of ice formation in macromolecular crystallography. *Acta crystallographica Section D, Biological crystallography* *56*, 996-1001.
- Ruby, S.W., and Abelson, J. (1988). An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* *242*, 1028-1035.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* *74*, 5463-5467.
- Saschenbrecker, S., Bracher, A., Rao, K.V., Rao, B.V., Hartl, F.U., and Hayer-Hartl, M. (2007). Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. *Cell* *129*, 1189-1200.
- Sauter, C., Basquin, J., and Suck, D. (2003). Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic acids research* *31*, 4091-4098.
- Sauterer, R.A., Goyal, A., and Zieve, G.W. (1990). Cytoplasmic assembly of small nuclear ribonucleoprotein particles from 6 S and 20 S RNA-free intermediates in L929 mouse fibroblasts. *The Journal of biological chemistry* *265*, 1048-1058.
- Schagger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical biochemistry* *166*, 368-379.
- Schedlbauer, A., Gandini, R., Kontaxis, G., Paulmichl, M., Furst, J., and Konrat, R. (2011). The C-terminus of ICln is natively disordered but displays local structural preformation. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* *28*, 1203-1210.
- Schrank, B., Gotz, R., Gunnensen, J.M., Ure, J.M., Toyka, K.V., Smith, A.G., and Sendtner, M. (1997). Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proceedings of the National Academy of Sciences of the United States of America* *94*, 9920-9925.
- Schreiber, G., Haran, G., and Zhou, H.X. (2009). Fundamental aspects of protein-protein association kinetics. *Chemical reviews* *109*, 839-860.
- Segault, V., Will, C.L., Sproat, B.S., and Luhrmann, R. (1995). In vitro reconstitution of mammalian U2 and U5 snRNPs active in splicing: Sm proteins are functionally interchangeable and are essential for the formation of functional U2 and U5 snRNPs. *The EMBO journal* *14*, 4010-4021.
- Selenko, P., Sprangers, R., Stier, G., Buhler, D., Fischer, U., and Sattler, M. (2001). SMN tudor domain structure and its interaction with the Sm proteins. *Nature structural biology* *8*, 27-31.
- Shpargel, K.B., Praveen, K., Rajendra, T.K., and Matera, A.G. (2009). Gemin3 is an essential gene required for larval motor function and pupation in *Drosophila*. *Molecular biology of the cell* *20*, 90-101.
- Smart, O.S., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vornrhein, C., and Bricogne, G. (2012). Exploiting structure similarity in refinement: automated NCS and target-

structure restraints in BUSTER. *Acta crystallographica Section D, Biological crystallography* *68*, 368-380.

Sprinzak, E., Cokus, S., Yeates, T., Eisenberg, D., and Pellegrini, M. (2009). Detecting coordinated regulation of multi-protein complexes using logic analysis of gene expression. *BMC Systems Biology* *3*:115.

Staley, J.P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* *92*, 315-326.

Stark, H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods in enzymology* *481*, 109-126.

Stark, H., Dube, P., Luhrmann, R., and Kastner, B. (2001). Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* *409*, 539-542.

Steitz, J.A., Dreyfuss, G., Krainer, A.R., Lamond, A.I., Matera, A.G., and Padgett, R.A. (2008). Where in the cell is the minor spliceosome? *Proceedings of the National Academy of Sciences of the United States of America* *105*, 8485-8486.

Strong, M., Sawaya, M.R., Wang, S., Phillips, M., Cascio, D., and Eisenberg, D. (2006). Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* *103*, 8060-8065.

Sugimoto, K., Mori, Y., Makino, K., Ohkubo, K., and Morii, T. (2003). Functional reassembly of a split PH domain. *Journal of the American Chemical Society* *125*, 5000-5004.

Suhre, K., and Sanejouand, Y.H. (2004). ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic acids research* *32*, W610-614.

Sulzenbacher, G., Gruez, A., Roig-Zamboni, V., Spinelli, S., Valencia, C., Pagot, F., Vincentelli, R., Bignon, C., Salomoni, A., Grisel, S., *et al.* (2002). A medium-throughput crystallization approach. *Acta crystallographica Section D, Biological crystallography* *58*, 2109-2115.

Tama, F., Valle, M., Frank, J., and Brooks, C.L., 3rd (2003). Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proceedings of the National Academy of Sciences of the United States of America* *100*, 9319-9323.

Tirion, M.M. (1996). Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Physical review letters* *77*, 1905-1908.

Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., and Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nature structural & molecular biology* *18*, 1414-1420.

Vagin, A., and Teplyakov, A. (1998). A translation-function approach for heavy-atom location in macromolecular crystallography. *Acta crystallographica Section D, Biological crystallography* *54*, 400-402.

Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta crystallographica Section D, Biological crystallography* *66*, 22-25.

- van Dam, A., Winkel, I., Zijlstra-Baalbergen, J., Smeenk, R., and Cuypers, H.T. (1989). Cloned human snRNP proteins B and B' differ only in their carboxy-terminal part. *The EMBO journal* **8**, 3853-3860.
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. *Journal of computational chemistry* **26**, 1701-1718.
- Vekilov, P.G. (2003). Solvent entropy effects in the formation of protein solid phases. *Methods in enzymology* **368**, 84-105.
- Vekilov, P.G., Feeling-Taylor, A.R., Yau, S.T., and Petsev, D. (2002). Solvent entropy contribution to the free energy of protein crystallization. *Acta crystallographica Section D, Biological crystallography* **58**, 1611-1616.
- Wahl, M.C., Will, C.L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701-718.
- Weber, G., Trowitzsch, S., Kastner, B., Luhrmann, R., and Wahl, M.C. (2010). Functional organization of the Sm core in the crystal structure of human U1 snRNP. *The EMBO journal* **29**, 4172-4184.
- Will, C.L., and Luhrmann, R. (2011). Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* **3**.
- Woolfson, D.N., Evans, P.A., Hutchinson, E.G., and Thornton, J.M. (1993). Topological and stereochemical restrictions in beta-sandwich protein structures. *Protein engineering* **6**, 461-470.
- Yau, S.T., Petsev, D.N., Thomas, B.R., and Vekilov, P.G. (2000). Molecular-level thermodynamic and kinetic parameters for the self-assembly of apoferritin molecules into crystals. *Journal of molecular biology* **303**, 667-678.
- Yeh, J.I., Biemann, H.P., Prive, G.G., Pandit, J., Koshland, D.E., Jr., and Kim, S.H. (1996). High-resolution structures of the ligand binding domain of the wild-type bacterial aspartate receptor. *Journal of molecular biology* **262**, 186-201.
- Yong, J., Kasim, M., Bachorik, J.L., Wan, L., and Dreyfuss, G. (2010). Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Molecular cell* **38**, 551-562.
- Zhang, L., and Buck, M. (2013). Molecular simulations of a dynamic protein complex: role of salt-bridges and polar interactions in configurational transitions. *Biophysical journal* **105**, 2412-2417.
- Zhang, R., So, B.R., Li, P., Yong, J., Glisovic, T., Wan, L., and Dreyfuss, G. (2011). Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146**, 384-395.
- Zhou, L., Hang, J., Zhou, Y., Wan, R., Lu, G., Yin, P., Yan, C., and Shi, Y. (2014). Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* **506**, 116-120.
- Zimmerman, S.B., and Minton, A.P. (1993). Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annual review of biophysics and biomolecular structure* **22**, 27-65.

## 9. Anhang

### 9.1 Nucleoside und Nukleinbasen

A	Adenosin/ Adenin
C	Cytidin/ Cytosin
G	Guanidin/ Guanin
T	Thymidin/ Thymin
U	Uridin/ Uracil

### 9.2 Aminosäuren

<b>A</b>	Ala	Alanin	<b>M</b>	Met	Methionin
<b>C</b>	Cys	Cystein	<b>N</b>	Asn	Asparagin
<b>D</b>	Asp	Aspartat	<b>P</b>	Pro	Prolin
<b>E</b>	Glu	Glutamat	<b>Q</b>	Gln	Glutamin
<b>F</b>	Phe	Phenylalanin	<b>R</b>	Arg	Arginin
<b>G</b>	Gly	Glycin	<b>S</b>	Ser	Serin
<b>H</b>	His	Histidin	<b>T</b>	Thr	Threonin
<b>I</b>	Ile	Isoleucin	<b>V</b>	Val	Valin
<b>K</b>	Lys	Lysin	<b>W</b>	Trp	Tryptophan
<b>L</b>	Leu	Leucin	<b>Y</b>	Tyr	Tyrosin

### 9.3 Publikationen

Teile dieser Arbeit gingen in folgende Publikationen ein:

Pelz, J.-P., Schindelin, H., van Pee, K., Kuper, J., Kisker, C., Diederichs, K., Fischer, U. and Grimm, C. (2015): „Crystallizing the 6S and 8S spliceosomal assembly intermediates – a complex project.“ *Acta Crystallographica Section D*. (in Revision)

Grimm, C., Chari, A., Pelz, J. P., Kuper, J., Kisker, C., Diederichs, K., Stark, H., Schindelin, H., and Fischer, U. (2013): „Structural basis of assembly chaperone- mediated snRNP formation.“ *Molecular Cell*. 49(4):692-703

## 9.4 Primärsequenz-Alignment verschiedener pICln-Orthologe

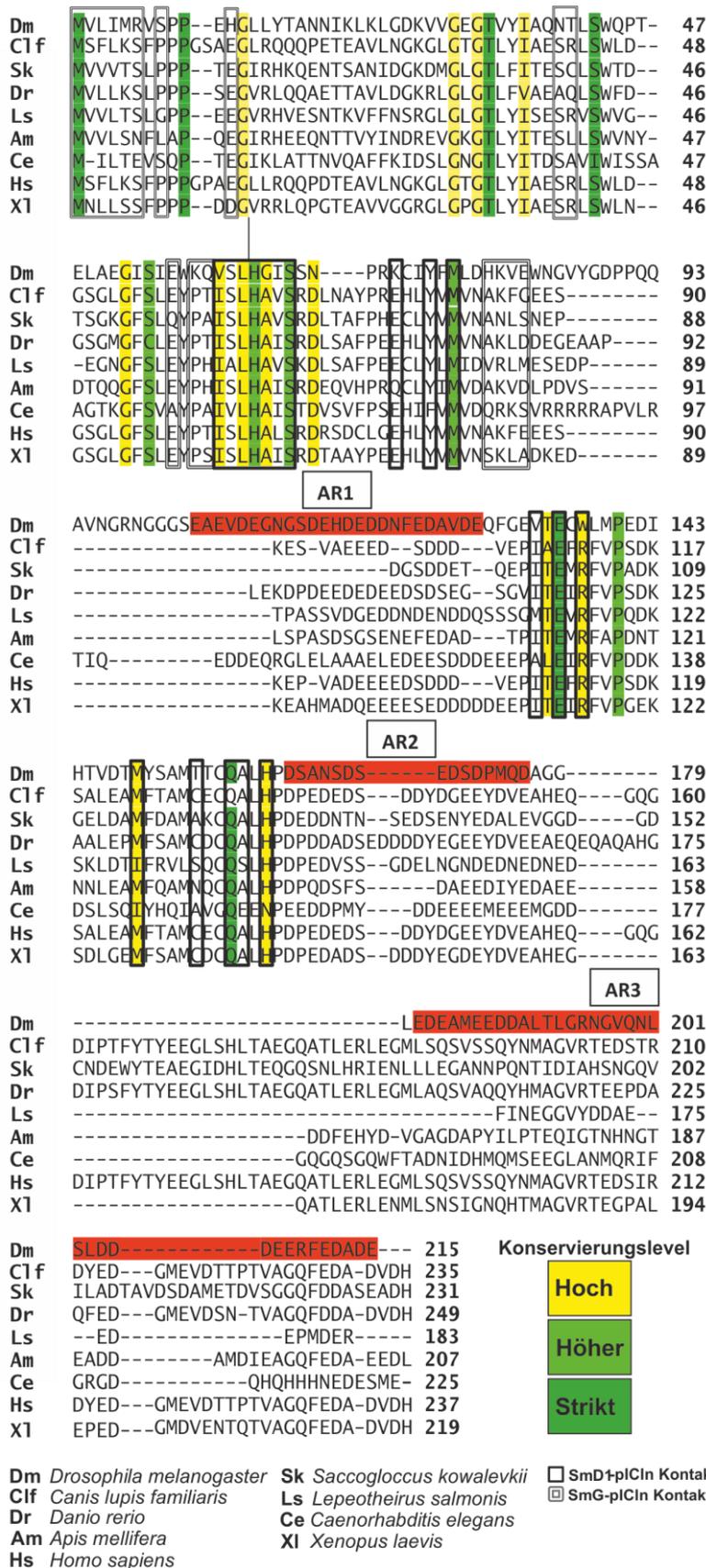


Abbildung 27 Sequenzalignment verschiedener pICln-Orthologe. Die Lokalisierung der sauren Regionen (AR) ist für *D. melanogaster* pICln durch die rote Markierung gezeigt. Die Sequenzkonservierung ist entsprechend Abbildung 14a, Seite 52, eingefärbt. Reste, die in Kontaktflächen zu SmD1 (schwarz) und SmG (grau, doppelt) involviert sind umrahmt.

## 9.5 Lebenslauf

### Hochschulstudium

01/2015 - 05/2015	Verfassung der Dissertation mit dem Titel „Strukturbiologische Untersuchungen zur Chaperone-vermittelten Zusammenlagerung von spleißosomalen U-snRNPs.
10/2010 - 12/2014	Studien zur Erlangung des naturwissenschaftlichen Doktorgrades Lehrstuhl für Biochemie, Prof. Dr. Utz Fischer Theodor-Boveri-Institut für Biowissenschaften Julius-Maximilians-Universität Würzburg
21.09.2010	Erwerb des Diploms in Biologie an der Julius-Maximilians-Universität Würzburg mit der Note „gut“.
11/2009 - 09/2010	Diplomarbeit mit dem Titel „Structural Basis of the Interaction between the La-related protein 4B Larp4B) and the cytosolic Poly(A)-binding protein (PABPC1)“. Note: „sehr gut“ Lehrstuhl für Biochemie, Prof. Dr. Utz Fischer Theodor-Boveri-Institut für Biowissenschaften Julius-Maximilians-Universität Würzburg
10/2007 - 09/2010	Hauptstudium der Biologie Fakultät für Biologie Julius-Maximilians-Universität Würzburg
30.09.2007	Erwerb des Vordiploms in Biologie an der Technischen Universität Darmstadt mit der Note „gut“.
10/2005 - 09/2007	Grundstudium der Biologie Fakultät für Biologie Technische Universität Würzburg

### Höhere Schullaufbahn

17.06.2005	Erwerb der Allgemeinen Hochschulreife mit der Note: „1,9“
07/1996 - 06/2005	Besuch des Adolf-Reichwein-Gymnasiums Heusenstamm

### Teilnahme an wissenschaftlichen Fortbildungskursen

02. - 07.07.2014	EMBO Practical Course „The Structural Characterization of Macromolecular Complexes“, EMBL Grenoble, Frankreich
07/2014 - 09/2014	Applied Pharmaceutical and Structural Bioinformatics, Fernkurs, Prof. Dr. Jarl Wikberg, Universität Uppsala, Schweden

### Teilnahme an wissenschaftlichen Kongressen

13. - 14.09.2012	2. Jahrestagung der LARP-Society, Banyuls-sur-mer, Frankreich
19. - 22.03.2012	19. Jahrestagung der Deutschen Gesellschaft für Kristallographie (DGK Conference), Freiberg
13. - 16.10.2010	Murnau Conference on Structural Biology, Murnau am Staffelsee
02. - 03.09.2010	1. Jahrestagung der LARP-Society, Imperial College, London, England
18. - 20.03.2010	EMBL Conference “The Complex Life of mRNA – From Synthesis to Decay, Heidelberg

## 9.6 Danksagung

- Ich danke Prof. Dr. Utz Fischer für das in mich gesetzte Vertrauen und die fachliche Betreuung während der Anfertigung dieser Arbeit,  
Prof. Dr. Alexander Buchberger für seine Tätigkeit als Zweitgutachter,  
Prof. Dr. Jürgen Seibel als drittem Prüfer des Promotionskolloquiums.
- Dr. Clemens Grimm möchte ich für die enge Zusammenarbeit, die fachliche Betreuung, insbesondere bei kristallographischen Fragestellungen und die vielen Gespräche, besonders während den langen Autofahrten nach Grenoble, danken.
- Ich danke Dr. Ashwin Chari, Prof. Dr. Holger Stark, Prof. Dr. Kay Diederichs und Katharina van Pee für Ihre wertvollen Beiträge zu unterschiedlichen Teilaspekten des Projektes.
- Ebenfalls möchte ich Prof. Dr. Hermann Schindelin, Prof. Dr. Caroline Kisker und Dr. Jochen Kuper für die Unterstützung bei der Strukturlösung des 8S-Komplexes und für Ihr Mitwirken bei der Entstehung beider Veröffentlichungen danken. Ebenfalls möchte ich mich dafür bedanken, dass mir die Nutzung der umfangreichen kristallographischen Einrichtungen des Rudolf-Virchow-Zentrums ermöglicht wurde.
- Ein ganz besonderer Dank geht an Emilia, Erika, Lissy, Sonja und Jürgen für Ihre technische Unterstützung. Hierbei natürlich ein ganz besonderer Dank für die regelmäßige Bereitstellung von Flüssigmedium und Agarplatten.
- Farah und Tanja möchte ich für Ihre organisatorische Unterstützung danken. Ganz besonders bleiben mir aber auch Eure lieben Worte und Aufmunterungen im Herbst 2014 in Erinnerung.
- Alle Kollegen am Lehrstuhl bleiben mir besonders wegen der guten und hilfsbereiten Arbeitsatmosphäre in der allerbesten Erinnerung. Danke für die konstruktiven, wenn auch manchmal hitzigen Diskussionen, besonders diejenigen, die bei dem einen oder anderen Bier stattfanden. Viel Erfolg und auch das nötige Glück an alle nachfolgenden Doktoranden!
- Den Spielern des SV Werder Bremen möchte ich dafür danken, dass sie mir bei all den Enttäuschungen der letzten Jahre zumindest die Schmach erspart haben, während meiner Zeit am Lehrstuhl in die 2. Bundesliga abzustiegen.
- Lieben Dank Euch guten Freunden, ob ihr mich schon seit der Schulzeit begleitet, oder in den letzten Jahren dazu gekommen seid. Ihr alle habt dafür gesorgt, dass Würzburg nicht nur Arbeitsstätte, sondern eine wunderbare Heimat geworden ist.
- Meinen Eltern und Schwiegereltern möchte ich für die Unterstützung während des Studiums und der Doktorarbeit danken. Ohne Eure regelmäßigen Aufmunterungen und die finanzielle Unterstützung wäre das Projekt wohl noch etwas steiniger geworden.
- Ein besonderer Dank gilt meiner Frau Julia. Dass ich immer auf Deine Liebe, Dein Verständnis und Deine Unterstützung zählen kann, ist die allergrößte Motivation!