

**Darstellung und Charakterisierung der humanen
3'-Prozessierungsfaktoren mPSF und CF II**

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

Der

Naturwissenschaftlichen Fakultät I Biowissenschaften

der Martin-Luther-Universität

Halle-Wittenberg,

vorgelegt

von Herrn Peter Schäfer

geb. am 25.10.1987 in Schwedt/Oder

Gutachter:

Prof. Dr. Elmar Wahle

apl. Prof. Dr. Ralph Golbik

Prof. Dr. Niels H. Gehring

Eingereicht: Halle (Saale), 16.11.2020

Verteidigt: Halle (Saale), 27.05.2021

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Eukaryotische Genexpression.....	1
1.2	Sequenzmotive in der 3'-UTR von eukaryotischen prä-mRNAs.....	2
1.3	3'-Prozessierung von prä-mRNAs in Säugetierzellen.....	4
1.4	Bekannte Faktoren der 3'-Prozessierung.....	5
1.4.1	CPSF.....	7
1.4.2	CstF.....	8
1.4.3	CF I.....	9
1.4.4	CF II.....	10
1.4.5	RNA-Polymerase II.....	13
1.4.6	Poly(A)-Polymerase und Poly(A)-Bindungsprotein.....	14
1.5	Zielstellung.....	15
2	Ergebnisse.....	16
2.1	mPSF bindet das Polyadenylierungssignal und ist ausreichend für eine spezifische Polyadenylierung.....	16
2.2	WDR33 und CPSF4 (30k) binden gemeinsam das Polyadenylierungssignal.....	20
2.3	Rekonstitution des humanen <i>Cleavage Factor</i> II (CF II) aus hPcf11 und hClp1.....	23
2.3.1	Herstellung von CF II und hClp1 aus Insektenzellen.....	24
2.3.2	CF II liegt als Dimer aus hPcf11 und hClp1 vor.....	27
2.4	Rekombinantes CF II ist in der prä-mRNA-Spaltung aktiv.....	29
2.4.1	Ein N-terminal verkürztes hPcf11 (Δ N769) in CF II ist hinreichend für eine Rekonstitution der 3'-Spaltaktivität.....	31
2.5	Die 5'-PNK Aktivität ist für die Komplementation der Spaltung von prä-mRNA nicht notwendig.....	33
2.5.1	Inaktivierung der 5'-PNK-Aktivität durch Mutagenese von hClp1.....	33
2.5.2	Charakterisierung der kinetischen Parameter der 5'-Polynukleotidkinase-Aktivität von hClp1.....	36
2.6	Humanes Pcf11 bindet bevorzugt an Guanin-reiche RNAs.....	40
2.6.1	hPcf11 bindet RNA-Substrate mit einer Mindestlänge von 14 Nukleotiden.....	40

2.6.2	hPcf11 bindet an eine Guanin-reiche Sequenz in der SV40 <i>late</i> RNA abwärts der Spaltstelle.....	44
2.6.3	Lineare Sequenzen und nicht G-Quadrupel Strukturen werden durch hPcf11 gebunden.....	46
2.6.4	Selektionsversuche mit randomisierten RNA 14-mer Sequenzen (Bind'n'Seq).....	50
2.6.5	Humanes Pcf11 bindet RNA hauptsächlich mittels seiner C-terminalen Zinkfinger-Domänen an die RNA	53
2.7	Expression und Reinigung des <i>Cleavage stimulatory Factor</i> (CstF) und des <i>Cleavage Factor I</i> (CF I).....	56
3	Diskussion.....	61
3.1	Der CPSF Subkomplex mPSF besteht aus vier Untereinheiten und bindet das Polyadenylierungssignal spezifisch und mit hoher Affinität.....	61
3.2	WDR33 und CPSF4 binden gemeinsam das Polyadenylierungssignal (PAS).....	62
3.3	hCpl1 und hPcf11 rekonstituieren CF II-Aktivität <i>in-vitro</i> und interagieren über konservierte Bereiche miteinander.....	63
3.4	Die 5'-PNK-Aktivität des hClp1 ist für die CF II-Aktivität in der 3'-Spaltreaktion nicht notwendig.....	65
3.5	hClp1 ist eine aktive PNK und unterscheidet sich in der ATP-Affinität von ceClp1 aus <i>C. elegans</i>	66
3.6	Der CF II-Komplex bindet hauptsächlich über hPcf11 an G-reiche RNA-Sequenzen.....	68
3.7	hPcf11 bindet die RNA in einer linearen Konformation und nicht als G-Quadruplex-Struktur.....	71
4	Zusammenfassung.....	72
5	Materialien.....	75
5.1	Chemikalien	75
5.2	Verbrauchsmittel	77
5.3	Geräte.....	78
5.4	Chromatographiesäulen und -materialien.....	79
5.5	Software.....	79
5.6	DNA und Proteine.....	80
5.6.1	Bacmide und Baculoviren.....	80

5.6.2	Plasmide und Transkriptionsvorlagen.....	81
5.6.3	DNA-Primer	82
5.6.4	Synthetische und enzymatisch hergestellte RNAs	83
5.6.5	Antikörper	85
5.6.6	Proteinpräparationen.....	86
5.6.7	Kommerziell erworbene Proteine	87
5.7	Zellstämme und Medien	87
5.7.1	Bakterien.....	87
5.7.2	Eukaryotische Zelllinien.....	88
5.7.3	Kulturmedien.....	88
5.7.4	Medienzusätze und Antibiotika.....	89
6	Methoden.....	90
6.1	Mikro- und molekularbiologische Methoden.....	90
6.1.1	Sterilisation.....	90
6.1.2	Herstellung und Transformation elektrokompetenter Bakterien	90
6.1.3	Bakterielle Glycerinkulturen	90
6.1.4	Verwendung des MultiBac Systems	91
6.1.5	Klonierung und Multiplikation von Zielgenen in Transfervektoren.....	91
6.2	Zellkultur	94
6.2.1	Aussaat und Propagation von Insektenzellen.....	94
6.2.2	Bestimmung der Zelldichte von Zellsuspensionen	95
6.2.3	Herstellung von tiefgefrorenen Vorratskulturen (<i>cryo stocks</i>).....	95
6.2.4	Transfektion von Sf21-Zellen mit Bacmid-DNA.....	96
6.2.5	Baculovirale Proteinexpression in Sf21/High5-Zellen	97
6.3	Methoden der Nukleinsäure-Biochemie	97
6.3.1	Radioaktive 5'-Markierung von Nukleinsäuren mittels T4-Polynukleotidkinase... 98	
6.3.2	Affinitätsreinigung von Nukleinsäuren mittels DEAE-Methode	99
6.3.3	Synthese von (radioaktiv markierten) RNAs durch <i>in vitro</i> Transkription.....	99
6.3.4	Zirkulardichroismus (CD)-Messungen von Ribonukleinsäuren	100

6.3.5	Agarose-Gelelektrophorese.....	101
6.3.6	Denaturierende Harnstoff-Polyacrylamid-Gelelektrophorese.....	101
6.4	Proteinbiochemische Methoden	102
6.4.1	Konzentrationsbestimmung von Proteinen.....	102
6.4.2	Glycin-SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE).....	103
6.4.3	Tricin-SDS-Polyacrylamid-Gelelektrophorese (Tricin-SDS-PAGE)	104
6.4.4	Kolloidale Färbung von SDS-PA-Gelen mittels Coomassie-Brilliant-Blau.....	105
6.4.5	Silberfärbung von SDS-Polyacrylamid-Gelen.....	105
6.4.6	Western-Blot-Analyse von Zielproteinen	106
6.4.7	Methoden der Proteinfractionierung und -Reinigung	107
6.5	Protein-Aktivitätsassays und RNA-Bindungsassays.....	113
6.5.1	Filterbindungsassay	113
6.5.2	UV-Vernetzung von Proteinen an RNAs.....	114
6.5.3	Limitierte Proteolyse von WDR33.....	115
6.5.4	<i>Electrophoretic Mobility Shift Assay</i> (EMSA).....	115
6.5.5	Native Polyacrylamid-Gelelektrophorese	116
6.5.6	5'-Polynukleotidkinase Assays mit CF II und hClp1	117
6.5.7	RNA 3'-Spaltassay.....	118
6.5.8	RNA-Selektion aus Pools mit randomisierten 14-mer Sequenzen (Bind'n'Seq) ..	119
7	Literaturverzeichnis.....	122
7.1	Wissenschaftliche Publikationen.....	122
7.2	Dissertationen und Abschlussarbeiten.....	134
7.3	Internetquellen	135
8	Tabellen- und Abbildungsverzeichnis.....	137
8.1	Abbildungsverzeichnis.....	137
8.2	Tabellenverzeichnis.....	139
8.3	Abkürzungsverzeichnis	139
9	Anhang.....	VII
9.1	Zusätzliche Abbildungen.....	VII

9.2	Lebenslauf.....	XXXII
9.3	Publikationen.....	XXXIII
9.4	Eidesstattliche Erklärung	XXXIV
9.5	Danksagung.....	XXXV

1 Einleitung

1.1 Eukaryotische Genexpression

In allen Zellen wird die kodierte Erbinformation in Form von Desoxyribonukleinsäure (DNA) in eine kurzlebige Kopie, die Ribonukleinsäure (RNA), transkribiert. Dies gestattet eine räumliche Trennung von Erbinformation und dem Ort der Proteinsynthese. In Eukaryoten wird zusätzlich durch die Trennung der RNA-Synthese im Nukleus und der Proteinsynthese im Zytoplasma eine direkte Kopplung beider Vorgänge verhindert. Die RNA-Polymerase II synthetisiert unter anderem die Transkripte der proteinkodierenden DNA-Sequenzen. Dabei wird in Metazoen ein *messenger* RNA (mRNA)-Vorläufer synthetisiert. Diese prä-mRNA wird in weiteren Schritten modifiziert und prozessiert. So wird am 5'-Ende schon kurz nach Synthesebeginn ein 7-Methylguanosin-*Cap* (m7GpppN) angebracht (Ho und Shuman, 1999; Jove und Manley, 1984; McCracken u. a., 1997; Rasmussen und Lis, 1993). Während der weiteren Synthese werden die nicht kodierenden Sequenzen innerhalb des translatierten Bereichs (Introns) durch das Spleißosom enzymatisch herausgeschnitten und die verbleibenden kodierenden Sequenzen (Exons) zusammengefügt (Shi, 2017).

Die kodierende Sequenz wird an beiden Enden durch nicht-translatierte Bereiche variabler Länge (UTR) mit vorwiegend regulatorischer Funktion flankiert. Als Abschluss der Synthese erfolgt die Termination sowie die Abspaltung eines Teils des 3'-UTR hinter einer spezifischen Signalsequenz (Birse u. a., 1998; Connelly und Manley, 1988; Lunde u. a., 2010) und der Synthese eines Poly(A)-Schwanzes am neu entstandenen 3'-Ende (Abbildung 1.1).

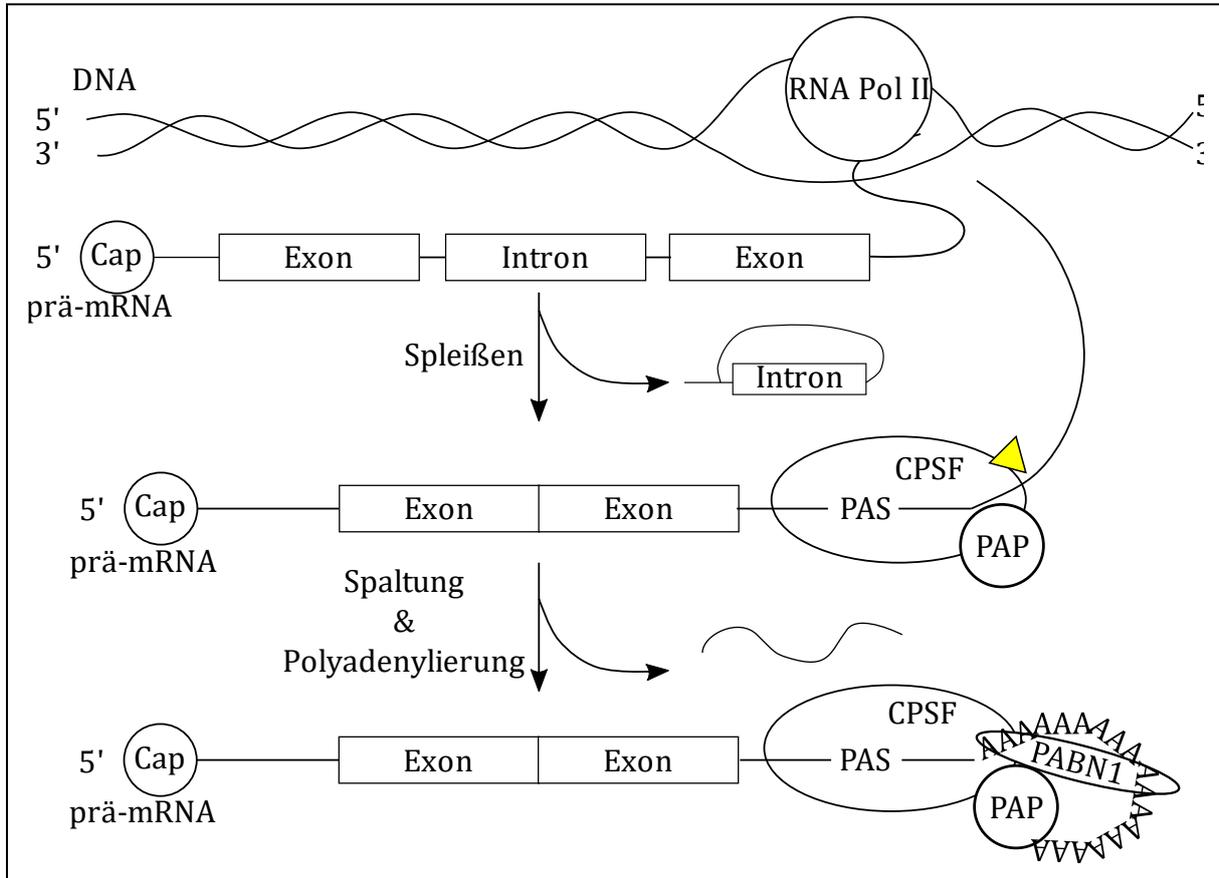


Abbildung 1.1 Synthese und Prozessierung der mRNA.

Vereinfachte schematische Darstellung der zentralen Schritte in der 3'-Prozessierung eukaryotischer prä-mRNAs. RNA Polymerase II (RNA Pol II) transkribiert den prä-mRNA Strang. Nach dem Capping erfolgt co-transkriptional das Spleißen durch das Spleißosom. Die Spaltung an der Spaltstelle (gelber Pfeilkopf) nach dem Polyadenylierungssignal (PAS) wird durch den zentralen Faktor CPSF katalysiert. Abschließend erfolgt die Polyadenylierung des 3'-Endes durch die kanonische Poly(A)-Polymerase (PAP) und der entstehende Poly(A)-Schwanz wird durch das nukleäre Poly(A)-Bindungsprotein (PABN1) bedeckt.

1.2 Sequenzmotive in der 3'-UTR von eukaryotischen prä-mRNAs

Die Prozessierung aller eukaryotischen prä-mRNAs wird durch Sequenzelemente innerhalb der RNA gesteuert, welche mit unterschiedlicher Spezifität und Affinität durch die beteiligten Proteinkomplexe erkannt und gebunden werden.

Das wichtigste Motiv ist wohl das kanonische Polyadenylierungssignal (PAS). Dessen Sequenz „AAUAAA“ ist hoch konserviert, selbst zwischen entfernt verwandten Eukaryoten. Es wird durch die beiden Untereinheiten CPSF4 und WDR33 des CPSF-Komplexes erkannt und spezifisch gebunden (Clerici u. a., 2018; Keller u. a., 1991; Schönemann u. a., 2014; Shi, 2014; Sun u. a., 2018). Schon der Austausch einzelner Basen in dieser Sequenz führt *in vitro* zu einer schwächeren Bindung des CPSF-Komplexes an die RNA und somit zu einer verringerten Prozessierung der zugehörigen Spaltstelle innerhalb der prä-mRNA. Die Punktmutation Uracil zu Guanin in dieser

Sequenz führt *in vitro* zu einem nahezu vollständigen Verlust der Bindung des CPSF. So kann diese „AAGAAA“ Sequenz als Spezifitätskontrolle in RNA-3'-Spaltassays verwendet werden (Bienroth u. a., 1991). *In vivo* werden Abweichungen von der Konsensussequenz nicht so stark diskriminiert. Die Abweichung „AUUAAA“ macht mit dem Konsensus „AAUAAA“ etwa 70–75% aller PAS in Säugern aus. Weitere 20% enthalten Abweichungen einzelner Nukleotide innerhalb dieser beiden Sequenzen. Etwa 10% der kartierten PAS können zudem nicht von dem Konsensus abgeleitet werden (Gruber u. a., 2016; Tian und Gruber, 2012). Das PAS liegt in einem Bereich zwischen 10-40 Nukleotiden stromaufwärts der eigentlichen Spaltstelle, wobei eine starke Häufung bei -21 Nukleotiden der Spaltstelle festgestellt wurde. Die Spaltstelle wird in den meisten Fällen durch ein benachbartes Adenosin gekennzeichnet und die RNA in 3'-Richtung davon durch die Endonuklease CPSF3 gespalten (Chen u. a., 1995; Wang u. a., 2018).

Weitere Elemente sind das stromaufwärts des PAS gelegene Motiv „UGUA“ (*upstream element* - USE), welches durch den Faktor CF I gebunden wird (Brown und Gilmartin, 2003; Martin u. a., 2012). Ähnlich werden auch die stimulierenden Effekte von CstF und hFip1 vermittelt. Die Sequenzmotive sind weniger konserviert und in ihrer Basenabfolge auch weniger spezifisch. CstF bindet abwärts des PAS eine „U/GU“-reiche Sequenz (*downstream element* - DSE; (Pérez Cañadillas und Varani, 2003; Yang u. a., 2018), während hFip1 einen U-reichen Sequenzabschnitt auf der 5'-Seite des PAS bevorzugt bindet (Lackford u. a., 2014). Durch die Interaktion dieser Faktoren mit CPSF führen die Protein-RNA-Interaktionen zu einer kooperativen Bindung des CPSF an dem RNA-Abschnitt (Murthy und Manley, 1995). Somit also auch zu einer verstärkten Prozessierung der prä-mRNA. Dabei enthält nicht jede RNA alle Motive (Gruber u. a., 2016). Auch können die Motive unterschiedlich stark konserviert gegenüber einer Konsensussequenz sein. Somit kann durch das Zusammenspiel aus Vorhandensein des Bindemotivs, umgebender RNA-Sequenz und der Interaktion der bindenden Proteine die Prozessierung der prä-mRNA ausgiebig kontrolliert werden (Ogorodnikov u. a., 2016).

Neben den genannten bekannten Motiven zeigen statistische Analysen auch weitere Sequenzmotive, welche in der Nähe des PAS mit vermehrter Häufigkeit vorkommen (HU u. a., 2005). Dies legt die Vermutung nahe, dass auch diese Sequenzen in Verbindung mit möglicherweise unbekanntem RNA-Bindungsproteinen auf die Prozessierung der prä-mRNA wirken könnten. Das Schema in der Abbildung 1.2 zeigt die wichtigsten Sequenzmotive und deren ungefähre Position relativ zum PAS.

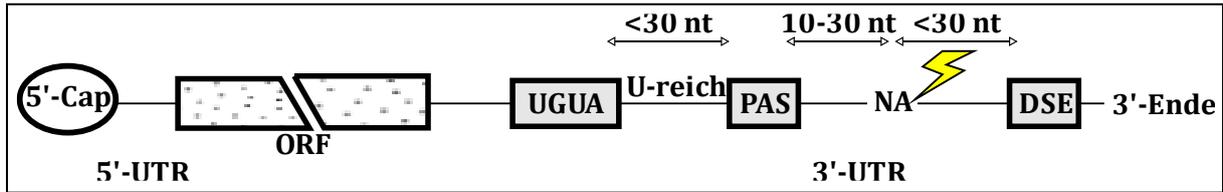


Abbildung 1.2 Schematische Darstellung bekannter 3'-UTR Sequenzmotive in prä-mRNAs. Das Upstream Element „UGUA“ wird durch CF I erkannt. Das kanonische Polyadenylierungssignal (PAS) „AAUAAA“ wird durch CPSF erkannt. Zwischen den beiden Elementen wird ein U-reicher Abschnitt durch hFip1 gebunden. Die Spaltstelle liegt etwa 21 nt hinter dem PAS und ist in der Regel ein (N)A-Dinukleotid (gelber Blitz). Das Downstream Element (DSE) wird durch CstF gebunden.

1.3 3'-Prozessierung von prä-mRNAs in Säugetierzellen

In dieser Arbeit liegt das besondere Augenmerk auf dem letzten Schritt der mRNA-Reifung in humanen Zellen, der 3'-Prozessierung. Der CPSF Kernkomplex mit seinen mindestens sieben Untereinheiten bindet das spezifische Polyadenylierungssignal (PAS) und ist der zentrale Faktor in den letzten beiden Schritten der 3'-Prozessierung, die Spaltung der RNA und die anschließende Polyadenylierung. Mit CPSF3 besitzt der CPSF-Komplex die katalytische Aktivität zur endonukleolytischen Spaltung der prä-mRNA (Mandel u. a., 2006; Shi und Manley, 2015; Xiang u. a., 2014). Diese Teilreaktionen der endonukleolytischen Spaltung der prä-mRNA ist stark reguliert und erfolgt nur nach der spezifischen Erkennung des PAS. Die Erkennung wird zusätzlich durch weitere stimulierende Komplexe beeinflusst und die Spaltung der RNA aktiviert. Das abgespaltene Fragment stromabwärts der Spaltstelle wird im Anschluss an die Spaltung rasch abgebaut (Millevoi und Vagner, 2009; Proudfoot, 2011; Wahle und Rügsegger, 1999; Zhao u. a., 1999), wohingegen das 5'-Produkt als stabile mRNA erhalten bleibt. Bislang sind drei weitere heteromere Komplexe bekannt, welche die Spaltstelle zusätzlich definieren und dabei helfen CPSF die Spaltung katalysieren. Die unterstützenden Komplexe CstF und CF I binden Sequenzmotive nahe der Spaltstelle und halten durch Proteininteraktion mit CPSF dieses in der Nähe des Polyadenylierungssignals (Abbildung 1.2). Ein weiterer Faktor ist CF II. Er stimuliert durch Interaktion mit den anderen Spaltfaktoren zusätzlich die Spaltung, ist in Säugern aber weitgehend nicht charakterisiert (de Vries, 2000). Diese zentralen Faktoren werden im späteren Verlauf genauer beschrieben und eingeordnet (Tabelle 1.1). Hinzu kommen noch viele der Spaltung zugeordnete Proteine, welche mit den bekannten Proteinen assoziieren oder aufgrund von Homologien zu anderen Organismen als wahrscheinliche Faktoren gelten. So zum Beispiel Rbbp6 (Di Giammartino u. a., 2014) oder UBE3D (Pearson u. a., 2019). Eine Übersichtsdarstellung bekannter Proteine ist in der Abbildung 1.3 zu sehen.

Der Poly(A)-Schwanz wird direkt im Anschluss an die Spaltung durch Poly(A)-Polymerase synthetisiert. Diese benötigt für die PAS-abhängige Substraterkennung die Bindung an CPSF und

für eine hohe Prozessivität die zusätzliche Bindung des Poly(A)-Bindungsprotein (PABPN1), um das RNA-Substrat zu polyadenylieren. Dadurch kann in menschlichen Zellen innerhalb von Sekunden ein Poly(A)-Schwanz definierter Länge von 250 Nukleotiden synthetisiert werden (Kühn u. a., 2009, 2017; Ogorodnikov u. a., 2016).

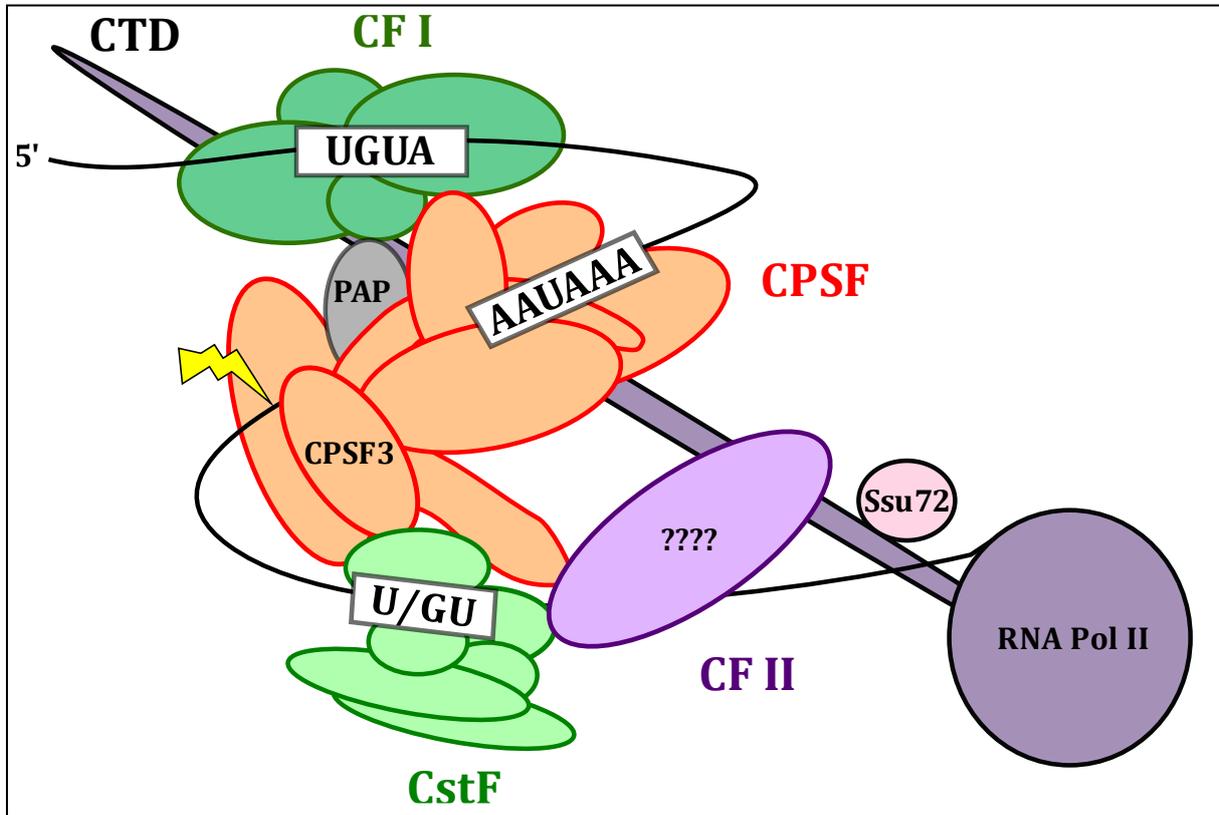


Abbildung 1.3 Schematische Darstellung bekannter 3'-Prozessierungsfaktoren während der Transkription. Die C-terminale Domäne (CTD) der RNA-Polymerase II (RNA Pol II) dient als Plattform für die 3'-Spaltfaktoren. Sie wird durch Phosphatasen wie Ssu72 im Phosphorylierungsstatus beeinflusst. Cleavage Factor I (CF I) bindet oberhalb der Spaltstelle an das UGUA-Motiv der RNA. Der Cleavage and Polyadenylation Specificity Factor (CPSF) bindet das Polyadenylierungssignal (PAS) und katalysiert mittels CPSF3 die Spaltung an der Spaltstelle (gelber Blitz). Cleavage stimulatory Factor (CstF) bindet unterhalb der Spaltstelle an ein GU/U-reiches Element. Cleavage Factor II (CFII) interagiert in unbekannter Zusammensetzung mit den anderen Faktoren und stimuliert die Aktivität des CPSF.

1.4 Bekannte Faktoren der 3'-Prozessierung

In der Tabelle 1.1 findet sich eine vergleichende Übersicht der im Folgenden beschriebenen Prozessierungsfaktoren aus Hefe und Mensch. Dabei erkennt man deutlich, dass der Hefekomplex weniger Komponenten aufweist als der menschliche Komplex. Zudem sind vermutlich noch nicht alle Komponenten des menschlichen Komplexes bekannt, denn eine *in vitro* Rekonstitution der Reaktion, wie für den Hefe-Komplex (Hill u. a., 2019), ist noch nicht beschrieben. Einige der identifizierten Homologen zu Hefeproteinen konnten noch nicht eindeutig belegt werden (markiert durch *). Im Allgemeinen sind die Säugetierproteine größer als die homologen Proteine

der Hefe und einige zeigen zudem unterschiedliche Zuordnungen in den Proteinkomplexen (Übersicht der Domänenstruktur der humanen Proteine s. Anhang Abbildung 9.4).

*Tabelle 1.1 Vergleichende Gegenüberstellung von Proteinen der 3'-Prozessierung aus Menschen und Bäckerhefe. Farbliche Unterlegung soll die Zugehörigkeit zu einem 3'-Spaltfaktor symbolisieren. Durch Sequenzhomologie vermutete, aber noch nicht experimentell bestätigte, Homologe in Menschen zu bekannten Hefeproteinen sind markiert durch *.*

<u>Homo Sapiens</u>		<u>Saccharomyces cerevisiae</u>		
CF I	CPSF5 (CF I 25k) CPSF7 (CF I 59k) CPSF6 (CF I 68/72k)			
CstF	CstF1 (50k) CstF2 τ (64k τ) CstF2 (64k) CstF3 (77k)	Rna15p Rna14p	CF IA	
	CF II?	hClp1 hPcf11		
		Hrp1p	CF IB	
CPSF	mPSF CPSF4 (30k) CPSF1 (160k) WDR33 hFip1	Yth1p Cft1/Yhh1p Psf2p Fip1p	PF/ Polymerase-Modul	CPF
	mCF CPSF3 (73k) CPSF2 (100k) Rbbp6* Symplekin	Ysh1p/Brr5 Cft2/Ydh1p Mpe1p Pta1p	CF II/ Nuklease-Modul	
	Ssu72 WDR82 PNUTS * PP1 α CstF2 τ (64k τ)*	Ssu72p Swd2p Ref2p Glc7p Pti1p	Phosphatase-Modul	
	Poly(A)-Polymerase	Pap1p		
	Poly(A)-Bindungsprotein (PABPN1)			
	RNA-Pol II Untereinheit Rpb1 (CTD)	RPB1p (CTD)		

1.4.1 CPSF

Der *Cleavage and polyadenylation specificity factor* (CPSF) ist der zentrale Faktor in der 3'-Prozessierung von prä-mRNAs. Zum einen erkennt und bindet er die kanonische Erkennungssequenz des Polyadenylierungssignals „AAUAAA“. Zum anderen enthält er auch die Endonuklease CPSF3, welche die prä-mRNA an der Spaltstelle schneidet und so ein definiertes 3'-Hydroxyende an der mRNA erzeugt. Dieses kann dann im Anschluss an die Spaltung für die Polyadenylierung durch die Poly(A)-Polymerase genutzt werden.

Zunächst wurde angenommen, dass der Säuger CPSF-Komplex aus nur vier Untereinheiten besteht, CPSF1–4 (160k, 100k, 72k, 30k) (Bienroth u. a., 1991; Murthy und Manley, 1992). Diese waren aber rekombinant nicht in der Lage die Spaltungs- und Polyadenylierungsreaktion zu rekonstituieren. Das humane Homologe des Hefe-Fip1p wurde später durch Sequenzvergleiche identifiziert. Es zeigte sich eine stabile Assoziation an die bekannten CPSF-Untereinheiten sowie eine Affinität für AU-reiche RNA Sequenzen (Kaufmann u. a., 2004).

Auch Symplekin und WDR33 wurden als mögliche Komponenten des CPSF Komplexes identifiziert (Shi u. a., 2009). Dies legt nahe, dass CPSF in Säugern mindestens aus den sieben oben genannten Untereinheiten besteht. Der CPSF-Komplex kann grob in zwei funktionelle Subkomplexe aufgeteilt werden, welche den oben beschriebenen Teilreaktionen Spaltung und Polyadenylierung entsprechen.

Zum einen wird die endonukleolytische Spaltung der prä-mRNA durch den Unterkomplex des sogenannten *mammalian Cleavage Factor* (mCF) vermittelt. Dazu zählen die Nuklease CPSF3 sowie die assoziierten Proteine CPSF2 und Symplekin (Sun u. a., 2020). CPSF2 und 3 gehören zur Proteinfamilie der β -CASP Nukleasen, einer Untergruppe der Metallo- β -Lactamase Proteinsuperfamilie. Dabei bilden CPSF2 und 3 ein fest assoziiertes Heterodimer, welches auch nur in dieser Kombination aus beiden Proteinen aktiv ist (Kolev u. a., 2008; Kyburz u. a., 2003; Sullivan u. a., 2009). Die katalytische Aktivität befindet sich jedoch höchstwahrscheinlich nur in CPSF3, da Aminosäureabweichungen innerhalb der Zink-koordinierenden Domäne des CPSF2 dieses katalytisch inaktivieren. Die Kristallstruktur eines CPSF3-Fragmentes zeigt eine Faltung mit aktivem Zentrum (Mandel u. a., 2006). Eine katalytische Aktivität konnte für das CPSF3 allein noch nicht eindeutig nachgewiesen werden. Biochemische Versuche liefern aber weitere starke Hinweise auf die Aktivität im Komplex. Mutationen des aktiven Zentrums sind letal bei *in vivo* Versuchen in Hefe und Spaltung ist Zink-abhängig (Ryan u. a., 2004). Zudem ist CPSF3 an der Spaltung der Histon-RNA beteiligt (Sun u. a., 2020).

An CPSF2 und 3 bindet in Säugern das Protein Symplekin (Sullivan u. a., 2009). Diesem wird vor allem eine unterstützende Rolle in der Stabilität des Komplexes zugeschrieben. Es kontaktiert viele der Prozessierungsfaktoren (Kolev und Steitz, 2005; Ruepp u. a., 2011; Xiang u. a., 2010). Das

Hefe-Homologe Pta1 ist im Unterschied zum Säugerprotein lose an diese beiden Proteine gebunden und dort dem Phosphatase-Modul zuzuordnen (Lidschreiber u. a., 2018). Erfolgreiche *in vitro* Versuche mit Homologen aus Säugern sind bislang nur für rekombinante Proteine der Histon-Prozessierung beschrieben (Sun u. a., 2020).

Der *mammalian Polyadenylation Specificity Factor* (mPSF) – bestehend aus den Proteinen CPSF1, WDR33, hFip1, CPSF4 – vermittelt die „AAUAAA“-Bindung und zusammen mit der Poly(A)-Polymerase und dem Poly(A)-Bindungsprotein die spezifische und längenabhängige Polyadenylierung der mRNA. Dies konnte im Verlauf der Promotion von L. Schönemann (Schönemann u. a., 2014b; Dissertation 2017) und der Arbeit von (Shi, 2014) zum ersten Mal gezeigt werden.

Welche der vier Untereinheiten jedoch die Bindung an das Polyadenylierungssignal vermitteln, war zum Anfang der Arbeit nicht bekannt und Gegenstand der Arbeit. So sollte der rekonstituierte Faktor mPSF hinsichtlich der Funktion, Struktur und Bindung charakterisiert werden.

1.4.2 CstF

Der *Cleavage Stimulation Factor* (CstF) ist ein Komplex aus drei verschiedenen Proteinen. Diese liegen jeweils als Homodimer im Komplex vor, sodass am Ende sechs Polypeptide den CstF-Komplex konstituieren. Die größte Untereinheit CstF3 bildet über die N-terminalen HAT-N und HAT-C Domänen eine stabile dimere Struktur (Bai u. a., 2007). Durch Kontakte im C-terminalen Bereich zu den beiden weiteren Untereinheiten CstF1 und CstF2 nimmt es im Komplex vor allem eine stabilisierende Funktion ein. *CryoEM*-Strukturen zeigen außerdem Interaktionen von CstF3 mit CPSF1 und WDR33 (Zhang u. a., 2020), sowie mit hFip1 (Kaufmann u. a., 2004) und der RNA Pol II CTD (Takagaki und Manley, 2000). Die 64 kDa große Untereinheit (CstF2) interagiert über eine Hinge-Region mit den letzten 30 Aminosäuren im C-Terminus des CstF3 (Moreno-Morcillo u. a., 2011). Ein N-terminales *RNA Recognition Module* (RRM) in CstF2 bindet an GU/U-reiche RNA-Sequenzen, dem sogenannten *Downstream Element* (DSE), welches sich innerhalb von 40 nt in 3'-Richtung nach der Spaltstelle befindet Abbildung 1.2 (Beyer et al., 1997; Yoshio Takagaki & Manley, 1997). Diese RNA-Interaktion kann durch Bindung des CstF2 an CstF3 beeinflusst und verstärkt werden (Grozdanov u. a., 2018). In Säugern existiert eine zweite Variante des CstF2, das CstF2 τ . Anders als das auf dem X-Chromosom codierte CstF2, ist das 2 τ -Protein auf Chromosom 19 codiert (Dass u. a., 2001; Wallace u. a., 1999). Es zeigt eine etwa 75%ige Übereinstimmung der Aminosäuresequenz und scheint das normale CstF2 ersetzen zu können. Allein die Sequenzspezifität ist von GU-reich zu U-reich verschoben (Monarez u. a., 2007). CstF2 kann über die Hinge-Region an den N-Terminus von Symplekin binden, allerdings scheint diese Interaktion mit der CstF3 Bindung zu konkurrieren (Takagaki und Manley, 2000). Des Weiteren bindet CstF2 an hClp1 (de Vries, 2000) und stellt so eine weitere Vernetzung der stimulierenden Faktoren der

3'-Prozessierung her. Die Proteine CstF2 und 3 sind hoch konservierte Proteine in Eukaryoten, CstF1 hingegen existiert z.B. in Hefe nicht (s. Tabelle 1.1). An einen Interaktionsbereich zwischen HAT-C Domäne und CstF2 Interaktion des CstF3 bindet die 50 kDa große Untereinheit (CstF1) über dessen WD40 Domäne im C-terminalen Bereich (Yang u. a., 2018). Das CstF1 bildet zudem eine dimere Struktur über eine Interaktion der N-terminalen Domäne. Es wird vermutet, dass die Bindung des CstF1 die Beweglichkeit des CstF3 und CstF2 Komplexes verringert und die dimere Struktur des Komplexes so zusätzlich stabilisiert wird.

1.4.3 CF I

Der *Cleavage Factor I* (CF I) ist ein Heteromer aus zwei Homodimeren. Es ist zum einen aus der 25 kDa großen Untereinheit CPSF5 (CF I 25k, NUDT21) aufgebaut, welche ein Homodimer über die Interaktion der N- und C-terminalen Enden bildet (Yang u. a., 2011). CPSF5 besitzt RNA bindende Eigenschaften und erkennt über die zentrale NUDIX-Domäne spezifisch die Sequenz „UGUA“, welche etwa 40 nt in 5'-Richtung der Spaltstelle innerhalb der Substrat-RNA gehäuft auftritt (Brown und Gilmartin, 2003; Yang u. a., 2010). Die Erkennung dieses Elements trägt zusätzlich zur Definition der 3'-Spaltstelle bei, wobei CF I die Nutzung distaler PAS begünstigt (Gruber und Zavolan, 2019; Zhu u. a., 2018). Über die C-terminale Region interagiert das CPSF5 mit der zweiten Komponente des Komplexes (Yang u. a., 2011).

Die zweite Komponente in CF I bildet eine größere Untereinheit, welche wieder in zwei Kopien im Komplex vorliegt. Dabei existieren in Säugern zwei Paraloge, zum einen das 59 kDa große CPSF7 (CF I 59k) und zum anderen das 68 kDa große CPSF6 (CF I 68k). CPSF6 kann wiederum auch als alternative Spleißvariante mit zusätzlichem Exon vorkommen, bezeichnet als CF I 72k (Rüegsegger u. a., 1996, 1998). Die großen Untereinheiten binden vor allem über ein N-terminal gelegenes RRM den C-Terminus des CPSF5. Dieses RRM trägt auch unspezifisch zur RNA-Bindung bei, indem es CPSF5 kooperativ in der Bindung unterstützt (Dettwiler u. a., 2004). Des Weiteren kann CPSF6/7 über eine C-terminal gelegene RS-Domäne an SR-Proteine binden. Dadurch wird eine Bindung an hFip1 vermittelt und eine Interaktion zu mPSF vermittelt (Venkataraman, 2005; Zhu u. a., 2018).

1.4.4 CF II

Der *Cleavage Factor II* wurde, wie auch CF I und CstF, durch Fraktionierung von nukleärem Extrakt aus HeLa Zellen identifiziert (Rüegsegger u. a., 1996; Takagaki u. a., 1989; de Vries, 2000). Dabei wurde die aktivierende Eigenschaft der Fraktionen auf die Spaltreaktion durch Komplementation mit anderen Extraktfraktionen identifiziert. Diese Säulenfraktionen waren allerdings keine homogenen Fraktionen reiner Proteine, sondern zumeist noch Gemische aus verschiedensten Proteinkomplexen.

Für den humanen *Cleavage Factor II* blieb die genaue Zusammensetzung des Komplexes unbekannt. Es konnten allerdings zwei humane Homologe der Hefeproteine Pcf11p und Clp1p in dieser Fraktion identifiziert werden (de Vries, 2000). Pcf11 und Clp1 bilden mit den Homologen zu CstF 64k und CstF 77k in Hefe den CF IA Subkomplex (Gordon u. a., 2011). Eine Assoziation der beiden humanen Proteine CstF 64k und 77k mit den beiden genannten CF II Proteinen ist allerdings nicht beschrieben. Zudem besitzt der Säugetierkomplex CstF zusätzlich eine 50 kDa große Untereinheit und kann in der Fraktionierung aus HeLa-Extrakt leicht von den CF II Faktoren abgetrennt werden. Es bestand also die Vermutung, dass im Menschen die beiden Proteine hPcf11 und hClp1 einen separaten Komplex bilden, welcher ausreicht, die Spaltaktivität des CF II zu vermitteln. In der Dissertation von A. Knoth (2009) konnten zudem in einem Flag-hClp1-*Pulldown* aus HEK293-Zellen bereits identifizierte Proteine der CF II-Fraktion in der Massenspektrometrie bestätigt werden. Darunter auch hPcf11 und Komponenten des tRNA *Splicing*-Komplexes. Die Interaktion des CF II mit weiteren Komplexen der 3' Prozessierung scheint eher transient zu sein. In einem Reinigungsversuch mit RNA-basierten Protein-*Pulldown* aus HeLa-Extrakt und Analyse gereinigter Proteine über Massenspektrometrie (Shi 2009) waren die Komponenten Pcf11 und Clp1 im Gegensatz zu den anderen Zentralen Faktoren CPSF, CstF und CF I unterrepräsentiert.

1.4.4.1 hClp1

Das humane Clp1 Protein ist die ausführlichere untersuchte Komponente des CF II-Komplexes. Das Protein ist hoch konserviert in Eukaryoten (s. Anhang Abbildung 9.5). Es besitzt eine zentrale 5'-Polynukleotidkinase (PNK) Domäne (Abbildung 1.4 A). Diese wird flankiert von den N- und C-terminalen Bereichen, welche in Hefe sowohl an der ATP-Bindung als auch an Protein-Protein Wechselwirkung beteiligt sind (Gordon u. a., 2011; Gross und Moore, 2001; Holbein u. a., 2011). Das hClp1 besitzt eine Spezifität für einzelsträngige RNA-Substrate und ist in der Lage, diese am 5'-OH Ende zu phosphorylieren (Weitzer und Martinez, 2007). Das *C. elegans*-Clp1 ist mit einem GC-Dinukleotid im aktiven Zentrum kristallisiert worden (Dikfidan u. a., 2014). In derselben Arbeit wurden auch die kinetischen Parameter des ceClp1 erfasst. So wurden die Parameter k_{cat} mit $2,9 \text{ s}^{-1}$, K_M für ATP mit $\sim 100 \text{ nM}$ und K_M für das GC-Dinukleotid mit $\sim 109 \text{ }\mu\text{M}$ ermittelt. Die

enzymatische Aktivität ist für die meisten Spezies vermutet und für einige gezeigt. Die kinetischen Parameter sind jedoch meist unbekannt. Das Hefe-Homolog besitzt im Gegensatz zu den anderen Homologen diese 5'-PNK-Aktivität aufgrund einiger Aminosäureunterschiede im aktiven Zentrum nicht (s. Anhang Abbildung 9.5, rote Pfeile).

Eine Deletion des hClp1 ist embryolethal (Karaca u. a., 2014; Weitzer u. a., 2015). Zudem gibt es beschriebene neurodegenerative Krankheiten, welche auf einer verringerten katalytischen Funktion des hClp1 beruhen (Hanada u. a., 2013; Karaca u. a., 2014; Weitzer u. a., 2015). Möglicherweise führt dort eine beobachtete Akkumulation von tRNA Fragmenten der Tyrosin Vorläufer-tRNA und daraus resultierende Sensitivität gegenüber oxidativem Stress zu dem Krankheitsbild.

Für das humane Clp1 wurde eine Assoziation mit dem humanen tRNA *Splicing*-Endonuklease-Komplex beschrieben (Paushkin u. a., 2004). Dabei ist bis heute ungeklärt, welche Funktion hClp1 innerhalb des Komplexes genau hat. Die klassische Prozessierung der tRNA-Introns durch den tRNA-Endonuklease TSEN-Komplex hinterlässt am 3'-Ende des 5'-Exons ein 2'-3'-zyklisches Phosphat. Dieses kann in der Ligation der tRNA-Exons durch HSPC117 verwendet werden. Eine 5'-PNK ist also theoretisch nicht notwendig. Die Eigenschaft 5'-OH Enden des 3'-Exons zu phosphorylieren wurde jedoch durch Weitzer und Kollegen (2007) hier als notwendiger Schritt der tRNA Reifung vorgeschlagen. So könnte auch in Säugern ein Mechanismus analog zum Hefe-Weg existieren (Lyons u. a., 2018). In Hefe wird der tRNA-Vorläufer in Hefe zunächst durch den Sen-Komplex gespalten. Dieser Schritt hinterlässt ein 2'-3'-zyklisches Phosphat am 3'-Ende, welches im Anschluss durch Trl1 zu 2'-Phosphat aufgespalten wird. Das 5'-Ende wird unter Verwendung von GTP durch Trl1 phosphoryliert, danach durch Trl1 mittels ATP adenyliert. Das so aktivierte 5'-Ende wird dann an das 3'-OH durch Trl1 ligiert. Abschließend wird das übrig gebliebene 2'-Phosphat durch Tpt1 auf NAD übertragen. Für keines der beteiligten Proteine nach der tRNA-Spaltung durch den Sen-Endonuklease-Komplex, außer für die vorgeschlagene Funktion des hClp1 als Kinase, sind im Menschen bislang Homologe beschrieben.

Das humane Pcf11 ist im tRNA-*Splicing*-Komplex nicht zu finden. Es scheint also parallel einen Komplex in der 3'-Prozessierung zu geben, in dem das hClp1 in Assoziation mit hPcf11 vorliegt, und ein Komplex von hClp1 in Assoziation mit dem tRNA-*Splicing*-Komplex.

In Hefe sind ausschließlich Interaktionen von yClp1 mit anderen 3'-Prozessierungsfaktoren bekannt. So interagiert yClp1 in einer Bindetasche zwischen dessen zentraler und C-terminaler Domäne mit den Aminosäuren 475–499 des yPcf11 (Abbildung 1.4 B, Noble u. a., 2007). Des Weiteren wurde auch eine Interaktion von yClp1 mit Ysh1, dem CPSF3-Homologen aus Hefe, beschrieben (Kyburz u. a., 2003).

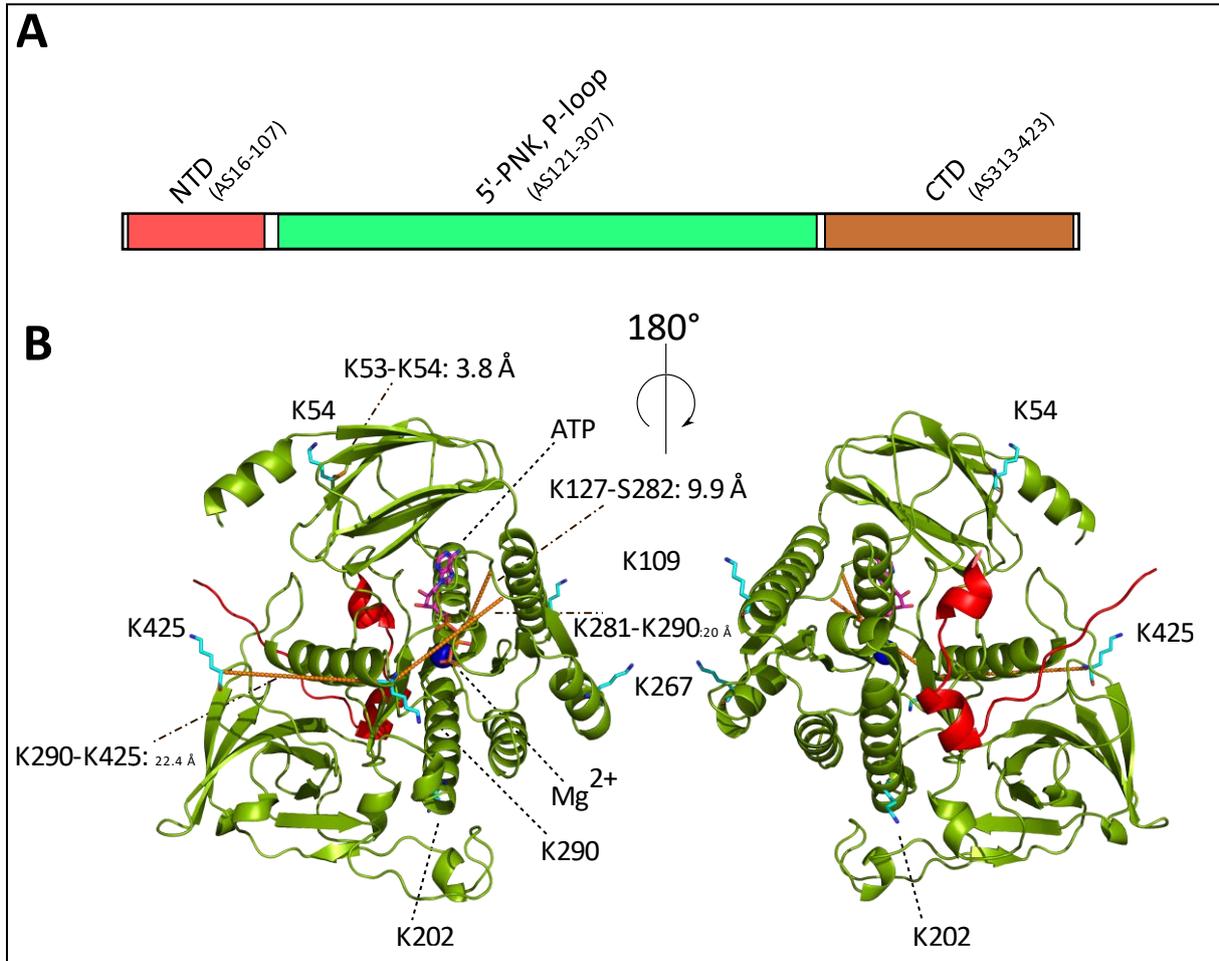


Abbildung 1.4 Domänen und Interaktionen des humanen Clp1.

A) Domänenstruktur des hClp1 mit Aminosäureangaben; NTD: N-terminale Domäne, 5'-PNK: 5'-Polynukleotidkinase Domäne mit ATP- und Protein- bindenden P-Loop Domäne, CTD: C-terminale Domäne. **B)** Sequenzhomologie-Modell der hClp1-Struktur (grün, berechnet mit Robetta Server) mit einem hinein modellierten Peptid aus yPcf11 (rot, PDB Nummer: 4O14). Vernetzbare Lysine sind gekennzeichnet durch eine gestrichelte Linie (Erstellt durch C. Tüting, entnommen aus Schäfer u. a., 2018)

1.4.4.2 hPcf11

Das humane Pcf11 ist im Vergleich zu seinem Hefe-Homolog wenig charakterisiert. RNA-Interferenz-Versuche gegen hPcf11 in HeLa-Zellen zeigten schwere Defizite in der 3'-Prozessierung von mRNAs und der Transkriptionstermination verglichen mit entsprechenden Kontrollzellen (West und Proudfoot, 2008).

Mit seinen ~172 kDa ist das hPcf11 im Vergleich zum ~50 kDa großen yPcf11 fast 3,5-mal größer. Darüber hinaus sind nur wenige Bereiche der beiden Proteine in ihrer Aminosäuresequenz konserviert (s. Anhang Abbildung 9.6). Wie in Abbildung 1.5 zu sehen ist, besitzen beide Varianten eine konservierte, N-terminal gelegene, (RNA Polymerase II) CTD-Interaktionsdomäne (CID) (Barillà u. a., 2001; Kecman u. a., 2018; Meinhart und Cramer, 2004). Die Interaktion von Pcf11 CID und RNA Pol II CTD führt *in vitro* zu einem Zerfall des Elongationskomplexes und der

Termination der Transkription (Zhang, 2005; Zhang und Gilmour, 2006). Des Weiteren hat die CID eine schwache RNA-bindende Aktivität, welche mit der CTD-Interaktion konkurriert und so auch zur Transkriptionstermination beitragen soll (Hollingworth u. a., 2006).

Neben dieser N-terminalen CID existiert im C-terminalen Bereich des Proteins ein konservierter Abschnitt (Abbildung 1.5). Dieser umfasst den Interaktionsbereich mit hClp1 (Noble u. a., 2007), welcher auf beiden Seiten von, ebenfalls konservierten, Zinkfingern flankiert wird (Guéguénat u. a., 2017; Yang u. a., 2017). Das yPcf11 hat zudem noch beschriebene Interaktionen mit den beiden CstF-Homologen Rna14 und -15 aus Hefe. Ob diese Interaktionen auch für die humanen Proteine existieren, war zu Beginn der Arbeit unbekannt.

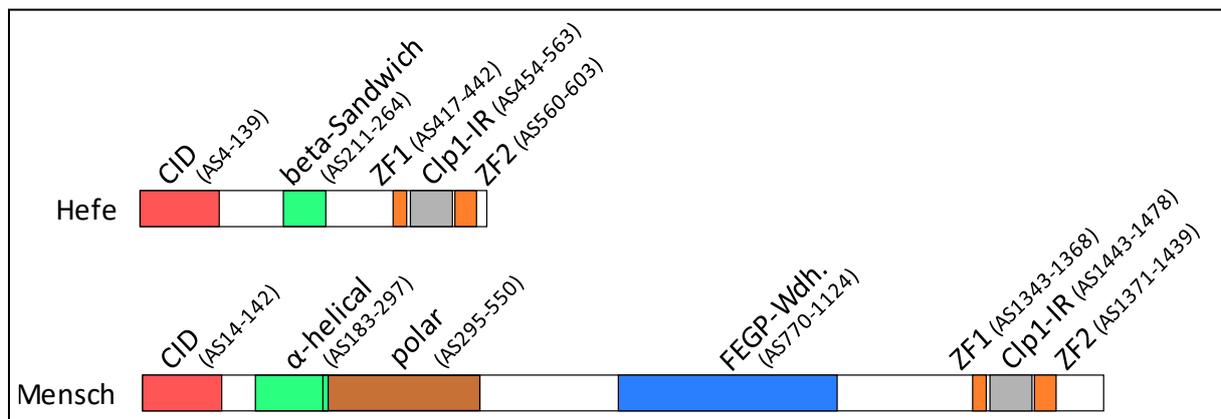


Abbildung 1.5 Domänenstruktur mit Angabe der Aminosäurepositionen des Bäckerhefe-Pcf11 im Vergleich mit dem humanen Pcf11.

N-terminal begonnen beinhaltet die Sequenz des yPcf11 folgende Strukturen und Bereiche: CID: (RNA-Pol II)-CTD-Interaktionsdomäne, Beta-Sandwich Domäne: interagiert mit Rna14 und -15 in CFIA, ZF: (RNA-bindende) Zinkfinger Domäne, Clp1-IR: Clp1-interagierende Region. Das hPcf11 beinhaltet neben den konservierten Bereichen (gleiche Namensgebung wie in yPcf11) auch folgende Bereiche: alpha-helikale und polare Region: Bereich ohne bekannte Funktion, FEGP-Wdh.: Sequenzwiederholungen der Aminosäuren FEGP mit unbekannter Funktion.

1.4.5 RNA-Polymerase II

Die RNA-Polymerase II, genauer gesagt die C-terminale Domäne (CTD) ist ein beschriebener Faktor der 3'-prä-mRNA Prozessierung (Hirose und Manley, 1998). Die Polymerase synthetisiert zunächst in ihrer Funktion den neuen RNA-Strang. Während der Synthese von prä-mRNAs bildet die CTD der DNA-abhängigen RNA-Polymerase II jedoch eine Plattform für die Assoziation vieler an der mRNA-Prozessierung beteiligter Proteine. So bringt diese Assoziation die Proteine in räumliche Nähe zum RNA-Substrat. Die RNA-Pol II CTD besteht hauptsächlich aus Aminosäurewiederholungen mit der Konsensus-Sequenz „YSPTSPS“. Die Sequenz ist weitgehend konserviert in Eukaryoten, wobei die Anzahl der Wiederholungen dieser Sequenz Speziesabhängig ist. Während der Transkription werden in diesem Sequenzmotiv an Serin2/7 durch CDK7, Serin5/Threonin4 durch CDK9 und dem Tyrosin1 durch Abl1/2 Phosphorylierungen

angebracht (Schüller u. a., 2016; Suh u. a., 2016). Der Phosphorylierungsstatus der CTD-Aminosäuresequenz „YSPTSPS“ ist dabei abhängig vom Transkriptionsstatus (Initiation, Elongation, Termination). Je nach Phosphorylierung binden bevorzugt an die CTD entsprechende Faktoren des Initiations-, Elongations- oder Terminationsschrittes (Corden, 2016; Davidson u. a., 2014; Hirose und Manley, 1998; Suh u. a., 2016). Die Phosphorylierungen können durch Phosphatasen auch wieder reversibel entfernt werden. Insbesondere die Phosphatasen PP1 α und Ssu72 sind am Übergang der Elongation in die Termination beteiligt, indem sie den Phosphorylierungsstatus der CTD verändern. PP1 α entfernt die Tyrosin1- und Serin2-Phosphorylierung, während Ssu72 das Serin5 dephosphoryliert (Dichtl u. a., 2002; Kecman u. a., 2018; Schreieck u. a., 2014). Entsprechende Homologe finden sich auch in Hefe (Tabelle 1.1).

1.4.6 Poly(A)-Polymerase und Poly(A)-Bindungsprotein

Die kanonische Poly(A)-Polymerase (PAPol α) ist ein zentraler Faktor in der 3'-Prozessierung eukaryotischer prä-mRNAs. Sie katalysiert die Synthese des Poly(A)-Schwanzes an das 3'-OH Ende, welches in der Spaltungsreaktion entsteht (Bardwell u. a., 1990; Christofori und Keller, 1989; Wahle, 1991), ist aber als stimulierender Faktor auch für die Spaltungsreaktion selbst notwendig. Allein besitzt die Poly(A)-Polymerase eine schwache Substratspezifität und bindet in Kristallstrukturen des Homologen aus Hefe die letzten drei Nukleotide der Substrat-RNA (Balbo und Bohm, 2007). Die Polymerase zeigt *in vitro* nur eine langsame, distributive Synthese des Poly(A)-Schwanzes. Die Interaktion mit weiteren Spaltungsfaktoren, wie CPSF und PABPN1, sorgt für die beobachtete PAS abhängige Substratspezifität und die nötige Affinität zur gespaltenen prä-mRNA und dem wachsenden Poly(A)-Schwanz. So wird eine effiziente und prozessive Katalyse des Poly(A)-Schwanzes erreicht (Dettwiler u. a., 2004; Kaufmann u. a., 2004; Kerwitz u. a., 2003). Die Poly(A)-Polymerase wird in der Synthese des Poly(A)-Schwanzes durch die genannte Interaktion mit dem nukleären Poly(A)-Bindungsprotein PABPN1 in der Prozessivität stimuliert. PABPN1 bindet an den wachsenden Poly(A)-Schwanz und bedeckt diesen auf einer Länge von etwa 30 Nukleotiden (Kühn u. a., 2003). Gleichzeitig wird durch diese Interaktion in Säugern auch die Länge des entstehenden Poly(A)-Schwanzes kontrolliert (Bienroth u. a., 1993; Kühn u. a., 2009, 2017).

1.5 Zielstellung

Das Ziel dieser Arbeit war einen Beitrag zur vollständigen Rekonstitution der 3'-Prozessierungsreaktion von prä-mRNA aus dem Säugersystem zu leisten. Diese Reaktion wird durch eine Reihe von Proteinkomplexen gesteuert und durch den zentralen Faktor CPSF katalysiert. Die gesamte 3'-Prozessierungsreaktion konnte aber bislang *in vitro* nicht für das Säugersystem rekonstituiert werden.

So wird im ersten Teil die rekonstituierte Teilreaktion der Polyadenylierung mit Säugerproteinen betrachtet (Schönemann u. a., 2014). Die an der RNA-Bindung beteiligten Komponenten des mPSF Subkomplexes waren noch nicht zweifelsfrei bekannt. In der Literatur waren die Komponenten CPSF1 und 4 als mögliche Kandidaten diskutiert. So sollten die Proteine des mPSF hinsichtlich ihrer RNA-Bindung, Affinität und Spezifität charakterisiert werden.

Wie zuvor beschrieben, sind an der Spaltungsreaktion neben mPSF und Poly(A)-Polymerase noch viele weitere Proteine beteiligt. Die zentrale Endonuklease CPSF3 in Assoziation mit CPSF2 katalysiert dabei die eigentliche Spaltreaktion. CPSF3 zeigt jedoch allein betrachtet eine geringe Spaltaktivität, welche durch stimulierende Komplexe wie CstF, CF I und CF II verstärkt oder möglicherweise aktiviert wird. Nach Rekonstitution des Minimalkomplexes für die Polyadenylierung durch Schönemann et al. 2014 und dessen Charakterisierung der RNA-Bindung im ersten Teil, sollte im Weiteren der Fokus auf der Spaltreaktion liegen.

Insbesondere der CF II-Komplex war noch nicht hinsichtlich der Untereinheiten und Stöchiometrie klar definiert. Die Komponenten der HeLa-NXT Fraktion CF II, wie hClp1 und hPcf11, wurden durch de Vries et al. beschrieben (de Vries, 2000). Unklar war, welche der beschriebenen Proteine ausreichen, um die Aktivität des CF II zu rekonstituieren. Die putativen Untereinheiten hClp1 und hPcf11 waren aufgrund ihrer Homologen aus Hefe wahrscheinliche Komponenten des CF II. Die humanen Proteine waren jedoch nicht biochemisch charakterisiert. Den CF II-Komplex zu definieren und die rekombinanten Proteine hinsichtlich ihrer Funktion und Aktivität in der 3'-Spaltungsreaktion zu testen, stellten ein zweites zentrales Ziel dieser Arbeit dar. Dies sollte weiter dazu beitragen, die 3'-Spaltreaktion zu definieren, sowie die gesamte Reaktion *in vitro* aus rekombinanten Faktoren nachzubilden.

2 Ergebnisse

2.1 mPSF bindet das Polyadenylierungssignal und ist ausreichend für eine spezifische Polyadenylierung

Während der Arbeiten zur Rekonstitution der Polyadenylierungsaktivität des CPSF-Komplexes wurden im Zuge der Dissertation von L. Schönemann verschiedene CPSF-Subkomplexe generiert und auf Aktivität getestet. Dafür wurde die genutzten Proteine mit baculoviraler Expression in Sf21-Zellen (Zelllinie aus dem Ovar von *Spodoptera frugiperda*) generiert. Mit dem MultiBac System (Berger u. a., 2010; Bieniossek u. a., 2008; Fitzgerald u. a., 2006) wurden die entsprechenden Leserahmen in baculovirale Genome inseriert, die Proteine viral exprimiert und anschließend über Affinitäts-Fusionspeptide gereinigt. Die Expression in Insektenzellen ermöglicht hierbei eine ähnliche posttranslationale Modifikation der Proteine im Vergleich zu menschlichen Zellen. Mit dieser Methode wurden drei CPSF Komplexe mit vier (CPSF1–4), mit fünf (CPSF1–4, hFip1) und mit sechs (CPSF1–4, hFip1, WDR33) Untereinheiten durch L. Schönemann exprimiert und über einen FLAG-Fusionspeptid an CPSF1 gereinigt (Abbildung 2.1 A). Die verwendete Nomenklatur 4k, 5k und 6k entspricht in der Abbildung 2.1 A und B der Anzahl der Proteine im jeweiligen CPSF-Komplex.

Mit den drei oben beschriebenen Proteinkomplexen konnten nun RNA-Bindungsstudien durchgeführt werden. Hierfür wurden im Rahmen dieser Arbeit Filterbindungsversuche durchgeführt. Als Substrat wurde eine kurze RNA gewählt welches das spezifische Polyadenylierungssignal (PAS) „AAUAAA“ enthält. Funktionales CPSF sollte dieses RNA-Substrat, genannt Wigley10 (W10 AAUAAACCCA) (Wigley et al. 1990), spezifisch und affin binden. Als Spezifitätskontrolle diente eine RNA mit einem U zu G Basenaustausch, genannt Wigley10 Δ (W10 Δ AAGAAACCCA). Der Basenaustausch führt zu einem starken Abfall der Affinität des CPSF zur angebotenen RNA. Das radioaktiv markierte RNA-Substrat wurde in einer Standard-Bindungsreaktion dem jeweiligen Komplex (4k, 5k oder 6k) als Bindungspartner angeboten. Nach Inkubation wurde die Reaktion über eine positiv geladenen Nitrozellulose-Membran filtriert. Dabei bleiben Proteine und die daran gebundene RNA an der Membran haften, freie RNA wird hingegen nicht gebunden. Im Anschluss kann die Menge der gebundenen RNA über Szintillationsmessung quantifiziert werden. In den Bindungsversuchen mit den drei Kombinationen 4k (CPSF1–4), 5k (CPSF1–4, hFip1) und 6k (CPSF1–4, hFip1, WDR33) zeigten 4k und 5k keine detektierbare Bindung an das angebotene RNA-Substrat W10. Nur der verwendete Komplex aus 6 Proteinen (6k) zeigte eine spezifische und affine Bindung an die W10-RNA, jedoch keine Bindung an die Spezifitätskontrolle W10 Δ . Somit zeigte sich, dass aktives CPSF neben den bekannten kanonischen Untereinheiten CPSF1 bis 4 (4k) auch WDR33 beinhalten musste, um RNA zu binden (Abbildung 2.1 B). Das humane Fip1 konnte keine Bindung an das RNA-Substrat

vermitteln. Da es in der Kombination aller sechs Untereinheiten enthalten war, welche die spezifische Bindung des W10-Substrates erlaubte, wurde es auch als Teil des CPSF-Komplexes angenommen. Die apparente Affinität des Komplexes aus den sechs CPSF-Untereinheiten (6k) für die W10 RNA lag bei etwa 20 nM. Die RNA-Kontrolle W10 Δ hatte mindestens eine um den Faktor 20 erhöhte Affinität und lag unter der Detektionsgrenze.

Der CPSF-Komplex aus 6 Untereinheiten (6k) wurde im Anschluss auf Funktionalität im Vergleich mit den anderen beiden Komplexen 4k und 5k in der Polyadenylierung getestet. Hierfür wurde ein Aktivitätstest mit einer enzymatisch synthetisierten „vorgespaltene“ RNA (L3pre) (Christofori und Keller, 1989) durch L. Schönemann durchgeführt. Dieses Substrat ist eine enzymatisch generierte RNA, welche durch sogenannte „run-off-transcription“ hergestellt wird. Es handelt sich um ein Fragment der L3prä-RNA, welches das PAS („AAUAAA“) enthält und nach der Spaltstelle abbricht. Dadurch wird ein Ende der RNA nach der nativen Spaltstelle des L3-RNA generiert und die Spaltreaktion im Vorfeld der Polyadenylierungsreaktion umgangen. Der Assay mit den drei Proteinkomplexen 4k, 5k und 6k zeigte auch hier eine PAS-abhängige Stimulation der Reaktion lediglich für die Kombination aller sechs Proteine (Schönemann 2014). Erst in Anwesenheit von WDR33 war Stimulierung der PAP sichtbar. Somit wurde gezeigt, dass WDR33 eine essenzielle Funktion in der RNA-Bindung des CPSF-Komplexes übernehmen muss und dadurch die Aktivität des Komplexes gewährleistet.

In vitro Rekonstitutionen der Polyadenylierungsreaktion (L. Schönemann) mit Subkomplexen aus dimeren Kombinationen des CPSF-Komplexes aus sechs Untereinheiten (6k) zeigten, dass nur vier Proteine für eine PAS-abhängige Aktivität hinreichend sind: CPSF1, CPSF4, WDR33 und hFip1 (Abbildung 2.1 C). Die vier Proteine konnten als Kernkomplex der spezifischen Polyadenylierung identifiziert werden. Dieser CPSF-Subkomplex aus vier Untereinheiten wird im Weiteren als *mammalian polyadenylation specificity factor* (mPSF) bezeichnet. Fehlt nur eines der vier Proteine in der Reaktion, konnte keine Stimulation der Polyadenylierung mehr beobachtet werden (Schönemann u. a., 2014). In der Filterbindung bindet der mPSF-Komplex mit einer Dissoziationskonstanten von 1,5 nM spezifisch die angebotene W10 RNA, welche die PAS-Sequenz beinhaltet. Die Kontrolle mit dem U zu G Austausch im PAS (W10 Δ) wurde hingegen nicht gebunden (Abbildung 2.1 D). Warum der Komplex eine etwa 10-fach niedrigere Bindungskonstante aufwies als der aus sechs Proteinen bestehende 6k-CPSF-Komplex mit 20 nM ist unklar (Vergleich Abbildung 2.1 B). Dabei sind jedoch auch die Unsicherheiten der in-Gel-Quantifizierung und des möglichen Anteils an inaktivem Protein zu berücksichtigen.

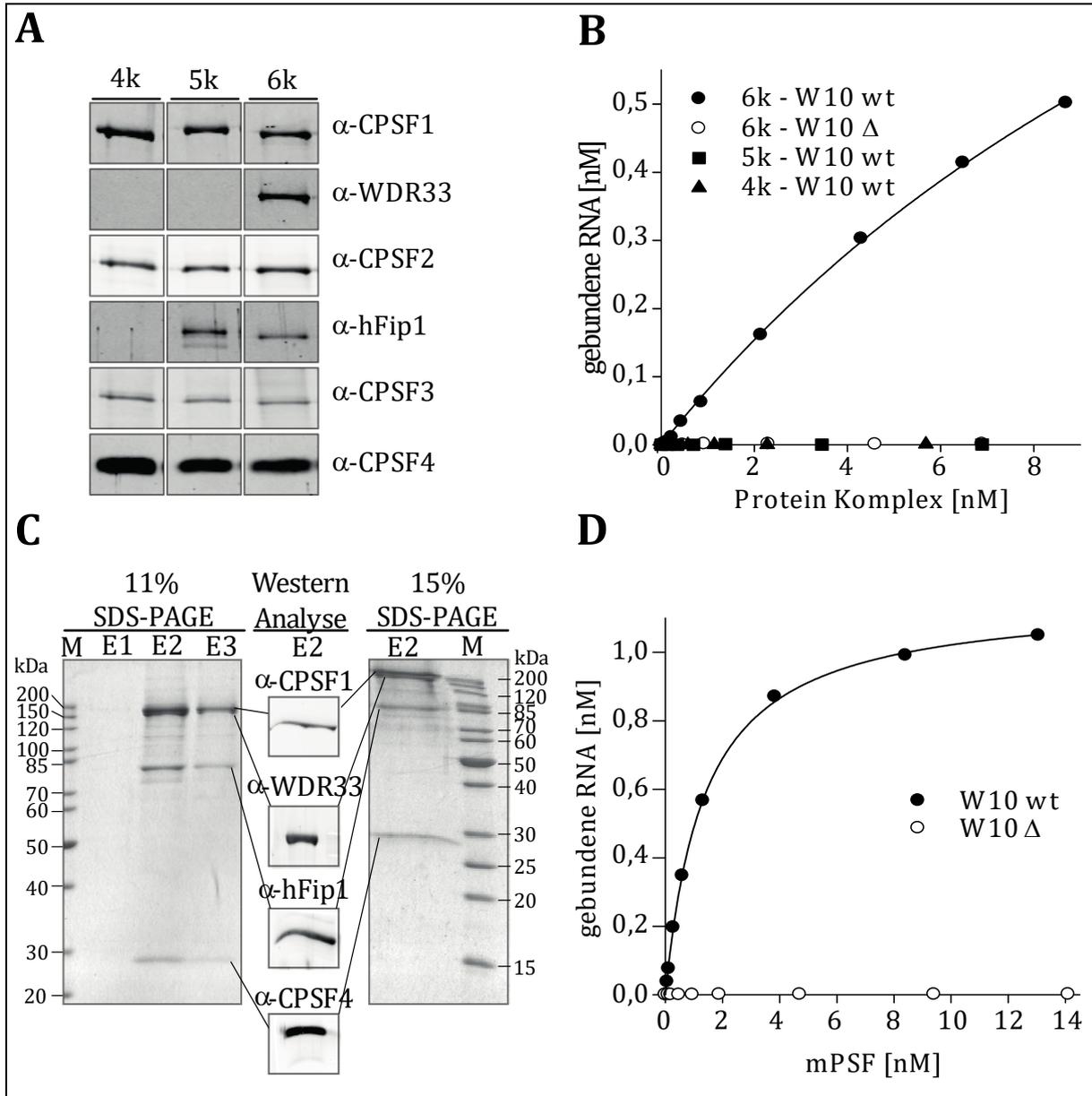


Abbildung 2.1 Expression und Filterbindungsexperimente von CPSF-Variationen modifiziert aus Schönemann u.a. 2014.

A) CPSF-Subkomplex (4k: CPSF1-4; 5k: CPSF1-4+hFip1; 6k: CPSF1-4 +hFip1+WDR33) Western-Blot-Analyse durch L. Schönemann. Die Elutionsfraktionen der jeweiligen Flag-Affinitätschromatographie nach Sf21-Koexpressionen. Verwendete primäre Antikörper sind links angegeben. Alle Signale entstammen einem Blot.

B) CPSF-Subkomplex Bindung der Proteinpräparationen 4k, 5k und 6k aus A) in Nitrozellulose-Filterbindungsassays. Jweils 1,5 nM W10-RNA (wt „AAUAAACCCA“ oder Δ -Variante „AAGAAACCCA“) wie in Material und Methoden beschrieben. **C)** mPSF FLAG-Reinigung durch L. Schönemann. Coomassie-gefärbte 11% bzw. 15% SDS-PA-Gele der FLAG-Elutionsfraktionen und in der Mitte eine Western-Blot-Analyse der Elutionsfraktion 2 (E2). Verwendete primäre Antikörper sind über den jeweiligen Western-Signalen angegeben, wobei alle gezeigten Signale von einem Blot stammen. **D)** mPSF (CPSF1, WDR33, hFip1, CPSF4)-Bindung an Nitrozellulose Filter. Proteinpräparation E2 aus C) in Bindungsreaktionen an 1,5 nM W10-RNA (wt oder Δ -Mutante) wie in Material und Methoden beschrieben.

Die intern radioaktiv markierte L3pre-RNA aus den Polyadenylierungsversuchen zeigte in einer nativen Gelelektrophorese (EMSA) eine PAS spezifische Bindung des mPSF-Komplexes an die angebotene RNA (Abbildung 2.2). Ein Vergleich der mPSF Präparation mit der CPSF-Fraktion IV, gereinigt aus Kalbsthymus (CPSF IV, Wahle 1995), zeigte hierbei ein etwas langsamer migrierendes Signal als für CPSF IV. Die untersuchte mPSF-Präparation migrierte in einer definierten Bande im EMSA. Die aus der Menge gebundener RNA errechnete Dissoziationskonstante lag auch hier im Bereich um 2 nM und war somit in Übereinstimmung mit den Daten aus der Filterbindung.

Die Funktionalität des mPSF-Komplexes in der Stimulation der Poly(A)-Polymerase (L. Schönemann, s. Anhang Abbildung 9.13) konnte ebenfalls gezeigt werden. In Verbindung mit PABPN1 zeigte der mPSF-Komplex eine zum CPSF IV vergleichbare Steigerung der PAP-Prozessivität. Auch die Länge des angefügten Poly(A)-Schwanzes an die Substrat-RNA war mit einer Länge von etwa 250–300 Nukleotiden vergleichbar mit früheren Daten, welche mit CPSF Fraktionen aus Kalbsthymus und PABPN1 gewonnen wurden (Kühn 2009).

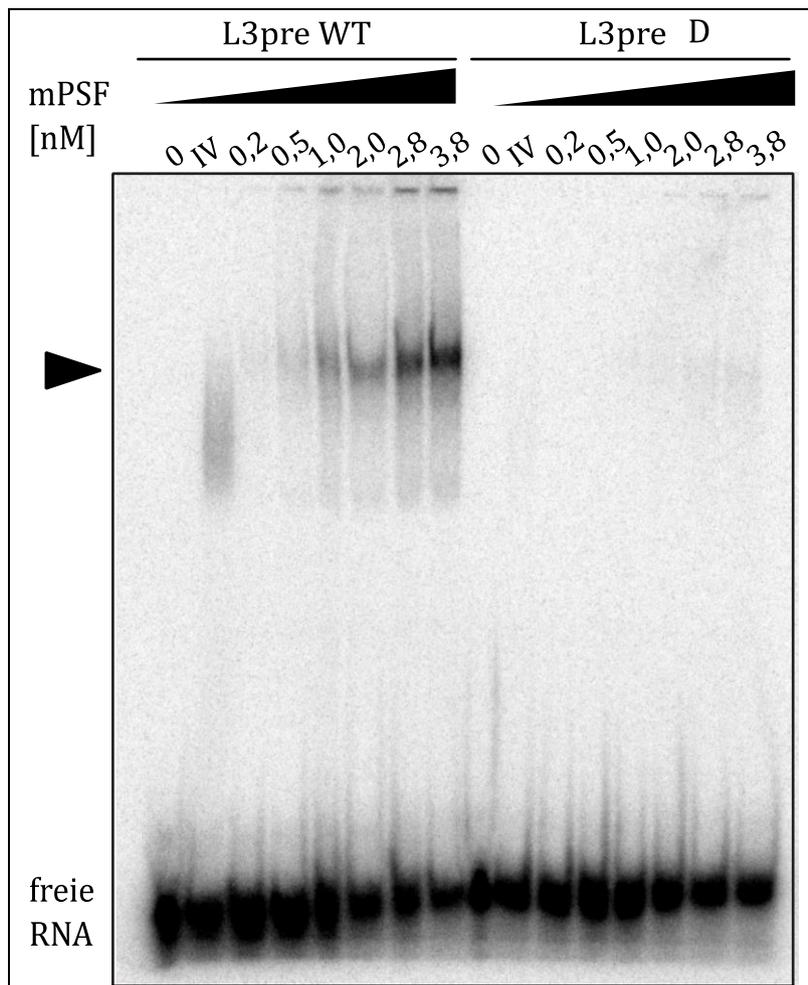


Abbildung 2.2 mPSF-EMSA mit Flag-Elutionsfraktion E2. Steigende Konzentrationen der mPSF-Präparation gezeigt in Abb. 2.1 C wurden eingesetzt wie über der jeweiligen Spur angegeben. In den Bindungsreaktionen wurden 0,5 nM L3pre-RNA oder L3pre Δ -RNA angeboten und im EMSA wie beschrieben in Material und Methoden aufgetrennt. Die Spur mit Namen „IV“ enthielt in der Bindungsreaktion CPSF-Fraktion IV, gereinigt aus Kalbsthymus (Bienroth u. a., 1991) als Kontrolle. Graphik wurde entnommen und modifiziert aus Schönemann u. a. 2014.

2.2 WDR33 und CPSF4 (30k) binden gemeinsam das Polyadenylierungssignal

Der mPSF-Komplex mit seinen vier Untereinheiten CPSF1, CPSF4, hFip1 und WDR33 bildet den Minimalkomplex des CPSF-Komplexes, welcher ausreichend ist, um zusammen mit PAP und PABPN1 die Polyadenylierung zu katalysieren. Zudem bindet mPSF spezifisch das PAS. Die Proteine, welche diese RNA-Bindung vermitteln waren jedoch nicht zweifelsfrei bekannt. Bisher war aufgrund von UV-Vernetzung von CPSF-Proteinfractionen an RNA angenommen worden, dass sowohl CPSF1 als auch CPSF4 gemeinsam die PAS binden und die Spezifität für die kanonische Sequenz AAUAAA vermitteln, da Signale bei 160 kDa und 30 kDa detektiert wurden (Jenny u. a., 1994; Keller u. a., 1991; Murthy und Manley, 1995). Das WDR33 war bisher jedoch ein unbekannter Faktor und migriert in der SDS-PAGE auf fast derselben Höhe wie das 160 kDa große CPSF1. Die hier gemachten RNA-Bindungsstudien mit CPSF-Subkomplexen und mPSF sowie die Aktivitätsassays mit den verschiedenen Subkomplexen ließen darauf schließen, dass auch WDR33 direkt an der Bindung des Signals beteiligt sein könnte. Ausgehend davon wurde auch mit mPSF eine UV-Vernetzung mit einem RNA-Oligo durchgeführt. Das verwendete Wigley10 RNA-Oligonukleotid war zum einen radioaktiv 5'-³²P markiert. Zum zweiten erfolgte der Austausch des Uracil in der Sequenz (5'-AAUAAACCCA-3') zu 5-Jodo-Uracil (benannt iW10). Nach einer Bindungsreaktion wurde das RNA-Oligo dann bei 312 nm UV-Licht an die Proteine des mPSF vernetzt. In diesem Verfahren werden nur Aminosäuren mit direktem Kontakt zur RNA vernetzt. Im Ergebnis zeigten sich in der SDS-PAGE starke Signale bei einem Molekulargewicht von 30 kDa und 170 kDa (Abbildung 2.3 A, Load). Eine Kontrolle mit unmarkiertem, radioaktivem W10-Oligonukleotid zeigte die gleichen Resultate. Diese zeigten allerdings eine geringere Signalstärke. Daraus lässt sich schließen, dass ein Großteil des detektierten Signals nach UV-Vernetzung durch die Vernetzung mit dem 5-Jodo-Uracil in dem PAS zustande kommen muss. In einer angeschlossenen Reinigung mit einer Ni-NTA-Matrix in Gegenwart von 8 M Harnstoff konnte aus der UV-Vernetzungsreaktion das radioaktiv markierte His-WDR33 angereichert werden (Abbildung 2.3 A, Elution E). Somit bindet das WDR33 direkt an das Polyadenylierungssignal. In Schönemann et al. 2014 konnte zudem durch PAR-CLIP gezeigt werden, dass WDR33 mit erhöhter Häufigkeit direkt an oder in direkter Umgebung des Polyadenylierungssignals bindet.

Eine Affinitätsreinigung aus UV-vernetztem mPSF mit polyklonalem Antikörper gegen CPSF4 zeigte eine spezifische Anreicherung des vernetzten iW10 Oligonukleotids bei einem Molekulargewicht von 30 kDa (Abbildung 2.3 B). Die Reinigung nach der UV-Vernetzung erfolgte dabei unter leicht denaturierenden Bedingungen bei 2 M Harnstoff. Dies bestätigte die zuvor gezeigte direkte Bindung des CPSF4 an das PAS. Eine versuchte Affinitätsreinigung mit Antikörper gegen CPSF1 zeigte keine Anreicherung eines radioaktiven Signals auf der entsprechenden Höhe bei 170 kDa (nicht gezeigt). Daraus konnte abgeleitet werden, dass vermutlich ausschließlich

WDR33 und CPSF4 an der Bindung des PAS beteiligt sind. Was im Gegensatz zu den erwarteten Proteinen CPSF1 und CPSF4 stand.

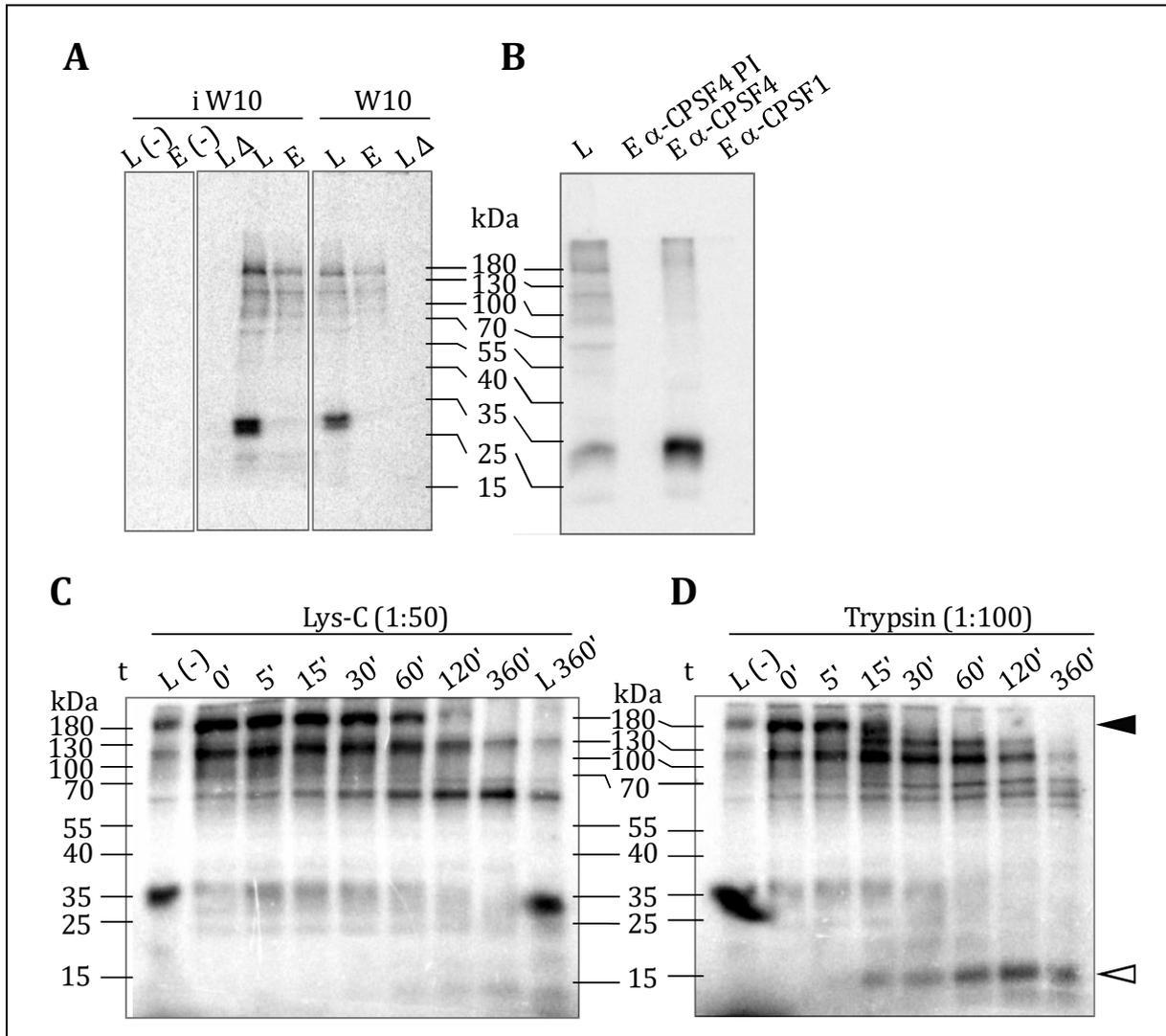


Abbildung 2.3 Crosslinking von W10-RNA an mPSF

A) UV-Vernetzung bei 312 nm von 10 nM MycHis₆-WDR33 in mPSF mit 10 nM radioaktiv markierter und 5-Jodo-Uracil modifizierter iW10-RNA (jeweils wt „AAUAAACCCA“ und Δ „ACUCAACCCA“) oder mit 10 nM nicht modifizierter W10-RNA (jeweils wt „AAUAAACCCA“ und Δ „AAGAAACCCA“). Aus den Wt RNA-Reaktionen wurde MycHis₆-WDR33 über eine Ni-NTA-Matrix unter denaturierenden Bedingungen gereinigt. Load (L) enthielt 5% der Reaktion, Eluate (E) enthielten 50% der Fraktion. (-) sind Reaktionen ohne Protein zur Kontrolle. Die gezeigten Gelfragmente sind zusammen geschnitten aus einem Gel. **B)** UV-Vernetzung von 10 nM mPSF mit 10 nM wt iW10-RNA bei 312 nm analog zu A) mit anschließender Antikörper-Affinitätsreinigung mittels anti-CPSF4- und anti-CPSF1-Antikörper unter leicht denaturierenden Bedingungen. Deutlich sieht man nur eine Anreicherung eines Signals in der anti-CPSF4-Reaktion, wenn man dies im Vergleich zur Spezifitätskontrolle mit Prä-Immuneserum betrachtet. Die aufgetragenen Mengen in den Spuren waren wie zuvor in A). **C)** Tricin-SDS-Gel der UV-Vernetzung von 10 nM iW10-RNA mit 10 nM mPSF (inklusive His₆-WDR33) wie zuvor. Anschließend wurde die Reaktion mit der Protease Lys-C im Unterschuss (Masseverhältnis 1:50) bei 0,75 M Harnstoff verdaut. Die Kinetik erfolgte mit Probenentnahme zu den gekennzeichneten Zeitpunkten und abschließender denaturierender Reinigung über eine Ni-NTA Matrix. Load (L) ist 10% der Reaktion ohne Affinitätsreinigung und die Spuren der Zeitpunkte entsprechen jeweils 100% der Ni-NTA-Elution. Der schwarze Pfeilkopf markiert das Vollängen-WDR33 und der weiße Pfeilkopf das Produkt bei 15 kDa. **D)** Tricin-SDS-Gel des Trypsin-Verdau (Masseverhältnis 1:100) analog zu C).

Die Elutionsfraktion der WDR33 Ni-NTA-Reinigung zeigte auch Signale bei geringeren Molekulargewichten als 170 kDa (Abbildung 2.3 A). Dies deutete auf einen partiellen proteolytischen Abbau des WDR33 hin, in dessen Folge Fragmente entstanden, welche noch immer das iW10 Oligonukleotid banden. Daraufhin wurde das UV-vernetzte WDR33 unter leicht denaturierenden Bedingungen (0,75 M Harnstoff) der limitierten Proteolyse durch verschiedene Proteasen ausgesetzt und anschließend unter denaturierenden Bedingungen (8 M Harnstoff) über eine Ni-NTA-Matrix gereinigt. In den Abbildung 2.3 C und D sieht man das Resultat dieser Verdauung für die Proteasen Lys-C (C) und Trypsin (D). Dabei sieht man deutlich den allmählichen Abbau des Signals des vollständigen WDR33 bei 180 kDa und die Anreicherung kleinerer Signale. Hier stechen besonders die Signale bei etwa 120, 60 und 15 kDa (weißer Pfeilkopf) hervor. Betrachtet man hierzu die Domänenstruktur des WDR33 (Abbildung 2.4), lässt sich vermuten, dass es sich dabei um Fragmente geschnitten nach G/P/R-reicher Region, Kollagen ähnlicher Domäne und in WD40-Wiederholungen handeln könnte. Aus dem Vergleich der eingegebenen Fraktion (L) und der Reaktion nach 360 Minuten vor Reinigung (L 360') kann man erkennen, dass WDR33 durch die Protease angegriffen wurde. Das CPSF4-spezifische Signal bei 30 kDa blieb nahezu konstant. Das Hexahistidin-Fusionspeptid liegt am N-Terminus des WDR33. Dadurch handelt es sich bei allen durch die Ni-NTA-Agarose angereicherten Banden um C-terminal verkürzte Produkte aus der Proteolyse. In Verbindung der beobachteten Signale bei etwa 15–35 kDa mit den berechneten Molekulargewichten der WDR33-Fragmente aus Tabelle 2.1 kann man schlussfolgern, dass die RNA-Bindung des WDR33 über die 7 WD40-Wiederholungen am N-Terminus vermittelt werden muss. Wahrscheinlich sogar in den ersten beiden WD40-Wiederholungen.

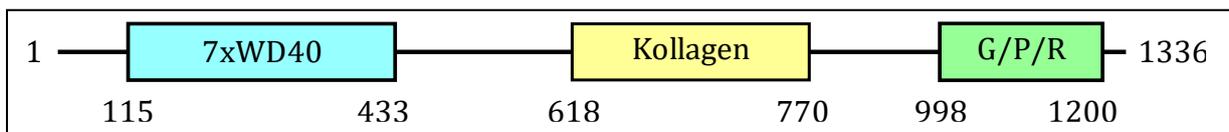


Abbildung 2.4 Schematische Darstellung der Domänen und strukturierten Bereiche des humanen WDR33 mit Angaben der Aminosäurepositionen.

N-terminal begonnen sind folgende Domänen und Bereiche gezeigt: WD40 – WD40-Wiederholungen (AS 117–412); Kollagen – Kollagen ähnliche Faltung der Region (AS 618–770); G/P/R – Glycin-, Prolin-, Arginin-reiche Region (AS 998–1200).

Tabelle 2.1 Übersicht der Domänen und strukturierten Bereiche des humanen WDR33. Domänen und Bereiche mit Angabe der Aminosäurepositionen, der Länge des Aminosäureabschnittes und des berechneten Molekulargewichtes ab dem N-Terminus des humanen WDR33 als Hilfestellung zur Einordnung der Signale der limitierten Verdauung des mPSF in Abbildung 2.3 C und D.

Domäne	AS-Position	Länge [AS]	berechnetes MW ab N-Terminus
WD 1	117–156	40	19,2 kDa
WD 2	159–198	40	24,2 kDa
WD 3	200–239	40	29,0 kDa
WD 4	242–283	42	33,7 kDa
WD 5	286–325	40	38,7 kDa
WD 6	329–369	41	43,3 kDa
WD 7	373–412	40	48,3 kDa
Kollagen-like	618–770	153	86,4 kDa
G/P/R-reich	998–1200	203	132,2 kDa

2.3 Rekonstitution des humanen *Cleavage Factor II* (CF II) aus hPcf11 und hClp1

Die humanen Proteine hPcf11 und hClp1 wurden zusammen mit anderen Proteinen als CF II aus nukleärem Extrakt gereinigt (Rüegsegger u. a., 1996; de Vries, 2000). Die Hefe-Proteine yPcf11 und yClp1 assoziieren in dem Subkomplex CF IA mit den Homologen der CstF-Untereinheiten 64k und 77k und wurden in Reinigungen mit diesen gemeinsam eluiert (Gross und Moore, 2001). Das humane Cpl1 wurde aus zytoplasmatischem Extrakt zusammen mit der tRNA-*Splicing*-Endonuklease gereinigt (Weitzer und Martinez, 2007). Bislang ist unklar, aus welchen Proteinen das humane CF II besteht. So wurde im Vorfeld der Arbeit bereits durch Anne Knoth, im Rahmen ihrer Promotion, eine Reinigung aus stabil transformierten HeLa-Zellen durchgeführt (Dissertation A. Knoth). Hierbei wurde ein His-FLAG-hClp1 Konstrukt transformiert und das Protein durch zwei aufeinander folgende Affinitätsreinigungen gereinigt. Bei einer anschließenden massenspektrometrischen Analyse zeigte sich eine Assoziation von hClp1 mit der tRNA-*Splicing*-Endonuklease (Paushkin u. a., 2004) sowie mit hPcf11. Da die tRNA-*Splicing*-Endonuklease vermutlich kein Faktor der 3'-Prozessierung von mRNAs ist, ergab sich daraus die Annahme, dass CF II möglicherweise nur aus hClp1 und hPcf11 bestehen könnte.

Die Darstellung der Proteine in Bakterien wie *E. coli* ist bisher nur mit verkürzten Varianten gelungen. Aus diesem Grund wurden im Rahmen dieser Arbeit Proteinkomplexe aus hPcf11 und hCpl1 sowie hClp1 allein in Insektenzellen hergestellt (Berger u. a., 2004; Bieniossek u. a., 2008; Fitzgerald u. a., 2006).

2.3.1 Herstellung von CF II und hCpl1 aus Insektenzellen

Die offenen Leserahmen (ORFs) von hPcf11 und Clp1 wurden zur Verfügung gestellt von G. Martin/W. Keller (Basel). Die ORFs wurden, wie in Material und Methoden beschrieben, subkloniert und mit Affinitäts-*Tags* versehen. Abschließend erfolgte die Integration in das baculovirale Genom (Bacmid). Dieses Bacmid ist eine Art Fusion aus Plasmid und Virusgenom. Es ermöglicht die Replikation des viralen Genoms in *E. coli* und die Replikation des Virus in infizierten Sf21 Zellen. Mit den generierten Bacmiden wurden Sf21-Zellen transfiziert. So wurden Viren für CF II (Strep-hPcf11, His-hClp1) gewonnen und zur Expression der Proteine in Sf21-Zellen genutzt. Etwa 72 h nach Infektion wurde die Sf21-Zellkultur geerntet ($\sim 1,6 \times 10^9$ Zellen). Das erhaltene Zellysate wurde zunächst über eine Ni-NTA-Affinitätssäule gegeben und das CF II durch den N-terminalen His-tag an hClp1 gereinigt. Deutlich erkennt man in Abbildung 2.5 A die gemeinsame Elution der beiden Proteine in den Elutionsfraktionen der Säule. Im Anschluss daran wurden die Elutionsfraktionen E2 und E3 vereint und mit salzfreiem Puffer verdünnt. Die so verdünnte Fraktion wurde über eine MonoQ-Säule fraktioniert. Wie in der Abbildung 2.5 B zu sehen ist, eluierten die Proteine hPcf11 und hClp1 gemeinsam in Fraktionen von der Matrix. Durch die Reinigung erfolgte auch eine Abtrennung des überschüssigen hClp1 aus der IMAC in den Waschfraktionen der MonoQ-Säule. Die Elutionsfraktionen wurden hinsichtlich der 5'-Polynukleotidkinase-Aktivität von hClp1 überprüft und gegen die UV-Absorptionssignale und Proteinkonzentrationen abgeglichen. Die Abbildung 2.5 D zeigt dabei, dass CF II-Protein (UV-Signal und Coomassie-Gel) mit der detektierten 5'-PNK-Aktivität (gestrichelte Linie) übereinstimmt. In der SDS-PAGE der MonoQ-Säule sind neben dem vollständigen hPcf11 bei 180 kDa auch verschiedene weitere Banden mit geringerem Molekulargewicht zu erkennen. Diese waren bis zu einer gewissen Größe im Western-Blot als Fragmente des hPcf11 nachweisbar (Abbildung 2.5 C). Im letzten Reinigungsschritt wurde eine Größenausschlusschromatographie über eine Superose6-Säule realisiert. Dazu wurden insgesamt 1,2 ml aus den MonoQ-Fraktionen 16, 18 und 19 vereint und auf die Säule geladen. Einige Verunreinigungen und Abbaubanden konnten vom hPcf11-hClp1 Proteinkomplex abgetrennt werden (Abbildung 2.5 G). Auch hier migrierten Protein und 5'-PNK-Aktivität zusammen über die Säule (Abbildung 2.5 E, UV-Signal und gestrichelte Linie). Der Komplex zeigte in der Auftrennung eine homogene Population bei einem errechneten Molekulargewicht von 1-0,8 MDa (Abbildung 2.5 F), welche klar abgetrennt vom Ausschlussvolumen der Säule war. Dieses Molekulargewicht lag weit oberhalb der theoretischen 220 kDa für einen globulären, dimeren Komplex. Für weitere Analysen wurden die Fraktionen #9–11 vereint und über Ultrafiltration aufkonzentriert (1 μ M).

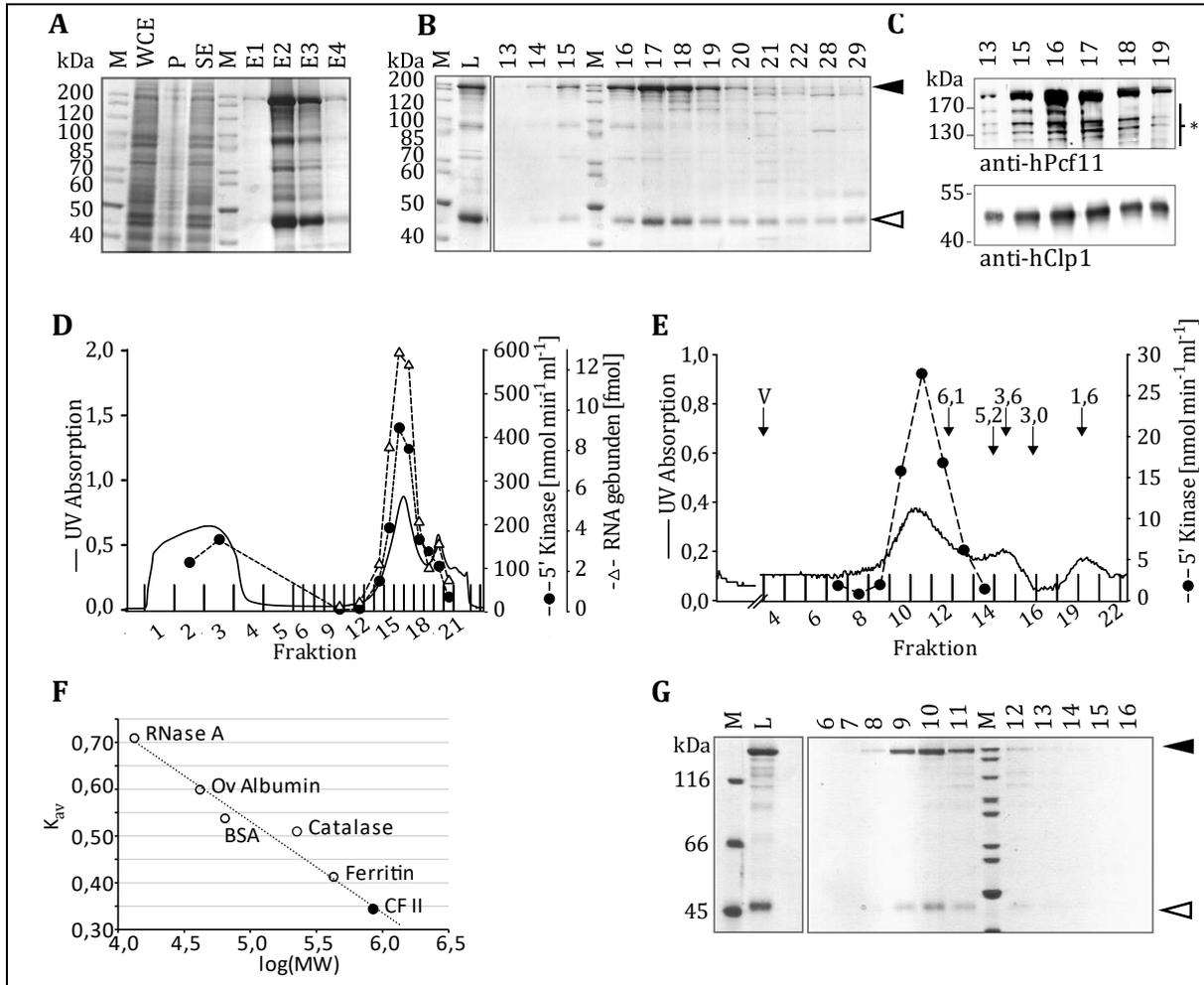


Abbildung 2.5 Reinigung von rekombinantem CF II

A) Coomassie gefärbtes SDS-PA-Gel der Ni-NTA-Affinitätsreinigung aus der CF II-Expression in SF21-Zellen. Fraktionen sind in den Spuren gekennzeichnet von links nach rechts: M – Unstained Protein Ladder, WCE – Zellaufschluss, P – Pellet nach Zentrifugation, SE – löslicher Überstand, E1–4 Ni-NTA-Elutionsfraktionen.

B) Coomassie gefärbtes SDS-PA-Gel der Reinigung der Ni-NTA E2/3-Fraktion über eine MonoQ-Säule. Elution erfolgte mit einem 0,2–1 M KCl-Gradienten über 15 ml. Die Fraktionsnummern sind über der Spur angegeben. Humanes Pcf11 bei 170 kDa (oben) markiert durch einen schwarzen Pfeilkopf und His₆-hClp1 bei 48 kDa (unten) markiert durch einen weißen Pfeilkopf.

C) Western-Blot-Analyse der angegebenen Fraktionen der MonoQ-Säule mit primären Antikörpern aus Kaninchen gegen hPcf11 (oben) und hClp1 (unten). Der Stern markiert Abbaubanden des hPCF11, welche durch den Antikörper erkannt werden.

D) UV-Chromatogramm (durchgezogene Linie) der MonoQ-Reinigung aus B) sowie enzymatische Tests über die Säulenfraktionen für die 5'-PNK-Aktivität von hClp1 (Punkte) in nmol markierte rC₁₄ Substrat-RNA pro min und pro ml sowie RNA-Bindung des CF II an SV40 late-RNA in fmol.

E) UV-Chromatogramm (durchgezogene Linie) einer 24 ml Superose6-Größenausschlusschromatographie der MonoQ-Fraktion #17 sowie enzymatische Tests über die Säulenfraktionen für die 5'-PNK Aktivität von hClp1 (Punkte). Die Stoke's Radii der eingesetzten Kalibrierproteine sind angegeben in nm und gekennzeichnet durch Pfeile.

F) Eichgerade der Marker-Proteine (leerer Kreis) aus der Größenausschlusschromatographie in der Auftragung des log₁₀(MW) gegen die K_{AV}-Werte mit linearer Regression. Der CF II-Komplex ist mit gefülltem Kreis eingetragen in die Eichgerade.

G) Coomassie gefärbtes SDS-PA-Gel der Größenausschlusschromatographie der MonoQ-Fraktion #17. Die Beschriftung und Pfeilkennzeichnung wie zuvor.

Der Virus für die Expression des His-hClp1 wurde genutzt, um eine 400 ml Kultur Sf21-Zellen zu infizieren. Etwa 72 h nach Infektion erfolgte die Ernte der Expressionskultur ($0,6 \times 10^8$ Zellen). Das so generierte hClp1 zeigte ein ähnlich gutes Expressionsverhalten wie jenes im CF II-Komplex mit hPcf11. Die Affinitätsreinigung über Ni-NTA-Agarose zeigte auch hier eine gute Anreicherung des His-hClp1 gegenüber dem Zelllysats (Abbildung 2.6 A). Nach der Ni-NTA-Affinitätsreinigung konnte allerdings keine Ionenaustauschchromatographie über MonoQ oder MonoS durchgeführt werden, da das Protein unspezifisch an die Matrix band und über den gesamten Gradienten eluierte. Stattdessen wurde die IMAC-Fraktion E2 direkt auf eine Superdex200 pg (pg - *prep grade*) Säule geladen und durch Größenausschlusschromatographie von verbliebenen Verunreinigungen weitgehend abgetrennt (Abbildung 2.6 B und C). Wie in Abbildung 2.6 B ersichtlich ist, befand sich ein geringer Teil des hClp1 im Ausschlussvolumen und in Fraktionen mit hohem Molekulargewicht. Dabei handelte es sich vermutlich um Aggregate des Proteins, da die Hauptfraktionen des hClp1 bei einem Molekulargewicht von etwa 45–40 kDa eluierten (Fraktion #16–19, Abbildung 2.6 D). Dies entspricht in etwa dem errechneten, nativen Molekulargewicht von 48,7 kDa. Auch die Kinaseaktivität des hClp1 migrierte mit den Fraktionen des gereinigten hClp1 (Abbildung 2.6 C, gestrichelte Linie). In den Fraktionen des Ausschlussvolumens (V) war hierbei nur eine geringe Aktivität zu beobachten. Aufgrund einer Phosphatase-Kontamination in der IMAC Elution E2 konnte kein 5'-Kinaseassay zur Kontrolle durchgeführt werden.

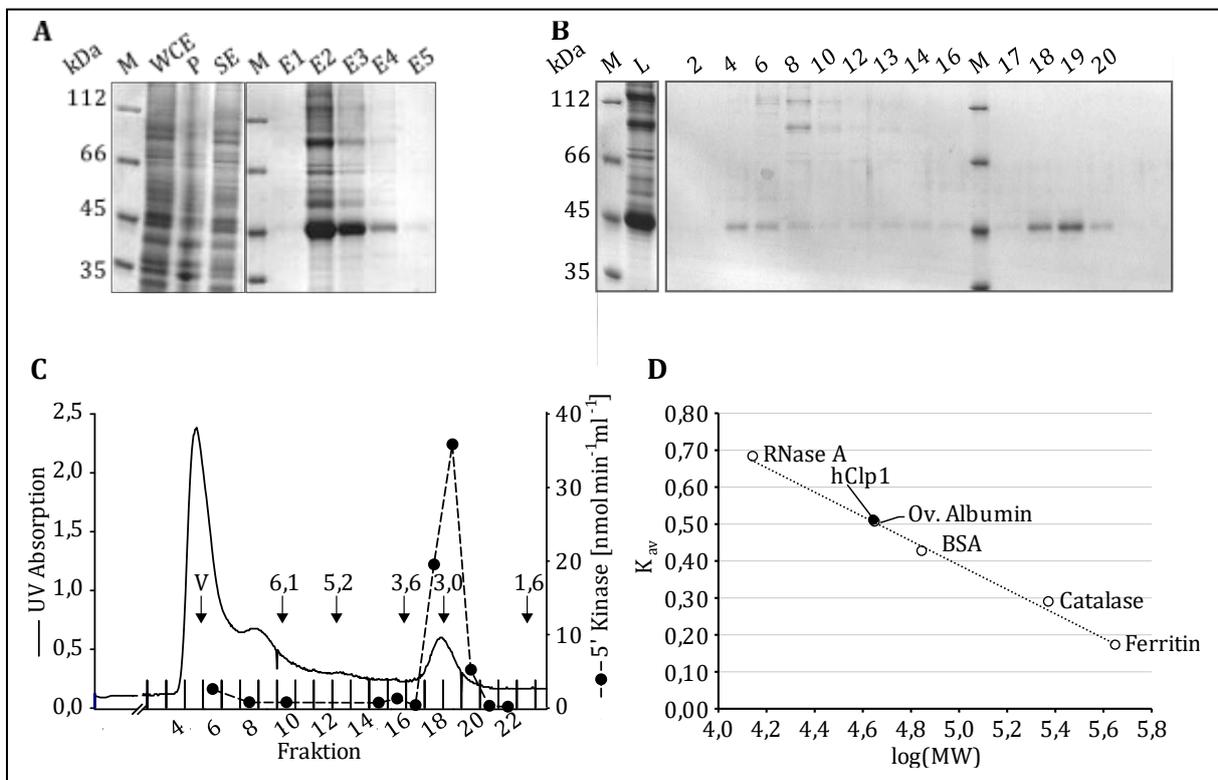


Abbildung 2.6 Reinigung von rekombinantem hClp1.

A) Coomassie gefärbtes SDS-PA-Gel der Ni-NTA-Affinitätsreinigung aus der His₆-hClp1-Expression in Sf21-Zellen. Fraktionen von links nach rechts aufgetragen mit: M – Unstained Protein Ladder, WCE – Zellaufschluss, P – Pellet nach Zentrifugation, SE – löslicher Überstand, E1–5 – Ni-NTA Elutionsfraktionen. Deutlich sichtbar ist eine Anreicherung des hClp1 bei einem Molekulargewicht von 48 kDa. **B)** Coomassie gefärbtes SDS-PA-Gel der 120 ml Superdex200 pg-Größenausschlusschromatographie der Ni-NTA-Fraktion E2. Die entsprechenden Fraktionsnummern sind über der Spur angegeben. Der Puffer war versetzt mit 0,1 mM ATP zur Stabilisierung des hClp1 während der Reinigung. **C)** UV-Chromatogramm (durchgezogene Linie) der Superdex200 pg-Größenausschlusschromatographie sowie enzymatische Tests über die Säulenfraktionen für die 5'-PNK-Aktivität von hClp1 (Punkte) in nmol markierter rC₁₄ Substrat-RNA pro min und pro ml. Die Stoke's Radii der eingesetzten Kalibrierproteine sind angegeben in nm und markiert durch Pfeile. **D)** Eichgerade der log₁₀(MW) der Marker-Proteine gegen die K_{AV}-Werte der verwendeten Marker-Proteine (leerer Kreis) aus der Größenausschlusschromatographie mit eingezeichnetem Wert für hClp1 (gefülltem Kreis).

2.3.2 CF II liegt als Dimer aus hPcf11 und hClp1 vor

Die aus dem Graphen der Abbildung 2.5 E ersichtliche Diskrepanz zwischen theoretischem (Heterodimer=220 kDa) und ermitteltem apparentem Molekulargewicht (1–0,8 MDa) des CF II warf die Frage nach der möglichen Stöchiometrie des Komplexes auf. Hierzu wurde zunächst ein Glycerin-Dichtegradient (12 ml) zwischen 40 und 20% vorbereitet und mit 100 µl (~230 µg Protein) der MonoQ-Fraktion #19 überschichtet. Dies wurde nach der Zentrifugation in 36 Fraktionen abnehmender Dichte fraktioniert. Das eluierte Protein wurde sowohl über UV-Absorption als auch über SDS-PAGE mit Silberfärbung verfolgt (Abbildung 2.7 A und B). Des Weiteren wurde die hClp1 Aktivität über einen 5'-PolynukleotidKinaseassay nachgewiesen und migrierte mit dem Proteinkomplex aus hPcf11 und hClp1 (Abbildung 2.7 A). Anhand parallel verwendeter Proteinstandards konnte somit in Abbildung 2.7 A das Sedimentationsverhalten charakterisiert werden. Der ermittelte Sedimentationskoeffizient (S) lag bei etwa 6,9 S.

Zusammen mit dem aus der Größenausschlusschromatographie (Abbildung 2.5 D) extrapolierten Stoke's Radius (R_S) von etwa 7,1 nm ließ sich über eine vereinfachte Formel (Erickson, 2009 Formel 7.1b; Siegel und Monty, 1966) das Molekulargewicht (M) ableiten.

$$M = 4205 * (S * R_S) [Da]$$

Das durch diese Formel erhaltene Molekulargewicht aus den oben genannten Messwerten ist rund 206 kDa. Beachtet man ein theoretisches Molekulargewicht von 220 kDa für ein hPcf11-hClp1 Dimer, so ergibt sich daraus eine wahrscheinliche Zusammensetzung als Heterodimer. Fasst man die Reinigungen des CF II über mehrere Säulen zusammen, lässt sich sagen, dass der Komplex aus hPcf11 und hClp1 stabil miteinander assoziiert blieb und ähnlich der Hefe-Faktoren in CF IA im Verhältnis eins zu eins assoziierte (Gordon u. a., 2011; Guéguéniat u. a., 2017). Das beobachtete hohe apparente Molekulargewicht in der Gelfiltration ist dabei wohl das Ergebnis einer ausgedehnten, nicht globulären Struktur der Proteine.

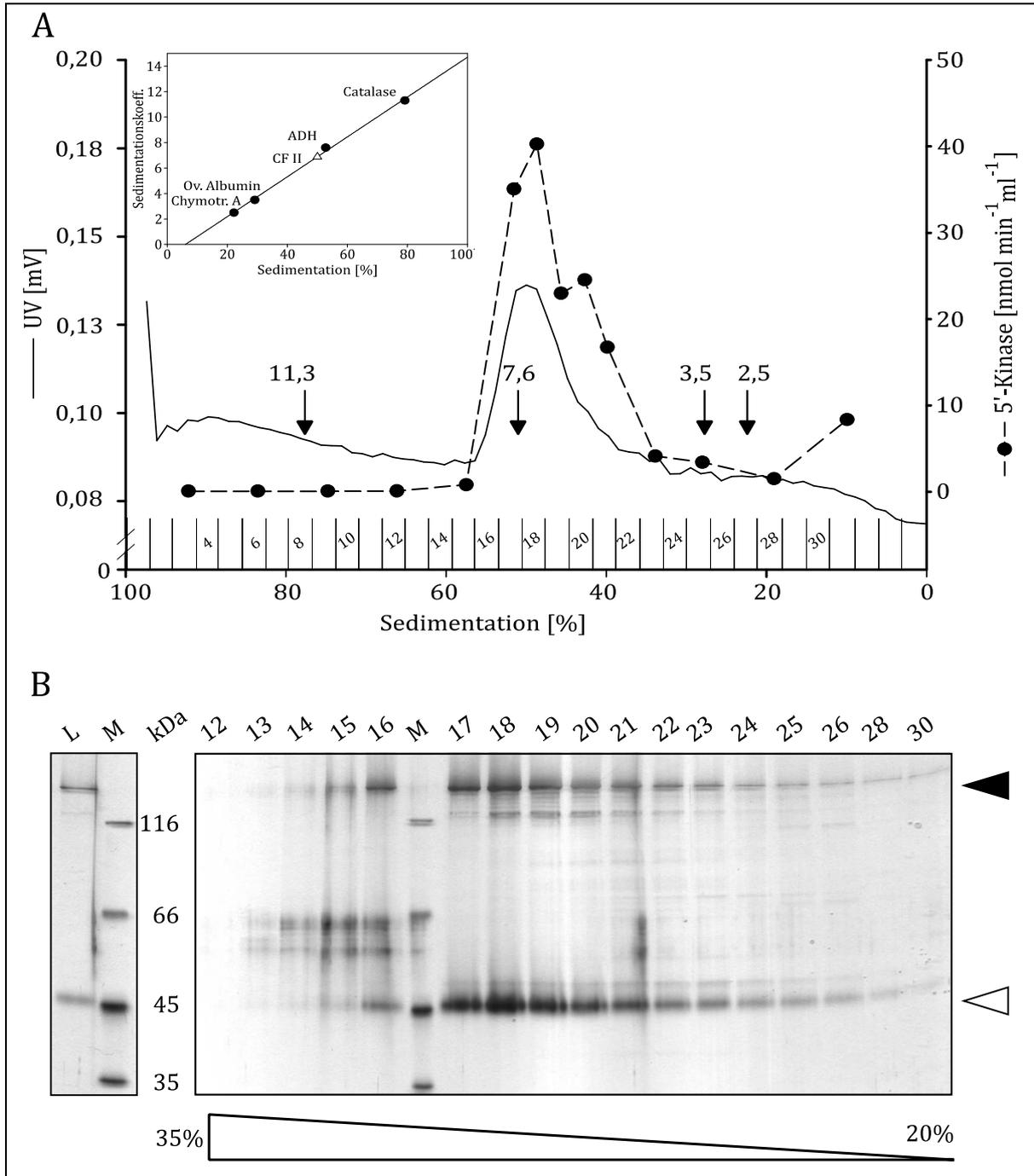


Abbildung 2.7 Dichtegradientenzentrifugation der CF II-MonoQ-Fraktion #19.

Ein Glyceringradient von 20–40% (v/v) in Gradienten-Puffer wurde mit einem Aliquot der Fraktion #19 beladen. Die Zentrifugation erfolgte in einem Beckmann SW40-Ti Swing Out-Rotor für 48h bei 4–8°C und bei 275.000 xg. **A**) Das erhaltene UV-Chromatogramm (durchgezogene Linie) der Auftrennung der MonoQ-Fraktion im Dichtegradienten, aufgenommen vom Boden zur Oberfläche des Gradienten. Die 5'-PNK Aktivität des hClp1 wurde auch hier über die Fraktionen des Gradienten hinweg in nmol je min und je ml (gestrichelte Linie mit Punkten) gemessen. Schwarze Pfeile markieren die genutzten Größenstandards mit der Angabe ihrer jeweiligen Sedimentationskoeffizienten. Eine Eichgerade der Standards ist oben links eingefügt (Punkte) und der CF II-Wert ist darin eingetragen (Dreieck). **B**) Silber-gefärbte SDS-PAGE mit Angabe der Fraktionen des Dichtegradienten über den Spuren. Gut sichtbar ist, dass hPcf11 (schwarzer Pfeilkopf) und hClp1 (weißer Pfeilkopf) als gemeinsamer Komplex in den Fraktion #17–#19 migrierten. Die Glycerinkonzentration des Gradienten ist unter dem Gel angegeben. Die Bildausschnitte wurden aus demselben Gel zusammengefügt.

2.4 Rekombinantes CF II ist in der prä-mRNA-Spaltung aktiv

Die oben beschriebenen gereinigten CF II-Fractionen aus der MonoQ-Säule wurden nun verwendet, um einen teilweise fraktionierten nukleären Extrakt (NXT) aus HeLa-Zellen zu komplementieren und auf Funktionalität des Proteinkomplexes aus hPcf11 und hClp1 in der prä-mRNA-Spaltung zu testen. In diesen modifizierten 3'-Spaltassays wurde eine mit α -[³²P]-UTP intern markierte Substrat-RNA angeboten und die Polyadenylierung nach der Spaltung durch zugesetztes Cordycepintriphosphat (3'-dATP) verhindert. Dem fraktionierten NXT wurde dann rekombinantes Protein hinzugefügt, um die 3'-Spaltaktivität wiederherzustellen. Nach einer ein- bis zweistündigen Inkubation der Spaltreaktion bei 30°C wurden die Spaltprodukte in einer denaturierenden Harnstoff-PAGE aufgetrennt. Als Grundlage diente in den Komplementationen ein fraktionierter Extrakt aus einer früheren Arbeit (Dissertation L. Schönemann). In der Fraktionierung wurde nukleärer Extrakt aus HeLa-Zellen zunächst mit Ammoniumsulfat fraktioniert und dann über eine Gelfiltration aufgetrennt (Abbildung 2.8 A). In Komplementationsassays der 3'-Spaltreaktion mit den Fraktionen der Gelfiltration zeigte sich, dass sowohl Poly(A)-Polymerase (PAP) als auch CF II von den restlichen Spaltungsfaktoren weitgehend abgetrennt wurden (L. Schönemann, s. Anhang Abbildung 9.12). Den verwendeten Proteinfractionen der NXT-Gelfiltration wurde stets Poly(A)-Polymerase, CstF und CF I zugesetzt, um die 3'-Spaltaktivität des CF II nicht durch fehlende Spaltfaktoren zu limitieren. Die Komplementation der Gelfiltrationsfraktion X.2 (GXT) mit Fraktionen aus der MonoQ-Fraktionierung des rekombinanten CF II zeigte eine deutliche 3'-Spaltaktivität (Abbildung 2.8 C). Diese Aktivität korrespondierte mit dem Elutionsprofil von CF II (Abbildung 2.8 C unten). In Abbildung 2.8 C ist auch ersichtlich, dass die Fraktion aus der Gelfiltration des nukleären Extrakts (GXT) deutliche Anzeichen einer unspezifischen Nuklease aufwies. Das gebildete spezifische 5'-Spaltprodukt der verwendeten L3 Substrat-RNA ist bei ca. 200 nt Länge in den Spuren 7–11 gut ersichtlich. In der Spezifitätskontrolle mit mutiertem Polyadenylierungssignal (Mutation AAUAAA zu AAGAAA) in der Substrat-RNA konnte das Spaltprodukt in den Spuren 19–23 nicht beobachtet werden. Deutlich sieht man auch, dass die 3'-Spaltaktivität nur nach der Zugabe von PAP und CF II zu beobachten ist und ohne eines der beiden nicht beobachtet werden kann (Kontrollen Spuren 2–4). Somit lässt sich ableiten, dass der rekombinante heterodimere Komplex aus hPcf11 und hClp1 ausreicht, um die Aktivität des CF II zu vermitteln und dessen 3'-Spaltaktivität rekonstituiert.

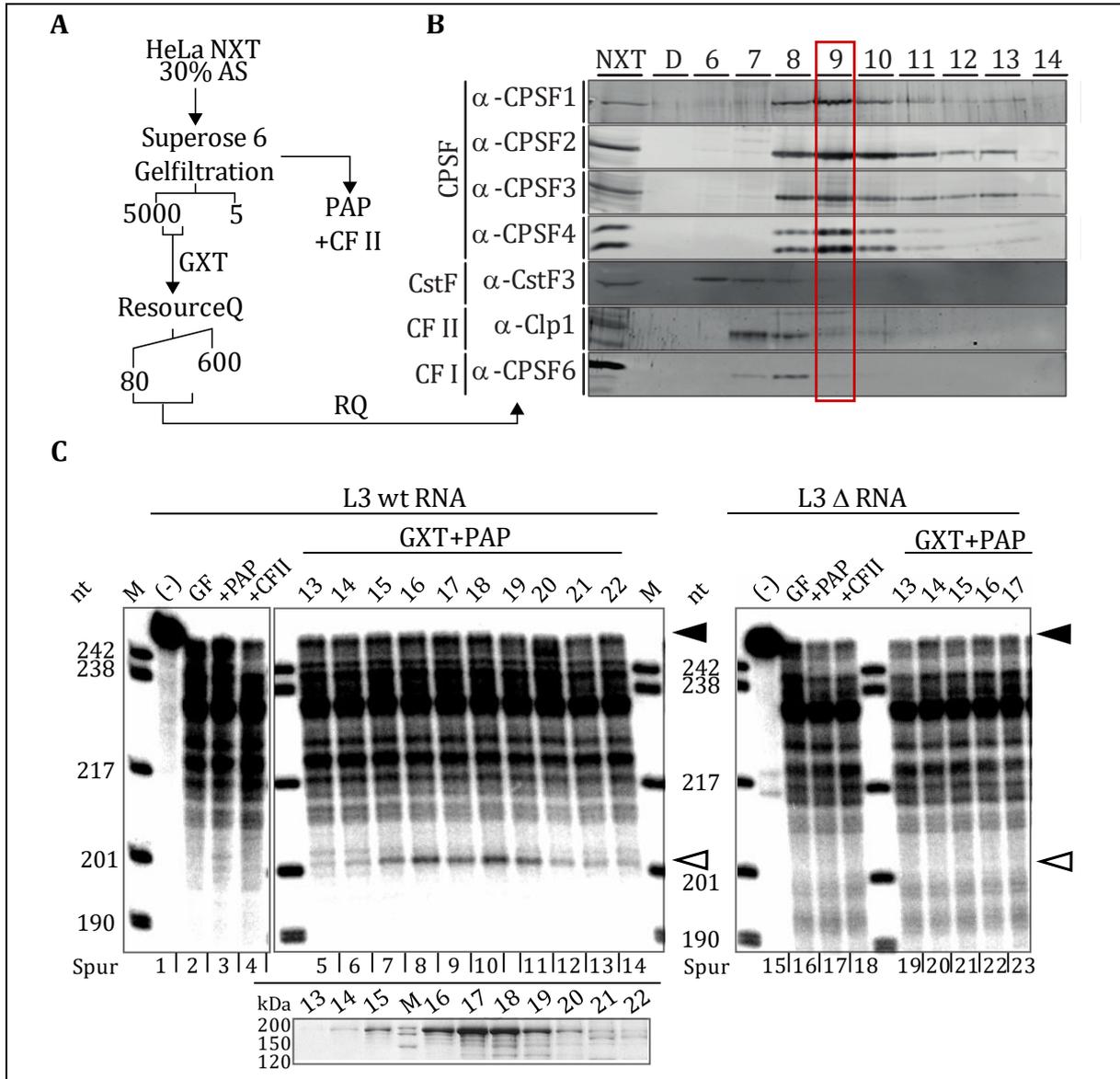


Abbildung 2.8 HeLa NXT-Fraktionierung und anschließender L3-RNA 3'-Spaltassay mit CF II-Komplementation

A) Schema der NXT-Reinigungsschritte aus HeLa Kernextrakt mit Kennzeichnung des Trennverfahrens und der weiter genutzten Fraktionen. Abkürzen sind in folgender Reihenfolge genutzt: 30% AS – 30% Ammoniumsulfat-Fällung bei 0°C aus NXT; Superose6-Gelfiltration – Größenausschlusschromatographie mit Angabe der Molekulargewichte in kDa mit Auftrennung in Poly(A)-Polymerase (PAP) und CF II sowie in GXT-gelfiltrierte NXT-Fraktion; ResourceQ (RQ) – Anionenaustauschchromatographie mit Angabe des KCl-Gradienten in mM KCl. **B)** Western-Blot-Analyse der NXT-Fraktionierung in den Fraktionen der ResourceQ-Säule (durch L. Schönemann). Die Fraktionen sind oben benannt (NXT = HeLa- Kernextrakt, D = Durchfluss der Säule). Die verwendeten primären Antikörper sind an der linken Seite aufgeführt, ebenso wie die Zugehörigkeit der Untereinheiten zu den Faktoren der 3'- Prozessierung. Alle gezeigten Signale stammen von einem Blot mit sequenzieller Detektion der Signale. **C)** Phosphorimager-Detektion eines CF II-Komplementationsassays der 3'-Spaltreaktion der GXT-Fraktion. Durchführung wie in Material und Methoden beschrieben mit je 3 µl GXT und je 100 fmol Poly(A)-Polymerase (PAP) sowie je 2 µl der oben angegebenen CF II-MonoQ-Fraktion pro Reaktion. Unterhalb des Assays sind nochmal die Coomassie-gefärbten hPcf11-Banden der MonoQ-Fractionen als Ausschnitt abgebildet. Die Substrat-RNA war entweder intern radioaktiv markierte L3- oder L3 Δ-RNA. (-) ist eine Kontrolle ohne Proteine. Die Pfeilköpfe markieren Ausgangssubstrat (schwarz) und 5'-Spaltprodukt (weiß). Bildausschnitte sind zusammengefügt aus einem Gel.

2.4.1 Ein N-terminal verkürztes hPcf11 (Δ N769) in CF II ist hinreichend für eine Rekonstitution der 3'-Spaltaktivität

In der Absicht die 3'-Spaltaktivität des CF II auf bestimmte Proteinbereiche einzugrenzen, wurden systematisch Deletionsvarianten von hPcf11 im Komplex mit hClp1 von C. Tüting generiert (Schäfer 2018). In hPcf11 wurden verschiedene Domänen und distinkte Strukturen identifiziert, auf deren Grundlage N-terminale Deletionsvarianten kloniert wurden. Die Domänen und Bereiche sind dargestellt in Abbildung 2.9 A. Sie umfassen N-terminal begonnen die gut konservierte CTD-Interaktionsdomäne (AS 14–142, CID), einen α -helikalen *Coiled-coil*-Bereich (AS 183–297, α -helikal), eine stark polare und basische Region (AS 295–550, polar) und damit überlappend eine Lysin-reiche Region (AS 500–770, XL), welche wohl mit hClp1 interagiert und an dieses quervernetzt werden kann. Mittig in der Proteinsequenz liegt ein nicht strukturierter Bereich aus Aminosäurewiederholungen der Abfolge „FEGP“ (AS 770–1124, FEGP-Wdh.). Der Proteinbereich mit den Aminosäurewiederholungen „FEGP“ bildet wahrscheinlich eine unstrukturierte Region innerhalb des Proteins, welche zudem zahlreiche Arginin-Methylierungen enthält (C. Tüting; UniProt).

Das Muster wiederholt sich alle acht bis 13 Aminosäuren und scheint bisher noch nicht in der Literatur beschrieben (s. Anhang Abbildung 9.7). Als distinktes Merkmal in der Proteinsequenz könnte der Bereich daher eine Funktion vermitteln. Am C-terminalen Ende der Sequenz liegt ein konservierter Bereich, welcher die hClp1-Interaktion vermittelt (AS 1371–1439, hClp1-IR) und von zwei Zinkfinger-Domänen flankiert wird (ZF, AS 1343–1368; AS 1443–1478). Die dargestellten hPcf11-Verkürzungen (Abbildung 2.9 A) im CF II-Komplex wurden analog zum vollständigen CF II-Komplex gereinigt. Dabei konnte stets eine stabile Assoziation des hClp1 mit dem jeweiligen hPcf11-Varianten beobachtet werden.

Für die Komplementation wurde in den 3'-Spaltassays zum Test auf 3'-Spaltaktivität der hPcf11-Verkürzungen eine NXT-Fraktion genutzt, welche nach der Gelfiltration noch über eine anschließende ResourceQ-Säule fraktioniert wurde. Durch die Anionenaustauschchromatographie ergaben sich Fraktionen, in denen auch die restlichen Faktoren CstF, CF I und CPSF teilweise aufgetrennt wurden (Abbildung 2.8 B). In den mit der ResourceQ-Fraktion (RQ) durchgeführten 3'-Spaltassays war zu beobachten, dass Wildtyp-CF II, wie zuvor, 3'-Spaltaktivität besaß (Abbildung 2.9 B, Spuren 4–7).

Die N-terminal verkürzte Variante hPcf11 Δ N769 (Spuren 8–11), welche mit den FEGP-Wiederholungen beginnt, besaß auch 3'-Spaltaktivität. Diese Aktivität schien allerdings etwas schwächer als mit dem vollständigen hPcf11. Die weiter verkürzte Variante hPcf11 Δ N1123, welche direkt nach den Wiederholungen beginnt, zeigte keine Stimulation der Spaltung mehr (Spuren 12–15). Auch noch weiter verkürzte CF II-Varianten waren nicht spaltaktiv. Die

Ergebnisse mit den verkürzten hPcf11-Varianten zeigen, dass die N-terminal gelegene CID des hPcf11 für die Spaltreaktion nicht essenziell ist. Eine 3'-Spaltaktivität mit der bakteriell in *E. coli* exprimierten FEGP-Wiederholungsregion allein war nicht zu beobachten (nicht gezeigt).

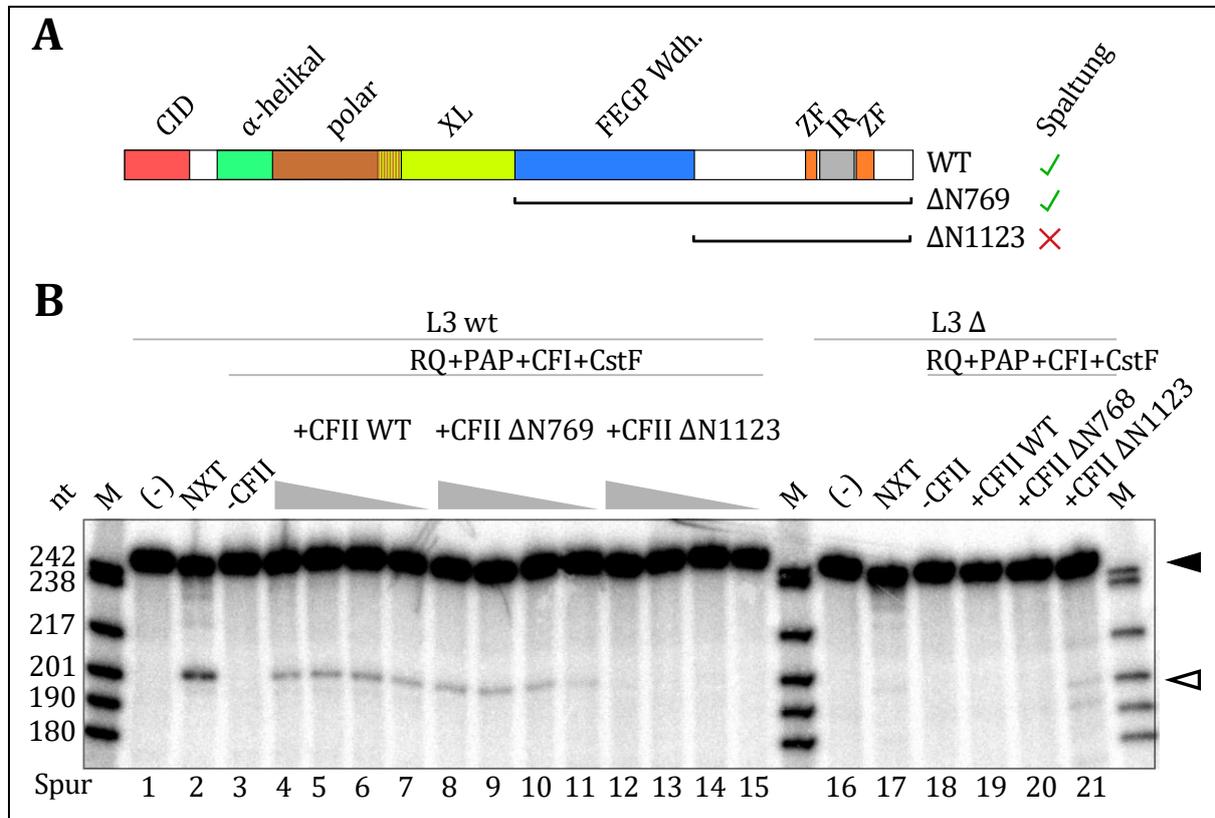


Abbildung 2.9 3'-Spaltassays mit verkürzten hPcf11-Varianten im CF II-Komplex

Erstellt durch C. Tüting (2018) und modifiziert übernommen aus Schäfer u.a. 2018. **A)** Schematische Darstellung des hPcf11 mit hervorgehobenen Domänen und definierten Regionen. CID – CTD Interaktionsdomäne (AS 14–142); XL – hClp1 crosslink Region (AS 500–770); FEGP – Region mit Motivwiederholung von FEGP (AS 770–1124); ZF – Zinkfingerdomäne eins (AS 1343–1368) und zwei (AS 1443–1478); IR – hClp1-Interaktionsregion (AS 1371–1439). Die darunter liegenden Balken markieren die verwendeten N-terminal verkürzten Varianten des hPcf11 im CF II. Auf der rechten Seite ist eine kurze Zusammenfassung der 3'-Spaltaktivität der CF II-Variante angegeben. **B)** 3'-RNA-Spaltassay mit Komplementation von 1 µl der NXT-ResourceQ-Fraktion #9 (RQ Abbildung 2.8 B) durch mit CF II sowie dessen Verkürzungen. Die zugesetzte CF II-Variante wurde in absteigenden Mengen zwischen 100 und 5 fmol pro Reaktion titriert. Als Substrat-RNA wurde 0.8 nM L3 genutzt und die Reaktion mit den darüber angegebenen Proteinen komplementiert. Die hPcf11-Verkürzung mit den FEGP-Wdh. (CF II ΔN769) zeigte Aktivität in der Spaltung. Das hPcf11-Fragment ohne FEGP-Wdh. (CF II ΔN1123) zeigte dies hingegen nicht. Der schwarze Pfeilkopf zeigt die Substrat-RNA, der weiße Pfeilkopf zeigt das 5'-Spaltprodukt.

2.5 Die 5'-PNK Aktivität ist für die Komplementation der Spaltung von prä-mRNA nicht notwendig

2.5.1 Inaktivierung der 5'-PNK-Aktivität durch Mutagenese von hClp1

Als weiterer Aspekt sollte betrachtet werden, ob die im menschlichen Clp1 vorhandene 5'-Polynukleotidkinase-Aktivität essenziell für die Spaltaktivität ist (Betrachtung der 5'-PNK-Aktivität in Kapitel 2.5.2). Das homologe Protein aus Hefe (yClp1) besitzt diese Aktivität aufgrund einiger Mutationen im aktiven Zentrum nicht. Auf Grundlage ausführlicher Mutationsstudien des homologen ceClp1 aus *C. elegans* (Dikfidan u. a., 2014) und der beschriebenen Mutation in hClp1 (Weitzer und Martinez, 2007) wurden entsprechende Mutationen in das humane Clp1 eingefügt. Die Identifizierung der entsprechenden Aminosäuren war durch die hohe Sequenzübereinstimmung zwischen den einzelnen Homologen möglich (Sequenzvergleich s. Anhang Abbildung 9.5). Aus der menschlichen cDNA Sequenz wurde durch zielgerichtete Mutagenese eine Punktmutante mit dem Aminosäureaustausch Lysin-127 zu Alanin (K127A) erzeugt. Im Falle des ceClp1 führte dies zu einem beschriebenen Abfall der Aktivität um den Faktor 150, im humanen Protein nur zu einer Verringerung der Aktivität um etwa den Faktor 30 (Daten s. Anhang Abbildung 9.9). Dies war nicht ausreichend, um eine Restaktivität während der Spaltreaktion auszuschließen. Darum wurden zwei Tripelmutanten generiert (Schema Abbildung 2.10 A). Diese trugen die Punktmutationen K127A, R288A und R293L (im Weiteren abgekürzt KR) oder D151A, R288A und R293L (abgekürzt DR). Dadurch sollten sowohl eine ATP-Bindung als auch der Phosphattransfer gestört werden.

Die hClp1-Mutanten (KR und DR) wurden in Kombination mit vollständigem hPcf11 in Insektenzellen exprimiert und analog zum Wildtyp gereinigt. Dabei zeigte sich in Abbildung 2.10 B, dass die Mutationen keinen nachteiligen Einfluss auf das Verhalten des Komplexes während der Reinigung hatten. Des Weiteren zeigte sich eine ähnliche Affinität des Komplexes für angebotene Substrat-RNA (s. Anhang Abbildung 9.10). Die 5'-PNK-Aktivität wurde über ein Kinaseassay (Material und Methoden 6.5.6) bestimmt und mit dem Wildtyp verglichen. Im Graphen der Abbildung 2.10 B ist dabei ersichtlich, dass eine Verringerung der Aktivität, um mindestens den Faktor 500 zu beobachten war (Nachweisgrenze). Für eine genauere Bestimmung der Restaktivität der Punktmutanten waren die Reaktionsbedingungen sowie die Sensitivität der Methode nicht geeignet. Für die weiteren Komplementationsassays war dies auch nicht notwendig, da die mögliche Restaktivität vermutlich nicht relevant war.

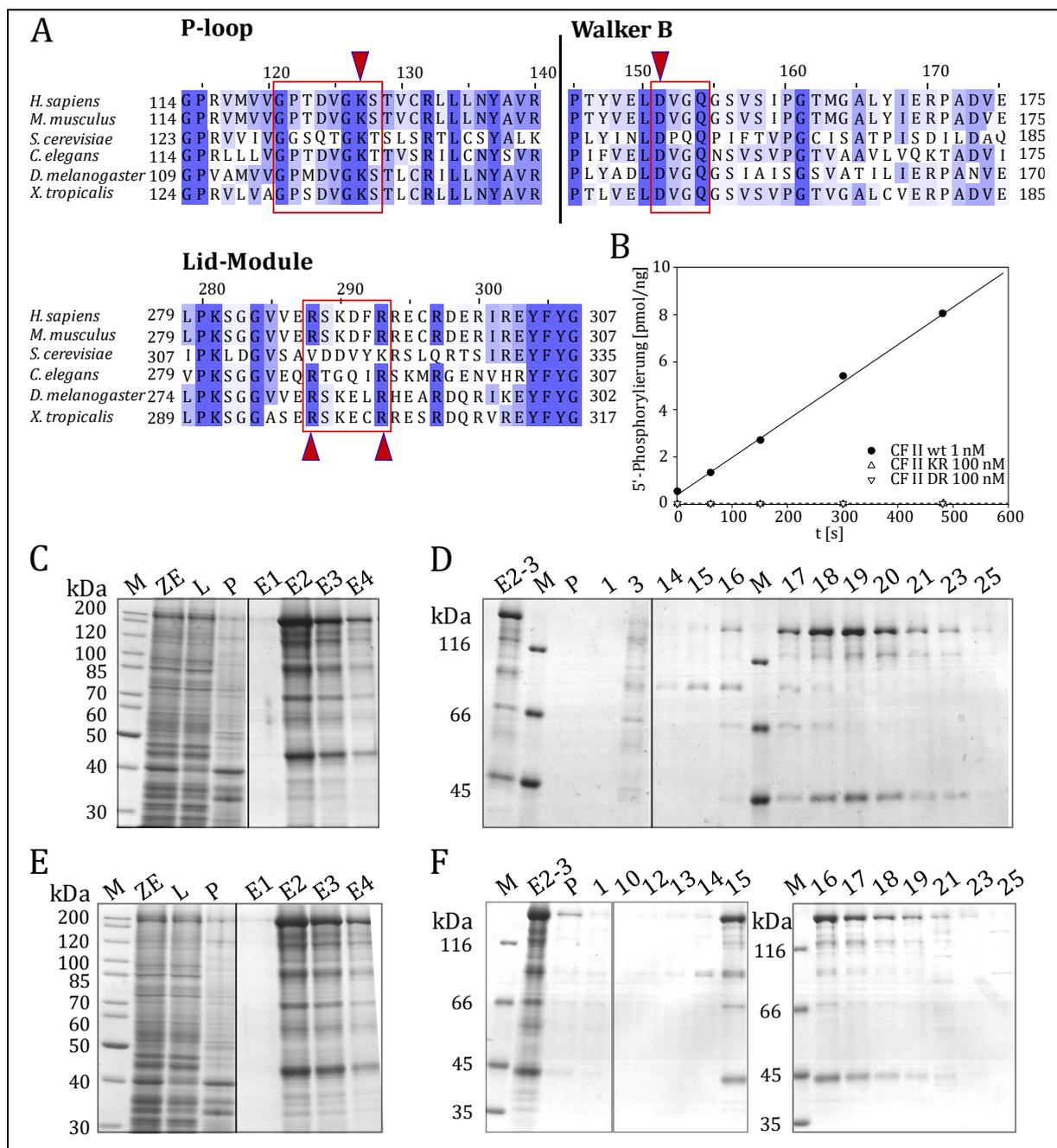


Abbildung 2.10 Aktivität und Reinigung Kinase-inaktiver Mutanten des hClp1 im CF II-Komplex

A) Ausschnitte eines Aminosäure-Sequenzvergleiches von hClp1-Homologen aus verschiedenen Spezies in den Bereichen der eingeführten hClp1-Mutationen K127A, D151A, R288A und R293L. Der P-loop der zentralen 5'-PNK-Domäne enthält mit K127 eine zentrale Aminosäure, die γ -Phosphat-Transfer von ATP auf das RNA-Substrat stabilisiert. Das Walker B-Motiv der zentralen 5'-PNK-Domäne stellt mit D151 die katalytische Aminosäure bereit, welche die 5'-Hydroxyl-Gruppe der RNA deprotoniert. Das Lid-Module mit R288 und R293 stabilisiert das gebundene ATP in der aktiven Tasche und stabilisiert die negativen Ladungen des ATP. Rote Pfeilköpfe markieren hier die genannten und mutierten Aminosäuren. **B)** 5'-Kinaseassay mit CF II und hClp1 Mutanten KR (K127A, R288A, R293L) und DR (D151A, R288A, R293L) in CF II. Die 5'-Phosphorylierung wurde mit den angegebenen Konzentrationen CF II und $10 \mu\text{M}$ rC₁₄-RNA als Substrat durchgeführt. **C)** Coomassie-gefärbte SDS-PAGE der Ni-NTA-Affinitätsreinigung aus der Expression von CF II mit der hClp1-Mutante KR in SF21-Zellen. **D)** Coomassie-gefärbte SDS-PAGE der MonoQ-Reinigung des CF II KR der Ni-NTA-Elutionsfraktionen E2 und E3 aus C). **E)** Coomassie-gefärbte SDS-PAGE der Ni-NTA-Affinitätsreinigung aus der Expression von CF II mit der hClp1-Mutante DR in SF21-Zellen. **F)** Coomassie-gefärbte SDS-PAGE der MonoQ-Reinigung der CF II DR Ni-NTA-Elutionsfraktionen E2 und E3 aus E).

Die MonoQ-Fractionen der Mutanten (s. Abbildung 2.10 D und F) wurden in einem Komplementationsassay eingesetzt. Ausgangsmaterial war hierbei eine ResourceQ-Fraktion (A10) des fraktionierten HeLa-Kernextraktes, welche hauptsächlich CPSF enthielt (s. Abbildung 2.8 B). Wie schon beschrieben, wurde diese mit CF I, PAP und CstF komplementiert, um die CF II-Abhängigkeit der 3'-Spaltung zu gewährleisten. In den Spuren 4–7 der Abbildung 2.11 erkennt man deutlich, dass sowohl der Wildtyp des hClp1 im CF II Komplex die Spaltaktivität vermittelte, als auch die beiden generierten hClp1 Mutanten KR (Spuren 8–11) und DR (Spuren 12–15). Dabei war kein nennenswerter Unterschied in dem Ausmaß der Spaltaktivität zu beobachten. Eine Kontrollreaktion in den Spuren 16–19 mit hClp1 zeigte hier, dass hClp1 allein die Spaltaktivität nicht komplementieren kann. Dies war allerdings schon aufgrund der Ergebnisse mit den verkürzten hPcf11 Varianten zu erwarten. Die Ergebnisse zeigen, dass die 5'-PNK-Aktivität für die 3'-Spaltung von Prä-mRNA nicht notwendig ist. Diese Beobachtung, zusammen mit der starken Konservierung des Proteins über verschiedene Spezies hinweg, lässt die Vermutung zu, dass die 5'PNK-Aktivität eine weitere essenzielle Rolle außerhalb der 3'-Prozessierung der mRNAs besitzt. Dies könnte in Assoziation mit der tRNA-*Splicing*-Endonuklease der Fall sein.

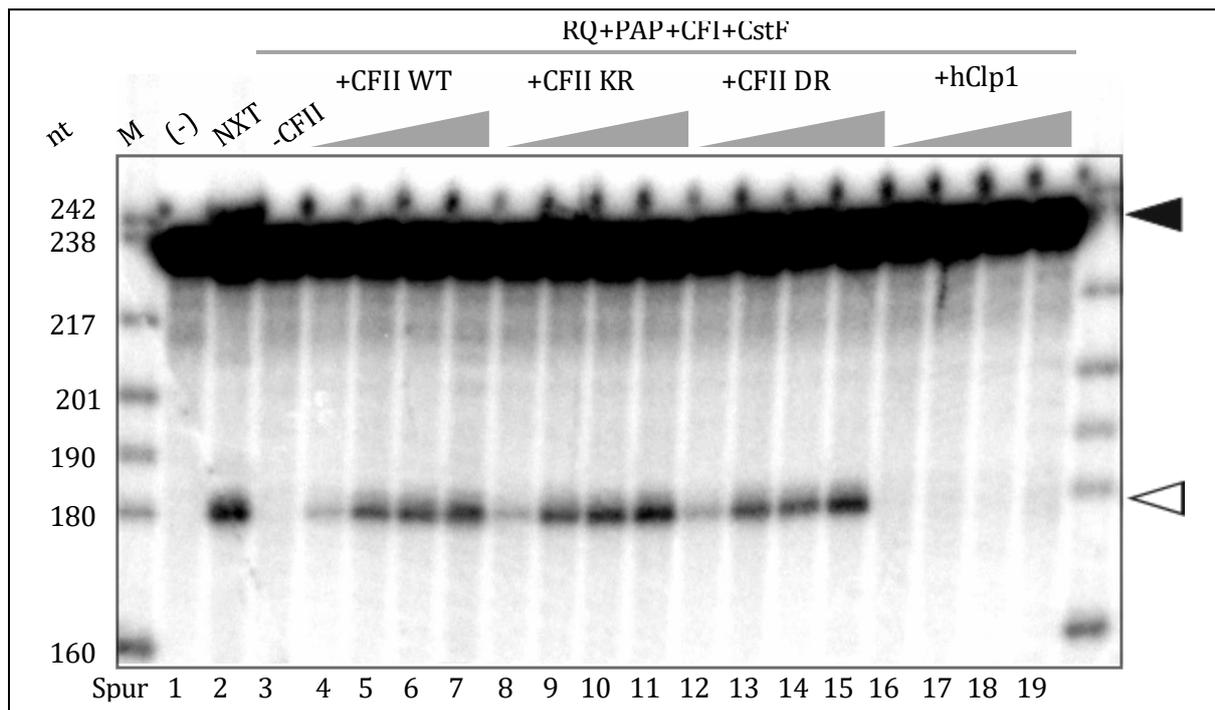


Abbildung 2.11 3'-Spaltsassay durch Komplementation eines fraktionierten nukleären Extraktes mit CF II und hClp1-Mutanten in CF II.

In dem 3'-Spaltsassay wurden wie beschrieben je 1 µl der RQ-Fraktion #9 aus der NXT-Fraktionierung mit den angegebenen Proteinen komplementiert. In den Reaktionen wurden CF II-Varianten in steigenden Mengen eingesetzt. Als Ausgang diente CF II Wt (2–200 fmol) im Vergleich zu den hClp1-Mutanten KR (K127A, R288A, R293L) oder DR (D151A, R288A, R293L) im CF II-Komplex (2–200 fmol). Alle drei CF II-Varianten zeigten vergleichbare Spaltaktivität. Eine Titration von isoliertem hClp1 Wt zeigte keine Spaltaktivität. Die Pfeilköpfe markieren die ungespaltene SV40 late-RNA (schwarz) sowie das 5'-Spaltprodukt (weiß). Spaltaktivität konnte in der Kontrollreaktion ohne CF II nicht detektiert werden. Die Aktivität des hClp1 hat scheinbar keinen Einfluss auf den Schritt der 3'-Spaltung. Die Spaltung war abhängig vom „AAUAAA“-Signal (nicht gezeigt).

2.5.2 Charakterisierung der kinetischen Parameter der 5'-Polynukleotidkinase-Aktivität von hClp1

Die 5'-Kinaseaktivität des hClp1 wurde bislang qualitativ durch Weitzer und Martinez (Weitzer und Martinez, 2007) untersucht. Sie fanden für das menschliche Clp1 eine Substratspezifität für doppelsträngige siRNA-Substrate mit 3'-Überhang und Spleißprodukte der tRNA-Prozessierung. Jedoch wurde keine systematische Analyse möglicher weiterer Substrate oder kinetischer Parameter durchgeführt. Dikfidan und Kollegen (Dikfidan u. a., 2014) beschrieben in ihrer Arbeit zum *C. elegans* Homolog (ceClp1) beschrieben, dass dieses bevorzugt einzelsträngige RNA phosphoryliert, welche ein 5'-GC enthält. Als Substrat akzeptiert wird zudem dsRNA mit 5'-Überhang und sehr ineffizient auch andere dsRNA-Varianten. DNA-Substrate wurden hingegen fast gar nicht als Substrat akzeptiert. Die durch Dikfidan und Kollegen für das ceClp1 ermittelten kinetischen Parameter beliefen sich dabei auf einen K_M für ATP von rund 0,26 mM und einen K_M für das RNA-Substrat von 100 μ M mit einem k_{cat} von etwa 2,9 s^{-1} .

Im Weiteren sollten nun auch die kinetischen Parameter des humanen Clp1 ermittelt und in Relation zum ceClp1 gesetzt werden. Zur Beantwortung der Frage nach Substratspezifität, K_M und k_{cat} wurden mit hClp1 und mit CF II 5'-PNK-Assays durchgeführt. Hierzu wurde als erstes eine Reihe von einzelsträngigen RNA-Substraten in steigenden Konzentrationen bei konstanter Protein- und ATP-Konzentration in Kinaseassays angeboten. Die ATP-Konzentration betrug dabei konstant 1 mM. Zur Detektion phosphorylierter Produkte wurde anteilig γ -[32 P]-ATP hinzugesetzt. Die entstandenen phosphorylierten RNA-Produkte wurden anschließend über 20%ige Harnstoff-PAGE vom ATP getrennt, über Phosphorimaging detektiert und mit ImageQuant 5.1 quantifiziert.

In Voruntersuchungen wurden Versuche mit dem CF II-Komplex durchgeführt, um eine eventuelle Substratspezifität einzugrenzen. Dabei zeigte sich, dass bei einigen der angebotenen Substrate Substratüberschusshemmungen in den Reaktionen beobachtet werden konnten (Abbildung 2.12 A). Die angebotenen rC₁₄- und rU₁₄-Homopolymere zeigten eine Reaktion, die einer Michelis-Menten-Kinetik entsprach. Am wahrscheinlichsten ist hierbei wohl eine störende intermolekulare Strukturbildung der RNA-Moleküle, welche die 5'-Hydroxygruppe der Substrate für hClp1 unzugänglich machte. Eine Auswertung der erhaltenen Messwerte (SigmaPlot 12.5), angepasst auf das Modell für Substratüberschusshemmung:

$$v = \frac{v_{max}}{1 + \frac{K_m}{[S_0]} + \frac{[S_0]}{K_i}}$$

... ((Schellenberger u. a., 1989) S. 101, s. Anhang Tabelle 9.1), zeigte hierbei K_M -Werte für die gehemmten Substrate in einem Bereich von 0,5–1,1 μM . Die V_{max} -Werte waren jedoch schnell abfallend und daher für eine Michaelis-Menten-Kinetik ungeeignet.

Zusätzlich zur Substratspezifität wurde untersucht, ob es eine Längenabhängigkeit der Effizienz der Phosphorylierung in verschiedenen RNA-Homopolymer-Substraten gibt. Dazu wurden Homopolymere von Poly(A)/-(U)/-(C) durch alkalische Lyse in Fragmente unterschiedlicher Längen gespalten und im Überschuss dem CF II in einer PNK-Reaktion als Substratpool angeboten. Als Vergleich diente hierbei die T4-Polynukleotidkinase (T4-PNK). Dabei zeigte sich nach der Auftrennung auf einer Harnstoff-PAGE (gezeigt in Abbildung 2.12 B für die drei RNA-Homopolymere), dass im Gegensatz zur T4-PNK-Reaktion keine bevorzugte Substratlänge innerhalb der angebotenen Population sichtbar war.

Als getestetes Substrat mit höchstem V_{max} -Wert stellte sich in den Vorversuchen das rC_{14} -Homooligomer heraus. Dies wurde genutzt, um die weiteren Assays für hClp1 und die Bestimmung der ATP-Affinität durchzuführen. Wie oben erwähnt, wurden die kinetischen Gleichgewichtsparameter über die Michaelis-Menten Gleichung

$$v = \frac{v_{\text{max}} * [S]}{[S] + K_M}$$

... (Schellenberger u. a., 1989) für das rC_{14} -Substrat ermittelt. In

Tabelle 2.2 sind die ermittelten Parameter ersichtlich. In der Abbildung 2.12 C sind die Michaelis-Menten-Kinetiken mit den hyperbolen Anpassungen für das rC_{14} -Substrat mit den Proteinpräparationen von CF II und hClp1 zu sehen. Der K_M (rC_{14}) lag für CF II bei $2,5 \pm 0,4 \mu\text{M}$ und für hClp1 in zwei Einzelbestimmungen bei $2,4 \mu\text{M}$ und $1,6 \mu\text{M}$. In der linearen Regression der Lineweaver-Burk Auftragung der Werte erkennt man in Abbildung 2.12 D eine Annäherung der K_M -Werte für beide Proteinpräparationen. So kann man annehmen, dass sich die Affinitäten für das RNA-Substrat vermutlich nicht unterschieden. Anmerkend sollte erwähnt werden, dass die verwendete Konzentration des ATP von 1 mM die Reaktion nicht in Sättigung verlaufen ließ. Die erforderlichen Konzentrationen an radioaktivem ATP wären für die Durchführung unter sättigenden Bedingungen zu hoch gewesen. Der k_{cat} war mit $1,6/0,7 \text{ s}^{-1}$ bei hClp1 etwa ein Drittel des erreichten k_{cat} im CF II-Komplex mit $4,5 \pm 1,4 \text{ s}^{-1}$.

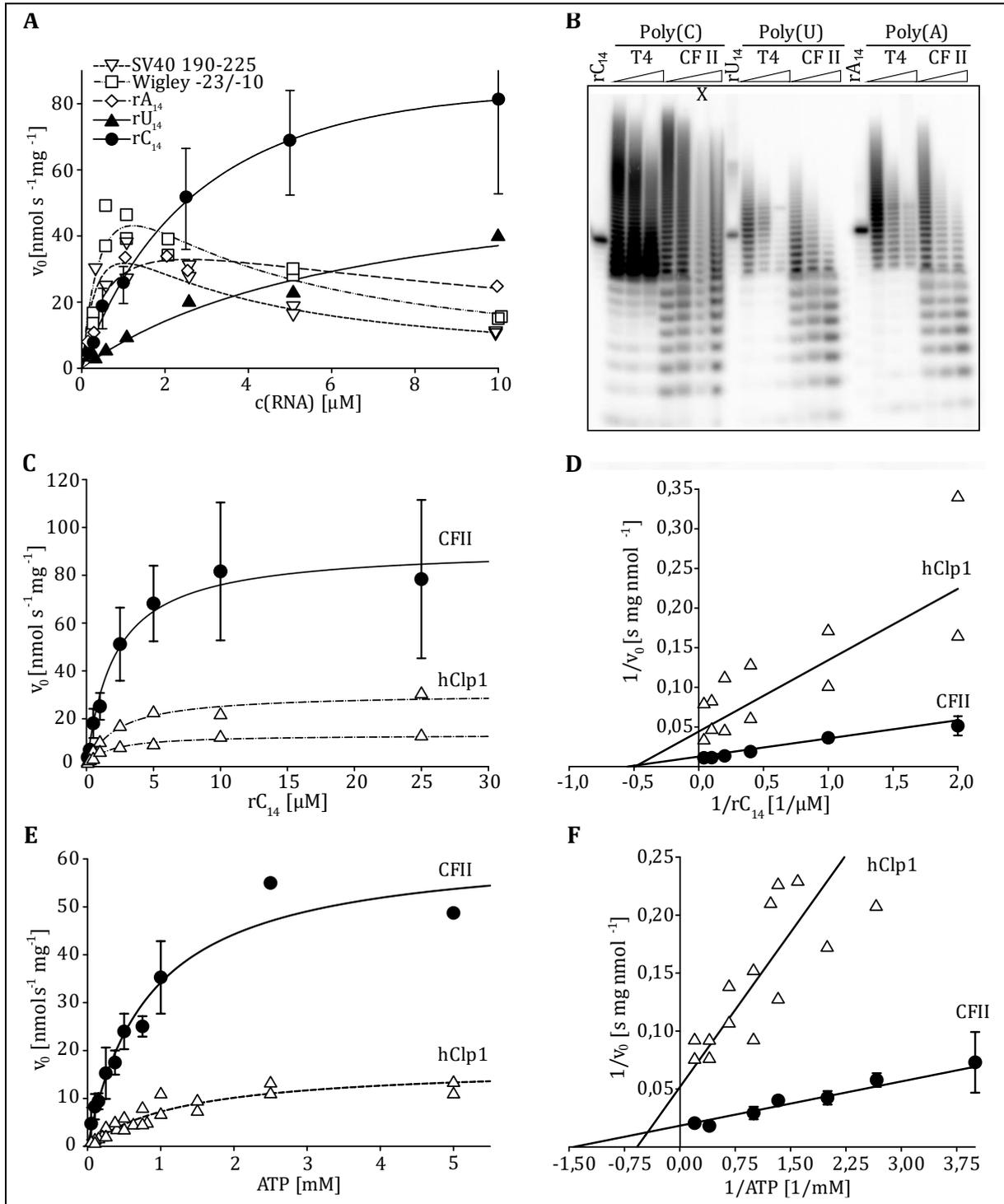


Abbildung 2.12 5'-PNK-Kinetiken von CF II und hClp1.

A) Reaktionsgeschwindigkeiten der 5'-PNK-Reaktionen von 1 nM CF II mit Titration von verschiedenen Substraten bei einer konstanten ATP-Konzentration von 1 mM (γ -ATP Anteil). Es wurde eine Anpassung an das Modell der Substratüberschusshemmung vorgenommen. **B)** 15%ige Harnstoff-PAGE von 5'-PNK Reaktionen dreier RNA-Homopolymere absteigender Länge innerhalb der Population (alkalische Hydrolyse), wobei die Balken die Dauer der alkalischen Lyse angeben. Verglichen wurde die Längenabhängigkeit mit 1,8 μ g RNA pro Reaktion in 5'-PNK-Reaktionen mit 250 nM CF II pro Reaktion und 5 U T4-PNK (NEB) pro Reaktion. Als Größenstandard diente ein jeweils 5'-markiertes rC/U/A₁₄-RNA-Oligomer. X= Spur defekt.

C) Aufgetragen sind Reaktionsgeschwindigkeiten von 5'-PNK-Reaktionen mit 2 nM CF II und hClp1, ausgewertet durch Harnstoff-PAGE und Phosphorimaging, angepasst nach dem Michaelis-Menten Modell. Reaktionen enthielten eine konstante ATP-Konzentration (γ -ATP Anteil) von 0,75 mM und steigende Substratkonzentration rC₁₄-RNA. CF II-Datenpunkte mit Fehlerbalken repräsentieren hier Dreifachbestimmungen. Aufgrund sinkender hClp1-Aktivität sind dafür nur Einzelbestimmungen aufgetragen. **D)** Lineweaver-Burke-Linearisierung der Daten aus C) mit Regressionsgeraden für CF II und hClp1. Gemeinsame Regression der Datenpunkte der einzelnen hClp1-Kinetiken. **E)** Aufgetragen sind Reaktionsgeschwindigkeiten von 5'-PNK-Reaktionen mit 5 nM CF II und hClp1, ausgewertet durch DEAE-Methode und Szintillationsmessung, angepasst nach Michaelis-Menten-Modell. Reaktionen enthielten konstante Substratkonzentration von 20 μ M rC₁₄-RNA und steigende ATP-Konzentration (γ -ATP Anteil). CF II-Datenpunkte mit Fehlerbalken aus 3 Einzelbestimmungen. Die hClp1-Bestimmungen sind wieder einzeln aufgetragen. **F)** Lineweaver-Burke-Linearisierung der ATP-Titration aus E) mit Regressionsgeraden für CF II und hClp1. Auch hier gemeinsame Regression der hClp1 Einzelbestimmungen.

Für die Bestimmung der kinetischen Konstanten bezüglich des ATP wurde 5 nM hClp1 mit 20 μ M rC₁₄ als Substrat eingesetzt. ATP wurde dabei in einem Bereich von 0,1 bis 5 mM titriert. Das nicht umgesetzte ATP wurde über die DEAE-Methode abgetrennt und die eingebaute Radioaktivität quantifiziert. Für ATP zeigte sich, wie in

Tabelle 2.2 zu sehen, dass CF II mit 700 \pm 200 μ M im Vergleich zu hClp1 mit 2500 \pm 1100 μ M eine leicht erhöhte Affinität besaß. In den Graphen der Abbildung 2.12 E und F kann dies sowie eine erhöhte Reaktionsrate des CF II auch schon gut anhand der Kurvenverläufe abgelesen werden. Die erhöhte Affinität zu ATP in den CF II-Reaktionen könnte durchaus aus einem stabilisierenden Effekt des hPcf11 auf hClp1 resultieren, welcher die Affinität für ATP steigert. Anteile von inaktivem Protein sind hierbei irrelevant, denn diese hätten zwar Einfluss auf Reaktionsrate, aber nicht auf die apparente Affinität zum Substrat. Auch in den ATP-Titrationen zeigte sich, dass die Umsatzrate für die hClp1-Präparation mit 1,3 \pm 0,5 s⁻¹ niedriger war als die Umsatzrate der CF II-Präparation mit 2,8 \pm 0,8 s⁻¹.

Tabelle 2.2 Michaelis-Menten-parameter der 5'-PNK Aktivität von CF II und hClp1.

Kinetische Parameter der 5'-PNK-Reaktion nach Michaelis-Menten für hClp1- und CF II-Reaktionen. Werte mit Fehlerabweichung resultieren aus gemittelten Werten von mindestens drei verschiedenen Assays (n=3-4). Werte aus Einzelbestimmungen sind durch „/“ getrennt.

Substrat	Protein	k_{cat} [s ⁻¹]	K_M [μ M]	k_{cat}/K_M [1/s*M]
ATP	CF II	2,8 \pm 0,8	700 \pm 200	4300 \pm 1400
	hClp1	1,3 \pm 0,5	2500 \pm 1100	500 \pm 100
rC ₁₄	CF II	4,5 \pm 1,4	2,5 \pm 0,4	1,8x10 ⁶ \pm 0,4x10 ⁶
	hClp1	(1,6/0,7)	2,4/1,6	(0,7x10 ⁶ /0,4x10 ⁶)

Für die rC_{14} -Versuche muss angemerkt werden, dass die hClp1-Präparation im Laufe der Lagerung an Aktivität verlor und die Werte für die Bestimmung der rC_{14} Parameter Einzelbestimmungen darstellen (Abbildung 2.12 C). Die tatsächlichen kinetischen Parameter für rC_{14} könnten für k_{cat} und K_M durchaus höher liegen und eher denen des CF II-Komplexes entsprechen.

2.6 Humanes Pcf11 bindet bevorzugt an Guanin-reiche RNAs

Für yPcf11 konnte neben der Assoziation zu yClp1 eine schwache unspezifische RNA-Bindung in der N-terminal gelegenen CID gezeigt werden (Hollingworth u. a., 2005; Noble u. a., 2005; Zhang, 2005). Zudem wurden durch verschiedene Gruppen zwei Zinkfinger-Domänen beschrieben. Der C-terminale Zinkfinger bildet dabei einen neuen Typ des Zinkfingers, der zuvor nicht bekannt war (Guéguéniat u. a., 2017; Yang u. a., 2017). Diese Zinkfinger sind im humanen Protein konserviert (s. Anhang Abbildung 9.6). In RNA-Bindungsversuchen mit C-terminalen Fragmenten des yPcf11, konnte allerdings keine spezifische RNA-Bindung beobachtet werden (Guéguéniat u. a., 2017).

2.6.1 hPcf11 bindet RNA-Substrate mit einer Mindestlänge von 14 Nukleotiden

Die mögliche RNA-Bindung von humanem Pcf11 an Substrat-RNA wurde in dieser Arbeit zunächst über Filterbindungsversuche untersucht. In diesen Filterbindungen wurden die Spaltsubstrat-RNAs L3 und SV40 late (s. Anhang Abbildung 9.1 und Abbildung 9.3) in einer Bindungsreaktion angeboten. Es zeigte sich für den humanen CF II Komplex eine Bindung an beide RNA-Spaltsubstrate (Abbildung 2.13 A). Die Bindung an die 231 nt lange SV40 *late*-RNA konnte mit einer apparenten Dissoziationskonstante (K_{50}) von etwa 0,5 nM ermittelt werden. Die Affinität zur 247 nt langen Adenovirus L3-RNA wurde mit 0,9 nM ermittelt. In den Versuchen wurde eine Proteinkontamination der CF II-Präparation ausgeschlossen, da die RNA-Bindung des CF II über die MonoQ-Säule hinweg dem CF II-Komplex zugeordnet werden konnte (siehe Abbildung 2.5 D). Das galt jedoch nur für den CF II-Komplex. Eine starke Bindung von hClp1 an RNA konnte nicht beobachtet werden. Die Versuche mit hClp1 zeigten eine um mindestens den Faktor 150 verringerte Bindung im Vergleich zu CF II. Zudem konnte in einem RNA-Bindungstest über die zugehörigen hClp1-Frakturen der Superdex200-Säulenfraktionen diese Affinität nicht eindeutig dem hClp1 Protein zugeordnet werden (nicht gezeigt). Eine mögliche Interpretation der bisherigen Ergebnisse ist, dass die Bindung des CF II an RNA vermutlich über das hPcf11 vermittelt wird. Im Weiteren wird die RNA-Bindung des CF II hinsichtlich der Spezifität und der Mindestlänge der Substrat-RNA charakterisiert.

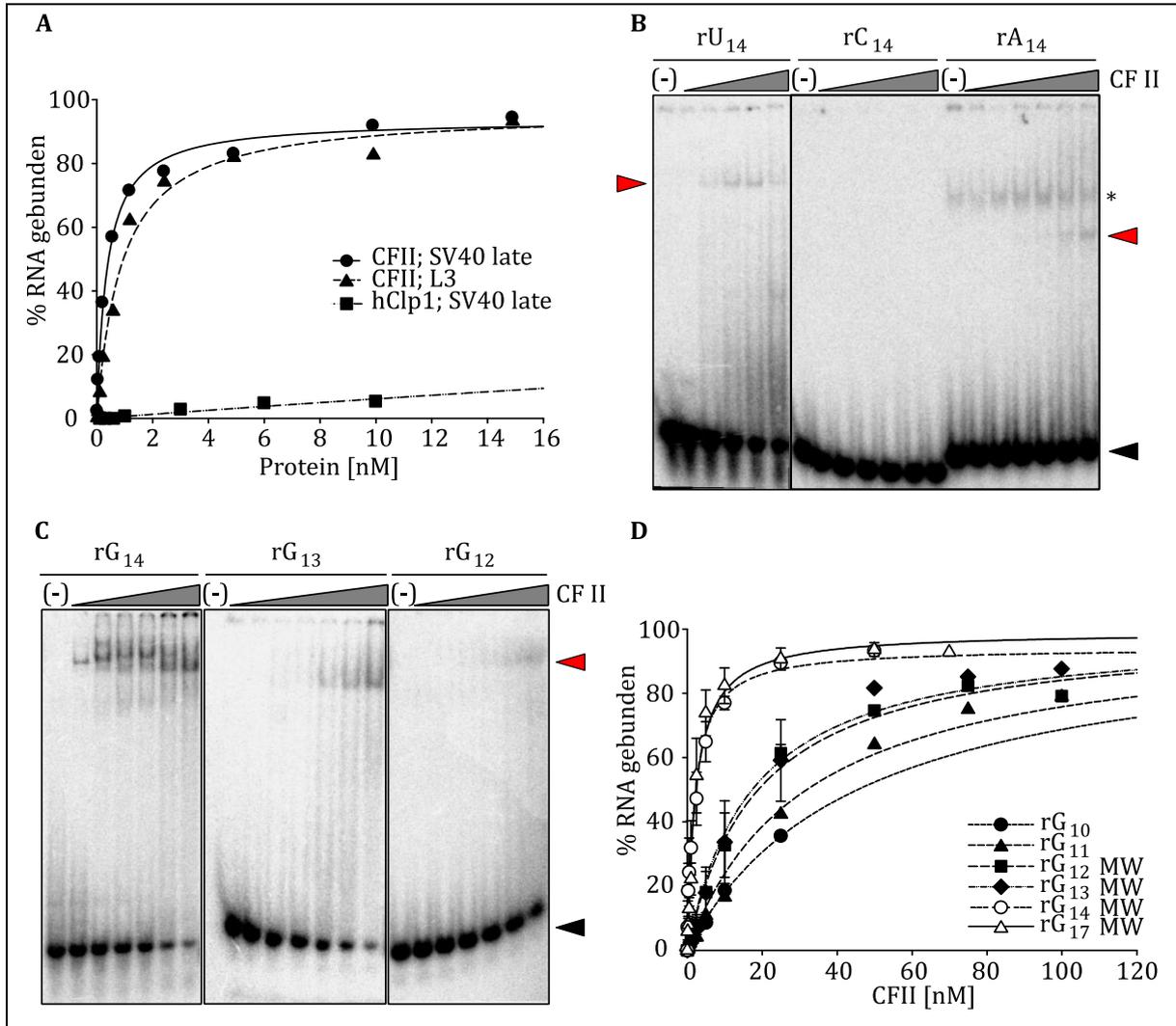


Abbildung 2.13 Bindungsversuche von CF II an verschiedene Substrate.

A) Filterbindungsversuche von CF II an SV40 late- und L3-Substrat-RNA. Die Affinitäten der beiden RNAs für CF II sind vergleichbar mit 0,5 nM und 0,9 nM. Die hClp1-Präparation bindet RNA nur mit sehr schwacher Affinität. **B)** Nativer Gelshift von steigenden Konzentrationen CF II (0–50 nM) mit 0,1 nM rU/C/A₁₄ RNA-Homooligomeren. Der rote Pfeilkopf markiert RNA im Komplex mit CF II, der schwarze Pfeilkopf markiert freie RNA. Der Stern markiert rA₁₄-Sekundärstrukturen, welche auch in der RNA-Kontrolle ohne Protein (–) sichtbar sind. **C)** Nativer Gelshift steigender Konzentrationen CF II (0–25 nM) und 0,1 nM rG₁₄ sowie CF II (0–100 nM) und 0,1 nM rG₁₃/rG₁₂. Der schwarze Pfeilkopf markiert die ungebundene RNA, der rote Pfeilkopf die gebundene RNA im CF II-Komplex. Deutlich sichtbar ist der Anstieg der Affinität von 13 ($K_{50} \sim 20$ nM) zu 14 ($K_{50} \sim 2$ nM) Nukleotiden Länge der Homopolymere. **D)** Hyperbole Anpassung der in-Gel Quantifizierung der gebundenen RNA aus nativen Gelshifts von CF II mit Guanin-RNA-Homopolymeren steigender Länge von 10 bis 17 Nukleotiden. Sichtbar ist auch hier, dass die Affinität des CF II zum Substrat stark zunimmt von 13 zu 14 Nukleotiden Länge. Einige der Kurven resultieren aus Mittelwerten (MW) von Dreifachbestimmungen mit dargestellter Standardabweichung.

Aufgrund der Beobachtung guter RNA-Bindung in Filterbindungsversuchen sollte nun die Sequenzspezifität des hPcf11 weiter eingegrenzt werden. Filterbindung erwies sich jedoch in Verbindung mit einigen kurzen RNA-Sequenzen als ungeeignet, da diese auch ohne Proteinassoziation an die Nitrozellulose banden. So wurde als alternative Methode die native PAGE (EMSA) verwendet. Die Detektion der RNAs erfolgte über eine 5'-[³²P]-Markierung der verwendeten RNA-Oligonukleotide. Meist wurde T4-PNK zu Markierung genutzt, in einigen Fällen wurden besonders G-reiche Sequenzen mittels CF II (hClp1) markiert, da dieses eine verbesserte Ausbeute gegenüber der Reaktion mit T4-PNK erbrachte. Die eingesetzten RNA-Oligonukleotide wurden stets unmittelbar vor der Bindungsreaktion für drei Minuten auf 95 °C erhitzt und auf Eis abgekühlt, um eventuelle Sekundärstrukturen aufzulösen. Standardmäßig wurde in den Bindungsreaktionen eine Endkonzentration der Substrat-RNA von 0,1 nM verwendet. Für schwach bindende RNAs ($K_{50} \geq 5$ nM) wurden auch Konzentrationen bis 0,5 nM eingesetzt. Die verwendete CF II-Präparation war ein Pool der MonoQ-Fraktionierung. Somit wurde gewährleistet, dass beim Wechsel von Aliquots stets die gleiche Proteinpräparation verwendet wurde. Nach der Trennung der Bindungsreaktion auf der nativen PAGE erfolgte die Detektion und Quantifizierung der RNA über Phosphorimaging. Dabei erfolgte die Auswertung der quantifizierten Mengen gebundener und freier RNA in relativen Prozentwerten, um die Vergleichbarkeit der verschiedenen Experimente untereinander zu gewährleisten.

Als Ergebnis dieser native PAGE-Experimente zeigte sich eindeutig eine starke Präferenz für die Bindung von CF II an G-reiche Sequenzen (Abbildung 2.13 B und C) mit einem apparentem K_{50} um 2 nM. Mit deutlich geringerer Affinität wurden Oligo rU₁₄ mit einer Dissoziationskonstanten von rund 70 nM und in noch geringerem Maße Oligo rA₁₄ mit ~320 nM gebunden. Der Bindungsreaktion mit rC₁₄ zeigte keine detektierbaren Mengen an gebundener RNA (Abbildung 2.13 B). Zudem konnte, wie in Abbildung 2.13 C und D sichtbar, ermittelt werden, dass die starke Bindung von Oligo-(G) mit einem K_{50} von etwa 2 nM ab einer Mindestlänge von 14 Nukleotiden erfolgte. Die Affinitäten nehmen unterhalb einer Länge von 14 stark ab. Schon eine Verkürzung um nur ein Nukleotid (Oligo-rG₁₃) führt, wie in Abbildung 2.13 D und Tabelle 2.3 erkennbar, zu einem drastischen Abfall der Affinität um den Faktor 10–15. Im Gegensatz dazu wird das rG₁₇ Oligonukleotid mit annähernd gleicher Affinität von 1,9 nM gebunden. Die für RNA-bindende Proteine untypisch lange Bindungssequenz von 14 Nukleotiden könnte auf mehrere Kontaktstellen des Proteins zum Substrat hindeuten, welche im Zusammenspiel eine affine Bindung gewährleisten. Denkbar wäre neben der Sequenz auch die Notwendigkeit einer RNA-Struktur, welche mit kürzeren Oligonukleotiden nicht gewährleistet werden kann.

Tabelle 2.3 Apparente Dissoziationskonstanten aus den nativen Gelshifts der Homopolymer-Reihen mit CF II

Name	apparenter K_{50} [nM]
rG ₁₀	45
rG ₁₁	41
rG ₁₂	21±9
rG ₁₃	21±12
rG ₁₄	2.0±1,3
rG ₁₇	1,9±0,9
rU ₁₄	~70
rC ₁₄	>1000
rA ₁₄	~320

Zusätzlich zu den Filterbindungsexperimenten und EMSA konnte ein direkter Nachweis der Bindung von hPcf11 an RNA durch UV-Vernetzung an radioaktiv markiertes rU₁₄ erbracht werden. In der Bindungsreaktion zur Einstellung des Gleichgewichts wurde 0,5 µM CF II aus einer MonoQ Fraktion mit 2 µM des radioaktiv markierten rU₁₄ inkubiert. Anschließend wurde Reaktion für eine Minute mit UV-Strahlung der Wellenlänge 253 nm bestrahlt und somit chemisch verknüpft. Das Protein-RNA Produkt wurde danach über eine Immunopräzipitation mittels polyklonalem anti-Pcf11 Antikörper vom Rest abgetrennt und über eine denaturierende SDS-PAGE analysiert. Dabei zeigte sich, wie in der Abbildung 2.14 erkennbar, dass ein spezifisches Signal für einen hPcf11-RNA Komplex bei etwa 170 kDa detektiert werden konnte. Zudem konnten einige niedermolekulare Signale abgetrennt werden. Dabei könnte es sich um C-terminale Pcf11-Fragmente handeln, da der Antikörper gegen den FEGP-Repeat gerichtet ist. Im Vergleich dazu zeigte die Kontrolle mit prä-Immunsereum (PI) nur schwache Hintergrundsignale. Auch wenn die Affinität des hPcf11 zu rU₁₄ nicht so hoch ist wie für rG₁₄, konnte hiermit eine eindeutige und direkte Assoziation des Proteins mit dem eingesetzten Oligonukleotid nachgewiesen werden. Zudem zeigte rU₁₄ eine gute Reaktionsrate in der UV-Vernetzung.

In PAR-CLIP-Experimenten mit yPcf11 wurden mögliche Bindungsstellen an Prä-mRNAs etwa 130 Nukleotide abwärts der Spaltstelle kartiert (Baejen u. a., 2017). Auch ChIP-Experimente in *S. cerevisiae* und *S. pombe* zeigten Bindung des yPcf11 an Sequenzen abwärts der Spaltstelle von prä-mRNA Substraten (Baejen u. a., 2017; Mayer u. a., 2012; Wittmann u. a., 2017).

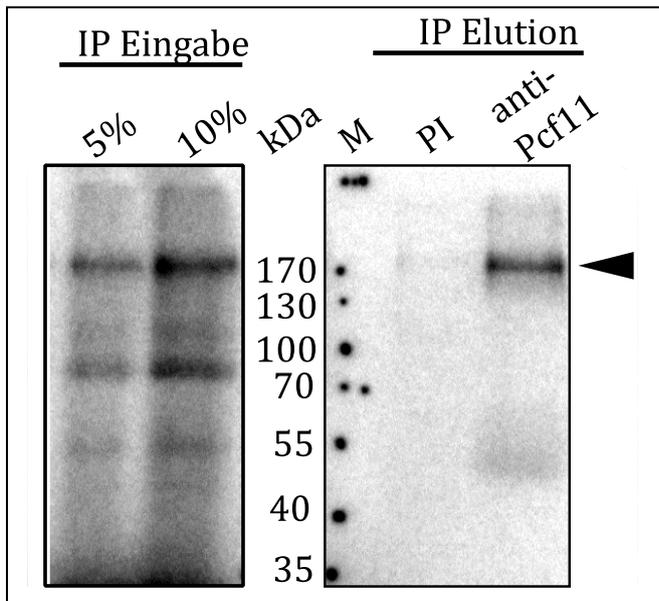


Abbildung 2.14 UV-Vernetzung und Antikörper-Affinitätsreinigung von CF II mit 5'-markierter rU_{14} -RNA. Ni-NTA gereinigtes CF II wurde nach einer Bindungsreaktion durch UV-Licht bei 254 nm an radioaktiv markierte RNA vernetzt. Das CF II wurde anschließend über einen anti-hPcf11-Antikörper affinitätsgereinigt (IP) und durch SDS-PAGE das gebundene Protein aufgetrennt. Die Detektion der RNA erfolgte über Phosphorimaging. Deutlich sichtbar ist die Anreicherung eines spezifischen Signals bei 170 kDa in der anti-Pcf11-Spur (schwarzer Pfeil) und eines schwächeren Signals bei 50 kDa. Dies entspräche ungefähr den Molekulargewichten von hPcf11 (172 kDa) und hClp1 (48 kDa). Aufgetragen wurden 50% der Elution. Präimmunsereum (PI) diente als Spezifitätskontrolle.

Auf Grundlage dieser Informationen, zusammen mit der Präferenz für G-reiche Sequenzen, konnte in der SV40 *late*-RNA ein Bereich abwärts der Spaltstelle ausgemacht werden, welcher diesen Kriterien entspricht. Diese wurde bereits in Analysen der SV40 *late*-RNA durch Qian und Wilusz (Qian und Wilusz, 1991) als notwendig für die Spaltung identifiziert.

2.6.2 hPcf11 bindet an eine Guanin-reiche Sequenz in der SV40 *late* RNA abwärts der Spaltstelle

Die Feststellung, dass hPcf11 im CF II an G-reiche RNA-Sequenzen mit mindestens 14 nt Länge bindet und die Beobachtung, dass auch die SV40 *late* RNA gebunden wurde, ergab die Frage nach der Bindestelle innerhalb der SV40 *late* RNA. Für die Suche nach der Bindestelle wurden zunächst durch Einsatz von α - $[^{32}\text{P}]$ -UTP intern radioaktiv markierte Fragmente der SV40 *late* RNA in einer *in-vitro* Transkription generiert und in Filterbindungsversuchen mit CF II als Substrat angeboten (Abbildung 2.15 A). Es zeigte sich eine etwa vierfach höhere Affinität des 3'-Fragmentes von nt 150–230 ($K_{50} = 1,3 \text{ nM}$) gegenüber des 5'-Fragmentes von nt 1–149 ($K_{50} = 4,2 \text{ nM}$). Unter Beachtung, dass das 3'-Fragment auch noch kürzer war, wurden aus diesem ~40 nt lange Fragmente abgeleitet. Diese Fragmente der SV40 *late*-RNA, genannt 150–190, 171–210 und

191–230, bildeten dabei Abschnitte mit jeweils 20 Nukleotiden Überlappung der Sequenz (s. Anhang Abbildung 9.1). Damit sollte einer Zerteilung einer möglichen Bindestelle vorgebeugt werden. In einem Filterbindungsversuch zeigte das 3'-Fragment 191–230 mit einem K_{50} von ~ 6 nM eine deutlich höhere Affinität gegenüber den anderen beiden Fragmenten (Abbildung 2.15 B). In dieser Nummerierung befindet sich U191 +15 Nukleotide abwärts der Spaltstelle in SV40 *late* (Abbildung 2.15 C, (Sheets u. a., 1987)). Der RNA-Abschnitt 190–230 zeigte auch in Gelshifts hohe Affinitäten an CF II ($K_{50} < 1$ nM) (Tabelle 2.4, s. Anhang Abbildung 9.2). Ausgehend davon wurden synthetische Oligonukleotide abgeleitet, welche zunehmend verkürzte Varianten dieser Sequenz darstellten und sich an der Nummerierung in Abbildung 2.15 C orientierten. Dabei zeigte sich, dass die 36 Nukleotide lange Sequenz 190–225 die höchste Affinität aufwies. Die Länge der RNA lässt hierbei jedoch mehrere mögliche Bindestellen zu.

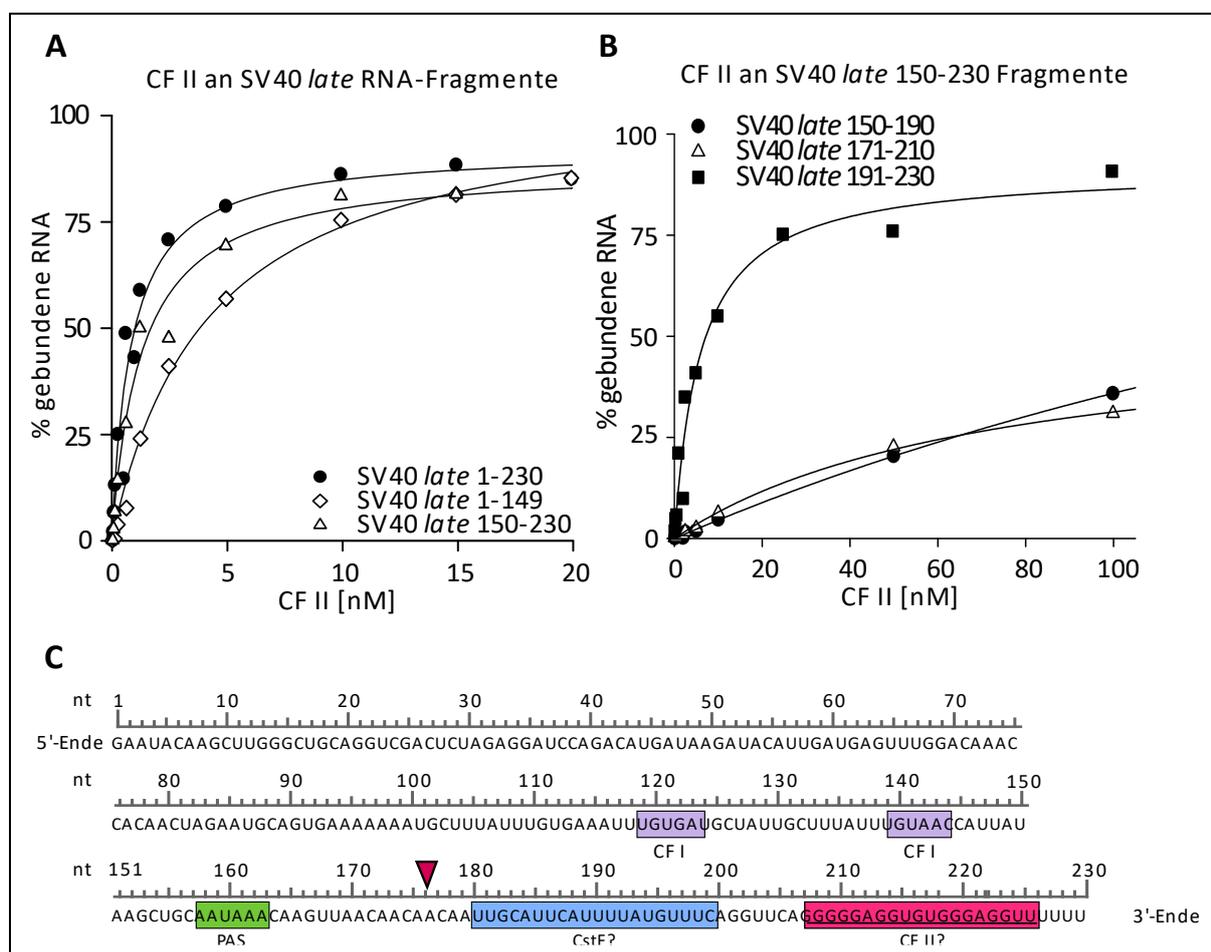


Abbildung 2.15 Filterbindungsversuche von CF II mit SV40 *late*-RNA und Übersicht der SV40 *late* RNA-Sequenz. **A**) Filterbindung von CF II an SV40 *late* ($K_{50} = 0,9$ nM) und an zwei SV40 *late*-Fragmente, genannt SV40 *late* 1–149 mit 150 nt Länge ($K_{50} = 4,2$ nM) und SV40 *late* 150–230 mit 81 nt Länge ($K_{50} = 1,3$ nM). **B**) Filterbindung von CF II an drei SV40 *late*-Fragmente des Abschnittes zwischen nt 150 und 230, welche je ~ 40 nt lang und je 20 nt überlappend in der Sequenz waren. SV40 *late* 191–230 zeigte die höchste Affinität zu CF II ($K_{50} = 6$ nM). SV40 *late* 171–210 ($K_{50} = 72$ nM) und SV40 *late* 150–190 ($K_{50} = 320$ nM) zeigten im Vergleich mindestens zehnfach schlechtere Affinitäten. **C**) Basensequenz der verwendeten SV40 *late* Substrat-RNA mit Markierung bekannter oder möglicher Bindesequenzen von 3'-Prozessierungsfaktoren. Der rote Pfeil markiert die Spaltstelle. Pink dargestellt ist die ermittelte, mögliche Guanin-reiche Bindestelle für CF II.

In Tabelle 2.4 zeigt sich, dass sich mit Verkürzung der Sequenz auf den G-reichen Bereich auch die Affinität verringerte. Die 14 Nukleotide lange Sequenz 208–221 wies mit einem apparenten K_{50} von 3,4 nM allerdings eine nur um den Faktor 10–15 schwächere Bindung auf als das 36-mer SV40 *late* 190–225 (s. Anhang Abbildung 9.2). Das 14-mer 208–221 lag damit nur knapp unter der ermittelten Affinität zu rG₁₄ mit 2,0 nM. Unter 14 Nukleotide verkürzte Sequenzen Länge zeigten, ähnlich der beschriebenen Oligo(G) Reihe, deutlich verringerte Affinitäten (nicht gezeigt). Somit scheint die SV40 *late* 208–221-Sequenz die beste Bindestelle innerhalb der SV40 *late* RNA zu sein.

Tabelle 2.4 Apparente Dissoziationskonstanten aus nativen Gelshifts von CF II an SV40 late-Sequenzabschnitte. Auflistung der getesteten SV40 late-Fragmente mit Kennzeichnung (unterstrichen), der durch Qian (Qian und Wilusz, 1991) beschriebenen, für die Spaltung essenzielle Sequenz, der SV40 late-RNA. Werte mit Standardabweichungen wurden hierbei aus mindestens drei Versuchen ermittelt. Die anderen sind Mittelwerte aus Doppelbestimmungen. Die markierte Sequenz stellte auch gleichzeitig die minimale Bindestequequenz des CF II dar.

Name	Sequenz	apparenter K_{50} [nM]
191–230	gggUUUAUGUUUCAGGUUCAGGGGGAGGUGUGGGAGGUUUUUU	0,4±0,2
190–225	UUUUAUGUUUCAGGUUCAGGGGGAGGUGUGGGAGGU	0,2±0,1
202–224	GGUUCAGGGGGAGGUGUGGGAGG	0,8
208–225	GGGGGAGGUGUGGGAGGU	2,6
208–224	GGGGGAGGUGUGGGAGG	5,2
208–221	GGGGGAGGUGUGGG	3,4

2.6.3 Lineare Sequenzen und nicht G-Quadrupel Strukturen werden durch hPcf11 gebunden

Alle bis hierhin getesteten Oligonukleotide, welche eine gute Affinität zu CF II aufwiesen, sind sehr G-reich. Prinzipiell können RNAs mit dem idealisierten Sequenzmotiv (GGN)₄ sogenannte G-Quadruplex-Strukturen (G₄) ausbilden. Dabei handelt es sich um eine Sekundärstruktur der RNA (Abbildung 2.16 A), welche nicht auf den klassischen von Watson und Crick beschriebenen Wasserstoffbrückenpaaren zwischen G und C oder A und U beruht. Hierbei bilden Guanine ein Quartett über Wasserstoffbrücken, welche einer Hoogsteen-Basenpaarung folgen. Beschrieben wurden diese Strukturen und deren Eigenschaften schon mehrfach vor allem in DNA-Sequenzen (Bochman u. a., 2012). Aber auch RNAs können G-Quadrupel ausbilden. Die RNAs bilden langsam sehr stabile G₄-Strukturen mit charakteristischen Eigenschaften. Strukturierte RNAs sind durch die Faltung kompakter als in ihrer linearen Form (Wang u. a., 2017). Das Guanin-Quartett bildet sich aus vier Abschnitten der RNA des Musters (GGN)₄ in einer planaren Ebene über die Hoogsteen-Basenpaarungen, wobei jede Wiederholung einen „Eckpfeiler“ darstellt. Je nach

Anzahl der aufeinander folgenden Guanin-Basen einer Wiederholung sind die Guanin-Quartette in mehreren Ebenen (Abbildung 2.16 A) gestapelt. Die Quartett-Struktur wird durch Kaliumionen zusätzlich stabilisiert. Interaktion mit anderen einwertigen Ionen hat offenbar keinen stabilisierenden oder destabilisierenden Effekt. Auch die Abwesenheit einwertiger Ionen destabilisiert die Struktur nicht. Als weiteres charakteristisches Merkmal zeigen G-Quadruplex-Strukturen in paralleler Faltung ein spezifisches Maximum in der Zirkulardichroismus-Spektroskopie (CD) bei etwa 265 nm (Tang und Shafer, 2006; Vorlíčková u. a., 2012). Unstrukturierte RNAs zeigen in der CD-Spektroskopie ein Maximum bei 272 nm, wodurch sie von G-Quadruplex-Strukturen unterschieden werden können. Die RNA-Struktur der G-Quadrupel kann auch für Proteininteraktionen eine Rolle spielen. Sie können zum einen die Assoziation RNA-bindender Proteine verhindern, welche lineare einzelsträngige RNA binden. Zum anderen können sie als Bindungsstellen für spezifisch G-Quadruplex bindende Proteine und Komplexe dienen (Wang u. a., 2017).

Die RNAs der Oligo-rG-Reihe länger als 10 Nukleotide, sowie die SV40 *late*-Fragmente und weitere genutzte G-reiche Oligonukleotide können, zumindest theoretisch, diese Strukturen inter- oder intramolekular ausbilden (Bedrat u. a., 2016). So wurde für weitere Studien zur möglichen Bindung an G-Quadruplex-Strukturen eine idealisierte 17-mer Sequenz aus der SV40 *late* 208–224 RNA abgeleitet (s. Schema Abbildung 2.15 C). Diese Sequenz „Wt2x“ stellt eine Wiederholung der G-reichen Region 208–215 dar, welche durch ein Uracil verbunden wurde (Sequenzen Tabelle 2.5). Außerdem wurde eine Variante mit verringertem G-Gehalt abgeleitet, welche allerdings noch immer die Fähigkeit zur G₄-Bildung besitzt ($\Delta 2x$). Eine weitere Sequenz ($\Delta\Delta 2x$) mit genauso stark verringertem Guanin-Gehalt, allerdings ohne Möglichkeit zur G₄-Bildung, diente als Kontrolle des G₄-bildenden Oligonukleotids $\Delta 2x$. Mit der Möglichkeit G₄-Strukturen anhand des charakteristischen Maximums bei 265 nm im CD-Spektrum von unstrukturierter RNA zu unterscheiden und den drei RNA-Oligonukleotiden Wt2x, $\Delta 2x$ und $\Delta\Delta 2x$ als Testsystem wurde die vorliegende *in vitro* Konformation dieser RNAs untersucht. Parallel dazu erfolgten in Gelshifts Bindungsstudien der RNAs an CF II.

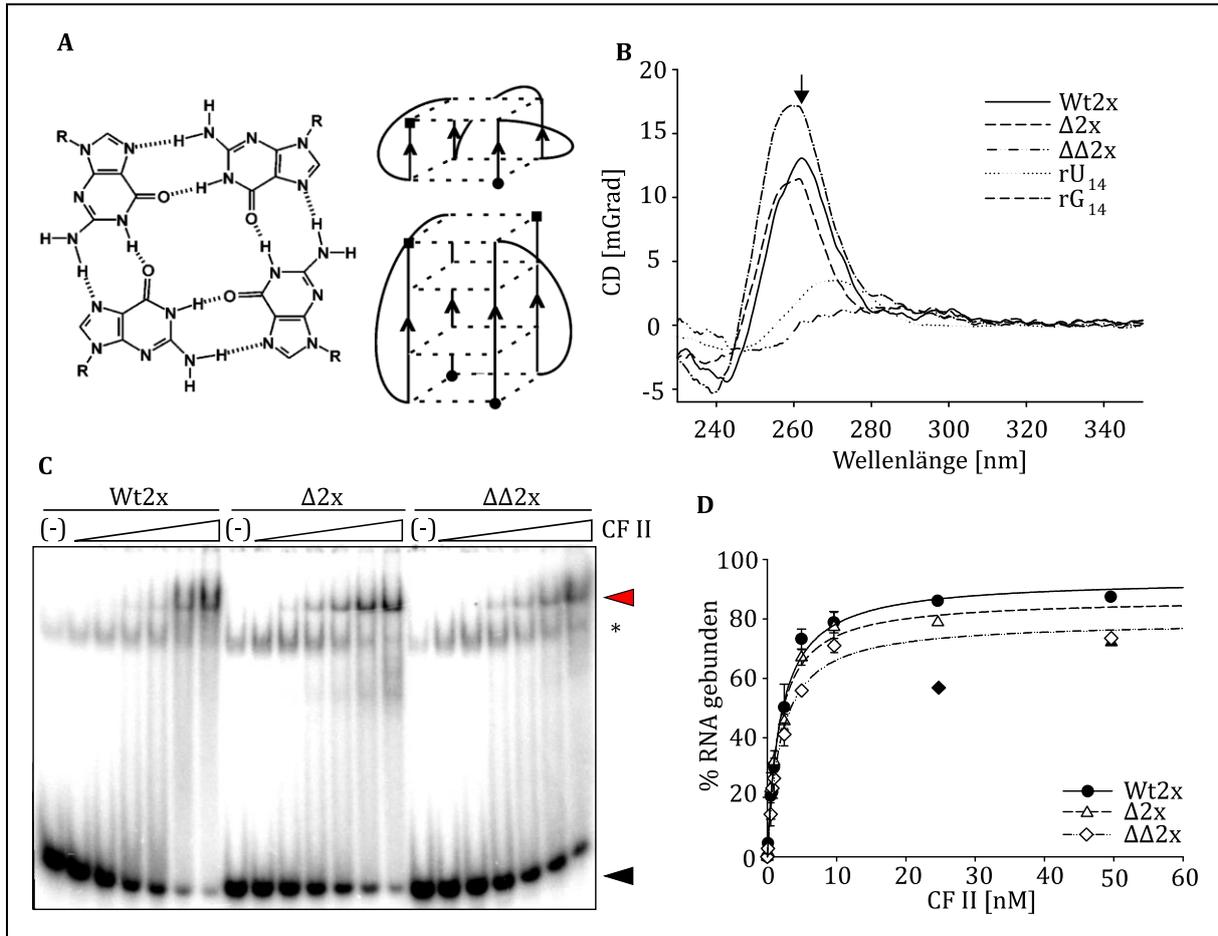


Abbildung 2.16 CD-Spektren der 17mer Oligonukleotide und Gelshifts

A) Schema einer möglichen Guanin-Quadruplex Hoogsteen-Basenpaarung und 3D-Struktur mit verschiedenen planaren Ebenen der gepaarten Basen, entnommen aus (Vorlíčková u. a., 2012). **B)** Spektren aus Zirkulardichroismus (CD)-Experimenten mit 2 μM an RNA-Oligomeren der idealisierten SV40 late RNA-Sequenzen Wt2x, $\Delta 2x$ und $\Delta\Delta 2x$. Als Vergleiche dienten das G₄-bildende rG₁₄-Oligomer und das unstrukturierte rU₁₄-Oligomer. Charakteristisch ist das G₄-Quadrupel Signal bei 262 nm (schwarzer Pfeil), deutlich zu sehen bei rG₁₄, Wt2x und $\Delta 2x$. Im Gegensatz dazu die unstrukturierte rU₁₄-RNA mit einem Maximum bei 272 nm, vergleichbar mit dem Signal der $\Delta\Delta 2x$ -RNA. Das Maximum der $\Delta 2x$ -RNA war im Gegensatz zum Wt2x-Maximum sensitiv gegenüber Erwärmung auf 95 °C. **C)** Nativer Gelshift von steigenden Konzentrationen CF II (0–10 nM) mit 0,1 nM der drei Oligomere Wt2x, $\Delta 2x$ und $\Delta\Delta 2x$. Die Affinität unterschied sich nicht sichtbar zwischen den drei RNAs und war unabhängig von G-Quadruplex-Potenzial oder Guanin-Gehalt. Der rote Pfeilkopf markiert den CF II-RNA-Komplex, der schwarze Pfeilkopf freie RNA. Der Stern kennzeichnet RNA-Sekundärstrukturen der nicht gebundenen RNA-Oligomere. **D)** Graphische Auftragung der Phosphorimager-Quantifizierung gebundener RNA des nativen Gelshifts aus C) mit hyperboler Anpassung der Datenpunkte.

CD-Spektren der RNA-Oligonukleotide wurden in RNA-Bindungspuffer mit 60 mM KCl zwischen 340 nm und 220 nm aufgenommen. Um die RNA-Konzentrationen möglichst gering zu halten und gleichzeitig ausreichende Signalintensität zu erreichen, wurde eine Küvette mit 1 cm Pfadlänge verwendet. Die Küvette und Messkammer wurden auf 20 °C temperiert. Eventuelle Sekundärstrukturen der RNAs wurden vor der Messung durch kurzes Erhitzen auf 95 °C und Abkühlung auf Eis vor der Messung aufgelöst. G₄-Strukturen blieben aufgrund ihrer Stabilität

dabei jedoch wahrscheinlich erhalten. Wie man in Abbildung 2.16 B erkennen kann, zeigten die RNA-Oligos Wt2x und $\Delta 2x$ im CD-Spektrum mit einem Dichroismus von +10 bis +15 m° bei 262 nm ein ähnlich starkes Maximum wie die Kontroll-RNA rG₁₄. Dies bestätigte, dass die RNA-Oligos Wt2x und $\Delta 2x$ bei Raumtemperatur und den hohen RNA Konzentrationen (2 μ M) eine stabile G₄-Struktur ausbildeten. Das RNA-Oligo $\Delta\Delta 2x$ hingegen zeigte, ähnlich der Kontroll-RNA rU₁₄, unter den gegebenen Bedingungen ausschließlich ein für lineare RNA charakteristisches Maximum bei 272 nm. Somit wurde die Erwartung bestätigt, dass die RNA $\Delta\Delta 2x$ keine G-Quadruplex-Strukturen ausbilden kann.

Tabelle 2.5 Apparente Dissoziationskonstanten aus nativen Gelshifts von CF II an die 17-mer Varianten SV40 late 208–224 Sequenz.

Die Sequenz Wt2x stellt eine exakte Wiederholung der ersten acht Nukleotide dar, wobei die Wiederholung durch ein Uracil verbunden wird. Die Varianten $\Delta 2x$ und $\Delta\Delta 2x$ besitzen einen um das gleiche Verhältnis verringerten Guanin-Gehalt. Das 17-mer $\Delta\Delta 2x$ besitzt im Gegensatz zu den anderen 17-meren keine Möglichkeit G₄-Strukturen auszubilden. Die Affinitäten sind Mittelwerte aus Doppelbestimmungen.

Name	Sequenz	apparenter K ₅₀ [nM]	G ₄ -Struktur
208–224	GGGGGAGGUGUGGGAGG	5,2	ja
Wt2x	GGGGGAGGUGgGGGAGG	2,0	ja
$\Delta 2x$	GGGaaAGGUGgGaaAGG	1,9	ja
$\Delta\Delta 2x$	GaGaGAGGUGaGaGAGG	2,3	nein

Aus der Tabelle 2.5 und den zugehörigen Graphen in Abbildung 2.16 D ist ersichtlich, dass die drei getesteten Oligonukleotide ähnlich gut mit einem K₅₀ um 2 nM an CF II binden. Zudem wird die Affinität trotz abnehmendem Guanin-Gehalt nicht nennenswert reduziert und entspricht annähernd der Affinität von rG₁₇ mit 1,9 nM. Die Affinitäten der G₄-bildenden Oligos Wt2x und $\Delta 2x$ unterschieden sich zudem nicht von der Affinität zum unstrukturierten Oligo $\Delta\Delta 2x$. Somit ist anzunehmen, dass alle drei RNAs in der unstrukturierten, einzelsträngigen Form gebunden wurden und keine Bindung an eine definierte Struktur stattfand. Alle drei RNAs migrierten im EMSA auf etwa der gleichen Höhe und es waren keine schnelleren Populationen zu erkennen (Abbildung 2.16 C).

All diese Beobachtungen lassen die Schlussfolgerung zu, dass hPcf11 die RNA vermutlich in einzelsträngiger Form und nicht als G₄-Struktur bindet. Weiter kann aus den Gelshifts abgeleitet werden, dass bei den im Gelshift verwendeten RNA-Konzentrationen G₄-Strukturen vermutlich nicht stabil vorliegen. Eine weitere Beobachtung ist, dass auch Sequenzen mit verringertem

Guanin-Gehalt ähnlich affin gebunden wurden. Auch ein Uracil-reiches Oligonukleotid mit geringem Guanin-Gehalt, abgeleitet aus CstF-SELEX Experimenten (Beyer et al 1997), zeigt eine gute Bindung mit einem K_{50} von etwa 4,4 nM (Tabelle 2.6). Oligonukleotide gänzlich ohne Guanin-Basen waren hingegen Substrate mit geringer Affinität (s. Homopolymere Tabelle 2.3).

2.6.4 Selektionsversuche mit randomisierten RNA 14-mer Sequenzen (Bind'n'Seq)

Die bisher verwendeten Methoden und Oligonukleotide waren für die gezielte Suche nach der Bindungssequenz innerhalb der SV40 *late* RNA ausgewählt. Die Ergebnisse der 17-mer Oligonukleotide Wt2x, Δ 2x und $\Delta\Delta$ 2x zeigten jedoch, dass auch RNAs mit verringertem Guanin-Gehalt gebunden werden. Es sollte in der Folge nun ein mögliches Bindungsmotiv für CF II innerhalb der RNA mittels eines ungerichteten statistischen Verfahrens gesucht werden. In Kooperation mit Thomas Treiber, Nora Treiber und Gunter Meister an der Universität Regensburg wurde versucht, Bindungsmotive für hPcf11 zu ermitteln. Hierzu wurden Selektionen aus zufälligen RNA Populationen (Bind'n'Seq) (Lambert u. a., 2015) durchgeführt. Aus einer Population zufälliger 14-mer RNA-Sequenzen mit konstanten GUUU 3'-Enden wurden in einer ersten Runde, die am besten bindenden Oligonukleotide selektiert. Dazu wurde zunächst das vollständige CF II in einer RNA-Bindungsreaktion angeboten und anschließend über den His-tag an Clp1 zusammen mit den gebundenen RNAs wieder isoliert. Mittels Phenol/Chloroform-Isolation wurden die RNAs wieder gereinigt. Als zweiter Ansatzpunkt für eine Selektion von RNAs wurde auch das isolierte hPcf11 Δ N1184 Fragment genutzt. Das Fragment wurde nach der RNA-Bindungsreaktion über Filterbindung isoliert. Die daran gebundene RNA wurde in einer anschließenden denaturierenden Elution wieder aus der Bindungsreaktion gereinigt und isoliert. Die gereinigten RNA-Oligonukleotide wurden an Adapter legiert, revers transkribiert, amplifiziert und über T7 RNA-Polymerase wieder in RNA transkribiert (Schema in Abbildung 2.17 A). Diese Population enthielt durch die T7-Transkription nun zusätzlich ein konstantes GGG 5'-Ende. Diese vorselektierte Population wurde in einer zweiten Selektionsrunde abermals als Bindungspartner angeboten und die gebundenen RNAs wurden wie zuvor aufgearbeitet. Die erhaltenen Sequenzen wurden nach beiden Durchläufen in einem Next Generation Sequenzierungsverfahren (NGS) ermittelt und ihre Häufigkeit quantifiziert (Schema in Abbildung 2.17 A).

Auf Grundlage einer ersten statistisch ausgewerteten Selektion mit CF II als Bindungspartner wurde die am häufigsten gebundene Sequenz dieser Analyse für weitere Bindungsstudien ausgewählt („BS2“, Tabelle 2.6). In der dargestellten Sequenz sind die methodisch bedingten, nicht zufälligen Sequenzen der RNAs unterstrichen. Die ermittelte Affinität des CF II zur BS2-RNA war mit 11 nM etwa 3–5-mal geringer als für den SV40 *late* 208–224-Abschnitt oder als für das

Wt2x-Oligonukleotid. Eine um das konstante 3'-Ende des BS2-Oligonucleotids verkürzte Sequenz (BS2-Δ1), zeigte keine Verringerung der Affinität. Das weiter um das konstante GGG 5'-Ende gekürzte BS2-Δ3 zeigte im Gegensatz eine stark verringerte Affinität zu CF II. Die Erhöhung des K_{50} um etwa das 10-fache auf 105 nM zeigte die starke Beteiligung dieses Guanin-reichen Abschnittes an der Bindung. Substituierte man diesen Abschnitt durch ein CCC 5'-Ende (BS2-Δ2), so wurde auch hier die Affinität stark verringert. Somit konnte gezeigt werden, dass nicht die Länge des RNA-Liganden, sondern die Guanin-reiche 5'-Sequenz wichtig für die Bindung des BS2-Δ1 Oligos ist.

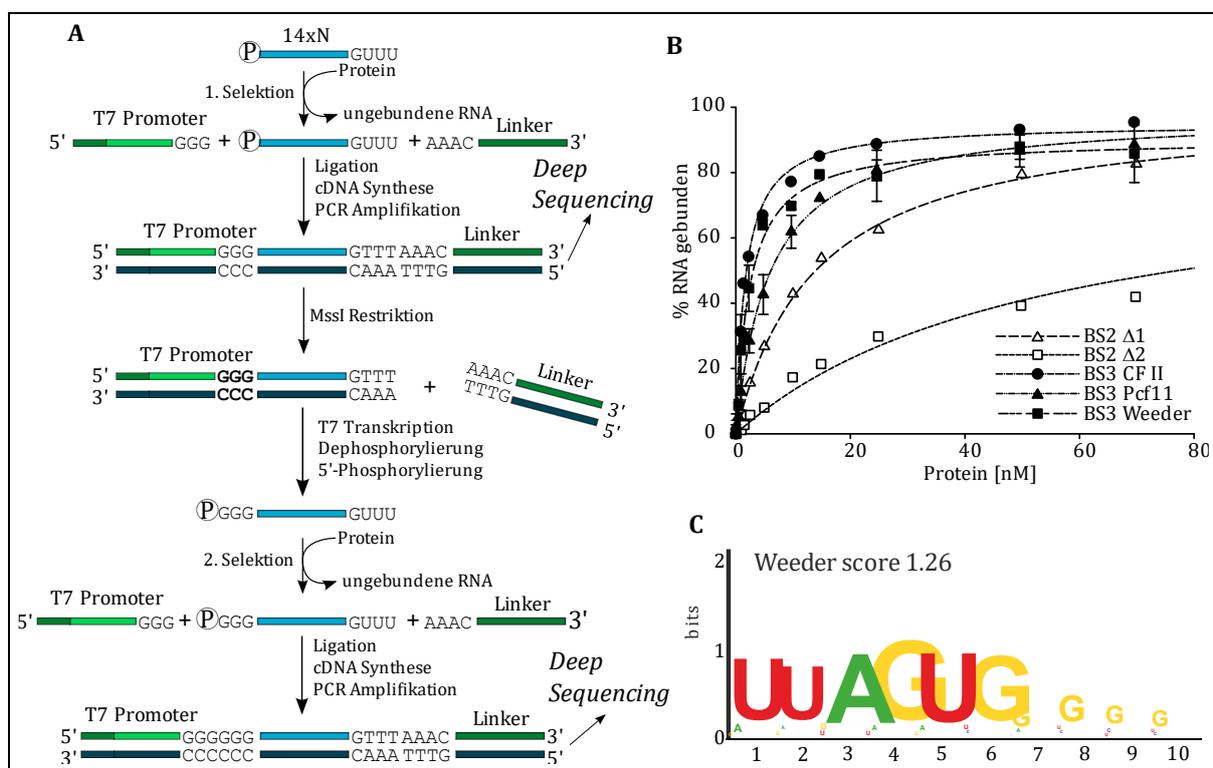


Abbildung 2.17 Bind'n'Seq (BS) Selektionsversuche und Ermittlung einer Konsensussequenz der CF II-RNA-Bindung

A) Schema des Bind'n'Seq Arbeitsablaufes durch Thomas Treiber (nach Treiber u.a. 2017), welcher die Anreicherung von randomisierten 14-meren mit CF II und hPcf11 ΔN1184 durchführte. Aus einem Pool an RNA-18-meren mit zufälliger 14-mer-Sequenz erfolgt die Bindung an das Zielprotein. Die gebundenen RNAs wurden isoliert, revers transkribiert und die cDNAs Sequenziert. Aus diesem selektierten Pool erfolgt eine T7-Transkription und eine erneute Selektion mit dem Zielprotein. Am Ende erfolgte die statistische Auswertung der gebundene Sequenzen. **B)** Auftragung der quantifizierten Daten von RNA-Bindung an CF II in nativen Gelshifts mit ausgewählten Oligonukleotiden, welche in Selektionsversuchen angereichert wurden. Die Bindungskurven resultieren aus einer hyperbolen Anpassung der Datenpunkte. Die Affinitäten zwischen den Oligonukleotiden der BS3-Reihe waren vergleichbar. Die Substitution von „GGG“ am 5'-Ende des BS2 Δ1-Oligomers zu „CCC“ im BS2 Δ2-Oligomer führte zu einem starken Affinitätsverlust. Die Gemeinsamkeit aller RNAs ist ein Guanosingehalt $\geq 50\%$. **C)** Sequenz-Konsensus der angereicherten RNAs aus Bind'n'Seq-Versuchen mit Pcf11 ΔN1184, ermittelt durch T. Treiber mit „Weeder“ (Zambelli u. a., 2014).

Zur weiteren Bestätigung der oben genannten Beobachtungen wurden die häufigsten gebundenen RNA-Sequenzen eines weiteren Selektionsversuchs, ohne die konstante 3'-Sequenz, synthetisiert. Die in Tabelle 2.6 „BS3 CFII“ und „BS3 Pcf11“ genannten RNAs sind dabei aus den Selektionsversuchen mit CF II und hPcf11 ΔN1184 hervorgegangen. Das letzte Oligo dieser Reihe „BS3 Weeder“ wurde aus einer mit der Software „Weeder“ (Zambelli u. a., 2014) abgeleiteten Konsensussequenz erstellt (Abbildung 2.17 C). Dieser Konsensus wurde dabei aus einer statistischen Analyse der Selektion mit hPcf11 ΔN1184 erstellt. Man sieht hier eindeutig die Anreicherung von Guaninen in der Sequenz. Alle drei der BS3-Oligos stellten bessere Liganden als die BS2-Oligos dar (Tabelle 2.6). Mit einer Affinität von 1,7 nM befand sich das BS3 CF II-Oligonukleotid sogar in dem gleichen Bereich wie rG17 und Wt2x Oligonukleotide. Allerdings zeigte der Guanin-Gehalt von nur 50%, dass offensichtlich nicht nur ein hoher Guanin-Gehalt ausschlaggebend für die Affinität war. Das ähnlich Guanosin-reiche BS3 Pcf11-Oligo band um den Faktor 2,5 weniger affin an CF II. Die Konsensussequenz BS3 Weeder band mit 2,4 nM etwa so gut wie das in der Sequenz sehr ähnliche Oligonukleotid SV40 *late* 208–225 oder die etwas weniger G-reiche RNA ΔΔ2x (Vergleich in Tabelle 2.4 und Tabelle 2.5).

Tabelle 2.6 Apparente Dissoziationskonstanten aus nativen Gelshifts von CF II an selektierte Sequenzen aus der Anreicherung zufälliger RNA-Populationen (Bind'n'Seq). Aufgelistet sind verschiedene Oligonukleotide und deren Varianten, welche aufgrund ihrer hohen Anreicherung in Bind'n'Seq-Experimenten ausgewählt wurden. Unterstrichene Sequenzabschnitte sind konstante Sequenzen, welche aus der verwendeten Bind'n'Seq-Methode resultierten. Auffällig ist der Abfall der Affinität der getesteten RNAs, sobald das 5'-„GGG“ in BS2-Δ1 ausgetauscht (BS2-Δ2) oder deletiert (BS2-Δ3) wurde. Affinitäten mit Standardabweichung resultieren aus n=3.

Name	Sequenz	apparenter K ₅₀ [nM]
CstF SELEX 2a	UGUGUUUUUAUUGUGU	4,4
BS2	<u>GGGUUU</u> AAGUGGGUCAC <u>GUUUAAAC</u>	11
BS2-Δ1	<u>GGGUUU</u> AAGUGGGUCAC <u>G</u>	13
BS2-Δ2	CCCUUU AAGUGGGUCAC <u>G</u>	54
BS2-Δ3	UUUAAGUGGGUCAC <u>G</u>	105
BS3 CFII	<u>GGGUU</u> ACGCGUGAGGAC <u>G</u>	1,7
BS3 Pcf11	<u>GGGUUU</u> AGUGCGCACG <u>G</u>	5,9±0,9
BS3 Weeder	<u>GGGUUU</u> AGUGGGGGG <u>G</u>	2,4±0,6

Zusammenfassend ließ sich aus den Selektionsversuchen keine eindeutige Bindungssequenz ableiten. Gemeinsam haben die guten Binder einen hohen Guanin-Gehalt. Ausgehend von der Affinität bildete die Region 208–211 aus der SV40 *late*-RNA offenbar eine annähernd ideale Bindungsstelle für CF II. Außerdem benötigt CF II mit 14 Nukleotiden Bindungslänge eine ungewöhnlich lange Bindungssequenz der RNA. So kann vermutet werden, dass hPcf11, ähnlich wie andere RNA bindende Prozessierungsfaktoren, kein eindeutiges Motiv bindet, sondern eher Präferenzen in der Basenzusammensetzung der Bindungssequenz und in Länge der Sequenz hat.

2.6.5 Humanes Pcf11 bindet RNA hauptsächlich mittels seiner C-terminalen Zinkfinger-Domänen an die RNA

Wie in den vorigen Abschnitten gezeigt, bindet CF II an G-reiche Sequenzen. Das humane Clp1 allein bindet hingegen kaum an RNA (s. Abbildung 2.13 A). Hinzu kommt der beobachtete spezifische *Crosslink* von hPcf11 mit rU14, welcher einen direkten Kontakt von RNA und hPcf11 impliziert. Also sollte die RNA-bindende Aktivität des CF II durch hPcf11 vermittelt werden. Tatsächlich ist für das Hefe-Homologe eine unspezifische RNA-Bindung der CID gezeigt (Hollingworth u. a., 2006; Zhang, 2005). Zudem besitzt das Pcf11 im C-terminalen Bereich die schon beschriebenen Zinkfingerdomänen. Diese können Protein-Protein- und auch Protein-RNA- oder Protein-DNA-Wechselwirkungen vermitteln. Für yPcf11 konnte zwar RNA-Bindung, aber keine Spezifität der Zinkfinger beobachtet werden (Guéguéniat u. a., 2017). Für das humane Pcf11 lag es nahe, die beiden Zinkfinger auch gezielt auf RNA-Bindung hin zu untersuchen. Dazu wurden Bindungsreaktionen an RNA mit den durch C. Tüting erstellten verkürzten hPcf11-Varianten durchgeführt (Schema Abbildung 2.18 A) und eine Auswertung mittels EMSA-Versuchen und Autoradiographie.

In der Abbildung 2.18 B und C sieht man, dass die angebotene RNA SV40 190–225 mit einem K_{50} von 0,3 nM am besten durch Wildtyp-CF II (Wt) gebunden wurde (s. Tabelle 2.7). Ohne die ersten 499 Aminosäuren des hPcf11, welche die CID und eine polare Region umfassen, wurde die RNA durch CF II $\Delta N499$ mit einer Affinität von 1,4 nM gebunden. Dies war in etwa ein Faktor fünf schwächer als durch Wildtyp-CF II (Abbildung 2.18 D, Tabelle 2.7). Das deutet auf eine Beteiligung der CID an der RNA Bindung hin, ähnlich der beschriebenen Bindungseigenschaft der CID des yPcf11. Wie schon in Abbildung 2.9 gezeigt, war hPcf11 $\Delta N499$ in CF II aber ausreichend, um die Spaltung zu aktivieren und die Aktivität des CF II zu rekonstituieren.

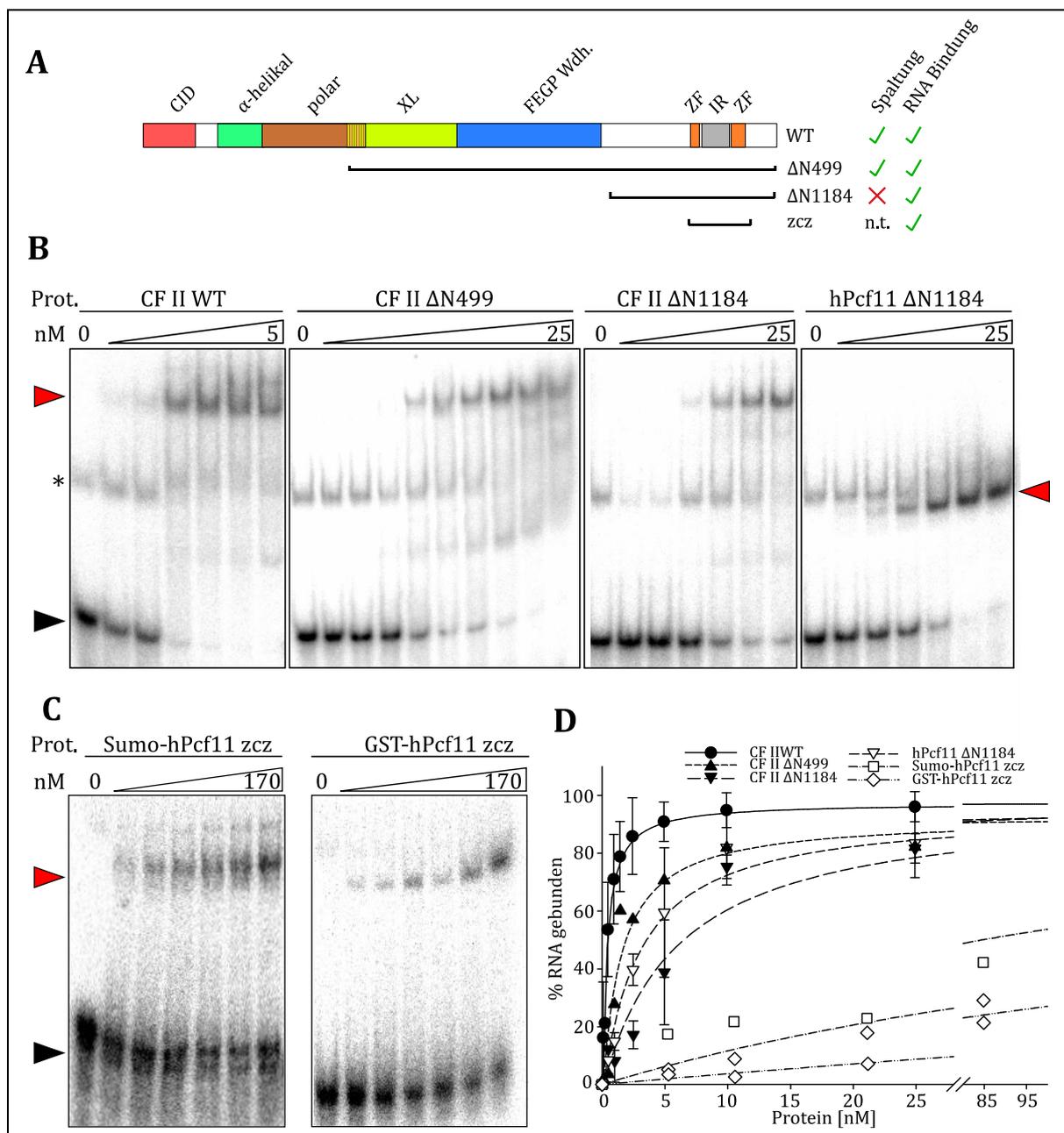


Abbildung 2.18 hPcf11 bindet RNA über die Zinkfinger Domänen

A) hPcf11-Schema mit Domänen und definierten Regionen wie zuvor dargestellt. CID – CTD Interaktionsdomäne (14–142); XL – hClp1 crosslink Region (500–770); FEGP – Region mit Motivwiederholung von FEGP (770–1124); ZF – Zinkfingerdomäne 1343–1368 und 1443–1478; IR – hClp1 Interaktionsregion (1371–1439). Die unter dem Schema liegenden Balken markieren die verwendeten verkürzten hPcf11-Varianten. Daneben aufgelistet sind zwei Kurzübersichten zur Spaltaktivität und RNA-Bindung der aufgelisteten hPcf11-Varianten (n.t. – nicht getestet). **B)** Native Gelshifts von 0,1 nM SV40 late 208–221 RNA-Substrat und steigenden Konzentrationen an CF II-Varianten, wie angegeben. Gut sichtbar ist der zunehmende Abfall der Affinität zum RNA-Substrat von CF II WT (WT-hPcf11) mit zunehmender Verkürzung der hPcf11-Varianten ΔN499 und ΔN1184 in CF II. Vergleich mit isoliertem hPcf11 ΔN1184. Ein roter Pfeil markiert den Protein-RNA Komplex, ein schwarzer Pfeil freie RNA und der Stern kennzeichnet RNA-Sekundärstrukturen. **C)** Nativer Gelshift mit 0,5 nM SV40 late 191–230-RNA und dem hPcf11-Fragment „zcz“ (Zinkfinger-Clp1 Interaktion-Zinkfinger; AS 1339–1506) als Fusion mit zwei verschiedenen Affinitätspeptiden in den angegebenen Konzentrationen. **D)** Bindungskurven der in-Gel quantifizierten RNA-Bindung der verschiedenen CF II und hPcf11 Varianten mit hyperbolischer Anpassung. Erkennbar ist die abnehmende Affinität zur RNA mit zunehmender Verkürzung des hPcf11. Die RNA-Bindung bleibt jedoch erhalten. Standardabweichungen aus $n \geq 3$.

Eine noch kürzere Variante des hPcf11 in CFII, hPcf11 Δ N118, zeigte verglichen mit CF II Δ N499 eine zusätzliche Verschlechterung der Affinität um den Faktor vier. Im Vergleich fällt auf, dass isoliertes hPcf11 Δ N1184 allein offenbar etwas besser gebunden hat als in der entsprechenden CF II-Variante (CF II Δ N1184). Ob das einen wirklichen Unterschied in der Affinität zwischen hPcf11 allein im Vergleich zum Komplex mit hClp1 darstellte, oder auf Anteile inaktiven Proteins in der Präparation hindeutete, ist offen. Dies bestätigte jedoch, dass hClp1 nicht nennenswert an der Bindung der SV40 *late*-RNA beteiligt war. Die CFII Δ N1184-Variante konnte, im Gegensatz zur längeren CF II Δ N499 Variante, die Spaltreaktion vermitteln (s. Abbildung 2.9). Man kann an den Bindungsdaten sehen, dass der Verlust der Spaltaktivität nicht auf einem Verlust der RNA-bindenden Aktivität des hPcf11 beruht

Tabelle 2.7 Apparente Dissoziationskonstanten aus nativen Gelshifts mit verkürzten hPcf11-Varianten mit SV40 late 190–225-RNA als Bindungspartner.

Parameter der hyperbolischen Anpassung der quantifizierten Bindungsdaten aus Abbildung 2.18 D. Die Deletion der N-terminale CID des hPcf11 im CF II-Komplex führt zu einem Abfall der Affinität zur RNA etwa um den Faktor 10. Weitere Verkürzungen sind nur etwas schwächer affin. Die ,zcz'-Fragmente binden mit etwa 300 bis 1000-fach geringerer Affinität an das RNA-Substrat.

Protein	K ₅₀ [nM]
CF II Wt	0,2±0,1
CF II Δ N499	1,4
CF II Δ N1184	5,9
hPcf11 Δ N1184	2,9
Sumo-hPcf11 zcz	~70
GST-hPcf11 zcz	~230

Abschließend wurde eine stark verkürzte hPcf11-Variante (hPcf11 zcz, s. Schema Abbildung 2.18 A) auf Bindung an RNA getestet. Diese besteht nur noch aus den Zinkfingerdomänen und der dazwischen liegenden hClp1 Interaktionsregion (AS 1339–1506). Dazu wurde von C. Tüting gereinigtes Protein genutzt. Aus Gründen der Stabilität und Reinigung wurde dieses Fragment zweimal unabhängig voneinander mit zwei verschiedenen N-terminalen Fusionspeptiden exprimiert (Glutathion-S-Transferase (GST); *Small Ubiquitin like Molecule* mit Hexahistidin-Peptid (His-SUMO)). Die Reinigung erfolgte über die entsprechenden Affinitätsmatrices mit anschließender MonoQ Säule. Die RNA-bindende Aktivität konnte auch hier in beiden Fällen über beide Säulen den Fusionspeptiden zugeordnet werden (Daten nicht gezeigt, C. Tüting). Mit den entsprechenden MonoQ-Fractionen wurden dann Bindungsexperimente analog zu den vorigen EMSA-Versuchen durchgeführt. In der Abbildung 2.18 C und der Tabelle 2.7 kann man deutlich

erkennen, dass beide Fusionsproteine das SV40 *late*-RNA-Fragment 190–225 mit deutlich verringerter Affinität binden. So war aufgrund der begrenzten Proteinkonzentrationen keine Titration bis in die Sättigung möglich. Daher konnten die Dissoziationskonstanten nur extrapoliert werden. Sie waren in der Anpassung um mindestens den Faktor 200 gegenüber dem Wildtyp-CF II erhöht.

Der Unterschied in den Affinitäten kann auch mathematisch über die Berechnung der Gibbs-Energie im Bindungsgleichgewicht beschrieben werden. Über die Formel:

$$\Delta G^\circ = -R * T * \ln(k)$$

$$\Delta G^\circ = -8,314 \frac{J}{mol * K} * 298,15 K * \ln\left(\frac{K_{50}}{c_0}\right)$$

... können diese berechnet werden. Der Unterschied des ΔG° von einem K_{50} um 0,3 nM (54,35 kJ/mol) zu einem ΔG° des K_{50} um 70 nM (40,84 kJ/mol) entspricht als $\Delta\Delta G$ etwa 13,5 kJ/mol. Dies entspricht in etwa der freien Energie einer Wasserstoffbrückenbindung mit einem ΔG von 10–20 kJ/mol. Alles in allem kann man so auch abschätzen, dass etwa 70% der freien Bindungsenergie durch die beiden Zinkfinger vermittelt werden.

2.7 Expression und Reinigung des *Cleavage stimulatory Factor* (CstF) und des *Cleavage Factor I* (CF I)

Für die Rekonstitution der Spaltungsreaktion wurden neben CF II auch weitere Proteine in Sf21-Insektenzellen exprimiert und gereinigt. Hierzu wurde zunächst der *Cleavage stimulatory Factor* (CstF) durch das baculovirale System exprimiert, wobei die drei Untereinheiten auf einem einzelnen Virus codiert waren. Das Heterohexamer CstF, bestehend aus je zwei Kopien der Untereinheiten 77k, 64k und 50k, wurde zunächst mittels des N-terminalen 8x His-Fusionspeptides an CstF-77k in einer IMAC gereinigt (Abbildung 2.19 A). Die Elutionsfraktionen E2 und E3 wurden vereint (~12 mg Protein), und gegen einen Puffer mit 50 mM KCl dialysiert. Zwei Drittel davon (~9 mg) wurden auf eine ResourceQ-Säule (1 ml) geladen und über einen KCl-Salzgradienten bei etwa 100–150 mM KCl (Abbildung 2.19 B, Fraktion #11–16) eluiert. Deutlich sieht man im Vergleich der beiden Säulen, dass die 64k Untereinheit in der IMAC unterrepräsentiert im Vergleich zu den anderen Untereinheiten vorlag. Der Überschuss der beiden anderen Untereinheiten konnte aber durch den Anionenaustauscher vom vollständigen Komplex abgetrennt werden. Die CstF-Hauptfraktion (#12) zeigte in einer Gelfiltration einen homogenen Peak und keine Aggregation. Allerdings lag das scheinbare Molekulargewicht bei etwa 2 MDa und somit, wie im Falle des CF II, weit oberhalb des erwarteten Molekulargewichts von 389 kDa. Eine Analyse mittels Rechtwinkel-Lichtstreuung im Anschluss an eine Gelfiltration

zeigte ein natives Molekulargewicht um 388 kDa und korrespondierte damit in dem erwarteten Bereich des Hexamers mit 389 kDa (s. Anhang Abbildung 9.15). Somit ist das Protein kein Aggregat, sondern vermutlich nur stark abweichend zur idealen globulären Struktur, von der bei der Kalibrierung in der Gelfiltration ausgegangen wird. Des Weiteren wurde die Funktionalität des gereinigten Proteins im Hinblick auf die RNA-Bindung getestet. Dazu wurde aus SELEX-Ergebnissen (Beyer u. a., 1997) eine ideale Sequenz abgeleitet und als Bindungspartner angeboten (Abbildung 2.19 F). Dieses synthetische RNA-16-mer, genannt CstF SELEX 2a, abgeleitet aus der Konsensus Sequenz des „Pool 2a“ (Beyer u. a., 1997) besitzt dabei die Möglichkeit, beide CstF-64k Kopien zu kontaktieren und den Komplex somit als Ganzes zu binden. Die hyperbole Anpassung der RNA-Bindungsversuche beruhte somit auf der Vereinfachung eines RNA-Bindungsereignisses pro CstF-Hexamer. In den Abbildung 2.19 F und G erkennt man deutlich die hochaffine Bindung des Hexamers an das in Filterbindungsversuchen angebotene 5'-markierte RNA-Oligonukleotid mit einer apparenten Dissoziationskonstanten von etwa 1,7 nM.

Zudem zeigt die Linearisierung der Auftragung im Lineweaver-Burk-Graphen in Abbildung 2.19 G, dass die Annahme eines Bindungsereignisses pro CstF-Molekül mit dem Modell vereinbar ist. Die Präparation wurde für die Komplementation des fraktionierten nukleären Extraktes verwendet.

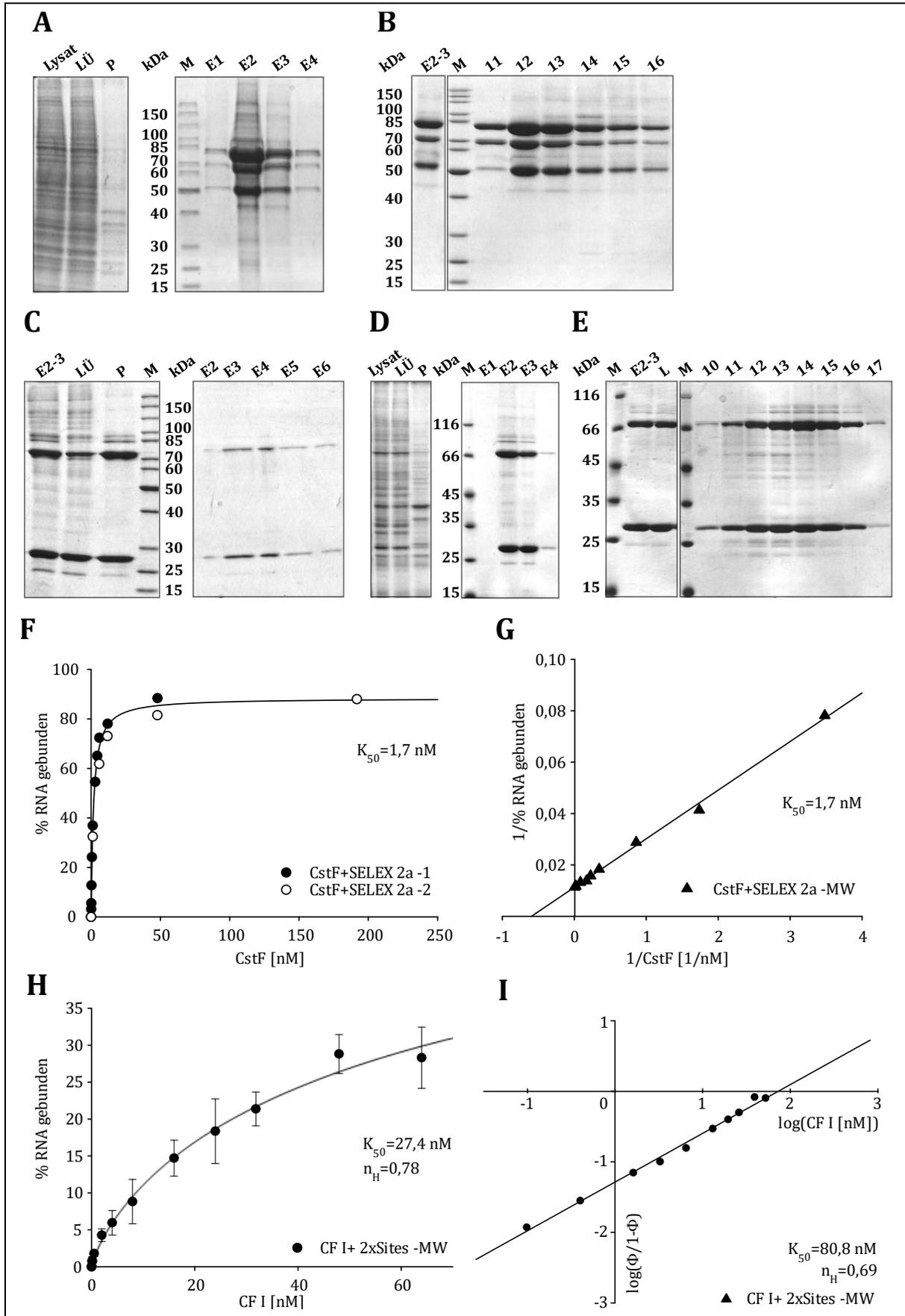


Abbildung 2.19 Reinigung und RNA Bindung von CstF und CF I

A) Coomassie-gefärbtes SDS-PA-Gel der Ni-NTA-Affinitätsreinigung von CstF (His-CstF77k) aus Sf21-Expression. Aufgetragen sind von links nach rechts: M – Unstained Protein Ladder; Lysat – Zellaufschluss; LÜ – löslicher Überstand; P – Pellet; E1–4 – Ni-NTA-Elutionsfraktionen. **B)** Coomassie-gefärbtes SDS-PA-Gel der Reinigung der vereinten CstF-Ni-NTA-Fraktion E2 und E3 über eine MonoQ-Säule mit einem 75–1000 mM KCl-Gradienten. Gut sichtbar ist die Abtrennung von nicht stöchiometrischen Protein-Anteilen in den frühen Fraktionen. **C)** Coomassie-gefärbtes SDS-PA-Gel der Strep-Tactin-Affinitätsreinigung von CF I aus Sf21-Expression (Strep-CF I 68k). Beschriftung sind wie zuvor, E2–6 sind D-Desthiobiotin Elutionen von der Strep-Tactin-Matrix. **D)** Coomassie-gefärbtes SDS-PA-Gel der Ni-NTA-Affinitätsreinigung von CF I (His-CF I 25k). **E)** Coomassie-gefärbtes SDS-PA-Gel der Reinigung der vereinten CF I Ni-NTA-Fractionen E2 und E3 über eine MonoQ-Säule mit einem Gradienten von 160–1000 mM KCl. Alle Gelausschnitte aus A) bis E) sind jeweils aus einem Gel zusammengefügt. **F)** Filterbindungsversuche des gereinigten CstF-Komplexes (MonoQ #12) an ein radioaktiv markiertes RNA 15-mer (SELEX 2a). Die Bindungskurve wurde ermittelt durch hyperbole Anpassung an zwei Einzelbestimmungen mit großer Übereinstimmung der Bindungsdaten und daraus abgeleitetem apparentem K_{50} . **G)** Lineweaver-Burke-Linearisierung der gemittelten Daten aus F) mit linearer Regression und abgeleitetem K_{50} . Der lineare Verlauf der Datenpunkte bestätigt die Annahme, dass ein RNA-Bindungsereignis pro Komplex stattfand. **H)** Filterbindungsversuche des Strep-Tactin-gereinigten CF I (E3) an ein radioaktiv markiertes RNA 19-mer (2xSites). Die Bindungskurve wurde ermittelt durch pseudo-hyperbole Anpassung nach Hill an die Mittelwerte aus Dreifachbestimmung. Daraus abgeleitet wurde der K_{50} und der Hill-Faktor n_H . **I)** Hill-Linearisierung der Daten aus H) für CF I zeigt einen linearen Verlauf der Datenpunkte. Die lineare Regression ergab einen Anstieg (n_H) von 0,7 und somit unter eins. Dies lässt auf eine mögliche negative Kooperativität der Bindungsereignisse der CF I-Dimere 25k/68k an die RNA schließen.

Auch der *Cleavage Factor I* (CF I) wurde ähnlich wie der CstF-Komplex gereinigt. Das Heterotetramer, bestehend aus je zwei Kopien der 25k und der 68k Untereinheit, wurde durch einen gemeinsamen Virus in Sf21-Zellen exprimiert und in einem ersten Schritt in einer IMAC gereinigt. Über ein N-terminales 8x His-Fusionspeptid an der CF I 25k-Untereinheit wurde dabei das Tetramer angereichert. In einer anschließenden Reinigung wurden Teile der IMAC-Elutionsfraktionen E2 und E3 über das N-terminale StrepII-Fusionspeptid an CF I 68k (Abbildung 2.19 C) gereinigt. In den Elutionsfraktionen E3 bis E4 der Strep-Tactin-Säule fanden sich dabei keine in der Coomassie-Färbung erkennbaren Kontaminationen (Abbildung 2.19 C). CF I war in diesen Fraktionen jedoch nur in geringen Konzentrationen vorhanden. Um die Funktionalität des Komplexes zu testen, wurde das Strep-Tactin-gereinigte CF I für Filterbindung in Kombination mit einem 19-mer (CF I 2xSites) genutzt. Dieses RNA-Oligonukleotid war, wie das CstF-Oligonukleotid, an eine SELEX-Konsensus-Sequenz angelehnt (Brown und Gilmartin, 2003) und bot innerhalb der RNA-Sequenz zwei „UGUA“-Bindungsstellen für die beiden in CF I enthaltenen 25k/68k Paare an. Dies sollte in den Filterbindungsversuchen die Assoziation eines RNA-Moleküls pro CF I-Komplex ermöglichen. In der Abbildung 2.19 H sieht man die Auftragung der gebundenen RNA gegen die Komplex-Konzentration. Die Daten wurden mit einer pseudo-hyperbolen Anpassung nach Hill bearbeitet. Dies erlaubte die Abschätzung der apparenten Dissoziationskonstanten mit rund 30 nM. Aus dem gewonnenen Parameter $n_H=0,7$ bis 0,8 ließ sich zudem eine negative Kooperativität in der RNA-Bindung der Protein-Untereinheiten ableiten.

Somit schien die Bindung eines 25k/68k-Dimers die Bindung des zweiten zu behindern. Die linearisierte Hill-Darstellung der Daten in Abbildung 2.19 I zeigte durch den linearen Verlauf der Datenpunkte mit einem Anstieg $n_H=0,7$, welcher von dem Wert eins abwich, dass wahrscheinlich eine negative Kooperativität der Bindungsereignisse der CF I-Dimere 25k/68k an die beiden „UGUA“-Motive der RNA vorlag. Möglicherweise wäre ein anderes Substrat hier geeigneter die RNA-bindung des CF I-Komplexes nachzuweisen.

In einer zweiten Präparation des CF I-Komplexes wurde das Zelllysate wieder in einer IMAC gereinigt. In der Abbildung 2.19 D sieht man die Anreicherung des CF I-Komplexes während der IMAC. Die Ni-NTA-Elution E2 zeigte in einer Gelfiltration drei homogene Peaks und keine Aggregation (s. Anhang Abbildung 9.16 oben). Allerdings lag auch bei diesem Komplex das scheinbare Molekulargewicht des Hauptpeaks mit etwa 0,8 MDa über dem erwarteten Wert von 186 kDa für ein Heterotetramer. Wie im Falle des CstF wurde auch für CF I eine Analyse mittels Rechtwinkel-Lichtstreuung im Anschluss an eine Gelfiltration durchgeführt. Diese zeigte ein natives Molekulargewicht um 227 kDa für den Hauptpeak und korrespondierte mit leichter Abweichung mit dem erwarteten Bereich des Hexamers um 184 kDa (s. Anhang Abbildung 9.16 unten). Die anderen beiden detektierten Peaks der Gelfiltration mit den ermittelten Molekulargewichten um 140 kDa und 50 kDa korrespondieren mit den theoretischen Molekulargewichten der Homodimere aus CF I-68k und CF I-25k. Die Ni-NTA Elutionsfraktion E2 wurde im Anschluss an die IMAC mit salzlosem Puffer auf eine Salzkonzentration von 160 mM KCl verdünnt und auf eine ResourceQ-Säule geladen. Die Elution des CF I-Komplexes durch den verwendeten KCl-Gradienten erfolgte in den Fraktionen 11–17 (Abbildung 2.19 E). Überschüssiges CF I-25k konnte dabei in den Durchflussfraktionen 1–5 abgetrennt werden. Diese Präparation wurde im Weiteren für die Komplementation des fraktionierten Extraktes verwendet.

3 Diskussion

Ziel dieser Arbeit war, die 3'-Prozessierung von Prä-mRNAs durch Betrachtung und Charakterisierung der beteiligten Komponenten mPSF und CF II weiter zu definieren und die zugrundeliegenden Abläufe besser zu beschreiben, als es bislang möglich war.

Im ersten Teil wurde die RNA-Bindung und spezifische Interaktion des mPSF-Komplexes mit dem Polyadenylierungssignal betrachtet. Im zweiten Teil der Arbeit ist es gelungen den Spaltfaktor CF II aus den beiden Komponenten hPcf11 und hClp1 zu rekonstituieren und dessen Aktivität in der 3'-Prozessierung nachzuweisen. Zudem wurde die RNA-bindende Eigenschaft des hPcf11 charakterisiert und die kinetischen Parameter der 5'-Polynukleotidkinase-Aktivität des hClp1 beschrieben.

3.1 Der CPSF Subkomplex mPSF besteht aus vier Untereinheiten und bindet das Polyadenylierungssignal spezifisch und mit hoher Affinität

Die Rekonstitution der Polyadenylierung durch L. Schönemann (Schönemann u. a., 2014, Dissertaton 2017) zeigte, dass die vier kanonischen Untereinheiten des CPSF (CPSF1–4) für die Polyadenylierungsreaktion nicht ausreichend sind. Die Untereinheiten CPSF2 und 3 sind für die Polyadenylierung zudem nicht notwendig. Gleichzeitig sind hFip1 und WDR33 essenzielle Komponenten der Polyadenylierung. Zusammen mit den Untereinheiten CPSF1 und CPSF4 bilden WDR33 und hFip1 den Polyadenylierungs-Minimalkomplex des *mammalian Polyadenylation Specificity Factors*.

Verschiedene Kombinationen aus möglichen CPSF-Untereinheiten wurden für die Bestimmung des mPSF-Komplexes zunächst in Filterbindungsversuchen auf Bindung an das synthetische RNA-10-mer W10 („AAUAAACCCA“) getestet. Dieses W10 RNA-Dekamer besteht im Wesentlichen aus dem Polyadenylierungssignal „AAUAAA“ und einer kleinen Verlängerung am 3'-Ende („CCCA“). Kombinationen aus vier Untereinheiten (4k: CPSF1–4) und fünf Untereinheiten (5k: CPSF1–4, hFip1) zeigten in den RNA-Bindungsversuchen kaum RNA-Assoziation. Ein Komplex aus sechs Untereinheiten (6k: CPSF1–4, hFip1, WDR33) zeigte affine Bindung an die W10-RNA mit einer Affinität von etwa 20 nM. Eine entsprechende RNA-Spezifitätskontrolle mit U zu G Substitution im PAS (W10Δ) wurde hingegen kaum gebunden. Die Kombination 4k und 5k (ohne WDR33) zeigten im Vergleich eine mindestens um einen Faktor 20–50 schlechtere Affinität für das W10 RNA-Dekamer. Aus diesen Versuchen konnte also geschlossen werden, dass WDR33 notwendig war, um eine RNA mit PAS affin und spezifisch zu binden. L. Schönemann konnte in weitergehenden Polyadenylierungsversuchen bestimmen, dass in Verbindung mit PAP und PABPN1 der

minimalkomplex mPSF, bestehend aus CPSF1, WDR33, hFip1 und CPSF4, ausreichend für eine PAS-abhängige Adenylierung der Substrat-RNA (L3pre) war.

Auch der mPSF-Komplex wurde in dieser Arbeit mittels Filterbindungsversuchen auf RNA-Bindung an die W10-RNA hin getestet. Betrachtet man die Affinität des Minimalkomplexes mPSF für das W10 RNA-Dekamer, so ist diese mit einem apparenten K_D von etwa 1,5 nM noch höher als in der 6k-Konstellation (+CPSF2 und +CPSF3) mit 20 nM. Dies ist in guter Übereinstimmung mit später publizierten Bindungsdaten für mPSF durch Clerici und Kollegen (Clerici u. a., 2017), welche eine Affinität von $0,65 \pm 0,90$ nM ermittelten. Die Affinität für die W10 Δ -Kontrolle war hier wieder mindestens um den Faktor 20 geringer. Die Unterschiede in der Affinität des mPSF- und des 6k-Komplexes (mPSF+CPSF2 und 3) könnten den jeweiligen Anteil des aktiven Proteins in der Reinigung widerspiegeln. RNA-Titrationsen, um den Anteil des aktiven Proteins zu bestimmen, wurden nicht durchgeführt. In der Gelfiltration beider Präparationen über eine Superose6 Säule zeigte zumindest die mPSF-Präparation ein gutes Verhalten mit definiertem Maximum und homogenem Peak (Schönemann Dissertation 2017, s. Anhang Abbildung 9.11). Im EMSA (Abschnitt 2.1, Abbildung 2.2) zeigte mPSF ebenfalls einen definierten Komplex. Eine Titration des Komplexes ergab eine ähnliche Affinität wie in der Filterbindung.

3.2 WDR33 und CPSF4 binden gemeinsam das Polyadenylierungssignal (PAS)

Die UV-Vernetzung des mPSF an das kurze, mit 5-Jodo-Uracil markierte W10-RNA-Dekamer zeigte zwei Signale, zum einen ein starkes Signal bei etwa 30 kDa und zum anderen ein schwächeres Signal bei etwa 170–180 kDa (Kapitel 2.2, Abbildung 2.3 A und B). Das 30 kDa große Signal konnte mit einem Antikörper gegen CPSF4 unter leicht denaturierenden Bedingungen spezifisch angereichert werden. Somit konnte die bekannte Interaktion von CPSF4 mit dem Polyadenylierungssignal (Barabino u. a., 1997, 2000) bestätigt werden. Die erste Annahme war natürlich, dass CPSF1, welches als zweiter Bindungspartner an das PAS vermutet wurde, durch das Signal bei 170 kDa repräsentiert wurde (Bienroth u. a., 1991; Jenny u. a., 1994; Murthy und Manley, 1992, 1995). Dies konnte allerdings mit einer Affinitätsreinigung mit Hilfe von anti-CPSF1 Antikörper nicht bestätigt werden (Kapitel 2.2, Abbildung 2.3 C). Stattdessen zeigte eine Affinitätsreinigung über Ni-NTA Agarose unter denaturierenden Bedingungen, dass das radioaktive Signal zusammen mit His-WDR33 gereinigt werden konnte (Kapitel 2.2, Abbildung 2.3 C). Diese Beobachtung stimmte mit der gezeigten Notwendigkeit von WDR33 für die PAS-spezifische Stimulation der Polyadenylierung überein. Bis dahin war allerdings nicht postuliert, dass WDR33 die RNA direkt kontaktieren könnte. Diese Interaktion wurde immer dem ebenfalls ca. 170 kDa großen Protein CPSF1 zugeschrieben.

Zur genaueren Kartierung der RNA-Bindung von WDR33 wurden verschiedene limitierte Proteolysen mit Lys-C und Trypsin für eine Eingrenzung der interagierenden Aminosäuren durchgeführt. Dabei waren alle gereinigten Produkte C-terminale Verkürzungen des Proteins, da die Affinitätsreinigung über den jeweiligen N-terminalen His-Tag erfolgte. Nach sechsständiger Inkubation akkumulierte vor allem beim Trypsin-Verdau ein Fragment von etwa 15–20 kDa Größe. Dies deutet auf eine Interaktion der RNA mit einem Proteolyse-stabilen Fragment der ersten 180 Aminosäuren des WDR33 hin. Betrachtet man die Struktur des WDR33-Proteins, so entspräche dies einem Abschnitt innerhalb der ersten beiden WD40-Wiederholungen des WDR33 hin. Tatsächlich zeigte eine später veröffentlichte Kristallstruktur der WD40-Wiederholungen in Verbindung mit CPSF4 und einem kurzen PAS-RNA-Oligonukleotid, dass die Aminosäuren Phe153, Lys117 und Ile156 der ersten WD40-Wiederholung die Nukleotide U3 und A6 des PAS kontaktieren (Clerici u. a., 2017, 2018; Sun u. a., 2018).

Proteine mit WD40-Wiederholungen sind im Allgemeinen eher Vermittler von Protein-Protein-Interaktionen, indem sie einen stabilen Rahmen für die Assoziation weiterer Proteine bilden (Zou u. a., 2016). Die Spezifität der Interaktion wird dabei meist durch Bereiche außerhalb der eigentlichen Wiederholungen definiert. Somit war die direkte Beteiligung der WD40-Wiederholungen an der RNA Bindung eher ungewöhnlich. Allerdings gibt es einige weitere RNA-bindende Vertreter der WD40-Familie. So bindet WDR79 das CAB-Motiv in *small Cajal body specific* RNAs (scaRNAs), welche an der *small nuclear* RNA (snRNA)-Modifikation beteiligt sind (Tycowski u. a., 2009). Als Komponente des SMN-Komplexes bindet zum einen Germin5 das Sm-Motiv in der U4-snRNA an der Kontaktfläche seiner beiden N-terminalen WD40-Domänen (Jin u. a., 2016). Zum anderen wurde auch das SMN-Protein STRAP (*Serine-threonine kinase receptor-associated protein*) in seinen WDR40-Domänen an RNA UV-vernetzt (Castello u. a., 2016). In der gleichen Studie wurden auch Rack1 (*Receptor of activated protein C kinase 1*) und eIF-3I (*Eukaryotic translation initiation factor 3 subunit I*) in ihren WD40-Wiederholungen an RNA UV-vernetzt (Castello u. a., 2016).

3.3 hCpl1 und hPcf11 rekonstituieren CF II-Aktivität *in-vitro* und interagieren über konservierte Bereiche miteinander

Unter den zentralen Faktoren der 3'-Prozessierung von mRNAs waren sowohl CstF als auch CF I bereits aus rekombinanten Proteinen rekonstituiert worden (Rüeggsegger u. a., 1998; Takagaki und Manley, 1994; Yang u. a., 2011, 2018). Selbst der zentrale Faktor CPSF konnte zumindest als Subkomplex mPSF für die Polyadenylierungsreaktion rekombinant dargestellt werden (Clerici u. a., 2017, 2018; Schönemann u. a., 2014; Sun u. a., 2018). Nur der CF II-Komplex blieb

dahingehend nicht charakterisiert. Als zentrale Komponenten des CF II-Komplexes waren hPcf11 und hClp1 wahrscheinlich. Diese zwei Proteine waren sowohl in CF II-Fraktionen aus HeLa Kernextrakten (de Vries, 2000), als auch aus hClp1 Affinitätsreinigungen aus HEK 293 Zellen (Dissertation A. Knoth, 2009) angereichert. Die genannten Präparationen enthielten aber zusätzlich noch weitere potenzielle Kandidaten. In dieser Arbeit ist es gelungen ein stabiles, rekombinantes Dimer der beiden Proteine hPcf11 und hClp1 zu reinigen. Die Proteine des CF II-Dimers waren, wie auch yPcf11 und yClp1 im CF IA-Komplex der Hefe, fest miteinander assoziiert und ließen sich über mehrere Säulen und verschiedene Trennverfahren gemeinsam reinigen. In der Größenausschlusschromatographie zeigte der CF II-Komplex stets ein viel größeres apparentes Molekulargewicht als dies theoretisch für ein Dimer zu erwarten gewesen wäre. Somit hätte der gereinigte CF II-Komplex nicht nur ein Dimer der beiden Proteine sein können. Auch multiple Assoziationen, ein nicht stöchiometrisches Verhältnis oder eine teilweise Aggregation wären denkbar gewesen. Die Untersuchung des Sedimentationskoeffizienten in der Dichtegradientenzentrifugation zeigte einen S-Wert von 6,9. Zusammen mit dem aus der Gelfiltration ermittelten Stoke's-Radius von 7,1 nM konnte daraus abgeleitet werden, dass das tatsächliche Molekulargewicht bei etwa 210 kDa lag. Damit war ein Hetero-Dimer der beiden Proteine hClp1 und hPcf11 (48 kDa und 172 kDa) die am wahrscheinlichsten vorliegende Variante. Das hPcf11 Protein besitzt vermutlich eine sehr unstrukturierte und nicht-globuläre Form und sorgte so für das hohe apparente Molekulargewicht in der Größenausschlusschromatographie. In der Arbeit konnte durch Komplementation von teilfraktioniertem nukleärem Extrakt gezeigt werden, dass ein CF II-Komplex aus hClp1 und hPcf11 ausreicht, um die 3'-Spaltaktivität des CF II zu vermitteln. Zudem konnte auch eine Dosisabhängigkeit der Spaltaktivität des eingesetzten rekombinanten CF II-Komplexes in der RQ-Fraktion 9 des fraktionierten NXT gezeigt werden. Die Spaltaktivität migrierte in den Komplementationen mit dem CF II-Komplex. Zudem konnte die 5'-PNK Aktivität des hClp1 übereinstimmend verfolgt werden und die Proteine hPcf11 und hClp1 konnten durch Western-Blot Analysen spezifisch nachgewiesen werden.

In der Dissertation von C. Tüting (2020) wurde durch Quervernetzung und Massenspektrometrie nachgewiesen, dass im rekombinanten CF II das hClp1 über Oberflächenkontakte der zentralen PNK- und der C-terminalen Domäne mit einem konservierten Bereich zwischen den Aminosäuren 1371–1439 des hPcf11 interagiert (Anhang Abbildung 9.8 und Abbildung 9.6). Dies wurde schon durch einen Sequenzvergleich der interagierenden Aminosäuren aus Kristallstrukturen des yClp1 mit einem yPcf11-Fragment zwischen den AS 454–563 nahegelegt (Noble u. a., 2007). Eine Mutation der Interaktionsregion in hPcf11 wurde im Rahmen der Arbeit nicht generiert, da hPcf11 als Einzelprotein unlöslich in den SF21-Zellen vorlag. Die Information, inwiefern eine Assoziation von hPcf11 mit hClp1 die Spaltaktivität beeinflusst, wäre zusätzlich von Interesse.

Im Weiteren wurden durch C. Tüting verschiedene verkürzte Varianten von hPcf11 im CF II-Komplex generiert. Diese wurden in der Arbeit in 3'-Spaltassays genutzt, um fraktionierten NXT zu komplementieren. Hier zeigten nur einige dieser Varianten Aktivität in Spaltassays. So musste die hPcf11-Variante im CF II nicht nur die beiden Zinkfinger und die hClp1-Interaktionsregion beinhalten, sondern auch die als FEGP-Wiederholung bezeichnete Region (hPcf11 Δ N 769). Wie diese FEGP-Wiederholungen an der Aktivität des CF II beteiligt sind, ist nicht klar. Ähnliche Strukturen konnten in einem Sequenzvergleich nur in hPcf11-Orthologen anderer Wirbeltiere identifiziert werden (s. Anhang Alignment). Relativ sicher ist, dass dieser Bereich als unstrukturierte Region vorliegt (*in silico* Vorhersage) und methyliert wird (The UniProt Consortium, 2019). Eine Protein-Protein-Interaktion der FEGP-Region des hPcf11 mit weiteren Spaltfaktoren ist wohl am wahrscheinlichsten. Die Aktivität der N-terminal verkürzten Variante hPcf11 Δ N 769 in CF II zeigte zudem, dass die gut konservierte N-terminale CID für die Aktivität in der 3'-Spaltung, in Übereinstimmung mit Daten aus Hefe (Sadowski u. a., 2003), nicht notwendig war.

3.4 Die 5'-PNK-Aktivität des hClp1 ist für die CF II-Aktivität in der 3'-Spaltreaktion nicht notwendig

Das humane Clp1 ist eine aktive 5'-Polynukleotidkinase (PNK) (de Vries, 2000), wie auch viele homologe Varianten aus anderen Organismen. In Hefe allerdings ist das Clp1-Protein inaktiv (Ramirez u. a., 2008). Dafür sorgen einige Abweichungen in der Aminosäurekette des γ Clp1. Hauptsächlich sorgen die Abweichungen von Arg316 und Arg321 zu Val316 und Lys321 in dem *Lid-Module* der PNK-Domäne dafür, dass der Phosphoryl-Transfer von ATP auf das Substrat erschwert wird (Dikfidan u. a., 2014). Daraus ergab sich die Annahme, dass die PNK-Aktivität eventuell für die Funktion von hClp1 in der 3'-Prozessierung von mRNAs nicht notwendig ist. Zur Überprüfung der Annahme wurden zwei verschiedenen Punktmutanten generiert, welche entweder den Phosphoryl-Transfer-Übergangszustand nicht mehr stabilisieren sollten (KR: K127A, R288A, R293L) oder im Phosphattransfer (DR: D151A, R288A, R293L) beeinträchtigt sein sollten. Das Lysin 127 interagiert während des Phosphoryl-Transfer-Übergangszustandes mit den β - und γ -Phosphatgruppen des ATPs. Arginin 288 und 293 sind Teil des *Lid-Modules* und kompensieren die vorübergehenden, negativen Ladungen während des Phosphoryl-Transfers. Das Aspartat 151 ist die katalytische Aminosäure, welche die 5'-Hydroxyl-Gruppe der RNA deprotoniert (Dikfidan u. a., 2014). Als Folge der eingeführten Mutationen „KR“ und „DR“ in hClp1 zeigten beide CF II-Varianten keine 5'-PNK-Aktivität mehr. Eine Mutation der ATP-Bindung sollte dabei weitgehend vermieden werden. In Hefe stört eine direkte Mutation der ATP-Bindungsstelle die Assoziation von γ Cpl1 und γ Pcf11, als Folge sogar die Assoziation von γ Pcf11 mit Rna15 und -14 (Ghazy u. a.,

2012; Holbein u. a., 2011; Vilain u. a., 2012). Während der Reinigung der beiden hClp1-Mutanten „KR“ und „DR“ im CF II-Komplex zeigten sich keine Unterschiede in Expression oder Reinigungen. Die Komplexe lagen stabil assoziiert miteinander vor, zeigten keinen verstärkten Abbau in den Zellextrakten und konnten wie das Wildtyp-CF II über zwei Säulen als Komplex gereinigt werden. Somit konnte angenommen werden, dass die Faltung und Stabilität der Komplexe nicht beeinträchtigt waren. In der Komplementation der Spaltung zeigte sich, dass beide hClp1-Punktmutanten in CF II in gleicher Weise wie das Wildtyp-CF II die Spaltaktivität vermittelten. Damit war auch in der untersuchten *in vitro* Situation die 5'-PNK-Aktivität des hClp1 in der prä-mRNA 3'-Prozessierung des Menschen nicht notwendig. Aufgrund der Limitation, auf fraktionierten nukleären Extrakt zurückgreifen zu müssen, kann hierbei natürlich nicht ausgeschlossen werden, dass kontaminierende Proteine unterhalb der Detektionsgrenze des Western-Blots vorhanden waren. Allerdings zeigten entsprechende Kontrollen jeweils die Notwendigkeit des zugesetzten CF II-Komplexes. Endogene Proteine konnten also nicht in ausreichenden Mengen vorhanden sein, um eine messbare 3'-Spaltaktivität zu vermitteln.

3.5 hClp1 ist eine aktive PNK und unterscheidet sich in der ATP-Affinität von ceClp1 aus *C. elegans*

Im Verlauf der Arbeit wurden auch die kinetischen Parameter des humanen Clp1 erfasst und die Substratspezifität für RNA und Affinitäten für RNA und ATP ermittelt. In der Literatur waren diese für das humane Clp1 noch nicht beschrieben. Als Vergleich kann hierbei das in Sequenz und Struktur sehr ähnliche Homologe aus *C. elegans* herangezogen werden. Für das ceClp1 wurden durch Dikfidan und Kollegen ausführliche Studien hinsichtlich der kinetischen Parameter des ceClp1 durchgeführt (Dikfidan u. a., 2014). Der ermittelte K_M des ceClp1 für ein GC-Dinukleotid lag bei etwa 109 μM mit einem k_{cat} von 2,9 s^{-1} . Für das zweite Substrat (ATP) wurde ein K_M von etwa 260 μM ermittelt mit einem k_{cat} von 2,9 s^{-1} . Für das humane hClp1 wurden dem CF II-Komplex zunächst verschiedene RNA-Substrate angeboten. Dabei zeigte sich, dass bei einigen Substraten eine Art Substratüberschusshemmung einsetzte. Dies galt für die Substrate Wigley -23/-10 (14-mer mit PAS), das SV40 *late* Fragment 190–225 (36-mer) und rA₁₄. Möglich wäre für diese Substrate, dass bei höheren Konzentrationen die Ausbildung von intermolekularen Strukturen begünstigt wurde, welche die RNA-Substratbindung des hClp1 beeinträchtigten. Zumindest für die Guanin-reiche SV40 *late* 190–225-RNA wurde in CD-Spektren in dieser Arbeit die Ausbildung von stabilen RNA-Strukturen nachgewiesen. Die Substrate rU₁₄ und rC₁₄ zeigten im untersuchten Bereich bis 20 μM keine Anzeichen für Überschusshemmung. Das für die weiteren Versuche genutzte rC₁₄-Substrat hatte einen K_M von 2,5±0,4 μM im Michaelis-Menten-Modell. Die ermittelte Wechselzahl k_{cat} liegt dann bei 4,5±1,4 s^{-1} . Der Vergleich zum isolierten

hClp1 war für die Bestimmung der RNA-Affinität schwierig. Das Protein verlor im Laufe der Lagerung an Aktivität. Die K_M für das rC₁₄-Substrat lagen in zwei Einzelbestimmungen bei 2,4 μM und 1,6 μM . Ein größerer Anteil an inaktivem Protein sollte hier zwar nicht zu einer Veränderung des K_M führen, die Werte könnten aber die experimentelle Abweichung widerspiegeln. Betrachtet man die linearisierte Darstellung der Daten für die RNA-Titrationsen, so erkennt man in der Tendenz einen ähnlichen x-Achsenabschnitt. Dies würde für einen ähnlichen RNA- K_M von CF II und hClp1 sprechen, welcher aus dem x-Achsenabschnitt abgeleitet werden kann. Durch den Aktivitätsverlust des hClp1 bedingt, konnte die Wechselzahl für das isolierte hClp1 nicht eindeutig bestimmt werden. Betrachtet man allerdings den chronologisch früheren Versuch, so war der k_{cat} mit 1,6 s^{-1} im Vergleich zu $4,5 \pm 1,4 \text{ s}^{-1}$ in den CF II-Versuchen schon um den Faktor zwei bis drei verringert. Der zweite Wert zeigte schon einen deutlichen Abfall der maximalen Reaktionsrate zu einem k_{cat} von 0,7 s^{-1} . Auch die unterschiedlichen y-Achsenabschnitte der Linearisierung der beiden RNA-Titrationsen des hClp1 sprechen für einen Abfall der Reaktionsgeschwindigkeit. Eine Wiederholung der Versuche zur RNA-Titration mit hClp1 war aufgrund des Aktivitätsverlustes nicht mehr möglich. So kann nicht unterschieden werden, ob die höhere Wechselzahl des hClp1 in CF II durch den stabilisierenden Effekt des hPcf11 bedingt war oder durch eine Beeinflussung der Reaktionsrate über die hPcf11-Interaktion. Zusammenfassend ist die Affinität für das RNA-Substrat mit 2,5 μM (CF II) und 2,4/1,6 μM (hClp1) um etwa den Faktor 40–50 höher als die Affinität des ceClp1 für das angebotene GC-Dinukleotid ($109 \pm 11 \mu\text{M}$) oder für das GAAA-Oligonukleotid ($99 \pm 25 \mu\text{M}$). Die Wechselzahl des CF II war mit $4,5 \pm 1,4 \text{ s}^{-1}$ mit dem homologen Protein ceClp1 mit $2,9 \pm 0,1 \text{ s}^{-1}$ leicht erhöht, aber durchaus vergleichbar. Der k_{cat} des hClp1 war unter den gegebenen Einschränkungen mit 1,6 s^{-1} und 0,7 s^{-1} leicht niedriger als der des ceClp1.

Betrachtet man nun die ATP-spezifischen Parameter des CF II, so erkennt man deutliche Unterschiede zwischen hClp1 im CF II-Komplex und dem isolierten hClp1. Der ATP- K_M lag für CF II bei $0,7 \pm 0,2 \text{ mM}$ und somit deutlich unter dem ATP- K_M des isolierten hClp1-Proteins mit $2,5 \pm 1,1 \text{ mM}$. Die ATP- K_M war für das isolierte hClp1 somit etwa dreimal so hoch wie im CF II-Komplex. Auch die Wechselzahl der ATP-Kinetiken des CF II mit einem k_{cat} von $2,8 \pm 0,8 \text{ s}^{-1}$ unterschied sich etwa um den Faktor zwei zum k_{cat} des isolierten hClp1 mit $1,3 \pm 0,5 \text{ s}^{-1}$. Auch die linearisierte Darstellung der ATP-Kinetiken zeigte sowohl abweichende x-Achsenabschnitte als auch y-Achsenabschnitte. Somit erschien das hClp1 im Komplex mit hPcf11 nicht nur stabilisiert, sondern auch die ATP-Bindung schien begünstigt gewesen zu sein. Denkbar wäre eine Stabilisierung der hClp1-Struktur in der ATP-gebundenen Form. Ein Einfluss der ATP-bindenden Aminosäuren des yClp1 auf die Assoziation an yPcf11 ist zumindest beschrieben (Ghazy u. a., 2012).

Die Daten der RNA-Kinetiken wurden erfasst, indem nach der 5'-PNK-Reaktion das nicht umgesetzte ATP mittels denaturierender PAGE abgetrennt und das radioaktive Signal der

phosphorylierten RNA über Phosphorimaging quantifiziert wurde. Die Daten der ATP-Kinetiken wurden erfasst, indem nach der 5'-PNK-Reaktion das nicht umgesetzte ATP mittels DEAE-Methode ausgewaschen wurde. Die Quantifizierung der phosphorylierten RNA erfolgte danach mit einer Szintillationsmessung. Auch wenn die 5'-PNK-Reaktion immer nach dem gleichen Muster erfolgte, so unterschieden sich DEAE-Methode und Harnstoff-PAGE in der Konfidenz der erfassten Daten. Daraus ergibt sich eine eingeschränkte Vergleichbarkeit der beiden Quantifizierungsmethoden hinsichtlich der ATP-Affinität und der ATP-Reaktionsrate gegenüber der Affinität RNA-Substrate und deren Reaktionsraten. Zudem konnten die rC₁₄-Kinetiken nicht unter ATP-gesättigten Bedingungen durchgeführt werden, da der Anteil und die Detektion des beigefügten radioaktiv markierten [γ -³²P]-ATP limitierend wirkte. So musste für die Reaktionen mit hohen ATP-Konzentrationen unter sättigender RNA-Konzentration eine große Menge [γ -³²P]-ATP beigefügt werden, um den Anteil des radioaktiven Phosphats über einen detektierbarem Schwellenwert zu halten. Dies führte jedoch in zunehmendem Maße auch zu Hintergrundsignalen, welche die Detektion der umgesetzten RNA erschwerte. Ein gekoppelter kalorimetrischer Assay, welcher optisch die ADP-Zunahme misst, hätte hierfür möglicherweise eine zusätzliche Herangehensweise dargestellt. Allerdings wären die Substratumsätze mit den submikromolaren Konzentrationen der RNA hierbei wohl schlecht messbar gewesen.

3.6 Der CF II-Komplex bindet hauptsächlich über hPcf11 an G-reiche RNA-Sequenzen

Während der Arbeit mit dem rekombinanten CF II konnte eine hochaffine Bindung an die L3- und SV40 *late*-RNA festgestellt werden. Das isolierte hClp1 zeigte eine solche Bindung hingegen nicht. Über mehrere Reinigungsschritte konnte die RNA-Bindung der Proteinfractionen in Übereinstimmung mit der Menge des enthaltenen CF II-Komplexes beobachtet werden. Als mögliche RNA-bindende Komponente ließ sich hPcf11 im CF II-Komplex vermuten. Eine UV-Vernetzung mit einem radioaktiv markierten rU₁₄ RNA-Oligomer und anschließende Antikörper-Affinitätsreinigung des hPcf11 zeigte in der Tat eine direkte Interaktion des hPcf11 mit der angebotenen RNA. Eine schwache RNA-Bindung war für die konservierte CID des yPcf11 bereits beschrieben (Hollingworth u. a., 2006; Zhang, 2005). Für die beiden C-terminalen Zinkfinger-Domänen war bis dahin keine RNA-Interaktion beschrieben. N-terminal verkürzte hPcf11-Varianten in CF II, ohne die CID, zeigten auch eine Stimulation der Spaltungsreaktion. So wurden diese Varianten auch auf RNA-Bindung hin getestet.

Dabei konnte gezeigt werden, dass die RNA-Bindung hauptsächlich durch die beiden C-terminalen Zinkfinger-Domänen vermittelt wird. Diese flankieren im C-terminalen Bereich des hPcf11 die

hClp1-Interaktionsregion und sind weitgehend zwischen verschiedenen Spezies konserviert. Mit zunehmender N-terminaler Verkürzung nahm auch die Affinität zur angebotenen RNA ab (Abbildung 3.1). Die Affinität zur angebotenen SV40 *late* 190–225-RNA der kurzen Zinkfinger-Fragmente (Fragment zcz: AS 1339–1506) war mit einem K_{50} von 70 nM (Sumo-Fusion) und von 230 nM (GST-Fusion) etwa um den Faktor 150–300 geringer, als die Affinität des volllängen-Pcf11 mit einem K_{50} von $0,2 \pm 0,1$ nM. Dieser Unterschied entspricht in der Bindungsenthalpie jedoch in etwa der Energie einer Wasserstoffbrücke. Die Spezifität zu Guanin-reichen RNA-Substraten blieb auch in den N-terminalen Verkürzungen des hPcf11 erhalten. Die CID scheint somit die RNA nur unspezifisch zu kontaktieren. Mutationen in den Zinkfinger-Domänen des yPcf11 legen nahe, dass die RNA-Bindung für die Funktion des yPcf11 im CF II-Komplex wichtig ist. Im Besonderen eine Mutation im zweiten Zinkfinger führte zu einer starken Beeinträchtigung im Zellwachstum und zu Defekten in der prä-mRNA-Prozessierung (Guéguéniat u. a., 2017).

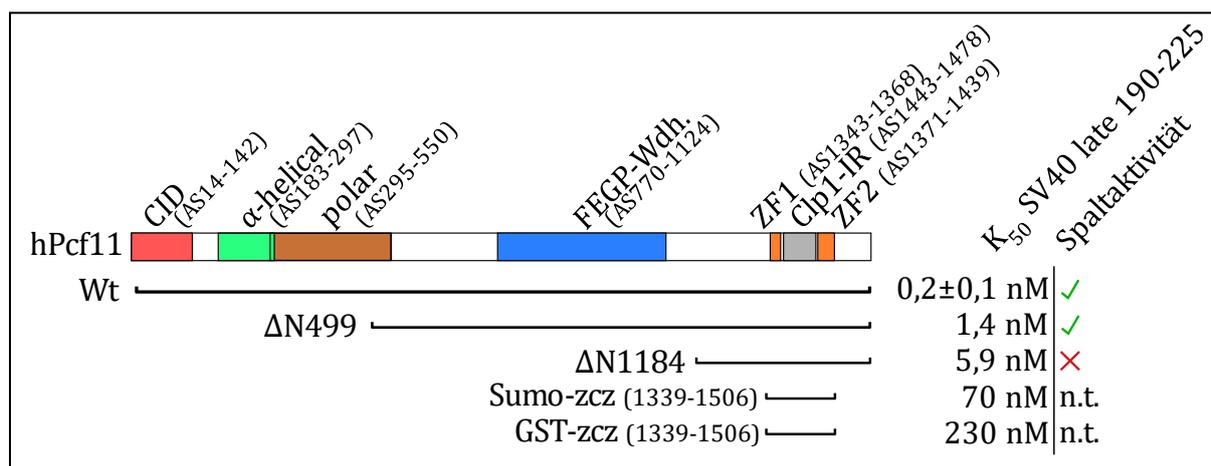


Abbildung 3.1 Schematische Darstellung der Domänen und Bereiche des hPcf11 und der Aktivität im CF II. Schematische Darstellung der Domänen und Bereiche in der humanen Pcf11-Primärstruktur mit detaillierter Angabe der Aminosäurepositionen. Die Darunter eingezeichnet sind die in der RNA-Bindung getesteten N-terminalen hPcf11-Verkürzungen im CF II-Komplex, sowie die isolierten ,zcz'-Fragmente mit zugehörigem K_{50} , für die verwendete SV40 *late* 190–225-RNA. Rechts sind zusätzliche Informationen angegeben, ob eine Spaltaktivität mit den CF II-Varianten nachgewiesen wurde.

Im Verlaufe der Arbeit konnte gezeigt werden, dass der CF II-Komplex über hPcf11 bevorzugt G-reiche Sequenzen bindet. Dabei konnte durch Bindungsstudien mit länger werdenden Oligo(G)-RNAs eine Bindungslänge von 14 Nukleotiden ermittelt werden. Dies erlaubte eine hoch affine Bindung mit einem K_{50} um 2 nM, wobei jede Verkürzung der RNA zu einem deutlichen Abfall der Affinität führte. Die duale Zinkfingerstruktur des hPcf11 mit der hClp1-Interaktionsregion in der Mitte, könnte die ungewöhnlich lange RNA-Bindungslänge von 14 Nukleotiden erklären.

Normalerweise würde man pro Zinkfinger eine Interaktion mit maximal 4–5 RNA-Nukleotiden erwarten (Hudson u. a., 2004).

Mit unterschiedlichen Methoden wurde zudem versucht, eine spezifische Bindungssequenz zu ermitteln. Zum einen wurde ein gerichteter Ansatz verfolgt, in dem die Sequenz der SV40 *late*-RNA als Ausgangspunkt diente. Bindungsversuche mit zunehmend verkürzten Fragmenten der RNA zeigten hierbei den G-reichen, 14 Nukleotide langen Sequenzabschnitt 208–221 als wahrscheinliche Bindungsstelle mit einem K_{50} von 3,4 nM für das CF II. Im Vergleich zur SV40 *late* 190–225-RNA (36 nt) mit einem K_{50} von $0,2 \pm 0,1$ nM stellte dies nur eine Verringerung der Affinität um etwa den Faktor 10–20 dar und bewegt sich somit im Bereich der Affinität für rG₁₄. Ein ungerichteter Ansatz über das „Bind’n’Seq“ Verfahren mit randomisierten 14-meren (Treiber u. a., 2017) zeigte eine ähnliche Präferenz für G-reiche Sequenzen. Eine größere Übereinstimmung mit der SV40-Sequenz konnte aber nicht beobachtet werden. Somit konnte also, mit den verwendeten Methoden, keine eindeutige Bindungssequenz identifiziert werden.

In seiner biologischen Funktion interagiert der CF II-Komplex mit 3'-UTRs von Prä-mRNAs und vermittelt die 3'-Spaltung der RNA durch CPSF3. Nicht alle 3'-UTR enthalten G-reiche Sequenzen, doch sind diese statistisch angereichert in der Umgebung von etwa ± 100 Nukleotiden um die Spaltstelle (HU u. a., 2005). Zudem konnte in Hefe die yPcf11-Bindung stromabwärts der Spaltstellen kartiert werden (Baejen u. a., 2017). Somit könnte CF II durch die Interaktion mit G-reichen RNA-Abschnitten, als Spaltfaktor die Spaltstelle in der 3'-Prozessierung zusätzlich definieren und regulatorisch auf die Auswahl der genutzten Spaltstelle einwirken. Diese Annahme wird durch die Beobachtung untermauert, dass hPcf11 die genutzte Spaltstelle der eigenen RNA und als Folge die eigene Translation reguliert (Kamieniarz-Gdula u. a., 2019). Zudem zeigten *Knock-down* Versuche auch, dass vor allem proximale PAS durch Interaktion mit hPcf11 bevorzugt prozessiert wurden und somit die Länge der 3'-UTR in der Population der Transkripte, bei hohem hPcf11-Proteinlevel, abnimmt. Ogorodnikov und Kollegen zeigten im hPcf11-*Knock-down* auch eine Zunahme an Transkripten mit distal genutzten Polyadenylierungssignalen. Dies hatte einen negativen Einfluss auf die Proliferation und induzierte die Differenzierung von Neuroblastoma-Zelllinien, nach dem induzierten hPcf11-*Knock-down* (Ogorodnikov u. a., 2018). Im Allgemeinen gelten Proteine, welche 3'-UTRs von Transkripten regulieren, als Proto-Onkogene, da meist über regulatorische Interaktionen innerhalb der 3'-UTR eines Transkriptes die post-transkriptionelle Kontrolle vermittelt wird. So kann die 3'-UTR Zelllokalisierung, RNA-Stabilität und Translation der RNA beeinflussen. Meist erfolgt dies über die Interaktion mit miRNAs oder mit RNA-bindenden Proteinen (Mayr, 2017). Kurze 3'-UTRs bieten also weniger Möglichkeiten diese Kontrolle zu vermitteln. Wird über eine hohe hPcf11-Konzentration also die 3'-UTR-Länge vieler Transkripte verkürzt, so ist dies oft ein Indikator für proliferative und undifferenzierte Zellen.

3.7 hPcf11 bindet die RNA in einer linearen Konformation und nicht als G-Quadruplex-Struktur

Guanin-reiche Sequenzen sind in der Lage Basenpaarungen einzugehen, die nicht auf einer Watson-Crick-Basenpaarung, sondern auf einer Hoogsteen-Basenpaarung beruhen. Sequenzen des Formates (GGN)₄ sind in der Lage, sogenannte G-Quadruplex-Strukturen auszubilden. Diese Form der Basenpaarung ist sowohl für DNA als auch RNA beschrieben (Vorlíčková u. a., 2012; Zarudnaya u. a., 2003). Diese Strukturen können durchaus Ziel von Proteininteraktionen sein oder eben diese Interaktion durch die hergestellte Konformation verhindern (Wang u. a., 2017). Die ermittelte CF II-Zielsequenz aus der SV40 *late*-RNA (nt 208–221) sowie einige Sequenzen der „Bind’n’Seq“ Ergebnisse besaßen die Voraussetzung, G-Quadruplex-Strukturen auszubilden. Die Möglichkeit, dass hPcf11 eben diese Struktur und nicht eine bestimmte Sequenz kontaktierte, konnte durch Bindungsversuche mit drei Sequenzvariationen des SV40 *late*-RNA-Abschnittes 208–221 weitgehend ausgeschlossen werden. Mittels Zirkulardichroismus konnte bestätigt werden, dass Zwei der in der Bindung angebotenen RNA-Oligonukleotide (Wt2x, Δ2x) eine G-Quadruplex-Struktur ausbilden können. Für diese Oligonukleotide Wt2x und Δ2x konnten die typischen Maxima bei 262 nm beobachtet werden (Vorlíčková u. a., 2012). Ein drittes angebotenes Oligonukleotid gleicher Länge (ΔΔ2x) konnte aufgrund der Basenabfolge keine G-Quadruplex-Strukturen ausbilden und zeigte ein Maximum bei 272 nm, welches charakteristisch für unstrukturierte RNA ist. Die Sequenz des ΔΔ2x unterschied sich vom Δ2x nur durch die Reihenfolge der Basen, nicht jedoch im Guanin-Gehalt. Somit wurde das Muster (GGN)₄ im Oligonukleotid ΔΔ2x vermieden. Die nativen Gelshift-Experimente zeigten jedoch eine ähnliche Affinität von hPcf11 für alle drei angebotenen RNAs. Die RNAs zeigten auch ein ähnliches Migrationsverhalten. Somit liegt die Schlussfolgerung nahe, dass die RNAs bei den niedrigen Konzentrationen in den Bindungsversuchen unstrukturiert vorlagen, hPcf11 die RNA in der linearen Form kontaktierte und nicht eine spezifische räumliche Konformation der angebotenen RNA benötigte. Zudem zeigten auch Bindungsversuche an weitere RNAs, dass der Guanin-Gehalt der RNA weitgehend, aber nicht vollständig, verringert werden konnte, ohne die Affinität des hPcf11 zum Substrat stark zu verringern.

4 Zusammenfassung

In dieser Arbeit konnte gezeigt werden, dass der durch L. Schönemann (2014) beschriebene Subkomplex mPSF des humanen 3'-Prozessierungsfaktors CPSF spezifisch und mit hoher Affinität das kanonische Polyadenylierungssignal „AAUAAA“ bindet. Eine zusammenfassende Darstellung der humanen 3'-Prozessierungsfaktoren findet sich in Abbildung 4.1. Für eine Bindung des Polyadenylierungssignals mit einem K_{50} von etwa 2 nM waren alle vier beschriebenen mPSF-Untereinheiten (CPSF1, hFip1, WDR33 und CPSF4) notwendig. Ein W10 RNA-Dekamer mit einem U zu G Basenaustausch im PAS wurde mit mindestens zwanzigfach geringerer Affinität gebunden. Dabei zeigten sowohl Filterbindungs- als auch native *Shift*-Experimente übereinstimmende Ergebnisse. Diese Ergebnisse konnten durch unabhängige Arbeiten anderer bestätigt werden (Clerici u. a., 2017; Sun u. a., 2018).

In dieser Arbeit wurde weiterhin gezeigt, dass die mPSF-Untereinheit CPSF4 durch UV-Vernetzung an das radioaktiv markierte W10-RNA-Dekamer geknüpft werden konnte. Nach der UV-Vernetzung erfolgte eine Antikörper-Affinitätsreinigung von CPSF4 zusammen mit dem vernetzten RNA-Oligomer. Die Auswertung zeigte ein spezifisch angereichertes Signal für CPSF4 und bestätigte eine direkte RNA-Interaktion. Eine RNA-Interaktion von CPSF4 im Bereich des PAS war schon lange bekannt, besteht CPSF4 doch zum größten Teil aus fünf Zinkfinger-Domänen (Barabino u. a., 1997; Jenny u. a., 1994). Die affine Bindung an das sehr kurze W10-RNA-Dekamer zeigt, dass die RNA-Bindung im Wesentlichen nur in der PAS-Sequenz erfolgen konnte.

Auf ähnliche Weise konnte gezeigt werden, dass auch die mPSF-Untereinheit WDR33 an das gleiche radioaktiv markierte W10-RNA-Dekamer vernetzt werden konnte. Nach der UV-Vernetzung an die RNA erfolgte die Anreicherung eines spezifischen WDR33-Signals durch denaturierende Reinigung des N-terminalen His-Tags. Damit wurde zum ersten Mal die RNA-Bindung von WDR33 nachgewiesen. Darüber hinaus wurde durch UV-Vernetzung und Antikörper-Affinitätsreinigung gezeigt, dass die, lange als PAS-bindend postulierte, Untereinheit CPSF1 offenbar nicht mit der W10-RNA vernetzt wurde. In einer Reihe limitierter Proteolysen des UV-vernetzten mPSF und anschließender denaturierender Ni-NTA-Affinitätsreinigung des His-WDR33 konnte gezeigt werden, dass wahrscheinlich die RNA-Bindung durch WDR33 innerhalb der ersten beiden WD40-Wiederholungen im N-terminalen Bereich des Proteins erfolgt.

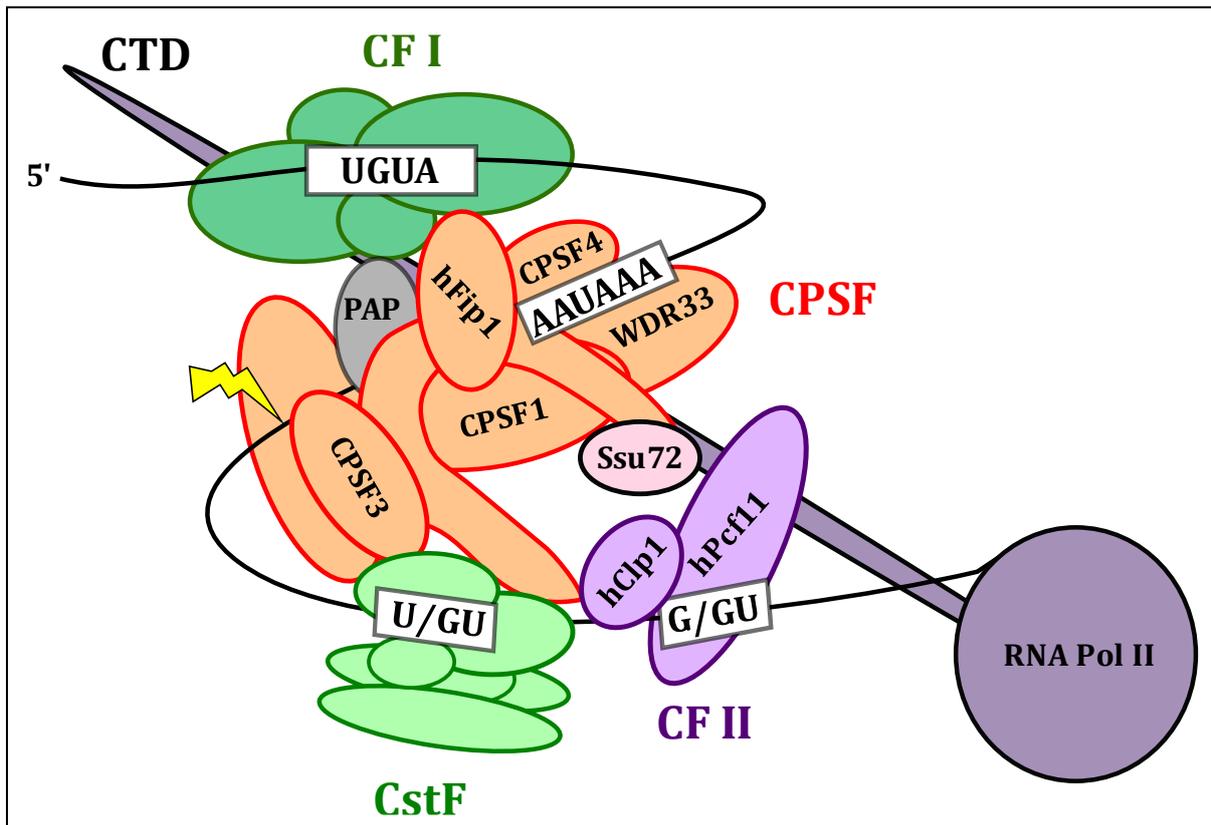


Abbildung 4.1 Aktualisierte schematische Darstellung des 3'-Prozessierungskomplexes auf der prä-mRNA. Die C-terminale Domäne (CTD) der RNA-Polymerase II (RNA Pol II) dient als Plattform für die 3'-Spaltfaktoren. Sie wird durch Phosphatasen wie Ssu72 im Phosphorylierungsstatus beeinflusst. Cleavage Factor I (CF I) bindet oberhalb der Spaltstelle an das „UGUA“-Motiv der RNA. Der Cleavage and Polyadenylation Specificity Factor (CPSF) kann in zwei katalytische Subkomplexe unterteilt werden, mPSF und Spaltmodul. mPSF – bestehend aus den vier Proteinen CPSF1, WDR33, hFip1 und CPSF4 – bindet das Polyadenylierungssignal (PAS) über die Untereinheiten WDR33 und CPSF4. Das Spaltmodul katalysiert mittels CPSF3 die Spaltung an der Spaltstelle (gelber Blitz). Der Cleavage stimulatory Factor (CstF) bindet unterhalb der Spaltstelle an ein GU/U-reiches Element und vermittelt über Interaktion mit CPSF zusätzlich die Spaltaktivität. Der Cleavage Factor II (CF II), bestehend aus hClp1 und hPcf11, interagiert über die Zinkfingerdomänen in hPcf11 mit G/GU-reichen Sequenzen und vermittelt durch Interaktion mit CPSF die 3'-Spaltaktivität und durch Interaktion mit der CTD die Transkriptions-Termination.

Im zweiten Teil der Arbeit wurde gezeigt, dass der humane 3'-Prozessierungsfaktor CF II aus den beiden Proteinen hClp1 und hPcf11 rekonstituiert werden kann (Abbildung 4.1). Damit gelang es zum ersten Mal, diesen Prozessierungsfaktor biochemisch zu definieren und auf diese zwei Untereinheiten einzugrenzen. In CF II-depletierten und teilfraktionierten Kernextrakt aus HeLa-Zellen konnte der rekombinante CF II-Komplex die 3'-Spaltaktivität rekonstituieren. Die Spaltaktivität des CF II nahm in den Komplementationen in einer dosisabhängigen Weise zu. Die isolierte Untereinheit hClp1 konnte die 3'-Spaltaktivität in einer Vergleichsreaktion nicht vermitteln. Die Expression des isolierten hPcf11 war leider nicht möglich. Die Spaltaktivität wurde auch durch N-terminal verkürzte hPcf11-Varianten im CF II-Komplex vermittelt, wobei

hPcf11 Δ N769 hier die kürzeste, aktive Variante darstellte. In Übereinstimmung mit Daten zum yPcf11 war die N-terminale CID dementsprechend für die Spaltaktivität nicht notwendig (Sadowski u. a., 2003). Die für die Aktivität notwendige FEGP-Wiederholungsregion des hPcf11 stellt ein interessantes Merkmal dar, welches in Homologen aus Hefe nicht vorhanden ist. Interaktionsstudien mit anderen 3'-Prozessierungsfaktoren wären hierfür von besonderem Interesse. Weiterhin wurde in 3'-Spaltreaktionen mit den 5'-PNK inaktiven hClp1-Mutanten „KR“ (K127A, R288A, R293L) und „DR“ (D151A, R288A, R293L) in CF II gezeigt, dass die Spaltaktivität des CF II auch ohne die 5'-PNK-Aktivität vermittelt wurde. Auch diese Ergebnisse sind in Übereinstimmung mit Daten aus Versuchen mit Hefe CF IA, da yClp1 in seiner natürlich vorkommenden Variante keine 5'-PNK-Aktivität besitzt (Dikfidan u. a., 2014; Noble u. a., 2007).

Die RNA-bindende Eigenschaft des CF II ist eine unerwartete Entdeckung. Diese RNA-Bindungsaktivität konnte auf hPcf11 eingegrenzt werden. Auch wenn durch verschiedenen Ansätze, wie Bindungsstudien an SV40 *late* -RNA-Fragmenten (Abschnitt 2.6.2) und Bind'n'Seq-Analysen (Abschnitt 2.6.4), keine eindeutige Konsensus-Sequenz gefunden wurde, so zeigten G-reiche Sequenzen mit einer Mindestlänge von 14 Nukleotiden die besten Affinitäten zum hPcf11 im CF II-Komplex. Sie wurden mit einem K_{50} von ~ 2 nM durch CF II gebunden. Übereinstimmend damit wurde in später erschienen PAR-CLIP Studien eine Bindung von hPcf11 an G-reiche Abschnitte in 3'-UTRs von mRNAs gefunden (Kamieniarz-Gdula u. a., 2019; Ogorodnikov u. a., 2018).

5 Materialien

5.1 Chemikalien

Chemikalie	Hersteller
2-Mercaptoethanol, ≥99%	Sigma-Aldrich
3'-O-Me-m ⁷ G(5')ppp(5')G RNA Cap Structure Analog	New England Biolabs
2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure (HEPES, Ultrapure)	Gerbu
α- ³² P-UTP, 3000 Ci/mmol	Perkin Elmer
Acrylamid, 4X (vierfach kristallisiert)	Serva
Acrylamid 30% (w/v)	Roth
Acrylamid:Bisacrylamid 37,5:1, 30% (w/v)	Serva
Acrylamid:Bisacrylamid 29:1, 40% (w/v)	Serva
Acrylamid:Bisacrylamid 19:1, 40% (w/v)	Serva
Agar-Agar, Kobe 1, für die Mikrobiologie Carl	Roth
Agarose, UltraPure	Invitrogen
Ammoniumsulfat	Merck
Ammoniumcarbonat	Merck
Ammoniumformiat, 97%	ACROS Organics
Ampicillin, Natriumsalz, 99%	Carl Roth
Baktotrypton (Pepton aus Casein), tryptisch verdaut, für die Mikrobiologie	Carl Roth
Bovines Serumalbumin (BSA), Fraktion V, für biochemische Zwecke	Merck
Cellfectin II reagent (Transfektionsreagenz)	Invitrogen
Coomassie-Brilliant-Blau R-250, pure	Serva
Cordycepintriphosphat	Sigma-Aldrich
D-Desthiobiotin	IBA
Diethylpyrocarbonat (DEPC), 97% NMR	Sigma-Aldrich
Dithiothreitol (DTT), für Molekularbiologie	Gerbu
Essigsäure, 100%, p.a.	Carl Roth
Ethanol, ≥99,8%	Sigma-Aldrich
Formaldehyd, mindestens 36,5%, p.a.	Sigma-Aldrich
γ- ³² P-ATP, 6000 Ci/mmol	PerkinElmer
Gentamycinsulfat, für die Biochemie	Carl Roth
Glycerin, ≥99,5%, p.a.	Carl Roth
Glycin, premium grade	Serva

Chemikalie	Hersteller
Glycogen, from oyster, research grade	Serva
Harnstoff, ≥99,5%, p.a.	Carl Roth
Hefeextrakt, SERVABACTER	Serva
Isopropanol, ≥99,8%, p.a.	Carl Roth
Isopropyl-β-D-thiogalactopyranosid (IPTG)	Peqlab
Kalibrier-Kit für Gelfiltrationen (HMW)	Pharmacia
Kalibrier-Kit für Gelfiltrationen (LMW)	Pharmacia
Kanamycinsulfat	Carl Roth
Kreatinphosphat, Dinatriumsalz	Sigma-Aldrich
L-Arabinose, ≥99%	Sigma-Aldrich
Leupeptin	Roche
Magermilchpulver Sucofin	(Edeka)
Methanol, ≥99%	Carl Roth
N, N'-Methylenbisacrylamid (Bisacrylamid), ≥99,5%	Gibco
N, N, N', N'-Tetramethylethyldiamin (TEMED), p.a.	Merck
Natriumchlorid, p.a.	Carl Roth
Tetra-Natriumdiphosphat	Merck
Natriumdodecylsulfat (SDS)	Serva
Nonidet P-40-Ersatzprodukt (NP-40)	Fluka
PageRuler prestained protein ladder	Fermentas
PageRuler unstained protein ladder	Fermentas
Pepstatin	Roche
Phenol, Roti-Phenol (TE-gesättigt)	Carl Roth
Phenol, Roti-Aqua-Phenol (Wasser-gesättigt)	Carl Roth
Phenylmethylsulfonylfluorid (PMSF), ≥99%	Sigma-Aldrich
Plasmid Midi Kit	Qiagen
Plasmid Mini Prep Kit PeqGOLD I	Peqlab
Plasmid Mini Prep Kit PeqGOLD II	Peqlab
Proteaseinhibitor-Cocktail (PIC), EDTA-frei, in DMSO	Sigma-Aldrich
Ribonukleosidtriphosphate (NTPs)	Fermentas
Roentogen liquid (Entwicklerlösung)	Tetenal
Roentogen superfix (Fixierlösung)	Tetenal
Roti-Quant Bradford-Reagenz, für lösliche Proteine	Carl Roth
SOB-Trockenmedium, für Molekularbiologie	Carl Roth
Spectinomycin-dihydrochlorid-pentahydrat	Sigma-Aldrich

Chemikalie	Hersteller
Tetracyclin-hydrochlorid, 96%	Calbiochem
Trypanblau-Lösung, 0,4% in PBS	BioRad
Tween-20, für die Synthese	Merck
X-Gal	Peqlab

5.2 Verbrauchsmittel

Material	Hersteller
Chromatographiesäulen aus Polypropylen, 10 ml, mit Fritte	BioRad
Cryokonservierungsgefäße, 2 ml	Greiner
DE81-Papier	Whatman
Einweg-Zählkammern für Zellzähler TC20	BioRad
Einwegpipetten für die Zellkultur, 2 bis 50 ml	Corning
Filme für Autoradiographie, Amersham Hyperfilm MP	GE Healthcare
Filterpapier	Whatman
Nitrozellulose Membranfilter NC20 und RC58, Porengröße 0,2 µm	Whatman
Parafilm „M“	Pechiney
Photospektrometrie-Einwegküvetten, Plastibrand	Brand
Protran Nitrozellulosemembran	Whatman
Reaktionsgefäße, Quali-Mikrozentrifugengefäße, 1,5 ml	Kisker
Reaktionsgefäße, 2 ml, <i>safe-lock</i>	Eppendorf
Reaktionsgefäße, 5 ml, <i>safe-lock</i>	Eppendorf
Schnappdeckelröhrchen, 12 ml, Polypropylen	Greiner
Schraubdeckelgefäße, 15 und 50 ml	Greiner
Szintillationscocktail Irgasafe Plus	PerkinElmer
Ultrafiltrationseinheit Amicon Ultra-2, MWCO30, 2 ml	Merck
Ultrafiltrationseinheit Amicon Ultra-4, MWCO30, 4 ml	Merck
Zellkulturflaschen, 25, 75 und 150 cm ²	TPP
Zellkulturtestplatten, 6 Vertiefungen	TPP
Zellschaber, 13 mm Breite	TPP

5.3 Geräte

Gerät	Hersteller
Automatischer Zellzähler TC20	BioRad
Blotting-Apparatur, 20 x 20 cm, <i>semi dry</i>	OWL
Brutschränke	Heraeus Instruments
Cryopräservationskontainer 5100, „Mr. Frosty“	Nalgene
Elektroporationsgerät Eporator	Eppendorf
Elektroporationsküvetten, Elektrodenabstand 1 mm	CellProjects
Fluoreszenz- und Autoradiographie-Imager Typhoon 9200	GE Healthcare
<i>storage phosphor</i> Autoradiographiefilme	Molecular Dynamics
FPLC-System ÄKTA <i>purifier</i>	GE Healthcare
Fraktionssammler Frac-950	GE Healthcare
Gradient Master (Gradienten-Mischer)	BioComp
Infrarot <i>imaging</i> System Odyssey Classic	Li-Cor
Kammer für horizontale Gelelektrophorese	OWL
Kammern für vertikale Gelelektrophorese, Gelbreiten 10, 20 und 35 cm	OWL
Kühl-Tischzentrifuge 5180 R	Eppendorf
Kühl-Tischzentrifuge 5427 R	Eppendorf
Kühlzentrifuge Avanti J-25	Beckman
Rotor JA-25.50	Beckman
Rotor JA-16.250	Beckman
Mikroskop Axiovert 25	Carl Zeiss
MultiMode <i>reader</i> Tristar ² S, LB 942	Berthold Technologies
Mikrotestplatten mit 96 Vertiefungen, für Absorptionsmessungen	Greiner
Sterilfiltrationsapparaturen, 250 und 500 ml	Nalgene
Schüttelinkubatoren für Bakterien- und Insektenzellkulturen	Infors
Spektrophotometer U-3000	Hitachi
UV-Vis Quarzglasküvetten, 20, 80 und 200 µl (1 cm)	Hellma
Sterilbank HERASafe HS 15	Heraeus Instruments
Szintillationszähler Tri-Carb 2100TR, Liquid Scintillation Analyzer	Packard
Thoma-Zählkammer und geschliffene Deckgläschen	Brand
Tischzentrifuge, Biofuge <i>pico</i>	Heraeus Instruments
Ultraschallgerät und -sonde, Branson Sonifier 250	Branson
CD-Spektrometer J-815	Jasco
UV Stratalinker® 1800	Stratagene

Gerät	Hersteller
Viscotek TDA 305	Malvern Panalytical Instruments
Ultrazentrifuge L8-70M Rotor SW40-Ti	Beckmann

5.4 Chromatographiesäulen und -materialien

Material	Hersteller
Ni-NTA-Agarose, 50% Suspension	Qiagen
Poly(A)-Sephrose 4B Chromatographiematerial	Pharmacia
ResourceQ Anionenaustauschersäule, 1 ml	GE Healthcare
MonoQ Anionenaustauschersäule, 1 ml	GE Healthcare
Superdex200 pg HR 10/30 Größenausschlusschromatographiesäule, 210 ml	Pharmacia
Superose 6 Chromatographiematerial, <i>prep grade</i>	Pharmacia
Superose 6 3.2/300 Größenausschlusschromatographiesäule, 2,4 ml	GE Healthcare
Superose 6 32/300 Größenausschlusschromatographiesäule, 24 ml	GE Healthcare
Strep-Tactin-Sephrose, 50% Suspension	IBA
Protein A Sepharose CL-4B	GE-Healthcare

5.5 Software

Programm	Publisher
Inkscape, Version 0.92.0	Free Software Foundation, Inc. [13]
ICE, Version 1.0.8.1 (2014) Technologies	Berthold
ImageQuant, Version 5.2 (1999)	GE Healthcare
GraphRobot	Lei A. Wang, [11]
JalView	Waterhouse u.a. [12]
Lasergene, Version 4.00 (1999)	DNASTAR
Odyssey, Version 2.1.12 (2006)	Li-Cor
OMNISEC Software	Malvern Panalytical
Photoshop CS6, Version 13.0.1 (2012)	Adobe
SigmaPlot, Version 12.5 (2011)	Systat Software
SnapGene Viewer, Version 3.1.4 (2017)	GSL
Biotech	

Programm	Publisher
Unicorn (Chromatographiesteuerung), Version 5.11 (2006)	GE Healthcare
UV Solutions, Version 2.2 (2010)	Hitachi
Spectra Manager™ II Software (CD-Spektren)	Jasco
Weeder	(Zambelli u. a., 2014)
WEBLogo	(Crooks u. a., 2004)
Microsoft Office 2019	Microsoft, Inc.

5.6 DNA und Proteine

5.6.1 Bacmide und Baculoviren

Name	Kodierte ORFs	Ursprung
pBac-MycHis-WDR33- hFip1^{loxP}-F160-30^{Tn7}	MycHis-WDR33 (human) hFip1 (human) Flag-CPSF160 (human) CPSF30 (human)	Lars Schönemann
pBac-YFP^{loxP}	YFP	Lars Schönemann
pBac-YFP^{loxP}-CFII_m (KR)	YFP	diese Arbeit
K127A/R288A/R293L^{Tn7}	Strep-hPcf11 (human) His-hClp1 K127A/R288A/R293L (human)	
pBac-YFP^{loxP}-CFII_m (DR)	YFP	diese Arbeit
D151A/R288A/R293L^{Tn7}	Strep-hPcf11 (human) His-hClp1 D151A/R288A/R293L (human)	
pBac-YFP^{loxP}-CFII_m^{Tn7}	YFP Strep-hPcf11 (human) His-hClp1 (human)	diese Arbeit
pBac-YFP^{loxP}-CFII_m-499^{Tn7}	YFP His-hClp1 (human) Pcf11 (500-1555) (human)	Christian Tüting

Name	Kodierte ORFs	Ursprung
pBac-YFP^{loxP}-CFI_m-769^{Tn7}	YFP His-hClp1 (human) Pcf11 (770-1555) (human)	Christian Tüting
pBac-YFP^{loxP}-CFI_m-1184^{Tn7}	YFP His-hClp1 (human) Pcf11 (1185-1555) (human)	Christian Tüting
pBac-YFP^{loxP}-CFI_m-1352^{Tn7}	YFP His-hClp1 (human) Pcf11 (1353-1555) (human)	Christian Tüting
pBac-YFP^{loxP}-Clp1^{Tn7}	YFP His-hClp1 (human)	diese Arbeit
pBac-mCherry^{loxP}-CstF^{Tn7}	Strep-CstF 50k (human) CstF 64k (human) His-CstF77k (human)	Matthias Lorbeer
pBac-YFP^{loxP}-CFI_m^{Tn7}	His-CFI _m 25k (human) Strep-CFI _m 68k (human)	Matthias Lorbeer

5.6.2 Plasmide und Transkriptionsvorlagen

Name	Klon	Ursprung
pKL-Flag-hCPSF1-hCPSF4	11	Gudrun Scholz, Uwe Kühn
pUCDM-YFP-MCS2	1	Gudrun Scholz, Uwe Kühn
pBADZ-HisCre	21	Gudrun Scholz, Uwe Kühn
pFL-Strep-Pcf11	1	diese Arbeit
pFL-His-Clp1-Strep-Pcf11	48	diese Arbeit
pFL-His-Clp1	2	diese Arbeit
pFL-His-Clp1 R288A /R293L-StrepII-Pcf11	8	diese Arbeit
pFL-His-Clp1-K127A/R288A /R293L-StrepII-Pcf11	2	diese Arbeit
pFL-His-Clp1 D151A/R288A R293L-StrepII-Pcf11	1	diese Arbeit
pGM10-His-Clp1 R288A R293L	4	diese Arbeit
pFL-His-Clp1 D151A	6	diese Arbeit
pFL-His-Clp1 K127A	15	diese Arbeit

Name	Klon	Ursprung
pBR322, <i>HpaII</i>-verd., dephos.		Gudrun Scholz, Uwe Kühn
pSP64-SV40 <i>late</i>		Anne Knoth
<i>DraI</i>-verdaut		
pSP64-SV40<i>late</i> Δ		Anne Knoth
<i>DraI</i>-verdaut		
pSP64-L3	15*	Anne Knoth
<i>DraI</i>-verdaut		
pSP64-L3 Δ	16	Anne Knoth
<i>DraI</i>-verdaut		
pSP64-L3pre		Anne Knoth
<i>DraI</i>-verdaut		
pSP64-L3pre Δ		Anne Knoth
<i>DraI</i>-verdaut		
T7-SV40 <i>late</i> 191-230	-	Invitrogen

5.6.3 DNA-Primer

Name	Sequenz 5'→3'	Ursprung/ Hersteller
NcoI StrepII for	CATGTGGTCTCACCCACAATTTGAAAAAGG	Invitrogen
NcoI StrepII rev	CATGCCTTTTTCAAATTGTGGGTGAGACCA	Invitrogen
StrepII BamHI for	GATCCATGTGGTCTCACCCACAATTTGAAAAA	Invitrogen
StrepII BamHI rev	GATCTTTTTCAAATTGTGGGTGAGACCACATG	Invitrogen
BamHI His8 for	GATCCATGTCTCATCATCATCATCATCACCACGGC	Invitrogen
BamHI His8 rev	GATCGCCGTGGTGATGATGATGATGATGATGAGACATG	Invitrogen
Sall StrepII SacII	TCGACATGTGGTCTCACCCACAATTTGAAAAGGCCCGC	Invitrogen
SacII StrepII Sall	GGGCCTTTTTCAAATTGTGGGTGAGACCACATG	Invitrogen
NheI Clp1 for	AGGATCCGCTAGCATGGCTCACCATCACCATC	Invitrogen
NsiI Clp1 rev	TTTATGCATCTACTTCAGATCCATGAACC	Invitrogen
T7-SV40 <i>late</i> 191-230	AAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAC CCTATAGTGAGTCGTATTA	Invitrogen
T7 Primer for	TAATACGACTCACTATAGGG	Invitrogen
clp1 K127A forw	CCCCACTGATGTGGGCGCGTCGACAGTGTGTCGCCTTC	Invitrogen
clp1 K127A rev	GAAGGCGACACACTGTTCGACGCGCCACATCAGTGGGG	Invitrogen

Name	Sequenz 5'->3'	Ursprung/ Hersteller
clp1 D151A NheI forw	TATGTGGAGCTAGCTGTGGGCCAGGGT	Invitrogen
clp1 D151A NheI rev	ACCCTGGCCCACAGCTAGCTCCACATA	Invitrogen
clp1 R288A R293L forw	GGGTGTGGTGGAGGCCTCCAAGGACTTCCTACGTGAATGTA GGG	Invitrogen
clp1 R288A R293L rev	CCCTACATTCACGTAGGAAGTCCTTGGAGGCCTCCACCACA CCC	Invitrogen

5.6.4 Synthetische und enzymatisch hergestellte RNAs

Name	Sequenz 5'->3'	Ursprung/ Hersteller
rG₁₀	GGGGGGGGGG	biomers.net GmbH
rG₁₁	GGGGGGGGGGG	biomers.net GmbH
rG₁₂	GGGGGGGGGGGG	biomers.net GmbH
rG₁₃	GGGGGGGGGGGGG	biomers.net GmbH
rG₁₄	GGGGGGGGGGGGGG	biomers.net GmbH
rG₁₇	GGGGGGGGGGGGGGGG	biomers.net GmbH
rU₁₄	UUUUUUUUUUUUUUU	biomers.net GmbH
rC₁₄	CCCCCCCCCCCCCCC	biomers.net GmbH
rA₁₄	AAAAAAAAAAAAAAAA	Poly(A)- Hydrolyse
Wigley10 wt (W10)	AAUAAACCCA	MWG Eurofins
Wigley10 Δ (W10 Δ)	AAgAAACCCA	MWG Eurofins

Name	Sequenz 5'->3'	Ursprung/ Hersteller
5-Jodo-Uracil- Wigley10 wt (iW10)	AA- ⁵ JU-AAACCCA	biomers.net GmbH
5-Jodo-Uracil- Wigley10 Δ (iW10 Δ)	Ac- ⁵ JU-cAAACCCA	biomers.net GmbH
Wigley -23/-10 wt	CUGCAAUAAACAAGUUAA	MWG Eurofins
SV40 late 191-230	gggUUUAUGUUUCAGGUUCAGGGGGAGGUGUGGGAGGUU UUUU	T7-RNA Pol. Transkript
SV40 late 190-225	UUUUAUGUUUCAGGUUCAGGGGGAGGUGUGGGAGGU	biomers.net GmbH
SV40 late 202-224	GGUUCAGGGGGAGGUGUGGGAGG	biomers.net GmbH
SV40 late 208-225	GGGGGAGGUGUGGGAGGU	biomers.net GmbH
SV40 late 208-224	GGGGGAGGUGUGGGAGG	biomers.net GmbH
SV40 late 208-221	GGGGGAGGUGUGGG	biomers.net GmbH
Wt2x	GGGGGAGGUGgGGGAGG	biomers.net GmbH
Δ2x	GGGaaAGGUGgGaaAGG	biomers.net GmbH
ΔΔ2x	GaGaGAGGUGaGaGAGG	biomers.net GmbH
CstF SELEX	UGUGUUUUUAUUGUGU	MWG Eurofins
BS2	<u>GGGUUUAAGUGGGUCACGUUUAAAC</u>	biomers.net GmbH
BS2-Δ1	<u>GGGUUUAAGUGGGUCACG</u>	biomers.net GmbH
BS2-Δ2	CCCUUUAAGUGGGUCAC <u>G</u>	biomers.net GmbH

Name	Sequenz 5'→3'	Ursprung/ Hersteller
BS2-Δ3	<u>UUUAAGUGGGUCACG</u>	biomers.net GmbH
BS3 CFII	<u>GGGUUACGCGUGAGGACG</u>	biomers.net GmbH
BS3 Pcf11	<u>GGGUUUAGUGCGCACGUG</u>	biomers.net GmbH
BS3 Weeder	<u>GGGUUUAGUGGGGGGGUG</u>	biomers.net GmbH

5.6.5 Antikörper

Name	Ursprung	Verdünnung	Herkunft
α-CF I-25, „CF I 30k“	Kaninchen	1:2000	AG Keller, Universität Basel
α-Pcf11, #274	Kaninchen	1:2000	AG Keller, Universität Basel
α-Clp1, #171	Kaninchen	1:2000	AG Keller, Universität Basel
α-CPSF30, SA1368	Kaninchen	1:2000	Gudrun Scholz, Uwe Kühn
α-CPSF73, #009	Kaninchen	1:2000	AG Keller, Universität Basel
α-CPSF100, SA1366	Kaninchen	1:2000	Gudrun Scholz, Uwe Kühn
α-CPSF160, #105	Kaninchen	1:2000	AG Keller, Universität Basel
α-CstF-77, monoklonal	Maus	1:2000	Abnova
α-Flag, M2, monoklonal	Maus	1:1000	Sigma-Aldrich
α-hFip1, SZ1005	Kaninchen	1:2000	AG Keller, Universität Basel
α-Strep, monoklonal	Maus	1:1000	IBA
α-Symplekin, A301-465A	Kaninchen	1:2000	Biomol
α-WDR33, A301-152A	Kaninchen	1:2000	Biomol
α-Kaninchen, IRDye 800CW	Esel	1:15000	Li-Cor
α-Maus, IRDye 800CW	Ziege	1:15000	Li-Cor

5.6.6 Proteinpräparationen

Name, Fraktion	Stammkonzentration [μM]	Ursprung
CF I, I Ni-NTA-Strep #E4	0,5 μM	diese Arbeit
CF I, V Ni-NTA RQ #12-14	11,5 μM , 13,6 μM , 15 μM	diese Arbeit
CstF, I Ni-NTA_RQ #A11-13	1,5 μM , 3 μM , 2 μM	diese Arbeit
CF II, VII Ni-NTA_MonoQ #18-20	5 μM , 11 μM , 8 μM	diese Arbeit
Methyliertes BSA	16,4 mg/ml	Uwe Kühn
Poly(A)-Polymerase wt	10,93 μM	Anne Knoth, Gudrun Scholz
mPSF, I. Flag E2	0,69 μM	Lars Schönemann
CPSF IV, präpariert aus Kalbsthymus		Gudrun Scholz, Uwe Kühn
CPSF (7er-Komplex), II_poly(A) #A9	0,5 μM	Lars Schönemann
CF II K127A/R288A/R293L, I Ni-NTA_Mono Q #18	1,8 μM	diese Arbeit
CF II D151A/R288A/R293L, I Ni-NTA_Mono Q #B16	2,8 μM	diese Arbeit
hClp1, III Ni-NTA_Superdex200 #18	2,0 μM	diese Arbeit
CFII ΔN 499, I Ni-NTA_MonoQ #14	4,8 μM	Christian Tüting
CFII ΔN 769, I Ni-NTA_MonoQ #20	1,7 μM	Christian Tüting
CFII ΔN 1123, I Ni-NTA_MonoQ #11	3,3 μM	Christian Tüting
hPcf11 ΔN 1184, Ni-NTA MonoQ #19-23	1,5 μM	Christian Tüting
His-Sumo-hPcf11 zcz	1,2 μM	Christian Tüting
His-Sumo-FEGP Ni-NTA_MonoQ #29	0,8 μM	Christian Tüting
GST-hPcf11 zcz GSH-Seph_MonoQ #30	4,8 μM	Christian Tüting

5.6.7 Kommerziell erworbene Proteine

Name	Hersteller
Calf intestinal phosphatase (CIP)	Roche
Cre-Rekombinase und Reaktionspuffer	New England Biolabs
Proteinase K, lyophilisiert	Merck
Murine RNase-Inhibitor, RNi	New England Biolabs
Restriktionsendonukleasen und Reaktionspuffer	New England Biolabs
SP6 RNA-Polymerase und Reaktionspuffer	Roche
T7 RNA-Polymerase und Reaktionspuffer	New England Biolabs
T4 5'-Polynukleotidkinase (PNK) und Reaktionspuffer	New England Biolabs
T4 DNA-Ligase	New England Biolabs
HeLa Nukleärer Extrakt	IPRA Cell (Belgien)

5.7 Zellstämme und Medien

5.7.1 Bakterien

Stamm	Genotyp	Ursprung/Hersteller
BW23474	<i>E. coli</i> , F-, $\Delta(\text{argF-lac})169$, $\Delta\text{uidA4}::\text{pir-116}$, <i>recA1</i> , <i>rpoS396</i> (Am), <i>endA9</i> (del-ins)::FRT, <i>rph-1</i> , <i>hsdR514</i> , <i>rob-1</i> , <i>creC510</i> [Wanner & Haldimann, 2004]	Propagation von Donorplasmiden des MultiBac-Systems (pUCDM und pSPL), von T. J. Richmond, ETH Zürich
DH10B	<i>E. coli</i> , F-, <i>mcrA</i> , $\Delta(\text{mrr-hsdRMS-mcrBC})$, $\phi 80\text{lacZ}\Delta\text{M15}$, ΔlacX74 , <i>recA1</i> , <i>endA1</i> , <i>araD139</i> , $\Delta(\text{ara, leu})7697$, <i>galU</i> , <i>galK</i> , $\lambda\text{-rpsL}$, <i>nupG</i> /pMON14272/ pMON7124	
DH10-MultiBac	<i>E. coli</i> , F-, <i>mcrA</i> , $\Delta(\text{mrr-hsdRMS-mcrBC})$, $\Phi 80\text{lacZ}\Delta\text{M15}$, ΔlacX74 , <i>recA1</i> , <i>endA1</i> , <i>araD139</i> , $\Delta(\text{ara, leu})7697$, <i>galU</i> , <i>galK</i> , $\lambda\text{-rpsL}$, <i>nupG</i> /pMON14272/ pMON7124	Stamm enthält das MultiBac-Bacmid, von T. J. Richmond und Kollegen, ETH Zürich

Stamm	Genotyp	Ursprung/Hersteller
DH10-	<i>E. coli</i> , F-, <i>mcrA</i> , $\Delta(mrr\text{-}hsdRMS\text{-}mcrBC)$, $\Phi 80lacZ\Delta M15$,	Stamm DH10-
MultiBac^{Cre}	$\Delta lacX74$, <i>recA1</i> , <i>endA1</i> , <i>araD139</i> , $\Delta(ara, leu)7697$, <i>galU</i> , <i>galK</i> , λ - <i>rpsL</i> , <i>nupG</i> /pMON14272/ pMON7124	MultiBac transformiert mit Plasmid pBADZ- HisCre

5.7.2 Eukaryotische Zelllinien

Name	Ursprung/Genotyp	Hersteller
Sf21	IPLB-Sf21AE, Immortalisierte Zelllinie aus Ovarzellen von <i>Spodoptera frugiperda</i> [Vaughn <i>et al.</i> , 1977]	Invitrogen
High-Five	BTI-TN-5B1-4, Immortalisierte Zelllinie aus Ovarzellen von <i>Trichoplusia ni</i> [Wickham <i>et al.</i> , 1992]	Elena Conti, Abteilung Strukturbiologie, MPI Martinsried

5.7.3 Kulturmedien

Name	Zusammensetzung	Ursprung/Hersteller
LB	10 g/l Baktotrypton, 5 g/l Hefeextrakt, 10 g/l Natriumchlorid, pH mit Natriumhydroxid auf 7,5 eingestellt, autoklaviert	Carl Roth
LB-Agar	10 g/l Baktotrypton, 5 g/l Hefeextrakt, 10 g/l Natriumchlorid, pH mit Natriumhydroxid auf 7,5 eingestellt, 15 g/l Agar-Agar, autoklaviert	Carl Roth
SOC	30,7 g/l SOB-Trockenmedium (Roth), 20 mM Glukose, autoklaviert	Carl Roth
TY (Niedrigsalz)	10 g/l Baktotrypton, 5 g/l Hefeextrakt, 5 g/l Natriumchlorid, pH mit Natriumhydroxid auf 7,5 eingestellt, autoklaviert	Carl Roth
ExCell 420 SFM	Serumfreies Medium für die Kultivierung von Insektenzellen, mit L- Glutamin, sterilfiltriert	Sigma-Aldrich

5.7.4 Medienzusätze und Antibiotika

Name	Arbeitskonzentration	Ursprung/Hersteller
Ampicillin	100 µg/ml	Carl Roth
Chloramphenicol	30 µg/ml	Merck
Gentamycin	10 µg/ml	Carl Roth
IPTG	0,5 mM	Peqlab
Kanamycin	30 µg/ml	Carl Roth
L-Arabinose	1 g/l	Sigma-Aldrich
Spectinomycin	50 µg/ml	Sigma-Aldrich
Tetracyclin	12,5 µg/ml	Calbiochem
X-Gal	200 µg/ml	Peqlab
Zeocin (Lösung)	25 µg/ml	Invitrogen

6 Methoden

6.1 Mikro- und molekularbiologische Methoden

6.1.1 Sterilisation

Lösungen, Nährmedien und Materialien wurden, soweit möglich, durch Autoklavieren bei 121 °C für 30 Minuten sterilisiert. Nicht hitzebeständige Lösungen wurden durch Filtration über Nitrozellulosefilter mit 0,2 µm Porengröße sterilisiert. Plastikgeräte wurden durch UV-Bestrahlung oder Bacillol sterilisiert.

6.1.2 Herstellung und Transformation elektrokompenter Bakterien

Für die Transformation von Plasmid-DNA in Bakterien wurden elektrokompente *E. coli* Bakterien verwendet. Dazu wurden zunächst Bakterien auf selektiven LB-Agar Platten ausplattiert und vereinzelt. Nach entsprechender Inkubation bei 37 °C wurde eine Einzelkolonie in LB-Medium mit dem entsprechenden Selektionsmittel überführt und eine Vorkultur über Nacht bei 37 °C angezogen.

Aus dieser Vorkultur wurden 1–5 ml genutzt, um eine 400 ml LB-Kultur zu inokulieren. Diese wurde bei 37 °C inkubiert und bis zu einer OD₆₀₀ von 0,6–0,8 angezogen. Nach Erreichen der optischen Dichte wurde die Kultur 15 Minuten auf Eis gekühlt, geerntet und 15 Minuten bei 4 °C und 4000xg zentrifugiert. Das Pellet wurde in 1 Volumen (400 ml) eiskalter 10%iger Glycerinlösung resuspendiert. Dieser Schritt wurde dreimal mit geringer werdenden Volumina wiederholt (200 ml; 40 ml; 0,8 ml). Abschließend wurden die kompetenten Zellen in 50 µl Aliquots in Stickstoff eingefroren und bei -80 °C gelagert.

Die Transformation dieser kompetenten Zellen erfolgte dann durch Elektroporation. Dazu wurde ein Aliquot der Zellen mit 1–5 ng Plasmid DNA (100–200 ng Bacmid DNA) versetzt und 5–10 Minuten auf Eis inkubiert. Anschließend erfolgte die Elektroporation in dafür geeigneten Küvetten mit einem Stromstoß von 5 ms Dauer (1800 V bei 1 mm Spaltbreite; 2500 V bei 2 mm Spaltbreite). Anschließend wurden die Zellen in 950 µl 37 °C warmen SOC-Medium resuspendiert und für ca. eine Stunde bei 37 °C inkubiert. Anschließend wurden die Zellen in verschiedenen Verdünnungen auf selektiven LB-Agaplatten ausgestrichen und positive Klone selektiert.

6.1.3 Bakterielle Glycerinkulturen

Für die langfristige Lagerung von Bacmid-DNA wurden *E. coli* Kulturen mit dem entsprechenden Bacmid als Glycerinkulturen bei -80 °C gelagert. Dafür wurden 750 µl einer Übernachtskultur mit 250 µl steriler Glycerinlösung (60% v/v) vermischt und in Stickstoff eingefroren.

6.1.4 Verwendung des MultiBac Systems

Das MultiBac-System ist ein Klonierungs- und Expressionssystem zur Erzeugung rekombinanter Proteine in Insekten-Zellkultur (Berger u. a., 2010; Fitzgerald u. a., 2006). Im folgenden Abschnitt wird die allgemeine Arbeitsweise für diese spezialisierte Form gentechnischer Arbeit beschrieben.

6.1.5 Klonierung und Multiplikation von Zielgenen in Transfervektoren

Für die Klonierungsarbeit mit dem MultiBac-System können neben den vier in Abbildung 6.1 A gezeigten Vektoren pKL, pFL, pUCDM und pSPL auch pFBDM und pKDM verwendet werden. Alle Plasmide enthalten zwei *multiple cloning sites* (MCS), jeweils unter Kontrolle der viralen Promotoren polH oder p10. So können Zielgene schnell und zielgerichtet mittels Restriktion und Ligation (NEB Enzyme) in die Vektoren kloniert werden. Die Plasmide werden je nach Replikationsursprung in Donor- (pUCDM und pSPL; konditionaler R6K γ -Replikationsursprung) und Akzeptorplasmide (pKL, pKDM und pFBDM, pFL; Cole1-, bzw. BR322-Replikationsursprung) unterschieden. Zur Generierung von Vektoren mit mehr als zwei ORFs können zwei verschiedene Ansätze verfolgt werden. Zum einen können mittels Restriktion die Multiplikationsmodule (M) aus den Akzeptorplasmiden ineinander kloniert werden (Abbildung 6.1 C), zum anderen besteht die Möglichkeit zur Fusion der Donorplasmide mit den Akzeptorplasmiden pFL und pKL mittels *in vitro* Cre-loxP Rekombination (loxP-Elemente in den entsprechenden Vektoren s. Abbildung 6.1 B).

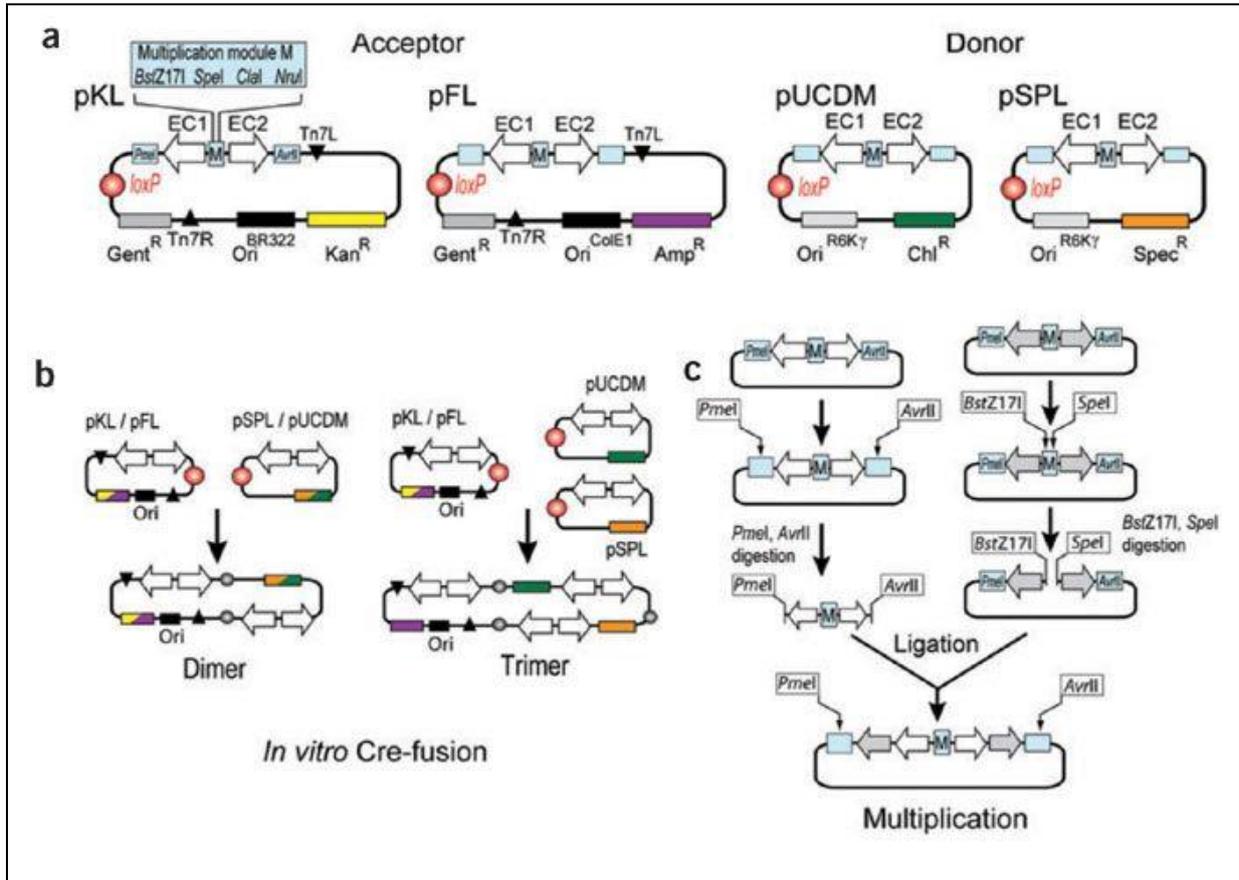


Abbildung 6.1 Schematische Darstellung der MultiBac-Klonierungsvektoren.

A) Schematische Übersicht der (*loxP*) Akzeptorplasmide pKL und pFL sowie der (*loxP*) Donorplasmide pUCDM und pSPL. **B)** Schematische Darstellung der *loxP*-Integration durch die Cre-*lox*-Rekombinase eines oder zweier Donorplasmide in ein Akzeptorplasmide zur Generierung von Vektoren mit mehreren ORF. **C)** ORF-Multiplikationsstrategie in Akzeptorplasmiden durch Restriktionsschnittstellen. Verschiedene Multiplikationsmodule können durch Verschachtelung ineinander integriert werden und Akzeptorplasmide mit mehreren ORFs generieren. Entnommen aus (Fitzgerald u. a., 2006).

6.1.5.1 In vitro Fusion von Transfervektoren via Cre-*loxP* Rekombination

Die Cre-*loxP* Rekombination erfolgte im Wesentlichen nach dem vorgegebenen Protokoll. Dazu wurde folgender Reaktionsansatz für eine *in vitro* Reaktion vorbereitet:

Cre-*loxP*-Rekombination

- 1x Cre-Reaktionspuffer
- 500 ng Donorplasmid (je)
- 500 ng Akzeptorplasmid
- 1 U Cre-Rekombinase
- ddH₂O ad 10 µl

Dies wurde 30 Minuten bei 37 °C inkubiert und 1–5 µl der Reaktion in elektrokompetente *E. coli* transformiert. Positive Klone wurden über die geeignete Kombination aus Antibiotika selektiert. Die erfolgreiche Integration wird hier durch den konditionellen R6K γ -Replikationsursprung der Donorplasmide kontrolliert. Dieser wird in den *E. coli* Zellen nicht repliziert.

6.1.5.2 *In vivo* Fusion von Transfervektoren mit Bacmiden via Cre-loxP Rekombination

Für die *in vivo* Reaktion der Cre-loxP-Rekombination müssen zunächst elektrokompetente DH10-MultiBac Zellen hergestellt werden, welche die Cre-Rekombinase exprimieren. Dazu wurden DH10-MultiBac Zellen zunächst mit dem Plasmid pBADZ-HisCre transformiert und mittels Selektion auf Zeocin, Kanamycin und Tetracyclin identifiziert. Die folgende Herstellung der elektrokompetenten Zellen erfolgte im Wesentlichen nach dem oben beschriebenen Arbeitsablauf (6.1.2) unter Verwendung geeigneter Antibiotika. Die Expression der Cre-Rekombinase wird dabei bei einer OD₆₀₀ von 0,25 über die Zugabe von 0,1% (w/v) Arabinose induziert und ab OD₆₀₀ von 0,5 geerntet und weiterbearbeitet.

Konstrukte in Donorplasmiden können anschließend via *in vivo* Cre-loxP-Rekombination in das MultiBac-Baculovirusgenom integriert werden. Dazu werden 250 ng der entsprechenden Plasmide in elektrokompetente DH10-MultiBacCre-Zellen transformiert. Die Erholungszeit in SOC-Medium wird im Anschluss auf sechs bis acht Stunden ausgedehnt, um die Rekombinationsreaktion zu gestatten. Die Selektion positiver Klone erfolgt durch entsprechende Antibiotika und Blau-Weiß-Selektion (blaue Kolonien sollten hierbei gewählt werden). Positive Klone können dann entweder direkt zur Gewinnung von Bacmid-DNA (Abschnitt 5.10.9) oder zur erneuten Herstellung elektrokompetenter Zellen für die Integration weiterer Transfervektoren via Tn7-Transposition genutzt werden.

6.1.5.3 Integration von Transfervektoren in Bacmide via Tn7-Transposition

Elektrokompetente DH10B-MultiBac *E. coli* werden mit 250 ng der entsprechenden Akzeptorplasmide transformiert. Im Anschluss erfolgt eine Erholung der Transformanten in SOC-Medium über sechs bis acht Stunden bei 37 °C im Schüttler. In dieser Zeit erfolgt die Tn7-Transposition der Plasmide in das Bacmid-Genom durch die in den DH10B-MultiBac-Bakterien exprimierten Tn7-Transposase. Im Anschluss erfolgt die Selektion positiver Transformanten auf geeigneten LB-Agar Platten mit den benötigten Antibiotika. Eine Blau-Weiß Selektion (weiß = positiv) gestattet hierbei die erfolgreiche Integration in das Bacmid zu erkennen. Positive

Kolonien wurden im Anschluss zweimal wie zuvor vereinzelt, um Mischklone oder falsch positive Klone auszuschließen.

6.1.5.4 Isolation rekombinanter Bacmid-DNA

Bacmid-DNA wurde aus positiven Transformanten über alkalische Lyse mit anschließender Isopropanol-Fällung isoliert. Als Vorbereitung dafür wird eine 5–6 ml Übernachtskultur des Klons in LB-Medium mit den entsprechenden Antibiotika inokuliert.

Puffer I	Puffer II (Lysepuffer)	Puffer III (Neutralisationspuffer)
50 mM Tris-HCl, pH 8,0 10 mM EDTA 100 µg/ml RNase A	200 mM Natriumhydroxid 1% (w/v) SDS	3 M Kaliumacetat, pH 5,5

Die Zellen werden geerntet (15 Minuten, 4000xg, 4 °C), der Überstand verworfen und das Pellet in 250 µl Puffer I resuspendiert. Durch Zugabe von 250 µl Puffer II erfolgt im Anschluss die Lyse für eine bis zwei Minuten bei Raumtemperatur. Dies wird durch Zugabe von 350 µl Puffer III neutralisiert. Proteine und genomische DNA werden im Anschluss pelletiert (20 Minuten, 16.000xg). Der Überstand wird anschließend mit 0,8 Volumen Isopropanol versetzt, für 15 Minuten auf Eis inkubiert und die Bacmid-DNA abschließend, durch Fällung (20 Minuten, 16.000xg), präzipitiert. Der Überstand wird verworfen und das DNA-Pellet mit 80%igem Ethanol gewaschen, möglichst steril getrocknet und in 50 µl steriler HEPES-Lösung (10 mM, pH 7) resuspendiert. Diese kann dann bei 4 °C für etwa ein Jahr gelagert und für die Transfektion von Sf21-Zellen genutzt werden.

6.2 Zellkultur

Die Arbeiten mit der Sf21 und High Five™ Insekten-Zellkultur orientierte sich maßgeblich an den Vorgaben des Vertreibers ThermoFischer [1]. Generell muss jedoch beachtet werden, dass nur Sf21 Zellen für Virusgeneration aus Bacmiden und Viruspropagation geeignet sind. Proteinexpression kann in beiden Zellkulturen erfolgen.

6.2.1 Aussaat und Propagation von Insektenzellen

Ausgehend von gefrorenen Stammkulturen (-180 °C, *Cryo Stock*) sollten Sf21 Kulturen zunächst als semiadhärente Kulturen ausgesät werden. Dazu werden $\sim 1 \cdot 10^7$ Zellen im Wasserbad

aufgetaut und in einer 75 cm² Kulturflasche in 10 ml ExCell Medium resuspendiert. Nach 30 Minuten bis 60 Minuten Inkubation bei 27 Grad sollten die Zellen sich abgesetzt haben und der Überstand wird abgenommen. Nach Zugabe 25 ml frischen Mediums werden die Zellen bei 27 Grad kultiviert. Bei Erreichen von 90–100% Konfluenz wurden die Zellen mit einem Zellschaber von der Oberfläche gelöst und im Verhältnis 1:4 in neues Medium verdünnt. Dabei darf die Zelldichte nicht unter ~20% Konfluenz fallen.

Sind genügend Zellen vorhanden kann die adhärenente Kultur in eine Suspensionskultur überführt werden. Bei adaptierten Sf21 mit genügend Zellen kann ausgehend vom *Cryo Stock* direkt zu diesem Schritt übergegangen werden. Etwa $2,5-3 \cdot 10^7$ Zellen sollten dazu vorhanden sein. Diese werden auf eine Zelldichte von $1 \cdot 10^6$ Zellen/ml verdünnt und in einen sterilen Erlenmeyerkolben entsprechender Größe überführt. High 5-Kulturen können, ohne adhärenentes Wachstum, direkt in Suspensionskultur überführt werden. Die Zelldichte sollte dabei $0,8 \cdot 10^6$ Zellen/ml in der ersten Passage nicht unterschreiten.

Die Kultivierung der Insektenzellen in ExCell-Medium erfolgt bei 27,5 °C und 125 rpm. Dabei sollte alle zwei bis drei Tage die Zelldichte und Viabilität der Zellen überprüft werden. In der fortlaufenden Kultivierung sollten die Zellen im Anschluss auf $0,8 \cdot 10^6$ (Sf21) oder $0,5 \cdot 10^6$ (High 5) mit frischem Medium verdünnt werden.

6.2.2 Bestimmung der Zelldichte von Zellsuspensionen

Die Zelldichte wurde in dieser Arbeit mit zwei Methoden bestimmt. Zum einen wurde durch die manuelle Zählung in einer Thoma-Neubauer Zählkammer unter dem Lichtmikroskop die Zelldichte bestimmt und zum anderen durch die automatische Zählung mit dem TC20 Zellzähler (Biorad, [2]). In beiden Verfahren wurde für die Zählung eine Vitalfärbung zur Bestimmung der Viabilität durchgeführt. Dafür wurde die Zellsuspension 1:1 mit einer Trypanblau-Lösung (0,4% w/V in PBS) gemischt. Dieser Farbstoff kann Zellmembranen nicht durchdringen. Intakte Zellen erscheinen somit nicht gefärbt, während tote Zellen blau eingefärbt werden. Ein Verhältnis von intakten zu toten Zellen von mindestens 90% ist dabei ein Indikator für eine stabile Kultur.

6.2.3 Herstellung von tiefgefrorenen Vorratskulturen (*cryo stocks*)

Zur langfristigen Lagerung von Insektenzellen wurden diese als *cryo-stock* in flüssigem Stickstoff gelagert. Dafür wurden Zellen in einer frühen Passage einen Tag nach dem Umsetzen gezählt und durch Zentrifugation für 10 Minuten bei 200 xg geerntet. Das Zellpellet wurde dann auf $2 \cdot 10^7$ Zellen/ml mit Cryopräservationsmedium (45% altes Medium, 45% frisches Medium, 10% DMSO)

resuspendiert. Je $2-4 \cdot 10^7$ Zellen wurden dann in Cryokonservierungsgefäße aliquotiert und im Cryopräservationskontainer „Mr. Frosty“ auf -80 °C abgekühlt (-1 °C/min).

6.2.4 Transfektion von Sf21-Zellen mit Bacmid-DNA

Isolierte Bacmid-DNA wurde mittels Cellfectin II (Invitrogen, [3]) in Sf21-Zellen transfiziert. Dazu wurden im Wesentlichen die Herstellerangaben beachtet.

Pro Transfektion wurden $1,5 \cdot 10^6$ Sf21-Zellen in einer 6-Well Platte ausgesät und für mindestens 30 Minuten bei $27,5\text{ °C}$ inkubiert, um ein adhärentes Anheften an den Boden zu gestatten. Parallel dazu wurde ein Transfektionsgemisch aus $200\text{ }\mu\text{l}$ Ex-Cell Medium, $8\text{ }\mu\text{l}$ Cellfectin II und $10\text{ }\mu\text{l}$ Bacmid-Präparation hergestellt und für 20 Minuten bei Raumtemperatur inkubiert. Nach der Inkubation wurde dies auf einen Milliliter mit dem Medium aufgefüllt, der Zellüberstand im 6-Well abgenommen und die ein Milliliter Bacmid-Lösung auf die Zellen gegeben. Nach drei bis fünf Stunden Inkubation bei $27,5\text{ °C}$ wurde die Bacmid-Lösung abgenommen und drei Milliliter neues Medium auf die Zellen gegeben. Nach 72 Stunden Inkubation bei $27,5\text{ °C}$ konnte der Überstand vorsichtig abgenommen werden. Dieser Überstand enthielt die gesuchten Baculoviren und konnte als initialer Virustiter (V_0) bei 4 °C für mehrere Monate gelagert werden. Die verbliebenen Zellen wurden mit neuem Medium überschichtet und nach weiteren zwei Tagen Inkubation auf das gesuchte Protein hin untersucht.

6.2.4.1 Kontrolle der Transfektion

Der Erfolg der Transfektion wurde auf zwei möglichen Wegen überprüft. Zum einen kann indirekt mittels Fluoreszenz noch in lebenden Zellen überprüft werden, ob ein Reporter-Fluoreszenzprotein in den Zellen exprimiert wird, welches im Bacmid-Genom integriert vorliegen muss. Die YFP-Detektion erfolgte dabei unter folgenden Einstellungen im Fluoreszenzscanner: Anregung 523 nm , Detektion 526 nm , Detektorspannung 600 V . Eine Detektion spricht hierbei für eine erfolgreiche Infektion der Zelle mit dem Baculovirus.

Zum anderen kann ein direkter Nachweis der gesuchten Zielproteine über SDS-PAGE und Western-Blot erfolgen. Hierzu werden die Zellen nach der fünftägigen Inkubation mit 1 ml PBS gewaschen und vom Well-Boden gelöst. Die Zellen werden durch kurze Zentrifugation bei 200 xg geerntet und anschließend mit $100\text{ }\mu\text{l}$ Lysis-Puffer versetzt und aufgeschlossen, indem sie für 30 Minuten auf Eis inkubiert werden.

Lysepuffer:

300 mM NaCl
30 mM Tris-HCl, pH 7,5
1% (V/V) NP-40
1 mM DTT und PMSF
1 µg/µl Leupeptin und Pepstatin

6.2.4.2 Viruspropagation und Virustiter

Für die Vermehrung des Virus aus dem initialen Virustiter V_0 wurden 100 ml SF21-Zellen in Suspensionskultur bei $0,8 \cdot 10^6$ Zellen/ml vorgelegt. Diese wurden dann mit 0,1–0,5% des Kulturvolumens an V_0 versetzt und so infiziert. Das Wachstum, die Viabilität und die Zellgröße wurden täglich verfolgt und die Kultur stets auf $1 \cdot 10^6$ Zellen/ml verdünnt. Sobald ein Arrest des Wachstums zu beobachten war, wurden nur noch die Viabilität und die Zellgröße überwacht. Bei $\leq 50\%$ lebenden Zellen und einem mittleren Durchmesser von ≥ 20 µm wurden die Zellen geerntet (fünf Minuten, 1000 xg) und der Überstand als neuer Virustiter weiterverwendet.

6.2.5 Baculovirale Proteinexpression in Sf21/High5-Zellen

Die Expression von Proteinen und Komplexen in Sf21-Zellen erfolgte, indem Zellen mit einer Dichte von $1,5 \cdot 10^6$ Zellen/ml ausgesät wurden. Diese wurden dann mit einer Virusmenge versetzt, die eine Multiplizität von mindestens eins gewährleistete ($MOI \geq 1$). Dadurch wird eine möglichst gleichzeitige Infektion aller Zellen gestattet. Nach 68–72 Stunden Inkubation bei 27 °C und 125 rpm wurden die Zellen gezählt, die Viabilität überprüft und eine erfolgreiche Infektion anhand des Zelldurchmessers sichergestellt. Die Zellen wurden für 10 Minuten bei 250 xg zentrifugiert und das Pellet in PBS gewaschen. Die pelletierten Zellen können dann bis zur weiteren Verwendung bei -80 °C gelagert werden.

6.3 Methoden der Nukleinsäure-Biochemie

Standardmethoden wie Nukleinsäure Präzipitation durch Ethanol- oder Isopropanol und Phenol-/Chloroform-Extraktion wurden nach Protokollen aus (Sambrook und Russell, 2001) ausgeführt. Plasmid Präparationen aus *E. coli* erfolgten durch Kits von Peqlab und Qiagen. Die Konzentrationsbestimmung von Nukleinsäuren erfolgte durch photometrische Bestimmung der Absorption bei 260 nm Wellenlänge. Dabei wurde für DNA bei einer Absorption von Eins als eine Konzentration von 50 ng/µl angenommen und für RNA eine Konzentration von 40 ng/µl.

6.3.1 Radioaktive 5'-Markierung von Nukleinsäuren mittels T4-Polynukleotidkinase

Zur Markierung von RNA-Oligonukleotiden und pBR322-DNA-Fragmenten (*Hpa*II-Verdau, Marker) wurde die Bakteriophagen T4 5'-Polynukleotidkinase (T4-PNK, NEB) verwendet. Einige sehr kurze RNAs (<12 nt) sowie die G-reichen RNAs G₁₄ und G₁₇ wurden aufgrund der besseren Ausbeute mit CF II radioaktiv markiert. Radioaktive Markierungen von RNAs mit der T4-PNK und CF II wurden in der Regel in Reaktionen von 20 µl durchgeführt:

T4-PNK Reaktion	CF II-PNK Reaktion
1x T4 PNK-Puffer	30 mM HEPES pH 8,0
DNA (0,5–1 µg) oder RNA (0,1–10 pmol)	100 mM Kaliumchlorid
8 U RNasin	3 mM Magnesiumacetat
20 U T4 PNK	1 mM DTT
20 µCi $\gamma^{32}\text{P}$ -ATP	10% (w/v) Saccharose
H ₂ O <i>ad.</i> 20 µl	0,02% NP-40
	5 µM Zinkchlorid
	RNA (0,1–10 pmol)
	0,4 U RNasin
	0,8 µM CF II
	20 µCi $\gamma^{32}\text{P}$ -ATP
	H ₂ O <i>ad.</i> 20 µl

Die Reaktion wurde für mindestens eine Stunde bei 37 °C inkubiert, anschließend mit Wasser auf 200 µl aufgefüllt und einer Ethanol-fällung unterzogen. Bei Nukleinsäuren unter 15 nt Länge wurde dabei 4 Volumen abs. Ethanol statt der Standardmenge von 2,5 Volumen verwendet, um eine vollständige Präzipitation zu gewährleisten. Das Pellet wurde abschließend in 20 µl HEPES-Lösung (10 mM, pH 7,0; RNase-frei) aufgenommen.

Der radioaktive pBR322-Marker wurde in Elektrophoresen eingesetzt und die Menge an Radioaktivität in den Marker-Spuren wurde in etwa der Radioaktivität der zu analysierenden Proben angepasst. Die Konzentrationen der radioaktiv markierten 5'-Enden der RNAs wurde durch DEAE-Methode und Szintillationsmessung bestimmt.

6.3.2 Affinitätsreinigung von Nukleinsäuren mittels DEAE-Methode

Der Transfer von ^{32}P auf 5'-Hydroxylgruppen von Nukleinsäuren nach 5'-PNK-Reaktionen wurde durch die DEAE-Methode, gekoppelt mit Szintillationsmessung, kontrolliert. Dazu wurde im Wesentlichen die Methode aus (Stayton und Kornberg, 1983) modifiziert und verwendet. DE81-Anionenaustausch-Papier wurde mit 10 mM tetra-Natriumdiphosphat vorbehandelt, mit abs. Ethanol gewaschen und getrocknet. Nach der PNK-Markierung wurde ein Aliquot der Reaktion auf vorbehandeltes DE81-Papier getropft und getrocknet. Im Anschluss erfolgten mehrere Waschschrte mit DEAE-Waschpuffer, in denen freies $\gamma^{32}\text{P}$ -ATP und Phosphatreste ausgewaschen wurden. Auf die gewaschenen DE81-Papiere wurden abschließend 2 ml Szintillationscocktail (Perkin Elmer) gegeben und die Radioaktivität der markierten RNAs quantifiziert.

DEAE-Waschpuffer

0,3 M Ammoniumformiat, pH 7.8

10 mM tetra-Natriumdiphosphat

0,5% (w/v) NP-40

0,1% (w/v) SDS

6.3.3 Synthese von (radioaktiv markierten) RNAs durch *in vitro* Transkription

Für *in vitro* 3'-Spaltassays wurden intern radioaktiv markierte Substrat-RNAs generiert. Diese wurden mittels SP6- oder T7-RNA-Polymerase (NEB) in *In-Vitro*-Transkription mittels *run-off* Transkription von linearisierten Plasmid-DNAs hergestellt. Die RNA-Substrate (L3, SV40 *late* und SV40 *late* Fragmente >39 nt) wurden in 40 μl Reaktionen synthetisiert. Eine 40 μl Transkriptionsreaktion erfolgte nach dem folgenden Schema:

SP6-Transkription

1x T7-Puffer

1x NTP-Mix

40 μCi α - ^{32}P -UTP

8 mM 3'-O-Me-m⁷G(5')ppp(5')G Cap-Analogon

2,5 mM DTT

0,01% (v/v) NP-40

2 μg DNA (linearisiertes Plasmid)

40 U RNasin

50 U SP6-Polymerase

H₂O ad. 40 μl

5x NTP-Mix

2,5 mM ATP

2,5 mM CTP

0,5 mM GTP

0,5 mM UTP

Gelelutionspuffer

500 mM Ammoniumacetat

10 mM Magnesiumchlorid

0,1 mM EDTA

0,1% (w/v) SDS

T7-Transkription

1x T7-Puffer
1x NTP-Mix
40 μCi α - ^{32}P -UTP
8 mM 3'-O-Me-m⁷G(5')ppp(5')G Cap-Analogon
2,5 mM DTT
2 μg DNA (linearisiertes Plasmid)
40 U RNasin
50 U T7-Polymerase
H₂O ad. 40 μl

Die Reaktionen wurden für drei bis fünf Stunden, für kurze Transkripte über Nacht, bei 37°C inkubiert. Anschließend wurde die Reaktion mit Wasser auf 100 μl aufgefüllt, die RNA mittels Ethanol-fällung präzipitiert und das RNA-Pellet in 10–20 μl HEPES-Lösung [10 mM, pH 7,0] aufgenommen. Diese wurde dann mit Formamid-Ladepuffer versetzt und die Transkriptionsprodukte anschließend mittels denaturierender Harnstoff-Gelelektrophorese aufgetrennt. Die RNAs wurden mittels Autoradiographie detektiert und Produkte entsprechender Länge wurden aus dem Gel ausgeschnitten, mit 400 μl Gelelutionspuffer versetzt und über Nacht bei Raumtemperatur inkubiert. Die eluierte RNA wurde mit Phenol-/Chloroform-Extraktion gereinigt, einer Ethanol-fällung unterzogen und in 50–100 μl HEPES-Lösung [10 mM, pH 7,0] wieder aufgenommen. Durch Bestimmung der spezifische Radioaktivität des eingesetzten UTPs (Zerfälle/(min*fmol UTP)) konnte die Ausbeute und Konzentration der statistisch eingebauten UMPs und daraus die Konzentration der gereinigten RNA berechnet werden.

6.3.4 Zirkulardichroismus (CD)-Messungen von Ribonukleinsäuren

Die RNA-CD-Messungen erfolgten bei Raumtemperatur (20 °C) in CD-Puffer. Die Spektren wurden in einem Jasco J-815 CD-Spektrometer aufgenommen, welches eine temperierbare Küvettenkammer besaß. Als Küvette diente eine 100 μl Quarzglasküvette (Hellman) mit einer Pfadlänge von 1 cm. Die Auswertung und Darstellung der Daten erfolgte in der zugehörigen Software Spectra Manager II (Jasco).

2x CD-Puffer

100 mM Tris pH 8,0
200 mM Kaliumchlorid oder Lithiumchlorid
20% Glycerin
0,02% NP-40

Vor der Verdünnung für die Messungen wurden die RNAs in 10 mM HEPES-Puffer pH 8,0 für fünf Minuten bei 95 °C denaturiert und dann auf Eis abgekühlt. Die RNAs wurden dann in 2x CD-Puffer auf eine Konzentration von 2 µM verdünnt und sofort im Anschluss das CD-Spektrum zwischen 320 nm und 230 nm aufgenommen. Zur Kontrolle auf eine stabile Konformation der RNA wurden mehrere Messungen bei steigender Inkubationszeit durchgeführt und auf Abweichungen überprüft. Als Referenzen in der CD-Messung wurden das G₄-bildende rG₁₄ und das unstrukturierte rU₁₄ RNA-Oligonukleotid in einer Konzentration von 2 µM verwendet.

6.3.5 Agarose-Gelelektrophorese

Kontrollen, PCR-Fragmente und Restriktionsanalysen von Plasmid-DNAs wurden mittels horizontaler Agarose-Gelelektrophorese analysiert. Je nach Größe der erwarteten DNA-Fragmente wurden Agarosegele mit 0,8 bis 1,5% (w/v) Agarose-Anteil in 0,5x TBE verwendet. Für native Kontrollen von RNA-Fragmenten und -Produkten wurde 1x TBE verwendet. Die Elektrophorese erfolgte bei 15–20 V/cm Gel.

0,5x TBE	1x TBE
45 mM Tris-Base	90 mM Tris-Base
45 mM Borsäure	90 mM Borsäure
0,5 mM EDTA	1 mM EDTA

6.3.6 Denaturierende Harnstoff-Polyacrylamid-Gelelektrophorese

Die gelelektrophoretische Auftrennung von (radioaktiv markierten) RNAs wurde durch denaturierende Harnstoff-PAGE realisiert. Es wurden 200x200x0,4 mm-Gele (bis 30 Proben) und 350x450x0,4 mm-Gele (bis 56 Proben) verwendet. Die benötigten Gellösungen wurden jeweils mit frisch eingewogenem Harnstoff hergestellt und die Acrylamid- und die TBE-Konzentration wurden durch Verdünnung einer 40% (w/v) Acrylamid-Stammlösung und einer 10x TBE-Stammlösung eingestellt.

Gellösung

8,3 M Harnstoff
 1x TBE
 5 bis 20 % (w/v) Acrylamid:Bisacrylamid (19:1)

Formamid-Ladepuffer

80 % (v/v) deionisiertes Formamid
 10 mM EDTA
 0,25% (w/v) Xylencyanol
 0,25% (w/v) Bromphenolblau
 1x TBE

Harnstoff-Ladepuffer

8,3 M Harnstoff

Für die Polymerisation von 100 ml Gellösung wurden 500 µl APS [10% (w/v)] und 100 µl TEMED hinzugegeben, kurz gemischt und das Gel zügig und stetig zwischen die vorbereiteten Glasplatten gegossen. Um eine vollständige Polymerisation des Gels zu erreichen, wurde die Polymerisation etwa eine Stunde lang erlaubt, bevor das Gel weiterverwendet wurde. Die Gelelektrophorese unter denaturierenden Bedingungen wurde weiterhin erreicht, indem das Harnstoff-PA-Gel bei der Elektrophorese eine Betriebstemperatur von mindestens 40 °C besaß. Deshalb wurden die Gele vor dem Laden der Proben einem mindestens 30-minütigen Vorlauf unterzogen. Vorlauf und Elektrophorese von 200x200x0,4 mm-Gelen wurden bei konstanten 20 W durchgeführt. Bei Gelen von 350x450x0,4 mm wurde eine Leistung von konstanten 50 W angelegt. Bei der Auftrennung kurzer Nukleinsäuren unter ca. 10 nt Länge führen Salze und Farbstoffe wie Xylencyanol und Bromphenolblau oft zu verzerrenden Effekten im Migrationsverhalten der Nukleinsäuren. Für die Analyse dieser kurzen RNAs wurde ein Harnstoff-Ladepuffer verwendet.

6.4 Proteinbiochemische Methoden

6.4.1 Konzentrationsbestimmung von Proteinen

6.4.1.1 Photometrische Bestimmung

Liegt eine reine Proteinlösung vor, kann die Konzentration photometrisch durch Absorptionsmessung bei einer Wellenlänge von 280 nm bestimmt werden. In Proteingemischen kann, ähnlich wie für Nukleinsäuren, ein statistischer Durchschnittswert für die Absorption und Häufigkeit der Aminosäuren Phenylalanin, Tryptophan und Tyrosin von 1 g/l bei einer $A_{280\text{nm}}=1$ angenommen werden. Für isolierte Proteine in Lösung kann ein spezifischer Extinktionskoeffizient ϵ anhand der Aminosäuresequenz bestimmt werden. Dieser kann dann durch Einsetzen in das Lambert-Beersche Gesetz ($A = \epsilon * c * d$) zur Bestimmung der Konzentration c genutzt werden.

6.4.1.2 Konzentrationsbestimmung nach Bradford

Die Konzentration von Proteinen in komplexen Lösungen, z.B. in Verbindung mit Nukleinsäuren, kann nicht photometrisch bestimmt werden. Zur Bestimmung der Gesamtproteinkonzentration wurde deshalb die Proteinbestimmung nach Bradford verwendet. Dies erfolgte in 96-Well Mikrotiter-Platten nach den Vorgaben des Herstellers der RotiQuant 5x Bradford-Lösung (Roth). Die 5x Lösung wurde im Verhältnis 2:7,5 mit Wasser verdünnt. Die Proteinprobe wurde in entsprechender Verdünnung in einem Volumen von 50 µl vorgelegt und mit 200 µl Bradford-Reagenz gemischt. Nach 5–10 Minuten Inkubation erfolgte die Messung der Absorption bei 595 nm im MultiMode-reader (Berthold). Als Standardprotein zur Erstellung einer Referenz-Eichgeraden wurde BSA-Fraktion IV mit bekannter Konzentration verwendet.

6.4.1.3 Densitometrische Konzentrationsbestimmung aus Coomassie-gefärbten SDS-Polyacrylamid-Gelen

Die Bestimmung der Konzentration einzelner Proteine in Proteingemischen erfolgte über die densitometrische Quantifizierung Coomassie-gefärbter Proteine in SDS-Polyacrylamid-Gelen mittels der ImageQuant Software. Als Referenz wurde wieder eine Eichgerade über parallel, auf demselben Gel aufgetragene, BSA-Spuren bekannter Konzentration erstellt. Diese Methode enthält eine relative Unsicherheit durch die unterschiedliche Anfärbbarkeit von Proteinen, durch den verwendeten Farbstoff Coomassie-Brilliant-Blau R250.

6.4.2 Glycin-SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE)

Die denaturierende elektrophoretische Auftrennung von Proteinen wurde durch das diskontinuierliche Gelsystem nach Laemmli (LAEMMLI, 1970) realisiert.

Sammelgel:

3,6% (w/v) Acrylamid:Bis. (37,5:1)
62,5 mM Tris-HCl, pH 6,8
0,05% (w/v) SDS

Trenngel:

6 bis 15% (w/v) Acrylamid:Bis. (37,5:1)
375 mM Tris-HCl, pH 8,8
0,1% (w/v) SDS

SDS-Probenpuffer [2x]:

10% (v/v) Glycerin
62,5 mM Tris-HCl, pH 6,8
2% (w/v) SDS
715 mM 2-Mercaptoethanol
0,125% (w/v) Bromphenolblau

Laemmli-Laufpuffer [1x]:

50 mM Tris-Base
380 mM Glycin
0,1% (w/v) SDS

Für die Elektrophorese wurden standardmäßig 100x200x1 mm Gele verwendet. Diese wurden aus 20 ml Trenngel-Lösung der gewünschten PAA-Konzentration gegossen. Die Polymerisation wurde durch Zugabe von 150 µl APS (10% w/v) und 60 µl TEMED gestartet und bis zum Abschluss mit Isopropanol überschichtet. Nach der Polymerisation des Trenngels wurde aus 10 ml Sammelgel-Lösung die Polymerisation mit 80 µl APS und 40 µl TEMED gestartet und über das Trenngel gegossen. Nach Abschluss der Polymerisation wurde das Gel mit 1x Laemmli-Laufpuffer verwendet und bei 15–20 V/cm (150–200 V) die Elektrophorese durchgeführt. Die Proben wurden vor dem Laden mit 2x Probenpuffer versetzt und fünf Minuten bei 95 °C denaturiert.

6.4.3 Tricin-SDS-Polyacrylamid-Gelelektrophorese (Tricin-SDS-PAGE)

Die denaturierende elektrophoretische Auftrennung von Proteinen nach der limitierten Proteolyse von WDR33 (Abschnitt 6.5.3) erfolgte mittels Tricin-SDS-PAGE (Schägger, 2006). Der Vorteil dieser Methode gegenüber der standardmäßig genutzten Glycin-SDS-PAGE liegt in der verbesserten Auftrennung von Proteinen kleiner als 30 kDa. In der Page wurden folgende Puffer und Lösungen verwendet:

Sammelgel:	Trenngel:
4% (w/v) Acrylamid:Bis. (32:1)	10% (w/v) Acrylamid:Bis. (32:1)
1x Gelpuffer	1x Gelpuffer
	1% (w/v) Glycerin

Gelpuffer [3x]:	Kathodenpuffer	Anodenpuffer
1 M Tris	100 mM Tris	100 mM Tris
330 mM HCl	100 mM Tricin	22,5 mM HCl
0,1% (w/v) SDS	0,1% SDS	pH 8,9
pH 8,45	pH 8,25	

Für die Elektrophorese wurden standardmäßig 100x200x1 mm Gele verwendet. Diese wurden aus 20 ml Trenngel-Lösung (10% PA) gegossen. Die Polymerisation wurde durch Zugabe von 120 µl APS (10% w/v) und 40 µl TEMED gestartet und bis zum Abschluss mit Isopropanol überschichtet. Nach der Polymerisation des Trenngels wurde aus 10 ml Sammelgel-Lösung die Polymerisation mit 40 µl APS und 20 µl TEMED gestartet und über das Trenngel gegossen. Nach Abschluss der Polymerisation wurde das Gel mit Kathoden- und Anoden-Laufpuffer verwendet. Die Elektrophorese wurde bei und bei 3 V/cm (30 V) gestartet und als die Proben in das Gel migriert waren auf eine Gesamtleistung von 10 W erhöht. Die Proben wurden vor dem Laden mit

2x Probenpuffer versetzt und fünf Minuten bei 95 °C denaturiert. Nach vier Stunden wurde die Elektrophorese beendet und das Gel im Anschluss für zwei Stunden bei 60 °C unter Vakuum getrocknet.

6.4.4 Kolloidale Färbung von SDS-PA-Gelen mittels Coomassie-Brilliant-Blau

Proteinbanden in SDS-Polyacrylamid-Gelen wurden mit Coomassie-Brilliant-Blau gefärbt. Bei dieser Methode lagert sich der blaue Coomassie-Farbstoff an basische Aminosäureseitenketten an und färbt diese somit an. Zur Färbung wurde das Gel nach der Elektrophorese für 15 Minuten in Fixierlösung gelegt und anschließend für etwa zwei Stunden mit Färbelösung gefärbt. Abschließend wurde das Gel mit Entfärber so lange entfärbt, bis das gewünschte Verhältnis von Bandenintensität zu Hintergrundfärbung erreicht war.

Fixierlösung	Färbelösung	Entfärber
25% (v/v) Isopropanol	10% (v/v) Essigsäure	10% (v/v) Essigsäure
10% (v/v) Essigsäure	0,01% (w/v) Coomassie-Brilliant-Blau R-250	

6.4.5 Silberfärbung von SDS-Polyacrylamid-Gelen

Als sensitive Methode zur Anfärbung von Proteinen wurde die Silberfärbung verwendet. Dabei werden Proteinmengen bis zu einer Nachweisgrenze von 10 ng durch Silberionen angefärbt, indem diese an negativ geladene Aminosäureseitenketten anlagern und dort zu elementarem Silber reduziert werden.

Fixierlösung	Sensibilisierungslösung
50% (v/v) Ethanol	0,02% (w/v) Natriumthiosulfat
12% (v/v) Essigsäure	
0,05% (v/v) Formaldehyd	
6% (w/v) Natriumcarbonat	
Färbelösung	Entwicklerlösung
0,2% (w/v) Silbernitrat	0,0004% (w/v) Natriumthiosulfat
0,076% (v/v) Formaldehyd	0,05% (v/v) Formaldehyd

Nach der Elektrophorese wurde das Gel (100x200x1 mm) zunächst in 100 ml Fixierlösung gegeben und für mindestens eine Stunde (oder über Nacht) schüttelnd inkubiert. Danach wurde

das Gel mit 250 ml 20%iger Ethanol Lösung (v/v) dreimal für acht Minuten gewaschen. Das Gel wurde dann für zwei Minuten mit 100 ml Sensibilisierungslösung inkubiert und anschließend zweimal je eine Minute in Wasser gewaschen. Zur Färbung wurde das Gel für 20 Minuten in 100 ml Färbelösung inkubiert, dann zweimal kurz in Wasser geschwenkt und abschließend für die Entwicklung mit 100 ml Entwicklerlösung versetzt. Die Entwicklung wurde mit 100 ml 12%iger Essigsäurelösung (v/v) gestoppt.

6.4.6 Western-Blot-Analyse von Zielproteinen

Für die Analyse von Zielproteinen in Proteingemischen wurde die Western-Blot-Analyse (Burnette, 1981) angewendet. Proteine wurden zuerst in der SDS-PAGE aufgetrennt und anschließend im *Semi-Dry*-Verfahren auf eine Nitrozellulosemembran übertragen.

1x Transferpuffer	TN-Puffer	Membrane-Strip-Puffer
50 mM Tris-Base	20 mM Tris-HCl, pH 7,5	0,2 M NaOH pH 9–10
40 mM Glycin	150 mM Natriumchlorid	
0,37% (w/v) SDS		
20% (v/v) Methanol		
Ponceau S-Lösung	TN-Tween	Blocking-Reagenz
0,5% (w/v) Ponceau S	20 mM Tris-HCl, pH 7,5	2,5% (w/v) Magermilchpulver
0,05% (v/v) Tween-20	150 mM Natriumchlorid	in TN-Puffer
in TN-Puffer		
1% (v/v) Essigsäure		

Die Verwendung der Blot-Apparatur erfolgte nach Angaben des Herstellers. Der *Semi-Dry*-Transfer wurde für 90 Minuten bei konstanten 3 mA/cm² Gel-Fläche durchgeführt. Nach der Elektrophorese wurde die Membran mit Ponceau S-Lösung reversibel gefärbt, um Markerbanden und Laufspuren markieren zu können und Proteinmengen auf dem Blot abzuschätzen. Zur Sättigung überschüssiger Bindungskapazitäten der Membran wurde diese für eine Stunde in Blocking-Reagenz inkubiert. Primäre Antikörper wurden in geeigneter Verdünnung (in TN-Tween mit 0,5% (w/v) Magermilchpulver) mit der geblockten Membran für mindestens zwei Stunden bei Raumtemperatur inkubiert. Nach drei- bis viermaligem Waschen der Membran mit TN-Tween erfolgte eine halbstündige Inkubation mit sekundärem Antikörper in Dunkelheit (Fluoreszenzmarkierte Antikörper in TN-Tween, Li-Cor). Die Detektion des Signals erfolgte nach

Waschen mit TN-Tween und TN-Puffer nach Herstellerangaben auf dem Li-Cor Fluoreszenzscanner.

Nach der Detektion wurden Nitrozellulosemembranen mit Strip-Puffer behandelt (einmal fünf Minuten Wasser, 10–20 Minuten in Membrane-Strip-Puffer bei Raumtemperatur, einmal fünf Minuten TN-Puffer), wenn mit mehreren primären und sekundären Antikörpern geprobt wurde.

6.4.7 Methoden der Proteinfractionierung und -Reinigung

Die im Folgenden beschriebenen Methoden gelten für die in der Arbeit durchgeführten Reinigungen von rekombinanten Proteinen. Abweichungen von den hier beschriebenen Standardverfahren in Ablauf oder Pufferzusammensetzung werden in den jeweiligen Ergebnissen genannt. Detaillierte Informationen zu Puffern, Gradienten und Ausbeuten finden sich im Anhang.

6.4.7.1 Aufschluss von Insektenzellen nach Proteinexpression

Zell-Pellets aus baculoviralen Expressionen in SF21-Insektenzellen wurden stets zu einer Dichte von $2-4 \times 10^6$ Zellen/ml in geeignetem Lysepuffer resuspendiert. Der Aufschluss der Zellen erfolgte mittels Ultraschall-Sonde (5x 20 Sekunden, 50% Betriebszeit, 40% Intensität). Im Allgemeinen wurde für die Proteinreinigung der folgende Puffer verwendet. Der Kaliumchlorid-Salzgehalt und die Puffer-Salze wurden gegebenenfalls an die Methode angepasst.

Protein-Lysepuffer

50–300 mM Kaliumchlorid
 20–50 mM Tris oder HEPES pH 7,5 (bei RT)
 0–1 mM EDTA
 1 mM DTT
 10% (w/v) Saccharose oder (v/v) Glycerin
 1 mM PMSF
 1 mg/l Pepstatin und Leupeptin

Nach dem Aufschluss wurden unlösliche Bestandteile in einem Zentrifugationsschritt bei 20.000 xg für 30 Minuten bei 4 °C abgetrennt. Der lösliche Überstand nach der Zentrifugation wurde dann für die weitere Reinigung der enthaltenen Proteine verwendet.

6.4.7.2 Ni-NTA-Agarose Affinitätsreinigung

Rekombinante Proteine mit His-Fusionspeptiden wurden über die Affinität von Histidin zu immobilisierten Nickel-Metallionen gereinigt (IMAC). Dazu wurde Ni-NTA-Agarose (Qiagen, [4]) als Affinitätsmatrix genutzt. Die Reinigung folgte dem vorgeschlagenen Protokoll des Herstellers für native Reinigung von Fusionsproteinen. Für Ni-Affinitätsreinigungen wurde dem Lysepuffer 10 mM Imidazol zugesetzt, um unspezifische Bindung an die Ni-NTA-Matrix zu minimieren. Eine geeignete Menge Affinitätsmatrix wurde in dem Lysepuffer äquilibriert und dem löslichen Überstand des Zellaufschlusses hinzugegeben. Die Bindung an die Matrix erfolgte für zwei bis drei Stunden unter ständiger Bewegung. Das Säulenmaterial wurde anschließend in einer Zentrifugation (fünf Minuten, 250 xg, 4 °C) geerntet. Der Überstand als Durchfluss gesammelt und die Matrix in einem kleinen Rest des Durchflusses resuspendiert. Das Säulenmaterial wurde in eine Leer-Säule überführt. Im nächsten Schritt wurde mit 5 mal 10 Säulenvolumen ungebundenes Protein ausgewaschen. In den Waschsritten wurde die Imidazolkonzentration auf 20 mM erhöht. Für die native Elution des Fusionsproteins von der Matrix wurde dem Lysepuffer 500 mM Imidazol zugesetzt (pH Kontrolle zu 8,0). Die Elution des Zielproteins erfolgte dann in Schritten von 0,5–1 Säulenvolumen mit 10-minütiger Inkubation zwischen den Elutionsschritten.

6.4.7.3 Strep-Affinitätschromatographie

Die Reinigung von Proteinen mit Strep-Fusionspeptiden erfolgte über eine Strep-Tactin-Sepharose-Matrix (IBA, [5]). Dabei wird die nicht-kovalente und hoch affine Bindung des Strep-Peptids an das Säulenmaterial genutzt. Vor Applikation des Zell-Lysats wurde ein geeignetes Volumen an Strep-Tactin-Sepharose entsprechend der Herstellerangaben in Lysepuffer äquilibriert (IBA). Zellaufschlusses und Bindung an die Affinitätsmatrix erfolgten im Lysepuffer für zwei Stunden bei 8 °C unter stetiger Bewegung. Die Matrix wurde für fünf Minuten bei 250 xg und 4 °C geerntet. Der Überstand wurde als Durchfluss gesammelt. Ein Rest wurde genutzt, um die Matrix zu resuspendieren und die Suspension wurde in eine Leer-Säule überführt. Nach dem vollständigen Beladen der Säule folgten fünf Waschsritte mit je einem Säulenvolumen mit Lysepuffer. Die Elution erfolgte unter nativen Bedingungen durch Competition der Strep-Bindung mit 5 mM D-Desthiobiotin im Elutionspuffer, in sechs Schritten, mit jeweils 0,5 Säulenvolumen. Zwischen den Elutionsschritten erfolgte jeweils eine 10-minütige Inkubation. Die Strep-Tactin-Sepharose wurde nach der Chromatographie laut Herstellerangaben regeneriert.

6.4.7.4 Antikörper-Affinitätsreinigung

Die Affinitätsreinigung von Proteinen mit Antikörpern erfolgte über eine Protein A Sepharose CL-4B-Matrix (GE Healthcare, [14]) und richtete sich im Wesentlichen nach Herstellerangaben. Dabei wird die nicht-kovalente und hoch affine Bindung des Protein A an die schwere Kette der genutzten primären Antikörper genutzt, welche wiederum das gesuchte Antigen binden. Somit kann für verschiedene Affinitätsreinigungen die gleiche Matrix verwendet werden. Vor Applikation des jeweiligen Proteingemisches wurden 5 mg pro Reaktion an Protein A Sepharose entsprechend der Herstellerangaben gewaschen und äquilibriert. Danach erfolgte die Kopplung des jeweiligen primären Antikörpers (15 µl Serum oder Prä-Immunsereum) an die Protein A Sepharose für eine Stunde im Kühlraum bei 8 °C. Die Matrix mit dem Antikörper wurde dann in PBS gewaschen und im jeweiligen Bindungspuffer äquilibriert.

CPSF4+1	hPcf11
50 mM Tris/HCl pH 7,4	25 mM Tris-HCl, pH 8,5
150 mM Natriumchlorid	200 mM Kaliumchlorid
0,05 % NP-40	10% Saccharose
2 M Harnstoff	0,02% NP-40
	0,2 mM EDTA
	0,5 mM DTT

Nach der UV-Vernetzung wurden CPSF4 und -1 denaturiert, indem sie eins zu eins mit Bindungspuffer +8 M Harnstoff gemischt wurden. Die Lösung wurde dann durch Zugabe von kristallinem Harnstoff auf 8 M eingestellt und für 30 Minuten bei Raumtemperatur unter Bewegung inkubiert. Vor der Bindung des CPSF4 oder -1 an die Antikörper-Affinitätsmatrix erfolgte die Verdünnung mit Bindungspuffer (ohne Harnstoff) auf den 2 M Harnstoff-Gehalt. Die Bindung erfolgte im 2 M Harnstoff-Puffer für eine Stunde bei Raumtemperatur unter stetiger Bewegung. Die Matrix wurde für fünf Minuten bei 250 xg präzipitiert und dann dreimal mit dem Bindungspuffer gewaschen. Die Elution erfolgte mit 4xSDS-Puffer. Die Analyse erfolgte im Anschluss über SDS-PAGE und Autoradiographie. Die Bindungsreaktion mit hPcf11 erfolgte im Bindungspuffer, für eine Stunde bei 8 °C und ständiger Bewegung. Im Anschluss wurde auch hier die Matrix präzipitiert und dreimal im Bindungspuffer gewaschen. Die Elution und anschließende Analyse erfolgten analog zu CPSF4 und -1.

6.4.7.5 Anionenaustauschchromatographie

In dieser Arbeit wurden zwei verschiedene Anionenaustauscher-Säulen genutzt: eine 1 ml ResourceQ-Säule und eine 1 ml MonoQ-Säule. Hier ist an das Trägermaterial eine funktionale Gruppe verknüpft, welche über ein tertiäres Amin eine starke positive Ladung besitzt. Somit binden negativ geladene Ionen und Aminosäureseitenketten an das Material. Bei niedriger Salzkonzentration bindet das Protein an die Säulenmatrix und kann anschließend mit einem geeigneten Gradienten steigender Ionenstärke eluiert werden. Für die korrekte Applikation des Gradienten wurde bei den Reinigungen stets das Chromatographiesystem der ÄKTA *purifier* verwendet und mit der entsprechenden Software UNICORN gesteuert. Die Fraktionierung der Gradienten erfolgte automatisch durch einen angeschlossenen Fraktionssammler in 0,5 ml Schritten. Eine geringe Flussrate und ein langsam ansteigender Gradient sind bei dieser Methode von entscheidender Bedeutung für eine effiziente Chromatographie. In der Regel erfolgte die Trennung bei einem Gradienten von 80 mM bis 1 M Kaliumchlorid. In manchen Fällen wurden Säulen auch mit höherer Salzkonzentration beladen (150 bis 200 mM Kaliumchlorid) falls die Elutionsprofile der Zielproteine dies zuließen. Die Reinigung über einen Anionenaustauscher ermöglichte gleichzeitig die Entfernung von Nukleinsäuren aus einer Proteinpräparation, da diese erst bei hohen Salzkonzentrationen eluieren.

6.4.7.6 Größenausschlusschromatografie

In der Größenausschlusschromatographie (oder Gelfiltration) wird das Proteingemisch gemäß der Molekülvolumen und der Durchmesser aufgetrennt und so nach dem nativen Molekulargewicht getrennt. Daraus können Informationen zum Oligomerisierungs- oder Aggregationsstatus und dem Stoke's Radius von Proteinen gewonnen werden. Große Proteine und Proteinkomplexe migrieren schneller durch das Gel, während kleine Proteine vermehrt in die Poren der Gelmatrix eindringen und dadurch länger für die Migration über die Säule brauchen. Durch Kalibrierung einer Gelfiltrationssäule mit Proteinen bekannter Molekulargewichte kann eine Korrelation zwischen Elutionsvolumen (K_{AV}) und Logarithmus des Molekulargewichts eines Proteins hergestellt werden.

In dieser Arbeit wurden verschiedene vorgepackte Säulen ([6], [7]) zur Gelfiltration sowie auch eine selbstgepackte Superose6-Säule (100 ml, [8]) verwendet. Chromatographien wurden mit dem Chromatographiesystem ÄKTA *purifier* und der entsprechenden Software UNICORN gesteuert. Die verwendeten Puffer orientierten sich am Standard-Lysepuffer mit Anpassungen an das zu reinigende Material. Die Chromatographien und die Kalibrierung der Säulen entsprachen dabei den Angaben der Hersteller der Säulen, des Säulenmaterials und der Kalibrier-Kits

(Pharmacia, GE Healthcare, [9]). Die Fraktionierung des Eluats erfolgte automatisch durch einen angeschlossenen Fraktionssammler.

6.4.7.7 Rechtwinkel-Lichtstreuung

Das native Molekulargewicht der 3'-Prozessierungsfaktoren CstF und CF I wurden aufgrund der großen apparenten Molekulargewichte in der 24 ml Superose6-Säule zusätzlich über die Analyse mittels Rechtwinkel-Lichtstreuung bestimmt. Dafür wurde die Gelfiltrationssäule mit einem Viscotek TDA 305 Detektor gekoppelt. Die Analyse erfolgte bei Raumtemperatur und in RWLS-Puffer. Es wurden 100 µl CstF-Komplex (ResourceQ #12) einer Konzentration von 5,6 g/l und CF I-Komplex (Ni-NTA E2) einer Konzentration von 2 g/l auf die Säule geladen. Als Referenz wurde eine Kalibrierung mit 100 µl BSA einer Konzentration von 3 g/l durchgeführt.

RWLS-Puffer

20 mM Tris pH 8

200 mM Kaliumchlorid

5% (w/v) Saccharose

Die Gelfiltration erfolgte wie beschrieben und wurde in dem Detektor durch Messung von Brechungsindex, Rechtwinkel-Lichtstreuung und UV-Absorption verfolgt. Über die bekannten Proteinkonzentrationen kann dann über folgende Formeln:

$$\text{Refraktometer Signal} = K_{RI} * \frac{dn}{dc} * \text{Konzentration}$$

$$\text{UV Signal} = K_{UV} * \frac{dA}{dc} * \text{Konzentration}$$

... der Faktor $\left(\frac{dn}{dc}\right)$ bestimmt werden. Nimmt man nun diese beiden Faktoren, kann man aus der Formel:

$$\text{RWLS Signal} = K_{LS} * \text{Molekulargewicht} * \left(\frac{dn}{dc}\right)^2 * \text{Konzentration}$$

... das Molekulargewicht bestimmen. Die Kalibrierung nach dem BSA-Standard, Ermittlung der Parameter und Auswertung der Daten erfolgte in der zugehörigen Software OMNISEC (Malvern Panalytica).

6.4.7.8 Dichtegradientenzentrifugation

In dieser Arbeit wurden analytische Glyceringradienten angefertigt. Diese hatten jeweils ein Gradientenvolumen von 12 ml und ein Konzentrationsgefälle von 20–40% (v/v) Glycerin in Gradienten-Puffer.

Gradienten-Puffer

250 mM Kaliumchlorid

5 μ M Zinkchlorid

25 mM Dikaliumhydrogenphosphat pH 8,0

10% (w/v) Glycerin

1 mM DTT

Für den analytischen Versuch wurden 200 μ l CF II (MonoQ #19) dreimal gegen 200 ml Gradienten-Puffer bei 8 °C dialysiert. Die dialysierte Fraktion wurde dann 30 Minuten bei 20.000 xg und 4 °C zentrifugiert. Der Überstand wurde komplett auf den Gradienten überschichtet und somit geladen. Die Zentrifugation erfolgte in einem SW40-Ti Rotor (Beckmann) bei 275.000 xg für 48 Stunden bei 4°C. Um eine Störung der Gradienten zu verhindern, wurde zunächst mit minimaler Beschleunigung eine Geschwindigkeit von 3000 xg eingestellt (Einstellung Accel-4) und dann langsam bis 275.000 xg beschleunigt. Am Ende der Zentrifugation wurde langsam aktiv gebremst (Einstellung Decel-4). Die Gradienten wurden mit einer Glaskanüle vom Boden des Röhrchens her abgesaugt und in 36 Fraktionen absteigender Dichte fraktioniert (~350 μ l je Fraktion). Während der Fraktionierung wurde das UV-Profil (280 nm) des Gradienten aufgezeichnet, um die Position der Proteine im Gradienten zu lokalisieren. Zur Digitalisierung wurde ein 12 bit USB-A/D-Wandler an den analogen Schreiber (Pharmacia) adaptiert, der neben der UV-Absorption auch den Wechsel zwischen Fraktionen aufzeichnete. Die zur Digitalisierung nötige Software (GradientScan) wurde in Java durch M. Götze programmiert. Zur Bestimmung des Sedimentationskoeffizienten wurden parallel zur CF II-Präparation Gradienten mit Kalibrierproteinen (Catalase, Alkoholdehydrogenase – ADH, *O.v.* Albumin, Chymotrypsinogen A) angefertigt, von denen das Sedimentationsverhalten bekannt war. So konnte der Sedimentationskoeffizient des CF II anhand einer Kalibriergeraden ermittelt werden. Die Fraktionen wurden in flüssigem Stickstoff eingefroren und bei -80°C für weitere Aktivitätsassays gelagert.

6.5 Protein-Aktivitätsassays und RNA-Bindungsassays

6.5.1 Filterbindungsassay

Der Filterbindungsassay stellt eine schnelle Methode zur Analyse von RNA-Proteininteraktionen dar und wurde in dieser Arbeit genutzt, um die Affinität von verschiedenen Proteinpräparationen für unterschiedliche RNA-Substrate zu bestimmen. Die Bindungsreaktionen erfolgten in einem einheitlichen Bindungspuffer. Die Bindungsstudien mit mPSF, CstF und CF I erfolgten in 1x FBK 100-Puffer. Die Bindungsstudien mit CF II erfolgten in 1x FBL 100-Puffer. Die Bindung an radioaktiv markierte RNAs wurde für 20–30 Minuten bei Raumtemperatur in 40–100 µl Reaktionen gestattet. Die RNA wurde in der Reaktion stets in einer konstanten Konzentration angeboten, welche fünf- bis zehnfach unterhalb des apparenten K_{50} lag. Die Proteinpräparationen wurden dann in Reaktionen mit steigender Konzentration angeboten. Direkt im Anschluss an die Bindungsreaktion wurden 95% der Reaktion auf einen Nitrozellulosefilter gegeben und nicht gebundene RNAs mit 5 ml eiskaltem Waschpuffer über Vakuum abgesaugt. Die von den Proteinen auf den Filtern zurückgehaltene, radioaktive RNA wurde anschließend mittels Szintillationszählung quantifiziert.

2x FBK/-L 100-Puffer	Filterbindungs-Waschpuffer:
200 mM Kalium-/Lithiumchlorid	100 mM Kalium-/Lithiumchlorid
100 mM Tris pH 8,0	50 mM Tris pH 8,0
20% (v/v) Glycerin	10% (v/v) Glycerin
0,02% (v/v) NP-40	0,1 mM EDTA
0,2 mM EDTA	

Als Kontrolle der Bindungsreaktion wurden stets Reaktionen mitgeführt, zu welcher kein Protein gegeben wurde. Diese wurden zum einen als Kontrolle der Hintergrundbindung über die Filter abgesaugt und gewaschen, zum anderen wurde die Kontrolle nicht über einen Nitrozellulosefilter abgesaugt, sondern direkt im Szintillationszähler gemessen und diente zur Berechnung der spezifischen Radioaktivität der eingesetzten Substrat-RNA:

$$A_{spez} = \frac{\text{radioaktive Zerfälle}}{\text{min} * n_{RNA}}$$

Über diese spezifische Radioaktivität wurde dann die Menge proteingebundener RNA für die anderen Reaktionen berechnet:

$$n_{gebunden} = \frac{\text{radioaktive Zerfälle}}{\text{min} * A_{spez}}$$

Die Menge gebundener RNA entspricht der Menge an gebildeten RNA-Proteinkomplexen. Damit ließ sich die Menge an freiem Protein berechnen, welche nach der Bindungsreaktion übrigblieb:

$$n_{\text{freies Prot.}} = n_{\text{gesamt Prot.}} - n_{\text{gebunden}}$$

Die Auswertung der Datenpunkte erfolgte durch hyperbole Anpassung der graphischen Auftragung des freien Proteins (X-Achse) gegen die Menge an gebundener RNA (Y-Achse). Die Konzentration an freiem Protein bei halbmaximaler RNA-Bindung entspricht dem apparenten K_{50} , des jeweils getesteten Proteins zum gewählten RNA-Substrat.

6.5.2 UV-Vernetzung von Proteinen an RNAs

Die UV-Vernetzung von RNA an Proteine erfolgte für die Proteinpräparationen von mPSF (Flag-Reinigung E2) und CF II (Ni-NTA-Elution E2).

mPSF

Die Vernetzung von 10 nM mPSF erfolgte an 10 nM 5'-markierte RNA (W10/iW10) in einem Reaktionsvolumen von 100 μ l unter Polyadenylierungsbedingungen, aber ohne ATP. Die Reaktion wurde für 10 Minuten bei 37 °C vorinkubiert und anschließend mit UV-Licht der Wellenlänge 312 nm (iW10) oder 253 nm (W10) für fünf Minuten bestrahlt. Die Bestrahlung erfolgte auf Eis und im Abstand von 3 cm zur Lichtquelle nach (Kühn u. a., 2003). Die Produkte wurde ohne RNase-Verdau durch SDS-PAGE und Phosphorimaging analysiert, für Antikörper-Affinitätsreinigung verwendet oder in limitierten proteolytischen Verdauen analysiert.

Analysen des His-WDR33 im mPSF-Komplex nach der UV-Vernetzung erfolgten nach dem folgenden modifizierten Protokoll. Es wurden 100 μ l der UV-vernetzten Reaktion mit 1,1 ml eines Denaturierungspuffers (6 M Guanidinium-HCl, 100 mM NaH_2PO_4 , 10 mM Tris/HCl pH 8, 0,05% NP-40) gemischt und für 15 Minuten bei Raumtemperatur inkubiert. Danach wurde die Reaktion mit 20 μ l Ni-NTA-Agarose (Qiagen) versetzt und nochmals für 30 Minuten unter Bewegung inkubiert. Die Matrix wurde danach einmal im Denaturierungspuffer und zweimal im Denaturierungspuffer mit 8 M Harnstoff anstatt Guanidinium-HCl gewaschen. Das gebundene Protein wurde in dem 8 M Harnstoffpuffer eluiert, indem dieser mit 100 mM EDTA versetzt wurde. Die Produkte wurde ohne RNase-Verdau durch SDS-PAGE und Phosphorimaging analysiert

CF II

UV-Vernetzung von 0,5 μ M CF II erfolgte an 2 μ M 5'-markierte rU₁₄-RNA in einem Reaktionsvolumen von 50 μ l (25 mM Tris-HCl pH 8,5, 100 mM KCl, 10% Saccharose, 0,02% NP-40, 0,2 mM EDTA, 0,5 mM ATP, 5 mM MgCl_2 , 0,5 mM DTT). Die Reaktion wurde für fünf Minuten bei Raumtemperatur vorinkubiert und anschließend wie zuvor eine Minute lang bei einer

Wellenlänge von 253 nm UV-Vernetzt. Die Produkte wurde ohne RNase-Verdau durch SDS-PAGE und Phosphorimaging analysiert oder für Antikörper-Affinitätsreinigung verwendet.

6.5.3 Limitierte Proteolyse von WDR33

Die limitierte Proteolyse des WDR33 erfolgte im Anschluss an eine mPSF UV-Vernetzung im 500 µl Maßstab. Die Reaktion wurde auf eine Konzentration von 0,75 M Harnstoff mit Dentaurationpuffer eingestellt und für 10 Minuten bei 37 °C inkubiert. Entweder wurde die Endoprotease Lys-C (Promega) anschließend im Verhältnis von 1/50 zur Masse des WDR33 eingesetzt oder die Endoprotease Trypsin im Verhältnis von 1/100. Der Verdau erfolgte bei 37 °C, wobei zu den jeweiligen angegebene Zeitpunkten Aliquots entnommen wurden, welche dem Ausgangsvolumen von 50 µl der UV-Vernetzung entsprachen. Gestoppt wurde der Verdau dann durch Zugabe von 800 µl Stopp-Lösung. Die Ni-NTA-Affinitätsreinigung der His-WDR33-Fragmente erfolgte wie in der UV-Vernetzung von Proteinen (Abschnitt 6.5.2) beschrieben. Die Produkte wurden im Anschluss in einer 10%-PA Tricin- SDS-PAGE (Schägger, 2006) aufgetrennt und durch Phosphorimaging analysiert.

Denaturierungspuffer	Stopp-Lösung
50 mM Tris/HCl pH 7,4	20 mM Tris pH 8
150 mM Natriumchlorid	100 mM NaH ₂ PO ₄
0,05 % NP-40	0,05 % NP-40
8 M Harnstoff	8 M Harnstoff

6.5.4 Electrophoretic Mobility Shift Assay (EMSA)

Der *electrophoretic mobility shift assay* (EMSA) wurde in dieser Arbeit für die Auftrennung von Bindungsreaktionen von mPSF an L3prä-RNA verwendet. In dem Assay werden Protein-(RNA)-Komplexe im nativen Gel anhand ihrer Eigenladung aufgetrennt. Die Pufferbedingungen während der Bindungsreaktion entsprachen denen der *in vitro* 3'-Spaltreaktion (Abschnitt 6.5.7) ohne, dass bei diesem Assay PVA, Cordycepintriphosphat oder Poly(A)-Polymerase zugesetzt wurde. Die Bindungsreaktionen mit steigenden Konzentrationen mPSF-Komplex und 0,1 nM intern markierter L3prä- oder L3prä Δ-RNA wurden für 30 Minuten bei 30 °C inkubiert. Direkt im Anschluss an die Bindungsreaktion wurden 50% des Reaktionsgemisches (nicht gekühlt) auf ein natives Composite-Gel (s. unten) geladen und elektrophoretisch aufgetrennt. Das Gel lief vor dem Beladen mindestens 30 Minuten bei 400 V und bei 8 °C vor und wurde unter angelegter Spannung (180V) beladen. Die Elektrophorese erfolgte dann bei konstanten 400 V und bei 8°C. Nach der

Elektrophorese wurde das Gel im Vakuum bei 60°C getrocknet und die radioaktiven Signale über Phosphorimaging detektiert. In den Autoradiogrammen wurden anschließend die Signale der freien (nicht-gebundenen) RNA-Substrate, sowie der langsamer migrierenden RNA-Proteinkomplexe über ImageQuant quantifiziert.

Natives Composite-Gel

6% Acrylamid:Bisacrylamid (80:1)

0,5% (w/v) Agarose

0,5x TBE

Für ein Gel (20x20x0,1 cm) wurden etwa 50 ml Gellösung hergestellt. Diese wurde durch Mischen zweier Lösungen (25 ml Acrylamidlösung in 1x TBE und 25 ml 55°C warme 1% Agaroselösung in Wasser) hergestellt. Unverzüglich nach dem Mischen wurde die Polymerisation des Acrylamids durch Zugabe von 80 µl 10% (w/v) APS-Lösung und 80 µl TEMED gestartet und die Gellösung zwischen vorbereitete Glasplatten gegossen.

6.5.5 Native Polyacrylamid-Gelelektrophorese

Die native gelelektrophoretische Auftrennung von (radioaktiv markierten) RNAs und RNA-Proteinkomplexen aus RNA-Bindungsreaktionen wurde durch native TBE-PAGE realisiert. Es wurden 200x200x1 mm-Gele (bis 21 Proben) verwendet. Die benötigten Gellösungen wurden jeweils frisch hergestellt und die Acrylamid- und die TBE-Konzentration wurden durch Verdünnung einer 30% (w/v) Acrylamid-Stammlösung (60:1) und einer 10x TBE-Stammlösung eingestellt.

Gellösung

0,5x TBE

5% (w/v) Acrylamid:Bisacrylamid (60:1)

Für die Polymerisation von 50 ml Gellösung wurden 250 µl 10%ige (w/v) APS-Lösung und 80 µl TEMED hinzugegeben, kurz gemischt und das Gel zügig und stetig zwischen die vorbereiteten Glasplatten gegossen. Um eine vollständige Polymerisation des Gels zu erreichen, wurde die Polymerisation etwa eine Stunde lang erlaubt, bevor das Gel weiterverwendet wurde. Die Gelelektrophorese unter nativen Bedingungen wurde nach einem mindestens 30-minütigen

Vorlauf bei einer Betriebstemperatur von maximal 8 °C und in 0,5x TBE als Laufpuffer im Kühlraum durchgeführt. Die Proben wurden für die Elektrophorese bei einer angelegten Spannung von 5 V/cm (100 V) direkt aus der Bindungsreaktion geladen. Die Elektrophorese erfolgte danach bei 20 V/cm (400 V). Bei der Auftrennung wurden Farbstoffe wie Xylencyanol und Bromphenolblau als Referenz des Migrationsverhaltens der Nukleinsäuren neben den Proben geladen.

6.5.6 5'-Polynukleotidkinase Assays mit CF II und hClp1

CF II/hClp1-PNK Reaktion

variable Komponenten	konstante Komponenten	Stopp-Mix
2–5 nM CF II/hClp1	50 mM HEPES pH 8,0	100 mM EDTA
0,1–5 mM ATP	100 mM Kaliumchlorid	1% (w/v) SDS
25–6,5 µCi $\gamma^{32}\text{P}$ -ATP	3 mM Magnesiumacetat	
RNA (0,1–25 µM)	10% (w/v) Saccharose	
H ₂ O ad. 32,5 µl	0,02% NP-40	
	5 µM Zinkchlorid	
	0,4 U RNasin	
	1 mM DTT	

Die Reaktion wurde ohne Protein vorgelegt und für fünf Minuten bei 30 °C vorinkubiert. Die Reaktion wurde durch Zugabe des Proteins (CF II MonoQ-Fraktion 16 oder hClp1 Superdex200 pg-Fraktion 18) gestartet. Die 5'-PNK-Reaktion erfolgte dann bei 30°C in einer Zeitkinetik von null bis fünf Minuten für jede untersuchte RNA- oder ATP-Konzentration. Der Nullwert wurde in Doppelbestimmung erfasst, indem eine Reaktion ohne Protein vor Start der Kinetik abgenommen wurde. In ATP-Titrationsen wurde die rC₁₄-Konzentration konstant bei 20 µM gehalten. Die Reaktion wurde gestoppt, indem zu jedem Zeitpunkt 5 µl der Reaktion abgenommen und auf 5 µl vorgelegten Stopp-Mix gegeben wurden. Die Reaktion wurde danach auf Eis gelagert, bis das freie ATP durch die DEAE-Methode von der RNA abgetrennt wurde. Die verbliebene Radioaktivität und somit das inkorporierte γ -³²Phosphat wurde mittels Szintillationsmessung quantifiziert. In RNA-Titrationsen wurde die Reaktion bei konstanter ATP-Konzentration von 1 mM durchgeführt. Die Reaktionen wurden gestoppt, indem zu jedem Zeitpunkt 5 µl der Reaktionen abgenommen und auf 5 µl vorgelegten Formamid-Puffer gegeben wurden. Die Reaktion wurde danach auf Eis gelagert, am Ende für fünf Minuten bei 95 °C denaturiert und in einer 15%-PA Harnstoff-PAGE denaturierend aufgetrennt. So konnte das freie

ATP abgetrennt werden. Als Referenz zur Berechnung des inkorporierten γ - 32 P-Phosphat wurde eine Verdünnung des Null-Minuten-Ansatzes geladen und eine spezifische Aktivität des γ - 32 P-ATPs berechnet. Als Größenstandard wurde eine T4-PNK markierte Referenz der Substrat-RNA geladen. Die DEAE-Methode konnte hier wegen der geringen RNA-Konzentrationen nicht verwendet werden, da die Hintergrundsignale des radioaktiven ATP zu stark waren. Die Detektion der Signale erfolgte über Autoradiographie und Phosphorimager und die Quantifizierung in ImageQuant.

6.5.7 RNA 3'-Spaltassay

Zur Analyse von Proteinpräparationen auf 3'-Spaltaktivität wurden diese mit den spaltbaren, intern radioaktiv markierten RNA-Substraten L3 oder SV40 *late* inkubiert. Die Spezifität der Polyadenylierungssignal-abhängigen 3'-Spaltung wurde überprüft, indem auch RNA-Substrate mit mutiertem Polyadenylierungssignal (U zu G) angeboten wurden. Die durchgeführten Spaltungsreaktionen hatten ein Volumen von 20 μ l und besaßen die folgende Zusammensetzung:

1x Spaltreaktion	2,5x Cleavage Buffer (CB)
<hr/>	<hr/>
1x <i>Cleavage Buffer</i>	50 mM HEPES, pH 7,9
100 mM Kaliumchlorid (gesamt)	5 mM DTT
2,5 nM RNA-Substrat	2,5 mM Magnesiumchlorid
2 μ l (fraktionierter) NXT	50 mM Kreatinphosphat
1 bis 8 μ l Proteinpräparation(en)	10% (w/v) PVA
CDB ad. 20 μ l	1,25 mM Cordycepintriphosphat
	0,1% (v/v) Tween 20
Proteinase K-Puffer [2x]:	Cleavage Dilution Buffer (CDB)
<hr/>	<hr/>
200 mM Tris-HCl, pH 8,0	50 mM HEPES, pH 7,9
25 mM EDTA	200 mM Kaliumchlorid
300 mM Natriumchlorid 2% (w/v) SDS	1 mM DTT
	0,2 g/l methyliertes BSA
	0,5 mM EDTA
	10% (w/v) Glycerin
	0,1% (v/v) Tween 20

Die Spaltungsreaktionen wurden für zwei Stunden bei 30°C inkubiert und durch Zugabe von jeweils 80 µl Stopp-Mix (1x PK-Puffer, 0,1 g/l Glykogen, 0,05 g/l Proteinase K) beendet und für weitere 30 Minuten bei 37°C inkubiert. Anschließend wurden die RNAs mittels Ethanolfällung präzipitiert, in 2 µl HEPES-Puffer (10 mM, pH 7,0) und 4 µl Formamid-Ladepuffer aufgenommen und in einer Harnstoff-PAGE aufgetrennt. Die radioaktiven RNA-Fragmente im Gel wurden über Autoradiographie detektiert und falls nötig in ImageQuant densitometrisch ausgewertet. Bei der Berechnung von Spaltungsaktivitäten wurde der Signalverlust durch das 3'-Spaltprodukt berücksichtigt.

6.5.8 RNA-Selektion aus Pools mit randomisierten 14-mer Sequenzen (Bind'n'Seq)

Die RNA-Selektionen wurden durch T. Treiber zum einen mit hPcf11 ΔN1184 und zum anderen mit CF II Wt-Komplex mehrere Male wiederholt durchgeführt.

Es wurden 100 nM hPcf11 ΔN1184 in 400 µl Filterbindungspuffer mit 15 µg phosphorylierter 14-mer RNA der Sequenz (N)₁₄GUUU versetzt und für 30 Minuten bei Raumtemperatur unter Rotation inkubiert. Die Reaktion wurde dann über einen Nitrocellulose-Filter (GE Healthcare) abgesaugt, welcher mit Filterbindungspuffer mit 15 µg/ml Heparin benetzt war. Im Anschluss wurde mit 20 ml eiskaltem Filterbindungspuffer gewaschen. Der gebundene Protein-RNA-Komplex wurde ausgeschnitten und die RNA durch Zugabe von 250 µl RNA-Elutionspuffer durch 30-minütiges schütteln bei 70 °C eluiert.

Die CF II-Reaktionen wurden analog mit 100 nM Protein in 400 µl CF II-Bindungspuffer vorgelegt und durch Zugabe von 20µl Ni-NTA-Sepharose (GE Healthcare) an die Matrix gebunden. Dies wurde dann mit 15 µg des (N)₁₄GUUU-RNA-Pools versetzt und für 30 Minuten bei Raumtemperatur unter Rotation inkubiert. Die Ni-NTA-Beads wurden durch Zentrifugation (1000 g, zwei Minuten, 4°C) präzipitiert, dreimal mit eiskaltem CF II-Bindungspuffer gewaschen und die gebundene RNA durch Zugabe von 200 µl Ni-Elutionspuffer eluiert. In beiden Fällen wurden auch Kontrollen ohne Protein mitgeführt, um eine Hintergrund-Bindung von RNA auszuschließen.

Ni-Elutionspuffer	RNA-Elutionspuffer
10 mM Tris pH 7,0	50% Formamid
400 mM NaCl	1,8 M Natriumacetat
1 mM EDTA	2 mM EDTA
1% (w/v) SDS	0,2% (w/v) SDS

CF II- Bindungspuffer	Filterbindungspuffer
25 mM Tris pH 7,5	25 mM Tris pH 8,0
150 mM KCl	50 mM KCl
3 mM MgCl ₂	2 mM MgCl ₂
0,01% (v/v) NP-40	1 mM DTT
1 g/l BSA	0,01% NP-40
5% (v/v) Glycerin	10% (v/v) Glycerin
0,1 U/μl Ribolock	0,1 U/μl Ribolock (ThermoFischer)

Die jeweiligen eluierten RNAs der Selektionen wurden durch Phenol/Chloroform extrahiert und durch Ethanol-fällung präzipitiert. Die RNAs wurden dann an einen 3'-DNA-Adapter (AAACTGGAATTCTCGGGTGCCAAGG-Amino-C7) und an einen 5'-RNA-Adapter mit T7-Promoter (GUUCAGUAAUACGACUCACUAUAGGG) ligiert. Die Ligationsprodukte wurden dann mit dem *First Strand cDNA Synthesis Kit* (ThermoFischer) und dem Primer „GCCTTGCCACCCGAGAATTCCAGTTT“ revers-transkribiert. Anschließend erfolgte eine PCR-Amplifikation der cDNA-Sequenzen in der über die verwendeten Primer „*Barcodes*“ für die folgende *next generation sequencing* (NGS) Reaktion eingeführt wurden. Die PCR-Produkte wurden durch eine Elektrophorese über ein 6% PA Harnstoff PAGE aufgetrennt und Banden auf Höhe von 150 nt ausgeschnitten. Die DNA wurde dann über Nacht in 0,4 M NaCl eluiert und mittels Ethanol-fällung präzipitiert. Das DNA-Pellet wurde in Wasser gelöst und für die NGS und T7-Transkription verwendet.

Die RNA-Substrate der zweiten Selektionsrunde wurden aus 50 ng cDNA des ersten PCR-Pools generiert. Dazu wurde der PCR-Pool erneut mit den Primern „AATGATACGGCGACCACCGAGATCTACACGTTTCAGTAATACGACTCACTATAGG“ und „GCCTTGCCACCCGAGAATTCCAGTTT“ amplifiziert. Die Produkte wurden mit einem PCR-CleanUp Kit (Macherey-Nagel) gereinigt und durch Zugabe von 1,5 μl des Restriktionsenzym *FastDigest MssI* (ThermoFisher) verdaut, welches die Sequenz GTTTAAAC registriert. Diese entsteht nur bei der erfolgreichen Ligation des 3'-Adapters und sichert so eine qualitative Überprüfung dieses Schrittes. Die gespaltene DNA wurde dann in einer T7-Transkription in einen neuen RNA-Pool umgeschrieben, welcher die Sequenz GGG(N)₁₄GUUU enthält. Die generierte RNA wurde über eine 18% PA Harnstoff-PAGE wie zuvor gereinigt, dephosphoryliert mit FastAP (ThermoFisher) und abschließend mit T4-PNK am 5'-Ende phosphoryliert. Die Bindungsreaktionen mit hPcf11 ΔN1184 oder CF II Wt wurden dann mit 15 μg des einmal selektierten RNA-Pools wiederholt und so erneut selektiert und wie zuvor aufgearbeitet.

Die Sequenz-Bibliotheken wurden mit dem 150 *cycle MiSeq Reagent Kit* generiert, welchem der Sequenzier-Primer (5'-GATCTACACGTTTCAGTAATACGACTCACTATAGGG-3') zugegeben wurde.

Dies erfolgte über Sequenzierung mittels des MiSeq-Instruments (Illumina). Die Analyse der selektierten Sequenzen wurde auf die zweite Runde der Selektion begrenzt. Die erhaltenen Sequenzen wurden nach den eingefügten „Barcodes“ sortiert und gefiltert auf Sequenzen, welche sowohl den 3'-Adapter enthalten als auch mittels *MssI* prozessiert waren. Nach dem Zuschneiden auf die variablen (N)₁₄-Sequenzen (14 nt) wurden die Sequenzen auf eine statistische Anreicherung von Motiven mit der Software Weeder2 (Zambelli u. a., 2014) überprüft. Zwischen 14,000 und 44,000 verwertbare Sequenzen wurden so pro Experiment analysiert.

7 Literaturverzeichnis

7.1 Wissenschaftliche Publikationen

Baejen, C., Andreani, J., Torkler, P., Battaglia, S., Schwalb, B., Lidschreiber, M., Maier, K.C., Boltendahl, A., Rus, P., Esslinger, S., u. a. (2017). Genome-wide Analysis of RNA Polymerase II Termination at Protein-Coding Genes. *Mol. Cell* 66, 38-49.e6.

Bai, Y., Auperin, T.C., Chou, C.Y., Chang, G.G., Manley, J.L., und Tong, L. (2007). Crystal Structure of Murine CstF-77: Dimeric Association and Implications for Polyadenylation of mRNA Precursors. *Mol. Cell* 25, 863–875.

Balbo, P.B., und Bohm, A. (2007). Mechanism of Poly(A) Polymerase: Structure of the Enzyme-MgATP-RNA Ternary Complex and Kinetic Analysis. *Structure*.

Barabino, S.M.L., Hübner, W., Jenny, A., Minvielle-Sebastia, L., und Keller, W. (1997). The 30-kd subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are rna-binding zinc finger proteins. *Genes Dev.* 11, 1703–1716.

Barabino, S.M.L., Ohnacker, M., und Keller, W. (2000). Distinct roles of two Yth1p domains in 3'-end cleavage and polyadenylation of yeast pre-mRNAs. *EMBO J.* 19, 3778–3787.

Bardwell, V.J., Zarkower, D., Edmonds, M., und Wickens', M. (1990). The Enzyme That Adds Poly(A) to mRNAs Is a Classical Poly(A) Polymerase. *Mol. Cell. Biol.* 10, 846–849.

Barillà, D., Lee, B.A., und Proudfoot, N.J. (2001). Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.*

Bedrat, A., Lacroix, L., und Mergny, J.-L. (2016). Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* 44, 1746–1759.

Berger, I., Fitzgerald, D.J., und Richmond, T.J. (2004). Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* 22, 1583–1587.

Berger, I., Craig, A., Fitzgerald, D.J., Richmond, T.J., Fitzgerald, D.J., und Richmond, P.T.J. (2010). MultiBac Expression System User Manual (European Molecular Biology Laboratory (EMBL)).

Beyer, K., Dandekar, T., und Keller, W. (1997). RNA ligands selected by cleavage stimulation factor contain distinct sequence motifs that function as downstream elements in 3'-end processing of pre-mRNA. *J. Biol. Chem.* 272, 26769–26779.

Bieniossek, C., Richmond, T.J., und Berger, I. (2008). MultiBac: Multigene baculovirus-based eukaryotic protein complex production. *Curr. Protoc. Protein Sci.* 1–26.

- Bienroth, S., Wahle, E., Suter-Crazzolaro, C., und Keller, W. (1991). Purification and characterization of the cleavage and polyadenylation specificity factor involved in the 3' processing of messenger RNA precursors. *J. Biol. Chem.* *266*, 19768–19776.
- Bienroth, S., Keller, W., und Wahle, E. (1993). Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.* *12*, 585–594.
- Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W., und Proudfoot, N.J. (1998). Coupling termination of transcription to messenger RNA maturation in yeast. *Science (80-.)*. *280*, 298–301.
- Bochman, M.L., Paeschke, K., und Zakian, V.A. (2012). DNA secondary structures: Stability and function of G-quadruplex structures. *Nat. Rev. Genet.*
- Brown, K.M., und Gilmartin, G.M. (2003). A Mechanism for the Regulation of Pre-mRNA 3' Processing by Human Cleavage Factor Im. *Mol. Cell* *12*, 1467–1476.
- Burnette, W.N. (1981). "Western Blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* *112*, 195–203.
- Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.-M., Foehr, S., Curk, T., Krijgsveld, J., und Hentze, M.W. (2016). Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell* *63*, 696–710.
- Chen, F., MacDonald, C.C., und Wilusz, J. (1995). Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res.* *23*, 2614–2620.
- Christofori, G., und Keller, W. (1989). Poly(A) polymerase purified from HeLa cell nuclear extract is required for both cleavage and polyadenylation of pre-mRNA in vitro. *Mol. Cell. Biol.* *9*, 193–203.
- Clerici, M., Faini, M., Aebersold, R., und Jinek, M. (2017). Structural insights into the assembly and polyA signal recognition mechanism of the human CPSF complex. *Elife* *6*.
- Clerici, M., Faini, M., Muckenfuss, L.M., Aebersold, R., und Jinek, M. (2018). Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. *Nat. Struct. Mol. Biol.* *25*, 135–138.
- Connelly, S., und Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* *2*, 440–452.
- Corden, J.L. (2016). Pol II CTD Code Light. *Mol. Cell* *61*, 183–184.
- Crooks, G.E., Hon, G., Chandonia, J.M., und Brenner, S.E. (2004). WebLogo: A sequence logo

generator. *Genome Res.*

Dass, B., McMahon, K.W., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., und MacDonald, C.C. (2001). The Gene for a Variant Form of the Polyadenylation Protein CstF-64 Is on Chromosome 19 and Is Expressed in Pachytene Spermatocytes in Mice. *J. Biol. Chem.* *276*, 8044–8050.

Davidson, L., Muniz, L., und West, S. (2014). 3' end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes Dev.* *28*, 342–356.

Dettwiler, S., Aringhieri, C., Cardinale, S., Keller, W., und Barabino, S.M.L.L. (2004). Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *J. Biol. Chem.* *279*, 35788–35797.

Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H., und Keller, W. (2002). A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. *Mol. Cell.*

Dikfidan, A., Loll, B., Zeymer, C., Magler, I., Clausen, T., und Meinhart, A. (2014). RNA Specificity and Regulation of Catalysis in the Eukaryotic Polynucleotide Kinase Clp1. *Mol. Cell* *54*, 975–986.

Erickson, H.P. (2009). Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biol. Proced. Online* *11*, 32–51.

Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., und Berger, I. (2006). Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* *3*, 1021–1032.

Ghazy, M.A., Gordon, J.M.B., Lee, S.D., Singh, B.N., Bohm, A., Hampsey, M., und Moore, C. (2012). The interaction of Pcf11 and Clp1 is needed for mRNA 3'-end formation and is modulated by amino acids in the ATP-binding site. *Nucleic Acids Res.* *40*, 1214–1225.

Di Giammartino, D.C., Li, W., Ogami, K., Yashinski, J.J., Hoque, M., Tian, B., und Manley, J.L. (2014). RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3' UTRs. *Genes Dev.* *28*, 2248–2260.

Gordon, J.M.B., Shikov, S., Kuehner, J.N., Liriano, M., Lee, E., Stafford, W., Poulsen, M.B., Harrison, C., Moore, C., und Bohm, A. (2011). Reconstitution of CF IA from Overexpressed subunits reveals stoichiometry and provides insights into molecular topology. *Biochemistry* *50*, 10203–10214.

Gross, S., und Moore, C. (2001). Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proc. Natl. Acad. Sci.* *98*, 6080–6085.

Grozdanov, P.N., Masoumzadeh, E., Latham, M.P., und MacDonald, C.C. (2018). The structural basis of CstF-77 modulation of cleavage and polyadenylation through stimulation of CstF-64 activity.

Nucleic Acids Res.

Gruber, A.J., und Zavolan, M. (2019). Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* *20*, 599–614.

Gruber, A.J., Schmidt, R., Gruber, A.R., Martin, G., Ghosh, S., Belmadani, M., Keller, W., und Zavolan, M. (2016). A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. *Genome Res.* *26*, 1145–1159.

Guéguéniat, J., Dupin, A.F., Stojko, J., Beaurepaire, L., Cianférani, S., Mackereth, C.D., Minvielle-Sébastien, L., und Fribourg, S. (2017). Distinct roles of Pcf11 zinc-binding domains in pre-mRNA 3' end processing. *Nucleic Acids Res.* *45*, 10115–10131.

Hanada, T., Weitzer, S., Mair, B., Bernreuther, C., Wainger, B.J., Ichida, J., Hanada, R., Orthofer, M., Cronin, S.J., Komnenovic, V., u. a. (2013). CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* *495*, 474–480.

Hill, C.H., Boreikaitė, V., Kumar, A., Casañal, A., Kubík, P., Degliesposti, G., Maslen, S., Mariani, A., von Loeffelholz, O., Girbig, M., u. a. (2019). Activation of the Endonuclease that Defines mRNA 3' Ends Requires Incorporation into an 8-Subunit Core Cleavage and Polyadenylation Factor Complex. *Mol. Cell* *73*, 1217-1231.e11.

Hirose, Y., und Manley, J.L. (1998). RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* *395*, 93–96.

Ho, C.K., und Shuman, S. (1999). Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* *3*, 405–411.

Holbein, S., Scola, S., Loll, B., Dichtl, B.S., Hübner, W., Meinhart, A., und Dichtl, B. (2011). The P-Loop Domain of Yeast Clp1 Mediates Interactions Between CF IA and CPF Factors in Pre-mRNA 3' End Formation. *PLoS One* *6*, e29139.

Hollingworth, D., Kelly, G., Frenkiel, T.A., Noble, C.G., Taylor, I.A., und Ramos, A. (2005). Backbone assignment of PCF11 CTD binding domain [1]. *J. Biomol. NMR* *31*, 363.

Hollingworth, D., Noble, C.G., Taylor, I.A., und Ramos, A. (2006). RNA polymerase II CTD phosphopeptides compete with RNA for the interaction with Pcf11. *RNA* *12*, 555–560.

HU, J., Lutz, C.S., Wilusz, J., und Tian, B. (2005). Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA* *11*, 1485–1493.

Hudson, B.P., Martinez-Yamout, M.A., Dyson, H.J., und Wright, P.E. (2004). Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat. Struct. Mol. Biol.* *11*, 257–264.

- Jenny, A., Hauri, H.P., und Keller, W. (1994). Characterization of cleavage and polyadenylation specificity factor and cloning of its 100-kilodalton subunit. *Mol. Cell. Biol.* *14*, 8183–8190.
- Jin, W., Wang, Y., Liu, C.-P., Yang, N., Jin, M., Cong, Y., Wang, M., und Xu, R.-M. (2016). Structural basis for snRNA recognition by the double-WD40 repeat domain of Gemin5. *Genes Dev.* *30*, 2391–2403.
- Jove, R., und Manley, J.L. (1984). In vitro transcription from the adenovirus 2 major late promoter utilizing templates truncated at promoter-proximal sites. *J. Biol. Chem.* *259*, 8513–8521.
- Kamieniarz-Gdula, K., Gdula, M.R., Panser, K., Nojima, T., Monks, J., Wiśniewski, J.R., Riepsaame, J., Brockdorff, N., Pauli, A., und Proudfoot, N.J. (2019). Selective Roles of Vertebrate PCF11 in Premature and Full-Length Transcript Termination. *Mol. Cell* *74*, 158-172.e9.
- Karaca, E., Weitzer, S., Pehlivan, D., Shiraishi, H., Gogakos, T., Hanada, T., Jhangiani, S.N., Wiszniewski, W., Withers, M., Campbell, I.M., u. a. (2014). Human CLP1 Mutations Alter tRNA Biogenesis, Affecting Both Peripheral and Central Nervous System Function. *Cell* *157*, 636–650.
- Kaufmann, I., Martin, G., Friedlein, A., Langen, H., und Keller, W. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.* *23*, 616–626.
- Kecman, T., Kuś, K., Heo, D.-H., Duckett, K., Birot, A., Liberatori, S., Mohammed, S., Geis-Asteggiane, L., Robinson, C. V., und Vasiljeva, L. (2018). Elongation/Termination Factor Exchange Mediated by PP1 Phosphatase Orchestrates Transcription Termination. *Cell Rep.* *25*, 259-269.e5.
- Keller, W., Bienroth, S., Lang, K.M., und Christofori, G. (1991). Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J.* *10*, 4241–4249.
- Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., und Wahle, E. (2003). Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J.* *22*, 3705–3714.
- Kolev, N.G., und Steitz, J.A. (2005). Symplekin and multiple other polyadenylation factors participate in 3'-end maturation of histone mRNAs. *Genes Dev.* *19*, 2583–2592.
- Kolev, N.G., Yario, T.A., Benson, E., und Steitz, J.A. (2008). Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation. *EMBO Rep.* *9*, 1013–1018.
- Kühn, U., Nemeth, A., Meyer, S., und Wahle, E. (2003). The RNA binding domains of the nuclear poly(A)-binding protein. *J. Biol. Chem.* *278*, 16916–16925.

- Kühn, U., Gündel, M., Knoth, A., Kerwitz, Y., Rüdell, S., und Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear Poly(A)-binding protein regulating the interaction between Poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J. Biol. Chem.* *284*, 22803–22814.
- Kühn, U., Buschmann, J., und Wahle, E. (2017). The nuclear poly(A) binding protein of mammals, but not of fission yeast, participates in mRNA polyadenylation. *RNA* *23*, 473–482.
- Kyburz, A., Sadowski, M., Dichtl, B., und Keller, W. (2003). The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. *Nucleic Acids Res.*
- Lackford, B., Yao, C., Charles, G.M., Weng, L., Zheng, X., Choi, E.A., Xie, X., Wan, J., Xing, Y., Freudenberg, J.M., u. a. (2014). Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J.* *33*, 878–889.
- LAEMMLI, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* *227*, 680–685.
- Lambert, N.J., Robertson, A.D., und Burge, C.B. (2015). RNA Bind-n-Seq. In *Methods in Enzymology*, S. 465–493.
- Lidschreiber, M., Easter, A.D., Battaglia, S., Rodríguez-Molina, J.B., Casañal, A., Carminati, M., Baejen, C., Grzechnik, P., Maier, K.C., Cramer, P., u. a. (2018). The APT complex is involved in non-coding RNA transcription and is distinct from CPF. *Nucleic Acids Res.*
- Lunde, B.M., Reichow, S.L., Kim, M., Suh, H., Leeper, T.C., Yang, F., Mutschler, H., Buratowski, S., Meinhart, A., und Varani, G. (2010). Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.* *17*, 1195–1201.
- Lyons, S.M., Fay, M.M., und Ivanov, P. (2018). The role of RNA modifications in the regulation of tRNA cleavage. *FEBS Lett.* *592*, 2828–2844.
- Mandel, C.R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J.L., und Tong, L. (2006). Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* *444*, 953–956.
- Martin, G., Gruber, A.R., Keller, W., und Zavolan, M. (2012). Genome-wide Analysis of Pre-mRNA 3' End Processing Reveals a Decisive Role of Human Cleavage Factor I in the Regulation of 3' UTR Length. *Cell Rep.* *1*, 753–763.
- Mayer, A., Schrieck, A., Lidschreiber, M., Leike, K., Martin, D.E., und Cramer, P. (2012). The Spt5 C-Terminal Region Recruits Yeast 3' RNA Cleavage Factor I. *Mol. Cell. Biol.* *32*, 1321–1331.
- Mayr, C. (2017). Regulation by 3'-Untranslated Regions. *Annu. Rev. Genet.* *51*, 171–194.

- McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., und Bentley, D.L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* *11*, 3306–3318.
- Meinhart, A., und Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* *430*, 223–226.
- Millevoi, S., und Vagner, S. (2009). Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res.*
- Monarez, R.R., MacDonald, C.C., und Dass, B. (2007). Polyadenylation proteins CstF-64 and tauCstF-64 exhibit differential binding affinities for RNA polymers. *Biochem. J.* *401*, 651–658.
- Moreno-Morcillo, M., Minvielle-Sébastien, L., Fribourg, S., und MacKereth, C.D. (2011). Locked tether formation by cooperative folding of Rna14p monkeytail and Rna15p hinge domains in the yeast CF IA complex. *Structure.*
- Murthy, K.G.K., und Manley, J.L. (1992). Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.*
- Murthy, K.G.K., und Manley, J.L. (1995). The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev.* *9*, 2672–2683.
- Noble, C.G., Hollingworth, D., Martin, S.R., Ennis-Adeniran, V., Smerdon, S.J., Kelly, G., Taylor, I.A., und Ramos, A. (2005). Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nat. Struct. Mol. Biol.* *12*, 144–151.
- Noble, C.G., Beuth, B., und Taylor, I.A. (2007). Structure of a nucleotide-bound Clp1-Pcf11 polyadenylation factor. *Nucleic Acids Res.* *35*, 87–99.
- Ogorodnikov, A., Kargapolova, Y., und Danckwardt, S. (2016). Processing and transcriptome expansion at the mRNA 3' end in health and disease: finding the right end. *Pflügers Arch. Eur. J. Physiol.*
- Ogorodnikov, A., Levin, M., Tattikota, S., Tokalov, S., Hoque, M., Scherzinger, D., Marini, F., Poetsch, A., Binder, H., Macher-Göppinger, S., u. a. (2018). Transcriptome 3' end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat. Commun.* *9*, 1–16.
- Paushkin, S. V., Patel, M., Furia, B.S., Peltz, S.W., und Trotta, C.R. (2004). Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. *Cell.*
- Pearson, E.L., Graber, J.H., Lee, S.D., Naggert, K.S., und Moore, C.L. (2019). Ipa1 Is an RNA Polymerase II Elongation Factor that Facilitates Termination by Maintaining Levels of the Poly(A)

- Site Endonuclease Ysh1. *Cell Rep.* 26, 1919-1933.e5.
- Pérez Cañadillas, J.M., und Varani, G. (2003). Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *EMBO J.* 22, 2821–2830.
- Proudfoot, N.J. (2011). Ending the message: Poly(A) signals then and now. *Genes Dev.*
- Qian, Z.W., und Wilusz, J. (1991). An RNA-binding protein specifically interacts with a functionally important domain of the downstream element of the simian virus 40 late polyadenylation signal. *Mol. Cell. Biol.* 11, 5312–5320.
- Ramirez, A., Shuman, S., und Schwer, B. (2008). Human RNA 5'-kinase (hClp1) can function as a tRNA splicing enzyme in vivo. *RNA* 14, 1737–1745.
- Rasmussen, E.B., und Lis, J.T. (1993). In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes. *Proc. Natl. Acad. Sci. U. S. A.* 90, 7923–7927.
- Rüegsegger, U., Beyer, K., und Keller, W. (1996). Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. *J. Biol. Chem.* 271, 6107–6113.
- Rüegsegger, U., Blank, D., und Keller, W. (1998). Human pre-mRNA cleavage factor ImIs related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Mol. Cell* 1, 243–253.
- Ruepp, M.-D., Schweingruber, C., Kleinschmidt, N., und Schümperli, D. (2011). Interactions of CstF-64, CstF-77, and symplekin: Implications on localisation and function. *Mol. Biol. Cell* 22, 91–104.
- Ryan, K., Calvo, O., und Manley, J.L. (2004). Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA* 10, 565–573.
- Sadowski, M., Dichtl, B., Hübner, W., und Keller, W. (2003). Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. *EMBO J.* 22, 2167–2177.
- Sambrook, J., und Russell, D.W. (2001). *Molecular Cloning - Sambrook & Russel - Vol. 1, 2, 3.* Hum. Mutat.
- Schäfer, P., Tüting, C., Schönemann, L., Kühn, U., Treiber, T., Treiber, N., Ihling, C., Graber, A., Keller, W., Meister, G., u. a. (2018). Reconstitution of mammalian cleavage factor II involved in 3' processing of mRNA precursors. *RNA* 24, 1721–1737.
- Schägger, H. (2006). Tricine–SDS–PAGE. *Nat. Protoc.* 1, 16–22.
- Schellenberger, A., Fischer, G., Hübner, G., und Ulbrich, R. (1989). *Enzymkatalyse* (Jena: VEB Gustav Fischer Verlag Jena).

- Schönemann, L., Kühn, U., Martin, G., Schäfer, P., Gruber, A.R., Keller, W., Zavolan, M., und Wahle, E. (2014). Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. *Genes Dev.* *28*, 2381–2393.
- Schrieck, A., Easter, A.D., Etzold, S., Wiederhold, K., Lidschreiber, M., Cramer, P., und Passmore, L.A. (2014). RNA polymerase II termination involves C-terminal-domain tyrosine dephosphorylation by CPF subunit Glc7. *Nat. Struct. Mol. Biol.* *21*, 175–179.
- Schüller, R., Forné, I., Straub, T., Schrieck, A., Texier, Y., Shah, N., Decker, T.M., Cramer, P., Imhof, A., und Eick, D. (2016). Heptad-Specific Phosphorylation of RNA Polymerase II CTD. *Mol. Cell* *61*, 305–314.
- Sheets, M.D., Stephenson, P., und Wickens, M.P. (1987). Products of in vitro cleavage and polyadenylation of simian virus 40 late pre-mRNAs. *Mol. Cell. Biol.* *7*, 1518–1529.
- Shi, Y. (2014). CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev.* *28*, 2370–2380.
- Shi, Y. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell Biol.* *18*, 655–670.
- Shi, Y., und Manley, J.L. (2015). The end of the message: Multiple protein–RNA interactions define the mRNA polyadenylation site. *Genes Dev.*
- Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., Frank, J., und Manley, J.L. (2009). Molecular Architecture of the Human Pre-mRNA 3' Processing Complex. *Mol. Cell* *33*, 365–376.
- Siegel, L.M., und Monty, K.J. (1966). Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. *BBA - Biophys. Incl. Photosynth.*
- Stayton, M.M., und Kornberg, A. (1983). Complexes of Escherichia coli primase with the replication origin of G4 phage DNA. *J. Biol. Chem.* *258*, 13205–13212.
- Suh, H., Ficarro, S.B., Kang, U.B., Chun, Y., Marto, J.A., und Buratowski, S. (2016). Direct Analysis of Phosphorylation Sites on the Rpb1 C-Terminal Domain of RNA Polymerase II. *Mol. Cell* *61*, 297–304.
- Sullivan, K.D., Steiniger, M., und Marzluff, W.F. (2009). A Core Complex of CPSF73, CPSF100, and Symplekin May Form Two Different Cleavage Factors for Processing of Poly(A) and Histone mRNAs. *Mol. Cell* *34*, 322–332.
- Sun, Y., Zhang, Y., Hamilton, K., Manley, J.L., Shi, Y., Walz, T., und Tong, L. (2018). Molecular basis

for the recognition of the human AAUAAA polyadenylation signal. *Proc. Natl. Acad. Sci.* *115*, E1419–E1428.

Sun, Y., Zhang, Y., Aik, W.S., Yang, X.-C., Marzluff, W.F., Walz, T., Dominski, Z., und Tong, L. (2020). Structure of an active human histone pre-mRNA 3'-end processing machinery. *Science* (80-). *367*, 700–703.

Takagaki, Y., und Manley, J.L. (1994). A polyadenylation factor subunit is the human homologue of the *Drosophila* suppressor of forked protein. *Nature* *372*, 471–474.

Takagaki, Y., und Manley, J.L. (1997). RNA recognition by the human polyadenylation factor CstF. *Mol. Cell. Biol.* *17*, 3907–3914.

Takagaki, Y., und Manley, J.L. (2000). Complex Protein Interactions within the Human Polyadenylation Machinery Identify a Novel Component. *Mol. Cell. Biol.* *20*, 1515–1525.

Takagaki, Y., Ryner, L.C., und Manley, J.L. (1989). Four factors are required for 3'-end cleavage of pre-mRNAs. *Genes Dev.* *3*, 1711–1724.

Tang, C.F., und Shafer, R.H. (2006). Engineering the quadruplex fold: Nucleoside conformation determines both folding topology and molecularity in guanine quadruplexes. *J. Am. Chem. Soc.* *128*, 5966–5973.

The UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* *47*, D506–D515.

Tian, B., und Graber, J.H. (2012). Signals for pre-mRNA cleavage and polyadenylation. *Wiley Interdiscip. Rev. RNA* *3*, 385–396.

Treiber, T., Treiber, N., Plessmann, U., Harlander, S., Daiß, J.-L.L., Eichner, N., Lehmann, G., Schall, K., Urlaub, H., und Meister, G. (2017). A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Mol. Cell* *66*, 270-284.e13.

Tycowski, K.T., Shu, M.-D., Kukoyi, A., und Steitz, J.A. (2009). A Conserved WD40 Protein Binds the Cajal Body Localization Signal of scaRNP Particles. *Mol. Cell* *34*, 47–57.

Venkataraman, K. (2005). Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev.* *19*, 1315–1327.

Vilain, S., Esposito, G., Haddad, D., Schaap, O., Dobрева, M.P., Vos, M., van Meensel, S., Morais, V.A., de Strooper, B., und Verstreken, P. (2012). The yeast complex I equivalent NADH dehydrogenase rescues pink1 mutants. *PLoS Genet.*

Vorlíčková, M., Kejnovská, I., Sagi, J., Renčiuk, D., Bednářová, K., Motlová, J., und Kypr, J. (2012).

Circular dichroism and guanine quadruplexes. *Methods* 57, 64–75.

de Vries, H. (2000). Human pre-mRNA cleavage factor IIm contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J.* 19, 5895–5904.

Wahle, E. (1991). Purification and characterization of a mammalian polyadenylate polymerase involved in the 3' end processing of messenger RNA precursors. *J. Biol. Chem.*

Wahle, E., und Rügsegger, U. (1999). 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.*

Wallace, A.M., Dass, B., Ravnik, S.E., Tonk, V., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., und MacDonald, C.C. (1999). Two distinct forms of the 64,000 Mr protein of the cleavage stimulation factor are expressed in mouse male germ cells. *Proc. Natl. Acad. Sci. U. S. A.* 96, 6763–6768.

Wang, R., Zheng, D., Yehia, G., und Tian, B. (2018). A compendium of conserved cleavage and polyadenylation events in mammalian genes. *Genome Res.* 28, 1427–1441.

Wang, X., Goodrich, K.J., Gooding, A.R., Naeem, H., Archer, S., Paucek, R.D., Youmans, D.T., Cech, T.R., und Davidovich, C. (2017). Targeting of Polycomb Repressive Complex 2 to RNA by Short Repeats of Consecutive Guanines. *Mol. Cell* 65, 1056-1067.e5.

Weitzer, S., und Martinez, J. (2007). The human RNA kinase hClp1 is active on 3' transfer RNA exons and short interfering RNAs. *Nature* 447, 222–226.

Weitzer, S., Hanada, T., Penninger, J.M., und Martinez, J. (2015). CLP1 as a novel player in linking tRNA splicing to neurodegenerative disorders. *Wiley Interdiscip. Rev. RNA* 6, 47–63.

West, S., und Proudfoot, N.J. (2008). Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. *Nucleic Acids Res.* 36, 905–914.

Wittmann, S., Renner, M., Watts, B.R., Adams, O., Huseyin, M., Baejen, C., El Omari, K., Kilchert, C., Heo, D.-H., Kecman, T., u. a. (2017). The conserved protein Seb1 drives transcription termination by binding RNA polymerase II and nascent RNA. *Nat. Commun.* 8, 14861.

Xiang, K., Nagaike, T., Xiang, S., Kilic, T., Beh, M.M., Manley, J.L., und Tong, L. (2010). Crystal structure of the human symplekin–Ssu72–CTD phosphopeptide complex. *Nature* 467, 729–733.

Xiang, K., Tong, L., und Manley, J.L. (2014). Delineating the Structural Blueprint of the Pre-mRNA 3'-End Processing Machinery. *Mol. Cell. Biol.* 34, 1894–1910.

Yang, F., Hsu, P., Lee, S.D., Yang, W., Hoskinson, D., Xu, W., Moore, C., und Varani, G. (2017). The C terminus of Pcf11 forms a novel zinc-finger structure that plays an essential role in mRNA 3'-end processing. *RNA* 23, 98–107.

- Yang, Q., Gilmartin, G.M., und Doublie, S. (2010). Structural basis of UGUA recognition by the Nudix protein CFIm25 and implications for a regulatory role in mRNA 3' processing. *Proc. Natl. Acad. Sci.* *107*, 10062–10067.
- Yang, Q., Coseno, M., Gilmartin, G.M., und Doublie, S. (2011). Crystal Structure of a Human Cleavage Factor CFIm25/CFIm68/RNA Complex Provides an Insight into Poly(A) Site Recognition and RNA Looping. *Structure* *19*, 368–377.
- Yang, W., Hsu, P.L., Yang, F., Song, J.E., und Varani, G. (2018). Reconstitution of the CstF complex unveils a regulatory role for CstF-50 in recognition of 3-end processing signals. *Nucleic Acids Res.* *46*, 493–503.
- Zambelli, F., Pesole, G., und Pavesi, G. (2014). Using Weeder, Pscan, and PscanChIP for the Discovery of Enriched Transcription Factor Binding Site Motifs in Nucleotide Sequences. *Curr. Protoc. Bioinforma.* *47*, 2.11.1-2.11.31.
- Zarudnaya, M.I., Kolomiets, I.M., Potyahaylo, A.L., und Hovorun, D.M. (2003). Downstream elements of mammalian pre-mRNA polyadenylation signals: Primary, secondary and higher-order structures. *Nucleic Acids Res.* *31*, 1375–1386.
- Zhang, Z. (2005). CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes Dev.* *19*, 1572–1580.
- Zhang, Z., und Gilmour, D.S. (2006). Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol. Cell* *21*, 65–74.
- Zhang, Y., Sun, Y., Shi, Y., Walz, T., und Tong, L. (2020). Structural Insights into the Human Pre-mRNA 3'-End Processing Machinery. *Mol. Cell* *77*, 800-809.e6.
- Zhao, J., Hyman, L., und Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *TL - 63. Microbiol. Mol. Biol. Rev.*
- Zhu, Y., Wang, X., Forouzmand, E., Jeong, J., Qiao, F., Sowd, G.A., Engelman, A.N., Xie, X., Hertel, K.J., und Shi, Y. (2018). Molecular Mechanisms for CFIm-Mediated Regulation of mRNA Alternative Polyadenylation. *Mol. Cell*.
- Zou, X.-D., Hu, X.-J., Ma, J., Li, T., Ye, Z.-Q., und Wu, Y.-D. (2016). Genome-wide Analysis of WD40 Protein Family in Human. *Sci. Rep.* *6*, 39262.

7.2 Dissertationen und Abschlussarbeiten

- Dissertation A. Knoth (2008): Analysis of the mammalian pre-mRNA 3'-end processing cleavage complex and attempts of its reconstitution
- Bachelorarbeit L. Schönemann (2010): Arbeiten zur Rekonstitution des *cleavage and polyadenylation specificity factor* (CPSF)
- Diplomarbeit M. Lorbeer (2012): Darstellung und Charakterisierung der 3'-Prozessierungsfaktoren CF_m und CstF
- Masterarbeit L. Schönemann (2012): Rekombinante Expression des *cleavage and polyadenylation specificity factor* (CPSF) in Sf21-Insektenzellen.
- Bachelorarbeit M. Schmidt (2015): Bakterielle Expression der Carboxy-terminalen Domäne der RNA-Polymerase II und ihrer spezifischen Phosphatase Ssu72
- Dissertation L. Schönemann (2017): In-Vitro-Rekonstitution des 3'-Prozessierungsfaktors CPSF (*cleavage and polyadenylation specificity factor*)
- Dissertation C. Tüting (2020): Strukturelle und funktionelle Analyse des mRNA-3'-Prozessierungs-Komplex

7.3 Internetquellen

- [1] **ThermoFischer (2017):** Growth and maintenance of insect cell lines, Catalog numbers B82101, B82501 and B85502, Publication number MAN0000030, Revision B.0
https://tools.thermofisher.com/content/sfs/manuals/Insect_Cell_Lines_UG.pdf
 (Zugriff: 10.11.2020)
- [2] **BioRad (2011):** TC20 Automated Cell Counter, Instruction Manual., Catalog #145-0101
<http://www.bio-rad.com/webroot/web/pdf/lsr/literature/10024423.pdf>
 (Zugriff: 10.11.2020)
- [3] **ThermoFischer (2017):** Cellfectin II Reagent, Cellfectin II Transfection Protocol., Publication number MAN0007821, Revision 2.0
https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0007821_Cellfectin_II_Reagent_UG.pdf&title=VXNlc_iBHdWlkZTogQ2VsbGZlY3RpbjBJSSBSZWFnZW50
 (Zugriff: 10.11.2020)
- [4] **Qiagen (2003):** The QIAexpressionist™, A handbook for high-level expression and purification of 6xHis-tagged proteins, 5th Edition
<https://www.qiagen.com/de/resources/resourcedetail?id=79ca2f7d-42fe-4d62-8676-4cfa948c9435&lang=en>
 (Zugriff: 10.11.2020)
- [5] **IBA (2020):** Expression and purification of proteins using Strep-Tactin®, A comprehensive manual, Version PR02-0028
<https://www.iba-lifesciences.com/isotope/2/2-1201-002-Manual-Strep-Tactin-Purification.pdf>
 (Zugriff: 10.11.2020)
- [6] **GE Healthcare (2018):** HiLoad 16/600 Superdex 200 prep grade, Instruction 28992017 AD
<https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=15934>
 (Zugriff: 10.11.2020)
- [7] **GE Healthcare (2018):** Instruction 29115046 AD, Superose 6 Increase 5/150 GL, and Superose 6 Increase 10/300 GL
<https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=17307>
 (Zugriff: 10.11.2020)
- [8] **GE Healthcare (2005):** Instructions 52-1782-00 AF, Gel filtration Media, Superose 6 prep grade and Superose 12 prep grade
<https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=11544>
 (Zugriff: 10.11.2020)

- [9] **GE Healthcare (2011):** *Gel Filtration Calibration Kits, 28-4038-41*
<https://www.cytivalifesciences.co.kr/wp-content/uploads/2020/04/Gel-Filtration-Calibration-Kit-LMW-HMW.pdf>
(Zugriff: 10.11.2020)
- [10] **UniProtKB (2019):** *The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 47: D506-515 (2019)*
<https://www.uniprot.org/>
(Zugriff: 10.11.202)
- [11] **GraphRobot (2019):** *Lei A. Wang (2019). GraphRobot*
<https://www.graphrobot.com>
(Zugriff: 10.11.202)
- [12] **JalView:** *Waterhouse, A.M., Procter, J.B., Martin, D.M.A, Clamp, M. and Barton, G. J. (2009) "Jalview Version 2 - a multiple sequence alignment editor and analysis workbench" Bioinformatics 25 (9) 1189-1191 doi: 10.1093/bioinformatics/btp033*
<https://www.jalview.org/>
(Zugriff: 10.11.202)
- [13] **Inkscape (2017):** <https://inkscape.org/>
(Zugriff: 10.11.202)
- [14] **GE Healthcare (2011):** *Protein A Sepharose™ CL-4B, Instructions 71-7090-00 AF*
<https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=14030>
(Zugriff: 13.11.2020)

8 Tabellen- und Abbildungsverzeichnis

8.1 Abbildungsverzeichnis

Abbildung 1.1 Synthese und Prozessierung der mRNA.....	2
Abbildung 1.2 Schematische Darstellung bekannter 3'-UTR Sequenzmotive in prä-mRNAs.....	4
Abbildung 1.3 Schematische Darstellung bekannter 3'-Prozessierungsfaktoren während der Transkription.....	5
Abbildung 1.4 Domänen und Interaktionen des humanen Clp1.....	12
Abbildung 1.5 Domänenstruktur mit Angabe der Aminosäurepositionen des Bäckerhefe-Pcf11 im Vergleich mit dem humanen Pcf11.....	13
Abbildung 2.1 Expression und Filterbindungsexperimente von CPSF-Variationen modifiziert aus Schönemann u.a. 2014.....	18
Abbildung 2.2 mPSF-EMSA mit Flag-Elutionsfraktion E2.....	19
Abbildung 2.3 Crosslinking von W10-RNA an mPSF.....	21
Abbildung 2.4 Schematische Darstellung der Domänen und strukturierten Bereiche des humanen WDR33 mit Angaben der Aminosäurepositionen.....	22
Abbildung 2.5 Reinigung von rekombinantem CF II.....	25
Abbildung 2.6 Reinigung von rekombinantem hClp1.....	27
Abbildung 2.7 Dichtegradientenzentrifugation der CF II-MonoQ-Fraktion #19.....	28
Abbildung 2.8 HeLa NXT-Fraktionierung und anschließender L3-RNA 3'-Spaltassay mit CF II-Komplementation.....	30
Abbildung 2.9 3'-Spaltassays mit verkürzten hPcf11-Varianten im CF II-Komplex.....	32
Abbildung 2.10 Aktivität und Reinigung Kinase-inaktiver Mutanten des hClp1 im CF II-Komplex.....	34
Abbildung 2.11 3'-Spaltassay durch Komplementation eines fraktionierten nukleären Extraktes mit CF II und hClp1-Mutanten in CF II.....	35
Abbildung 2.12 5'-PNK-Kinetiken von CF II und hClp1.....	38
Abbildung 2.13 Bindungsversuche von CF II an verschiedene Substrate.....	41
Abbildung 2.14 UV-Vernetzung und Antikörper-Affinitätsreinigung von CF II mit 5'-markierter rU ₁₄ -RNA.....	44
Abbildung 2.15 Filterbindungsversuche von CF II mit SV40 late-RNA und Übersicht der SV40 late RNA-Sequenz.....	45
Abbildung 2.16 CD-Spektren der 17mer Oligonukleotide und Gelshifts.....	48
Abbildung 2.17 Bind'n'Seq (BS) Selektionsversuche und Ermittlung einer Konsensussequenz der CF II-RNA-Bindung.....	51
Abbildung 2.18 hPcf11 bindet RNA über die Zinkfinger Domänen.....	54

Abbildung 2.19 Reinigung und RNA Bindung von CstF und CF I.....	59
Abbildung 3.1 Schematische Darstellung der Domänen und Bereiche des hPcf11 und der Aktivität im CF II.....	69
Abbildung 4.1 Aktualisierte schematische Darstellung des 3'-Prozessierungskomplexes auf der prä-mRNA.....	73
Abbildung 6.1 Schematische Darstellung der MultiBac-Klonierungsvektoren.	92
Abbildung 9.1 Vergleichende Gegenüberstellung der SV40 late-Sequenz und verwendeten Fragmenten.	VII
Abbildung 9.2 Native Gelshifts von CF II (wt) und SV40 late-RNA Fragmenten.....	VIII
Abbildung 9.3 Vergleichender Gegenüberstellung der L3-Sequenz und der L3 prä-Sequenz.....	IX
Abbildung 9.4 Schematische Übersichtsdarstellung der Proteindomänen von bekannten 3'-Prozessierungsfaktoren.	XI
Abbildung 9.5 Alignment verschiedener homologer Clp1-Varianten aus ausgesuchten Spezies.	XIII
Abbildung 9.6 Alignment verschiedener homologer Pcf11-Varianten aus ausgesuchten Spezies.	XIX
Abbildung 9.7 Alignment der FEGP-Wiederholungen aus hPcf11.	XIX
Abbildung 9.8 Übersichtsdarstellung der CF II-Crosslinks.	XX
Abbildung 9.9 Vergleich der maximalen Reaktionsgeschwindigkeiten von CF II Wt und Mutanten.....	XXI
Abbildung 9.10 RNA-Bindung von CF II mit Mutationen in hClp1.	XXII
Abbildung 9.11 Größenausschlusschromatographie von mPSF.	XXIII
Abbildung 9.12 : Spaltungsaktivität nach Fraktionierung mittels Größenausschlusschromatographie.	XXIV
Abbildung 9.13 : Längenkontrolle in der Polyadenylierung mit mPSF.....	XXV
Abbildung 9.14 BSA-Größenausschlusschromatographie mit angeschlossener Rechtwinkel- Lichtstreuung.	XXVI
Abbildung 9.15 CstF-Größenausschlusschromatographie mit angeschlossener Rechtwinkel- Lichtstreuung.	XXVII
Abbildung 9.16 CF I-Größenausschlusschromatographie mit angeschlossener Rechtwinkel- Lichtstreuung.	XXVIII

8.2 Tabellenverzeichnis

Tabelle 1.1 Vergleichende Gegenüberstellung von Proteinen der 3'-Prozessierung aus Menschen und Bäckerhefe.....	6
Tabelle 2.1 Übersicht der Domänen und strukturierten Bereiche des humanen WDR33.....	23
Tabelle 2.2 Michaelis-Menten-parameter der 5'-PNK Aktivität von CF II und hClp1.....	39
Tabelle 2.3 Apparente Dissoziationskonstanten aus den nativen Gelshifts der Homopolymer-Reihen mit CF II.....	43
Tabelle 2.4 Apparente Dissoziationskonstanten aus nativen Gelshifts von CF II an SV40 late-Sequenzabschnitte.....	46
Tabelle 2.5 Apparente Dissoziationskonstanten aus nativen Gelshifts von CF II an die 17-mer Varianten SV40 late 208–224 Sequenz.....	49
Tabelle 2.6 Apparente Dissoziationskonstanten aus nativen Gelshifts von CF II an selektierte Sequenzen aus der Anreicherung zufälliger RNA-Populationen (Bind'n'Seq).....	52
Tabelle 2.7 Apparente Dissoziationskonstanten aus nativen Gelshifts mit verkürzten hPcf11-Varianten mit SV40 late 190–225-RNA als Bindungspartner.....	55
Tabelle 9.1 Parameter der Substratüberschusshemmung der CF II 5'-PNK-Reaktion.....	XXII

8.3 Abkürzungsverzeichnis

Abkürzung	Langschrift
A260	Absorption bei 260 nm
A280	Absorption bei 280 nm
Abb.	Abbildung
Δ	gr. Delta - allgemeine Bezeichnung für (Deletions-) Mutanten
EMSA	electrophoretic mobility shift assay
kDa	Kilo-Dalton
l	Liter
M	Molar (Mol/l)
MDa	Mega-Dalton
min	Minuten
nt	Nukleotide
p.a.	por analysi
U	<i>units</i> (Allgemeine Einheit für enzymatische Aktivität)
UV-Licht	ultraviolettes Licht
wt	Wildtyp
xg	Einheit für Schwerebeschleunigung

(v/v)	Volumen pro Volumen
(w/v)	Masse pro Volumen
A	Adenin
ATP	Adenosintriphosphat
C	Cytosin
CTP	Cytidintriphosphat
G	Guanin
GTP	Guanosintriphosphat
T	Thymin
TTP	Thymidintriphosphat
U	Uracil
UTP	Uridintriphosphat
CF I	<i>cleavage factor I</i> (aus <i>Homo sapiens</i>)
CF IA	<i>cleavage factor IA</i> (aus <i>Saccharomyces cerevisiae</i>)
CF IB	<i>cleavage factor IB</i> (aus <i>Saccharomyces cerevisiae</i>)
CF II	<i>cleavage factor II</i> (aus <i>Homo sapiens</i> und <i>Saccharomyces cerevisiae</i>)
CFP	<i>cyan fluorescent protein</i>
CPF	<i>cleavage and polyadenylation factor</i> (aus <i>Saccharomyces cerevisiae</i>)
CPSF	<i>cleavage and polyadenylation specificity factor</i> (aus <i>Homo sapiens</i>)
CstF	<i>cleavage stimulatory factor</i> (aus <i>Homo sapiens</i>)
PAP	Poly(A)-Polymerase
PABPN1	Nukleäres Poly(A)-bindendes Protein 1
PP1 α	α -Isoform der katalytischen Untereinheit der Proteinphosphatase 1
Rbbp6	<i>Retinoblastoma-binding protein 6</i>
WDR33	<i>WD-repeat protein 33</i>
YFP	<i>yellow fluorescent protein</i>
5'-SP	5'-Spaltprodukt
APS	Ammoniumperoxodisulfat
C-Terminus	Carboxy-Terminus (von Proteinen)
CTD	Carboxy-terminale Domäne
dd. H ₂ O	bidestilliertes Wasser
DEAE	Diethylaminoethyl (-Zellulose)
DNA	Desoxyribonukleinsäure
DSE	downstream element
<i>E. coli</i>	<i>Escherichia coli</i>
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure

HCl	Salzsäure
IPTG	Isopropyl- β -D-thiogalactopyranosid
LB	<i>lysogeny broth</i> (-Medium)
m ⁷ GpppN	7-Methylguanosin- <i>cap</i>
M β L (-Familie)	Metallo- β -Laktamase (-Familie)
MOI	multiplicity of infection
mRNA	messenger RNA
mRNP	<i>messenger</i> Ribonukleoproteinpartikel
N-Terminus	Amino-Terminus (von Proteinen)
PMSF	Phenylmethylsulfonylfluorid
Prä-mRNA	mRNA-Vorläufer
(r)ATP	(Ribose-)ATP
RNA	Ribonukleinsäure
RRM	RNA recognition motif
SDS	Natriumdodecylsulfat
(SDS-) PA-Gel	(SDS-) Polyacrylamid-Gel
(SDS-) PAGE	(SDS-) Polyacrylamid-Gelelektrophorese
snRNA	<i>small nuclear ribonucleic acid</i>
snRNP	<i>small nuclear ribonucleoprotein particle</i>
scaRNA	<i>Small Cajal body specific ribonucleic acid</i>
SOC	<i>super optimal broth</i> -Medium mit 20 mM Glukose
Stabw.	Standardabweichung
Tab.	Tabelle
Tris (-Base)	Trishydroxymethylaminomethan
Tris-HCl	Tris-Lösung, pH mit HCl eingestellt
TY	tryptone yeast (-Medium)
u.a.	und andere
USE	upstream auxiliary element
UTR	nicht-translatierte Region

9 Anhang

9.1 Zusätzliche Abbildungen

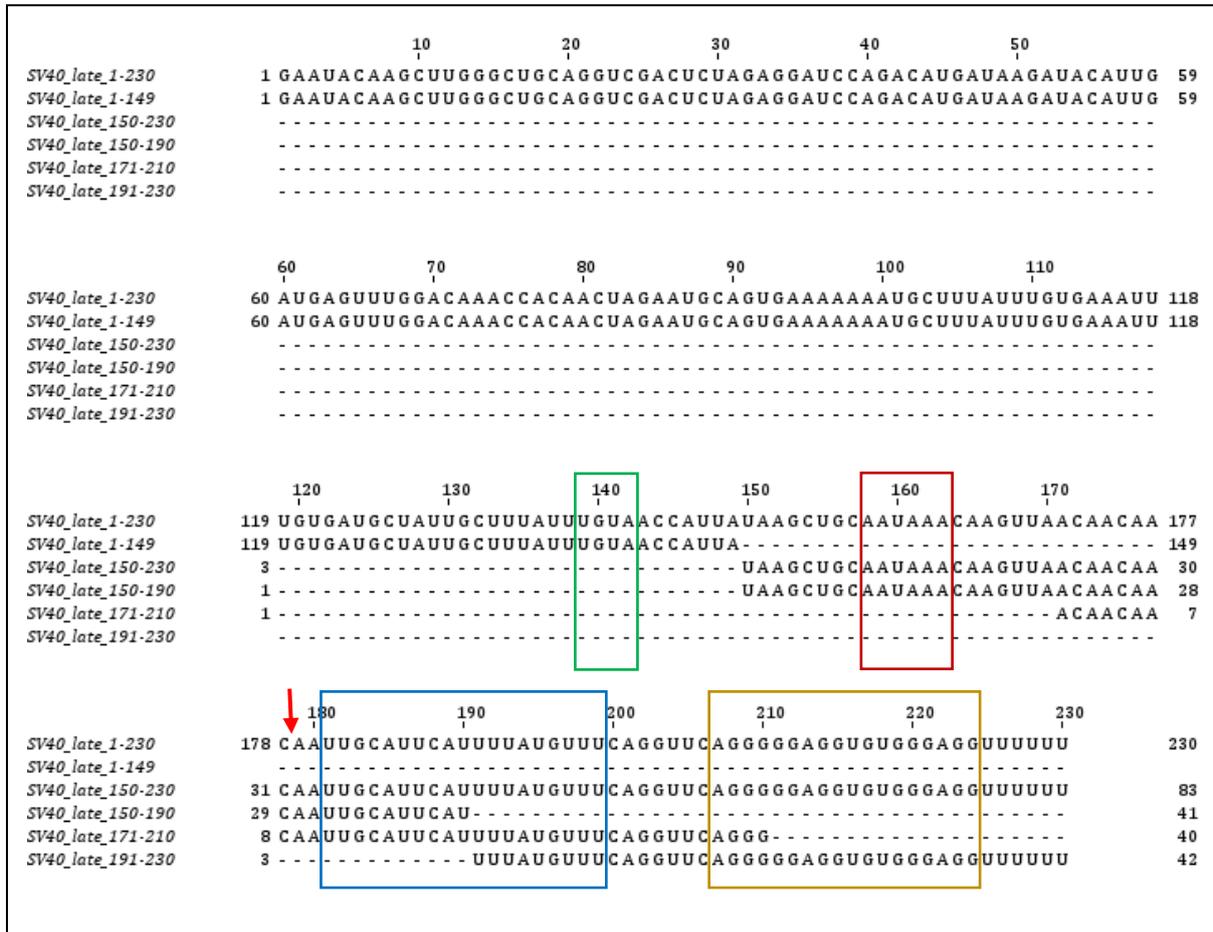


Abbildung 9.1 Vergleichende Gegenüberstellung der SV40 late-Sequenz und verwendeten Fragmenten. Fragmentbezeichnungen entsprechen der Basenposition der ersten und letzten Base der RNA-Sequenz mit Referenz zur vollständigen SV40 late-Sequenz. Grün umrandet ist das CF I-Bindungsmotiv „UGUA“. Rot umrandet ist das kanonische Polyadenylierungssignal „AAUAAA“ (PAS). Die Δ -Mutante würde an dieser Stelle die Sequenz „AAGAAA“ enthalten. Der Rote Pfeil markierte die 3'-Spaltstelle. In Blau umrandet ist die vermutete U/GU-reiche Bindungsstelle des CstF-Komplexes (DSE). Orange umrandet ist die ermittelte G/GU-reiche Bindestelle des CF II-Komplexes. Erstellt mit JalView.

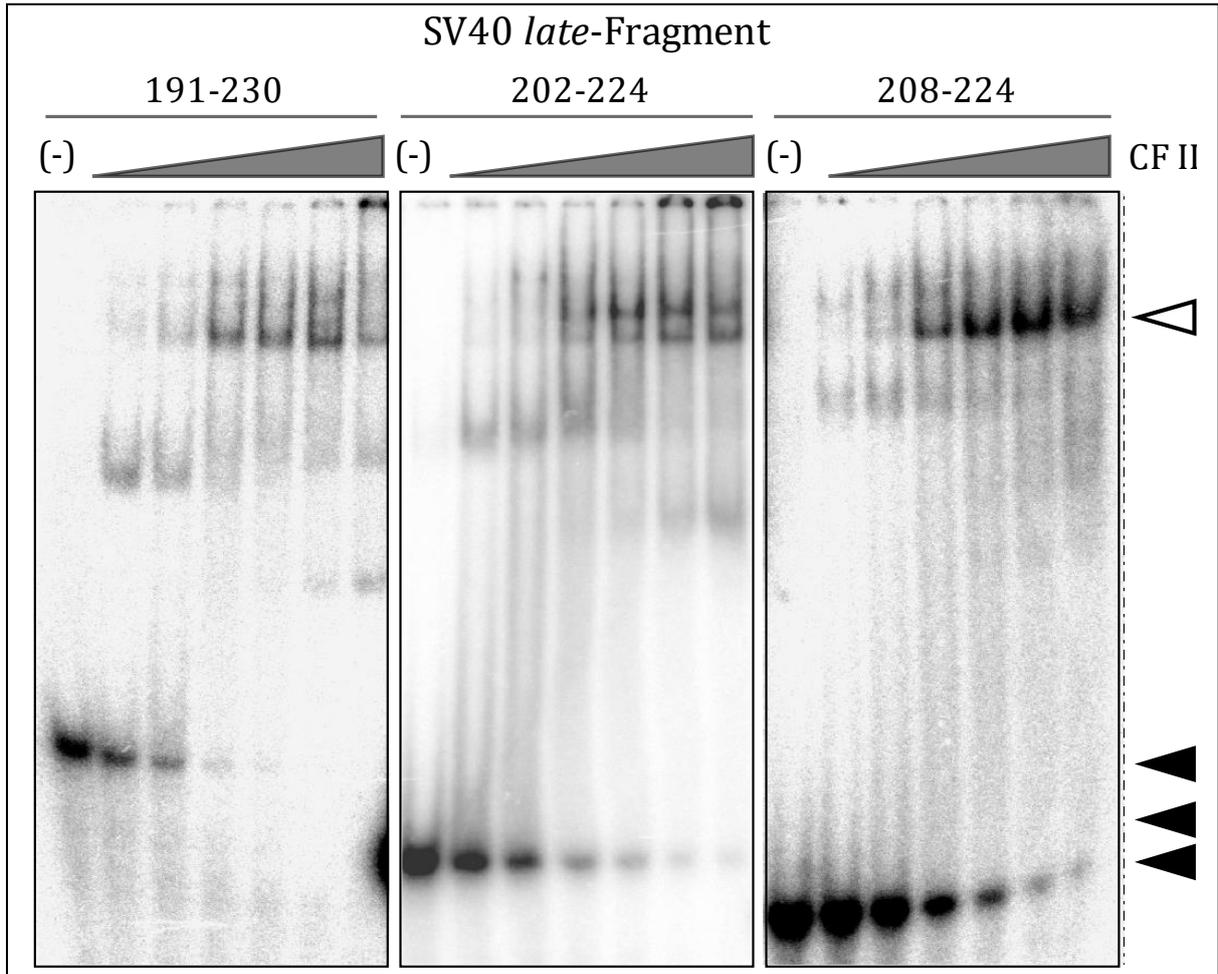


Abbildung 9.2 Native Gelshifts von CF II (wt) und SV40 *late*-RNA Fragmenten.

Nativer Gelshift in einem 5%igen PA-Gel nach einer Bindungsreaktion von CF II an die angegebenen radioaktiv markierten RNA-Fragmente der SV40 *late*-RNA. In der Bindungsreaktion SV40 *late* 190-225 waren 0,1 nM RNA und steigende Konzentrationen CF II zwischen 0 nM und 10 nM. In den anderen beiden Bindungsreaktionen waren 0,5 nM RNA und 0 nM bis 50 nM CF II-Komplex. Der weiße Pfeilkopf markiert den RNA-Proteinkomplex, die schwarzen die jeweilige freie RNA.

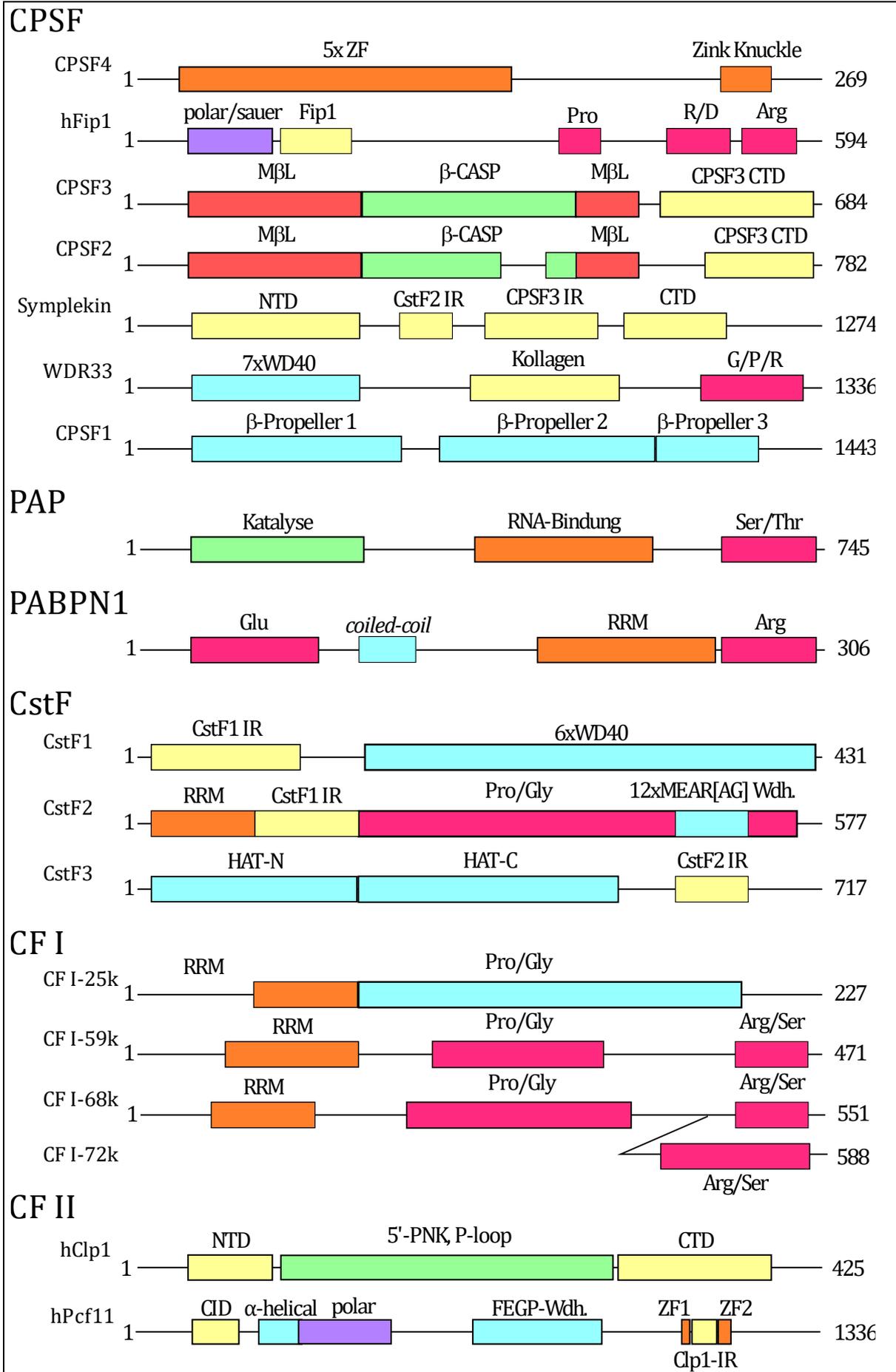


Abbildung 9.4 Schematische Übersichtsdarstellung der Proteindomänen von bekannten 3'-Prozessierungsfaktoren.

Orientierungshilfe für die Proteinstrukturen, die in der Einleitung beschrieben werden. Nicht-skalierte Darstellung der Proteine mit Angabe der ersten und letzten Aminosäureposition sowie schematische dargestellte Größe der Domänen und Bereiche. Strukturierte und benannte Domänen und Bereiche sind in Blau eingefärbt. Interaktionsregionen mit anderen Prozessierungsfaktoren und strukturierte Bereiche sind Gelb eingefärbt. RNA-bindende Domänen und Bereiche sind in Orange eingefärbt. Proteinbereiche mit einer Anreicherung an bestimmten Aminosäuren sind in Pink eingefärbt und Bereiche mit einer Anreicherung an geladenen Aminosäuren in Lila. Katalytische Domänen und Bereiche sind in Grün eingefärbt.

Anhang

		10	20	30	40	50	60																																																														
<i>H_sapiens</i>	1	-----MGE	EA	ND	DK-----	-----KPT	TK	FE	LE	RE	TE	LR	FE	VE	AS	Q-	SV	QL	EL	LL	T	MA	EI	44																																													
<i>B_taurus</i>	1	-----MGE	EA	ND	DK-----	-----KPT	TK	FE	LE	RE	TE	LR	FE	VE	AS	Q-	SV	QL	EL	LL	T	MA	EI	44																																													
<i>M_musculus</i>	1	-----MSE	E	SN	DD	DK-----	-----KPT	TK	FE	LE	RE	TE	LR	FE	VE	AS	Q-	SV	QL	EL	LL	T	MA	EI	44																																												
<i>G_gallus</i>	1	-----MAD	D	GG	DE	CK-----	-----KQ	VAK	FE	LE	RE	TE	LR	FE	VE	AS	Q-	TV	Q	ME	LL	T	MA	EI	44																																												
<i>D_rerio</i>	1	-----MTA	E	AA	EK	SV	EE	GL	SS	SS	GS	A-	GS	SG	TR	FD	L	DK	ET	LR	FE	VE	AS	Q-	RV	QL	EL	LL	T	MA	EI	56																																					
<i>X_tropicalis</i>	1	-----MS	E	DL	SAP	-----	GPS	AP	DS	SS	FT	GP	ST	K	FE	LE	RE	TE	LR	LE	VE	GT	D-	PV	R	VE	LV	S	GL	AE	V	54																																					
<i>D_melanogaster</i>	1	-----MSE	D	-----	-----	-----	Q	G	K	D	Y	T	LE	SD	SE	LR	FE	IE	Q	K	DA	K	V	L	V	LV	S	LV	S	GL	AE	V	39																																				
<i>C_elegans</i>	1	-----MSE	E	-----	-----	-----	N	V	Q	E	F	V	L	KE	D	CE	LR	FA	AG	DD	S-	DV	CE	LE	L	V	K	Y	AE	I	38																																						
<i>S_cerevisiae</i>	1	MA	S	L	P	G	I	D	E	H	T	T	S	E	L	I	T	G	D	-----	-----	-----	NE	W	H	K	L	V	I	PK	G	S	D	W	Q	I	D	L	K	A	E	G	-	KL	I	V	K	V	N	S	G	I	V	E	I	55													
		70	80	90	100	110	120	130																																																													
<i>H_sapiens</i>	45	FG	TE	L	TR	NK	K	F	T	F	D	A	G	A	K	V	A	V	F	T	-----	WH	G	C	S	V	Q	L	S	R	T	E	V	A	Y	V	SK	D	T	P	M	L	L	Y	L	T	H	T	A	L	E	Q	M	R	R	106													
<i>B_taurus</i>	45	FG	TE	L	TR	NK	K	F	T	F	D	A	G	A	K	V	A	V	F	T	-----	WH	G	C	S	L	Q	L	S	R	T	E	V	A	Y	V	SK	D	T	P	M	L	L	Y	L	T	H	T	A	L	E	Q	M	R	R	106													
<i>M_musculus</i>	45	FG	TE	L	TR	NK	K	F	T	F	D	A	G	A	K	V	A	V	F	T	-----	WH	G	C	S	L	Q	L	S	R	T	E	V	A	Y	V	SK	D	T	P	M	L	L	Y	L	T	H	T	A	L	E	Q	M	R	R	106													
<i>G_gallus</i>	45	FG	TE	L	TR	NK	K	F	T	F	D	A	G	A	K	V	A	V	F	T	-----	WH	G	C	T	V	Q	L	S	R	T	E	V	A	Y	V	SK	D	T	P	M	L	L	Y	L	T	H	T	A	L	E	Q	M	R	R	106													
<i>D_rerio</i>	57	FG	SE	L	NR	N	K	K	Y	T	F	G	P	G	S	K	I	A	V	F	T	-----	W	Q	C	G	V	A	L	S	G	K	T	E	V	A	Y	V	SK	D	T	P	M	L	L	Y	L	T	H	A	A	L	E	Q	M	R	R	118											
<i>X_tropicalis</i>	55	FG	TE	L	TR	NK	K	F	T	F	P	P	G	S	R	A	A	I	F	T	-----	WH	G	C	T	V	Q	L	W	G	S	P	D	V	A	Y	V	SR	D	T	P	M	L	L	Y	L	T	H	T	V	G	L	E	Q	M	R	R	116											
<i>D_melanogaster</i>	40	FG	TE	L	V	K	K	K	Y	E	F	G	V	G	A	K	V	A	I	F	T	-----	Y	Q	G	C	V	L	H	V	S	G	K	M	D	V	C	I	S	K	E	T	P	M	V	Q	Y	V	C	H	A	A	L	E	Q	F	R	M	E	101									
<i>C_elegans</i>	39	FG	TE	L	L	L	N	K	K	Y	T	F	P	A	K	S	R	V	A	A	F	T	-----	W	K	G	A	T	I	E	L	V	G	T	T	E	S	A	Y	V	A	E	S	T	P	M	V	I	L	I	H	A	A	M	E	V	R	K	100										
<i>S_cerevisiae</i>	56	FG	TE	L	A	V	D	D	E	Y	T	F	-	Q	N	W	K	F	P	I	Y	A	V	E	E	T	E	L	L	W	K	C	P	D	L	-----	T	T	N	T	I	T	V	K	P	N	H	T	M	K	Y	I	Y	L	H	F	M	L	E	K	I	S	M	119					
		140	150	160	170	180	190	200																																																													
<i>H_sapiens</i>	107	AE	KE	-----	E	E	R	G	P	R	V	M	V	G	P	T	D	V	G	K	S	T	V	C	R	L	L	L	N	Y	A	V	R	L	-	G	R	R	P	T	Y	V	E	L	D	V	G	Q	S	V	S	I	P	G	T	M	G	A	L	Y	I	168							
<i>B_taurus</i>	107	AE	KE	-----	E	E	R	G	P	R	V	M	V	G	P	T	D	V	G	K	S	T	V	C	R	L	L	L	N	Y	A	V	R	L	-	G	R	R	P	T	Y	V	E	L	D	V	G	Q	S	V	S	I	P	G	T	M	G	A	L	Y	I	168							
<i>M_musculus</i>	107	AE	KE	-----	E	E	R	G	P	R	V	M	V	G	P	T	D	V	G	K	S	T	V	C	R	L	L	L	N	Y	A	V	R	L	-	G	R	R	P	T	Y	V	E	L	D	V	G	Q	S	V	S	I	P	G	T	M	G	A	L	Y	I	168							
<i>G_gallus</i>	107	A	E	R	E	-----	D	E	R	G	P	R	V	M	V	G	P	T	D	V	G	K	S	T	V	C	R	L	L	N	Y	A	V	R	L	-	G	R	R	P	T	F	V	E	L	D	V	G	Q	S	V	S	I	P	G	T	M	G	A	L	Y	I	168						
<i>D_rerio</i>	119	A	E	K	D	-----	N	E	R	G	P	R	V	M	V	G	P	T	D	V	G	K	S	T	V	C	R	L	L	N	Y	A	V	R	L	-	G	R	R	P	T	L	V	E	L	D	V	G	Q	S	V	S	V	P	G	T	M	S	A	L	C	I	180						
<i>X_tropicalis</i>	117	A	E	R	E	-----	G	E	R	G	P	R	V	L	V	A	G	P	S	D	V	G	K	S	T	L	C	R	L	L	N	Y	A	V	R	L	-	G	R	R	P	T	L	V	E	L	D	V	G	Q	S	V	S	V	P	G	T	V	G	A	L	C	V	178					
<i>D_melanogaster</i>	102	A	E	E	K	-----	D	R	Y	G	P	V	A	M	V	G	P	M	D	V	G	K	S	T	L	C	R	I	L	L	N	Y	A	V	R	L	-	G	R	R	P	L	Y	A	D	L	D	V	G	Q	S	I	A	I	S	G	S	V	A	T	I	L	I	163					
<i>C_elegans</i>	101	R	E	E	Q	A	A	G	N	S	N	K	A	K	G	P	R	L	L	V	G	P	T	D	V	G	K	T	T	V	S	R	I	L	C	N	Y	S	V	R	Q	-	G	R	T	P	I	F	V	E	L	D	V	G	Q	N	S	V	S	V	P	G	T	V	A	A	V	L	168
<i>S_cerevisiae</i>	120	-----	N	F	E	G	P	R	V	V	I	V	G	G	S	Q	T	G	K	T	S	L	S	R	T	L	C	S	Y	A	L	K	F	N	A	Y	Q	P	L	Y	I	N	L	D	P	Q	Q	P	I	F	T	V	P	G	C	I	S	A	T	P	I	178							
		210	220	230	240	250	260	270																																																													
<i>H_sapiens</i>	169	ER	P	A	D	V	E	E	G	-----	F	S	I	Q	A	P	L	V	Y	H	F	G	S	T	T	P	G	T	N	I	K	L	Y	N	K	I	T	S	R	L	A	D	V	F	N	Q	R	C	E	V	N	R	R	A	S	222													
<i>B_taurus</i>	169	ER	P	A	D	V	E	E	G	-----	F	S	I	Q	A	P	L	V	Y	H	F	G	S	T	T	P	G	T	N	I	K	L	Y	N	K	I	T	S	R	L	A	D	V	F	N	Q	R	C	E	V	N	R	R	A	S	222													
<i>M_musculus</i>	169	ER	P	A	D	V	E	E	G	-----	F	S	I	Q	A	P	L	V	Y	H	F	G	S	T	T	P	G	T	N	I	K	L	Y	N	K	I	T	S	R	L	A	D	V	F	N	Q	R	C	E	V	N	R	R	A	S	222													
<i>G_gallus</i>	169	ER	P	A	D	V	E	E	G	-----	F	S	L	Q	A	P	L	V	Y	H	F	G	S	T	T	P	G	T	N	I	K	L	Y	N	K	I	T	S	R	L	A	D	V	F	N	Q	R	C	E	V	N	R	R	A	S	222													
<i>D_rerio</i>	181	ER	P	A	D	V	E	E	G	-----	F	S	V	Q	A	P	L	V	F	H	F	G	S	T	T	P	G	T	N	I	K	L	Y	N	K	L	T	S	L	A	D	A	F	S	R	C	E	V	N	R	R	A	S	234															
<i>X_tropicalis</i>	179	ER	P	A	D	V	E	E	G	-----	F	S	A	Q	A	P	L	V	Y	H	F	G	S	T	T	P	G	T	N	I	K	L	Y	N	K	L	T	S	R	L	A	H	V	F	N	L	R	C	D	S	N	R	R	A	S	232													
<i>D_melanogaster</i>	164	ER	P	A	N	V	E	E	G	-----	F	A	K	T	A	P	L	V	Y	H	F	G	H	K	S	P	S	G	N	S	V	L	Y	N	A	V	V	S	K	M	A	E	V	T	L	Q	S	L	N	S	N	K	R	T	K	217													
<i>C_elegans</i>	169	Q	K	T	A	D	V	I	D	-----	F	E	R	N	Q	P	I	V	F																																																		

	490	500	
<i>H_sapiens</i>	413 FLLI - MDI RFMDLK - - -		425
<i>B_taurus</i>	413 FLLI - MDI RFMDLK - - -		425
<i>M_musculus</i>	413 FLLI - MDI RFMDLK - - -		425
<i>G_gallus</i>	413 FLLI - MDI RFMDLK - - -		425
<i>D_reio</i>	431 TLLI - MDI RFI DLK - - -		443
<i>X_tropicalis</i>	425 VLLI - MDI RFMDLK - - -		437
<i>D_melanogaster</i>	409 ALLLWSELQFMDNHT - -		423
<i>C_elegans</i>	413 VLVF - SDI THLDDQIKR		428
<i>S_cerevisiae</i>	435 AMIL - TSYRYLE - - - -		445

Abbildung 9.5 Alignment verschiedener homologer Clp1-Varianten aus ausgesuchten Spezies. Eingefärbte Aminosäuren entsprechen einer Konservierung zwischen 50% (schwach) und 100% (stark) zwischen den untersuchten Sequenzen. Die roten Pfeile markieren die in dieser Arbeit mutierten Aminosäuren K127, D151, R288 und R293. Erstellt mit JalView aus Sequenzen der UniProt KB Datenbank und aligned mit dem Mafft-Server.

Anhang

	490	500	510	520	530	540	550											
<i>H. sapiens</i>	390 E SMRL	SDMNKRDPRLK	KKHLQDKTDG	KDDDDVKEKRR	TKAEKKDK	DEHMKSSSEH	RLA	GSRNKIINGIVQKQ 457										
<i>B. taurus</i>	391 E STRV	SDMSKRDPRLK	KKHLQDKTDS	KDDDTKDKRKT	AEKKDK	DEHMKSSSEH	RLV	GSRNKIINGIVQKQ 458										
<i>M. musculus</i>	391 E SMRL	SDMSKRDPRLK	KKHLQDKAE	GKDEDEK	KEKRR	TKAEKKDK	DEHMKSSSEH	RVI	GSRSKIINGIVQKQ 458									
<i>G. gallus</i>	385 E STKV	SEM SKRDPRLK	KKHLQEKPE	SKGDEVKE	KRRNTE	KKKDEH	KTGEH	RPV	GSRNKTVNGVQKQ 451									
<i>D. rerio</i>	346 E QSK	SAEVNKKDP	RLHKKMHER	MD SKDEDV	REKRRS	SEKKDR	DEGSK	NLDH	QKLS	SNRGLK	INGSVNKH 414							
<i>X. tropicalis</i>	370 ENVR	DLDIRR	DPRLHKKHL	RNKSDD	KDAEP	KEKRR	CDKKE	EH	KSEH	RVA	GNRNK	PVNGIVPKN 432						
<i>D. melanogaster</i>	379 SSTT	SSSSSDVR	--HKG	VTA	SSSNQ	PTV	KRS	GKIT	SKD	SDRY	V	RNGSP	--LG	SAKRK	SSSPSS	SPS 442		
<i>C. elegans</i>	274 EDLR	SLTK	KPPV	---	---	---	---	---	---	---	---	---	---	---	---	---	307	
<i>S. cerevisiae</i>	233	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	272	
	560	570	580	590	600	610	620											
<i>H. sapiens</i>	458 DTI	TEESEK	QGT	KPGRS	SSTRKR	SR	SR	SP	---	---	---	---	---	---	---	---	513	
<i>B. taurus</i>	459 DTI	TEESEK	P	G	T	K	P	G	R	S	S	T	R	K	R	S	R	SP
<i>M. musculus</i>	459 DMVT	EESEK	QGT	KPGRS	SSTRKR	SR	SR	SP	---	---	---	---	---	---	---	---	514	
<i>G. gallus</i>	452 DV	TEESEK	Q	G	K	Q	G	R	S	S	N	R	K	R	S	R	S	SP
<i>D. rerio</i>	415 EK	LE	T	F	L	K	Q	E	---	---	---	---	---	---	---	---	465	
<i>X. tropicalis</i>	433 ET	ST	N	P	D	K	---	---	---	---	---	---	---	---	---	---	484	
<i>D. melanogaster</i>	443 K	S	R	S	T	S	H	K	S	S	S	S	R	G	T	S	S	L
<i>C. elegans</i>	308	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	334	
<i>S. cerevisiae</i>	273	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	274	
	630	640	650	660	670	680												
<i>H. sapiens</i>	514 STP	KAGKI	RQ	SGAK	QS	---	---	---	---	---	---	---	---	---	---	---	566	
<i>B. taurus</i>	515 STP	KAGKI	RQ	SGV	KQS	---	---	---	---	---	---	---	---	---	---	---	566	
<i>M. musculus</i>	515 LAP	KAGKMR	Q	SGL	KQS	---	---	---	---	---	---	---	---	---	---	---	567	
<i>G. gallus</i>	509 STP	KIGKI	RQ	IGPK	QS	---	---	---	---	---	---	---	---	---	---	---	562	
<i>D. rerio</i>	466	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	510	
<i>X. tropicalis</i>	485 SL	PKI	GRL	R	Q	P	G	S	K	S	---	---	---	---	---	---	536	
<i>D. melanogaster</i>	494 SL	LAGA	SLP	ASA	KAP	KDLE	K	P	T	F	Q	I	L	E	P	H	550	
<i>C. elegans</i>	335 PAP	PL	S	S	A	P	P	---	---	---	---	---	---	---	---	---	359	
<i>S. cerevisiae</i>	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
	700	710	720	730	740	750												
<i>H. sapiens</i>	567 QET	---	TN	Q	S	T	K	S	G	T	E	P	K	E	N	V	W	Q
<i>B. taurus</i>	567 QET	---	AG	Q	S	T	K	S	S	T	E	P	K	E	N	V	W	S
<i>M. musculus</i>	568 QET	---	AG	Q	S	M	K	S	G	G	D	P	K	E	N	I	W	Q
<i>G. gallus</i>	563 QEI	---	AS	Q	S	T	K	A	S	P	D	P	K	E	N	A	E	N
<i>D. rerio</i>	511 AE	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	569
<i>X. tropicalis</i>	537 QDP	---	V	S	L	I	S	N	K	S	N	T	E	P	K	E	D	I
<i>D. melanogaster</i>	551 PPT	F	S	V	D	S	L	D	K	Q	D	T	T	S	G	H	I	S
<i>C. elegans</i>	360	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	378
<i>S. cerevisiae</i>	275	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	283
	760	770	780	790	800	810	820											
<i>H. sapiens</i>	630 GIL	SPRAPK	Q	Q	H	R	L	S	V	---	---	---	---	---	---	---	688	
<i>B. taurus</i>	630 GIL	SPRAPK	Q	Q	H	R	L	S	V	---	---	---	---	---	---	---	687	
<i>M. musculus</i>	631 GIL	SPRAPK	Q	Q	H	R	L	S	V	---	---	---	---	---	---	---	688	
<i>G. gallus</i>	629 GIL	SPRAPK	Q	Q	H	R	L	S	V	---	---	---	---	---	---	---	686	
<i>D. rerio</i>	570 RPP	T	P	R	T	P	K	---	---	---	---	---	---	---	---	---	626	
<i>X. tropicalis</i>	601 GVA	SPRAPK	Q	Q	H	W	L	S	V	---	---	---	---	---	---	---	658	
<i>D. melanogaster</i>	597	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	661	
<i>C. elegans</i>	379 PI	V	P	V	Q	A	P	P	Q	H	---	---	---	---	---	---	413	
<i>S. cerevisiae</i>	284	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	317	
	830	840	850	860	870	880	890											
<i>H. sapiens</i>	689 IR	QLF	QY	---	---	---	---	---	---	---	---	---	---	---	---	---	697	
<i>B. taurus</i>	688 IR	QLF	QY	---	---	---	---	---	---	---	---	---	---	---	---	---	731	
<i>M. musculus</i>	689 IR	QLF	QY	---	---	---	---	---	---	---	---	---	---	---	---	---	697	
<i>G. gallus</i>	687 IR	QLF	QY	---	---	---	---	---	---	---	---	---	---	---	---	---	729	
<i>D. rerio</i>	627 I	N	Q	L	F	Q	Y	---	---	---	---	---	---	---	---	---	668	
<i>X. tropicalis</i>	659 IR	QLF	QY	---	---	---	---	---	---	---	---	---	---	---	---	---	702	
<i>D. melanogaster</i>	662 I	N	K	L	Q	V	G	G	A	G	E	K	R	P	A	D	L	
<i>C. elegans</i>	414	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	418	
<i>S. cerevisiae</i>	318	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	330	
	900	910	920	930	940	950	960											
<i>H. sapiens</i>	732 KAK	L	K	R	T	Q	V	Q	H	S	F	P	R	L	D	L	D	
<i>B. taurus</i>	732 KAK	L	K	R	T	Q	V	Q	H	S	F	P	R	L	D	L	D	
<i>M. musculus</i>	730 KAK	L	K	R	T	Q	V	Q	H	S	F	P	R	L	D	L	D	
<i>G. gallus</i>	730 KAK	L	K	R	T	Q	V	Q	H	S	F	P	R	L	D	L	D	
<i>D. rerio</i>	669 KL	L	R	R	T	Q	L	H	R	P	V	Q	R	G	H	L		
<i>X. tropicalis</i>	703 KAK	L	K	R	T	Q	V	Q	H	A	F	A	D	R	E	Y		
<i>D. melanogaster</i>	730 D	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---		
<i>C. elegans</i>	419 K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---		
<i>S. cerevisiae</i>	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---		

Anhang

		1940	1950	1960	1970	1980	1990	2000		
<i>H_sapiens</i>	1338	Y T - - G I Q C Y S C G M R F T T S Q T D - - V Y A D H L D W H Y R - - - - - Q N R T E K D V S - - - - - R K V T H R R W							1384	
<i>B_taurus</i>	1468	Y T - - G I Q C Y S C G M R F T T S Q T D - - V Y A D H L D W H Y R - - - - - Q N R T E K D V S - - - - - R K V T H R R W							1514	
<i>M_musculus</i>	1336	Y T - - G I Q C Y S C G M R F T T S Q T D - - V Y A D H L D W H Y R - - - - - Q N R T E K D V S - - - - - R K V T H R R W							1382	
<i>G_gallus</i>	1475	Y T - - G I Q C Y S C G M R F T T S Q T D - - V Y A D H L D W H Y R - - - - - Q N R S E K D V S - - - - - R K I T H R R W							1521	
<i>D_riero</i>	1241	Y T - - G I Q C Y S C G M R F T A S Q T D - - V Y A D H L D W H Y R - - - - - Q N R S E K D I S - - - - - K K V T H R R W							1287	
<i>X_tropicalis</i>	1485	Y T - - G I Q C Y S C G M R F T K S Q T D - - I Y A D H L D W H Y R - - - - - Q N R T E K D V S - - - - - R K I T H R R W							1531	
<i>D_melanogaster</i>	1363	Y L - - G M Q C S S C G V R F P P E Q T I - - K Y S Q H L D W H F R - - - - - Q N R R E R D S T - - - - - R K A T S R K W							1409	
<i>C_elegans</i>	669	I T K R A D A C K F C G M R L D D S Q G K S K E W D H M D W H V K - - - - - Q N L A R H G S N N S - - - - - A A V P Y R Q W							721	
<i>S_cerevisiae</i>	420	- - - - - K C S V C G K R F G N S E S E K L L Q N E H L D W H F R I N T R I K G S Q N T A N T G I S N S N L N T T T T R K N I Q S R N W							482	
		→								
		2010	2020	2030	2040	2050	2060			
<i>H_sapiens</i>	1385	Y Y S L T D W I E F E E I A D L E E R A K S Q F F E - - - - - K V H - - - - - E E V V L K T Q E A A K E K E F Q - - - - -							1430	
<i>B_taurus</i>	1515	Y Y S L T D W I E F E E I A D L E E R A K S Q F F E - - - - - K V H - - - - - E E V V L K T Q E A A K E K E F Q - - - - -							1560	
<i>M_musculus</i>	1383	Y Y S L T D W I E F E E I A D L E E R A K S Q F F E - - - - - K V H - - - - - E E V V L K T Q E A A K E K E F Q - - - - -							1428	
<i>G_gallus</i>	1522	Y Y S V T D W I E F E E I A D L E E R A K S Q F F E - - - - - K A H - - - - - E E V V L K T Q E A A K E K E F Q - - - - -							1567	
<i>D_riero</i>	1288	Y Y S L T D W I E F E E I A D L E E R A K S Q F F E - - - - - K V H - - - - - E E V V K T Q E A A K E R E F Q - - - - -							1333	
<i>X_tropicalis</i>	1532	Y Y S L K D W I E F E E I A D L E E R A K S Q F F E - - - - - K V H - - - - - E E V I L K T Q E A A K E K E F L - - - - -							1577	
<i>D_melanogaster</i>	1410	Y Y D L N D W R Q Y E E I E D V E E R E K N - F L E A Q G Q - - - - - P G G V E A L D E L S Q Q R S L D - - S P V P							1459	
<i>C_elegans</i>	722	Y P S T S T W - - - - - L T P R A S D E T N E - - - - - Q E A D K P E E - - - - - P L P							750	
<i>S_cerevisiae</i>	483	Y L S D S Q W A A F K D D E I T S T K H K N D Y T D P H A N K N I D K S A L N I H A D E N D E G S V D N T L G S D R S N E L E I R G K Y V							551	
		→								
		2080	2090	2100	2110	2120	2130			
<i>H_sapiens</i>	1431	S V P A G P A G A V E S C E I C Q E Q F E Q Y W D E E E E E W H L K N A I R V D G K I Y H P S C Y E D Y Q - - - - -							1483	
<i>B_taurus</i>	1561	S V P A G P A G A V E S C E I C Q E Q F E Q Y W D E E E E E W H L K N A I R V D G K I Y H P S C Y E D Y Q - - - - -							1613	
<i>M_musculus</i>	1429	S V P A G P A G A V E S C E I C Q E Q F E Q Y W D E E E E E W H L K N A I R V D G K I Y H P S C Y E D Y Q - - - - -							1481	
<i>G_gallus</i>	1568	S V P A G P A G A V E R C E I C R E Q F E Q Y W D E E E E E W H L K N A I R M H E K I Y H P S C Y E D Y Q - - - - -							1620	
<i>D_riero</i>	1334	S V K A A A D V V H E L C E I C Q E Q F E M Y W E E E E E W H L K N A I R V D E K T Y H P S C Y E D Y K - - - - -							1386	
<i>X_tropicalis</i>	1578	S V P A G P A G A D E I C E I C Q E Q F E Q Y W D E E E E E W H L K N A M R M N K I Y H P S C Y E D Y K - - - - -							1630	
<i>D_melanogaster</i>	1460	T C A A G R D D V D H C C D M C H E K F E Q F Y N E E L E E W H L R S A I R V E D K I Y H P L C Y E D Y K A S L N P P T E V K S D Q D V D							1528	
<i>C_elegans</i>	751	G V A S S G V K T K E - C S V C G E K F D E Y Y D D D E E T W R L R D T V N V H G K I V H S A C A S D A A R S L - - - - -							805	
<i>S_cerevisiae</i>	552	V V P E T S Q D M A F K C P I C K E T V T G V Y D E E S G E W V W K N T I E V N G K Y F H S T C Y H E T S - - - - -							604	
		→								
		2140	2150	2160	2170	2180	2190	2200		
<i>H_sapiens</i>	1484	- - N T S S F - - - - - D C T P S P S K T P V E N - - - - - P L N I M L N I							1509	
<i>B_taurus</i>	1614	- - N T S S F - - - - - D C T P S P S K T P V E N - - - - - P L N I M L N I							1639	
<i>M_musculus</i>	1482	- - N T S S F - - - - - D C T P S P S K T P V E N - - - - - P L N I M L N I							1507	
<i>G_gallus</i>	1621	- - N T S S F - - - - - D C T P S P S K T P V E N - - - - - P L N I M L N I							1646	
<i>D_riero</i>	1387	- - N T S S F M - - - - - D T T P S P N K M L T E N - - - - - P L T A F L - -							1411	
<i>X_tropicalis</i>	1631	- - N T S S F L - - - - - D S T P S P S K T L L E N - - - - - P L N A M L N L							1657	
<i>D_melanogaster</i>	1529	M N N T D D N A M D T L I K V E D D E D E G T P N S S Q P I F E D D D D V I V L P N E E P S V T E I V D D D D E D E Y V P G N V T R A D							1597	
<i>C_elegans</i>	806	- - D N S S F F - - - - - N D S D - - - - -							815	
<i>S_cerevisiae</i>	605	- - Q N S S K - - - - - S N - - - - - S G K V G L D D							619	
		→								
		2210	2220	2230	2240	2250	2260	2270		
<i>H_sapiens</i>	1510	V K N E L Q E P C D S P - - - - -							1521	
<i>B_taurus</i>	1640	V K N E L Q E P C E S P - - - - -							1651	
<i>M_musculus</i>	1508	V K N E L Q E P C E S P - - - - -							1519	
<i>G_gallus</i>	1647	V K Q E T D E S C D S P - - - - -							1658	
<i>D_riero</i>	1412	K Q E Q D D E T S C S - - - - -							1422	
<i>X_tropicalis</i>	1658	V K E E V T E P T C S - - - - -							1668	
<i>D_melanogaster</i>	1598	M G N E S Q D K T E S N S E T K E P K K D Q V E L G E Q Q N E S A N E S D V E I Q E P N I P F T D L D T Y V E K E M D E A T R A A L L N							1666	
<i>C_elegans</i>										
<i>S_cerevisiae</i>	620	L K K L V T K - - - - -							626	
		→								
		2280	2290	2300	2310	2320	2330	2340		
<i>H_sapiens</i>	1522	- - K V K E E R I D - - - - - T P P A C T E E - - - - - S I A							1540	
<i>B_taurus</i>	1652	- - K V K E E R T D - - - - - T P P A C T E E - - - - - S I A							1670	
<i>M_musculus</i>	1520	- - K V K E E Q I D - - - - - A P P A C S E E - - - - - S V A							1538	
<i>G_gallus</i>	1659	- - Q V K E E P D D - - - - - T P A A C A E E - - - - - S T P							1677	
<i>D_riero</i>	1423	- - S I K E E P A D D - - - - - D M E A P M V K E E - - - - -							1441	
<i>X_tropicalis</i>	1669	- - V K E E P A D - - - - - Y T A S S C C E - - - - - I P							1685	
<i>D_melanogaster</i>	1667	V K I K E E P K D E Y E E D E D D G F E D V G T V V S L L P L P E D E I S I H S S E T Q T H T I G S S A S P A T I E R P A S V T S L S L P							1735	
<i>C_elegans</i>	816	- - I K K E E P F D - - - - -							823	
<i>S_cerevisiae</i>										
		→								
		2350	2360	2370	2380	2390	2400	2410		
<i>H_sapiens</i>	1541	- - - - - T P S E I K T E N D T V E S V - - - - -							1555	
<i>B_taurus</i>	1671	- - - - - T A T E I K T E N D T V E S V - - - - -							1685	
<i>M_musculus</i>	1539	- - - - - T P T E I K T E S D T V E S V - - - - -							1553	
<i>G_gallus</i>	1678	- - - - - A S I D V K T E P D - - E S V - - - - -							1690	
<i>D_riero</i>	1442	- - - - - V Q V K L E G E S Q T S A I V F - - - - -							1457	
<i>X_tropicalis</i>	1686	- - - - - L L T E I K K E P A E P E - - - - -							1698	
<i>D_melanogaster</i>	1736	A N E A D E L E Q G D A N T D T D A E L N G E K Q E A T H N L S A V G P A L P L A S I V N K I K I N I T K N T S S N S H N S A S N T T T T							1804	
<i>C_elegans</i>										
<i>S_cerevisiae</i>										

Abbildung 9.6 Alignment verschiedener homologer Pcf11-Varianten aus ausgesuchten Spezies. Eingefärbte Aminosäuren entsprechen einer Konservierung zwischen 15% (schwach) und 100% (stark) zwischen den untersuchten Sequenzen. Die ockerfarbenen Pfeile markieren die hoch konservierte CID. Die blauen Pfeile markieren die Bereichsgrenzen der FEGP-Wiederholungen, welche in Vertebraten konserviert scheinen. Die orangenen Pfeile markieren die beiden hoch konservierten Zinkfinger welche die Clp1-Interaktionsregion (grauer Pfeil) flankieren. Erstellt mit JalView aus Sequenzen der UniProt KB Datenbank und aligned mit dem Mafft-Server.

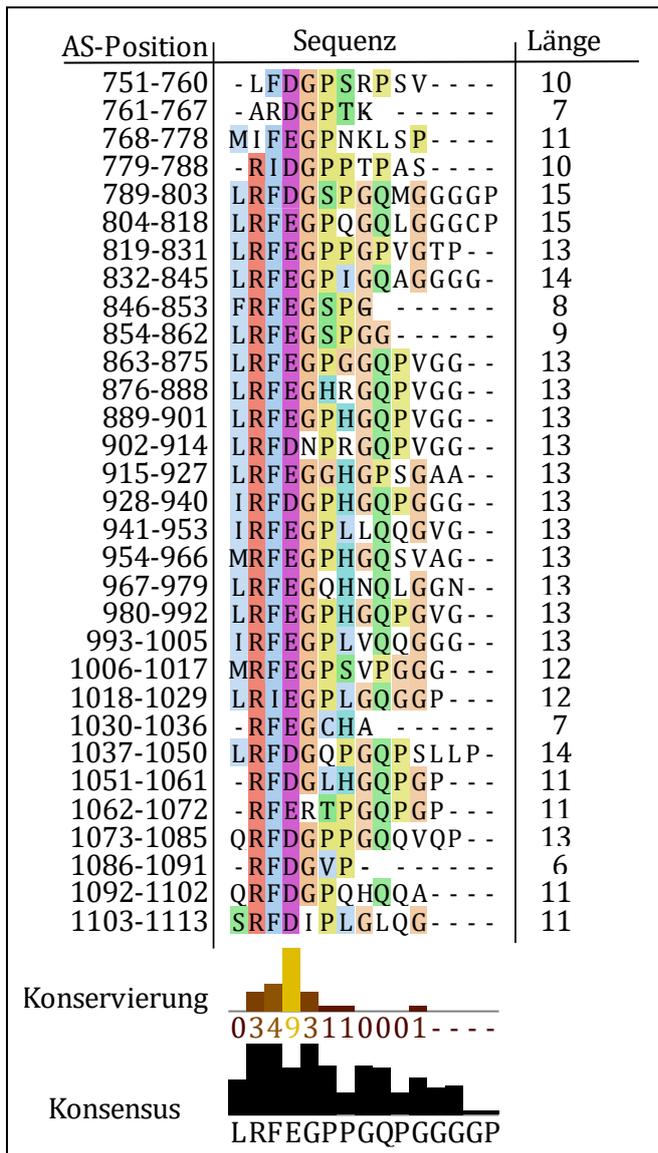


Abbildung 9.7 Alignment der FEGP-Wiederholungen aus hPcf11. Alignment der Region zwischen Aminosäure (AS) 751 und 1113 des hPcf11. Jede Wiederholung wurde bei der Aminosäureposition -2 oder -1 des Phenylalanins getrennt und dann mit dem Mafft-Server aligned. Eingefärbte Aminosäuren entsprechen einer Konservierung zwischen 15% (schwach) und 100% (stark) zwischen den untersuchten Sequenzen. Darunter die %-Konservierung als Ziffern zwischen 0 und 9 der jeweiligen Position und die ermittelte Konsensus-Sequenz der Wiederholungen mit Balken-Indikator der Übereinstimmung. Erstell mit JalView.

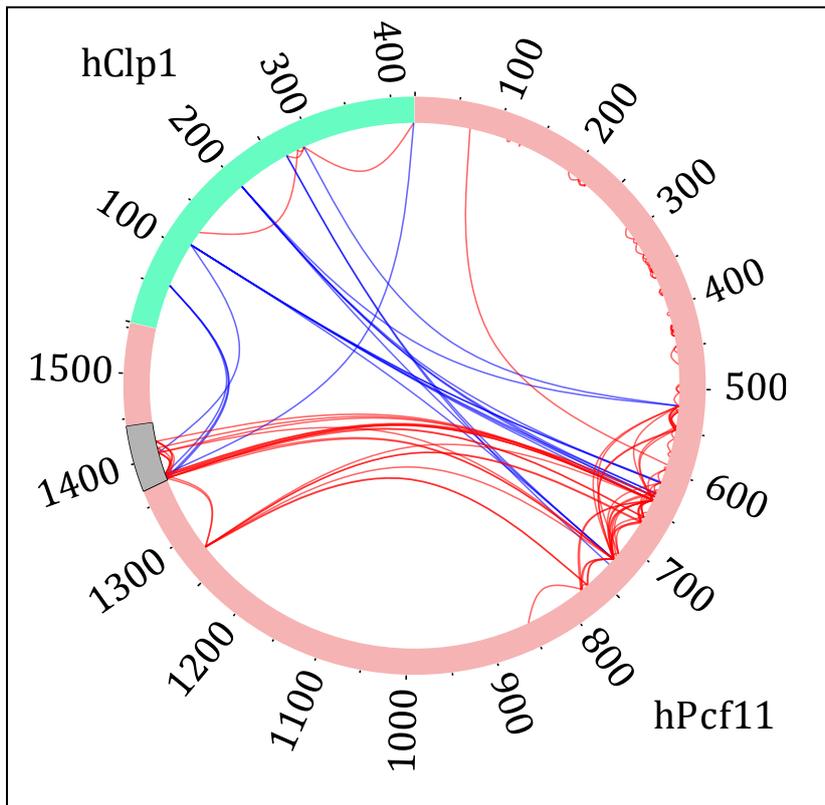


Abbildung 9.8 Übersichtsdarstellung der CF II-Crosslinks.

Übersicht der CF II-Crosslink-Daten und der anschließenden massenspektrometrischen Analyse durch C. Tüting. Die Kreisdarstellung repräsentiert die Primärstruktur der Untereinheiten hPcf11 (rot) und hClp1 (grün) mit den angegebenen Positionen der Aminosäuren. Blaue Linien repräsentieren intermolekulare und rote Linien intramolekulare Crosslinks. Die hClp1-Interaktionsregion in hPcf11 (basierend auf der Kristallstruktur in Noble u. a. 2007) ist grau unterlegt. Modifiziert übernommen aus Schäfer u. a. 2018.

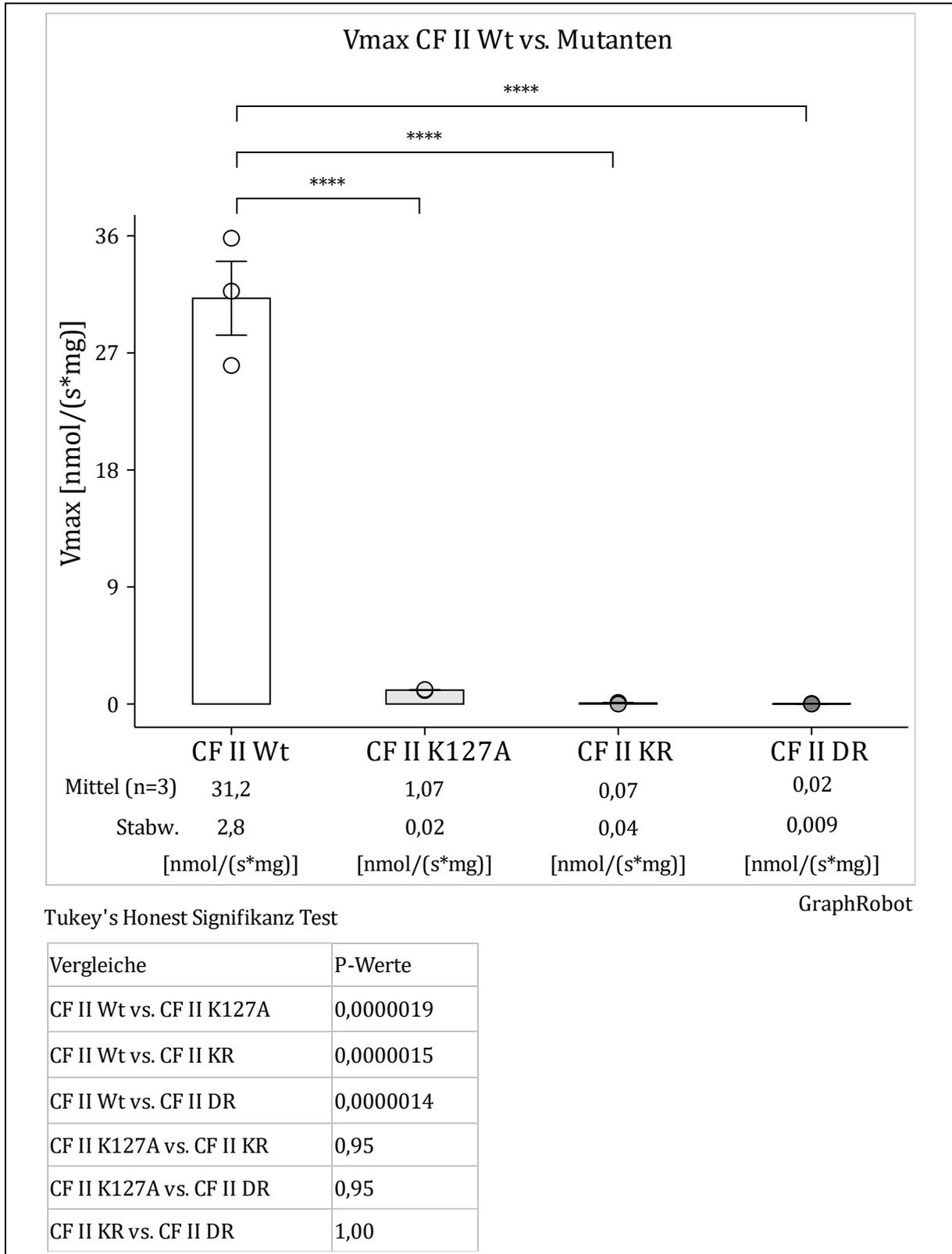


Abbildung 9.9 Vergleich der maximalen Reaktionsgeschwindigkeiten von CF II Wt und Mutanten. Balkendiagramm der maximalen Reaktionsgeschwindigkeit der 5'-PNK-Aktivität des CF II-Komplexes (hClp1) sowie den in der Arbeit verwendeten hClp1-Mutanten K127A, K127A+R288A+R293L und D151A+R288A+R293L in CF II. Werte aus 3 verschiedenen Versuchen mit unterschiedlichen Konzentrationen an Protein bei konstanten $10 \mu\text{M}$ rC_{14} und 1 mM ATP, normalisiert auf die Proteinmenge. Statistische Auswertung der Daten mit Angabe der Mittelwerte und der Standardabweichung (Stabw.) und Signifikanztest. Erstellt mit GraphRobot.

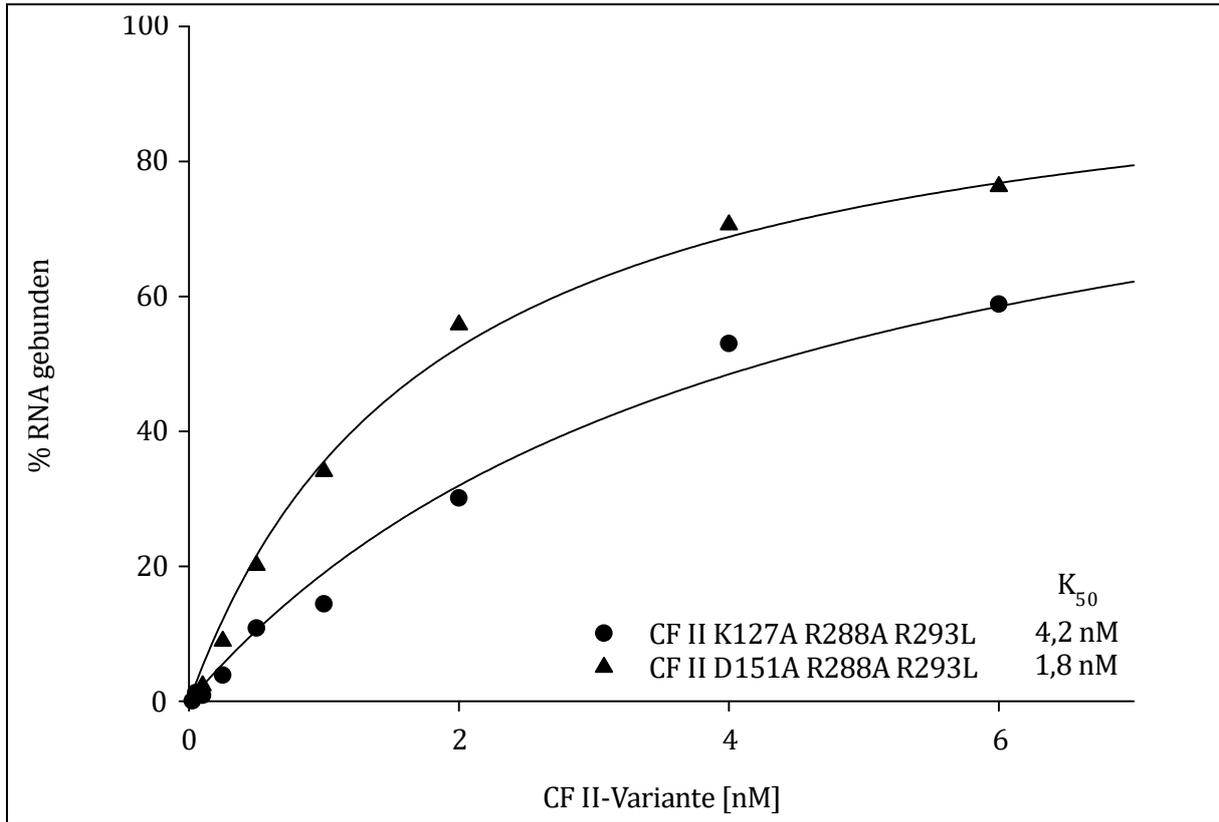


Abbildung 9.10 RNA-Bindung von CF II mit Mutationen in hClp1. Filterbindungsversuche der MonoQ CF II-Präparationen KR (hClp1 K27A, R288A, R293L) Fraktion #18 (Kreis) und DR (hClp1 D151A, R288A, R293L) Fraktion #16 (Dreieck) an 0,1 nM SV40 late-RNA.

Tabelle 9.1 Parameter der Substratüberschusshemmung der CF II 5'-PNK-Reaktion. Berechnung der kinetischen Parameter der 5'-PNK-Reaktion verschiedener verwendeter RNA-Substrate. Abgeleitet wurden die Parameter K_M , K_I und V_{max} aus dem Modell der Substratüberschusshemmung.

Substrat	K_M [μM]	K_I [μM]	V_{max} [$\text{nmol}\cdot\text{s}^{-1}\cdot\text{mg}^{-1}$]
Wigley -23/-10	1,1/0,5	1,7/2,6	120/77
SV40 late 190-225	0,9/0,3	1,7/1,8	71/70
rA ₁₄	0,6	10	50
rU ₁₄	6,7	/	62
rC ₁₄	3,6±2,7	57±23	125±72

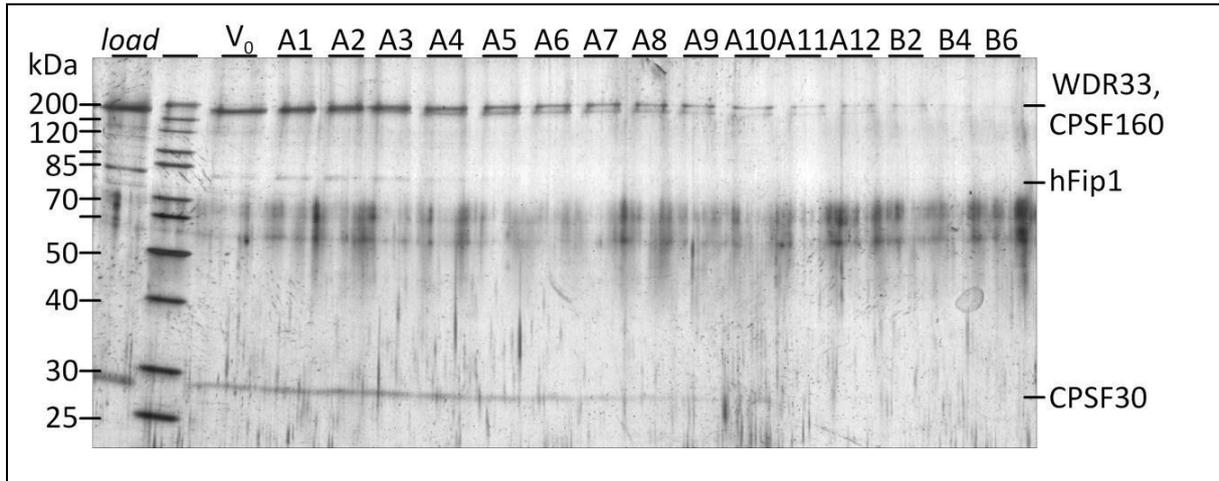


Abbildung 9.11 Größenausschlusschromatographie von mPSF. Silbergefärbtes 11% SDS-PA-Gel. Die untersuchten Fraktionen sind oben angegeben, die Molekulargewichte der Markerbanden am linken Bildrand zugeordnet und die Laufhöhen der mPSF-Untereinheiten an der rechten Seite gekennzeichnet. Die Identität der sichtbaren Proteinbanden wurde mittels Western-Blot-Analysen überprüft (Daten nicht gezeigt). Die Fraktion des Ausschlussvolumens ist mit „V₀“ benannt. (aus der Dissertation von Lars Schönemann 2017).

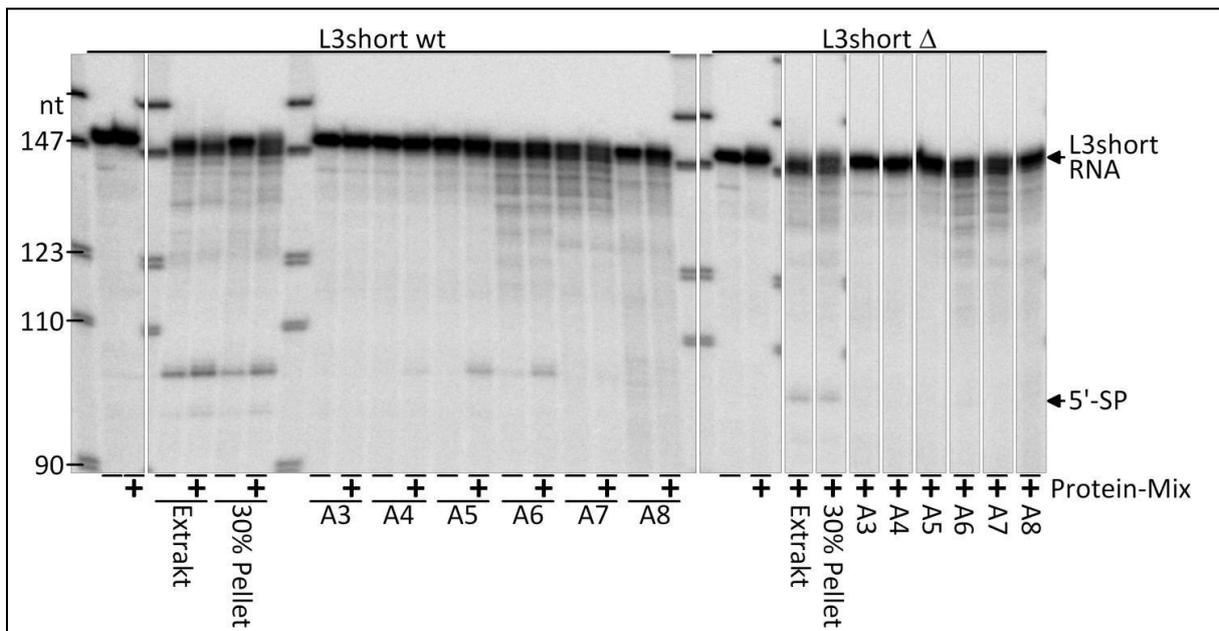


Abbildung 9.12 : Spaltungsaktivität nach Fraktionierung mittels Größenausschlusschromatographie. Autoradiogramm nach denaturierender Harnstoff-PAGE zur Analyse der Spaltungsaktivität von Fraktionen der Größenausschlusschromatographie (Superose 6) auf L3short wt und Δ RNA. Die Reaktionen in der ersten Spur von links für beide RNAs enthielten jeweils nur RNA, die zweite Gelspur zeigt Reaktionen mit Protein-Mix aus rekombinanter Poly(A)-Polymerase, CPSF (7er-Komplex), CstF, CF I, CF II, WDR82, Ssu72 und phosphorylierter CTD zu je 5 nM. In den nachfolgenden Gelspuren wurden Reaktionen mit Kernextrakt, nicht-dialysierter Pelletfraktion nach Präzipitation mit 30% Ammoniumsulfat und Fraktionen der Größenausschlusschromatographie („A3“ bis „A8“) durchgeführt. Für die L3short wt RNA sind die Aktivitäten für alle Fraktionen jeweils allein und nach Komplementation durch Protein-Mix gezeigt, für die L3short Δ RNA nur nach Komplementation mit Protein-Mix. Die Größen der Markerbanden (in Nukleotiden) sind für beide Abbildungsteile an der linken Seite zugeordnet. Auf der rechten Seite sind die Laufhöhen der nicht-gespaltenen L3short RNA sowie des erwarteten 5'-Spaltprodukts (5'-SP) angegeben. (aus der Dissertation von Lars Schönemann 2017).

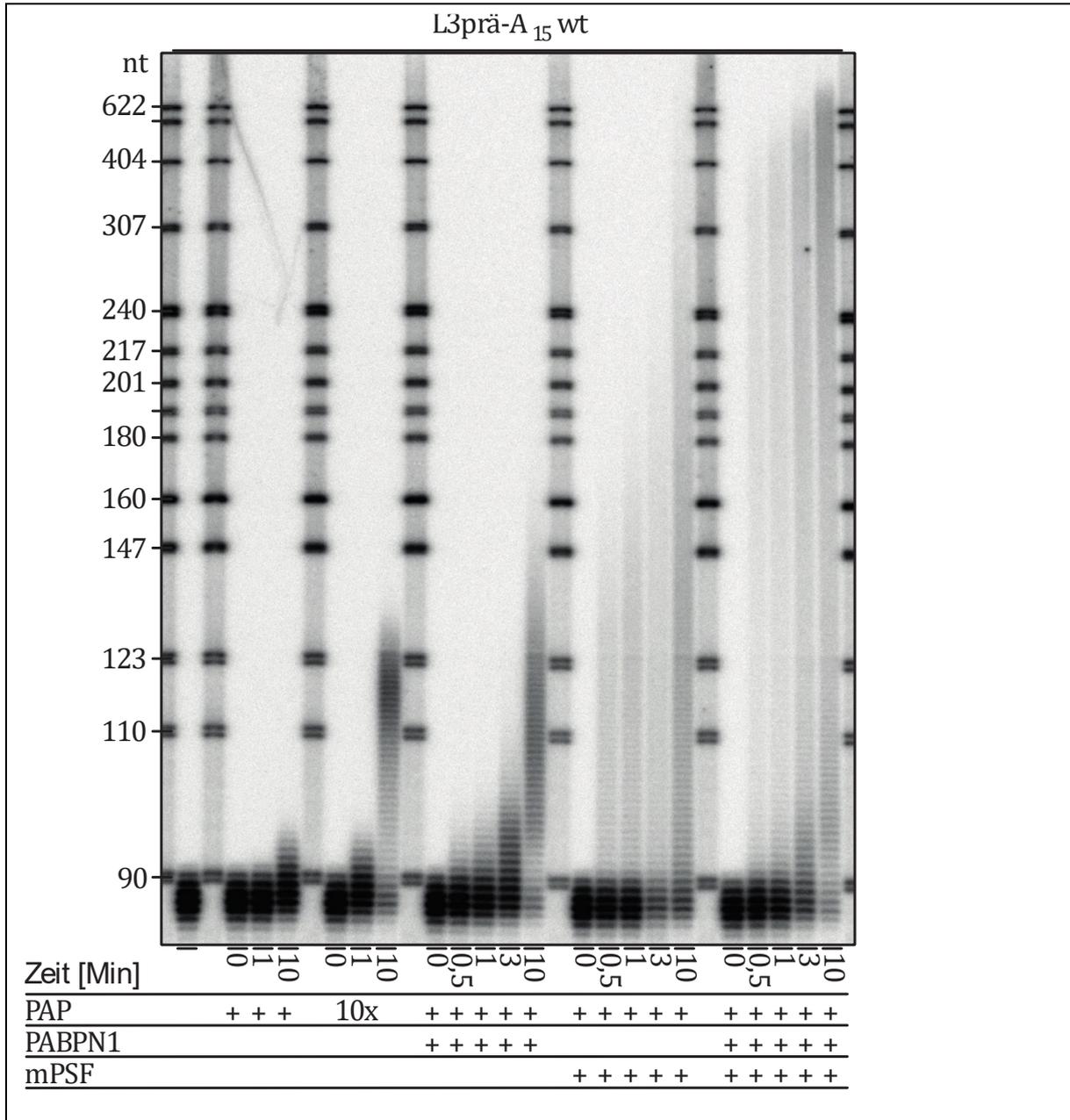


Abbildung 9.13 : Längenkontrolle in der Polyadenylierung mit mPSF. Autoradiogramm nach denaturierender Harnstoff-PAGE zur Analyse der Längenkontrolle der Polyadenylierung. Es wurden Polyadenylierungskinetiken mit 4 nM L3prä-A15 wt RNA, 0,4 nM Poly(A)-Polymerase (PAP), 4 nM mPSF und 60 nM rekombinantem PABPN1 aufgenommen. Die Reaktionszeiten sind unten (in Minuten) angegeben. Die erste Spur von links zeigt nur die eingesetzte RNA, es folgen drei Spuren mit der einfachen Menge an Poly(A)- Polymerase und drei Spuren mit der zehnfachen Menge an Poly(A)-Polymerase. Anschließend sind Kinetiken mit PABPN1, mPSF und beiden Proteinpräparationen zu sehen. Die Größen der Markerbanden (in Nukleotiden) sind an der linken Seite zugeordnete (entnommen aus der Dissertation von L. Schönemann 2017).

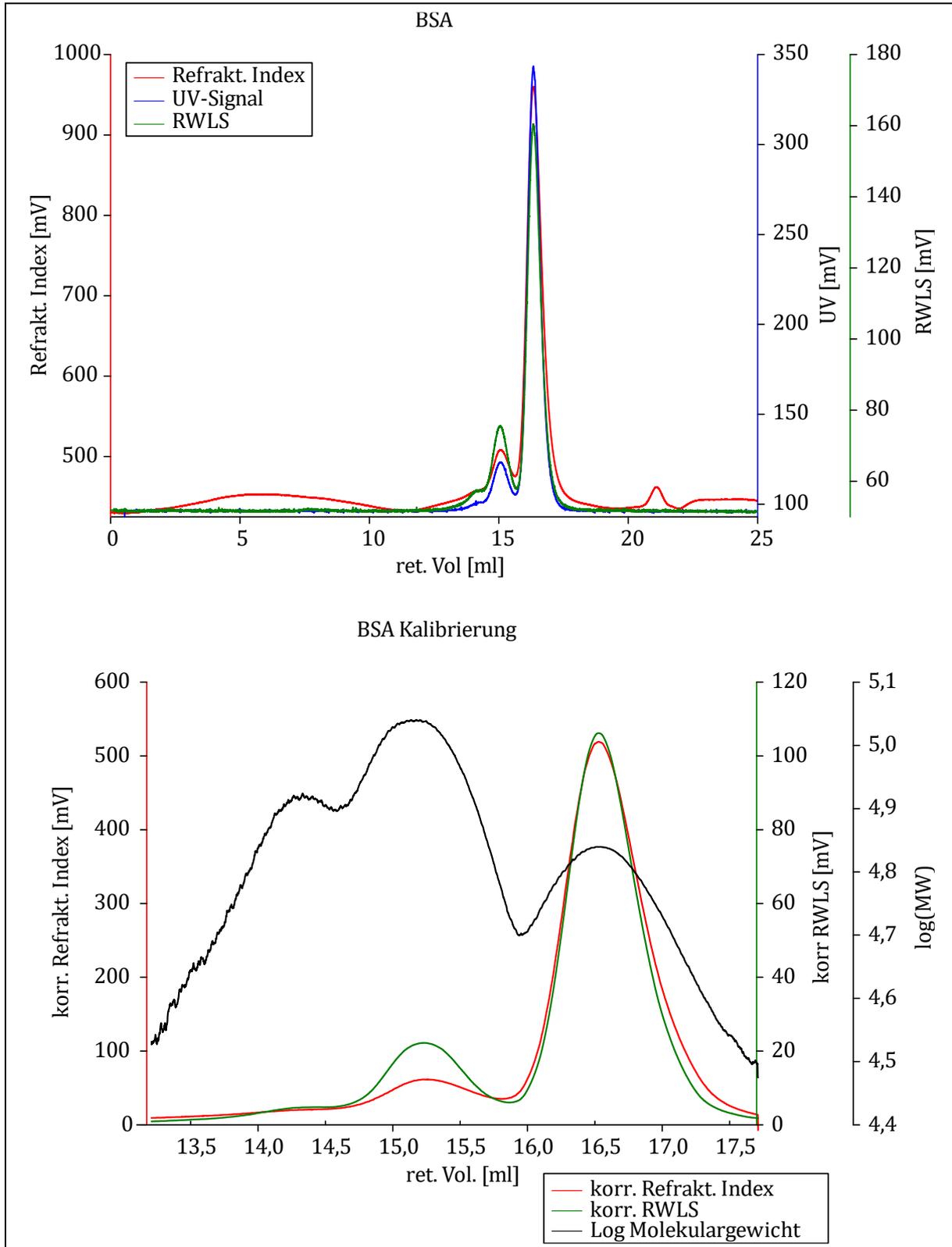


Abbildung 9.14 BSA-Größenausschlusschromatographie mit angeschlossener Rechtwinkel-Lichtstreuung. Die obere Graphik zeigt die Detektorsignale für UV (blaue Linie), Brechungsindex (rote Linie) und die Rechtwinkel-Lichtstreuung (grüne Linie) der BSA-Auftrennung (Albumin Fraktion IV). Die untere Graphik zeigt die mit der OMNISEC Software korrigierten Signale (selbe Farbgebung) und die berechneten $\log(MW)$ der Protein-Peaks. Theoretisches Molekulargewicht des Monomers ist 66 kDa.

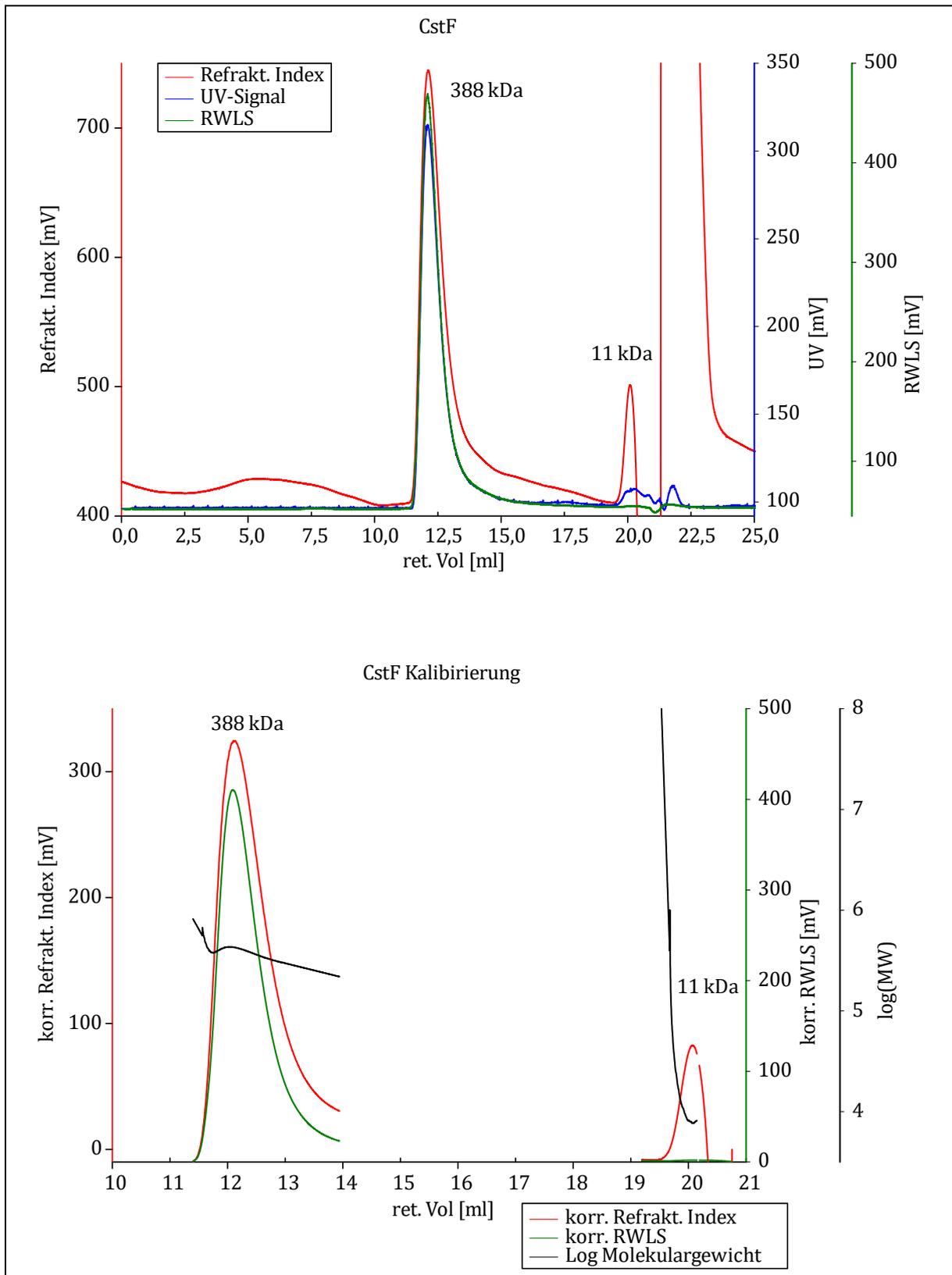


Abbildung 9.15 CstF-Größenausschlusschromatographie mit angeschlossener Rechtwinkel-Lichtstreuung. Die obere Graphik zeigt die Detektorsignale für UV (blaue Linie), Brechungsindex (rote Linie) und die Rechtwinkel-Lichtstreuung (grüne Linie) der CstF-Auftrennung der ResourceQ Fraktion #12. Die untere Graphik zeigt die mit der OMNISEC Software korrigierten Signale (selbe Farbgebung) und die berechneten $\log(MW)$ der Protein-Peaks. Theoretisches Molekulargewicht des Heterohexamers ist 382 kDa.

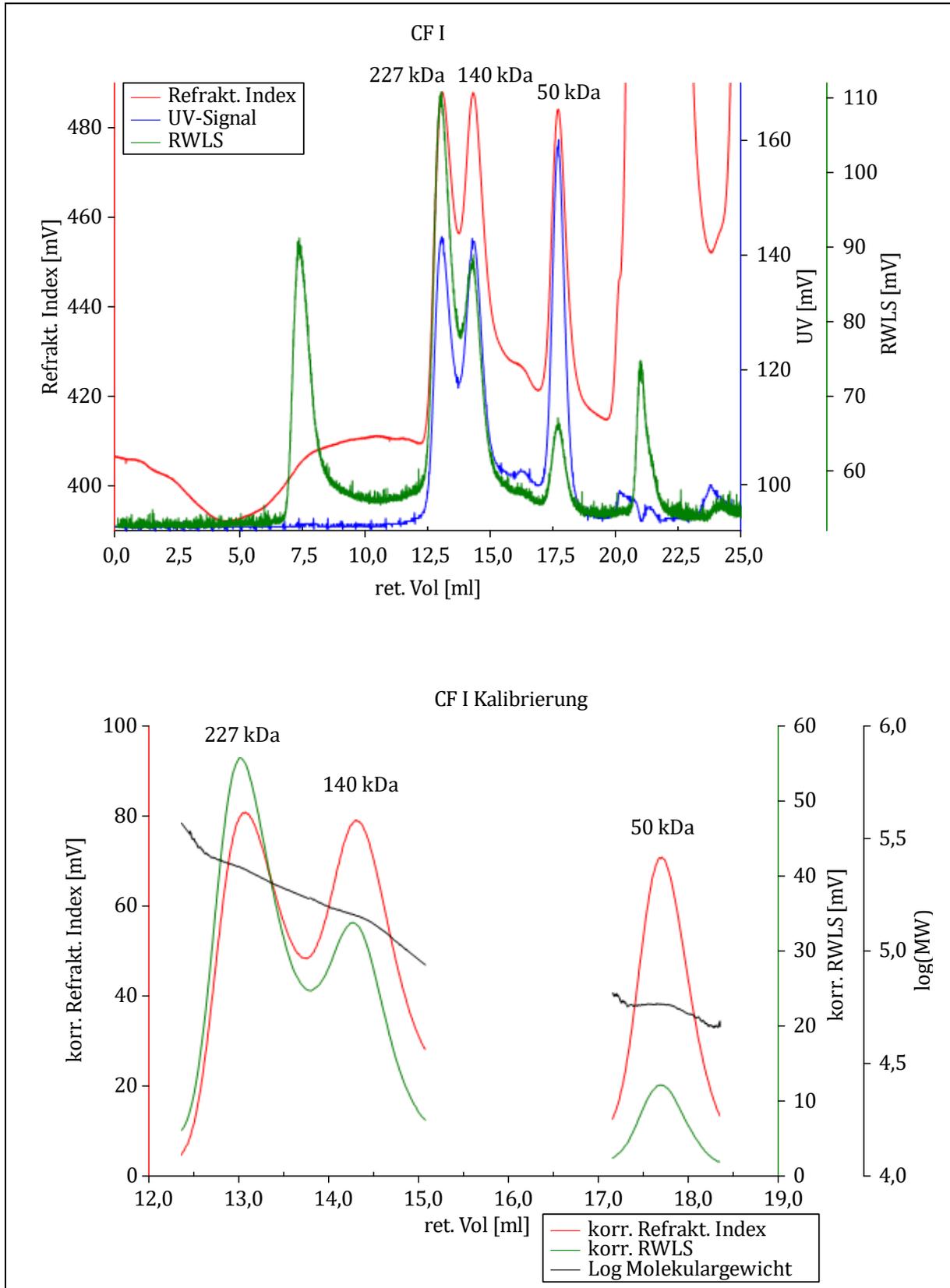


Abbildung 9.16 CF I-Größenausschlusschromatographie mit angeschlossener Rechtwinkel-Lichtstreuung. Die obere Graphik zeigt die Detektorsignale für UV (blaue Linie), Brechungsindex (rote Linie) und die Rechtwinkel-Lichtstreuung (grüne Linie) der CF I-Auftrennung der Ni-NTA Fraktion E2. Die untere Graphik zeigt die mit der OMNISEC Software korrigierten Signale (selbe Farbgebung) und die berechneten $\log(MW)$ der Protein-Peaks. Theoretisches Molekulargewicht des Heterotetramers ist 186 kDa. Die homodimere haben ein theoretisches Molekulargewicht von 136 kDa (CF I-68k) und 50 kDa (CF I-25k).

Reinigungen von CF II

Expression

Wirtszellen	verwendeter Virus	Expressionsdauer	Zellmenge
Sf21	pBac(Strep-hPcf11, His-hClp1) ^{Tn7-} -(YFP) ^{loxP}	72 h	1,6x10 ⁹

Ni-NTA Affinitätsreinigung

Material	Zellmenge	Lysepuffer	Hauptfraktionen	Proteinausbeute
3 ml Ni-NTA Agarose (Qiagen)	1,6x10 ⁹	300 mM Kaliumchlorid, 50 mM Tris pH 8,0, Saccharose	E2, E3, E4	5 mg

MonoQ-Affinitätsreinigung

Säule	Load	Salzgradient	Hauptfraktionen	Proteinausbeute
1 ml MonoQ (GE Healthcare)	Ni-NTA E2, E3	170-1000 mM Kaliumchlorid, 5 µM Zinkchlorid, 15 Säulenvolumen	#15-21	2 mg

Superose 6-Größenausschlusschromatographie

Säule	Load	Puffer	Hauptfraktionen	Proteinausbeute
100 ml Superose6	MonoQ #16, 18, 19 (0,5 mg)	Ni-NTA Lysepuffer	#09-11	260 µg

Reinigungen von hClp1

Expression

Wirtszellen	verwendeter Virus	Expressionsdauer	Zellmenge
Sf21	pBac(His-hClp1) ^{Tn7-} -(YFP) ^{loxP}	72 h	0,6x10 ⁹

Ni-NTA Affinitätsreinigung

Material	Zellmenge	Lysepuffer	Hauptfraktionen	Proteinausbeute
----------	-----------	------------	-----------------	-----------------

1 ml Ni-NTA Agarose (Qiagen)	0,6x10 ⁹	100 mM Kaliumchlorid, 50 mM Tris pH 7,5, Saccharose	E2, E3	3,5 mg
------------------------------	---------------------	---	--------	--------

Superdex200 pg-Größenausschlusschromatographie

Säule	Load	Puffer	Hauptfraktionen	Proteinausbeute
120 ml Superdex200 pg	Ni-NTA E2 (2,1 mg)	Ni-NTA Lysepuffer, 0,1 mM ATP	#16-19	510 µg

Reinigungen von CF I

Expression

Wirtszellen	verwendeter Virus	Expressionsdauer	Zellmenge
Sf21	pBAC(His-25k_Strep-68k) ^{Tn7} -(YFP) ^{loxP}	72 h	1,4x10 ⁹

Ni-NTA Affinitätsreinigung

Material	Zellmenge	Lysepuffer	Hauptfraktionen	Proteinausbeute
1 ml Ni-NTA Agarose (Qiagen)	1,4x10 ⁹	300 mM Kaliumchlorid, 50 mM Tris pH 7,5, Saccharose	E2, E3, E4	12 mg

ResourceQ-Affinitätsreinigung

Säule	Load	Salzgradient	Hauptfraktionen	Proteinausbeute
1 ml ResourceQ (GE Healthcare)	Ni-NTA E2 (8,2 mg)	160-1000 mM Kaliumchlorid, 15 Säulenvolumen	#11-17	7,3 mg

Strep-Tactin-Affinitätsreinigung

Säule	Load	Puffer	Hauptfraktionen	Proteinausbeute
0,2 ml Strep-Tactin-Sepharose (IBA)	Ni-NTA E3 (0,5 mg)	Ni-NTA Lysepuffer	E3, E4, E5	260 µg

Reinigungen von CstF

Expression

Wirtszellen	verwendeter Virus	Expressionsdauer	Zellmenge
Sf21	pBAC(His-77k, 64k, Strep-50k) ^{Tn7-} -(YFP) ^{loxP}	72 h	1,8x10 ⁹

Ni-NTA Affinitätsreinigung

Material	Zellmenge	Lysepuffer	Hauptfraktionen	Proteinausbeute
3 ml Ni-NTA Agarose (Qiagen)	1,8x10 ⁹	50 mM Kaliumchlorid, 50 mM Tris pH 7,5, Glycerin	E2, E3, E4	12 mg

ResourceQ-Affinitätsreinigung

Säule	Load	Salzgradient	Hauptfraktionen	Proteinausbeute
1 ml ResourceQ (GE Healthcare)	Ni-NTA E2+3 (9 mg)	50-1000 mM Kaliumchlorid, 15 Säulenvolumen	#11-16	7,7 mg

9.2 Lebenslauf

Persönliche Daten

Name:	Peter Schäfer
Akademischer Grad	Master of Science Biochemie
Anschrift:	Halle (Saale)
Geschlecht:	männlich
Nationalität:	deutsch

Ausbildung

März 2014 – November 2020	<p>Wissenschaftlicher Mitarbeiter, Doktorand, Martin-Luther-Universität Halle-Wittenberg, Institut für Biochemie und Biotechnologie, Allgemeine Biochemie (Prof. Dr. Elmar Wahle),</p> <ul style="list-style-type: none"> • Thema: Charakterisierung und Darstellung der humanen 3'-Prozessierungsfaktoren mPSF und CF II
Oktober 2011 – Januar 2014	<p>Masterstudium der Biochemie, Martin-Luther-Universität Halle-Wittenberg</p> <ul style="list-style-type: none"> • Schwerpunkte: Pflanzenbiochemie, Nukleinsäurebiochemie, Proteintechnologie <p>Masterarbeit: „Zur Rolle von FtsH-Proteasen in der Plastiden-Biogenese“</p>
Oktober 2008 – September 2011	<p>Bachelorstudium der Biochemie, Martin-Luther-Universität Halle-Wittenberg</p> <p>Bachelorarbeit: „Untersuchungen zur Expression und Funktion des PARP-6 Transkripts in Hodgkin-Lymphom Zellen“</p>
August 2000 – Juni 2007	<p>„Carl-Friedrich-Gauß-Gymnasium“, Schwedt/Oder</p>

Lehrgänge

März 2018	Grundlagenlehrgang GMP
-----------	------------------------

Halle, den 31.05.2021

Peter Schäfer

9.3 Publikationen

Schäfer P, Helm S, Köhler D, Agne B, Baginsky S. Consequences of impaired 1-MDa TIC complex assembly for the abundance and composition of chloroplast high-molecular mass protein complexes. *PLoS One*. 2019 Mar 13;14(3):e0213364. doi:10.1371/journal.pone.0213364.

Schäfer P, Tüting C, Schönemann L, Kühn U, Treiber T, Treiber N, Ihling C, Graber A, Keller W, Meister G, Sinz A, Wahle E. Reconstitution of mammalian cleavage factor II involved in 3' processing of mRNA precursors. *RNA*. 2018 Dec;24(12):1721-1737. doi:10.1261/rna.068056.118. Epub 2018 Aug 23.

Schönemann L, Kühn U, Martin G, Schäfer P, Gruber AR, Keller W, Zavolan M, Wahle E. Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. *Genes Dev*. 2014 Nov 1;28(21):2381-93. doi:10.1101/gad.250985.114. Epub 2014 Oct 9.

Vortrag:

“Reconstitution and characterization of the human Cleavage Factor II”. 13.12.2018, jährlicher Retreat des Graduiertenkollegs GRK 1591, Stangerode, Deutschland

Poster-Präsentation:

Peter Schäfer, Lars Schönemann, Christian Tüting, Uwe Kühn and Elmar Wahle. “The human pre-mRNA 3' processing factor Cleavage Factor II (CFII_m) consists of two subunits and binds RNA with high affinity.” 30.05.-03-06.2017, 22nd Annual Meeting of the RNA Society, Prag, Tschechien

9.4 Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und ohne fremde Hilfe verfasst habe, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Die eingereichte schriftliche Fassung der Arbeit entspricht der auf dem elektronischen Speichermedium. Weiterhin versichere ich, dass die vorliegende Arbeit noch nicht als Abschlussarbeit an anderer Stelle eingereicht wurde.

Halle (Saale), 16.11.2020

Peter Schäfer

9.5 Danksagung

Wie so viele andere Leistungen im Leben, ist auch diese Arbeit nicht ohne die Unterstützung, ohne die Weisheit und ohne die Nachsicht anderer Menschen entstanden. So möchte ich im Folgenden den Menschen danken, die mich maßgeblich in der Anfertigung dieser Arbeit unterstützt haben

Zunächst möchte ich mich bei meinem Doktorvater Prof. Elmar Wahle bedanken. Durch die Gelegenheit, die mir hier zur Anfertigung einer Dissertation gegeben wurde, ist diese Arbeit erst möglich geworden. Zudem wurden Ergebnisse stets kritisch hinterfragt, Probleme konstruktiv diskutiert und mögliche weitere experimentelle Aspekte des Themas vorgeschlagen. Auch die Veröffentlichungen, die aus der Arbeit resultierten, wären ohne seine Anleitung wohl so nicht möglich gewesen. Auch ein Dankeschön für die Geduld und die Motivation in dieser Zeit.

Des Weiteren möchte ich auch apl. Prof. Dr. Ralph Golbik und Prof. Niels Gehring für die Übernahme der Zweit- und Drittgutachten danken.

Natürlich gilt auch den (ehemaligen) Mitgliedern der Arbeitsgruppe mein Dank. Allen voran Dr. Uwe Kühn, der mich oft mit sehr viel Zeit und gutem Rat in praktischen wie theoretischen Fragen während der Arbeit unterstützt hat. Seine sehr genaue und akribische Art führte oftmals zu einem verbesserten Ergebnis, auch wenn dies nicht immer sofort von mir wertgeschätzt wurde. Zudem machte ihn seine ausgeprägte Kollegialität stets zu einem verlässlichen Ansprechpartner.

Mein besonderer Dank gilt auch Dr. Lars Schönemann, ohne dessen akribische Vorarbeit zur Definition des mPSF-Komplexes der erste Teil meiner Arbeit nie möglich gewesen wäre. Auch er hat mich besonders in der Anfangszeit unterstützt und durch ein klein wenig Konkurrenz auch motiviert, mich selbst zu verbessern. Auch den restlichen Mitgliedern des Arbeitskreises der 3'-Prozessierung möchte ich danken. Dr. Christian Tüting hat durch seine Klonierung und Expression der hPcf11-Deletionen, sowie durch die *Crosslinking*-MS-Versuche und *Molecular Modeling* zusätzlich zum Erfolg der Arbeit beigetragen. Gudrun Scholz hat ein ums andere mal schwierige Klonierungen spielend leicht durchgeführt. Moritz Schmidt hat uns in seinen Studienarbeiten durch Herstellung von anderen 3'-Spaltfaktoren unterstützt. Aber natürlich haben alle Mitglieder der AG Wahle mich auf die ein oder andere Weise unterstützt und motiviert.

Zuletzt möchte ich meiner Familie und meinen Freunden danken. Sie haben mich in meiner Ausbildung unterstützt, haben mein Studium finanziert (Danke!) und lange und geduldig auf ein Ergebnis meiner Dissertation gewartet. Sie haben mir auch Freude und Zuversicht gegeben, Korrektur gelesen (Danke Laura), Stress und Stimmungsschwankungen ertragen und trotzdem bedingungslos an mich geglaubt.

Vielen Dank euch allen.